

ВВЕДЕНИЕ В НАСТОЯЩУЮ МАТЕМАТИКУ

Николай Казимиров
Алексей Савватеев

Пособие для учителей математики
по мотивам курса
«100 уроков математики»
Алексея Савватеева



Москва,
Университет Дмитрия Пожарского
2022

УДК 372.851
ББК 74.262.21
К14

Издание подготовлено и выпущено в свет
при финансовой поддержке компании ООО «АЙ ТИ ВИ ГРУПП»

Казимиров Н. И., Савватеев А. В.

К14 Введение в настоящую математику: пособие для учителей математики по мотивам курса «100 уроков математики» Алексея Савватеева. — М.: Русский фонд содействия образованию и науке. Университет Дмитрия Пожарского, 2022. — 408 с.

ISBN 978-5-91244-307-7

Издание представляет собой развернутый и доработанный конспект лекций видеокурса «100 уроков математики» Алексея Савватеева, который был прочитан в Филипповской школе (Москва) в 2014–2018 гг.

В книге на разном уровне строгости и сложности излагается концепция числа. Начиная с простых геометрических образов, описывающих обычные арифметические действия, и заканчивая сложными алгебраическими понятиями, авторы знакомят читателя с началами теории чисел, теории групп, линейной алгебры и комплексного анализа.

Основное внимание в книге уделено следующим темам: движения и подобия прямой и плоскости, линейные уравнения в целых числах, арифметика остатков, кольцо многочленов, группа перестановок, комплексные числа, модели действительных чисел, теория пределов. В книге разбирается ряд известных математических фактов: Основная теорема арифметики, теорема Шаля, теорема Ферма при $n = 4$, неразрешимость задачи об удвоении куба, формула Эйлера, теорема Кантора.

Особое внимание в книге уделяется теоретико-групповому подходу к описанию математических концепций, подробно разбирается структура группы перестановок и связанные с этим задачи. Кроме того, достаточно подробно изучается аксиома полноты (принцип непрерывности) действительных чисел, а также производится построение вещественной и комплексной экспоненты.

Книга снабжена большим количеством вспомогательных чертежей и иллюстраций (более 100), а также задачами различной степени сложности для самостоятельного решения (более 800).

Содержание охватывает такие темы, как: геометрия, линейная алгебра, движения, теория групп, комплексные числа, математический анализ, многочлены, кольцо гауссовых чисел.

УДК 372.851
ББК 74.262.21

ISBN 978-5-91244-307-7

© Казимиров Н. И., Савватеев А. В., текст, иллюстрации, 2022
© Кузьмин А. А., переплет, 2022
© Оформление, Русский фонд содействия образованию и науке, 2022

Глава 1. Визуальная арифметика	8
1.1 Запись действий с отрезками	8
1.2 Понятие натурального числа	20
1.3 Визуальные доказательства	22
1.4 Соизмеримость отрезков, алгоритм Евклида	25
Упражнения	27
Глава 2. Движения прямой	34
2.1 Сдвиг, композиция сдвигов, группа	34
2.2 Отражение	40
2.3 Таблица композиций движений прямой	43
2.4 Теорема о гвоздях	46
2.5 Все конечные подгруппы движений прямой	47
Упражнения	50
Глава 3. Вокруг окружности	52
3.1 Движения окружности	52
3.2 Группа движений окружности	56
3.3 Наматывание прямой на окружность	60
Упражнения	63
Глава 4. Целые числа и ОТА	65
4.1 Целые числа. Кольцо	65
4.2 Кузнечик НОД и алгоритм Евклида	67
4.3 Простые числа и ОТА	69
Упражнения	74
Глава 5. Симметрии фигур	80
5.1 Симметрии правильного треугольника	80
5.2 Симметрии правильного многоугольника	82
5.3 Подгруппы движений окружности	84
5.4 Симметрии ромба, группа Клейна	87
Упражнения	89
Глава 6. Движения плоскости и пространства	90
6.1 Виды движений плоскости. Теорема Шаля	90
6.2 Сравнение движений прямой, окружности и плоскости	103

6.3	Пара слов о движениях сферы	105
6.4	Пара слов о движениях пространства	106
	Упражнения	109
Глава 7.	Линейные уравнения	112
7.1	Уравнение прямой на плоскости	112
7.2	Линейные уравнения в целых числах	117
	Упражнения	121
Глава 8.	Рациональность и соизмеримость	123
8.1	Построение рациональных чисел	123
8.2	Соизмеримость. Иррациональности	132
	Упражнения	138
Глава 9.	Арифметика остатков	140
9.1	Арифметика остатков	140
9.2	Свойства арифметики остатков	147
	Упражнения	150
Глава 10.	Многочлены	155
	Упражнения	163
Глава 11.	Перестановки	169
11.1	Теория множеств: отношения и функции	169
11.2	Обозначения перестановок	177
11.3	Пара слов о конечных группах	183
11.4	Знакопеременная группа	192
11.5	Структура группы перестановок	195
	Упражнения	205
Глава 12.	Комплексная арифметика и алгебра	217
12.1	Алгебра комплексных чисел	217
12.2	Гауссовы целые числа	223
	Упражнения	234
Глава 13.	Введение в линейную алгебру	244
13.1	Преобразования	244
13.2	Подобия прямой и плоскости	248
13.3	Линейное пространство	254
13.4	Линейные операторы	262
13.5	Арифметика матриц	267
13.6	Матрицы и комплексные числа	282
13.7	Действие линейных операторов	284

Упражнения	287
Глава 14. Алгебраические числа	299
14.1 Плотные линейно упорядоченные множества	299
14.2 Зазоры между рациональными числами	302
14.3 О построениях циркулем и линейкой	304
14.4 Многочлены и алгебраические числа	307
Упражнения	313
Глава 15. Континуум	317
15.1 Мощности множеств	317
15.2 Изоморфизмы	326
15.3 Действительные числа	328
15.4 Модели действительных чисел	343
Упражнения	352
Глава 16. Элементы математического анализа	356
16.1 Оценки и пределы	356
16.2 Экспонента	365
16.3 Комплексная экспонента	380
Упражнения	388
Приложение А. Задачи на индукцию	395
Предметный указатель	399
Обозначения	403
Список литературы	407

ПРЕДИСЛОВИЕ К КУРСУ

Данный курс рассчитан на целевую аудиторию, включающую сильных и мотивированных школьников (начиная с седьмого класса), студентов, учителей математики и физики и всех, кто интересуется математикой и планирует так или иначе связать с ней жизнь. Компоновка и содержание курса во многом повторяют «100 уроков математики», прочитанные А. В. Савватеевым в Филипповской школе (г. Москва) в 2014–2018 годах. Этот курс не является учебником и не может подменять собой программу общеобразовательной школы, однако он существенным образом расширяет математический кругозор как учеников, так и учителей математики.

Начальные главы курса представляют собой графическое введение в базовые математические концепции, такие как сложение, умножение, упорядочение чисел. Кроме того, на примере движений прямой, окружности и плоскости формируется понятие числа как преобразования.

Практически сразу появляется алгоритм Евклида, основная теорема арифметики и цепные дроби.

С первых же глав вводятся алгебраические понятия группы и кольца на примере композиций движений и отражений.

Главы 1–6 доступны отдельным сильно мотивированным школьникам 7–9-х классов.

Далее мы приступаем к построению рациональных чисел, решению линейных уравнений в целых числах, разрабатываем теорию делимости, изучаем кольцо вычетов по данному модулю. Довольно подробно в курсе изучаются перестановки, гауссовы целые числа, подобия плоскости.

Главы 7–12 доступны олимпиадным любителям из 7–9-х классов и матшкольникам из старших классов.

Начиная с 13-й главы углубление в математику становится необратимым. Мы рассматриваем векторные пространства, линейные операторы, матрицы и постепенно подбираемся к построению континуума всех вещественных чисел. Рассматривается понятие плотного множества, непрерывного упорядоченного поля, дается несколько формулировок аксиомы полноты.

Заканчивается курс введением в математический анализ и построением комплексной экспоненты с выводом формулы Эйлера.

Главы 13–16 доступны старшеклассникам, планирующим поступать в сильные физико-математические вузы — призёрам и победителям Всероссийской олимпиады.

Студенты физико-математических специальностей и учителя «продвинутой математики» — наша главная целевая аудитория. Данный курс будет для них помощником в самообразовании.

Большое внимание в курсе уделяется подготовке читателя к языку, методам и символике высшей математики.

Значительное количество задач заимствовано из листков, подготовленных для занятий в школе № 179 города Москвы, которые вели в 2017–2022 годах Сергей Дориченко, Дмитрий Ильинский, Сергей Шашков и др. Кроме того, часть задач заимствована из онлайн-проекта «**Де-ти и наука**», выпускающего новую версию лекций А. В. Савватеева «100 уроков математики для детей».

Авторы выражают отдельную признательность Михаилу Бочкарёву за добросовестное вычитывание пособия на предмет опечаток и недочетов. Ответственность за все оставшиеся недочеты авторы берут на себя!

Отметим также, что книга не состоялась бы без финансовой поддержки «ITV Group».

Визуальная арифметика

Аннотация

В данной главе закладывается фундамент арифметики с помощью визуальных образов. Действия с отрезками и прямоугольниками являются иллюстрацией действий с числами. Цель — дать наглядное обоснование законам арифметики и получить некоторые навыки арифметических операций и сравнений чисел.

Попутно вводится понятие натурального числа как количества применяемых операций композиции, а также как меры длины, площади, объема относительно заданной мерной единицы.

1.1. Запись действий с отрезками

Рассмотрим произвольную геометрическую (т.е. не имеющую толщины и идеально ровную) прямую, и на ней будем откладывать отрезки. Под отрезком мы будем понимать часть прямой, ограниченной с двух сторон выбранными точками — концами отрезка. А под откладыванием отрезка от заданной точки мы понимаем совмещение одного из концов этого отрезка с заданной точкой. Если при этом мы совмещаем левый конец отрезка, то производим откладывание вправо, а если правый — влево.

Откладывать отрезок можно либо от какой-то заданной точки на прямой, либо от уже отмеченного на ней отрезка. Во втором случае откладывание означает стыковку откладываемого отрезка с имеющимся без наложения, так что в итоге получается некоторый больший отрезок.

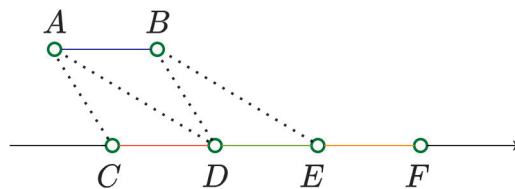


Рис. 1.1. Откладывание отрезка.

Например, на рис. 1.1 можно видеть отрезок AB , ограниченный точками A и B , а ниже, на произвольной прямой мы откладываем данный отрезок сначала от точки C вправо, затем от уже полученного отрезка CD , снова вправо, в результате чего получаем точку E и новый более длинный отрезок CE . Наконец, откладывая третий раз исходный отрезок AB от точки E вправо (или от отрезка CE вправо), получаем точку F и еще более длинный отрезок CF .

Таким образом, откладывание отрезков в одну сторону напоминает одинаковые шаги в выбранном направлении.

Заметим, что направления вправо и влево есть условные обозначения для двух возможных направлений на прямой. Важно один раз относительно данной прямой договориться, какое слово какому направлению соответствует, и в дальнейшем не менять эту договоренность во избежание путаницы.

Мы можем заметить также, что с каждым откладыванием мы получаем все более длинный отрезок, что следует из того, что он включает в себя ранее построенные отрезки. Вообще, если фигура A включает в себя фигуру B (или это включение можно увидеть с помощью наложения фигур друг на друга), то мы считаем фигуру A большей, чем фигура B .

На прямой величина (размер) отрезка — это его длина. При откладывании отрезка мы либо наращиваем длину результирующего отрезка, либо можем ее сокращать, если повернем наше откладывание в противоположном направлении.

Например, на рис. 1.1 мы шли от точки C , постоянно наращивая отрезок, и получили отрезок CF , однако теперь мы можем повернуть вспять и отложить отрезок AB влево от точки F , тем самым вновь вернуться в точку E . В результате таких откладываний мы проделаем путь $C \rightarrow D \rightarrow E \rightarrow F \rightarrow E$, а результирующим отрезком вновь окажется отрезок CE , т.е. длина результирующего отрезка уменьшится на последнем шаге откладываний.

Наблюдение: откладывание вправо есть прибавление длины, а откладывание влево — вычитание (уменьшение) длины.

Конечно, если в результате откладываний влево мы перескочим через исходную точку C и пойдем дальше, то результирующий отрезок снова станет постепенно расти, но мы для определенности договоримся движение вправо называть *прибавлением*, а движение влево — *вычитанием*.

Если точки A и B исходного отрезка совпадают, то такой отрезок превращается в точку. Тем не менее его также можно откладывать, но при таком откладывании мы никуда не сдвинемся. Откладывание

отрезка с совпадающими концами, который мы назовем *нулевым отрезком*, есть прибавление или вычитание нуля.

Мы можем комбинировать откладывание отрезков вправо и влево, т. е. производить *серию* последовательных откладываний отрезков (при этом мы можем пользоваться разными исходными отрезками, не обязательно только отрезком AB).

Результат *серии откладываний* равносителен одному откладыванию отрезка, соединяющего стартовую и финишную точки, причем финишная точка:

- может быть справа от стартовой (результатом является одно откладывание вправо, т. е. прибавление длины);
- может совпадать с ней (результатом оказалось нулевое откладывание);
- быть слева от стартовой точки (результатом является одно откладывание влево, т. е. вычитание).

Серии откладываний можно проиллюстрировать складным метром. Раскладывание колена на 180° означает прибавление его длины к общей серии откладываний, а складывание — вычитание его длины из общей серии откладываний. При этом от стартовой точки можно уйти как вправо, так и влево, или остаться на месте.

С помощью этой же линейки нетрудно продемонстрировать, что композиция откладываний:

- **ассоциативна**, т. е. одни и те же манипуляции с линейкой мы можем производить в разном порядке, например, сначала разложить два ее начальных колена, а затем разложить два колена в конце, и это будет ровно то же самое, как в случае раскладывания сначала конечного участка, а затем начального, — финальная конфигурация линейки будет одинаковой;
- **коммутативна**, т. е. можно взять две линейки, произвести над ними какие-то манипуляции по складыванию/раскладыванию, затем их состыковать, и результат (направление и число колен линейки от начальной точки до конечной) не будет зависеть от того, в каком порядке линейки будут состыкованы.

Кроме того, очевидно, что у каждого откладывания существует обратное, приводящее в результате к нулевому откладыванию. Для этого



нужно произвести ровно ту же самую серию откладываний, только поменять направление. Или, что то же самое, пройти по линейке в обратную сторону.

Далее любое откладывание будем записывать буквами a, b, c, \dots , имея в виду под ними как прибавления, так и вычитания.

Откладывание, противоположное a , будем обозначать $-a$. При этом комбинация откладываний соединяется знаком '+', а если встречается комбинация $a + (-b)$, то пишем проще: $a - b$.

Обратные откладывания — это просто перевернутые в обратную сторону «линейки»!

В контексте настоящей главы под словом **вектор** мы будем понимать произвольное откладывание, т. е., проще говоря, величину и направление сдвига на прямой, полученного с помощью одного или серии произведенных откладываний. При этом вектор «ничего не знает» о конкретных стартовой и конечной точках, он лишь является рецептом (или серией предписаний), содержащем список команд двух типов: отложить такой-то отрезок влево от текущей точки либо отложить такой-то отрезок вправо от текущей точки. В результате применения вектора к данной конкретной точке мы переходим в некоторую, вообще говоря, другую точку на прямой.

Произвольные векторы мы будем обозначать строчными латинскими буквами: a, b, c, \dots .

Если направление вектора — это направление влево (т. е. всякий раз при его откладывании от конкретной стартовой точки финишная точка оказывается левее стартовой), то вектор называется *отрицательным*, а если это направление вправо — *положительным*. **Нулевой вектор** — это вектор, который предписывает оставаться на месте (либо его серия команд приводит к возвращению в начальную точку). Нулевой вектор не имеет направления и обозначается цифрой 0.

Сложение

Теперь если мы производим сначала одну серию откладываний, обозначенную вектором a , а затем другую (или такую же) серию откладываний, обозначенную вектором b , то результат такого совместного откладывания векторов мы будем называть **суммой** векторов, используя обозначение $a + b$.

Несложно догадаться, что если мы рассматриваем векторы как рецепты (списки предписаний), то и сумма векторов также будет рецептом, составленным из последовательного выполнения двух исходных рецептов. Таким образом, *сумма векторов — это тоже некоторый вектор*.

Наконец, если векторы a и b таковы, что их композиция $a + b$ окажется нулевым вектором (т. е. итоговый список предписаний этих векторов в итоге оказался равносильным команде «стоять на месте»), то такие векторы называются **противоположными** друг другу. При этом мы вводим обозначения: $b = -a$ и $a = -b$. Таким образом, по определению мы получаем, что $a + (-a) = 0$.

Добавим, что для всякого вектора a , очевидно, существует противоположный, ведь для его определения достаточно взять инструкцию, предписывающую смещение в противоположную сторону.

Перечислим основные **свойства суммы векторов**:

S1 $(a + b) + c = a + (b + c)$ (ассоциативность).

S2 $a + b = b + a$ (коммутативность).

S3 $a + 0 = 0 + a = a$ (аддитивное свойство нуля).

Эти свойства выше уже были продемонстрированы в процессе рассмотрения операций со складной линейкой.

S4 Для всякого вектора a существует вектор $-a$ такой, что $a + (-a) = 0$ (противоположный элемент).

S5 Если $a = b$, то $a + x = b + x$.

Это свойство мы считаем очевидным из наших геометрических представлений.

S6 Если $a + x = b + x$, то $a = b$ (правило сокращения).

Добавим к исходному равенству в обеих частях элемент $-x$, обратный к x , и получим

$$(a + x) + (-x) = a + (x + (-x)) = a,$$

где мы воспользовались ассоциативностью операции сложения. Аналогично заключаем, что $(b + x) + (-x) = b + (x + (-x)) = b$. Таким образом, используя предыдущее правило, получаем $a = b$.

S7 Верно одно и только одно: либо $a = b$, либо $a = b + x$, либо $a = b - x$, где x — положительный вектор (правило трихотомии).

Это утверждение также требует наглядного интуитивного доказательства. Тут нужно понимать, что наша линейка представляет собой один связный отрезок, а значит, путешественник может пройти по ней из любой точки в любую точку. Отметим на ней точки A и B . Если A находится левее B , то нужно пройти положительный путь от A к B , это и есть наше x , а если наоборот, то отрицательный, т. е. $-x$.

Сложение векторов можно проиллюстрировать с помощью геометрических построений на плоскости.

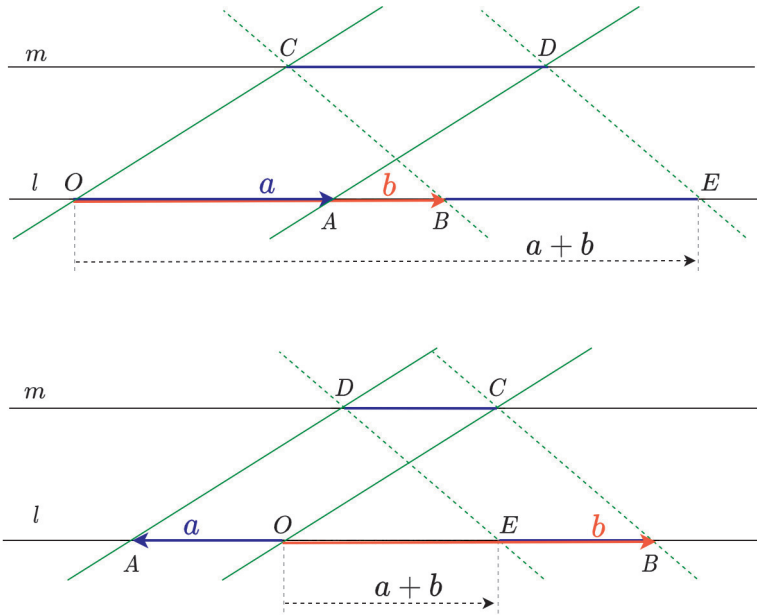


Рис. 1.2. Сумма векторов геометрически.

Возьмем какую-нибудь прямую l на плоскости и отметим на ней точку O в качестве начала отсчета. Отложим от этой точки на прямой l два вектора a и b , получим таким образом точки A и B . Затем выберем произвольную точку C , не лежащую на прямой l , и проведем через нее прямую m , параллельную l . Теперь проведем секущие OC и AD так, что $OC \parallel AD$. После чего проведем еще две секущие: BC и параллельную ей прямую DE . Нетрудно проверить, что точка E получается откладыванием вектора a от точки B (в силу свойств параллелограмма), следовательно, отрезок OE представляет собой результат откладывания вектора $a + b$ от точки O . На рис. 1.2 представлены два случая — сонаправленные и противоположно направленные векторы a и b .

Сравнение

Понятие отрицательного и положительного векторов позволяют ввести сравнение на векторах.

Для начала скажем, что положительный вектор больше нуля, обозначая это следующим способом: $x > 0$ или $0 < x$.

Далее, если $b = a + x$, где $x > 0$, то пишем $a < b$ (или $b > a$).

Свойства сравнения (очевидны):

O1 не верно, что $x < x$ (антирефлексивность);

О2 если $a < b$ и $b < c$, то $a < c$ (транзитивность);

О3 верно одно и только одно: либо $a = b$, либо $a < b$, либо $b < a$ (правило трихотомии);

О4 $a < b \Leftrightarrow a + x < b + x$, где $x > 0$.

Умножение

Перейдем к определению произведения двух векторов.

Построим перпендикулярно направленные оси Ox и Oy . На каждой оси — свой собственный мир векторов и линеек.

Произведение векторов a и b — это площадь, построенная на перпендикулярных отрезках, которые получены в результате откладывания вектора a от точки O по оси Ox и вектора b от точки O по оси Oy (см. рис. 1.3).

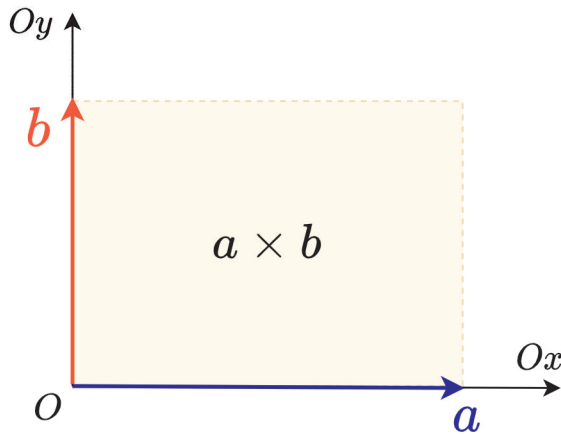


Рис. 1.3. Визуальное умножение.

Площадь — величина всегда неотрицательная, а векторы мы умеем откладывать вверх и вниз, вправо и влево, т. к. имеем две перпендикулярные линейки. Соответственно, у нас имеется 4 различных ситуации: вправо и вверх, вправо и вниз, влево и вверх, влево и вниз.

Здесь мы впервые сталкиваемся с таким понятием, как *ориентированная площадь*. Представим себе, что в нашу плоскость воткнута перпендикулярная ось, с конца которой мы наблюдаем за откладыванием векторов и подсчетом площадей. Когда первый вектор отложен вправо, а второй вверх, то площадь прямоугольника, наблюдаемая нами сверху, замечается по направлению от первого вектора ко второму. Такое направление (против часовой стрелки, или от оси Ox к оси Oy)

в математике принято считать положительным направлением. Поэтому соответствующую площадь мы считаем положительной.¹

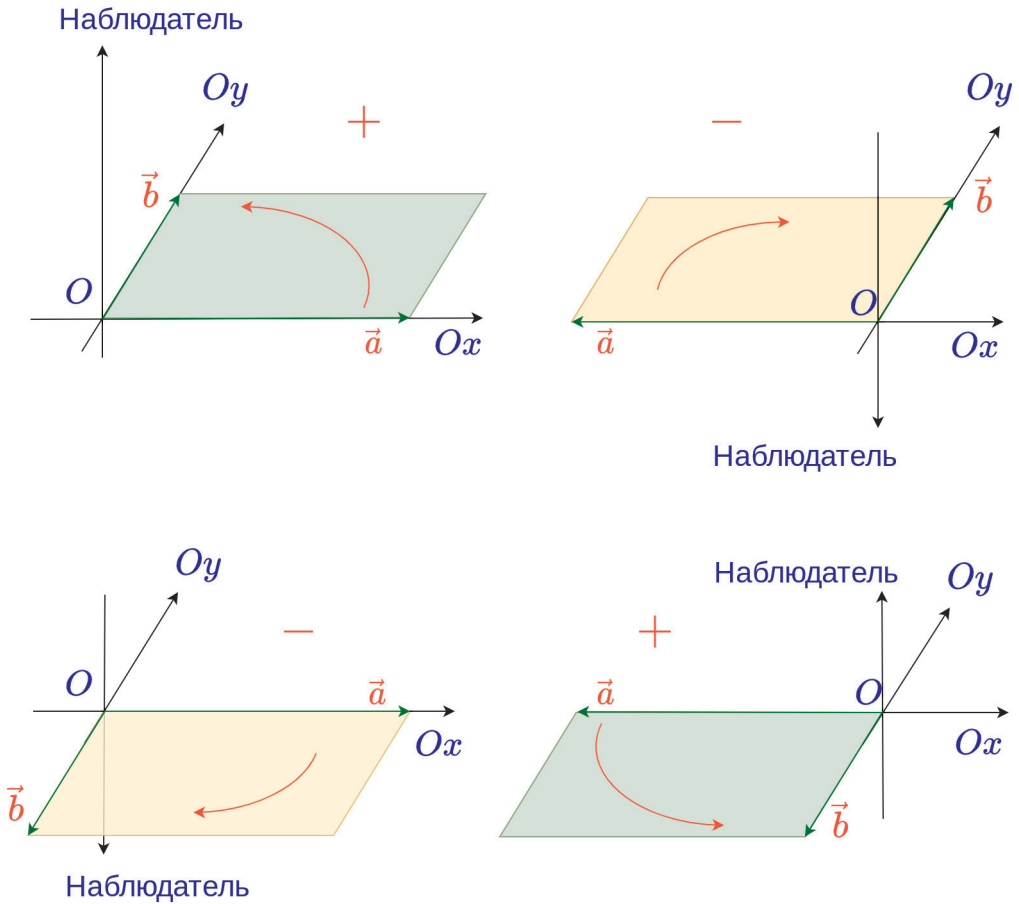


Рис. 1.4. Ориентированная площадь.

Теперь перевернем правый вектор, лежащий на оси Ox и смотрящий вправо, через начало координат, в результате чего он будет смотреть влево и станет отрицательным. Вместе с ним перевернется и наш прямоугольник, площадь которого мы наблюдаем. И вместе с ним перевернется и наблюдательная вышка. Ее направление станет противоположным. И, хотя мы по-прежнему с ее конца будем наблюдать площадь в правильном направлении, вся конструкция перевернется и сменит ориентацию в пространстве на противоположную, поэтому произведение отрицательного вектора на положительный станет также отрицательным.

Вектор наблюдателя в данном случае является не чем иным, как

¹ То, что именно такое направление следует называть положительным, — это всего лишь вопрос договоренностей о терминах.

указателем знака произведения двух исходных векторов a и b . Если стандартное расположение векторов a, b, n (где n — вектор наблюдателя) считать так называемой «правой тройкой» векторов, то при смене знака исходных векторов a и b расположение вектора n (при сохранении их взаимного расположения) будет меняться относительно его начального «правого» расположения — он тоже будет менять знак.

Проследим это на дальнейших шагах.

Теперь перекинем вектор, лежащий на оси Oy , через начало координат, так что он сменит знак на противоположный, и прямоугольник, построенный на этих векторах, снова окажется в правильном, положительном состоянии (о чем свидетельствует направление вектора наблюдателя), хотя обе его стороны есть отрицательные числа. Стало быть, произведение двух отрицательных чисел есть число положительное.

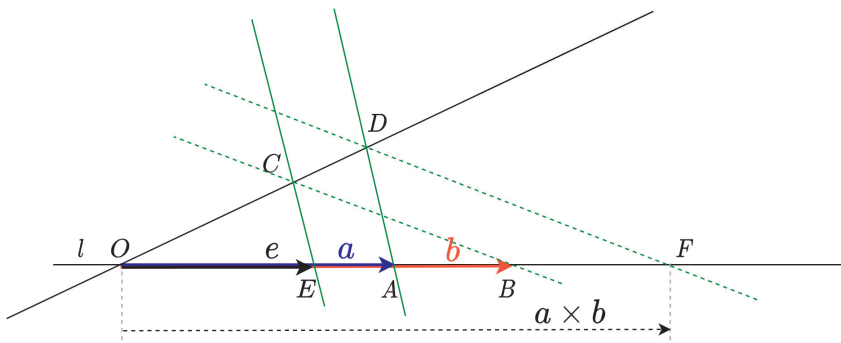
Наконец, еще одна манипуляция с вектором Ox в обратную сторону снова переведет конструкцию в отрицательное положение, и произведение снова станет отрицательным.

Итак, мы видим, что как только мы снабжаем плоскость дополнительным пространственным ориентиром, мы можем различать знак площади, сделать ее ориентированной в зависимости от того, куда направлены векторы, порождающие данную площадь.

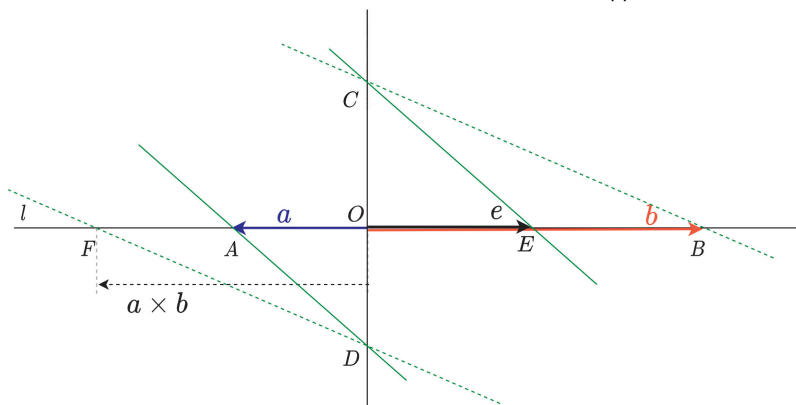
В дальнейшем мы столкнемся с еще более общей конструкцией, где будем строить ориентированную площадь на произвольном параллелограмме. Но и там ситуация распадется на два знаковых класса в зависимости от ориентации векторов.

Чтобы все-таки научиться представлять произведение векторов как вектор, т. е. изображать произведение $a \times b$ на той же самой прямой, на которой мы откладываем исходные векторы a и b , нам потребуется задать масштаб данной прямой, т. е. определить на нем единицу длины, или единичный вектор e . Тогда построить вектор $a \times b$ на прямой мы сможем с помощью известной из геометрии теоремы Фалеса.

Возьмем прямую l и отметим на ней начало отсчета O . Отложим от точки O векторы a и b , получим таким образом точки A и B . Кроме того, зафиксируем масштаб прямой l , отложив на ней единичный вектор $e = OE$. Выберем не лежащую на прямой l произвольную точку C и построим прямую OC . Затем построим две секущие этого угла: EC и параллельную ей прямую AD . После чего построим две новые секущие угла: CB и параллельную ей прямую DF . Точка F является результатом откладывания от точки O вектора, длину которого можно интерпретировать как произведение длин векторов a и b . Это следует из пропорций в подобных треугольниках с общей вершиной O . На рис. 1.5 представлено два случая произведения векторов: для одинаковых знаков и для противоположных.



Одинаковые знаки



Различные знаки

Рис. 1.5. Произведение векторов геометрически.

Как видим, и при такой интерпретации у нас получается смена знака произведения, причем оба варианта согласованы друг с другом.

Таким образом, знак умножения двух чисел определяется знаком (направлением) образующих векторов в соответствии со следующей *таблицей композиции знаков*:

	+	-
+	+	-
-	-	+

Данная таблица демонстрирует нам простейший пример **группы**, к определению которой мы вернемся.

Отметим некоторые свойства этой простейшей таблицы-группы. Во-первых, мы видим, что композиция знаков тоже есть знак, во-вторых, что композиция любого знака с плюсом ничего не меняет, т.е. плюс является нейтральным в операции композиции. В-третьих, что знак

минус сам себе обратен, так как умножение его самого на себя дает плюс. Все эти свойства являются определяющими свойствами математического понятия группы и будут «работать» на протяжении всего курса.

Отметим, что умножение можно иллюстрировать не только площадью, но и объемом, если нам необходимо перемножить три числа. Хотя сама по себе операция умножения бинарная (т. е. имеет два аргумента), один из множителей, в свою очередь, также может быть результатом операции умножения, так что в итоге получается умножение трех чисел.

Например, таким способом можно показать *ассоциативность умножения*, т. е. независимость результата умножения трех чисел от порядка выполнения двух операций умножения (рис. 1.6).

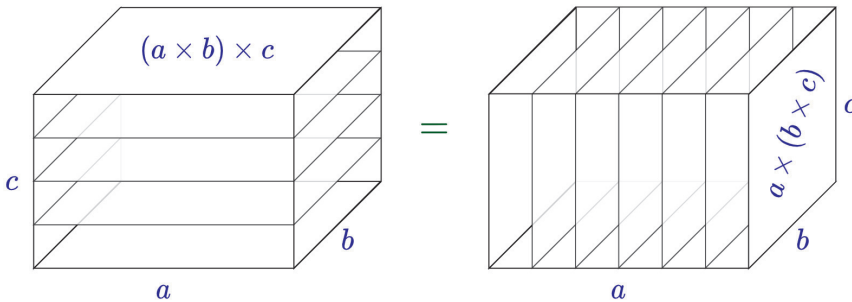


Рис. 1.6. Ассоциативность умножения.

Договоримся теперь масштабный вектор e обозначать более привычным символом — цифрой 1. Кроме того, сумму $1 + 1$ обозначим за 2, сумму $1 + 1 + 1$ — за 3, и т. д. Это позволит нам рассматривать плоскость в масштабе, заданном вектором e , как сетку из квадратов размером 1×1 , что позволит проще вычислять соответствующую площадь.

Теперь мы легко можем получить **свойства умножения**:

P1 $(a \times b) \times c = a \times (b \times c)$ (ассоциативность — рис. 1.6);

P2 $a \times b = b \times a$ (коммутативность — рис. 1.7);

P3 $a \times 1 = 1 \times a = a$ (нейтральный элемент по умножению);

P4 $a \times (b + c) = a \times b + a \times c$ (дистрибутивный закон — рис. 1.8);

P5 $0 \times a = 0$ (мультипликативное свойство нуля), поскольку

$$0 + a \times 0 = a \times 0 = a \times (0 + 0) = (a \times 0) + (a \times 0) \Rightarrow 0 = (a \times 0);$$

P6 если $a \times b = 0$, то $a = 0$ или $b = 0$ (отсутствие делителей нуля);

P7 если $a \times c = b \times c$ и $c \neq 0$, то $a = b$ (правило сокращения), поскольку

$$0 = a \times c - b \times c = (a - b) \times c,$$

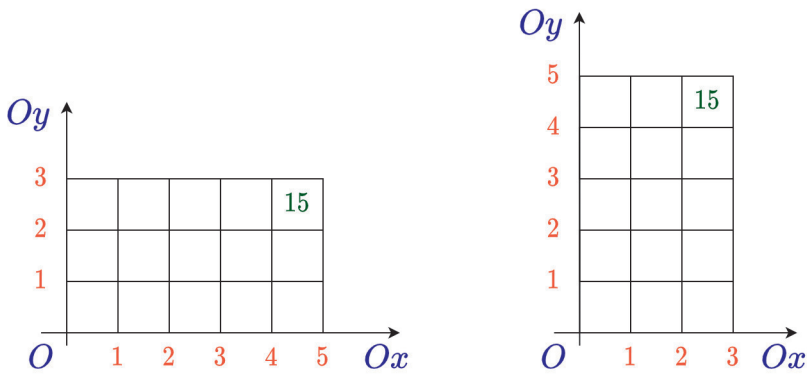


Рис. 1.7. Коммутативность умножения.

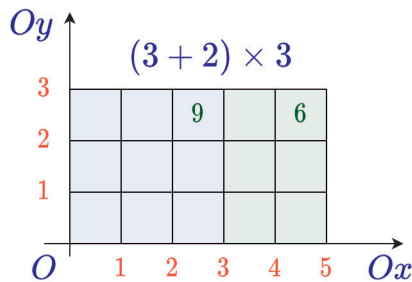


Рис. 1.8. Дистрибутивность операций сложения и умножения.

а так как $c \neq 0$, получаем, что $a - b = 0$, т. е. $a = b$;

P8 если $a < b$ и $c > 0$, то $a \times c < b \times c$, и обратно: если $a \times c < b \times c$ и $c > 0$, то $a < b$ (монотонность — рис. 1.9);

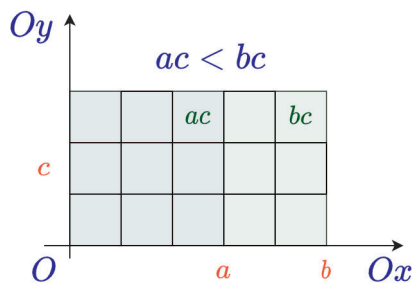


Рис. 1.9. Согласованность порядка с умножением.

Отметим, что такие свойства, как P2, P3, P5, P6, достаточно очевидно

получаются из рис. 1.5, в то время как остальные легче увидеть на объемных иллюстрациях.

Здесь и далее мы часто будем фиксировать некоторый конкретный масштаб на геометрической прямой, принимая за единичный отрезок некоторый заданный вектор. Обычно геометрическую прямую с выбранными на ней точкой отсчета и масштабной единицей принято называть числовой прямой и обозначать за \mathbb{R} . Если же мы рассматриваем плоскость с началом координат и масштабной сеткой, то для нее мы используем обозначение \mathbb{R}^2 . В дальнейшем мы уточним эти понятия и их символику.

1.2. Понятие натурального числа

Отметим, что между арифметическими операциями есть очевидная взаимосвязь, которая является следствием закона дистрибутивности (P4) и введения единицы длины. А именно, многократное сложение $a + a + a + a + \dots$, если воспользоваться графическим представлением, есть просто умножение отрезка a на отрезок длины $1 + 1 + 1 + \dots$, где единицы взяты ровно в том же количестве, в каком встречается a (рис. 1.10).

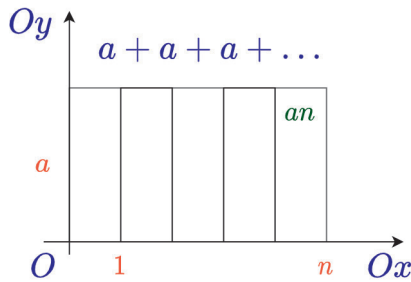


Рис. 1.10. Кратность сложения.

В итоге, если мы число a складываем n раз (где n — это обозначение суммы единиц), то это все равно, что мы производим умножение

$$a \times \underbrace{(1 + 1 + \dots + 1)}_{n \text{ раз}} = a \times n.$$

В случае многократного умножения действуем по аналогии: $aaa\dots$ обозначаем как a^n , понимая под n количество чисел a , которые мы перемножили.

Эта количественная роль чисел, которые выступают для обозначения кратности однотипных операций, и является смыслом натурально-

го числа. Натуральное число есть ответ на вопрос: «сколько раз?» или «сколько штук?»

Целые числа не могут похвастаться таким натуральным (предметным) определением. В целых числах к количеству добавляется еще и направление (назад или вперед, но об этом позже), а рациональное число отражает отношение между количествами.

В математике очень важно давать определения так, чтобы формулы имели единообразный вид во всех случаях. Это делает работу с формулами проще. Поэтому нулевая кратность должна быть определена так, чтобы все законы сложения и умножения продолжали работать, в частности, ассоциативный закон сложения, поэтому сумму из нуля слагаемых мы определим равной нулю:

$$\underbrace{a + \dots + a}_{0 \text{ раз}} = 0.$$

Аналогично поступаем и в случае степени, полагая произведение нулевого количества множителей равным 1:

$$\underbrace{a \times \dots \times a}_{0 \text{ раз}} = 1.$$

Почему именно 1? Потому что если мы начнем перемножать разные числа с разными кратностями

$$aa\ bbbb\ ccccc,$$

то результат не должен меняться, если кратность одной из переменных станет равной нулю. В данном примере любую из кратностей можно сделать нулевой, тогда эта переменная выпадет из записи, и результат останется верным, только если нулевая кратность умножения равна 1.

Многие правила в математике для крайних значений определяются с целью сохранить общий вид формул, если это не приводит к противоречию.

Итак, **натуральные числа** — это показатели кратности операций (сложения и умножения).

С другой стороны, ничто не мешает нам сумму единиц рассматривать саму по себе, т. е. как сумму единичных отрезков. В данном случае нет никакой разницы между суммой единичных отрезков и кратностью сложения единичного отрезка.

Такая интерпретация натурального числа вполне согласуется с операциями сложения и умножения, сохраняет все законы арифметики: ассоциативность, коммутативность, дистрибутивность.

Поэтому натуральные числа, привязанные к единичным отрезкам, можно также считать мерой длины, площади, объема и т. д.

Ноль мы будем считать натуральным числом, поскольку мы рассматриваем нулевую кратность для однородности законов арифметики.

Таким образом, *натуральные числа — это и кратности операций, и мера длины.*

Натуральные числа отвечают за соизмеримость и арифметическую кратность чисел (любых): говорят, что a **кратно** b (обозначение: $a:b$), если $a = bn$ или $a = (-b)n$ при некотором натуральном n .

При этом: ноль кратен любому числу, а нулю кратен только ноль! Действительно, $0:b$ означает, что при некотором n имеем $0 = bn$. Это верно как раз при $n = 0$. Предположим, что какое-то число кратно нулю: $a:0$, тогда при некотором n должно быть $a = 0n$. Но при любом n имеем $0n = 0$, так что только $a = 0$ будет кратно нулю.

Если a кратно b , то говорят также, что b **делит** a (обозначение: $b|a$), или что b является делителем a .

Если $a > 0$ кратно $b > 0$, то $a = kb = b + (k - 1)b$, где $k > 0$. Здесь $x = (k - 1)b$. Поэтому $a \geq b$. Так что для положительных векторов кратность означает превосходство в смысле сравнения. Аналогичные неравенства можно получить и для отрицательных векторов.

В дальнейшем, следуя традиции, мы часто будем пропускать знак умножения, записывая ab вместо $a \times b$, если это не будет вызывать неоднозначного толкования формулы.

1.3. Визуальные доказательства

Продолжая использовать геометрическую интерпретацию для описания действий с числами, покажем ряд примеров того, как с помощью простейших преобразований над геометрическими фигурами можно получить некоторые полезные формулы.

Квадрат суммы. Пусть даны два отрезка длины a и b . Построим квадрат $(a + b) \times (a + b)$ и внутри него квадраты $a \times a$ и $b \times b$ (см. рис. 1.11). Тогда оставшиеся прямоугольники будут иметь площадь $a \times b$, причем таких прямоугольников ровно 2. Приравнивая площадь всего квадрата к сумме его частей, получаем хорошо известную формулу квадрата суммы:

$$(a + b)^2 = a^2 + 2ab + b^2.$$

Теорема Пифагора. Разрежем тот же самый квадрат $(a + b) \times (a + b)$ немного иначе, а именно: половинки прямоугольников $a \times b$, т. е. прямоугольные треугольники с катетами a , b и гипотенузой c , разложим

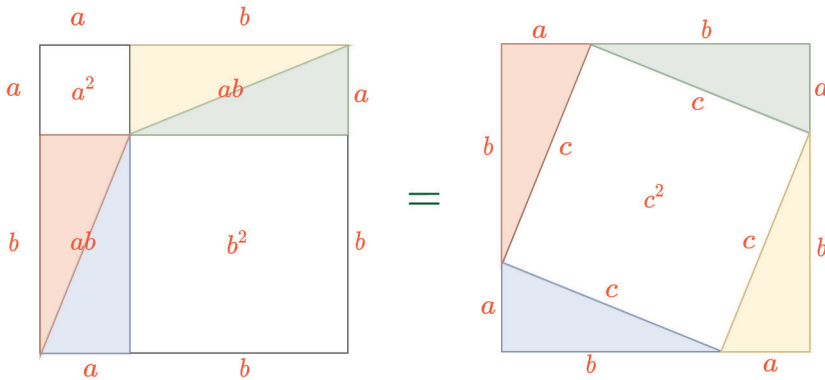


Рис. 1.11. Бином Ньютона и теорема Пифагора.

по углам исходного большого квадрата (рис. 1.11). Далее заштрихуем одинаковые треугольники в первом и втором способе разрезания квадрата $(a + b) \times (a + b)$. Поскольку исходные квадраты равны, равны и их площади, а также площади тех незаштрихованных остатков, которые получатся, если выбросить заштрихованные треугольники. Эти остатки представляют собой в первом случае сумму двух квадратов $a \times a$ и $b \times b$, а во втором случае — квадрат $c \times c$:

$$a^2 + b^2 = c^2,$$

тем самым мы получаем теорему Пифагора.

Разность квадратов. Пусть даны два числа a и b , причем $a > b$. Тогда мы можем построить $a \times a$, а внутри него выделить квадраты $b \times b$ и $(a - b) \times (a - b)$ (рис. 1.12). В результате такого разрезания останется еще два прямоугольника $b \times (a - b)$. После чего сложим вместе квадрат $(b - a) \times (b - a)$ и два прямоугольника $b \times (a - b)$, соединив их по стороне $a - b$. В итоге получится прямоугольник $(a - b) \times (a - b + b + b)$, т. е. $(a - b) \times (a + b)$. С другой стороны, их площадь равна площади квадрата $a \times a$ за вычетом площади квадрата $b \times b$, т. е. $a^2 - b^2$.

Таким образом получаем, что $(a - b)(a + b) = a^2 - b^2$.

Сумма подряд идущих чисел. Найдем сумму $1 + 2 + \dots + n$. Решим эту задачу сначала для случая четного n . Расставим прямоугольники $1 \times k$ лесенкой от самого маленького к самому большому (рис. 1.13), $k = 1, 2, \dots, n$. Далее заметим, что если последний столбик переложить на первый, предпоследний на второй и т. д., то будут получаться одинаковые столбики высотой $n + 1$. Получится $n/2$ прямоугольников длины $n + 1$, так что сумма равна $(n + 1) \times n/2$.

Если же n — нечетное, то применим тот же метод, оставив временно в сторону последнее слагаемое n . Получится $(n - 1)/2$ столбиков высотой

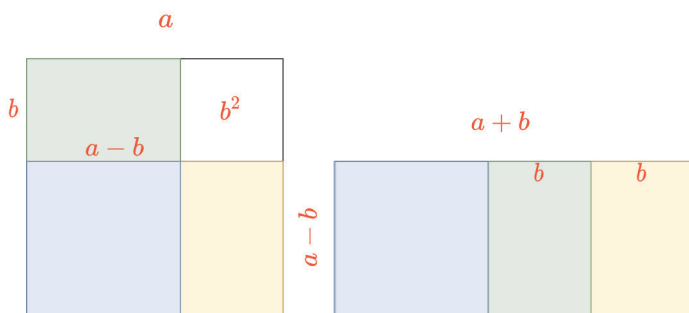


Рис. 1.12. Разность квадратов.

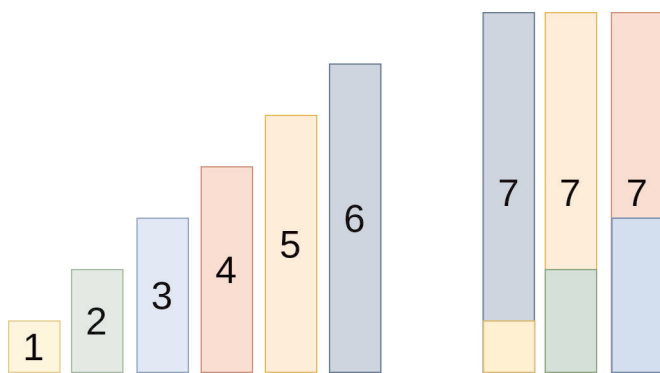


Рис. 1.13. Сумма арифметической прогрессии.

n и плюс еще один столбик высотой n . Итого получаем

$$\frac{n-1}{2} \times n + n = \left(\frac{n-1}{2} + 1 \right) \times n = n \times \frac{n+1}{2}.$$

Здесь, правда, хоть мы и оперируем только натуральными числами, все же нужно иметь некоторое представление о дробях, чтобы от суммы $(n-1)/2 + 1$ перейти к дроби $(n+1)/2$. Но более детальное изучение дробей мы пока отложим.

Бином Ньютона для $n = 3$: Дадим еще одно визуальное доказательство, а именно, получим формулу для бинома Ньютона $(a+b)^n$ в случае $n = 3$ (отметим, что чуть выше мы уже получили ее для случая $n = 2$).

Для этого нам потребуется включить пространственное воображение, взять кубик (например, сделанный из сыра) с ребром $a+b$ и разрезать его тремя взаимно перпендикулярными плоскостями так, чтобы у одного из углов кубика оказался вырезанным кубик размера $a \times a$. В итоге мы получим восемь параллелепипедов, стороны которых легко определить (это будут различные сочетания чисел a и b). См. рис. 1.14.

Таким образом получаем, что $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.

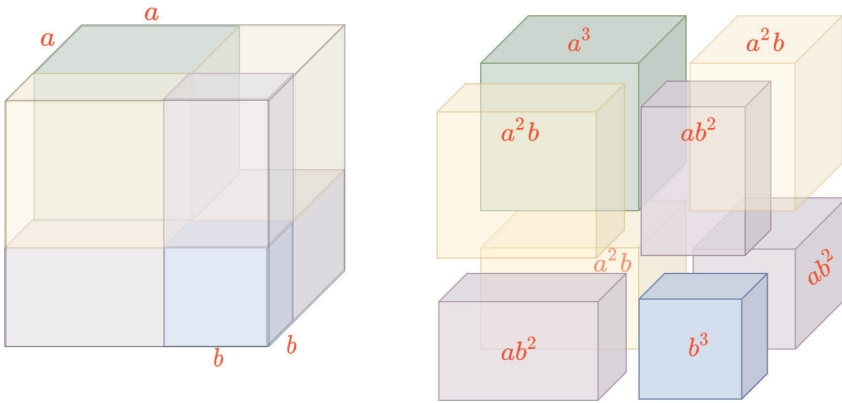


Рис. 1.14. Бином Ньютона.

1.4. Соизмеримость отрезков, алгоритм Евклида

Пусть у нас на числовой прямой сидят в точке O два кузнечика. Один умеет прыгать на длину a как вправо, так и влево, а второй — на длину b как вправо, так и влево.

Могут ли они через какое-то конечное количество прыжков попасть в одну и ту же точку, отличную от O ?

Ответ — да, если есть такая точка A , что отрезок OA кратен и a , и b одновременно, т. е. при некоторых натуральных n, m , не равных нулю, будет верно равенство $an = bm$:

$$\underbrace{a + a + \dots + a}_{n \text{ раз}} = \underbrace{b + b + \dots + b}_{m \text{ раз}}$$

Отрезки, имеющие общий кратный отрезок, называются **соизмеримыми**.

Из равенства $an = bm$ видно, что есть общий отрезок $c = a/m = b/n$, который целое число раз укладывается как в a , так и в b . Такой отрезок c называется **общим делителем** a и b . Максимальный положительный общий делитель a и b называется **наибольшим общим делителем** a и b , он обозначается $\text{НОД}(a, b)$. Ясно, что общий делитель существует тогда и только тогда, когда существует общий кратный отрезок.

Действительно, если общий делитель c существует, то он, являясь делителем a и b , имеет представление $c = a/m = b/n$ при некоторых положительных натуральных n, m , откуда $an = bm$, и отрезок OA длины an является искомым. Обратно, если существует отрезок OA , одновременно кратный как a , так и b , то имеет место представление $OA = an = bm$, откуда число $c = a/m = b/n$ является делителем как a , так и b , а значит, их общим делителем.

Может возникнуть вопрос: как найти отрезок c такой, что $ct = a$? Забегая вперед, скажем, что деление произвольного отрезка на равные части — геометрическая задача. Имея произвольную меру длины, мы можем применить теорему Фалеса для деления имеющегося отрезка на равные части (см. рис. 1.15).

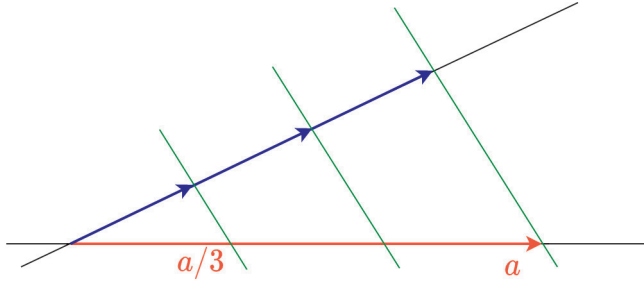


Рис. 1.15. Построение целой доли отрезка.

На этот отрезок можно выйти другим путем: строим прямоугольник $a \times b$ (предполагая, что $a < b$) и начинаем отсекаем в нем квадраты. Сначала отсекаем квадраты $a \times a$, пока можем, останется кусок $a \times b_1$ ($b_1 < a$), затем отсекаем квадраты $b_1 \times b_1$, пока можем, останется кусок $a_1 \times b_1$ ($a_1 < b_1$) и т. д. — см. рис. 1.16.

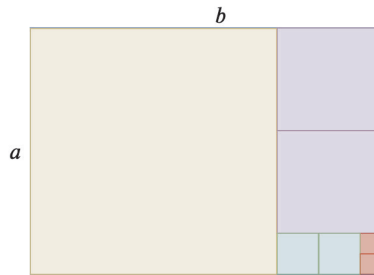


Рис. 1.16. Вычисление НОД(a, b).

Если исходные отрезки соизмеримы, то процесс остановится: исходный прямоугольник будет разбит на конечное число квадратиков. При этом можно заметить, что самые мелкие квадратик целое число раз укладываются во все более крупные просто по построению. Это значит, что прямоугольник $a \times b$ можно разбить на конечное число маленьких квадратиков, общее число которых будет nm штук.

Действительно, соизмеримость a и b означает, что стороны прямоугольника можно разбить на целое число частей длины $c = a/m = b/n$, а процесс разрезания этого прямоугольника на квадраты всегда будет происходить по сетке с шагом c , причем каждый раз шаг разрезания будет уменьшаться. Но у нас конечное число квадратиков $c \times c$, стало быть, процесс разрезания рано или поздно упрется в прямоугольник со

стороной c , который уже будет разделен на целое число квадратов без остатка. На этом алгоритм и закончится.

Длина стороны финального квадратика будет иллюстрировать НОД отрезков a и b , т. к. это максимальный квадрат, которым можно замостить прямоугольник $a \times b$.

Действительно, любой квадрат, которым можно замостить прямоугольник $a \times b$, целое число раз укладывается в квадрат $a \times a$ и, как следствие, в оставшийся прямоугольник $a \times (b - a)$, а значит, целое число раз укладывается в квадрат $(b - a) \times (b - a)$ и, как следствие, в оставшийся прямоугольник и т. д. То есть если каким-то квадратом можно замостить исходный прямоугольник, то им же можно замостить и финальный квадратик. Следовательно, этот квадратик наибольший из всех таких, которыми можно замостить прямоугольник $a \times b$.

Такой процесс постепенного спуска к НОД называется **алгоритмом Евклида**. К нему мы еще неоднократно вернемся с более формальной точки зрения.

Заметим, что числа a и b при этом вовсе не обязаны быть натуральными. Это — какие-то векторы на прямой. В том числе они могут быть отрицательными (при этом их НОД, если он существует, всегда выбирается со знаком «плюс»).

Упражнения

Обязательные упражнения

1.1° Нанести на прямой метки, соответствующие шагам вправо и влево, считая начальной точкой O , а все шаги равновеликими (т. е. каждый шаг равен выбранной единице длины). Дойти до точки 5, а затем от точки 5 до точки -5 . Записать последовательность шагов с помощью ± 1 , предполагая, что шаг вправо записывается как $+1$, шаг влево записывается как -1 .

1.2° Описать в терминах одномерного путешественника операции сложения: $5 + 3$, $8 - 4$, $3 - 5$, $-2 - 6$. Сколько шагов и в какую сторону он прошел и в каком порядке? Записать в каждом случае путь с помощью ± 1 и расставить скобки, объединяя в них указанные слагаемые.

1.3° Путь — это последовательность единичных шагов, обозначаемых $+1$ (шаг вправо) и -1 (шаг влево). Путь может начинаться в любой точке прямой. Записать пути, соответствующие операциям $-2+7$, $10-5$, $11 - 2 - 4$, $-8 + 3 + 10$.

1.4° Выберем точку O в качестве начала отсчета, затем нанесем на прямую точки, которые получаются в результате отсчета шагов влево и вправо, т. е. точки $\pm 1, \pm 2, \pm 3$ и т. д. Назовем эти точки *целыми*.

а) В какой точке окажется путешественник, если он стартует в точке -3 и проходит путь $4 - 1$? Изобразить графически.

б) В какой точке окажется путешественник, если он стартует в точке 1 и проходит путь $11 - 4 + 7$? Изобразить графически.

1.5° Два пути назовем *эквивалентными*, если, стартуя в одной и той же точке, они и закончатся в одной и той же точке. Эквивалентны ли пути $-2 + 7, 10 - 5, 11 - 2 - 4, -8 + 3 + 10$?

1.6° Путь a назовем *обратным* к пути b , если, стартовав там, где путь b заканчивается, он повторяет все шаги пути b в обратном порядке и с противоположным знаком (например, путь $1 + 1 + 1 - 1 - 1 - 1$ обратен к пути $-1 - 1 - 1 + 1 + 1 + 1$). Построить пути, соответствующие операциям $5 + 3, 8 - 4, 3 - 5, -2 - 6$, построить обратные к ним пути, выразить обратные пути в виде суммы или разности двух чисел (использовать те же цифры, что у исходного пути).

1.7° Изобразить ориентированные площади, соответствующие произведениям 3×5 и $5 \times 3, (-2) \times 6$ и $6 \times (-2), (-3) \times (-4)$ и $(-4) \times (-3)$.

1.8° Найти НОД(10, 6), НОД(11, 5), НОД(12, 9) методом прямоугольников.

1.9° Сколько и каких шагов должны сделать 10- и 6-шаговые кузнечики, чтобы попасть в точку НОД(10, 6)?

1.10° Доказать, что a и b соизмеримы тогда и только тогда, когда существует отрезок d такой, что отрезки a и b укладываются в нем целое число раз: $d = ka = lb$. Верно ли, что это также равносильно тому, что два путешественника могут встретиться в какой-то точке прямой, отличной от точки O ?

1.11° Верно ли, что отрезки a и b соизмеримы тогда и только тогда, когда a и $2b$ соизмеримы?

1.12° Сколько и каких квадратов получится в результате применения графического алгоритма Евклида к прямоугольнику со сторонами 75 и 21? А со сторонами 324 и 141?

1.13° Применяя операцию Евклида, прямоугольник разрежали на большой квадрат, два квадрата поменьше и два совсем маленьких. Найти отношение сторон исходного прямоугольника.

1.14° Доказать, что если стороны прямоугольника соизмеримы, то, применяя операцию Евклида, мы в конце концов разрежем его на квадраты (применить метод бесконечного спуска).

1.15° Доказать, что если применение графического алгоритма Евклида разрезает прямоугольник на некоторое конечное число квадратов, то стороны прямоугольника соизмеримы, и сторона самого маленького квадрата будет их наибольшей общей мерой.

1.16° Доказать, что любая общая мера соизмеримых отрезков a и b целое число раз укладывается в их наибольшей общей мере.

1.17° От прямоугольника отрезали квадрат и получили прямоугольник, подобный исходному. Соизмеримы ли стороны исходного прямоугольника? Чему равно отношение его сторон?

1.18° Докажите, что $\text{НОД}(a, b)$ существует и единственный, если целые a и b не равны одновременно нулю.

1.19° Докажите, что $\text{НОД}(a, b) = \text{НОД}(a - b, b) = \text{НОД}(r, b)$, где r — остаток от деления a на b .

1.20° Найдите наибольшую общую меру отрезков $15/28$ и $6/35$.

1.21° Какие расстояния можно отложить на прямой, имея шаблоны 6 см и 15 см?

1.22° Найдите возможные значения: а) $\text{НОД}(n, 12)$; б) $\text{НОД}(n, n + 1)$; в) $\text{НОД}(2n + 3, 7n + 6)$; д) $\text{НОД}(n^2, n + 1)$.

1.23° Найти с помощью графического метода сумму подряд идущих нечетных чисел от 1 до n , где n — нечетное.

1.24° Рассмотрим последовательность уголков. а) Сколько клеток в k -м уголке? б) Чему равна суммарная площадь первых k уголков? в) Чему

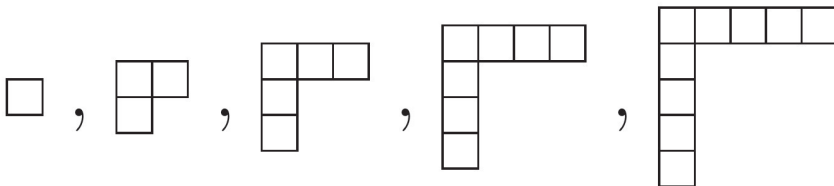


Рис. 1.17. Последовательность уголков.

равно k -е нечетное число и сумма первых k нечетных чисел? д) Чему равно k -е четное число и сумма первых k четных чисел? е) Вычислите сумму 100 последовательных нечетных чисел, начиная со 179.

1.25° Треугольные числа Диофанта (рис. 1.18) обозначим по порядку T_1, T_2, T_3, T_4 и т. д.

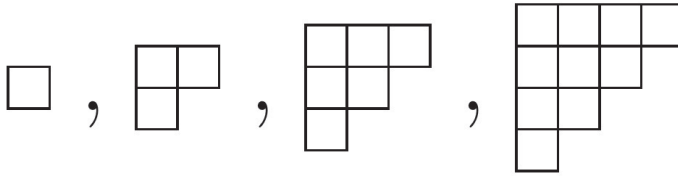


Рис. 1.18. Треугольные числа Диофанта.

- a) Сложите из двух последовательных треугольных чисел квадрат.
- b) Что получится при сложении T_n с T_n ?
- c) Выразив T_n через n , найдите $1 + 2 + \dots + n$.
- d) Докажите геометрически, что $T_{n+m} = T_n + T_m + nm$.

Принцип Дирихле: если в n клетках сидит не менее $n + 1$ кроликов, то найдется клетка, в которой сидит не менее двух кроликов.

1.26° В коробке 50 конфет трех видов. Докажите, что конфет какого-то вида не менее 17.

1.27° Какое наименьшее количество учеников должно быть в школе, чтобы гарантированно можно было найти трех учеников, отмечающих день рождения в один день?

1.28° Докажите, что существуют две различные степени семерки, оканчивающиеся на одну и ту же комбинацию из трех цифр.

1.29° Докажите *обобщенный принцип Дирихле*: если в n клетках сидит не менее k кроликов, то найдется клетка, в которой сидит не менее k/n кроликов. Как следует понимать это утверждение, если k не делится на n нацело?

1.30° Докажите, что если в n клетках сидит менее $\frac{n(n-1)}{2}$ кроликов, то найдутся две клетки, в которых сидит одинаковое количество кроликов (может быть, ни одного).

1.31° В коробке 70 карандашей. Докажите, что найдутся либо 9 карандашей одного цвета, либо 9 карандашей разных цветов.

1.32° В кинотеатре 7 рядов по 10 мест каждый. Группа из 50 детей сходила на утренний сеанс, а потом на вечерний. Докажите, что найдутся двое детей, которые на утреннем сеансе сидели в одном ряду и на вечернем тоже сидели в одном ряду.

Дополнительные упражнения

1.33 Выведите формулу суммы геометрической прогрессии $1 + x + x^2 + x^3 + \dots$ ($0 < x < 1$) путем домножения этой суммы на x . Найти:

- a) $\frac{1}{10} + \frac{1}{100} + \frac{1}{1000} + \dots$;
 b) $1 + 0.2 + (0.2)^2 + (0.2)^3 + \dots$;
 c) $\frac{1}{0.99} + \frac{1}{0.99^2} + \frac{1}{0.99^3} + \dots$

1.34 Исследовать ряды на сходимость:

- a) $1 + 1/3 + 1/5 + 1/7 + \dots$;
 b) $1 + 1/3^2 + 1/5^2 + 1/7^2 + \dots$;
 c) $\frac{1}{1001} + \frac{1}{2001} + \frac{1}{3001} + \dots + \frac{1}{1000n+1} + \dots$;
 d) $1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots$;
 e) $1 + \frac{2}{3} + \frac{3}{5} + \frac{5}{9} + \dots + \frac{n}{2n-1} + \dots$

1.35 Доказать, что если ряды $\sum_n a_n^2$ и $\sum_n b_n^2$ сходятся, то сходятся также и ряды:

$$\sum_n a_n b_n, \quad \sum_n (a_n + b_n)^2.$$

Здесь все $a_n, b_n \geq 0$.

1.36 Доказать сходимость ряда

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

где $0 \leq a_n < 10$.

1.37 Докажите геометрически, что $1 + \dots + (n-1) + n + (n-1) + \dots + 1 = n^2$.

1.38 Получите геометрически выражение для: a) $(a+b+c)^2$; b) $(a+b+c)^3$.

1.39 Объясните равенство на рис. 1.19 и получите формулу для суммы квадратов $1^2 + 2^2 + 3^2 + \dots + n^2$.

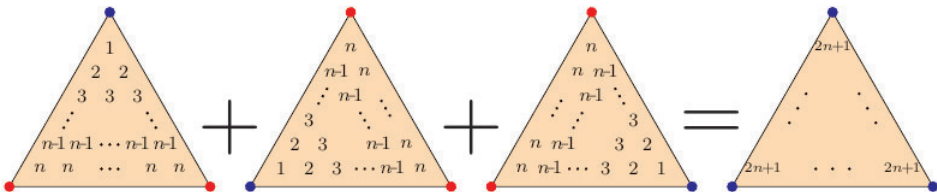


Рис. 1.19. Сумма квадратов – 1.

1.40 В соревнованиях по бегу участвуют 100 спортсменов. Известно, что среди любых 12 из них найдутся двое знакомых между собой. Докажите, что как бы ни раздали спортсменам стартовые номера (не обязательно от 1 до 100), найдутся два знакомых спортсмена, номера которых начинаются с одной и той же цифры.

1.41 Для награждения по итогам школьного конкурса имеется 70 конфет. При каком наибольшем количестве конкурсантов им можно будет раздать конфеты так, что все они получают разное и не меньшее 3 количество конфет?

1.42 На заводе 7 цехов, в которых работает 360 человек. Докажите, что в каких-то пяти из этих цехов работает не менее 258 человек.

1.43 Пять мальчиков собрали 53 гриба, причем известно, что никакие двое не собрали грибов поровну. Докажите, что какие-то трое из них собрали не менее 36 грибов.

1.44 Числа от 1 до 9 некоторым образом разбиты на три группы. Докажите, что произведение чисел в одной из групп не меньше 72.

1.45* С помощью рис. 1.20 получите еще один способ найти формулу для суммы квадратов.

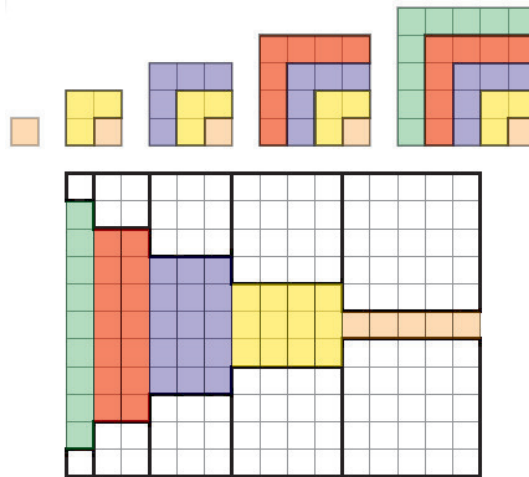


Рис. 1.20. Сумма последовательности квадратов.

1.46* Можно ли 100 гирь массами 1, 2, 3, ..., 99, 100 разложить на 10 кучек разной массы так, чтобы выполнялось условие: чем тяжелее кучка, тем меньше в ней гирь?

1.47* У 10 девочек было по 10 конфет. Каждая девочка подарила несколько конфет другим (конфеты, полученные в подарок, девочки

оставляют себе). В результате у всех девочек оказалось разное число конфет. Докажите, что какая-то из девочек подарила конфет не меньше, чем у нее их оказалось в конце.

1.48* На складе имеется несколько ящиков общей массой 10 тонн, причем масса каждого не превосходит тонны. Какое наименьшее количество трехтонок нужно заказать, чтобы точно суметь вывезти их все за один раз?

Движения прямой

Аннотация

В этой главе мы переходим к более формальной работе с точками и векторами на прямой. Целью является знакомство с понятиями «движение», «композиция движений». Проводится полный анализ видов движений и свойств их композиций.

Попутно вводится понятие группы и подгруппы в приложении к группе движений на прямой. Изучаются все конечные подгруппы движений прямой.

2.1. Сдвиг, композиция сдвигов, группа

Иллюстративная сказка. Представим себе очень длинную однорядную автомобильную парковку на территории какого-нибудь бизнес-центра. На этой парковке размечены места номерами 0, 1, 2 и т.д. (слева направо). В какой-то момент парковку достроили влево и решили, не мудрствуя лукаво, продолжить нумерацию отрицательными числами -1 , -2 , -3 и т.д. Получилась шкала примерно как на градуснике для измерения уличной температуры. Водителям, работающим в этом бизнес-центре, выдали парковочные талоны с номерами парковочных мест, т.е. такие же числа 0, ± 1 , ± 2 и т.д. В соответствии с талонами они занимают свои места, так что получается, что водитель с талоном номер 0 встает на место номер 0, водитель с талоном номер 1 — на место номер 1 и т.д.

Но потом появляется необходимость поменять бордюр и плитку там, где находятся два крайних левых парковочных места, пусть это будут номера -3 и -2 . Возникает потребность куда-то девать те автомобили, для которых зарезервированы номера -3 и -2 . Вместо того чтобы предложить обменять талоны -3 и -2 на резервные номера парковки, начальнику охраны приходит в голову гениальная идея: повесить на въезде плакат с надписью (красным фломастером на А4): «Внимание! Занимайте номер парковки на 2 больше, чем указан в вашем талоне!!»

Так что водитель, имеющий парковочный талон номер -3 , занимает место -1 , номер -2 — место 0, номер -1 — место 1 и т.д.



Рис. 2.1. Иллюстрация сдвига на прямой.

Иначе говоря, все автомобили должны теперь вставать на 2 места правее, т. е. произвести сдвиг относительно своего обычного места, указанного в парковочных талонах (см. рис. 2.1).

Отметим еще одну особенность истории с парковкой: несмотря на произведенное перемещение автомобилей, они по-прежнему остаются на парковке, не занимая места где-либо еще, например на проезжей части или тротуаре. Просто потому, что запас мест справа оказался достаточным для данной манипуляции.

А что если бы парковка была неограниченно расширяемой в обе стороны автоматически всякий раз, когда не хватает места? Как говорят математики, она была бы потенциально бесконечной.

Геометрически мы можем представить это так: у нас имеется прямая, на которой нанесена разметка числами $0, \pm 1, \pm 2$ и т. д. через равные расстояния между соседними точками. Прямая — бесконечная в обе стороны. Но вдруг возникает необходимость сдвинуть эту прямую вправо на 2 единицы. Для этого всем точкам прямой дается команда сдвинуться на вектор длины 2 вправо.

Однако, чтобы сохранить историю этого сдвига и проверить его правильность, следует сдвигать не саму прямую, а ее копию. В итоге мы получаем прямую-оригинал и прямую-образ. Если уж быть совсем точными, то у нас возникает то, что в математике называется функцией, т. е. соответствие между оригинальными точками и их точками-образами на копии прямой.

Так мы видим не только новое положение точек, после сдвига, но и как оно соотносится с прежним их положением!

Понятно, что сдвигать геометрическую прямую, не выходя за ее пределы, можно только вправо или влево, причем на вектор произвольной

длины, не обязательно на число 2, или 3, или им подобное. Тем более что цифровую разметку на прямую можно и вовсе не наносить.

Преобразование, состоящее в том, что все точки прямой сдвигаются на вектор a , называется **сдвигом** на вектор a . При этом, чтобы узнать, в какую точку (относительно исходной разметки) перейдет точка A , нужно от точки A отложить вектор a , т. е. найти сумму $A + a$. Это будет некоторая точка A' на этой же прямой (см. рис. 2.2).

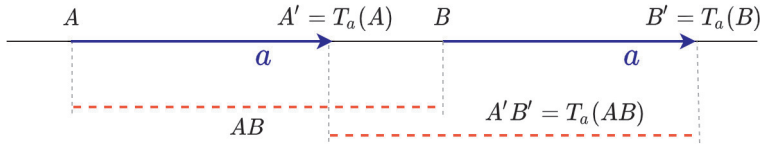


Рис. 2.2. Сдвиг прямой на вектор a .

Преобразование сдвига на вектор a обозначим T_a , а его действие на точку A обозначим $T_a(A)$. Так что

$$T_a(A) = A', \quad a = \overrightarrow{AA'}.$$

Можно говорить также о сдвиге не одной точки, а целого семейства точек. Например, если у нас изначально на прямой был отмечен отрезок AB , то результатом его сдвига на вектор a будет отрезок $T_a(AB) = A'B'$ (см. рис. 2.2). Поскольку все точки отрезка AB сдвинулись одновременно в одном и том же направлении на один и тот же вектор, то легко понять, что результатом будет отрезок той же самой длины в силу свойств сложения векторов.

Сдвиг является движением (не случайно это однокоренные слова!).

В математике существует более общее и строгое понятие, описывающее преобразования подобного рода. А именно, **движением** называется преобразование (прямой, плоскости и т. д.¹), сохраняющее расстояния: если между точками A и B было расстояние x , то после преобразования движения расстояние между точками A' и B' , в которые перешли исходные точки, тоже будет x , и так для любой пары точек! Для сдвига это очевидно, поскольку ко всем точкам прибавляется один и тот же вектор. Иначе говоря, то самое свойство инвариантности «размера» или длины отрезка при сдвиге, о котором мы говорили чуть выше, берется в качестве определения понятия движения в общем случае.

Математическое движение — это результат физического движения (есть только начальное и конечное состояние системы).

¹ Под «и т. д.» мы спрятали понятие *метрического пространства*, т. е. пространства с заданной на нем функцией расстояния между точками.

Сдвиг характерен тем, что он в качестве параметра имеет только вектор, т. е. величину и направление сдвига, но он никак не связан с исходной разметкой прямой!

Всевозможные сдвиги сами по себе можно рассматривать как некоторые математические объекты, над которыми можно производить определенные операции. Точнее, предположим, что у нас задано два сдвига T_a и T_b на векторы, соответственно, a и b . Мы можем применить их последовательно: взяв произвольную точку A на прямой, сначала сдвинем ее на вектор a , получим точку A' , а затем полученную точку A' сдвинем на вектор b , получим точку A'' . См. рис. 2.3.

Вопрос: какое преобразование переводит точки A в точки A'' и как оно зависит от исходных преобразований T_a и T_b ?

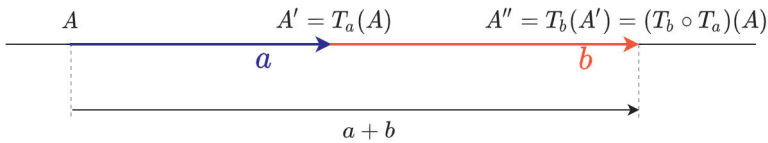


Рис. 2.3. Композиция сдвигов.

Последовательное применение преобразований так, что каждое следующее применяется к результатам предыдущего, называется **композицией преобразований** и обычно записывается при помощи символа-связки \circ (в порядке *справа налево!*). В нашем примере, где мы сначала применили сдвиг T_a , а затем сдвиг T_b , в итоге мы применили композицию $T_b \circ T_a$. Иначе говоря, для произвольной точки A на прямой образом ее сдвига является

$$(T_b \circ T_a)(A) = T_b(T_a(A)). \quad (2.1)$$

Композиция сдвигов, очевидно, соответствует сумме векторов:

$$T_b \circ T_a = T_{a+b}.$$

Более того, поскольку сдвиг задается вектором, а композиция сдвигов — суммой векторов, можно заметить, что все сдвиги прямой — это практически то же самое, что и сами векторы, на которые производятся эти сдвиги. Достаточно T_a заменить на a и \circ заменить на $+$, как все выражения для сдвигов превращаются в выражения для векторов и полностью повторяют их свойства.

В частности, композиция сдвигов перестановочна в силу коммутативности сложения векторов:

$$T_b \circ T_a = T_a \circ T_b.$$

Кроме того, композиция сдвигов ассоциативна, т. е. если мы имеем последовательность из трех и более сдвигов, то мы можем начать вы-

числять ее с любого места цепочки, постепенно сворачивая выражение, как с обычными числами:

$$T_a \circ T_b \circ T_c = (T_a \circ T_b) \circ T_c = T_a \circ (T_b \circ T_c),$$

т.е. сначала вычислить композицию последних двух, и результат подставить в первую, либо же наоборот — сначала вычислить композицию первых двух, и результат применить к последней. Это правило можно тиражировать на цепочку композиций любой длины. Результат при этом будет один и тот же, совершенно так же, как если бы мы складывали подряд несколько чисел.

Кратность сдвига (т.е. кратная композиция одного сдвига с самим собой) обозначается как степень

$$\underbrace{T_a \circ \dots \circ T_a}_n = T_a^n$$

и соответствует кратности сложения или сдвигу на вектор, получаемый умножением исходного вектора на степень кратности: $T_a^n = T_{an}$.

Есть, очевидно, и нулевой сдвиг $T_0 = \text{id}$ — это **тождественное преобразование**, которое ничего не меняет.

Наконец, для каждого сдвига T_a существует обратный сдвиг T_a^{-1} — это сдвиг на вектор $-a$, т.е. сдвиг в обратном направлении на ту же величину.

Нулевой сдвиг сам себе обратен.

Все перечисленные свойства говорят нам о том, что у исходных векторов и у сдвигов прямой есть нечто общее, проявляющееся в их «алгебре», т.е. в операциях над векторами и сдвигами.

Отсюда мы выходим на одно из основных понятий современной математики — понятие группы. Это такое множество G с заданной на нем одной бинарной операцией \circ , для которой выполняются следующие аксиомы.

G1 Результат групповой операции снова лежит в этом же множестве (например, композиция сдвигов есть сдвиг):

$$u, v \in G \Rightarrow u \circ v \in G.$$

G2 Групповая операция **ассоциативна** (сочетательный закон): для любых элементов u, v, w группы G :

$$(u \circ v) \circ w = u \circ (v \circ w)$$

(например, $(T_a \circ T_b) \circ T_c = T_a \circ (T_b \circ T_c)$).

G3 Существует **нейтральный элемент** id такой, что для любого элемента u имеет место равенство

$$u \circ \text{id} = u = \text{id} \circ u.$$

G4 Групповая операция **обратима**: для всякого элемента u существует обратный ему элемент v такой, что

$$u \circ v = \text{id} = v \circ u$$

(например, обратный сдвиг — это сдвиг в противоположную сторону: $T_a^{-1} = T_{-a}$). Элемент v в таком случае обозначается как u^{-1} и называется **обратным** к элементу u .

Нетрудно видеть, что множество всех сдвигов с операцией композиции, а также множество всех векторов с операцией сложения образуют группы.

Мало того, группа сдвигов (как и группа векторов) **коммутативна** (абелева), т. е. для ее групповой операции выполняется переместительный закон:

G5 $u \circ v = v \circ u$ для всех u, v из группы G .

Для удобства записи мы можем также определить кратность обратного сдвига:

$$T_a^{-n} = (T_a^{-1})^n = T_{-a}^n = T_{-an} = (T_{an})^{-1}.$$

Кроме того, мы можем вычислить обратный сдвиг к композиции сдвигов:

$$(T_a \circ T_b)^{-1} = T_b^{-1} \circ T_a^{-1}.$$

Это легко следует из общего свойства группы:

$$(u \circ v)^{-1} = v^{-1} \circ u^{-1}, \text{ т. к. } (u \circ v) \circ (v^{-1} \circ u^{-1}) = u \circ (v \circ v^{-1}) \circ u^{-1} = \text{id}.$$

На основе только одного сдвига T_a можно построить группу сдвигов специального вида, используя понятия кратности и обратного сдвига:

$$\langle T_a \rangle = \{T_a^n, T_a^{-n} \mid n = 0, 1, 2, \dots\}.$$

Такие группы, которые порождаются каким-то одним элементом с помощью его многократных композиций и операции обращения, называются **циклическими**.

Циклическая группа $\langle T_a \rangle$ является своего рода реализацией целых чисел \mathbb{Z} , к которым мы еще вернемся. С другой стороны, очевидно, что $\langle T_a \rangle$ — подмножество множества всех сдвигов, т. е. часть более широкой группы. Отсюда возникает понятие подгруппы.

Подгруппа — непустое подмножество группы, на котором групповая операция удовлетворяет групповым аксиомам, т. е. подгруппа сама является группой с той же операцией, которая задана в группе.

Подгруппу можно определить иначе: непустое подмножество $H \subseteq G$ группы G называется подгруппой, если для любых $a, b \in H$ имеет место $a \circ b^{-1} \in H$.

Оба определения эквивалентны. Действительно, если H удовлетворяет первому определению, т. е. само является группой, то требование $(a, b \in H) \rightarrow (a \circ b^{-1} \in H)$ выполняется по определению группы автоматически. Обратно, если данное требование выполнено, то, очевидно, что единица группы G принадлежит H ($a \circ a^{-1} \in H$), обратный элемент к элементу $b \in H$ также принадлежит H , т. к. $\text{id} \circ b^{-1} \in H$, ассоциативность операции наследуется от группы G , замкнутость операции $(a, b \in H) \rightarrow (a \circ b \in H)$ проверяется непосредственно, если воспользоваться тем, что $a \circ b = a \circ (b^{-1})^{-1}$.

Каждый сдвиг T_a порождает (с помощью его многократной композиции и взятия обратного сдвига) свою подгруппу в группе всех сдвигов.

Каждый вектор a порождает (с помощью его многократного суммирования и взятия обратного вектора) свою подгруппу в группе всех векторов.

Вспомним также таблицу знаков умножения, которую мы рассматривали, изучая произведение векторов:

	+	-
+	+	-
-	-	+

Если рассматривать знаки как элементы множества $G = \{+, -\}$ и в качестве операции над ними взять умножение, определенное данной таблицей, то мы получим пример группы из двух элементов, которая является коммутативной (впрочем, любая двухэлементная группа коммутативна).

2.2. Отражение

Продолжим нашу историю с движением автомобилей. Пусть на сей раз вместо парковки они готовятся к параду и должны занять свои места в ряду с номерами $0, \pm 1, \pm 2$ и т. д. Номера мест нанесены, как и ранее, на асфальт и представляют собой один ряд. У водителей автомобилей есть предписания, в которых указано, какие места нужно занять. Следующим шагом предписание является команда обменяться местами так, чтобы порядок автомобилей сменился на противоположный. Иначе

говоря, каждому нужно проехать полукруг и занять место, симметричное относительно заданного. Например, центром симметрии и соответствующих полукругов является место 0. Тогда автомобиль, стоящий на 1-м месте, должен переехать на место -1 , стоящий на 2-м месте, — на место -2 и т. д.

В результате мы получим перестроение на параде, при котором автомобили опишут полукруги и встанут в обратном порядке, причем нулевой автомобиль сохранит свое место (см. рис. 2.4).

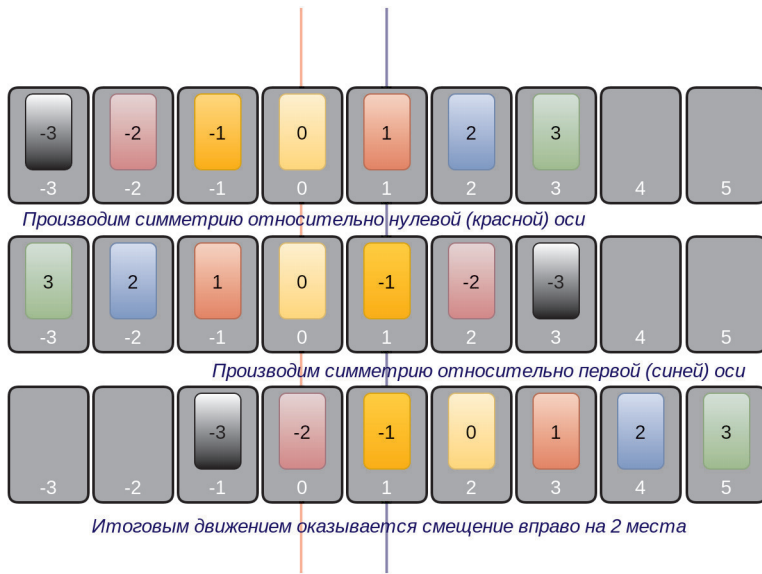


Рис. 2.4. Иллюстрация композиции двух отражений.

Заметим, что и в этом случае расстояние между автомобилями сохранится: как было ранее два автомобиля между 0-м и 3-м, так и останется. И так для всех пар автомобилей.

Такой вид движений называется **отражением**. Вся наша линейная парковка отразилась относительно нулевого места.

Особо отметим, что физически отражение всегда требует выхода за пределы исходной фигуры. Если сдвиг мы могли осуществить, находясь внутри парковочной сетки (предположим, что мы имеем дело с танками, которые могут на месте повернуться на 90 градусов, произвести перемещение, а затем развернуться снова, либо что имеем дело с параллельной парковкой, где автомобили стоят вдоль направления нумерации), то отражение никак не возможно выполнить, оставаясь в пределах исходной парковки — потребуется выезд на проезжую часть.

Отражение на геометрической прямой — то же самое. Сначала мы должны выбрать центр отражения, который останется на месте, затем перевернуть прямую в обратном направлении (снова имеем выход во

внешнее пространство, если представить отражение как физический процесс!).

Отражение с центром в точке O будем обозначать S_O . Отражение можно представить как огромное количество сдвигов, выполняемых одновременно. Для каждой точки — свой сдвиг, причем на разные векторы для разных точек. Так, в результате действия отражения S_O на точку A мы получим точку $A' = T_a(A)$, где вектор $a = 2\overrightarrow{AO}$. То есть мы производим сдвиг на расстояние $2|OA|$, только в противоположную от A сторону.

Как и в случае со сдвигом, отражение «помнит» исходную разметку прямой, а значит, мы всегда можем сказать, какая точка откуда пришла в свое новое состояние.

Отражение, в отличие от сдвига, намертво привязано к одной выделенной точке на прямой в *исходной разметке*, и полностью ею определяется! Мы можем рассмотреть два и более отражений, но все они должны быть заданы в одной исходной разметке прямой, чтобы не возникла путаница.

Композиция отражений и композиция отражения и сдвига определяются аналогично композиции сдвигов (см. (2.1)):

$$(S_O \circ S_A)(x) = S_O(S_A(x)), \quad (S_O \circ T_a)(x) = S_O(T_a(x)), \quad (T_a \circ S_O)(x) = T_a(S_O(x)),$$

т. е. применение операций выполняется справа налево.

Отражение обратно самому себе: $S_O \circ S_O = \text{id}$, т. е. $S_O^{-1} = S_O$.

В терминах парада, показанного на рисунке 2.4, все отражения задаются относительно разметки мест на асфальте! В этом случае водителям для выполнения операции отражения не нужно знать, где какие номера автомобилей находятся, им достаточно видеть номер своего парковочного места, знать номер места — центра симметрии, и выполнить перемещение на удвоенное расстояние, чтобы занять противоположное место.

Поэтому композиция двух отражений, т. е. их последовательное применение, легко вычисляется. Действительно, пусть точка A под действием отражения с центром O переходит в точку A' , а та, в свою очередь, под действием отражения с центром C переходит в точку A'' — см. рис. 2.5.

Нарисуем три дуги: первая с центром O и радиусом AO (синий пунктир на рис. 2.5), вторая с центром C и радиусом $A'C$ (черная линия), третья с центром O и радиусом OC (красный пунктир). Поскольку мы имеем две концентрические дуги, зазоры между их концами симметричны и равны по длине, т. е. отрезки CA' и AX равны. С другой стороны, равны и отрезки CA' и $A''C$. Следовательно, отрезок XC переходит

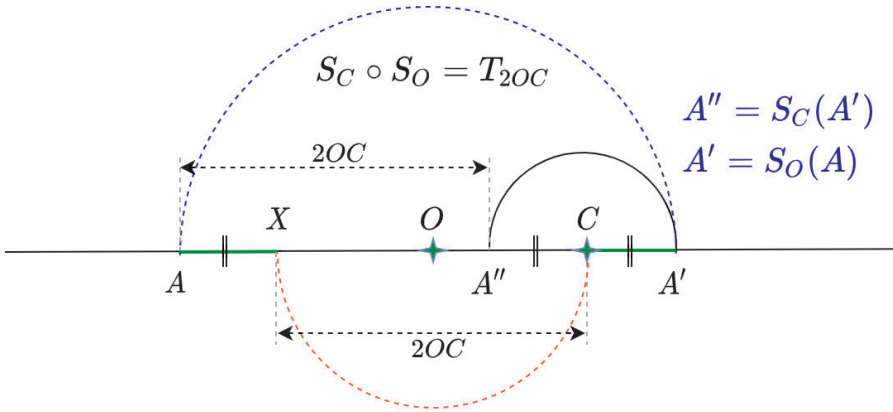


Рис. 2.5. Композиция двух отражений.

в отрезок AA'' сдвигом на расстояние, равное длине отрезка AX . Но длина XC равна, очевидно, $2OC$. Значит, $A'' = T_{2OC}(A)$.

Поскольку эти рассуждения верны для любой начальной точки A , мы получаем, что $S_C \circ S_O = T_{2OC}$. Если вспомнить общее групповое правило $(u \circ v)^{-1} = v^{-1} \circ u^{-1}$, то легко получить также, что

$$S_O \circ S_C = S_O^{-1} \circ S_C^{-1} = (S_C \circ S_O)^{-1} = T_{2OC}^{-1} = T_{-2OC} = T_{2CO}.$$

Таким образом, имеют место следующие равенства для композиции отражений:

$$S_O \circ S_C = T_{2CO}, \quad S_C \circ S_O = T_{2OC}. \quad (2.2)$$

Заметим, что композиция отражений является сдвигом и при этом не коммутативна! То есть отражения, производимые в разной последовательности, приводят, вообще говоря, к разным результирующим сдвигам, а именно — к противоположным.

2.3. Таблица композиций движений прямой

При изучении различных преобразований в математике очень важным элементом исследования является установление таблицы композиций различных видов преобразований. Простейшими примерами здесь являются таблицы сложения и умножения чисел, которые мы учим в школе. Аналогичное знание нам бы хотелось получить и для движений прямой.

Ранее мы уже выяснили, что получается при композиции двух сдвигов (тоже сдвиг, но на суммарный вектор) и при композиции двух отражений (снова сдвиг, причем на вектор, равный удвоенному сдвигу от центра первого отражения к центру второго).

Дальнейшей нашей программой будет выяснение результата разнородных композиций, т. е. композиции отражения и сдвига, а также полное описание всех движений прямой.

Композицию отражения и сдвига очень просто получить из ранее выведенных уравнений для композиции двух отражений:

$$S_O \circ T_a = S_{O-a/2}, \quad T_a \circ S_O = S_{O+a/2}. \quad (2.3)$$

Это легко проверить, если вместо a подставить $2CO$ и в равенствах (2.2) произвести необходимые домножения слева и справа на нужные обратные преобразования. Например,

$$S_O \circ S_C = T_{2CO} \Leftrightarrow S_O \circ S_O \circ S_C = S_O \circ T_{2CO} \Leftrightarrow S_C = S_O \circ T_{2CO},$$

откуда, полагая $a = 2CO$ и помня, что $C = O + OC = O - a/2$, получаем требуемое тождество $S_O \circ T_a = S_{O-a/2}$. Аналогичным образом проверяется и второе тождество.

Отметим, что здесь мы уже начинаем включать привычную для алгебры терминологию и приемы. Например, мы говорим, что умножили равенство $S_O \circ S_C = T_{2CO}$ на S_O слева, а на деле это означает, что мы применили преобразование S_O к результату преобразования, записанного в этом равенстве двумя способами. Поскольку, как мы видим, композиция преобразований не всегда коммутативна, очень важно различать умножение слева и справа!

Кроме того, мы воспользовались тем, что отражение само себе обратное, т. е. $S_O^{-1} = S_O$, в результате чего композицию $S_O \circ S_O$ можно отбросить как нулевое слагаемое или как единичный множитель.

Итак, композиция сдвига и отражения является отражением и при этом не коммутативна!

Теперь мы готовы выписать полную таблицу композиций отражений и сдвигов (см. табл. 2.1).

Таблица 2.1. Таблица композиций движений прямой

	T_a	S_O
T_b	T_{a+b}	$S_{O+b/2}$
S_C	$S_{C-a/2}$	T_{2OC}

Важно! Во всех таблицах композиций порядок действий следующий: берется элемент из левого столбца, следом за ним пишется элемент из верхней строки, результат такой композиции пишется в соответствующую им ячейку. При этом последовательность применения операций обратная: сначала выполняется операция из верхней строки, затем к ее результату применяется операция из левого столбца. Например, если в

левом столбце стоит преобразование F , а в верхней строке — преобразование G , то в соответствующей ячейке стоит композиция $F \circ G$, а ее применение к конкретной точке x выполняется по правилу $F(G(x))$.

Кратность отражения S_O^n определяется четностью числа n . В случае четного n это id , в случае нечетного — исходное S_O .

Стоит отметить, что пара $\{\text{id}, S_O\}$ образует самую маленькую нетривиальную группу движений (см. табл. 2.2).

Таблица 2.2. Группа одного отражения

	id	S_O
id	id	S_O
S_O	S_O	id

Видим, что таблица 2.2 полностью повторяет таблицу умножения знаков с точностью до переобозначений $+ \mapsto \text{id}$, $- \mapsto S_O$. Это легко объяснить тем, что отражение на прямой по сути и является сменой знака вектора, начало которого зафиксировано в центре отражения.

Суммируя, находим, что вообще все сдвиги и отражения вместе образуют группу (относительно операции композиции), т. е. для них выполняются аксиомы группы G1–G4. При этом данная группа не является коммутативной (не выполняется G5), поскольку, как мы видели, далеко не все композиции движений перестановочны.

Построим еще один пример группы. Рассмотрим класс всех сдвигов \mathbb{T} и класс всех отражений \mathbb{S} .

Мы можем определить композицию классов $\mathbb{T} \circ \mathbb{T}$, $\mathbb{T} \circ \mathbb{S}$, $\mathbb{S} \circ \mathbb{T}$ и $\mathbb{S} \circ \mathbb{S}$ как все возможные композиции движений из этих классов в указанном порядке. Иначе говоря, композиции классов — это их *умножение по Минковскому*, которое задается следующим образом:

$$\mathbb{T} \circ \mathbb{T} = \{t \circ t' \mid (t \in \mathbb{T}) \wedge (t' \in \mathbb{T})\}, \quad \mathbb{T} \circ \mathbb{S} = \{t \circ s \mid (t \in \mathbb{T}) \wedge (s \in \mathbb{S})\},$$

$$\mathbb{S} \circ \mathbb{T} = \{s \circ t \mid (s \in \mathbb{S}) \wedge (t \in \mathbb{T})\}, \quad \mathbb{S} \circ \mathbb{S} = \{s \circ s' \mid (s \in \mathbb{S}) \wedge (s' \in \mathbb{S})\}.$$

Из произведенных выше вычислений легко видеть таблицу композиций этих классов:

	\mathbb{T}	\mathbb{S}
\mathbb{T}	\mathbb{T}	\mathbb{S}
\mathbb{S}	\mathbb{S}	\mathbb{T}

и снова мы видим полную аналогию с таблицей знаков и таблицей для группы $\{\text{id}, S_O\}$. Данная таблица выражает самое общее правило композиции отражений и сдвигов, которое можно выразить словами:

композиция однородных движений есть сдвиг, композиция разнородных движений есть отражение.

Итак, мы теперь знаем, что сдвиги и отражения образуют группу движений прямой. Более того, мы понимаем структуру этой группы в общем виде.

Дальнейшая наша цель — доказать, что других движений прямой нет, т. е. что множество $\{T_a, S_O\}_{a, O}$ полностью исчерпывает все возможные движения прямой, т. е. представляет собой группу всех движений прямой.

2.4. Теорема о гвоздях

Анализ движений проводится на основе наблюдений за количеством неподвижных точек.

Напомним, что движение есть такое преобразование (прямой), которое сохраняет расстояния.

Пусть движение M таково, что оно оставляет на месте две точки $A \neq B$, т. е. $M(A) = A$ и $M(B) = B$ (см. рис. 2.6). Пусть $C' = M(C)$. M сохраняет расстояния AC и BC , откуда $AC = AC'$ и $BC = BC'$, откуда $C = C'$, т. е. $M(C) = C$ для любых точек C , т. е. $M = \text{id}$.

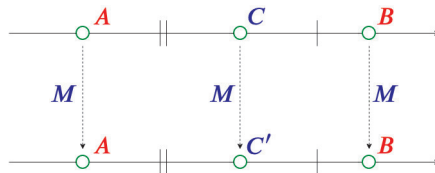


Рис. 2.6. Две неподвижные точки.

Пусть теперь движение M оставляет на месте ровно одну точку O (см. рис. 2.7). В этом случае $A' = M(A)$ и $A \neq A'$ и $AO = OA'$, тогда A' — отражение A относительно O . Следовательно, $M = S_O$.

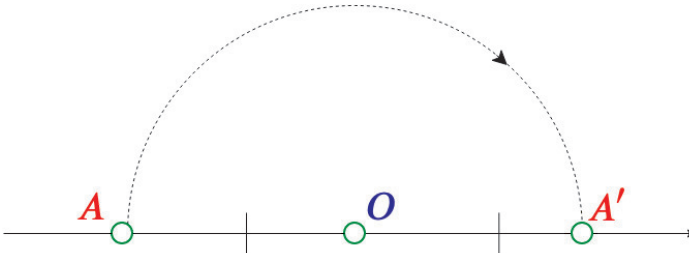


Рис. 2.7. Одна неподвижная точка.

Наконец, пусть движение M не оставляет на месте ни одной точки

и пусть $B = M(A)$ ($B \neq A$). Обозначим $x = AB$ (см. рис. 2.8). Тогда $T_x^{-1} \circ M(A) = A$, т. е. $T_x^{-1} \circ M$ оставляет на месте хотя бы одну точку A .

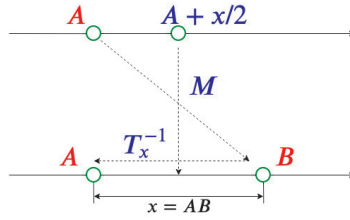


Рис. 2.8. Нет неподвижных точек.

Если оно оставляет на месте ровно одну точку A , то это некоторое отражение S_A , но тогда $M = T_x \circ S_A = S_{A+x/2}$. Получается, что M сохраняет точку $A + x/2$ на месте. Противоречие. Остается вариант, что $T_x^{-1} \circ M$ оставляет на месте как минимум две точки, но тогда $T_x^{-1} \circ M = \text{id}$, откуда $M = T_x \circ \text{id} = T_x$ — сдвиг.

Таким образом, все движения прямой — это либо сдвиги (в частности, id), либо отражения.

При этом любое движение — это либо одно отражение, либо композиция двух отражений.

2.5. Все конечные подгруппы движений прямой

Имея дело с группой движений, мы можем выделить несколько ее собственных подгрупп (определение и разбор понятия подгруппы см. на стр. 40):

- подгруппа всех сдвигов — бесконечная коммутативная группа;
- подгруппа, порожденная одним сдвигом $\langle T_a \rangle$ — тоже бесконечная коммутативная группа;
- подгруппа одного отражения $\{\text{id}, S_O\}$ — конечная группа, состоящая из двух элементов, также коммутативная;
- тривиальная подгруппа $\{\text{id}\}$.

Возникает вопрос: а существуют ли промежуточные по размеру конечные подгруппы группы движений? Попробуем описать все конечные подгруппы движений прямой.

Пусть H — конечная подгруппа группы движений прямой.

Ранее (см. стр. 40) мы рассматривали два определения подгруппы: 1) как подмножества, являющегося группой относительно той же операции, 2) как непустого подмножества H , в котором $ab^{-1} \in H$ при $a, b \in H$. В случае конечной подгруппы определение можно упростить еще больше, а именно заменить ab^{-1} на банальное ab .

Итак, непустое конечное (!) подмножество H группы G образует подгруппу группы G , если для любых $a, b \in H$ имеет место $ab \in H$, т. е. если H замкнуто относительно групповой операции.

Почему это так? Здесь нам на помощь приходит принцип Дирихле. Пусть в множестве H ровно n элементов ($n > 0$). Возьмем какой-то элемент $h \in H$ и рассмотрим все его натуральные степени относительно групповой операции: $h, h \circ h, h \circ h \circ h$ и т. д. Ясно, что мы можем построить сколь угодно длинные композиции, в том числе содержащие n и более вхождений элемента h . Возьмем тогда первые $n + 1$ таких композиций. Все они, по условию, являются элементами множества H , т. е. совпадают с одним из его n элементов. Но тогда в силу принципа Дирихле найдется как минимум две равных композиции. Пусть в одной из них k вхождений h , а в другой j , причем $k < j$:

$$\underbrace{h \circ h \circ \dots \circ h \circ h}_{k \text{ раз}} = \underbrace{h \circ h \circ \dots \circ h \circ h}_{j \text{ раз}}.$$

Пользуясь тем, что в группе G есть обратный элемент h^{-1} , домножим это равенство справа k раз на h^{-1} , в итоге получим

$$\text{id} = \underbrace{h \circ h \circ \dots \circ h \circ h}_{j-k \text{ раз}}.$$

Справа — композиция элементов из H , а значит, принадлежит H , откуда следует, что id исходной группы G находится в H . Мы можем еще раз умножить полученное равенство на h^{-1} , и получим

$$h^{-1} = \underbrace{h \circ h \circ \dots \circ h \circ h}_{j-k-1 \text{ раз}},$$

где $j - k - 1 \geq 0$. Справа либо композиция элементов из H , либо id , который также принадлежит H по доказанному. Но тогда и $h^{-1} \in H$. Так что всякий элемент входит в H вместе со своим обратным. А отсюда уже следует, что H удовлетворяет второму определению подгруппы, и по доказанному на стр. 40 является группой с той же операцией и единицей, что и в группе G .

Итак, пусть H — непустое конечное подмножество группы движений прямой, замкнутое относительно операции композиции. Как мы только что выяснили, H является подгруппой группы движений.

Таким образом, во-первых, id есть элемент H .

Во-вторых, никакой сдвиг T_a при ненулевом a не может быть элементом H , иначе в H окажутся все степени T_a , т. е. $\langle T_a \rangle \subseteq H$, и H будет бесконечной.

В-третьих, если в H есть хотя бы два различных отражения S_A и S_B ($A \neq B$), то и их композиция также находится в H , но это ненулевой сдвиг T_{2AB} , а все такие сдвиги мы исключили чуть выше. Следовательно, если в группе H и есть отражение, то только одно.

Таким образом, либо $H = \{\text{id}\}$ (тривиальная группа), либо $H = \{\text{id}, S_O\}$ при некотором отражении S_O . Это утверждение дает нам полное описание всех возможных *конечных* подгрупп группы движений прямой.

Упражнения

Обязательные упражнения

Пусть на прямой даны 4 точки A, B, C, D , поставленные друг за другом с одинаковым шагом (см. рис. 2.9).

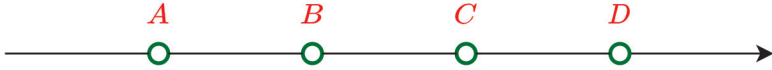


Рис. 2.9. Точки на прямой.

- 2.1° Куда перейдет точка A при отражении S_B ?
- 2.2° Куда перейдут точки B, C, D при преобразовании $T_{AB} \circ T_{CA}$?
- 2.3° Куда перейдут точки A, B, C при преобразовании $S_C \circ T_{AB}$?
- 2.4° Какое движение переводит A в C и B в D ?
- 2.5° Существует ли движение, которое переводит A в B и B в D ?
- 2.6° Опишите все движения, которые переводят A в C , используя только буквы A, B, C, D и обозначения сдвига и отражения.
- Введем координату на прямой, отметим там точки с целыми координатами: $\dots, -2, -1, 0, 1, 2, \dots$. Через S_n обозначим отражение относительно точки n , через T_n — сдвиг на число n .
- 2.7° Известно, что при некотором преобразовании G точка 0 переходит в 2, а 2 — в 3. Может ли оно быть движением? Каким?
- 2.8° Известно, что при некотором преобразовании G точка 0 переходит в 3, а 2 — в 1. Может ли оно быть движением? Каким?
- 2.9° Известно, что при некотором преобразовании G точка 0 переходит в 2, а при обратном преобразовании G^{-1} точка 3 переходит в -1 . Может ли G быть движением? Каким?
- 2.10° Дано движение G . Известно, что $G^{-1}(0) = 1$ и при этом у G^{-1} нет неподвижных точек. Чему равно G ?
- 2.11° Назовем **четностью движения** прямой четность количества отражений, с помощью которых это движение может быть выражено. Какова четность следующих движений: $S_0, T_x, T_x \circ T_y, S_0 \circ T_x, S_0 \circ S_1 \circ T_x \circ T_y, T_x^{-1}, S_0^{-1}, S_0 \circ S_1 \circ \dots \circ S_n$?
- 2.12° Доказать, что:
- а) Четность обратного движения G^{-1} совпадает с четностью исходного движения G .

- b) Четность композиции движений равна сумме четностей (по модулю 2) компонентов.
 c) Четность движения не зависит от его представления в виде композиций каких-либо движений.

2.13° Какое движение получится при композиции:

- a) $S_0 \circ S_1$;
 b) $S_0 \circ S_1 \circ S_2$;
 c) $S_0 \circ S_2 \circ S_1$?

2.14° Построить сдвиг на 7 единиц вправо с помощью композиции двух отражений.

2.15° Пусть G и Q — два движения прямой, причем $G \circ Q = \text{id}$ и $G \neq Q$. Какими могут быть G и Q ?

2.16° Вывести равенства $S_C \circ T_a = S_{C-a/2}$ и $T_b \circ S_O = S_{O+b/2}$ из соотношения $S_C \circ S_O = T_{2OC}$ алгебраическим путем.

2.17° Пользуясь аксиомами группы, доказать обобщенный закон ассоциативности, т. е. для любого n выполнено

$$(\dots((a_1 \circ a_2) \circ a_3) \circ \dots) \circ a_n = a_1 \circ (a_2 \circ (a_3 \circ \dots \circ (a_{n-1} \circ a_n) \dots)).$$

2.18° Доказать единственность нейтрального элемента группы, т. е. если $u \circ \text{id} = \text{id} \circ u = u$ и $u \circ \text{id}' = \text{id}' \circ u = u$, то $\text{id} = \text{id}'$.

Дополнительные упражнения

2.19* Каким движением является композиция

$$S_n \circ S_{n-1} \circ \dots \circ S_1 \circ S_0?$$

Ответ получить в зависимости от четности n .

2.20* При каких n сдвиг T_n выражается в виде композиций S_0 и S_1 ?

2.21* При каких n сдвиг S_n выражается в виде композиций S_0 и S_1 ?

2.22* Пусть G и Q — два движения прямой, причем $G \circ Q = Q \circ G$ и $G \neq Q$. Какими могут быть G и Q ?

2.23* Доказать, что никакая композиция движений S_n и T_m с целыми индексами n, m не может быть равна сдвигу T_x с нецелым x и отражению S_y с неположительным y .

Вокруг окружности

Аннотация

В этой главе мы расширяем сферу деятельности и переходим к движениям окружности. Снова изучаем виды движений, строим таблицу композиций, доказываем соответствующий вариант теоремы Шаля.

Попутно сопоставляем движения окружности с движениями прямой, выходим на отрицательные степени композиций и их арифметические свойства, как следствие, получаем целые числа другим путем.

По аналогии с натуральными числами говорим о том, что целые числа — это и степени композиций движений, и мера длины, только оснащенная знаком, т. е. направлением измерения длины.

3.1. Движения окружности

Иллюстративная сказка. Вспомним песенку «встаньте, дети, встаньте в круг!» Представим, что много-много детей в спортивном зале выстроились в круг и начали водить хоровод. При этом в центре круга стоит воспитательница, так что все дети держатся на равном от нее расстоянии и смотрят на нее. Какое они осуществляют движение в зале? Они ходит под одной и той же окружности то в одну сторону, то в другую. И если мы поместим себя на место воспитательницы, то поймем, что хоровод просто вращается вокруг одного центра то влево, то вправо.

Разнообразим описанное мероприятие. В зале на полу прочерчена линия, которая делит его пополам на два прямоугольника. По хлопку в ладоши дети перебегают на противоположную сторону зала, т. е. они все бегут по прямым параллельным линиям, перпендикулярно линии, нарисованной на полу (примерно так же люди переходят зебру при переходе улицы по зеленому сигналу светофора). При этом каждый из них считает число шагов до линии, а затем продолжает движение в том же направлении еще на столько же шагов. Дойдя до конца, все разворачиваются так, чтобы снова видеть воспитательницу.

В итоге они снова образуют круг, только вывернутый наизнанку, т. к. у каждого участника круга сосед, стоявший слева, теперь стоит справа, и наоборот! Такое действие с кругом называется отражением. При этом тот факт, что соседи поменялись местами, говорит нам о некотором необычном движении, а именно о движении, меняющем ориентацию (перепутали право и лево).

Формализуем это с геометрической точки зрения. Берем окружность. Какие у нее есть движения, переводящие ее саму в себя?

Прежде всего, повторим, что движение — это преобразование, сохраняющее расстояния (иначе это еще называется термином *изометрия*). Поэтому, если мы говорим о движении, переводящем фигуру (прямоугольник, квадрат, многоугольник, плоскость и т. д.) в саму себя, то это значит, что мы берем копию этой фигуры и накладываем ее на оригинал до полного совмещения контуров. При этом допускается вертеть ее как угодно, лишь бы наложение фигур оказалось идеальным — без выступов и впадин, без какой-либо деформации.

Для того чтобы уточнить смысл определения движения, нужно зафиксировать способ измерения расстояний на окружности. Расстоянием между точками окружности A и B мы будем называть длину меньшей из дуг, соединяющих эти точки (вместо длины дуги можно использовать величину соответствующего ей угла, измеренного в радианах или градусах).

Очевидно, что движениями окружности являются как минимум следующие: 1) вращение вокруг ее центра, а также 2) отражение относительно прямых, проходящих через ее центр.

В некотором смысле окружность — аналог прямой. Только эту прямую взяли за два конца и замкнули где-то на бесконечности.

Поэтому вращение окружности напоминает сдвиг прямой, а отражение окружности относительно прямой, проходящей через ее центр, напоминает отражение на прямой относительно точки (можно считать ее отражением относительно перпендикулярной прямой).

Если представить, что на окружности большого радиуса живут маленькие одномерные математики, то для них окружность будет практически не отличима от прямой, и движения окружности они будут воспринимать именно как движения прямой.

Поворот на угол α мы обозначим за R_α (положительный — против часовой стрелки), отражение относительно прямой, имеющей угол наклона φ , обозначим за S_φ ($0 \leq \varphi < 180^\circ$). Угол наклона прямой измеряется от заданного раз и навсегда радиуса окружности, который можно считать точкой отсчета (аналог нуля на прямой). В качестве радиуса нулевого угла мы выберем горизонтальный радиус справа от центра окружности.

Ось отражения S_φ мы будем обозначать l_φ (см. рис. 3.1).

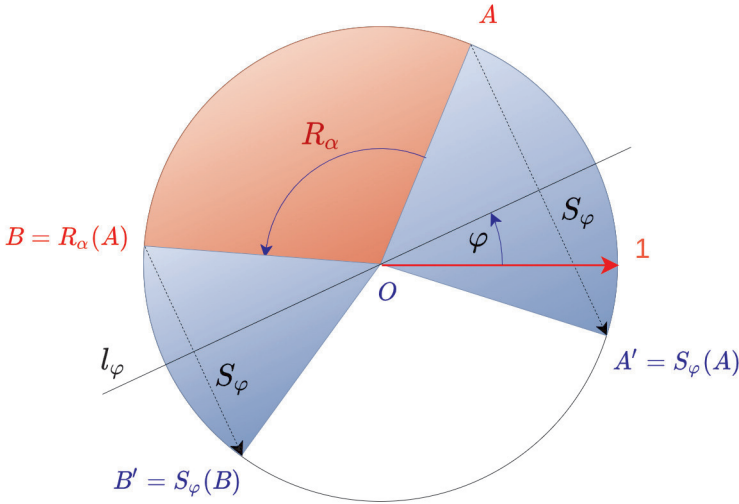


Рис. 3.1. Отражение и поворот окружности.

В полном соответствии с нашей аналогией между поворотами окружности и сдвигами прямой замечаем, что композиция поворотов есть поворот на суммарный угол: $R_\alpha \circ R_\beta = R_{\alpha+\beta}$.

Далее, у каждого поворота есть обратный: $R_\alpha^{-1} = R_{-\alpha}$, т. е. поворот в противоположном направлении.

Кроме того, повороты коммутируют: $R_\alpha \circ R_\beta = R_\beta \circ R_\alpha$.

А также есть нейтральный поворот $\text{id} = R_0$ на нулевой угол. Правда, здесь аналогия становится неточной, т. к. в случае сдвигов прямой нулевой сдвиг $\text{id} = T_0$ определялся единственным вектором — нулевым, в случае же вращений окружности один и тот же нейтральный поворот определяется целой бесконечной серией углов, кратных 360° . Иначе говоря, между углами α и преобразованиями поворота R_α соответствие не является взаимно однозначным.

Общая формула здесь такова:

$$R_\alpha = R_{\alpha \pm 360^\circ k},$$

где k — натуральное число. Поэтому обычно параметр α выбирается так, что $0^\circ \leq \alpha < 360^\circ$, однако при различных расчетах для удобства записи формул мы можем выходить за рамки этого диапазона, имея в виду указанное тождество. Например, поворот по часовой стрелке на 240° есть R_{-240° , и он же равен повороту на 120° против часовой стрелки, т. е. R_{120° , поскольку $120^\circ - (-240^\circ) = 360^\circ$, т. е. полный оборот.

Все перечисленные свойства поворотов говорят нам о том, что все повороты окружности образуют коммутативную группу относительно операции композиции.

Пользуясь групповой нотацией, скажем, что R_α^n представляет собой композицию поворота R_α с самим собой в количестве n штук. R_α^n при этом называется степенью поворота R_α . Кроме того, по определению, $R_\alpha^{-n} = (R_\alpha^n)^{-1}$, так что степень поворота может быть любым целым числом.

И точно так же, как в случае сдвигов прямой, можно заметить, что $R_\alpha^n = R_{\alpha n}$.

Заметим, что некоторые повороты дают id в некоторой степени, например, $R_{90^\circ}^4 = \text{id}$, $R_{60^\circ}^6 = \text{id}$ и т.д. Более того, можно дать точный критерий того, когда это происходит.

Если угол, выраженный в градусах, соизмерим с величиной 360° , то поворот на данный угол имеет конечную положительную натуральную степень, в которой он обращается в id .

Действительно, соизмеримость угла φ с углом 360° , как мы определяли ранее, означает, что существует некоторый угол ψ , кратный как φ , так и 360 , т.е.

$$\psi = \varphi m = 360n.$$

Но это и означает, что поворот R_φ , возведенный в степень m , даст угол, кратный 360° , т.е. id .

Если же угол α , выраженный в градусах, не соизмерим с величиной 360° , то никакая целая степень поворота R_α не равна id (ведь если бы получилось, что $R_\alpha^n = \text{id}$, то мы бы получили равенство $\alpha = k \cdot 360^\circ/n$ при некотором k , а это и означает соизмеримость углов).

Таким образом, критерием того, что R_α в некоторой степени даст id , является условие соизмеримости угла α , выраженного в градусах, с величиной 360° .

Возникает вопрос: а бывают ли несоизмеримые углы?

Ответ: да, такие углы есть. Например, это угол в 1 радиан.

По определению угол φ равен 1 радиану, если длина дуги окружности, соответствующая данному углу, в точности равна радиусу этой окружности.

Когда угол измеряется в радианах, имеется в виду, что мера угла есть длина соответствующей этому углу дуги единичной окружности. В частности, развернутому углу соответствует половина длины единичной окружности, обозначаемая числом π , так что угол 180° — это π радиан.

Если бы угол в 1 радиан был соизмерим с полным оборотом, то число π также оказалось бы соизмеримым с 1. Известно, однако, что это не

так! Но доказательство этого факта является сложной математической теоремой и останется за рамками нашего курса.

В свете сказанного, получаем, что сколько бы раз мы ни откладывали угол в 1 радиан на окружности, мы никогда не окажемся в точке, соответствующей нулевому углу. Поэтому группа вращений, порожденная степенями $R_{1\text{rad}}$, является бесконечной. Позже мы докажем теорему о том, что углы поворота из этой группы образуют плотное множество, т.е. этими углами можно с любой точностью приблизить поворот на любой угол. Визуально это означает, что если мы начнем ходить по окружности одинаковыми шагами длиной в радиус этой окружности и ставить каждый раз кляксы, то со временем эти кляксы равномерно покроют всю окружность (а учитывая, что они имеют ненулевой размер, они полностью закрасят окружность через какое-то конечное число шагов, но об этом — позже).

В зависимости от соизмеримости угла поворота с полным оборотом некоторые повороты порождают конечные циклические подгруппы в группе движений, а некоторые — нет. Дальнейшей нашей целью является изучение группы вращений в контексте изучения группы движений окружности, а также, как и в случае прямой, анализ конечных подгрупп движений окружности.

3.2. Группа движений окружности

Композиция двух отражений формально повторяет аналогичный результат для отражений прямой:

$$S_\psi \circ S_\varphi = R_{2(\psi-\varphi)}, \quad S_\varphi \circ S_\psi = R_{2(\varphi-\psi)}. \quad (3.1)$$

Например, второе равенство легко увидеть из картинки 3.2, где точка A переходит в A' под действием отражения S_ψ относительно оси l_ψ , а затем A' переходит в A'' под действием отражения S_φ относительно соответствующей оси l_φ .

Суммарный угол поворота точки A при переходе в точку A'' можно разбить на две пары углов так, что в каждой паре углы равны в силу свойств отражения (разные пары отмечены разным цветом), и в то же время угол между осями состоит как раз из суммы углов, принадлежащих разным парам. Нетрудно убедиться в аналогичном результате и в том случае, если точка лежит между осями отражений.

Итак, композиция отражений является поворотом на двойной угол между их осями. Отсюда видно также, что композиция отражений не коммутативна! Перестановка отражений приводит к смене направления вращения.

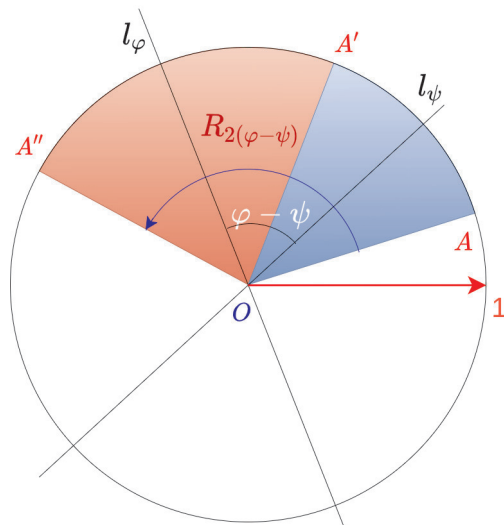


Рис. 3.2. Композиция двух отражений.

Теперь получим композицию отражения и поворота:

$$S_\varphi \circ R_\alpha = S_{\varphi-\alpha/2}, \quad R_\alpha \circ S_\varphi = S_{\varphi+\alpha/2}. \quad (3.2)$$

Действительно, рассмотрим композицию $S_\varphi \circ R_\alpha$. Пусть также $\psi = \varphi - \alpha/2$. Из (3.1) получаем, что

$$S_\varphi \circ S_\psi = R_{2(\varphi-\psi)} = R_\alpha,$$

после чего умножаем это равенство слева на S_φ и, пользуясь тем, что $S_\varphi \circ S_\varphi = \text{id}$, находим, что:

$$S_\psi = S_\varphi \circ R_\alpha,$$

откуда, производя замену $\psi = \varphi - \alpha/2$, окончательно получаем, что

$$S_\varphi \circ R_\alpha = S_{\varphi-\alpha/2}.$$

Аналогично доказывается второе равенство.

Итак, композиция отражения и поворота является отражением и при этом тоже не коммутативна!

Запишем полную таблицу композиций отражений и поворотов окружности:

	R_α	S_ψ
R_β	$R_{\alpha+\beta}$	$S_{\psi+\beta/2}$
S_φ	$S_{\varphi-\alpha/2}$	$R_{2(\varphi-\psi)}$

По аналогии с прямой обозначим за \mathbb{T} класс всех вращений окружности, за \mathbb{S} — класс всех отражений окружности.

Получаем аналогичную таблицу композиций классов:

	\mathbb{T}	\mathbb{S}
\mathbb{T}	\mathbb{T}	\mathbb{S}
\mathbb{S}	\mathbb{S}	\mathbb{T}

Снова наблюдаем все ту же группу умножения знаков!

Существуют ли другие движения окружности? Ответ: нет!

Анализ движений проводится, как и в случае прямой, на основе наблюдений за количеством неподвижных точек.

Для начала заметим, что если при движении окружности одна точка остается на месте, то неподвижной будет и диаметрально противоположная ей точка. Если бы это было не так, то, очевидно, расстояние между этими точками (равное половине дуги окружности) не сохранилось бы — оно стало бы меньше. А это невозможно при движении.

Поэтому при анализе движений окружности всегда нужно иметь в виду, что пары противоположных точек ведут себя одинаково — либо они обе стационарны, либо обе двигаются.

Пусть движение M таково, что оно оставляет на месте две точки $A \neq B$, не являющиеся диаметрально противоположными.

Тогда, во-первых, $M(A) = A$ и $M(B) = B$. Пусть C — еще какая-то точка и $C' = M(C)$. Здесь могут быть два варианта: либо C лежит на малой дуге AB , либо на большой. Эти дуги не могут быть равны по длине, т. к. A и B не являются противоположными (см. рис. 3.3). Точка C' тоже может лежать строго на одной из этих дуг.

Поскольку M сохраняет расстояния, дуги AC и AC' равны, дуги BC и BC' равны. А значит, равны и суммы длин дуг $AC + CB$ и $AC' + C'B$. Отсюда следует, что C и C' могут лежать только на одной и той же дуге. Но тогда, в силу равенства дуг AC и AC' точки C и C' также должны совпадать (они лежат на одной дуге и на равных расстояниях от концов). Таким образом, $M(C) = C$ для любых точек C , т. е. $M = \text{id}$.

Пусть движение M оставляет на месте ровно одну пару противоположных точек A и A' (см. рис. 3.3). Рассмотрим снова произвольную точку C на окружности, отличную от A и A' . И пусть $C' = M(C)$. Тогда $C \neq C'$ и $AC = AC'$. Отсюда следует, что C' — отражение точки C относительно оси AA' . Следовательно, $M = S_\varphi$, где φ — угол наклона прямой AA' .

Пусть движение M не оставляет на месте ни одной точки. Возьмем произвольную точку A на окружности, и пусть $B = M(A)$ (по условию $B \neq A$). Обозначим за α угол дуги AB (см. рис. 3.4).

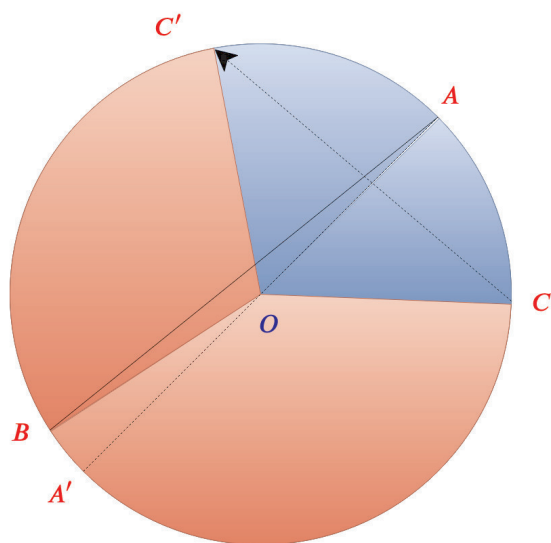


Рис. 3.3. Две неподвижные точки.

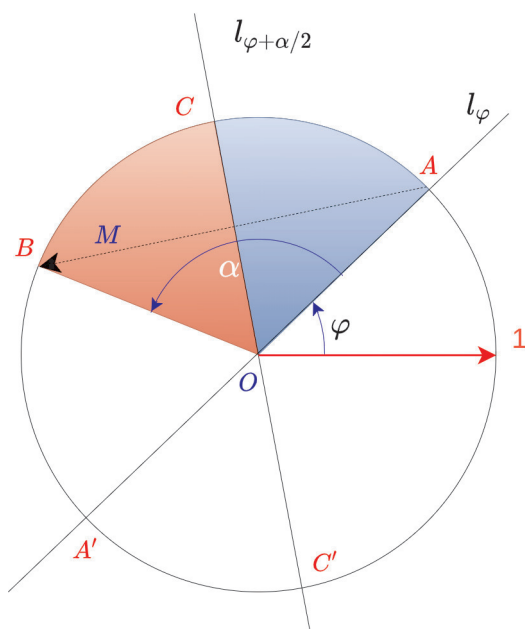


Рис. 3.4. Нет неподвижных точек.

Тогда $R_\alpha^{-1} \circ M(A) = A$, т. е. $R_\alpha^{-1} \circ M$ оставляет на месте хотя бы одну точку A (а точнее, пару противоположных точек A и A'). Если оно оставляет на месте ровно одну пару точек A и A' , то это некоторое

отражение S_φ (на рис. 3.4 ось отражения обозначена за l_φ), но тогда $M = R_\alpha \circ S_\varphi = S_{\varphi+\alpha/2}$ в силу второго равенства в (3.2). Получается, что M сохраняет точку C на месте (C есть середина дуги AB). Получаем противоречие с тем, что M не оставляет на месте ни одной точки.

Остается вариант, что $R_\alpha^{-1} \circ M$ оставляет на месте как минимум две точки, не являющиеся противоположными, но тогда $R_\alpha^{-1} \circ M = \text{id}$ (по доказанному ранее), откуда $M = R_\alpha \circ \text{id} = R_\alpha$ — поворот.

Таким образом, всякое движение окружности — это либо поворот (в частности, id), либо отражение относительно оси, проходящей через центр окружности.

При этом любое движение окружности — это либо одно отражение, либо композиция двух отражений.

3.3. Наматывание прямой на окружность

Совместим теперь окружность с прямой иным способом. Выделим на окружности точку O и начнем обход окружности (вращение) в положительном направлении, т. е. против часовой стрелки.

Выше мы видели, что углы поворота, кратные 360° , т. е. полному обороту, соответствуют тождественному движению, т. е. приведут нас в точку отправления O .

Однако, если с точки зрения математического движения ничего не изменилось, физически мы проделали путь, равный длине окружности. Для удобства будем считать, что радиус окружности есть единичный вектор, так что ее длина равна 2π , и с каждым полным оборотом мы будем «наматывать» расстояние 2π .

Более общо, расстояние, пройденное по окружности единичного радиуса, когда этот радиус заметает угол α , равно $\alpha(2\pi/360^\circ)$. Чтобы каждый раз не переводить единицы измерения радиуса в градусы и наоборот, углы также принято измерять в единицах длины — радианах. А именно, *угол в 1 радиан соответствует повороту, при котором точка проделает по окружности путь, равный по длине радиусу данной окружности*. Нетрудно видеть, что в градусах 1 радиан будет иметь выражение $360^\circ/(2\pi)$ или $180^\circ/\pi \approx 57^\circ$.

В дальнейшем условимся все углы измерять в радианах, если не оговорено иное.

Известно, что число π не соизмеримо с целыми числами (как уже отмечалось, этот факт является довольно сложной теоремой), так что поворот R_1 на 1 радиан ни в какой положительной степени не приведет нас снова в точку исхода O .

Зато поворот $R_{2\pi}$ в точности возвращает нас в точку отправления O .

При каждом таком повороте мы проделываем путь, равный углу поворота, т. е. 2π (радиус равен 1).

Следовательно, степени такого поворота $R_{2\pi}^n$ дадут прохождение пути длиной $2\pi n$.

Представим эту картину не с точки зрения жителей окружности, бегающих по замкнутой траектории, а с точки зрения жителей прямой, которая наматывается на окружность. С их точки зрения все выглядит несколько иначе и больше напоминает движение колеса по дорожному полотну: окружность катится по прямой и через равные промежутки касается точкой O данной прямой.

Если при этом два друга — один из мира окружности, второй из мира прямой, — двигаются с одинаковой скоростью в одном направлении, то они могут синхронизироваться в точке касания окружности и прямой и разговаривать друг с другом, постоянно двигаясь каждый по своему объекту, но вместе.

Нужно заметить при этом, что если колесо вращается по часовой стрелке, т. е. в отрицательном направлении, то вдоль прямой оно движется направо, т. е. в положительном направлении. Но фокус в том, что житель окружности для синхронизации с жителем прямой должен идти навстречу вращению колеса, т. е. тоже в положительном направлении! Таким образом, движения обоих друзей имеют одинаковый знак! На рис. 3.5 мы отметили синей стрелкой направление движения жителя окружности, а черной — встречное вращение самой окружности.

Итак, колесо катится, два друга беседуют, точка O то и дело, а именно через каждые 2π метров, соприкасается с прямой. Каждый раз, когда точка O касается прямой, наш ученый друг из мира прямой ставит на ней отметины и считает их по порядку, т. е. приравнивает к степени совершенного поворота колеса: в начальный момент времени это был 0, затем 1 оборот, затем 2 оборота и т. д.

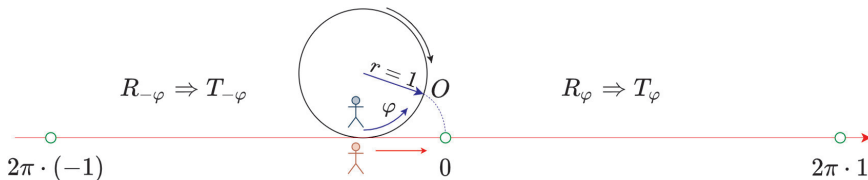


Рис. 3.5. Соприкосновение миров прямой и окружности.

Что же мы видим на прямой? Мы видим не что иное, как шкалу натуральных чисел, в точности соответствующую степеням вращений окружности. Число 2π , фигурирующее как коэффициент, является не более чем единицей измерения. Кто-то измеряет в метрах, кто-то — в радиусах, а мы измеряем в радиусах окружности.

Представим теперь, что в какой-то момент касания точки O с прямой физика мира изменилась, и вращение начало осуществляться в обратную сторону!

Наши друзья-ученые при этом продолжают совместное путешествие, но только назад. Они пойдут отсчитывать уже проставленные отметки на прямой в убывающем порядке, пока не вернутся в точку 0 . Но здесь процесс не остановится, и движение продолжится дальше.

Как все это записать на языке поворотов и сдвигов?

Предположим, что сначала окружность повернулась на n полных оборотов вперед, а затем на m полных оборотов назад (см. рис. 3.6).

Мы получаем итоговый поворот, записываемое как $R_{2\pi m}^{-1} \circ R_{2\pi n}$.

А что мы имеем с точки зрения движения на прямой?

Сначала был произведен сдвиг $T_{2\pi n}$, затем сдвиг $T_{-2\pi m}$.

И мы видим, что индекс, определяющий итоговый поворот и итоговый сдвиг, — один и тот же!

Причем если $n > m$, то сдвиг жителя прямой будет вправо на расстояние $2\pi(n - m)$, а поворот жителя окружности будет положительным на угол $2\pi(n - m)$.

Если же $n < m$, то сдвиг жителя прямой будет влево на расстояние $2\pi(m - n)$, а поворот жителя окружности будет отрицательным (по часовой стрелке) на угол $2\pi(m - n)$.

Ранее мы уже договаривались, что перед векторами, направленными влево, будем ставить знак '-'. Так же мы будем поступать и с углами вращений в отрицательную сторону.

Соответственно, при $n < m$ мы будем иметь итоговый сдвиг на прямой $T_{-2\pi(m-n)}$ и итоговый поворот на окружности $R_{-2\pi(m-n)}$, которые также можно записать в виде степеней:

$$T_{-2\pi(m-n)} = T_{2\pi}^{-(m-n)} \quad \text{и} \quad R_{-2\pi(m-n)} = R_{2\pi}^{-(m-n)}.$$

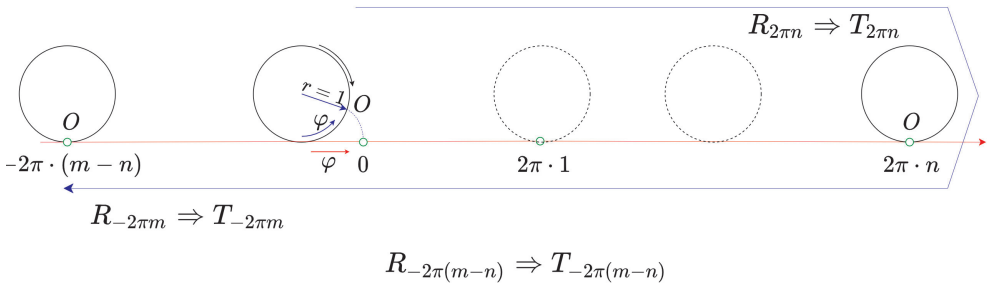


Рис. 3.6. Соответствие сдвигов и вращений прямой и окружности.

Осталось добавить маленький штрих к портрету, а именно: в случае $n < m$ под разностью $n - m$ будем понимать запись $-(m - n)$.

Тогда уже независимо от того, $n < m$, или $m < n$, или $n = m$, композиция поворотов и сдвигов сначала на n вправо и затем на m влево будет записываться одинаково:

$$T_{2\pi(n-m)} = T_{2\pi}^{n-m} \text{ и } R_{2\pi(n-m)} = R_{2\pi}^{n-m}.$$

В итоге мы приходим к тому, что называется **целыми числами**, включающими натуральные числа и отрицательные натуральные числа (при этом $-0 = 0$).

Сколько бы мы ни вращали окружность на 2π в ту или иную сторону с помощью поворота $R_{2\pi}$, мы совершаем поворот на целую степень полного оборота. При этом, как бы мы ни катали окружность по прямой, точка O будет ставить отметки в точках $2\pi k$, где k — целое число.

Тем самым окружность наносит на прямую метки, соответствующие целым числам, и мы видим почти механическую взаимосвязь между целыми числами в качестве меры длины и целыми числами в качестве степеней кратных композиций.

Упражнения

Обязательные упражнения

3.1° Выпишите все конечные подгруппы группы движений окружности порядка не выше 6, содержащие отражение S_0 (относительно горизонтальной оси).

3.2° Какова группа движений правильного треугольника, квадрата, пятиугольника?

3.3° Доказать, что $\pi > 3$.

3.4° Пусть G — движение окружности. Сколько у G может быть неподвижных точек (имеется в виду общее количество, найдите все возможные варианты)?

3.5° Пусть G — движение окружности и $G(A) = A$, $G(B) \neq B$. Какой вид может иметь G ?

3.6° Пусть диаметры l и k перпендикулярны. Найдите $S_l \circ S_k$.

3.7° Известно, что точка A переходит при движении G окружности в точку A' , диаметрально противоположную точке A . Каким может быть движение G ?

3.8° Движение назовем *четным*, если оно является композицией двух отражений, а в противном случае — *нечетным*. Верно ли, что:

a) Композиция четных движений — четное движение, композиция двух нечетных движений — четное движение, композиция четного движения с нечетным движением — нечетное движение?

b) G четно тогда и только тогда, когда G^{-1} нечетно?

3.9° Центральная симметрия — это какое движение?

3.10° Композицией каких отражений можно выразить центральную симметрию?

3.11° С помощью отражения относительно оси Ox (горизонтальной оси) и вращений выразить отражение относительно оси Oy (вертикальной оси).

3.12° Возьмем некоторый угол $\varphi > 0$. Найдите:

a) $S_0 \circ S_\varphi$;

b) $S_0 \circ S_\varphi \circ S_{2\varphi}$;

c) $S_0 \circ S_{2\varphi} \circ S_\varphi$;

d) $S_0 \circ S_\varphi \circ S_{2\varphi} \circ S_{3\varphi} \circ \dots \circ S_{n\varphi}$.

Чему равно последнее выражение, если $\varphi = \pi/2$, $\varphi = \pi$, $\varphi = 2\pi$?

3.13° Построить поворот на угол 90° при помощи двух отражений.

3.14° При каких n поворот на угол $n\varphi$ выражается в виде композиций S_0 и S_φ ?

3.15° Пусть G и Q — движения окружности, причем $G \circ Q = Q \circ G$. Какими могут быть G и Q ?

3.16° Пусть G и Q — движения окружности, причем $G \circ Q = \text{id}$. Какими могут быть G и Q ?

Целые числа и ОГА

Аннотация

Это — первая глава, где мы по-настоящему погружаемся в арифметику, используя тот понятийный аппарат, который был наработан в предыдущих главах. Здесь вводится обозначение множества целых чисел, дается строгое определение алгебраического понятия «кольцо», обосновывается алгоритм Евклида.

Ключевым моментом является доказательство теоремы о том, что НОД двух чисел можно записать в виде их линейной комбинации с целыми коэффициентами. Этот факт выводится как непосредственно из алгоритма Евклида, так и с помощью сумм Минковского.

Далее выводится основная теорема арифметики и некоторые ее следствия.

4.1. Целые числа. Кольцо

Итак, совмещение поворотов со сдвигами дает нам полную свободу перемещений в положительном и отрицательном направлении. При этом с точки зрения окружности ничего не меняется — происходит итоговое движение id , а с точки зрения прямой — происходит разметка точек с равным шагом. Ясно, что сам шаг при этом не имеет значения. Мы могли бы взять окружность радиуса R , и тогда шаг был бы равен $2\pi R$. В частности, можно взять радиус $R = 1/2\pi$, и тогда точки на прямой расположатся с шагом 1.

Такую же картину можно получить, если взять все точки, получаемые из выделенной точки 0 степенями сдвига на единичный вектор, используя положительные и отрицательные, т. е. целые, степени.

Как видим, целые числа, как и натуральные, можно интерпретировать и как степени движений (и вообще любых преобразований, имеющих обратные), и как векторы сдвигов на прямой, а значит, к ним применимы определенные ранее операции сложения, вычитания и умножения. При этом результат умножения получает такой знак, который определяется из таблицы умножения знаков.

Множество всех целых чисел принято обозначать буквой \mathbb{Z} . Вместе с операциями сложения (вычитания) и умножения структура $(\mathbb{Z}, +, \cdot)$ называется **кольцом целых чисел**. Кольцо — это структура, где можно складывать, вычитать и умножать (соблюдая обычные правила арифметики, см. ниже).

Понятие кольца является обогащением понятия группы, т. к. добавляется операция умножения. Всякое кольцо есть группа по операции сложения, т. е. кольцо можно считать частным случаем группы.

Ранее мы уже видели такие группы, как группа движений прямой, группа умножения знаков, группа композиций классов сдвигов и отражений, группа вращений окружности. Все они обладали одной операцией — композицией, которая соответствовала сложению параметров сдвигов и вращений.

Кроме того, мы ввели такое понятие, как кратность, заменяя тем самым многократное сложение умножением на целое число.

Кратность операций нельзя рассматривать как умножение сдвигов, вращений или отражений. Иначе говоря, если у нас имеется запись nT_v , обозначающая n -кратный сдвиг на вектор v , то в этой записи слева стоит натуральное число, отвечающее за кратность операции сдвига, а справа — собственно операция сдвига. В таком «произведении» множители принадлежат к разным понятиям (число и операция), а значит, запись nT_v нельзя рассматривать как умножение операций или как умножение чисел, т. е. как бинарную операцию над множеством однородных объектов. Поэтому движения в общем случае образуют только лишь группу, но не кольцо. (Забегая вперед, мы можем сказать, что движения с операцией умножения на целое число образуют модуль над кольцом целых чисел.)

Однако уже сами кратности, как самостоятельные сущности, можно и складывать, и умножать. Например, если мы рассмотрим сдвиг T_1 и композицию его кратностей $T_1^n \circ T_1^m$, то получим тот же сдвиг, но в суммарной кратности T_1^{n+m} , где $n, m \in \mathbb{Z}$. Ничто не мешает нам рассмотреть кратность m сдвига T_1^n , т. е. сдвиг $(T_1^n)^m$, а это уже будет не что иное, как сдвиг кратности nm , т. е. T_1^{nm} .

Иначе говоря, умножение на целых числах можно представить как кратности кратностей сдвигов!

Фиксируем понятие **кольца**. Это — множество K с двумя бинарными операциями $+$ (плюс) и \cdot (точка, «умножить»), которые подчинены следующим аксиомам:

R1 $a, b \in K \Rightarrow a + b \in K, a \cdot b \in K$ (замкнутость операций);

R2 $a, b, c \in K \Rightarrow (a + b) + c = a + (b + c), (a \cdot b) \cdot c = a \cdot (b \cdot c)$ (ассоциативность операций);

R3 существует элемент $0 \in K$ такой, что $a + 0 = 0 + a = a$ для всех $a \in K$ (аксиома нуля);

R4 для всякого элемента $a \in K$ существует противоположный $-a$ такой, что $a + (-a) = 0$ (аксиома противоположного элемента);

R5 для всех $a, b, c \in K$ имеем $(a + b) \cdot c = a \cdot c + b \cdot c$, $c \cdot (a + b) = c \cdot a + c \cdot b$ (правая и левая дистрибутивность);

R6 для всех $a, b \in K$ имеем $a + b = b + a$ (коммутативность сложения).

Обычно изучаются **кольца с единицей**, т. е. такие кольца, для которых выполняется требование:

R7 существует элемент $1 \in K$ такой, что $a \cdot 1 = 1 \cdot a = a$ для всех $a \in K$ (аксиома единицы);

а также **коммутативные кольца**, т. е. такие кольца, для которых

R8 для всех $a, b \in K$ имеем $a \cdot b = b \cdot a$ (коммутативность умножения).

Иначе говоря, в коммутативном кольце с единицей можно складывать, вычитать и умножать по обычным правилам.

Отметим также, что иногда в определении кольца допускается ослабление требований, а именно, отмена требования ассоциативности умножения: $(ab)c = a(bc)$. Такие кольца называют неассоциативными. Но мы будем придерживаться классического определения термина «кольцо», следуя в этом Кострикину [13].

4.2. Кузнечик НОД и алгоритм Евклида

Поработаем теперь непосредственно с целыми числами. Пусть у нас есть кузнечик, стоящий в точке 0, который умеет прыгать с шагом a и с шагом b в любую сторону. Числа a, b — натуральные.

Ясно, что он может попасть в любую точку вида $ka + mb$, где кратности k, m — целые. Можно ли проще описать, в какие точки он может попасть, а в какие — нет?

Определим множество $m\mathbb{Z}$ как множество всех произведений m на целые числа:

$$m\mathbb{Z} = \{mk \mid k \in \mathbb{Z}\}.$$

Пусть d — наименьшее положительное число, в которое кузнечик может попасть, т. е. оно имеет вид $d = ka + mb$ при некоторых k, m . Тогда он может попасть и в любое число вида nd , поскольку $nd = (nk)a + (nm)b$, где $n \in \mathbb{Z}$. Следовательно, кузнечик может попасть во все целые числа, кратные d (множество $d\mathbb{Z}$).

Но в любые другие целые числа он не сможет попасть. Действительно, если он попадает в какое-то число x , лежащее между двумя

соседними кратными d числами, т. е. в число $x = nd + y$, где $0 < y < d$, то тогда он может попасть в число y , т. е. остаток от деления x на d . Для этого после попадания в точку x ему нужно будет n раз применить алгоритм перехода на расстояние d , только в обратном направлении (если, например, из нуля в точку d он способен попасть, сделав s шагов на a и t шагов на b , то для сдвига на nd влево ему надо сделать sn шагов на $-a$ и tn шагов на $-b$). Но $y < d$ и притом положительное, а это противоречит выбору числа d . Таким образом, кузнечик попадает во все точки $d\mathbb{Z}$, и только в эти точки!

Что такое d на самом деле?

Заметим, что поскольку кузнечик может прыгнуть в точки a и b , а также, как мы только что показали, он прыгает только в точки, кратные d , то a и b сами кратны d , т. е. d является их общим делителем.

С другой стороны, если какое-то $q > d$ также является общим делителем a и b , то кузнечик прыгает только в точки, кратные q (т. к. он попадает только в точки вида $ka + mb$), но тогда он окажется неспособным попасть в точку $d < q$, т. к. d не кратно q .

Отсюда следует, что не существует делителя a и b большего, чем число d . Таким образом, d — наибольший общий делитель a и b , что обозначается следующим образом: $\text{НОД}(a, b)$.

Придем к тому же выводу, используя алгоритм Евклида (с отсечениями квадратов).

Пусть $a < b$. Вычтем из b столько a , сколько сможем: $b = k_0a + r_1$, где $0 \leq r_1 < a$. Далее, из a вычитаем столько r_1 , сколько сможем, если $r_1 > 0$. Получим $a = k_1r_1 + r_2$, где $0 \leq r_2 < r_1$. Снова, если $r_2 > 0$, вычитаем из r_1 столько r_2 , сколько можем: $r_1 = k_2r_2 + r_3$, где $0 \leq r_3 < r_2$. И так далее.

Видим, что всякий раз, когда $r_i > 0$, мы получаем $r_{i+1} < r_i$. Проблема в том, что это не может продолжаться бесконечно долго, т. к. от всякого натурального числа в сторону нуля можно спуститься за конечное число шагов (а ведь остатки у нас все положительные и целые!). Так что рано или поздно случится $r_{n+1} = 0$, и на этом алгоритм Евклида остановится! Это значит, что прямоугольник $a \times b$ можно замостить квадратами $r_n \times r_n$.

Если теперь раскрутить равенства $r_{i-1} = k_i r_i + r_{i+1}$ в обратную сторону, то мы получим, во-первых, что a и b кратны r_n , и, во-вторых, что $r_n = Ka + Mb$ при некоторых целых K, M . То есть r_n есть общий делитель исходных чисел a и b , и наш кузнечик способен попасть в точку r_n (а значит, и во все точки, ему кратные, т. е. в $r_n\mathbb{Z}$).

С другой стороны, если какое-то q является общим делителем a и b , то q делит $r_1 = b - k_0a$, делит $r_2 = a - k_1r_1$, делит $r_3 = r_1 - k_2r_2$ и т. д., наконец, делит r_n . Стало быть, $q \leq r_n$, и r_n — наибольший общий делитель a и b .

Итак, кузнечик способен попасть в $\text{НОД}(a, b)$, следовательно, $d \leq \text{НОД}(a, b)$. С другой стороны, выбор d таков, что $d = ka + mb$ при некоторых целых k, m , но тогда всякий делитель a и b является и делителем d , в частности $\text{НОД}(a, b)$ делит d , откуда $\text{НОД}(a, b) \leq d$. Таким образом, минимальный шаг, на который способен сдвинуться кузнечик, — это наибольший общий делитель чисел a и b . Поэтому кузнечика с ногами a и b можно назвать $\text{НОД}(a, b)$. Он способен прыгнуть (в несколько прыжков) во ВСЕ точки, кратные $\text{НОД}(a, b)$, и ТОЛЬКО в эти точки!

4.3. Простые числа и ОТА

У кузнечика НОД может случиться уникальная ситуация, когда даже при достаточно больших числах a и b он способен прыгнуть в любое целое число! Это верно в том и только том случае, когда $\text{НОД}(a, b) = 1$. Тогда говорят, что a и b взаимно просты. Например, 125 и 63 взаимно просты.

Для обозначения того факта, что числа a и b взаимно просты, мы будем пользоваться обозначением $a \perp b$ (см., напр., Кнут [7]).

Натуральное число называется **простым**, если оно имеет ровно два делителя среди натуральных чисел — единицу и самого себя. Целое число называется простым, если оно имеет ровно четыре делителя среди целых чисел: — ± 1 и самого себя со знаками \pm .

Если число a простое, а число b не кратно a , то a и b взаимно просты, т.к. a имеет только делители ± 1 и $\pm a$, но при этом $\pm a$ не делит b , значит, единственным положительным общим делителем a и b будет единица, т.е. $\text{НОД}(a, b) = 1$. Например, 101 — простое, так что в паре с любым другим числом (кроме кратного 101) они будут взаимно просты, и наш кузнечик сможет прыгнуть в любую целую точку! Например, он умеет прыгать на 101 и 62, значит, он умеет прыгать в любое целое число!

Любое натуральное число можно представить как произведение степеней простых чисел. Действительно, 1 есть произведение нулевых степеней простых чисел, например, 2^0 . Предположим, что для всех чисел от 1 до n утверждение о разложимости справедливо (внимание! индукция!) и рассмотрим число $n + 1$. Оно либо уже простое, либо делится на число меньше $n + 1$ и отличное от 1. Тогда $n + 1 = mk$, причем $m, k \leq n$, а они есть произведение степеней простых по предположению индукции, но тогда и $n + 1$ есть произведение степеней простых!

Простых чисел бесконечно много. Предположим, что это не так, и пронумеруем все простые числа:

$$p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, p_5 = 11, \dots, p_n.$$

Далее рассмотрим число $m = p_1 p_2 \dots p_n + 1$. Это число не является простым, т. к. оно больше p_n . Тогда по определению оно должно быть кратным какому-то числу d , которое больше 1 и меньше m . По доказанному выше число d можно представить как произведение степеней простых, причем как минимум одного простого в степени не меньше первой (иначе $d = 1$), т. е. d делится на какое-то простое число p_k из данного ряда.

Тогда m кратно числу p_k , т. е. $m = l p_k$. Отсюда следует, что

$$1 = m - p_1 p_2 \dots p_n = p_k (l - p_1 \dots p_{k-1} p_{k+1} \dots p_n),$$

т. е. единица кратна числу p_k . Противоречие.

Следовательно, наше предположение о том, что простых чисел конечный набор, ложно.

Если простое число p делит произведение чисел ab , то оно по крайней мере делит одно из них. Действительно, допустим, что p не делит a , тогда $\text{НОД}(p, a) = 1$, но тогда, как мы уже видели выше, $1 = kp + ma$ при некоторых целых k, m . Умножим это равенство на b : $b = kpb + tab$. Справа оба слагаемых делятся на p , значит, и b делится на p .

Из этого свойства легко получить **основную теорему арифметики** (ОТА): каждое натуральное число, большее 1, единственным образом представляется в виде произведения положительных степеней простых чисел:

$$n = p_1^{k_1} p_2^{k_2} \dots \quad (4.1)$$

Набор степеней k_1, k_2, \dots уникален для каждого числа $n > 1$. Действительно, если бы было два разложения, то после сокращения на одинаковые сомножители мы получили бы равенство

$$p_1^{k_1} p_2^{k_2} \dots p_m^{k_m} = q_1^{s_1} q_2^{s_2} \dots q_t^{s_t},$$

где $\{p_1, \dots, p_m\} \cap \{q_1, \dots, q_t\} = \emptyset$.

Но каждое простое слева делит произведение чисел справа, значит, делит один из его множителей (по доказанному выше), а значит, совпадает с одним из q_i , что по предположению невозможно. Следовательно, разложение по положительным степеням простых чисел единственно.

Мы могли бы сформулировать ОТА немного иначе, разрешив n принимать значение 1, а степени простых допускать равными нулю. В этом случае мы также получаем единственное представление вида (4.1), но с той оговоркой, что это произведение бесконечное, ибо в нем участвуют все простые числа, хотя с некоторого простого числа и до конца ряда степени равны нулю. Тогда число 1 можно записать как бесконечное произведение всех простых чисел в степени 0.

Поскольку мы не хотим оперировать бесконечными произведениями без особой надобности, формулировка ОТА только для $n \geq 2$ и только с положительными степенями простых множителей выглядит наиболее приемлемой.

Нужно также сделать специальную оговорку про \mathbb{Z} . Любое целое ненулевое число также единственным образом раскладывается по степеням простых натуральных чисел, но с точностью до знака \pm перед этим разложением. Можно пойти дальше и сказать, что любое ненулевое целое число раскладывается единственным образом по степеням простых целых чисел с точностью до знаков перед этими простыми числами, ведь $pq = (-p)(-q)$ и т. д. Но гораздо проще все сводить к разложению по положительным простым числам, выставляя общий знак числа перед ним. Поэтому, говоря об ОТА в целых числах, мы будем подразумевать ОТА в натуральных числах с точностью до выбора нужного знака перед произведением степеней положительных простых чисел.

Возьмем произвольное натуральное число $n \neq 0$. Перенумеруем все простые числа в порядке возрастания:

$$p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, \dots$$

В силу ОТА имеем единственное разложение в виде бесконечного произведения

$$n = p_1^{k_1} p_2^{k_2} \dots,$$

где k_1, k_2, \dots — некоторые натуральные числа (в частности, нули). Для удобства мы можем считать, что в этом произведении задействованы только простые числа p_1, \dots, p_n , т.к. $p_n > n$ и, следовательно, все простые с номерами n и более входят в это разложение в степени 0.

Вопрос: как с точки зрения разложений по степеням простых должны выглядеть делители числа n ?

Пусть $m \mid n$ и $m > 0$. Тогда $m = p_1^{t_1} p_2^{t_2} \dots$ при некоторых натуральных t_1, t_2, \dots . Кроме того, $n = qm$, где число q также имеет разложение по степеням простых $q = p_1^{s_1} p_2^{s_2} \dots$.

Отсюда получаем, что

$$p_1^{k_1} p_2^{k_2} \dots = p_1^{s_1} p_2^{s_2} \dots \cdot p_1^{t_1} p_2^{t_2} \dots$$

Отметим, что все указанные произведения конечны, т.к. начиная с некоторого номера j (например, при $j = n, n + 1, \dots$) показатели степеней $k_j = t_j = s_j = 0$ (иначе мы бы имели бесконечное произведение чисел, больших единицы). Поэтому мы можем воспользоваться коммутативностью и арифметическими свойствами степеней:

$$p_1^{k_1} p_2^{k_2} \dots = p_1^{s_1+t_1} p_2^{s_2+t_2} \dots$$

Мы имеем два разложения одного и того же числа, а оно в силу ОТА единственно, так что

$$k_1 = s_1 + t_1, \quad k_2 = s_2 + t_2, \quad \dots,$$

откуда легко видеть, что $s_1 \leq k_1$, $s_2 \leq k_2$ и т. д. Иначе говоря, делители числа n характеризуются тем, что их разложения по степеням простых включают те же самые простые числа, что и в разложении числа n , с равными или меньшими степенями.

Отсюда, в частности, легко получить формулу для количества положительных делителей числа n . Для этого нужно найти количество всех наборов s_1, s_2, \dots таких, что $0 \leq s_j \leq k_j$ для каждого номера j . В силу ОТА выбор s_j можно производить независимо друг от друга, не опасаясь получить на выходе одинаковые числа. Поэтому следует перемножить количество вариантов для каждой из степеней s_j . Например, для s_1 существует ровно $k_1 + 1$ вариант: $0, 1, 2, \dots, k_1$. Аналогично — для остальных степеней.

Итак, количество делителей числа n равно

$$(k_1 + 1)(k_2 + 1)(k_3 + 1) \dots$$

Это произведение конечно, т. к., начиная с некоторого номера j , степень $k_j = 0$.

Наконец, получим формулу для функции $\sigma(n)$, равной сумме всех положительных делителей числа n .

Как мы уже видели, всякий положительный делитель есть произведение вида $p_1^{s_1} p_2^{s_2} \dots$, где $0 \leq s_j \leq k_j$ для всех j . Иначе говоря, делители получаются как всевозможные произведения простых чисел, входящих в разложение числа n , во всех степенях, не превосходящих таковых в разложении числа n .

Для каждого p_j и степени k_j построим выражение $1 + p_j + p_j^2 + \dots + p_j^{k_j}$. Затем перемножим все такие выражения:

$$(1 + p_1 + p_1^2 + \dots + p_1^{k_1})(1 + p_2 + p_2^2 + \dots + p_2^{k_2}) \dots,$$

в результате мы получим сумму всех возможных произведений вида $p_1^{s_1} p_2^{s_2} \dots$ с допустимыми значениями степеней. Все произведения, которые получаются после раскрытия скобок, различны в силу ОТА.

В свернутом виде функция $\sigma(n)$ записывается так:

$$\sigma(n) = \prod_j (1 + p_j + p_j^2 + \dots + p_j^{k_j}).$$

И это произведение также конечно, т. к., начиная с некоторого номера j , будем иметь $k_j = 0$ и $(1 + p_j + p_j^2 + \dots + p_j^{k_j}) = 1$.

Основную теорему арифметики можно доказать разными способами. Покажем еще один способ, который использует множества и операции Минковского с этими множествами.

M1 Пусть $P, Q \subseteq \mathbb{Z}$. Суммой и разностью по Минковскому называются, соответственно, множества:

$$P \oplus Q = \{x + y \mid x \in P, y \in Q\}, \quad P \ominus Q = \{x - y \mid x \in P, y \in Q\}.$$

M2 Множества вида $a\mathbb{Z}$ замкнуты относительно операций сложения и умножения (более того, они являются подкольцами кольца \mathbb{Z}), поэтому для любых $P, Q \subseteq a\mathbb{Z}$ и любых $k, n \in \mathbb{Z}$ имеет место вложение:

$$kP \oplus nQ \subseteq a\mathbb{Z}.$$

M3 $a \mid b$ тогда и только тогда, когда $b\mathbb{Z} \subseteq a\mathbb{Z}$.

Действительно, если $a \mid b$, то $b = ka$. Если $x \in b\mathbb{Z}$, то $x = by = ak y \in a\mathbb{Z}$.

Пусть $b\mathbb{Z} \subseteq a\mathbb{Z}$. Так как $b \in b\mathbb{Z}$, то $b \in a\mathbb{Z}$, т.е. $b = ka$ при некотором целом k , то есть $a \mid b$.

M4 Решим вложение $P \ominus P \subseteq P$, где $P \subseteq \mathbb{Z}$.

1) Пустое множество удовлетворяет этому вложению.

2) Множество $P = \{0\}$ также удовлетворяет данному вложению.

3) Пусть $c \in P$ и $c \neq 0$. Тогда, во-первых, $0 \in P$, поскольку $c - c \in P$ в силу вложения $P \ominus P \subseteq P$.

Во-вторых, $-c \in P$, так как $-c = 0 - c$. Следовательно, вместе со всяким числом из P ему также принадлежит противоположное по знаку число. А это значит, что в P есть положительные числа.

Пусть $a = \min\{x \mid (x \in P) \wedge (x > 0)\}$, т.е. наименьший положительный элемент P . Нетрудно видеть, что $a\mathbb{Z} \subseteq P \ominus P$. Действительно, $a, 0, -a \in P$ по доказанному. Тогда $a + a = a - (-a)$, $a + a + a = a + a - (-a)$ и т.д. — все являются элементами $P \ominus P$. Аналогично — для сумм вида $-a - a$, $-a - a - a$ и т.д. То есть все числа вида ka , где $k \in \mathbb{Z}$, являются элементами множества $P \ominus P$. Откуда, в силу вложения $P \ominus P \subseteq P$ получаем, что $a\mathbb{Z} \subseteq P$.

Далее, если $P \setminus a\mathbb{Z}$ не пусто, то существует $x \in P \setminus a\mathbb{Z}$, причем $x = ka + d$, где $0 < d < a$ (то есть x не кратен a). Тогда $d = x - ka \in P \ominus P$, т.е. $d \in P$, что противоречит выбору a как минимального положительного элемента P . Следовательно, $P \subseteq a\mathbb{Z}$, что вместе с предыдущим вложением дает $P = a\mathbb{Z}$.

Таким образом, если $P \ominus P \subseteq P$, то либо $P = \emptyset$, либо $P = a\mathbb{Z}$ при некотором a (в том числе при $a = 0$ имеем $P = \{0\}$).

M5 $a\mathbb{Z} \oplus b\mathbb{Z} = \text{НОД}(a, b)\mathbb{Z}$.

Действительно, $P = a\mathbb{Z} \oplus b\mathbb{Z}$ удовлетворяет вложению $P \ominus P \subseteq P$, и значит, по свойству M4 $a\mathbb{Z} \oplus b\mathbb{Z}$ совпадает с множеством $c\mathbb{Z}$ при некотором c (причем если $a > 0$ или $b > 0$, то и $c > 0$), т. е.

$$a\mathbb{Z} \oplus b\mathbb{Z} = c\mathbb{Z}.$$

Отсюда, с одной стороны, следует, что $a\mathbb{Z}, b\mathbb{Z} \subseteq c\mathbb{Z}$, откуда (свойство M3) $c|a$ и $c|b$. С другой стороны, если $d|a$ и $d|b$, то $a\mathbb{Z}, b\mathbb{Z} \subseteq d\mathbb{Z}$, откуда (свойство M2) $c\mathbb{Z} \subseteq d\mathbb{Z}$, откуда (свойство M3) $d|c$. То есть любой делитель a и b делит число c , при этом c также является делителем a и b . Следовательно, $c = \text{НОД}(a, b)$.

M6 Если простое p делит произведение ab , то $p|a$ или $p|b$ (или p делит их обоих).

Предположим, что $p \nmid a$, тогда $\text{НОД}(p, a) = 1$ и (по свойству M5) $p\mathbb{Z} \oplus a\mathbb{Z} = \mathbb{Z}$. Отсюда $1 = kp + ma$ при некоторых целых k, m . Тогда $b = kbp + tab$, откуда следует, что $p|b$.

Если предположить, что $p \nmid b$, то аналогично получим $p|a$.

M7 Отсюда, как уже отмечалось выше, легко выводится Основная теорема арифметики (ОТА).

Упражнения

Обязательные упражнения

4.1° В какую ближайшую к нулю точку может попасть кузнечик, умеющий делать прыжки по числовой прямой длины 37 и 777, если он стартует в нуле?

4.2° Используя разложение на множители, решите уравнение:

$$n^3(n+1)^3 = 1728.$$

4.3° Кузнечик делает по числовой прямой прыжки длины 11 и 1331. Укажите точки, в которых он может оказаться: 99, 999, 1, 11, 111.

4.4° Два кузнечика на числовой прямой, стартуя из нуля, могут совершать любые комбинации прыжков: первый — длины 16 и 28, а второй — длины 9 и 15. В какой ближайшей к нулю точке они могут встретиться?

4.5° При каком минимальном целом $n > 0$ уравнение $120n = x^3$ будет иметь целочисленное решение?

- 4.6° Доказать, что любое простое число $p > 3$ имеет вид $6k+1$ или $6k+5$.
- 4.7° Доказать, что квадрат простого числа $p > 3$ при делении на 12 дает остаток 1.
- 4.8° Доказать, что любое общее кратное чисел a и b делится на их НОК.
- 4.9° Про натуральные числа a и b известно, что их НОД равен 15, а НОК равен 840. Найти a и b .
- 4.10° Доказать, что при $n > 2$ два числа $2^n - 1$ и $2^n + 1$ одновременно не могут быть простыми.
- 4.11° Какие натуральные числа делятся на 30 и имеют ровно 20 положительных делителей?
- 4.12° Рассмотрим целое число $n > 0$. Докажите, что количество упорядоченных пар натуральных чисел $\langle a, b \rangle$ таких, что $\text{НОК}(a, b) = n$, равно количеству натуральных делителей у числа n^2 .
- 4.13° Существуют ли целые x, y , для которых: а) $x^2 + y^2 = 99$; б) $x^2 + y^2 = 33333$; в) $x^2 + y^2 = 5600$?
- 4.14° а) [Решето Эратосфена] Выпишем целые числа от 2 до n . Подчеркнем 2 и сотрем числа, кратные 2. Первое неподчеркнутое число подчеркнем и сотрем кратные ему, и т. д., пока каждое число от 2 до n не будет подчеркнуто или стерто. Докажите, что мы подчеркнем в точности простые числа от 1 до n . б) Пусть очередное число, которое мы хотим подчеркнуть, больше \sqrt{n} . Докажите, что нестертые к этому моменту числа от 2 до n простые. в) Какие числа, меньшие 100, простые?
- 4.15° Числа a, b, c, n — натуральные, $\text{НОД}(a, b) = 1$, $ab = c^n$. Найдется ли такое целое x , что $a = x^n n$?
- 4.16° Решите в натуральных числах уравнение $x^{42} = y^{55}$.
- 4.17° Найдите разложение по степеням простых числа: а) 2022; б) 17!; в) $\binom{20}{10}$.
- 4.18° При каких натуральных k число $(k-1)!$ не делится на k ?
- 4.19° а) [Теорема Лежандра] Докажите, что простое число p входит в разложение по степеням простых числа $n!$ в степени $[n/p] + [n/p^2] + [n/p^3] + \dots$ (где $[x]$ — это целая часть числа x). С какого момента слагаемые в этой сумме станут равными нулю? б) Сколько у 2000! нулей в конце его десятичной записи? в) Может ли $n!$ делиться на 2^n ($n \geq 1$)?

4.20° Докажите, что существует бесконечное число простых чисел вида: а) $3k + 2$; б) $4k + 3$.

4.21° Сократить дробь $\frac{8547}{4144}$.

4.22° Даны 20 целых чисел, ни одно из которых не делится на 5. Докажите, что сумма двадцатых степеней этих чисел делится на 5.

4.23° Докажите, что из любых 52 целых чисел всегда можно выбрать два таких числа, что а) их разность делится на 51; б) их сумма или разность делится на 100.

4.24° Докажите, что: а) \overline{aaa} делится на 37 (черта означает позиционную запись числа цифрами); б) $\overline{abc} - \overline{cba}$ делится на 99 (где a, b, c — цифры).

4.25° Сформулировать и доказать признаки делимости на 2, 4, 5, 8.

4.26° Из числа $\overline{a_n \dots a_1 a_0}$ вычли сумму его цифр $a_n + \dots + a_1 + a_0$. а) Докажите, что получилось число, кратное 9. б) Выведите отсюда признаки делимости на 3 и на 9.

4.27° Написать на псевдоязыке алгоритм разложения числа по степеням простых.

4.28° Известно, что $n^2(m^2 + 1)(m + 1) = 9999$ при некоторых целых n, m . Найдите эти числа.

4.29° Произведение возрастов Машинных братьев равно 1664. Младший из братьев вдвое моложе старшего. Сколько у Маши братьев?

4.30° Пусть a и b — натуральные числа, не делящиеся на 10, такие, что $ab = 10000$. Чему равна их сумма?

4.31° В силу ОТА будем записывать положительное натуральное число m как последовательность \overline{m} степеней простых:

$$m = p_0^{\alpha_0} p_1^{\alpha_1} \dots p_k^{\alpha_k} \dots \iff \overline{m} = \langle \alpha_0, \alpha_1, \dots, \alpha_k, \dots \rangle, \quad (4.2)$$

где $p_0 < p_1 < p_2 < \dots$ — все простые числа, начиная с 2.

Докажите, что если $\overline{m} = \langle \alpha_0, \alpha_1, \dots, \alpha_k, \dots \rangle$ и $\overline{n} = \langle \beta_0, \beta_1, \dots, \beta_k, \dots \rangle$, то

$$\begin{aligned} \overline{nm} &= \langle \alpha_0 + \beta_0, \alpha_1 + \beta_1, \dots, \alpha_k + \beta_k, \dots \rangle, \\ \overline{\text{НОД}(n, m)} &= \langle \min(\alpha_0, \beta_0), \min(\alpha_1, \beta_1), \dots, \min(\alpha_k, \beta_k), \dots \rangle, \\ \overline{\text{НОК}(n, m)} &= \langle \max(\alpha_0, \beta_0), \max(\alpha_1, \beta_1), \dots, \max(\alpha_k, \beta_k), \dots \rangle. \end{aligned}$$

4.32° Докажите, что $\text{НОД}(n, m)\text{НОК}(n, m) = nm$.

4.33° Вычислите при помощи алгоритма Евклида:

a) НОД(91, 147); b) НОД(-144, -233); c) НОД(525, 231);

d) НОД(7 777 777, 7 777); e) НОД(10946, 17711); f) НОД($2^m - 1, 2^n - 1$).

4.34° Доказать, что все остатки r_k в алгоритме Евклида можно представить в виде линейной комбинации $ax + by$, подобрав подходящие целые x, y .

4.35° Доказать, что алгоритм Евклида завершается за конечное число шагов для любых стартовых целых положительных чисел a и b .

4.36° Докажите, что НОД(a, b) делится на любой общий делитель чисел a и b .

4.37° С помощью представления НОД в виде линейной комбинации исходных чисел докажите, что если НОД(a, b) = 1 и $ac : b$, то $c : b$.

4.38° Какие расстояния можно отложить от данной точки на прямой, пользуясь двумя шаблонами (без делений) длины a см и b см (где НОД(a, b) = d)?

4.39° Переставив цифры в числе N , получили в 3 раза меньшее число. Докажите, что $N : 27$.

4.40° Верен ли такой признак делимости на 27: число делится на 27 тогда и только тогда, когда сумма его цифр делится на 27?

4.41° Запись числа N составлено из записей подряд идущих чисел от 19 до 92:

$$N = 19202122 \dots 909192.$$

На какую максимальную степень тройки оно делится?

4.42° Докажите, что число $11 \dots 11$, запись которого состоит из 3^n единиц, делится на 3^n .

4.43° Докажите, что число делится на 11 тогда и только тогда, когда сумма его цифр, стоящих в четных разрядах, и сумма его цифр, стоящих в нечетных разрядах, отличаются на число, кратное 11.

4.44° Может ли $n!$ оканчиваться ровно на 4 нуля? А ровно на 5 нулей?

4.45° Через $\tau(m)$ обозначим количество всех положительных делителей числа m . Найти $\tau(p^k)$, где p — простое число. Верно ли, что $\tau(ab) = \tau(a)\tau(b)$ при условии НОД(a, b) = 1. Найти $\tau(n)$, если

$$n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$$

есть разложение числа n по степеням простых.

4.46° Через $\sigma(m)$ обозначим сумму всех положительных делителей числа m . Найти $\sigma(p^k)$, где p — простое число, k — целое положительное. Верно ли, что $\sigma(ab) = \sigma(a)\sigma(b)$ при условии $\text{НОД}(a, b) = 1$? Найдите $\sigma(n)$, где

$$n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$$

есть разложение числа n по степеням простых.

4.47° Натуральное число называется **совершенным**, если сумма всех его делителей, меньших его, равна ему самому. Например, 6 и 28 — совершенные числа. Докажите, что число $2^{n-1}(2^n - 1)$ будет совершенным, если $2^n - 1$ — простое число.

Через $\varphi(m)$ обозначим количество всех положительных целых чисел, не превосходящих m и взаимно простых с ним:

$$\varphi(m) = \#\{k \mid 0 < k \leq m \wedge k \perp m\}. \quad (4.3)$$

4.48° Вычислите значения функций φ , τ и σ для чисел 999, 512, 5!

4.49° Доказать, что $2^n - 1$ кратно трем тогда и только тогда, когда n — четное, и $2^n + 1$ кратно трем тогда и только тогда, когда n — нечетное.

4.50° Пусть f_n — «число Ферма» $2^{2^n} + 1$. Докажите, что $f_m \perp f_n$, если $m < n$.

4.51° Доказать, что если $2^n + 1$ — простое число, то n является степенью двойки.

4.52° Докажите, что

$$\text{НОД}(kn, km) = k\text{НОД}(n, m), \quad \text{НОК}(kn, km) = k\text{НОК}(n, m).$$

4.53° Написать алгоритм вычисления последней десятичной цифры выражения a^b на основе последней цифры числа a и представления числа b в виде $b = 4k + r$.

4.54° Найдите совершенное число, кратное 16.

4.55° Сколько существует различных разложений в виде суммы двух простых чисел для числа 22?

4.56° Пифагор назвал *содружественными* числа a и b такие, что a является суммой всех делителей числа b (без самого числа b), а число b является суммой всех делителей числа a (без самого числа a). Найдите число, содружественное числу 220.

4.57° Целое положительное число называется *свободным от квадратов*, если оно не делится на m^2 при любом целом $m > 1$. Установите необходимое и достаточное условие того, что число n свободно от квадратов через представление (4.2) числа n в виде показателей простых.

Дополнительные упражнения

4.58 Найдутся ли такие 10 разных целых чисел, ни одно из которых не квадрат целого числа, со свойством: квадратом целого числа будет произведение а) любых двух из них; б) любых трех них?

4.59 Докажите, что существует бесконечно много натуральных чисел, не представимых как сумма трех или менее точных квадратов.

4.60 Написать реализацию алгоритма Евклида на псевдоязыке программирования. А также алгоритм, выводящий линейное представление НОД через исходные два числа.

4.61 Напишите на псевдоязыке программирования алгоритм вычисления $\tau(n)$ для любого положительного целого числа.

4.62 Напишите на псевдоязыке программирования алгоритм вычисления $\sigma(n)$ для любого положительного целого числа.

4.63 Напишите на псевдоязыке программирования алгоритм вычисления $\varphi(n)$ для любого положительного целого числа.

4.64* а) Докажите, что для любого натурального N существует делящееся на N натуральное число, все цифры которого только 0 и 1. б) Найдется ли такое число вида $1 \dots 10 \dots 0$?

4.65* Шайка из K разбойников отобрала у купца мешок с N монетами. Каждая монета стоит целое число грошей. Оказалось, что какую монету ни отложи, оставшиеся монеты можно поделить между разбойниками так, что каждый получит одинаковую сумму. Докажите, что $N - 1$ делится на K .

4.66* Оценить скорость алгоритма из задачи 4.27° следующим образом: посчитать количество операций деления с остатком, производимых в ходе выполнения алгоритма.

Симметрии фигур

Аннотация

В этой главе мы снова возвращаемся к геометрии и занимаемся полным описанием групп движений правильных многоугольников, а заодно и всех конечных подгрупп движений окружности. В конце главы рассматривается пример группы движений ромба и вводится определение четверной группы Клейна.

5.1. Симметрии правильного треугольника

Вернемся на окружность и рассмотрим на ней вращение $R_{2\pi/3}$, т. е. на 120° .

Множество вращений $R^3 = \{R_{2\pi/3}, R_{2\pi/3}^2, R_{2\pi/3}^3\}$ образует циклическую группу $\langle R_{2\pi/3} \rangle$. Легко видеть, что

$$R^3 = \{\text{id}, R_{2\pi/3}, R_{4\pi/3}\}.$$

Зафиксируем точку A на окружности и найдем ее образы при действии этой группы: $B = R_{2\pi/3}(A)$, $C = R_{4\pi/3}(A)$. Набор точек $\{A, B, C\}$ образует то, что называется *орбитой* точки A при действии группы R^3 , а также представляет собой набор вершин правильного треугольника, вписанного в данную окружность, поскольку все углы этого треугольника равны.

Посмотрим теперь на треугольник ABC . Какие движения переводят его в самого себя? Очевидно, это делают вращения из группы R^3 по построению. Но есть еще и отражения $S^3 = \{S_A, S_B, S_C\}$ относительно осей, проходящих через центр окружности и вершины треугольника (см. рис. 5.1).

Нетрудно проверить, что объединение $R^3 \cup S^3$, состоящее из трех вращений и трех отражений, образует группу относительно операции композиции движений. Такая группа называется **группой симметрий правильного треугольника**.

Таблица 5.1 содержит все варианты композиции двух симметрий из данной группы.

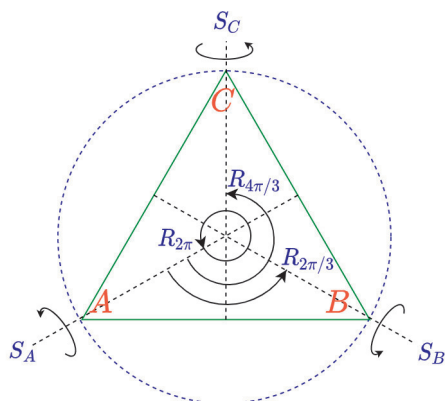


Рис. 5.1. Симметрии правильного треугольника.

Таблица 5.1. Композиции симметрий правильного треугольника

	id	$R_{2\pi/3}$	$R_{4\pi/3}$	S_A	S_B	S_C
id	id	$R_{2\pi/3}$	$R_{4\pi/3}$	S_A	S_B	S_C
$R_{2\pi/3}$	$R_{2\pi/3}$	$R_{4\pi/3}$	id	S_B	S_C	S_A
$R_{4\pi/3}$	$R_{4\pi/3}$	id	$R_{2\pi/3}$	S_C	S_A	S_B
S_A	S_A	S_C	S_B	id	$R_{4\pi/3}$	$R_{2\pi/3}$
S_B	S_B	S_A	S_C	$R_{2\pi/3}$	id	$R_{4\pi/3}$
S_C	S_C	S_B	S_A	$R_{4\pi/3}$	$R_{2\pi/3}$	id

Например, результат композиции $S_A \circ R_{4\pi/3}$ находится на строке S_A в столбце $R_{4\pi/3}$ и равен S_B , поскольку композиция операций вычисляется *справа налево*. Если перепутать порядок действий, то результат будет другой, а именно, S_C , что неверно.

На примере этой группы мы можем заметить, во-первых, что в группе можно выделить подгруппу вращений (верхний левый квадрат 3×3 таблицы 5.1). Во-вторых, что группа движений треугольника конечна и некоммутативна, поскольку ее таблица композиций несимметрична. Кроме того, в полном соответствии с таблицей умножения классов \mathbb{T} и \mathbb{S} , полученной для группы движений окружности, видим, что композиция вращений есть вращение, композиция вращения и отражения есть отражение, композиций двух отражений есть вращение.

Вопрос: есть ли еще какие-то движения окружности, переводящие правильный треугольник в себя?

Заметим, что при движении, переводящем треугольник в себя, вершины обязательно переходят в вершины. Действительно, вершины правильного треугольника обладают тем свойством, что расстояние меж-

ду любыми двумя из них является максимальным расстоянием внутри треугольника (какие две точки в треугольнике ни возьми, расстояние между ними не может быть больше, чем длина стороны этого правильного треугольника). Движение сохраняет расстояния, следовательно, максимальные расстояния также сохранятся, а значит, вершины могут перейти только в вершины.

Таким образом, преобразований правильного треугольника не может быть больше, чем всех возможных перестановок трех вершин:

$$\begin{pmatrix} A & B & C \\ A & B & C \end{pmatrix}, \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}, \begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix},$$

$$\begin{pmatrix} A & B & C \\ A & C & B \end{pmatrix}, \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}, \begin{pmatrix} A & B & C \\ B & A & C \end{pmatrix}.$$

Здесь мы воспользовались стандартным обозначением перестановки, в котором нижний ряд матрицы соответствует образам вершин, стоящих в верхнем ряду при данной перестановке. Подробнее мы изучим перестановки в главе 11.

Нетрудно видеть, что эти перестановки в точности соответствуют преобразованиям id , $R_{2\pi/3}$, $R_{4\pi/3}$, S_A , S_B , S_C . Так что данными преобразованиями исчерпываются все возможные движения, переводящие правильный треугольник в себя.

5.2. Симметрии правильного многоугольника

Рассмотрим еще один случай преобразований фигуры в себя. Пусть имеется правильный n -угольник. Тогда очевидными преобразованиями, сохраняющими форму и размеры фигуры, будут повороты на угол, кратный $2\pi/n$, а также отражения относительно некоторых осей, проходящих через центр данного n -угольника.

В случае четного n в многоугольнике все вершины разбиваются на пары противоположных, лежащих на общей оси симметрии, поэтому имеется $n/2$ осей симметрии, проходящих через вершины, и $n/2$ осей симметрии, проходящих через середины сторон. В случае нечетного n на каждую вершину приходится своя ось симметрии, которая одновременно является и осью симметрии, делящей противоположащую сторону пополам (см. рис. 5.2).

Будем обозначать отражение относительно оси, проходящей через вершину X , как S_X , а отражение относительно оси, проходящей через середину отрезка XY , как S_{XY} .

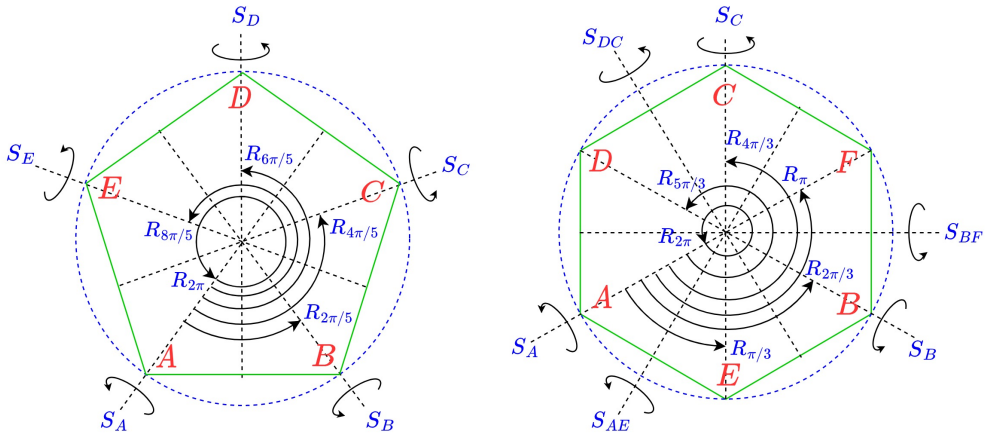


Рис. 5.2. Симметрии правильного n -угольника.

Тогда в общем виде все перечисленные выше движения правильного n -угольника, переводящие его в самого себя, можно записать в виде:

$$R_{2\pi k/n}, \quad S_X, S_{XY}, \quad X, Y \in \{A, B, C, \dots\}, \quad k \in \{1, \dots, n\}, \quad (5.1)$$

где X и Y — соседние вершины.

Например, для 5-угольника, изображенного на рис. 5.2, вращениями из этого списка будут: $R_{2\pi/5}, R_{4\pi/5}, R_{6\pi/5}, R_{8\pi/5}, R_{2\pi}$, а отражениями будут: S_A, S_B, S_C, S_D, S_E , которые можно также записать как $S_{CD}, S_{DE}, S_{EA}, S_{AB}, S_{BC}$.

Для 6-угольника, изображенного на рис. 5.2, вращениями из этого списка будут: $R_{\pi/3}, R_{2\pi/3}, R_{\pi}, R_{4\pi/3}, R_{5\pi/3}, R_{2\pi}$, а отражениями будут: $S_A, S_B, S_C, S_{AE}, S_{BF}, S_{CD}$.

Как и в случае треугольника, несложно показать, что этими $2n$ преобразованиями исчерпываются все движения, переводящие правильный многоугольник в себя. Действительно, движение многоугольника однозначно соответствует движению окружности, в которую он вписан, поскольку все точки дуги окружности, соединяющей соседние вершины многоугольника, сохраняют свое положение относительно этих вершин при любом движении.

Следовательно, все движения правильного многоугольника, переводящие его в себя, есть частные случаи движений окружности, которые мы ранее изучили. Но мы уже выяснили, что такими движениями могут быть либо повороты относительно центра окружности, либо отражения относительно осей, проходящих через центр данной окружности. Остается понять, какие повороты и отражения сохраняют правильный многоугольник, т. е. переводят его в себя.

Достаточно очевидно, что вращение на любой угол, не кратный углу между соседними вершинами, т. е. углу $2\pi/n$, не переводит многоугольник в себя, т. к. не произойдет совмещение его вершин.

То же самое можно сказать и об отражениях относительно осей, угол которых, отложенный от оси, соединяющей две противоположные вершины, не кратен половине угла $2\pi/n$, т. е. углу π/n . При таких отражениях вершина, ближайшая к оси отражения, не достигнет никакой другой вершины многоугольника, а значит, многоугольник не перейдет в себя при таком преобразовании.

Итак, группа движений правильного n -угольника состоит ровно из n поворотов и n отражений, указанных в (5.1).

Число $2n$ уже в случае квадрата ($2n = 8$) сильно уступает количеству всех перестановок n вершин, количество которых равно $n!$ (при $n = 4$ оно равно 24). С ростом n группа движений правильного n -угольника содержит очень мало элементов по сравнению с группой всех перестановок на n символах.

5.3. Подгруппы движений окружности

Правильные n -угольники дают приблизительное представление о подгруппах движений окружности. Приблизительное — именно в том смысле, что движения n -угольников с любой наперед заданной точностью (при достаточно большом n) будут представлять движения окружности.

Вопрос: все ли *конечные* подгруппы движений окружности задаются движениями правильных n -угольников?

Ответ: да, но с оговоркой. Некоторые конечные подгруппы совпадают с группами движений n -угольников, другие же являются их собственными подгруппами.

Действительно, пусть G — некоторая подгруппа движений окружности, причем конечная, т. е.

$$G = \{g_1, g_2, \dots, g_m\}.$$

Количество элементов конечной группы называют ее **порядком**.

Возьмем произвольный элемент g_k и рассмотрим множество всех его целых степеней, т. е. циклическую подгруппу

$$\langle g_k \rangle = \{\dots, g_k^{-1}, g_k^0, g_k, g_k^2, \dots\}.$$

Данное множество, очевидно, является подгруппой группы G , а значит, конечно. Но тогда среди степеней g_k точно есть два совпадающих значения: $g_k^s = g_k^t$ при $t \neq s$. Пусть для определенности $t > s$. Тогда,

умножая равенство на g_k^{-s} , получаем $g_k^{t-s} = g_k^0 = \text{id}$. Иначе говоря, g_k в некоторой положительной степени превращается в id .

Порядком элемента $g \in G$ называется минимальное целое положительное число s такое, что $g^s = \text{id}$. Как видим, для всякого $g_k \in G$ такой порядок существует.

При этом, как мы установили ранее, g_k — это либо поворот окружности, либо отражение относительно оси, проходящей через ее центр. В первом случае порядок может быть любым начиная с 1. В случае, когда порядок элемента g_k равен 1, получаем, что $g_k = \text{id}$, т. е. это поворот на нулевой угол (или угол 2π). Во втором случае, очевидно, что порядок g_k строго равен 2, т. к. отражение само себе обратно.

Если g_k — поворот и s — его порядок, то это поворот на угол $2\pi r/s$, где $r \perp s$. Действительно, с одной стороны, такой поворот в степени s дает угол $2\pi r$, т. е. id , с другой стороны, если бы r и s не были взаимно просты, то можно было бы сократить дробь r/s на $\text{НОД}(r, s)$, тем самым получив поворот на угол $2\pi r_1/s_1$, где $s_1 < s$, а значит, число s не было бы порядком поворота g_k .

Порядок элемента является одновременно и порядком циклической подгруппы, которую он порождает. Действительно, если s — порядок элемента g_k , то все g_k в степенях меньше s различны (иначе порядок оказался бы меньше s), а все бóльшие степени сводятся к меньшим сокращением на g_k^s . Так что в подгруппе $\langle g_k \rangle$ ровно s элементов!

Итак, мы видим, что в G есть подгруппы вида $\langle g_k \rangle$, которые либо тривиальны (состоят из одного элемента id), либо соответствуют группам вращения правильных многоугольников (если g_k — поворот, причем здесь стоит оговориться, что при $g_k = R_\pi$ многоугольника как такового нет, это вырожденный двуугольник), либо соответствуют группам отражений вида $\{\text{id}, S_\varphi\}$ при некотором угле наклона φ оси отражения.

Наша задача состоит в том, чтобы показать, что все эти подгруппы, а равно и сама группа G , есть подгруппы движений какого-то одного n -угольника.

В дальнейшем, когда не требуется явно указывать вид операции (R или S) мы будем опускать символ \circ операции композиции так же, как мы это делаем при записи арифметической операции умножения. Это связано с тем, что в общей теории групп групповая операция обычно воспринимается как умножение и чаще всего пропускается при записи.

Пусть $G' = \{g \in G \mid g \text{ — поворот или } \text{id}\}$. Ясно, что G' — подгруппа группы G . Предположим далее, что $G' \neq G$, т. е. в группе G существует хотя бы одно отражение h . В этом случае, как мы видели ранее, все элементы произведения Минковского hG' также являются отражениями. Предположим, что существует отражение $h' \in G \setminus (hG' \cup G')$. Ранее мы установили, что hh' есть поворот, причем $hh' = g \in G'$, т. к. $hh' \in G$.

Но тогда $h' = h^{-1}g = hg \in hG'$ (отражение обладает свойством $h = h^{-1}$), а это противоречит выбору h' .

Итак, если в группе G есть отражения, то все они находятся в одном классе hG' , причем этот класс не зависит от выбора отражения h . Иначе говоря, все отражения порождены каким-то одним отражением и всеми поворотами. При этом может оказаться, что в группе G есть только один поворот — id , а значит, там есть и только одно отражение.

Осталось разобраться с подгруппой G' всех поворотов.

Возьмем из G' самый маленький поворот, т. е. такой, чей угол наименьший, и обозначим его за g_0 . Угол поворота g_0 обозначим через x_0 , а порядок g_0 — через n . Тогда $x_0n = 2\pi r$ при некотором целом положительном r , взаимно простым с n .

Покажем, что на самом деле $r = 1$. Так как $r \perp n$, то $r\alpha + n\beta = 1$ при некоторых целых α, β (ранее мы получали это свойство из алгоритма Евклида). Тогда

$$\frac{2\pi}{n} = \frac{2\pi r}{n}\alpha + 2\pi\beta = x_0\alpha + 2\pi\beta.$$

Отсюда следует, что поворот $R_{2\pi/n}$ равен g_0^α . Но так как G' — группа, то $g_0^\alpha \in G'$, и мы в множестве G' находим поворот на угол $2\pi/n$, и если $r > 1$, то это противоречит выбору g_0 (так как угол x_0 в этом случае будет больше, чем $2\pi/n$). Следовательно, $r = 1$ и $x_0n = 2\pi$.

Пусть g — произвольный поворот из G' и его угол поворота равен $x > 0$ (если угол поворота отрицательный, то можно рассмотреть g^{-1} , который также принадлежит G'). По условию $x \geq x_0$, а если x не делится нацело на x_0 , то имеет место представление

$$x = kx_0 + y,$$

где $0 < y < x_0$. Углу y соответствует поворот $g' = g(g_0)^{-k}$, который, очевидно, принадлежит группе G' . Но тогда получается, что в группе G' имеется нетривиальный поворот на угол $y < x_0$, что противоречит выбору x_0 .

Следовательно, $y = 0$ и x кратно x_0 . А это значит, что произвольный поворот $g \in G'$ является степенью минимального поворота из G' .

Таким образом, подгруппа G' группы G состоит из поворотов, являющихся степенями поворота g_0 — того, у которого угол поворота наименьший среди поворотов группы G' и равен $2\pi/n$! В частности, отсюда следует и то, что порядок самой группы G' равен порядку этого наименьшего поворота g_0 , т. е. числу n .

Итак, произвольная конечная группа движений окружности:

- либо тривиальна, т. е. совпадает с $\{\text{id}\}$,

- либо является циклической группой поворотов $\langle g_0 \rangle$, совпадающей с группой поворотов правильного n -угольника, где n — порядок этой группы (включая вырожденный случай 2-угольника),
- либо является группой одного отражения $\{id, S_\varphi\}$,
- либо есть объединение $\langle g_0 \rangle \cup h\langle g_0 \rangle$, где h — некоторое отражение того же самого правильного n -угольника.

Наконец, заметим, что и тривиальная группа, и циклическая конечная группа поворотов порядка n , и группа одного отражения $\{id, S_\varphi\}$ (здесь важно отметить, что для согласования S_φ с многоугольником нужно, чтобы ось отражения проходила через вершину или середину стороны многоугольника), и наиболее полная группа $\langle g_0 \rangle \cup h\langle g_0 \rangle$ — все они являются подгруппами группы движений правильного n -угольника. Отсюда следует, что все конечные группы движений окружности являются подгруппами движений правильных многоугольников, лежащих на данной окружности.

5.4. Симметрии ромба, группа Клейна

Рассмотрим ромб, не являющийся квадратом. У такого ромба группа движений существенно меньше, чем у квадрата.

Движения ромба состоят из (см. рис. 5.3):

- двух отражений: относительно его диагоналей, обозначим эти отражения за S_1 и S_2 ;
- одного поворота: на угол π , обозначим этот поворот за R ;
- тождественного преобразования $id = R_{2\pi}$.

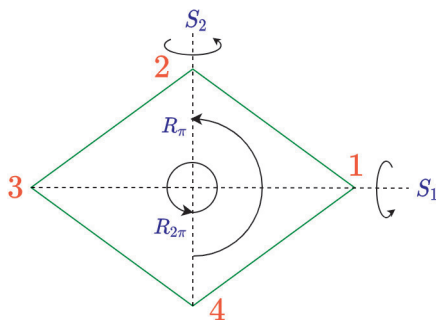


Рис. 5.3. Симметрии ромба.

Других движений ромба не существует. Докажем это.

Пронумеруем вершины ромба цифрами 1, 2, 3, 4 (1 и 3 противоположны), причем будем считать, что диагональ 1–3 — наибольшая из двух

(см. рис. 5.3). В этом случае расстояние между вершинами 1 и 3 является максимальным расстоянием между двумя точками в ромбе. Откуда следует, что диагональ 1–3 при любом движении переходит только в саму себя. Если эта диагональ остается на месте при движении, то данное движение является либо id , либо отражением относительно данной диагонали. Если эта диагональ переворачивается, то смотрим на вторую диагональ. Она может либо тоже остаться на месте, и тогда наше движение есть отражение относительно второй диагонали, либо тоже переворачивается, что соответствует повороту на угол π . Таблицу композиций группы движений ромба см. в табл. 5.2.

Таблица 5.2. Композиции движений ромба

	id	R	S_1	S_2
id	id	R	S_1	S_2
R	R	id	S_2	S_1
S_1	S_1	S_2	id	R
S_2	S_2	S_1	R	id

Отличие данной группы от группы движений правильного n -угольника состоит в том, что группа ромба является коммутативной.

Тем не менее это не единственное ее отличие от групп движений правильного многоугольника. Ведь в группе движений правильного многоугольника есть коммутативная подгруппа вращений. Например, группа вращений квадрата тоже имеет порядок 4. Но и тут мы находим отличие от группы движений ромба. Дело в том, что вращения квадрата есть степени одного поворота (а именно, поворота на 90°). То есть группа вращений квадрата — циклическая. А если мы посмотрим на таблицу умножения группы движений ромба, то заметим, что степени вращения R не дают ни одно из отражений, так же как и степени отражений не дают вращения. Это значит, что группа движений ромба не является циклической.

В то же время такая группа не уникальна по своей природе. Ее ипостаси мы еще встретим при изучении вычетов и перестановок. С точностью до переобозначений элементов это будет все та же группа движений ромба. Вообще, если у двух групп получается одна и та же таблица композиций (умножения) при некотором соответствии элементов одной группы элементам другой, то такие группы называются **изоморфными**. Общее название класса групп, изоморфных группе движений ромба, — «четверная группа Клейна», и общее обозначение ее таково: V_4 .

Упражнения

Обязательные упражнения

5.1° Пусть задан квадрат $ABCD$. Какими движениями исчерпывается группа его движений? Заполните таблицу композиций движений данного квадрата.

Движения плоскости и пространства

Аннотация

Данная глава продолжает тему групп движений. Здесь мы получаем теорему Шаля (для движений плоскости), а затем широкими мазками освещаем тему движений сферы и пространства.

Разделы о сфере и пространстве могут быть пропущены при первом ознакомлении с текстом.

6.1. Виды движений плоскости. Теорема Шаля

Иллюстративная сказка. Представим себе Красную площадь и парад Победы. По площади идут ровные коробки участников парада. Чтобы не нарушать красоту и геометрию движения коробок, солдаты маршируют так, чтобы в каждый момент времени между любыми двумя из них сохранялось одно и то же расстояние. Иначе говоря, они осуществляют движение прямоугольной фигуры по плоскости площади. Затем точно так же делают колонны боевой техники.

По площади все они движутся равномерно и прямолинейно, пока им не придется осуществить поворот сначала на Лобном месте на Васильевский спуск, а затем и на Кремлевскую набережную. И всякий поворот ради сохранения стройности шеренг и колонн нужно также осуществлять с сохранением расстояний между всеми участниками парада. Таким образом, движение плоскости можно иллюстрировать прямолинейным смещением и поворотом, а также их последовательными комбинациями.

Наконец, как и в случае окружности, возможно перестроение, при котором внутри одной коробки шеренги меняются ролями: первая шеренга становится последней, вторая — предпоследней и так далее. Здесь мы можем вспомнить аналогичную картинку с парковкой автомобилей, которую мы рассматривали ранее при описании движений прямой. Только вместо одного ряда машиномест у нас несколько шеренг:

при перестроении они проходят в противоположных направлениях на равные расстояния относительно центральной шеренги. В итоге получается такая же коробка, только стоявшие впереди участники оказываются сзади, и наоборот. Такое движение, переводящее коробку в себя и меняющее право-лево, либо перед-зад, называется осевой симметрией, и оно также, как и в случае окружности, меняет пространственную ориентацию коробки (см. рис. 6.1 — здесь справа мы видим такую же коробку построения, как и слева, только видим ее как бы с обратной стороны листа).

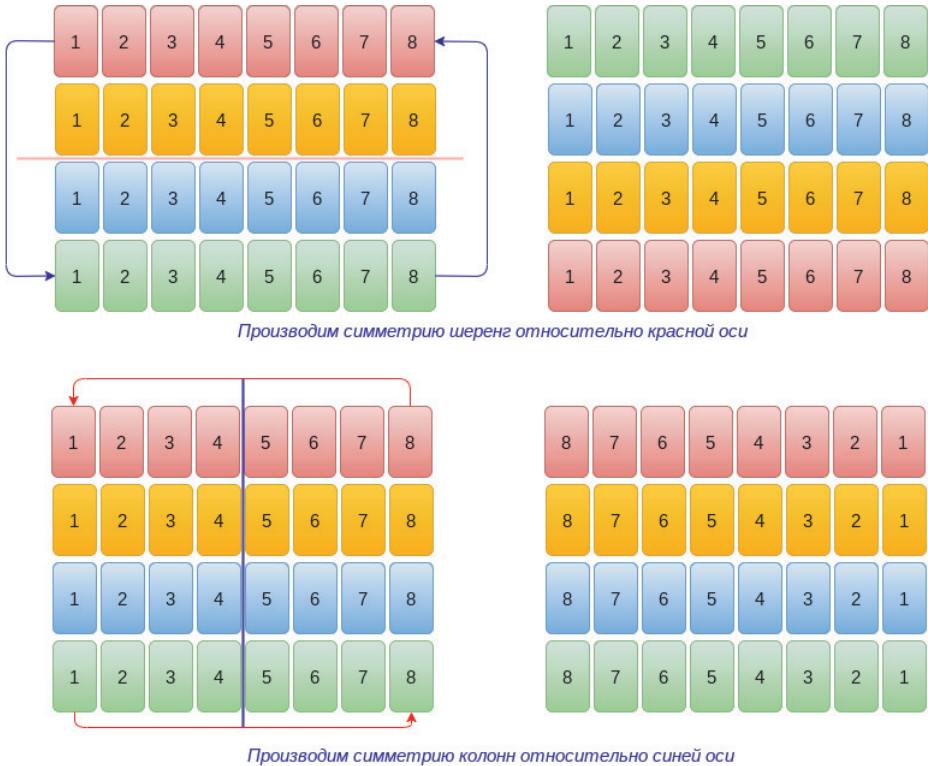


Рис. 6.1. Отражение плоскости.

Вернемся к геометрии. Из приведенной иллюстрации понятно, что на плоскости существует как минимум три вида движений: параллельный перенос на некоторый вектор v , поворот на некоторый угол α относительно центра вращения в некоторой точке O и отражение относительно некоторой оси l . Перенос мы обозначим за T_v , поворот — за R_α^O , отражение — за S_l . Параллельный перенос на вектор v — это такое движение, при котором все точки смещаются на вектор v (т. е. в одном и том же направлении на одно и то же расстояние). Поворот — это такое движение, при котором все точки сдвигаются по концентрическим

окружностям на один и тот же угол в одном и том же направлении. Симметрия относительно оси l — это такое движение, при котором все точки плоскости переходят в симметричные относительно данной оси (т. е. из точки на ось l опускается перпендикуляр, который затем продолжается на такое же расстояние, а все точки l при этом остаются на своих местах).

Для движений плоскости, являющихся параллельным переносом, действует сложение «по параллелограмму». А именно

$$T_v \circ T_u = T_{v+u},$$

где сумма векторов $v + u$ осуществляется по правилу параллелограмма (см. рис. 6.2). Из этого же правила следует, что композиция параллельных переносов есть параллельный перенос, причем порядок слагаемых не имеет значения, т. е. параллельные переносы коммутируют друг с другом.

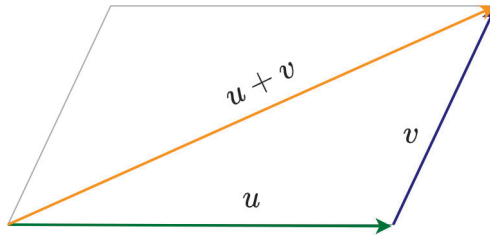


Рис. 6.2. Правило параллелограмма.

Для вращений существует аналогичное правило, если они имеют общий центр вращения:

$$R_\alpha^O \circ R_\beta^O = R_{\alpha+\beta}^O.$$

Ясно также, что в данном случае вращения коммутируют, т. к. они просто выполняют движения любой окружности с центром O .

Заметим также, что и параллельный перенос, и вращение, и отражение имеют обратные преобразования. Точнее, $T_v \circ T_{-v} = \text{id}$, $R_\alpha^O \circ R_{-\alpha}^O = \text{id}$, а также $S_l \circ S_l = \text{id}$. Таким образом, все эти движения обратимы. Отсюда же следует, что они являются взаимно однозначными отображениями, т. е. никакие две точки не переходят в одну и никакая точка не переходит в две и более точек.

Для вывода остальных свойств арифметики движений нам потребуются некоторые дополнительные сведения о движениях плоскости.

Для начала мы рассмотрим случай, когда движение плоскости сохраняет на месте хотя бы одну точку. Назовем такое движение буквой G . Выберем какую-нибудь неподвижную точку движения G и обозначим

ее за O . Теперь возьмем любую окружность с центром O и обозначим ее за S^1 . Что происходит с ней при таком движении? Ясно, что она переходит в себя, поскольку движение сохраняет расстояние, а все точки окружности равноудалены от центра O , и наоборот, все точки, равноудаленные от O , лежат на одной окружности.

Но тогда получается, что движение G порождает движение окружности S^1 . Покажем, что разные движения плоскости, сохраняющие нашу точку O , порождают разные движения окружности S^1 . Пусть движения плоскости G_1 и G_2 оставляют на месте точку O , но при этом не равны, т. е. как минимум одна точка A на плоскости переходит в первом случае в A_1 , а во втором — в A_2 , причем $A_1 \neq A_2$.

Здесь мы обратимся к одному из важнейших методов геометрии: *проецированию*. А конкретно, — к проецированию точки на окружность. Соединим точку O с точкой A и, в случае необходимости, продолжим отрезок до пересечения его с окружностью S^1 так, чтобы точки A и A' (точка пересечения) лежали по одну сторону от центра O . Аналогично построим точки A'_1 и A'_2 (рис. 6.3).

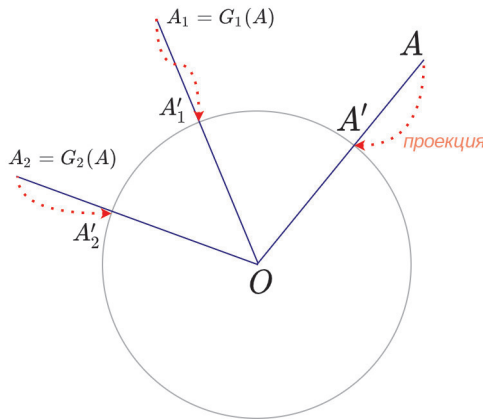


Рис. 6.3. Проецирование на окружность.

Далее заметим, что если бы $A'_1 = A'_2$, то, очевидно, A_1 и A_2 оказались бы на одной прямой, причем по одну сторону от O . Но в силу свойств движения, т. е. равенства расстояний $|OA_1| = |OA| = |OA_2|$, оказалось бы также, что $A_1 = A_2$, что противоречит нашему предположению.

Итак, разные движения плоскости, сохраняющие точку O , порождают и разные движения окружности S^1 с центром O .

Обратно, всякому движению окружности соответствует движение плоскости. Действительно, как мы знаем, движение окружности есть либо вращение, либо отражение относительно оси, проходящей через центр O . В первом случае возьмем в качестве движения G плоскости

вращение относительно центра O на тот же угол, во втором — отражение относительно той же оси.

Сказанное выше означает, что все движения плоскости, имеющие общую неподвижную точку O , взаимно однозначно определяются одноименными движениями окружности с центром в той же точке O . Говоря алгебраическим языком, множество движений плоскости содержит в себе группы, изоморфные группе движений окружности. Этих групп столько, сколько точек на плоскости.

Чтобы дать полное описание группы движений плоскости, нам потребуется описать все виды движений и построить их таблицу композиций так же, как мы это делали в случае движений прямой и окружности.

Редуцирование движений плоскости с неподвижной точкой к движениям окружности сразу же дает нам возможность понять, что такие движения бывают всего двух видов: повороты (в том числе id) и отражения относительно прямых. Причем первый случай получается тогда, когда у движения есть ровно одна неподвижная точка, либо вся плоскость неподвижна. А отражение окружности предполагает неподвижность двух диаметрально противоположных точек. Ясно, что при этом не только эти две точки плоскости останутся неподвижными.

Лемма 6.1. Если движение плоскости сохраняет неподвижными две различные точки, то оно сохраняет неподвижными все точки прямой, проходящей через данные две точки.

Доказательство. Пусть движение G плоскости оставляет на месте точки $A \neq B$.

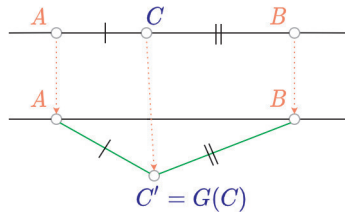


Рис. 6.4. Две неподвижные точки плоскости.

Для начала заметим, что движение G в этом случае прямую AB переводит саму в себя. Действительно, если бы это было не так, то три различные точки A, B, C , лежащие на этой прямой, перешли бы в треугольник ABC' , где $C' = G(C)$. Но тогда нарушается неравенство треугольника, при котором сумма двух его сторон всегда больше третьей, а для трех точек на одной прямой это не так (рис. 6.4).

Следовательно, G порождает движение прямой AB . Но для движения

прямой нам известна теорема «о гвоздях», доказанная в разделе 2.4, из которой следует, что если движение сохраняет две точки на месте, то оно сохраняет все точки этой прямой на месте! \square

Наконец, докажем вариант теоремы «о гвоздях» для плоскости.

Теорема 6.1 («о трех гвоздях»). *Если движение G плоскости оставляет на месте три точки, не лежащие на одной прямой, то $G = \text{id}$.*

Доказательство. Пусть даны три точки A, B, C , не лежащие на одной прямой, такие, что $G(A) = A$, $G(B) = B$ и $G(C) = C$. Возьмем окружность S^1 с центром в точке A . Движение G порождает на этой окружности либо поворот, либо отражение. Построим проекции точек B и C на данную окружность (см. рис. 6.5), получим точки B' и C' .

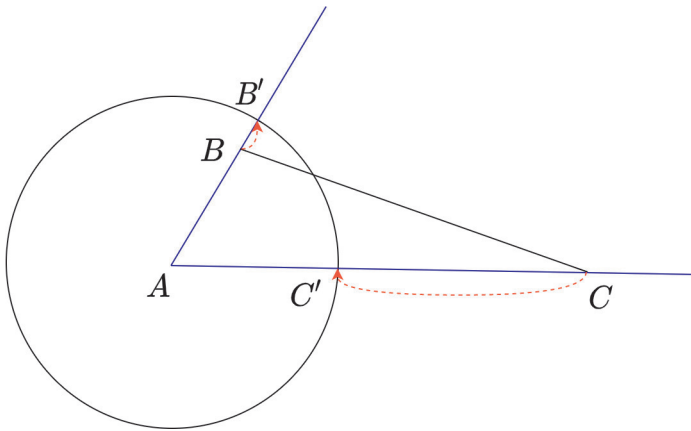


Рис. 6.5. Три неподвижные точки плоскости.

Из доказанной выше леммы следует, что движение G сохраняет на месте прямые AB и AC . Эти прямые различны, т.к. иначе все три точки A, B, C оказались бы на одной прямой. Следовательно, $B' \neq C'$, и эти точки не являются диаметрально противоположными. В то же время, это точки прямых AB и AC , а значит, они являются стационарными точками движения G , которое является, в частности, движением окружности S^1 .

Итак, мы имеем движение окружности, которое сохраняет на месте две различные точки, не являющиеся диаметрально противоположными. А про такое движение мы уже выяснили в разделе 3.2, что оно является id , т.е. поворотом на нулевой угол.

Остается показать, что если какое-то движение оставляет на месте окружность, то оно оставляет на месте и всю плоскость. Действительно, выше было показано, что разные движения плоскости, сохраняющие на месте некоторую точку A , порождают разные движения окруж-

ностей с центром в точке A . Поскольку движение id на окружности порождается движением id плоскости, из того, что рассматриваемое движение G является id на нашей окружности S^1 , следует, что оно является id и на всей плоскости. \square

Движение плоскости различные точки переводит в различные точки. Действительно, если бы это было не так, и мы бы имели $G(A) = G(B)$, хотя $A \neq B$, что противоречит принципу сохранения расстояний при движениях: $|AB| = |G(A)G(B)|$. Итак, движение плоскости является инъективным (одно-однозначным) отображением.

Посмотрим, что происходит, когда у движения G нет неподвижных точек. В этом случае мы не можем напрямую применить наши предыдущие рассуждения, выбрав окружность с неподвижным центром. Однако мы можем свести ситуацию к предыдущей следующим способом.

Возьмем произвольную точку O и положим $A = G(O)$. По предположению $A \neq O$. В таком случае, мы можем построить серединный перпендикуляр отрезка AO , который мы обозначим за l , и рассмотрим отражение S_l плоскости относительно него. Тогда, очевидно, композиция движений $S_l \circ G$ сохраняет на месте точку O , и мы оказываемся в рассмотренной ранее ситуации, только вместо движения G у нас теперь движение $S_l \circ G$.

Это значит, что движение $S_l \circ G$ является либо поворотом R_α^O с центром в точке O на угол α , либо отражением S_m относительно прямой m , проходящей через точку O . Но тогда, применяя слева симметрию S_l , находим, что

$$G = S_l \circ R_\alpha^O, \text{ либо } G = S_l \circ S_m.$$

В свою очередь, поворот R_α^O , если его рассматривать как движение окружности S^1 с центром O , как мы ранее выяснили, является композицией двух отражений относительно осей n, k , угол между которыми равен $\alpha/2$. Таким образом,

$$G = S_l \circ S_n \circ S_k, \text{ либо } G = S_l \circ S_m.$$

Итак, в том случае, когда движение не сохраняет на месте ни одной точки, оно является композицией двух или трех отражений.

Теорема 6.2. *Всякое движение плоскости есть композиция не более чем трех отражений.*

Отсюда, в частности, следует, что всякое движение плоскости является биекцией, т. е. взаимно однозначным отображением.

Выше мы дали обозначения трем видам движений плоскости: параллельный перенос, поворот и отражение относительно оси. Возникает

вопрос: исчерпываются ли движения плоскости только такими движениями или есть какие-то еще?

Пусть снова движение G таково, что оно не сохраняет на месте ни одной точки. Возьмем произвольную точку A и введем следующие обозначения: $B = G(A)$, $C = G(B)$. Поскольку G — движение (изометрия), имеем: $AB = BC$.

Предположим, что C лежит на прямой AB . Тогда вся прямая AB под действием движения G переходит в себя. Действительно, если D лежит на AB между A и B и переходит в D' вне прямой AB , то нарушится равенство $|AD| + |BD| = |BD'| + |CD'|$. Аналогично, нарушается равенство $|AD| - |BD| = |BD'| - |CD'|$ в случае, когда D лежит на прямой AB , но не между точками A и B .

Кроме того, если C лежит на прямой AB , то точки располагаются в порядке $A < B < C$. Действительно, поскольку $|AB| = |BC|$, то точка C может лежать либо симметрично точке A относительно B , и тогда $A < B < C$, либо $C = A$, но в этом случае середина отрезка AB окажется неподвижной точкой движения G , что противоречит условию.

Предположим далее, что C не лежит на прямой AB . В этом случае обозначим за X середину отрезка AB , за Y — середину отрезка BC , и через точки X, Y проведем прямую L . Необходимо показать, что L переходит в себя под действием движения G .

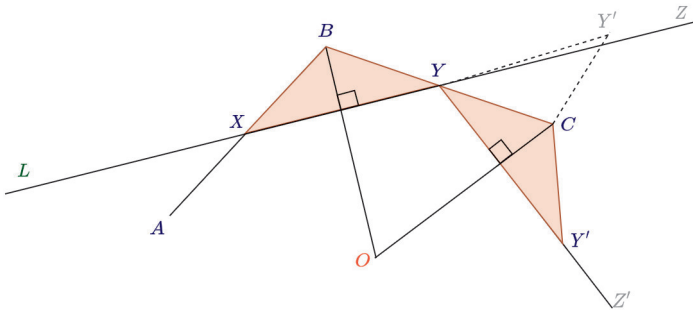


Рис. 6.6. Поиск прямой, переходящей в себя.

Пусть $Y' = G(Y)$. Ясно, что $X'Y = Y'Y'$, причем $X \neq Y'$, иначе бы середина $X'Y$ оказалась стационарной точкой G . Точки X и Y' могут лежать по одну или по разные стороны прямой BC . Если они лежат по разные стороны, то рассмотрим треугольники $\triangle X'Y'B$ и $\triangle Y'Y'C$ (см. пунктир на рис. 6.6). Они равны по трем сторонам в силу того, что G является движением, кроме того, они оба равнобедренные с основаниями $X'Y$ и $Y'Y'$, соответственно. Но тогда равны углы $\angle X'Y'B = \angle Y'Y'C$ при вершине Y' , отложенные в разные стороны прямой BC , т.е. эти углы являются вертикальными, откуда следует, что точка Y' находится

ся на прямой L . Мы снова получаем ситуацию трех точек в порядке $X < Y < Y'$, откуда следует, что вся прямая L переходит в себя под действием G .

Наконец, предположим, что X и Y' оказались по одну сторону прямой BC (см. рис. 6.6). В этом случае мы можем построить серединные перпендикуляры отрезков XY и YY' , которые пересекутся в некоторой точке O . Пусть $O' = G(O)$. Треугольник XOY равнобедренный, он переходит под действием G в треугольник $YO'Y'$, также равнобедренный. Это значит, что точка O' лежит на серединном перпендикуляре отрезка YY' , причем на таком же расстоянии от середины YY' , как и точка O от середины XY . Но отрезок BO пересекает отрезок XY . Значит, в силу свойств движения, отрезок CO пересекает отрезок YY' . И тогда для точки O' остается единственная возможность: $O = O'$. А это противоречит тому, что G не имеет стационарных точек.

Итак, предположение о том, что G не имеет неподвижных точек, приводит к тому, что найдется прямая L , которая под действием G переходит в себя. Это значит, что G действует на L как сдвиг (отражение имело бы неподвижную точку, а других движений прямой не существует).

Сдвиг на L задается вектором XY (или AB , если мы остались в первом варианте). Составим композицию $G \circ T_{-XY}$. Очевидно, что такое движение плоскости оставляет на месте прямую L . Как мы уже выяснили, это соответствует либо id , либо осевой симметрии S_L . Таким образом, $G = T_{XY}$, либо $G = S_L \circ T_{XY}$, причем вектор XY параллелен прямой L .

Итак, движение G , которое не сохраняет на месте ни одной точки, является либо параллельным переносом, либо нетривиальным параллельным переносом с последующим отражением относительно оси, параллельной вектору переноса. Во втором случае движение G называется **скользящей симметрией**.

Легко понять, что скользящая симметрия не сводится, вообще говоря, к другим видам движений. Действительно, id , поворот и осевая симметрия исключаются из-за наличия неподвижных точек. Параллельный перенос исключается, поскольку при нем отрезок AA' , где $A' = G(A)$ не может пересекать никакую прямую, параллельную направлению переноса, а при скользящей симметрии он пересекает ось симметрии (если A не лежит на этой оси).

Скользящую симметрию с осью симметрии l и параллельным переносом вдоль этой оси на вектор u будем обозначать W_u^l .

Для единства обозначений предусмотрим также вырожденный случай случай скользящей симметрии, а именно с нулевым вектором переноса, т. е. W_0^l , и положим ее равной обычному отражению S_l . Иначе

говоря, отражение относительно оси мы теперь считаем частным случаем скользящей симметрии точно так же, как id есть частный случай параллельного переноса.

Заметим также, что в случае скользящей симметрии ее компоненты коммутируют, т. е. если $W_u^l = S_l \circ T_u$, то $W_u^l = T_u \circ S_l$, это легко проверить геометрически.

Окончательно получаем следующую теорему.

Теорема 6.3 (Шаля). *Всякое движение плоскости (без разложения его на компоненты) есть движение одного из следующих классов:*

- a) *класс параллельных переносов (на произвольный вектор, в том числе нулевой), который мы обозначим символом \Rightarrow ;*
- b) *класс поворотов относительно произвольного центра, который мы обозначим символом \circ ;*
- c) *класс **скользящих симметрий**, который мы обозначим символом $\leftarrow\leftarrow$.*

Таблицу композиций для описанных в теореме Шаля классов см. в табл. 6.1.

Таблица 6.1. Таблица композиций движений плоскости

	\Rightarrow	\circ	$\leftarrow\leftarrow$
\Rightarrow	\Rightarrow	\circ	$\leftarrow\leftarrow$
\circ	\circ	\Rightarrow или \circ	$\leftarrow\leftarrow$
$\leftarrow\leftarrow$	$\leftarrow\leftarrow$	$\leftarrow\leftarrow$	\Rightarrow или \circ

Наша дальнейшая задача — обосновать таблицу 6.1, построив полную таблицу композиций движений плоскости.

Итак, нам нужно найти все попарные комбинации композиций переноса, поворота и скользящей симметрии, причем скользящую симметрию мы рассмотрим в двух вариантах — как отражение относительно оси и как собственно скользящую симметрию.

I. Ясно, что $T_u \circ T_v = T_{u+v}$, причем данная композиция коммутативна.

II. Рассмотрим композицию $T_v \circ R_\alpha^O$. Из свойств движений окружности мы знаем, что поворот можно представить как композицию двух отражений, а из свойств движений прямой — что сдвиг тоже можно представить как композицию двух отражений. Пусть $R_\alpha^O = S_2 \circ S_1$ и $T_v = S_4 \circ S_3$. При этом оси симметрий 1 и 2 пересекаются в точке O , а угол между ними равен $\alpha/2$ и откладывается от оси 1 к оси 2. Оси 3 и 4 параллельны друг другу и перпендикулярны направлению сдвига, причем расстояние между ними равно $v/2$, а направление сдвига — от оси 3 к оси 4.

Заметим, что выбор осей 1 и 2 произволен в рамках требований

общей точки пересечения и угла между ними. Сама пара осей может быть повернута как угодно, в частности, можно выбрать ось 2 так, чтобы она оказалась перпендикулярной вектору сдвига v . Кроме того, выбор осей 3 и 4 также произволен в рамках требований расстояния между ними и строго перпендикулярного положения к вектору сдвига. Саму пару осей можно двигать вдоль вектора v как угодно. Поэтому совместим ось 3 с осью 2.

Имеем следующее:

$$T_v \circ R_\alpha^O = (S_4 \circ S_3) \circ (S_2 \circ S_1) = S_4 \circ (S_3 \circ S_2) \circ S_1 = S_4 \circ S_1,$$

где мы сократили $S_3 \circ S_2$, т. к. это одно и то же отражение. Полученная композиция $S_4 \circ S_1$ является поворотом на тот же угол α , только с новым центром O_1 , который мы получили как пересечение осей 1 и 4 (см. рис. 6.7).

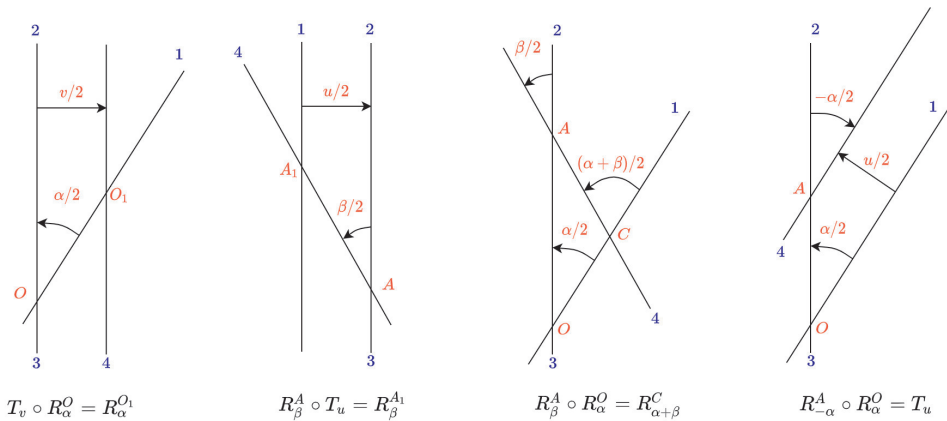


Рис. 6.7. Композиция поворотов и переносов.

III. Аналогично поступаем в случае композиции $R_\beta^A \circ T_u$ (см. рис. 6.7). При этом новый центр поворота смещается на вектор $-u/2$ и перпендикулярно ему на величину $|u/2|/\tan(\beta/2)$. Видим также, что композиция сдвига и поворота не коммутативна (кроме случая, когда перенос осуществляется на нулевой вектор).

IV. Рассмотрим композицию двух поворотов $R_\beta^A \circ R_\alpha^O$. Снова представим каждый поворот как композицию отражений, и снова, пользуясь произвольностью наклона пары осей отражений, расположим их так, чтобы вторая ось первого поворота и первая ось второго поворота совпали с прямой OA (см. рис. 6.7). Если углы таковы, что их сумма не кратна 2π , то оси 1 и 4 пересекутся, и точка C их пересечения окажется центром результирующего поворота. А угол поворота станет равным $\alpha + \beta$ (т. к. угол, измеренный от оси 1 к оси 4, есть дополнительный к углу ACO , а значит, равен сумме углов $\angle A$ и $\angle O$ треугольника ACO).

Если же $\alpha + \beta$ кратно 2π , то оси 1 и 4 будут параллельны, и композиция поворотов превратится в параллельный перенос на расстояние $2|AO| \cdot |\sin(\alpha/2)|$.

V. Рассмотрим композицию $G = T_v \circ S_l$. Чтобы понять, что это такое, воспользуемся тем же приемом, с помощью которого мы показали наличие еще одного вида движений плоскости — скользящей симметрии. Именно, возьмем на прямой l точку A и проследим за ее судьбой при двух итерациях преобразования G . Получатся точки $B = G(A)$ и $C = G(B)$. Далее, через точки X и Y , являющиеся серединами отрезков AB и BC , проведем прямую m . Эта прямая при преобразовании G переходит сама в себя, как было показано ранее. Все остальные точки сначала отражаются относительно m , затем смещаются в направлении вектора \vec{AC} , но на половину его длины. Обозначим $w = \vec{AC}/2$. Тогда получим, что $G = T_w \circ S_m$.

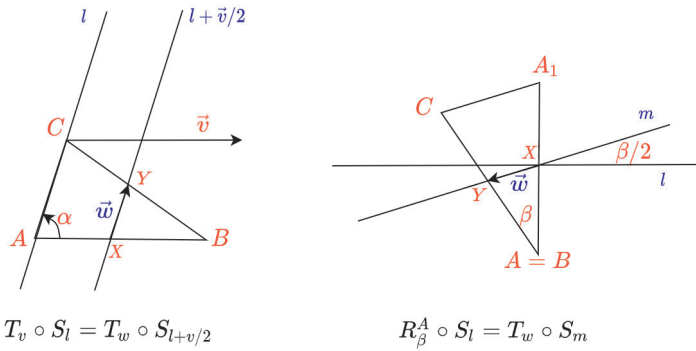


Рис. 6.8. Композиция отражения с переносом и поворотом.

Условимся записывать $w = \text{Pr}_l v$, что означает проекцию вектора v на ось l , а также $m = l + v/2$, что означает сдвиг оси l в направлении v на половину его длины. Тогда получим, что

$$T_v \circ S_l = T_{\text{Pr}_l v} \circ S_{l+v/2},$$

а это не что иное, как скользящая симметрия.

VI. В симметричном случае получим

$$S_l \circ T_u = T_{\text{Pr}_l u} \circ S_{l-u/2},$$

т. е. тоже скользящая симметрия с тем же сдвигом, но с симметричной относительно l осью симметрии.

Все сказанное справедливо, если угол наклона l не прямой и не нулевой относительно вектора v . Если окажется, что $l \perp v$, то $G = S_{l+v/2}$, поскольку смещение $\text{Pr}_l v = 0$. Если же $l \parallel v$, то исходное преобразование G уже есть скользящая симметрия с осью l и смещением v . Эти

случаи вписываются в общую формулу, если иметь в виду, что $l+v/2 = l$ в том случае, когда $l \parallel v$.

VII. Рассмотрим композицию $G = R_\beta^A \circ S_l$. В качестве отправной точки A_1 возьмем точку, симметричную центру вращения A относительно l . Тогда A_1 переходит в $A = B$, а далее B переходит в C (см. рис. 6.8). Снова проводим среднюю линию XY в полученном равнобедренном треугольнике. Тогда вектор $w = \overrightarrow{XY}$ есть смещение новой скользящей симметрии, а ось m , полученная вращением оси l вокруг точки X на угол $\beta/2$ — осью скользящей симметрии.

$$R_\beta^A \circ S_l = T_w \circ S_m, \quad w = 2\text{Pr}_m \overrightarrow{XA}, \quad X = \text{Pr}_l A.$$

Заметим, что в том случае, когда $A \in l$, получаем $w = 0$, и композиция G становится просто симметрией S_m .

VIII. В обратном случае получим

$$S_m \circ R_\alpha^O = T_{-w} \circ S_k, \quad w = 2\text{Pr}_k \overrightarrow{XO}, \quad X = \text{Pr}_m O.$$

Для внесения данных в итоговую таблицу обозначим за $l + \beta/2$ ось m , полученную из оси l поворотом на угол $\beta/2$ относительно точки X — проекции центра вращения A на ось l . Аналогично, $m - \alpha/2$. Тогда

$$R_\beta^A \circ S_l = T_w \circ S_{l+\beta/2}, \quad S_m \circ R_\alpha^O = T_{-w} \circ S_{m-\alpha/2}.$$

IX. Композиция двух симметрий $S_m \circ S_l$, как и в случае окружности, есть поворот $R_{2\angle lm}^{l \cap m}$, где $\angle lm$ есть угол от оси l до оси m , а центр вращения есть общая точка этих осей. Если угол между осями равен нулю, и при этом они не совпадают, а просто параллельны друг другу, то, как мы уже знаем, их композиция есть параллельный перенос в направлении от l к m (перпендикулярно им) на удвоенное расстояние между осями симметрии. Такой перенос обозначим $T_{2(m-l)}$.

Осталось рассмотреть композиции со скользящей симметрией.

X. Пусть $G = T_v \circ W_u^l$. Очевидно, что это на самом деле композиция $T_{u+v} \circ S_l$, и мы приходим к уже рассмотренному случаю композиции переноса и симметрии.

XI. Обратное, если G есть композиция $W_v^m \circ T_u$, то поскольку в скользящей симметрии перенос и симметрия коммутируют, запишем G как $S_m \circ T_{v+u}$, и мы снова имеем дело с изученной ситуацией.

XII. Пусть далее $G = R_\beta^A \circ W_u^l$. В таком случае

$$G = (R_\beta^A \circ T_u) \circ S_l = R_\beta^{A_1} \circ S_l,$$

т. е. снова приходим к известному случаю с заменой лишь центра вращения.

XIII. В обратном случае $W_v^m \circ R_\alpha^O$ все аналогично.

XIV. Наконец, пусть $G = S_m \circ W_u^l$. Здесь можно пойти разными путями для получения результата. Проще выглядит путь, при котором мы сначала найдем композицию двух симметрий:

$$G = (S_m \circ S_l) \circ T_u = \begin{cases} R_{2\angle ml}^O, & l \not\parallel m, \\ T_{2(m-l)+u}, & l \parallel m. \end{cases}$$

В первом случае центр симметрии O зависит от вектора u .

XV. Аналогично,

$$W_v^m \circ S_l = T_v \circ (S_m \circ S_l) = \begin{cases} R_{2\angle ml}^O, & l \not\parallel m, \\ T_{2(m-l)+v}, & l \parallel m, \end{cases}$$

где центр симметрии O зависит от вектора v .

XVI. Осталось рассмотреть случай композиции двух скользящих симметрий:

$$W_v^m \circ W_u^l = T_v \circ (S_m \circ S_l) \circ T_u = \begin{cases} R_{2\angle ml}^O, & l \not\parallel m, \\ T_{2(m-l)+u+v}, & l \parallel m, \end{cases}$$

где центр симметрии O зависит уже от обоих переносов u и v .

Сведем все полученные результаты в таблицу 6.2, в которой мы опустили явные выражения для некоторых векторов смещения и центров вращения для упрощения записи. Напомним также, что таблицу композиций следует читать «от строки к шапке», т. е. если строка начинается с S_m , столбец озаглавлен T_u , то на их пересечении стоит выражение для композиции $S_m \circ T_u$.

6.2. Сравнение движений прямой, окружности и плоскости

Отметим несколько общих свойств рассмотренных нами движений прямой, окружности и плоскости.

Во-первых, их всех можно свести к композиции отражений. Для одномерных объектов (прямая и окружность) — не более двух, для двумерных — не более трех.

Во-вторых, все движения можно разделить на два класса: сохраняющие и меняющие **ориентацию**. Те движения, которые сводятся к композиции четного числа симметрий, сохраняют ориентацию фигур, а те, которые сводятся к композиции нечетного числа симметрий, — меняют ориентацию фигур. Изменение ориентации означает, что право и лево меняются местами, т. е. мы как бы переходим в зазеркалье.

Таблица 6.2. Таблица композиций движений плоскости

id	T_u	R_α^O	S_l	W_u^l
T_u	T_{u+v}	$R_\alpha^{O_1}$	$W_{Prlv}^{l+v/2}$	$W_{Prl[u+v]}^{l+(u+v)/2}$
R_β^A	$R_\beta^{A_1}$	$R_{\alpha+\beta}^C$ ($\alpha + \beta \neq 0$) T_w ($\alpha + \beta = 0$)	$W_w^{l+\beta/2}$ ($A \notin l$) $S_{l+\beta/2}$ ($A \in l$)	$W_w^{l+\beta/2}$ ($A_1 \notin l$) $S_{l+\beta/2}$ ($A_1 \in l$)
S_m	$W_{Prlu}^{m-u/2}$	$W_w^{m-\alpha/2}$ ($O \notin m$) $S_{m-\alpha/2}$ ($O \in m$)	$R_{2\angle lm}^{l \cap m}$ ($l \nparallel m$) $T_{2(m-l)}$ ($l \parallel m$)	$R_{2\angle lm}^O$ ($l \nparallel m$) $T_{2(m-l)+u}$ ($l \parallel m$)
W_v^m	$W_{Prl[u+v]}^{m-(u+v)/2}$	$W_w^{m-\alpha/2}$ ($O_1 \notin m$) $S_{m-\alpha/2}$ ($O_1 \in m$)	$R_{2\angle lm}^O$ ($l \nparallel m$) $T_{2(m-l)+v}$ ($l \parallel m$)	$R_{2\angle lm}^O$ ($l \nparallel m$) $T_{2(m-l)+u+v}$ ($l \parallel m$)

При этом нужно отметить, что преобразования, меняющие ориентацию, обязательно требуют выхода в пространство более высокой размерности (для прямой — в плоскость, а для окружности и плоскости — в пространство размерности 3), если мы хотим осуществить их непрерывным движением.

В-третьих, есть и более глубинная связь движений прямой, окружности и плоскости. Мы уже отмечали, что окружность можно рассматривать как прямую, у которой склеили противоположные концы (где-то на бесконечности). И с этой точки зрения сдвиг на прямой является прямой аналогией вращения окружности — особенно в том случае, когда величина сдвига сильно меньше радиуса.

А отражение прямой при этом естественным образом превращается в отражение окружности. Только ось отражения должна проходить через место склейки двух бесконечностей. Остальные же отражения можно получить дополнительным сдвигом, т. е. вращением.

Далее, окружность находится на плоскости. И поэтому вращение окружности полностью аналогично вращению плоскости, если при этом совместить центры обоих вращений.

Еще проще увидеть совпадения понятий сдвига прямой и плоскости. В обоих случаях мы просто смещаем все точки на какой-то заданный вектор.

Тем не менее на плоскости появляется новый вид движения, который объединяет сдвиг и отражение относительно оси сдвига. Это — скользящая симметрия, т. е. отражение с последующим применением сдвига вдоль оси отражения. На одномерных объектах такое движение в принципе невозможно. На прямой отражение относительно этой же прямой ничего не дает, т. е. является *id*, а на окружности отражение относительно самой окружности вообще требует специального построения в геометрии плоскости (это преобразование называется *инверсией*).

6.3. Пара слов о движениях сферы

Имея опыт перехода от прямой к окружности, мы можем легко найти движения сферы, отправляясь от движений плоскости.

Представим себе сферу как плоскость, у которой бесконечно удаленный край был стянут в точку (метод «хинкали»). Во что превращаются при этом движения плоскости?

Сдвиг, он же параллельный перенос, превращается в такое движение, при котором все точки движутся по параллельным траекториям. С точки зрения географии это есть движение вдоль широтных линий. Да, проходят они при этом разное расстояние! Из-за чего, кстати, и

появляются силы Кориолиса, создающие океанические течения вроде Гольфстрима. Но расстояния между точками сохраняются, и это, несомненно, движение.

Вращение, которое, как мы помним, на окружности соответствует сдвигу на прямой, в случае сферы в прямом смысле слова совпадает со сдвигом! Дело в том, что вращение сферы вокруг оси — это вращение вокруг полюса, при котором угол поворота измеряется меридианом. Но ведь это же самое движение около экватора есть ровно то, что мы только что отнесли к сдвигам вдоль широтных линий.

Таким образом, сдвиг и вращение в случае сферы чудесным образом объединяются в один вид движений — *осевое вращение*. И это делает движения сферы чуть проще, чем движения плоскости, где сдвиг можно представить лишь как композицию двух вращений.

Далее, симметрия плоскости относительно прямой естественным образом переходит в *отражение* сферы относительно центральной секущей плоскости или, иначе говоря, относительно окружности большого круга. При такой симметрии полюса сферы меняются местами (полюса определяются пересечением со сферой прямой, пересекающей плоскость отражения в центре сферы и перпендикулярной ей), а плоскость отражения остается на месте.

Наконец, скользящая симметрия плоскости есть композиция сдвига и осевой симметрии, и ей на сфере соответствует *зеркальное вращение*, т. е. композиция отражения и вращения параллельно плоскости отражения.

Таким образом, все движения сферы распадаются на два класса: вращения и зеркальные вращения. При этом все движения есть композиция не более чем трех отражений.

Этот аналог теоремы Шаля для сферы можно доказать, используя очередную лемму о гвоздях, предполагая неподвижность пары противоположных точек (случай одной точки на плоскости), неподвижность целой окружности большого круга (случай двух точек на плоскости), отсутствие неподвижных точек.

6.4. Пара слов о движениях пространства

Наконец, мы можем от сферы перейти к пространству. На самом деле переход в пространство сопровождается лишь добавлением сдвига в пространстве, т. е. любое движение сферы можно рассматривать как движение пространства с одной неподвижной точкой — центром сферы. После чего можно применить сдвиг этого центра, и получить новые движения. Понятно, что никаких других движений тут быть не может.

Тем не менее классификация движений пространства становится сложнее примерно на столько же, насколько классификация движений плоскости превосходит классификацию движений окружности (и даже больше). А именно, в пространстве появляется **винтовое движение** (*винт*) как композиция осевого вращения и сдвига вдоль оси вращения. Это — обобщение скользящей симметрии на плоскости (если винт осуществляет поворот на 180° , мы как раз получаем скользящую симметрию).

Есть также и собственно **скользящая симметрия пространства**. Это — отражение относительно плоскости с последующим сдвигом вдоль направления, параллельного данной плоскости. Такое движение также является обобщением скользящей симметрии на плоскости.

Заметим, что более сложное движение — винт — включает в себя более простые. Так, если винт имеет нулевой сдвиг, то он сводится к осевому вращению, а если винт имеет нулевой поворот, то он сводится к сдвигу. Понятно, что в случае полного зануления параметров винта мы получим *id*.

Точно так же зеркальное вращение, как и в случае сферы, при нулевом повороте превращается в отражение.

Наконец, скользящая симметрия своим частным случаем имеет просто отражение относительно плоскости.

Таким образом, классификация движений пространства включает следующие виды движений:

- винт (в частности, сдвиг, осевое вращение, *id*);
- зеркальное вращение (в частности, отражение);
- скользящая симметрия (в частности, отражение).

Таблица 6.3. Сравнение движений

		Собственные движения (не меняют ориентацию)		Несобственные движения (меняют ориентацию)	
	Савиг	Поворот	Смещение поворота	Отражение	Смещенная симметрия
Прямая	савиг на чис- ло			относитель- но точки	
Окруж- ность		вращение		осевая симметрия	
Плоскость	параллель- ный перенос	относитель- но точки		осевая симметрия	скользящая симметрия (перенос+ симметрия)
Сфера	вращение вблизи экватора	вращение вблизи полюса		отражение относитель- но плоскости	зеркальное вращение (вращение + отражение)
Прост- ранство	параллель- ный перенос	вращение осевое	винт (перенос + вращение)	отражение относитель- но плоскости	скользящая симметрия (перенос+ отражение)

Упражнения

Обязательные упражнения

6.1° Докажите, что: а) параллельный перенос; б) поворот; в) осевая симметрия является движением.

6.2° Доказать, что параллельные переносы образуют группу.

6.3° Каковы конечные подгруппы группы параллельных переносов плоскости?

6.4° Доказать, что концентрические повороты плоскости образуют группу.

6.5° Каковы конечные подгруппы группы концентрических поворотов плоскости?

6.6° Доказать, что скользящие симметрии с параллельными осями отражения образуют группу.

6.7° Каковы конечные подгруппы группы скользящих симметрий с параллельными осями?

6.8° Назовем движение собственным, если оно сохраняет ориентацию. Какие движения являются собственными? Доказать, что собственные движения образуют нормальную подгруппу в группе движений.

6.9° Движение G переводит точку $(0, 0)$ в точку $(1, 0)$, а точку $(1, 0)$ — в $(1, 1)$. Может ли G быть: а) параллельным переносом; б) поворотом; в) скользящей симметрией? Укажите параметры соответствующих движений.

6.10° Вычислить движения:

а) сдвиг на вектор $(1, 0)$ с последующим поворотом относительно $O = (0, 0)$ на угол π ;

б) сдвиг на вектор $(0, 1)$ с последующей скользящей симметрией с осью Ox и сдвигом на вектор $(a, 0)$;

в) сдвиг на вектор $(1, 1)$ с последующим поворотом на угол π относительно точки $(0.5, 0.5)$;

г) поворот на угол $\pi/2$ относительно O с последующим сдвигом на вектор $(1, 1)$;

д) поворот на угол $\pi/3$ относительно O с последующей скользящей симметрией с осью OA , где $A = (1, 0.5)$ и сдвигом на вектор $(-1, 0)$;

f) поворот на угол $\pi/3$ относительно O с последующим поворотом на угол $\pi/3$ относительно $(0.5, 1)$.

6.11° Представить движение в виде композиции симметрий:

a) параллельный перенос на вектор $(2, 2)$;

b) поворот на угол $\pi/2$ относительно точки $(1, 1)$;

c) скользящая симметрия с осью Oy и сдвигом вдоль нее на вектор $(-1, 0)$.

6.12° Доказать, что движение a) прямую переводит в прямую; b) треугольник переводит в треугольник; c) окружность переводит в окружность; d) отрезок переводит в отрезок (при этом внутренние точки — во внутренние).

6.13° Вывести формулы таблицы композиций (хотя бы частично).

6.14° Найти композиции движений:

$$T_v^l \circ T_v^l, \quad T_v^l \circ T_v^l \circ T_v^l.$$

6.15° Всегда ли $U \circ V = V \circ U$, если U, V — движения плоскости? Привести поясняющие примеры.

6.16° Всегда ли композиция движений есть движение?

6.17° Пусть $W = U \circ V$, где V — поворот на угол $\pi/3$ с центром O , а U — сдвиг на вектор v . Найти такую точку x , что $W(x) = O$.

6.18° Пусть теперь $W = V \circ U$, где U, V — из предыдущей задачи. Решите уравнение $W(x) = O$.

6.19° К какому классу относятся движения W из двух предыдущих задач?

6.20° Пусть даны поворот R_α^O и сдвиг T_v , $\alpha \neq 0$, $v \neq 0$.

a) Построить такой равнобедренный треугольник $\triangle OAB$ с вершиной O , что $R_\alpha^O(A) = B$ и $\overrightarrow{BA} = v$.

b) Доказать, что A является неподвижной точкой композиции $T_v \circ R_\alpha^O$.

c) Доказать, что B является неподвижной точкой композиции $R_\alpha^O \circ T_v$.

6.21° Найдите обратное преобразование к $R_\alpha^O \circ T_v$ и $T_v \circ R_\alpha^O$. Какие точки x при этих обратных преобразованиях переходят в точку O ?

6.22° Чем является композиция двух осевых симметрий относительно: a) двух перпендикулярных прямых; b) двух параллельных прямых; c) двух прямых, образующих угол $\pi/3$?

6.23° Докажите, что композиция отражения относительно прямой l и параллельного переноса на вектор v является: а) осевой симметрией, если $v \perp l$; б) скользящей симметрией в любом случае.

6.24° Чем является композиция поворота с центром O на угол α и отражения относительно прямой l , если $O \in l$?

6.25° Найти композицию отражения относительно вертикальной оси и поворота на 180° относительно точки, не лежащей на оси симметрии.

6.26° (Теорема Наполеона) Пусть дан произвольный треугольник $\triangle ABC$. На его сторонах построим правильные треугольники и назовем их центры A', B', C' . Доказать, что полученный треугольник $\triangle A'B'C'$ — правильный. *Указание:* рассмотреть два вращения $R_{120^\circ}^{A'}$ и $R_{120^\circ}^{B'}$ и доказать, что $R_{120^\circ}^{B'} \circ R_{120^\circ}^{A'}(C') = C'$.

Линейные уравнения

Аннотация

Основная задача данной главы — дать полное описание решений линейных уравнений в целых числах. Попутно вводится уравнение прямой на координатной плоскости, хотя мы все еще подразумеваем, что работаем только с целыми числами.

7.1. Уравнение прямой на плоскости

Рассмотрим плоскость с координатными осями Ox и Oy . Что будет, если ее начать поворачивать? Во что переходит при этом ось Ox ?

Поскольку вращение — это движение, расстояние между точками сохраняется, и значит, никакие три точки, лежащие на прямой Ox , при повороте не могут перейти в точки, образующие невырожденный треугольник — они снова лягут на прямую, причем в том же самом порядке. Стало быть, Ox при вращении плоскости переходит в некоторую прямую.

Пусть центром вращения является точка $O = (0, 0)$ и ось Ox при вращении R_φ переходит в прямую l . Ясно, что l также проходит через начало координат O , т. к. это — стационарная точка вращения. Ранее (см. раздел 3.2) мы построили таблицу композиций движений окружности, из которой видно, что вращения образуют подгруппу группы движений окружности, т. е. все движения вида R_φ , где $0 \leq \varphi < 2\pi$, окружность переводят в окружность, сохраняя расстояния, при этом композиция вращений есть вращение, каждое вращение R_φ имеет обратное $R_{2\pi-\varphi}$, кроме того, вращения коммутируют, т. е. образуют абелеву группу.

Фиксируем на Ox точку $(1, 0)$ и посмотрим, куда она переходит под действием всех возможных вращений с центром O , т. е. описанной выше группы вращений. Поскольку расстояние от центра вращения сохраняется, ясно, что эта точка остается на окружности радиуса 1. В то же время, выбирая произвольную точку на этой окружности, мы легко укажем угол φ , на который нужно осуществить поворот плоскости

относительно центра O , чтобы точка $(1, 0)$ перешла в выбранную нами точку.

Итак, под действием группы вращений точка $(1, 0)$ переходит во все точки единичной окружности. Аналогично, если мы выберем произвольную точку $(r, 0)$ ($r > 0$), она будет переходить во все точки окружности радиуса r под действием группы вращений с центром в точке O .

В этом случае принято говорить, что группа вращений **действует** на плоскости, а множество всех значений, в которые она переводит выбранную точку, называют **орбитой** этой точки. В нашем примере орбитами являются концентрические окружности с центром O .

Можно доказать, что орбиты образуют классы эквивалентности, т. е. они попарно не пересекаются и в объединении дают всю область действия группы.

Фиксируем некоторое вращение R_φ , и пусть точка $(1, 0)$ при таком вращении перешла в точку $C = (x_0, y_0)$, лежащую на единичной окружности.

Возьмем произвольную точку $(r, 0)$, где $r > 0$, и проследим ее судьбу под действием того же вращение R_φ . Пусть $A = (x, y) = R_\varphi(r, 0)$. Ясно, что точки O, C, A лежат на одной прямой l .

Проведем вертикальные линии через абсциссы x_0 и x , а также горизонтальные линии через ординаты y_0 и y . Добавим новые точки пересечения B и D (см. рисунок 7.1).

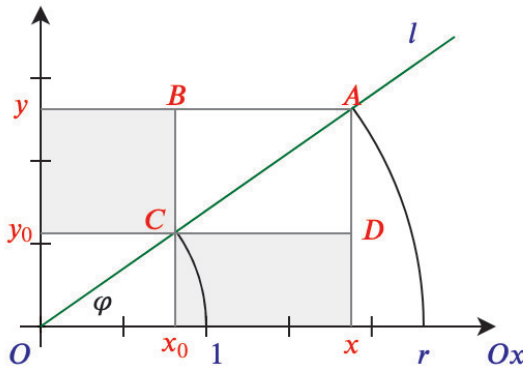


Рис. 7.1. Поворот оси Ox на угол φ .

Видим, что треугольники ABC и ADC равны по трем сторонам, также равны треугольники Oy_0C и Ox_0C и треугольники OyA и OxA . Отсюда легко установить равенство площадей $x_0(y - y_0) = y_0(x - x_0)$, откуда получаем

$$xy_0 - yx_0 = 0. \tag{7.1}$$

Итак, мы выбрали произвольное число r , взяли точку $(r, 0)$ на оси Ox , затем взяли ее образ относительно поворота R_φ и получили точку (x, y) на прямой l , для которой вывели уравнение (7.1). Но в силу произвольности выбора r (в том числе и отрицательных) и того факта, что все точки прямой l получаются как образы точек вида $(r, 0)$ при действии поворота R_φ , мы заключаем, что уравнение (7.1) описывает геометрическое положение всех точек прямой l .

Кроме того, уравнение (7.1) можно переписать в эквивалентной форме

$$\frac{x}{x_0} = \frac{y}{y_0},$$

если предположить, что $x_0 \neq 0$ и $y_0 \neq 0$. Из этого соотношения следует, что соответствующие прямоугольные треугольники Ox_0C и OxA подобны, а значит, все точки, удовлетворяющие уравнению (7.1), лежат на прямой l .

В частных случаях, когда $x_0 = 0$ либо $y_0 = 0$ (одновременно они не могут быть равны нулю), уравнение (7.1) также задает только такие точки, которые лежат на прямой l .

Следовательно, уравнение (7.1) задает геометрическое место тех и только тех точек, которые лежат на прямой l , т. е. на прямой, проходящей через начало координат и заданную точку $(x_0, y_0) = R_\varphi(1, 0)$.

Поэтому уравнение (7.1) называется **уравнением прямой l** .

Отметим, что точка (x_0, y_0) полностью определяется углом поворота φ , т. к. является образом точки $(1, 0)$ при повороте на угол φ . В то же время, произвольная точка на единичной окружности однозначно задает угол поворота в интервале от 0 до 2π . Таким образом, задать поворот с центром O и задать точку на единичной окружности — суть одно и то же.

По определению $x_0 = \cos \varphi$ и $y_0 = \sin \varphi$, а отношение $y_0/x_0 = \operatorname{tg} \varphi$.

Кроме того, отношение y_0/x_0 также однозначно определяет угол поворота, но только в интервале от 0 до π .

Наконец, поворот прямой (!) на угол $\pi + \alpha$ — это поворот на угол α с последующим отражением прямой l относительно точки O . Но отражение прямой относительно своей же точки дает нам ту же самую прямую с тем же самым уравнением для ее точек! Таким образом, прямая, проходящая через начало координат, полностью определяется тангенсом угла наклона, т. е. отношением y_0/x_0 .

Но раз все дело в отношении, то прямая задается любой точкой, координаты которой находятся в таком же соотношении, что и координаты точки (x_0, y_0) , лежащие на единичной окружности. Иначе говоря, одну и ту же прямую задают и все точки вида (rx_0, ry_0) , $(-rx_0, -ry_0)$,

если коэффициент $r > 0$. На рис. 7.2 мы обозначили эти точки, соответственно, C, A и $-C, -A$.

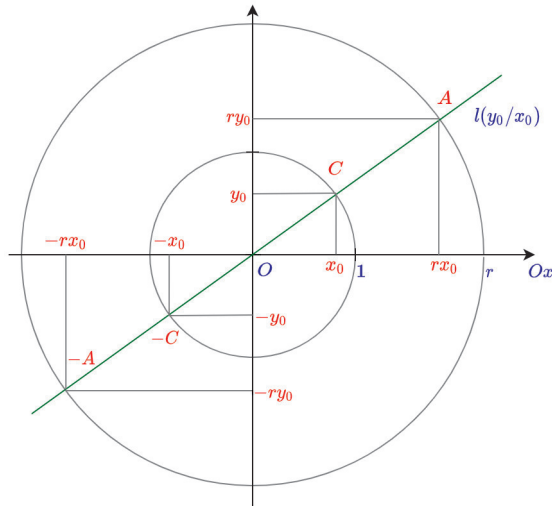


Рис. 7.2. Сохранение пропорций координат точек прямой.

Этот вывод можно получить и более формально, просто глядя на уравнение (7.1). Ведь если мы домножим обе его части на r , то ничего не изменится:

$$x(ry_0) - y(rx_0) = 0.$$

Итак, мы получили уравнение прямой, проходящей через начало координат. Рассмотрим вопрос чуть шире: что если прямая l не проходит через центр координат O ? В этом случае мы можем сдвинуть ее на некоторый вектор так, чтобы произвольно выбранная точка этой прямой перешла в точку O . Обозначим эту точку на прямой l за $S = (\Delta x, \Delta y)$, а сдвиг, соответственно, осуществим на вектор $(-\Delta x, -\Delta y)$ (см. рис. 7.3).

Тогда смещенные координаты $(x - \Delta x, y - \Delta y)$ уже будут пробегать прямую l' , проходящую через центр O , а ее уравнение нам известно:

$$(x - \Delta x)y_0 - (y - \Delta y)x_0 = 0,$$

или

$$xy_0 - yx_0 = c, \quad \text{где } c = y_0\Delta x - x_0\Delta y. \quad (7.2)$$

При этом коэффициенты (x_0, y_0) все так же отвечают за наклон прямой l и полностью определяются тангенсом угла наклона прямой l относительно положительного направления Ox , т. е. отношением y_0/x_0 .

Может показаться, что уравнение (7.2) сильно зависит от выбора точки S , поскольку свободный член c зависит от координат точки S .

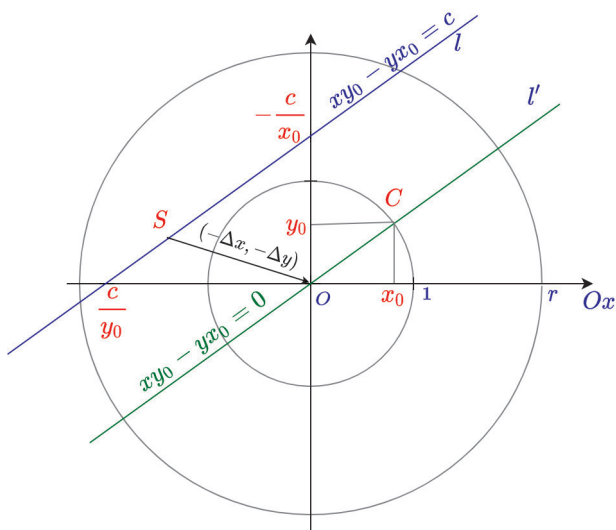


Рис. 7.3. Смещение прямой к началу координат.

Покажем, что это не так. Пусть $S' = (\Delta x', \Delta y')$ — какая-то другая точка прямой l . В этом случае она удовлетворяет найденному уравнению, т. е.

$$\Delta x' y_0 - \Delta y' x_0 = c,$$

но уравнение, найденное с помощью точки S' , будет иметь вид

$$x y_0 - y x_0 = y_0 \Delta x' - x_0 \Delta y',$$

откуда из предыдущего получаем, что вновь

$$x y_0 - y x_0 = c. \quad (7.3)$$

Таким образом, для нахождения c мы можем выбрать любую понравившуюся нам точку прямой l' , например, точку пересечения с одной из координатных осей.

Отметим также, что если взять произвольное уравнение вида (7.3), т. е. с произвольно выбранными коэффициентами x_0, y_0, c , потребовав только, чтобы x_0 и y_0 одновременно не обращались в ноль, то оно однозначно определит прямую на плоскости. Эта прямая будет проходить через точки $(c/y_0, 0)$ и $(-c/x_0)$ (см. рис. 7.3), если $x_0 \neq 0$ и $y_0 \neq 0$. При $x_0 = 0$ мы получим более простое уравнение $x y_0 = c$, которое определит вертикальную прямую, а при $y_0 = 0$ получаем уравнение $y x_0 = -c$, которое определит горизонтальную прямую.

Таким образом, уравнение (7.3) является уравнением произвольной прямой на координатной плоскости. Оно называется **линейным уравнением с двумя переменными**.

В случае, когда $x_0 \neq 0$, уравнение прямой (7.3) также можно переписать в виде

$$y = ax + b, \quad \text{где } a = \frac{y_0}{x_0}, \quad b = -\frac{c}{x_0}.$$

7.2. Линейные уравнения в целых числах

Поскольку мы пока владеем аппаратом только целых чисел (множество \mathbb{Z}), рассмотрим задачу о нахождении всех целых точек плоскости, через которые проходит заданная прямая. Под целыми точками плоскости мы будем понимать такие точки, обе координаты которых принадлежат \mathbb{Z} .

В общем виде **линейное уравнение в целых числах** выглядит следующим образом:

$$ax - by = c, \quad \text{где коэффициенты } a, b, c \in \mathbb{Z}. \quad (7.4)$$

Здесь мы сохранили форму уравнения (7.3) в продолжение предыдущей темы. В учебниках такое уравнение можно встретить в форме $ax + by + c = 0$. Понятно, что если переобозначить коэффициенты, сменив знаки у b и c , мы получим уравнение в форме (7.4). Ниже будет понятно, почему такая форма удобнее для исследования данного уравнения.

Наша задача: найти все такие x, y , тоже целые, которые удовлетворяют уравнению (7.4).

Сначала рассмотрим случай т.н. **однородного уравнения**:

$$ax - by = 0,$$

т.е. мы отбрасываем ту часть уравнения, которая не зависит от переменных x, y .

Как мы уже знаем, данное уравнение задает прямую, проходящую через начало координат, а ее наклон определяется отношением a/b .

Для начала проверим, нельзя ли данное отношение упростить. Если числа a, b имеют какой-то общий делитель, то разумно было бы на него сократить. И чтобы не проделывать это много раз, сократим их сразу на $\text{НОД}(a, b)$. Множество решений от этого не изменится, а само уравнение по-прежнему останется однородным и целочисленным:

$$\tilde{a}x - \tilde{b}y = 0, \quad \text{где } \tilde{a} = \frac{a}{\text{НОД}(a, b)}, \quad \tilde{b} = \frac{b}{\text{НОД}(a, b)}.$$

Таким образом, мы приходим к уравнению со взаимно простыми коэффициентами \tilde{a} и \tilde{b} .

Перепишем уравнение иначе: $\tilde{a}x = \tilde{b}y$. Заметим, что все числа здесь — целые. Причем $\tilde{b}y$ делится на \tilde{a} . Но так как \tilde{a} и \tilde{b} взаимно просты, то y делится на \tilde{a} . Это есть следствие того факта, который мы доказывали ранее в разделе 4.3: если простое число p делит произведение ab , то оно делит a или b (или их обоих). Поэтому если простое p делит \tilde{a} , то оно делит $\tilde{b}y$, но оно не может делить \tilde{b} , т. к. $\text{НОД}(p, \tilde{b}) = 1$, значит, оно делит y . Это значит, что все простые, составляющие число \tilde{a} , являются делителями y . В то же время эти простые не входят в \tilde{b} , поскольку $\text{НОД}(\tilde{a}, \tilde{b}) = 1$. Поэтому, если p^α входит в разложение \tilde{a} , то p^α также делит y . Следовательно, y делится на \tilde{a} , т. е.

$$y = k\tilde{a}$$

при некотором целом k .

Симметрично рассуждая, получаем, что x делится на \tilde{b} , т. е.

$$x = t\tilde{b}$$

при некотором целом t .

Подставим эти выражения в наше однородное уравнение:

$$\tilde{a}(t\tilde{b}) = \tilde{b}(k\tilde{a}),$$

откуда

$$t = k,$$

и больше никаких ограничений на выбор коэффициента k мы не имеем.

Таким образом, решениями уравнения $\tilde{a}x - \tilde{b}y = 0$ являются

$$\begin{cases} x = k\tilde{b} = kb/\text{НОД}(a, b), \\ y = k\tilde{a} = ka/\text{НОД}(a, b), \end{cases}$$

где $k \in \mathbb{Z}$. Эти же x и y являются решениями исходного однородного уравнения $ax - by = 0$.

Вернемся к неоднородному уравнению $ax - by = c$.

Для начала заметим, что если данное уравнение имеет решение в целых числах, то $ax - by$ делится на $\text{НОД}(a, b)$, а значит, c делится на $\text{НОД}(a, b)$. Поэтому, если c не делится на $\text{НОД}(a, b)$, то решений точно нет, т. е. в таком случае прямая $ax - by = c$ проходит мимо всех целых точек плоскости!

Покажем, что в случае делимости c на $\text{НОД}(a, b)$ решения обязательно есть, и опишем все такие решения.

Пусть $c = d\text{НОД}(a, b)$.

В разделе 4.3 мы установили, что $\text{НОД}(a, b)$ является линейной комбинацией чисел a и b , т. е. $\text{НОД}(a, b) = an + bm'$ при некоторых целых n и m' . Поскольку m' — целое, мы можем ввести новое обозначение $m = -m'$, откуда получим, что

$$\text{НОД}(a, b) = an - bm.$$

Отсюда следует, что пара чисел (dn, dm) удовлетворяет уравнению $ax - by = c$, поскольку $adn - bdm = d\text{НОД}(a, b) = c$.

Итак, представив $\text{НОД}(a, b)$ в виде линейной комбинации a и b , мы можем найти одно решение исходного уравнения.

Далее применим тот же прием, что и при изучении уравнений прямых — сдвинем прямую $ax - by = c$ так, чтобы точка (dn, dm) оказалась в начале координат. Для этого введем новые переменные

$$\hat{x} = x - dn, \quad \hat{y} = y - dm.$$

Тогда получаем, что $a\hat{x} - b\hat{y} = 0$. А такое уравнение мы уже решили выше, и его решением будет пара чисел $\hat{x} = kb/\text{НОД}(a, b)$ и $\hat{y} = ka/\text{НОД}(a, b)$, где k — любое целое число.

Собирая все вместе, находим общее решение исходного уравнения $ax - by = c$:

$$\begin{cases} x = kb/\text{НОД}(a, b) + dn, \\ y = ka/\text{НОД}(a, b) + dm, \end{cases}$$

где $d = c/\text{НОД}(a, b)$, n, m — коэффициенты в представлении $\text{НОД}(a, b) = an - bm$, k — это параметр решения, т. е. произвольное целое число.

Таким образом, решением линейного уравнения $ax - by = c$ в целых числах является сумма общего решения однородного уравнения $ax - by = 0$ и какого-нибудь частного решения исходного уравнения.

Основной трудностью при поиске частного решения является нахождение коэффициентов n и m представления $\text{НОД}(a, b)$.

Это представление можно найти с помощью алгоритма Евклида. Рассмотрим для примера уравнение

$$18x - 11y = 2.$$

Следуя алгоритму Евклида, получаем выкладки:

$$\begin{aligned} 18 &= 11 \cdot 1 + 7, \\ 11 &= 7 \cdot 1 + 4, \\ 7 &= 4 \cdot 1 + 3, \\ 4 &= 3 \cdot 1 + 1, \end{aligned}$$

где цветом выделены коэффициенты разложения. Последняя 1 — это и есть НОД(18, 11). Раскрутим алгоритм в обратную сторону:

$$\begin{aligned} 1 &= 4 - 3 = 4 - (7 - 4) = 4 \cdot 2 - 7 = (11 - 7) \cdot 2 - 7 = \\ &= 11 \cdot 2 - 7 \cdot 3 = 11 \cdot 2 - (18 - 11) \cdot 3 = \\ &= 11 \cdot 5 - 18 \cdot 3. \end{aligned}$$

Таким образом, наши искомые числа $n = -3$, $m = -5$. Напомним, что мы ищем представление НОД(18, 11) в виде $18n - 11m$, исходя из чего нужно правильно выбирать знаки перед коэффициентами. Говоря проще,

$$11 \cdot 5 - 18 \cdot 3 = 1.$$

Кроме того, $d = 2$, т.к. $c = 2$ и НОД(a, b) = 1. Откуда общее решение уравнения $18x - 11y = 2$ получаем в виде:

$$\begin{cases} x = 11k - 6, \\ y = 18k - 10, \end{cases}$$

где k — любое целое число. Проверим:

$$18(11k - 6) - 11(18k - 10) = 198k - 108 + 110 - 198k = 2.$$

Наконец, приведем еще один замечательный способ найти разложение НОД. Этот метод основан на представлении дробей в виде т.н. **цепных дробей**. Пусть дано уравнение

$$112x - 34y = 16.$$

Ищем приближение дроби $112/34$ следующим способом:

$$\frac{112}{34} = 3 + \frac{10}{34} = 3 + \frac{1}{3 + \frac{4}{10}} = 3 + \frac{1}{3 + \frac{1}{2 + \frac{2}{4}}} = 3 + \frac{1}{3 + \frac{1}{2 + \frac{1}{2}}}.$$

По сути дела, это — другая запись выкладок алгоритма Евклида, поскольку мы каждый раз последовательно выделяем неполное частное предыдущих остатков.

Как только мы дошли до хвоста вида $1/k$, мы останавливаемся, отбрасываем этот хвост и сворачиваем дробь обратно, получая приближение исходной дроби:

$$\frac{112}{34} \approx 3 + \frac{1}{3 + \frac{1}{2}} = \frac{23}{7}.$$

Далее, перемножая накрест эти дроби, получаем представление для НОД:

$$\text{НОД}(112, 34) = 112 \cdot 7 - 34 \cdot 23 = 2.$$

Искомые коэффициенты: $n = 7$, $m = 23$. Общее решение уравнения, таким образом, получаем в виде

$$\begin{cases} x = (34/2)k + (16/2) \cdot 7, \\ y = (112/2)k + (16/2) \cdot 23, \end{cases}$$

где k — любое целое число. Проверяем:

$$112(17k + 8 \cdot 7) - 34(56k + 8 \cdot 23) = 8(112 \cdot 7 - 34 \cdot 23) = 16.$$

Выше мы всюду рассматривали уравнения, в которых x идет с положительным коэффициентом, а y — с отрицательным. Иначе говоря, прямая, заданная таким уравнением, имеет наклон «вправо» от вертикальной оси. Но уравнение может быть, например, таким:

$$5x + 9y = 1.$$

Если мы хотим решать его по тем же формулам, то лучше перейти к новым переменным $\hat{x} = x$, $\hat{y} = -y$, и тогда мы получим уравнение

$$5\hat{x} - 9\hat{y} = 1.$$

Найдя его решения, мы просто меняем знак у \hat{y} и получаем решение исходного уравнения.

Упражнения

Обязательные упражнения

7.1° Решите в целых числах уравнения:

- | | |
|--------------------|-----------------------|
| a) $6x - 5y = 0$; | d) $4x + 7y = 41$; |
| b) $6x - 6y = 2$; | e) $7x - 5y = 21$; |
| c) $6x - 5y = 3$; | f) $19x + 17y = 15$. |

7.2° Найти все решения линейного уравнения в целых числах или доказать, что их нет: a) $5x - 9y = 2$; b) $225x + 81y = 18$; c) $10x - 18y = 3$.

7.3° Решите уравнения: a) $121x + 91y = 1$; b) $-343x + 119y = 42$; c) $111x - 740y = 11$.

7.4° Разложить в цепную дробь числа: a) $15/4$; b) $42/31$; c) $13/9$; d) $6/5$.

7.5° Используя разложение в цепную дробь решить уравнение в целых числах: a) $57x - 89y = 16$; b) $13x - 10y = 27$.

7.6° Покажите, как при помощи алгоритма Евклида можно по произвольным a и b найти такие k и l , что $ak + bl = \text{НОД}(a, b)$.

7.7° Найти линейное представление НОД с помощью алгоритма Евклида и методом цепных дробей:

$$\text{НОД}(5, 9), \quad \text{НОД}(18, 15), \quad \text{НОД}(225, 81).$$

7.8° Докажите, что уравнение $ax + by = d$ имеет решение в целых числах тогда и только тогда, когда $\text{НОД}(a, b) | d$. В частности, $\text{НОД}(a, b)$ — это наименьшее натуральное число, представимое в виде $ax + by$.

7.9° Кузнечик может прыгать на расстояние 15 и 7. Изначально он находится в точке 0. а) Найдите, как следует прыгать кузнечику, чтобы оказаться в точке 3. б) Найдите, за какое наименьшее число прыжков он может попасть в точку 6.

7.10° Пусть (x_0, y_0) — решение уравнения $ax + by = d$. Пусть a_0 и b_0 — такие числа, что $\text{НОД}(a, b)a_0 = a$, $\text{НОД}(a, b)b_0 = b$. Покажите, что каждое решение уравнения $ax + by = d$ имеет вид $x = x_0 + b_0 \cdot t$, $y = y_0 - a_0 \cdot t$, где t — целое число.

7.11° Известно, что пары чисел (x_1, y_1) и (x_2, y_2) являются решением уравнения $ax + by + c = 0$, где a, b, c — некоторые неизвестные целые коэффициенты. Найдите, выразив через (x_1, y_1) и (x_2, y_2) , чему равно a/b .

7.12° Решите в целых числах уравнение $2x + 3y + 5z = 1$.

7.13° Доказать, что уравнение $ax + by = ab$, где $a, b > 0$ и $\text{НОД}(a, b) = 1$, неразрешимо в натуральных числах.

7.14° Пусть m, n — целые и $(5m + 3n) \not\equiv 11$. Докажите, что: а) $(6m + 8n) \equiv 11$; б) $(9m + n) \equiv 11$.

7.15° Пусть в некоторой стране имеют хождение монеты достоинством только 14 и 23 тугрика. Продавец должен выдать сдачу покупателю в размере 1 тугрик. Считая, что у обоих имеется достаточное количество монет того и другого достоинства, указать способ, которым должен воспользоваться продавец для выдачи сдачи.

7.16° Найти цепную дробь для $\sqrt{3}$.

7.17° С помощью цепной дроби найти дробь

$$\frac{k}{r} \in \left[\frac{165}{256} - \frac{1}{512}; \frac{165}{256} + \frac{1}{512} \right]$$

при условии, что $r < 16$.

Рациональность и соизмеримость

Аннотация

В этой главе мы начинаем выход за пределы целых чисел, и прежде всего займемся построением чисел рациональных. Кроме того, мы увидим, что одними рациональными числами нельзя ограничиваться, т. к. существуют несоизмеримые с ним числа вроде корня из 2.

8.1. Построение рациональных чисел

До сих пор мы часто оперировали дробями, хотя нигде их не определяли. Разве что упоминали отношение y_0/x_0 как некоторый параметр, определяющий угол наклона прямой на координатной плоскости в главе 7. Причем для определения положения прямой важно именно отношение коэффициентов, а не их собственные значения.

Итак, рассмотрим прямую l , заданную уравнением $ax - by = 0$, где a, b — целые числа. Прямая l проходит через начало координат, т. е. точку $(0, 0)$.

Для начала пусть $a = 1$ и $b > 1$. Легко видеть, что такая прямая проходит через точки $(0, 0)$ и $(b, 1)$ (см. рис. 8.1).

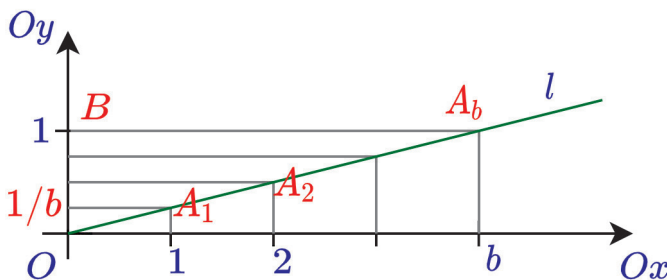


Рис. 8.1. Прямая $x - by = 0$.

На прямой l мы можем отметить точки A_1, A_2, \dots, A_b в местах пересечения этой прямой с вертикальными прямыми, имеющими уравнения $x = 1, x = 2, \dots, x = b$, соответственно.

Теперь рассмотрим треугольник OBA_b , где точка $B = (0, 1)$. В этом треугольнике через точки A_1, A_2, \dots мы можем провести линии, параллельные его горизонтальной стороне BA_b , которые отсекут на вертикальной стороне OB нашего треугольника отрезки.

Эти отрезки будут иметь одинаковую длину (двукратное применение теоремы Фалеса), т. к. точки на прямой l также расставлены с одинаковым шагом, что следует уже из выбора вертикальных секущих (они идут с шагом 1).

Итак, на вертикальной оси мы получили b одинаковых отрезков, сумма длин которых равна 1.

Здесь можно снова обратиться к сюжету с синхронно шагающими товарищами, только теперь один из них шагает по прямой Ox с шагом 1, а второй делает синхронные шаги по оси Oy так, чтобы прямые, проведенные перпендикулярно их линиям движения всегда скрещивались на прямой l . Ясно, что длины шагов этих товарищей будут отличаться, и это отличие зависит от наклона прямой l .

Какова же длина шагов на оси Oy ? Ответ: она равна одной b -й части единицы. И эта часть записывается как дробь $1/b$. Собственно, отношение $1/b$, как мы видели ранее, является определяющим для прямой l . Оно характеризует величину **наклона** этой прямой.

Мы можем взять сумму нескольких таких частей (пройти несколько шагов). Например, k ($k > 0$) частей размера $1/b$ дают в сумме отрезок длины в k раз больше, чем отрезок $1/b$. Такая часть записывается в виде дроби k/b .

Величину k/b можно получить и другим способом. Возьмем теперь прямую l' , заданную уравнением $kx - by = 0$ (см. рис. 8.2).

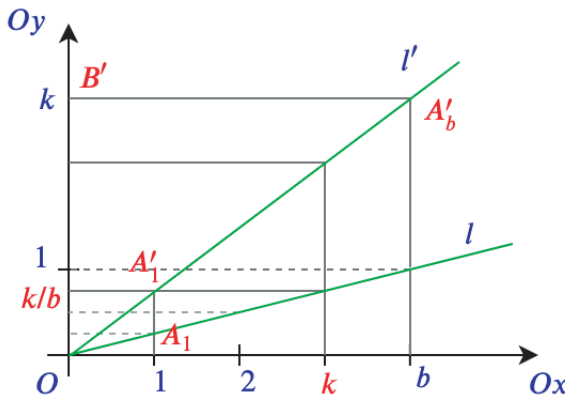


Рис. 8.2. Прямые $x - by = 0$ (l) и $kx - by = 0$ (l').

Эта прямая проходит через начало координат и точку (b, k) . Наклон этой прямой характеризуется отношением k/b .

Прделаем аналогичные предыдущему построения: проведем вертикальные линии с шагом 1, а затем горизонтальные линии от точек пересечения вертикальных с прямой l' , и посмотрим, какие отрезки у нас получатся на оси Oy .

Нетрудно видеть, что линия, соответствующая $x = k$, для прямой l отсекает на оси Oy метку, которую мы обозначили как k/b . Но ровно ту же самую метку покажет построение с помощью вертикальной линии $x = 1$ и прямой l' . Для этого достаточно сравнить уравнения этих прямых:

$$l : x - by = 0, \quad l' : kx - by = 0.$$

Если в первом вместо x подставить k , а во втором вместо x подставить 1, то получим одно и то же значение y . Отсюда и совпадение меток.

Пользуясь сюжетом с двумя товарищами, мы можем сказать, что наклоны прямых l и l' задают разный масштаб шагов второго товарища относительно первого. А именно, наклон прямой l' задает на оси Oy шаг (или масштаб) в k раз больше, чем наклон прямой l .

Получается, что наклон прямой, заданной уравнением $kx - by = 0$, задает умножение на число k того масштаба, который определяется наклоном прямой, заданной уравнением $x - by = 0$.

Чтобы упростить терминологию, просто скажем, что наклон k/b в k раз круче наклона $1/b$.

Рассмотрим теперь прямую, заданную уравнением $ax - by = 0$ с произвольными целыми коэффициентами a и b ($b \neq 0$). Ее наклон записывается в виде пропорции a/b . Мы можем провести построения, аналогичные предыдущему, и выяснить, что наклон прямой, заданной уравнением $(ka)x - by = 0$, также окажется в k раз круче наклона a/b , т. е. равным шагам на оси Ox при тех же геометрических построениях на оси Oy соответствуют шаги, которые в случае наклона $(ka)/b$ будут ровно в k раз больше, чем шаги, соответствующие наклону a/b .

Таким образом, рассматривая наклоны прямых, заданные пропорцией их коэффициентов, как некие *новые объекты*, мы можем ввести понятие умножения наклона на целое число. Если у нас есть наклон a/b прямой, заданной уравнением $ax - by = 0$, то результатом его умножения на число k является наклон $(ka)/b$ прямой, заданной уравнением $kax - by = 0$. Обозначая умножение числа на наклон точкой или пустым символом, запишем данное утверждение следующим равенством:

$$k \cdot (a/b) = (ka)/b,$$

причем теперь мы не будем ограничиваться только положительным k , считая его произвольным целым числом (в частности, нулем, при кото-

ром наклон станет нулевым, а прямая горизонтальной, и все шаги на оси Oy схлопнутся в одну точку — второй товарищ будет топтаться на месте, пока первый шагает вперед или назад).

Далее, число k — целое, стало быть, оно является суммой единиц (или минус единиц) в количестве $|k|$. А это значит, что умножение на k можно представить как многократное сложение:

$$k \cdot (a/b) = \underbrace{a/b + a/b + \dots + a/b}_{k \text{ раз}}$$

для положительного k и

$$k \cdot (a/b) = \underbrace{(-a)/b + (-a)/b + \dots + (-a)/b}_{-k \text{ раз}}$$

для отрицательного k . Умножение на -1 мы просто ввели по определению.

Далее, мы можем в этих суммах расставлять скобки как угодно и сворачивать внутри них сложение обратно в умножение, получая тем самым аддитивное свойство умножения наклона на целое число. Если $k = k_1 + k_2$, то

$$k \cdot (a/b) = k_1 \cdot (a/b) + k_2 \cdot (a/b).$$

Заметим, что это на самом деле закон дистрибутивности, являющийся одной из аксиом кольца и связывающий сложение с умножением.

Идем дальше. По уже установленным правилам оперирования с наклонами, нетрудно видеть, что

$$k_1 \cdot (a/b) = (k_1 a)/b, \quad k_2 \cdot (a/b) = (k_2 a)/b, \quad k \cdot (a/b) = (ka)/b,$$

откуда получаем, что

$$(k_1 a)/b + (k_2 a)/b = (ka)/b,$$

а поскольку произведения $k_1 a, k_2 a, ka$ — произвольные целые числа, мы получаем правило сложения наклонов

$$a_1/b + a_2/b = (a_1 + a_2)/b.$$

Важно: при сложении наклонов коэффициент b , т. е. знаменатель отношения и одновременно коэффициент перед переменной y в уравнении соответствующей прямой, должен быть одинаковым у обоих слагаемых! Только в этом случае мы получаем согласование операций сложения и умножения.

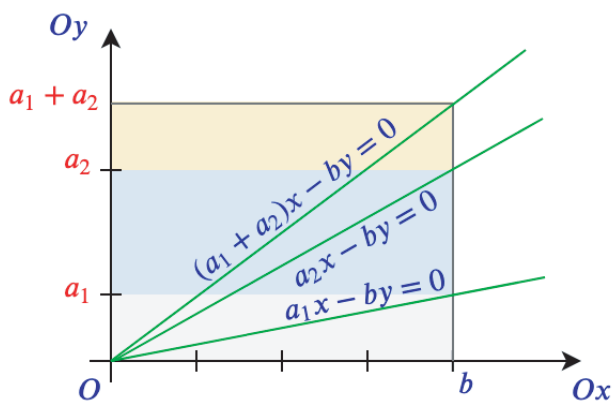


Рис. 8.3. Сложение наклонов.

Сложение наклонов прямых можно интерпретировать графически как сложение площадей прямоугольников с основанием b и высотой a_1 и a_2 . В результате получается прямоугольник с тем же основанием b и высотой $a_1 + a_2$. При этом прямые всегда проходят через точку $(0, 0)$ и через правый верхний угол прямоугольников (см. рис. 8.3). С помощью этой же картинке можно представить себе и умножение наклона прямой на целое число k . Для этого нужно растащить соответствующий этому наклону прямоугольник вверх k раз.

На самом же деле операции сложения, вычитания и умножения на целое число, производимые с коэффициентом перед x уравнения прямой, в точности повторяют таковые операции над целыми числами (поскольку это и есть целые числа!) и, соответственно, подчиняются всем аксиомам кольца целых чисел (дистрибутивный закон мы уже видели, а коммутативность сложения и умножения, существование нуля, единицы и обратных — все это наследие кольца \mathbb{Z}). [А вот и более научный термин для тех, кто собирается идти в математику глубоко: наклоны прямых с общим основанием b образуют *модуль над кольцом \mathbb{Z}* .]

Поэтому все наклоны вида a/b , $a, b \in \mathbb{Z}$, при фиксированном $b \neq 0$ с определенными выше операциями сложения и умножения образуют *кольцо* (изоморфное кольцу целых чисел).

Заметим теперь, что уравнение $x - by = 0$ прямой l можно переписать иначе: $kx - (bk)y = 0$. Чем оно отличается от уравнения $kx - by = 0$ прямой l' ? Очевидно, тем, что перед y появился коэффициент k . А теперь вспомним, что прямая l' задает наклон в k раз больше, чем прямая l . И это значит, что если мы хотим разделить наклон прямой l' на k , то мы должны умножить на k ее коэффициент перед y .

Итак, если мы хотим умножить наклон прямой с уравнением $ax - by = 0$ на целое число, то мы умножаем на это число коэффициент

перед x (прямая становится более крутой), а если мы хотим разделить наклон прямой на целое число, то мы умножаем на это число коэффициент перед y (прямая становится более полой):

$$k(a/b) = (ka/b), \quad (a/b)/k = a/(bk).$$

Делаем следующий шаг: умножение двух наклонов. На самом деле, наклон любой прямой вида $ax - by = 0$ мы можем выразить через композицию ранее определенных операций — умножения на целое число и деления на целое число:

$$a/b = a \cdot (1/1)/b,$$

откуда видим, что прямая $x - y = 0$ (соответствующая пропорции $1/1$) имеет наклон 45° и в операциях умножения может опускаться точно так же, как обычная единица. Таким образом, умножение наклонов прямых выглядит следующим образом

$$\begin{aligned} (a_1/b_1) \cdot (a_2/b_2) &= \\ &= a_1(1/b_1) \cdot (a_2/1)/b_2 = a_1(a_2/b_1)b_2 = (a_1a_2)/(b_1b_2). \end{aligned}$$

Отсюда нетрудно получить и процедуру деления наклонов прямых друг на друга, решив уравнение:

$$(a_1/b_1) = (a_2/b_2)(c/d),$$

т. е.

$$(a_1/b_1) = (a_2c)/(b_2d),$$

и эти наклоны задают одну и ту же прямую двумя разными способами. Умножим наклон в левой части равенства на единичный наклон $(1/1)$, представленный пропорцией $(a_2b_2)/(a_2b_2)$, получим

$$(a_1a_2b_2)/(a_2b_1b_2) = (a_2c)/(b_2d),$$

откуда видно, что можно выбрать следующее решение

$$c = a_1b_2, \quad d = a_2b_1,$$

а также любое ему пропорциональное в любое целое число раз (кроме нуля). Таким образом,

$$(a_1/b_1)/(a_2/b_2) = (a_1b_2)/(a_2b_1).$$

Наконец, чтобы научиться складывать произвольные наклоны прямых, мы должны уметь сводить сложение произвольных прямых к сложению прямых с одинаковым коэффициентом перед y , т. к. сложение мы определили выше только для данного случая.

Но и это не проблема:

$$\begin{aligned} (a_1/b_1) + (a_2/b_2) &= (a_1b_2)(b_1b_2) + (a_2b_1)/(b_1b_2) = \\ &= (a_1b_2 + a_2b_1)/(b_1b_2). \end{aligned}$$

Следующая картинка 8.4 показывает «арифметику наклонов» с произвольными параметрами. Здесь маленькие прямоугольники соответ-

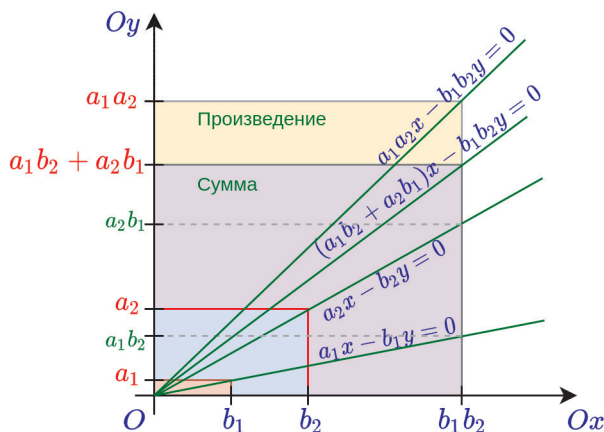


Рис. 8.4. Арифметика наклонов.

ствуют исходным прямым с уравнениями $a_1x - b_1y = 0$ и $a_2x - b_2y = 0$, пунктиром отмечены приведенные к общему основанию b_1b_2 прямоугольники, большой темный прямоугольник соответствует их сумме (буквально один приставлен сверху к другому), большой светлый прямоугольник – произведению (помножены основания и помножены высоты). На рисунке не нанесен масштаб (т. е. не указаны единицы на обеих осях). Дело в том, что в зависимости от выбора масштаба положение чисел на оси Oy может иметь различный порядок. Выбранное расположение нужно считать условным.

В целом картина представления рациональных чисел с помощью прямых с целочисленными коэффициентами выглядит следующим образом (рис. 8.5):

Итак, имея только множество целых чисел \mathbb{Z} , мы построили на плоскости всевозможные прямые, заданные линейными уравнениями с целыми коэффициентами, научились складывать, вычитать, умножать и делить их наклоны. Тем самым мы построили новую алгебраическую структуру, которая называется **полем рациональных чисел** и обозначается \mathbb{Q} .

На самом деле, в нашем построении есть еще и такая прямая, которая соответствует бесконечности. Это прямая, заданная уравнением

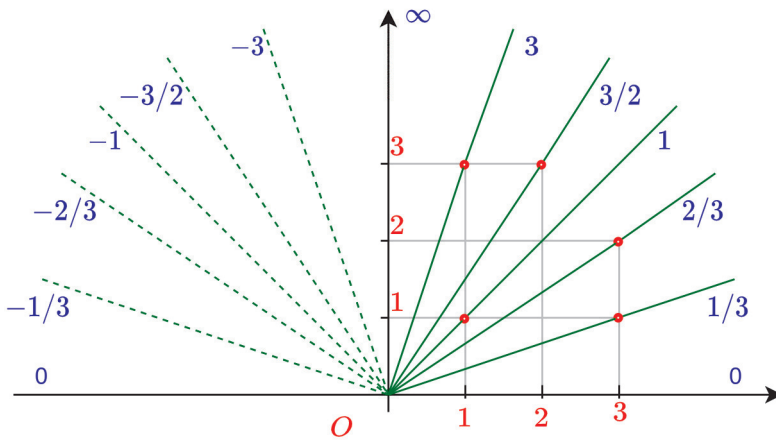


Рис. 8.5. Рациональные числа.

$x = 0$. Ее наклон можно записать как $1/0$ (хотя, строго говоря, вместо единицы можно подставить любое отличное от нуля целое число) или просто ∞ . Нулевой наклон определяется уравнением $y = 0$ и записывается пропорцией $0/1$ или просто 0 . В полном соответствии с установленными правилами, мы можем заметить, что если $a \neq 0 \neq b$, то

$$(a/b)(0/1) = 0/1 = 0, \quad (a/b)(1/0) = 1/0 = \infty,$$

$$(a/b)/(0/1) = 1/0 = \infty, \quad (a/b)/(1/0) = 0/1 = 0,$$

т. е. деление на ноль дает бесконечность, а деление на бесконечность дает ноль для любых ненулевых конечных наклонов прямых.

Но тут кроется проблема: $(1/0) \cdot (0/1) = 0/0$, что соответствует уравнению $0x - 0y = 0$. Такое уравнение не задает прямую, его решением является вся плоскость! Проще говоря, при умножении $0 \cdot \infty$ может получиться любое число!

Поэтому при определении поля бесконечный элемент не постулируется и, соответственно, деление на ноль не разрешено.

Поле рациональных чисел является представителем огромного количества различных полей, известных в математике. Как и кольцо, общее алгебраическое понятие поля задается списком аксиом, накладывающих некоторые дополнительные ограничения на кольцо. А именно — в поле операция умножения должна быть коммутативной и, кроме того, в поле разрешается делить на любой ненулевой элемент.

Приведем полный формальный список аксиом поля. Множество F с операциями $+$ и \cdot называется **полем**, если:

F1 $a, b \in F \Rightarrow a + b \in F, a \cdot b \in F$ (замкнутость операций);

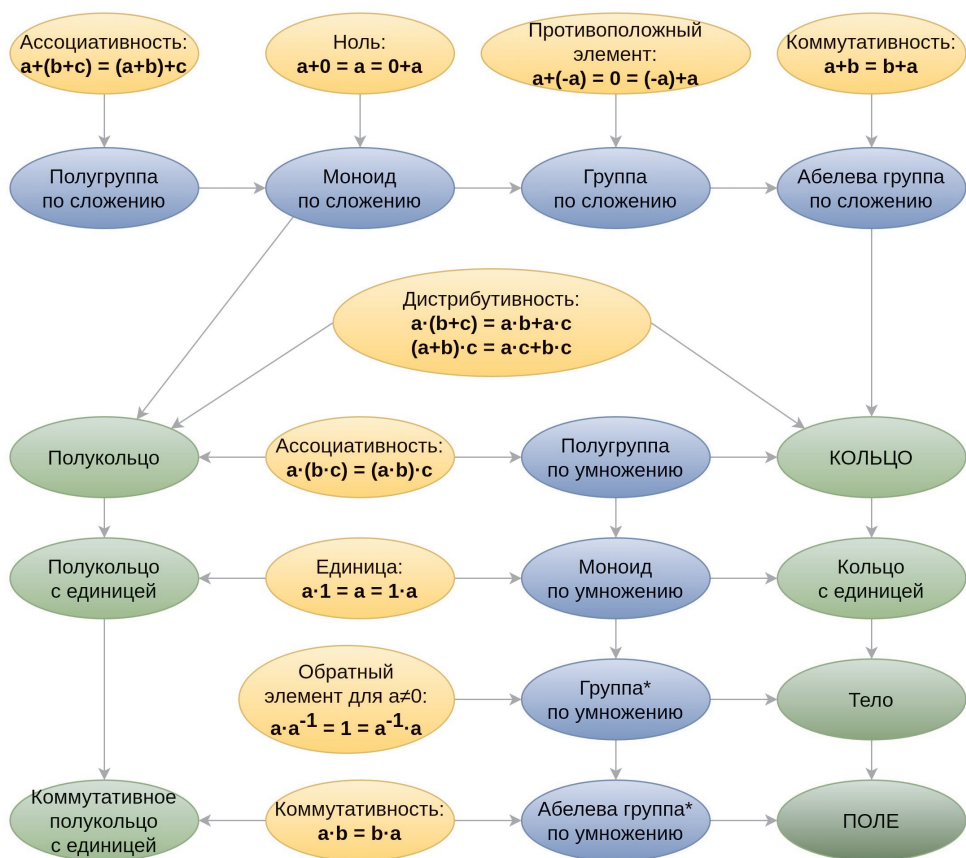


Рис. 8.6. Звездочка означает, что исключены нулевые элементы.

F2 $a, b, c \in F \Rightarrow (a + b) + c = a + (b + c), (a \cdot b) \cdot c = a \cdot (b \cdot c)$ (ассоциативность операций);

F3 для всех $a, b \in F$ имеем $a + b = b + a$ и $a \cdot b = b \cdot a$ (коммутативность операций);

F4 существует элемент $0 \in F$ такой, что $a + 0 = 0$ для всех $a \in F$ (аксиома нуля);

F5 для всякого элемента $a \in F$ существует противоположный $-a$ такой, что $a + (-a) = 0$ (аксиома противоположного элемента);

F6 существует элемент $1 \in F$ такой, что $a \cdot 1 = 1$ для всех $a \in F$ (аксиома единицы);

F7 для всякого элемента $a \in F$, если $a \neq 0$, то существует обратный a^{-1} такой, что $a \cdot a^{-1} = 1$ (аксиома обратного элемента);

F8 для всех $a, b, c \in F$ имеем $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$ (дистрибутивность).

Иначе говоря, поле — это коммутативное кольцо с единицей, в котором каждый ненулевой элемент обратим. На схеме 8.6 представлено формирование таких понятий, как поле и кольцо, из более простых свойств (или аксиом).

Заметка на будущее. Поле \mathbb{Q} в общей алгебре определяется как минимальное поле, содержащее натуральные числа. Кроме того, поле рациональных чисел может быть задано как поле частных кольца \mathbb{Z} .

8.2. Соизмеримость. Иррациональности

Рациональные числа мы определили как наклоны прямых, заданных уравнениями вида $ax - by = 0$, где a и b — произвольные целые числа, причем $b \neq 0$. Оказалось, что наклон прямой с таким уравнением однозначно определяется пропорцией коэффициентов данного уравнения, т. е. рациональным числом a/b . Мы используем термин «пропорция» намеренно, чтобы подчеркнуть, что и наклон, и соответствующее ему рациональное число не меняются, если числитель и знаменатель дроби a/b умножить на одно и то же целое число или разделить на общий делитель a и b .

Говоря алгебраическим языком, рациональные числа — это корни линейных уравнений, т. е. уравнений вида $a - bx = 0$ с целыми коэффициентами a и b .

Таким образом, выход в поле рациональных чисел происходит при попытке разрешить линейное уравнение, заданное над кольцом целых чисел.

Возникает естественное желание обобщить этот процесс. Что если мы рассмотрим линейное уравнение, но над полем рациональных чисел? Будет ли оно разрешимо?

Рассмотрим уравнение $rx - q = 0$, где $r, q \in \mathbb{Q}$. Тогда представим эти рациональные числа в виде дробей $r = a/b$, $q = c/d$, откуда

$$0 = rx - q = \frac{a}{b}x - \frac{c}{d} = \frac{adx - cb}{bd},$$

откуда ясно, что данное уравнение эквивалентно линейному уравнению $(ad)x - (cb) = 0$ с целыми коэффициентами, а значит, разрешимо в поле рациональных чисел.

Таким образом, поле \mathbb{Q} замкнуто относительно взятия решений линейных уравнений над ним. Термин «замкнутость», или **алгебраическая замкнутость** некоторой системы чисел, означает, что в этой системе можно найти корень алгебраического уравнения. При этом, в общем случае, **алгебраическое уравнение** — это приравненное к нулю

выражение, построенное из чисел данной числовой системы, некоторого набора переменных и разрешенных в этой системе чисел алгебраических операций (сложения и умножения). Другими словами, это уравнение, записанное в алфавите, состоящем из значков $=, +, \cdot$, переменных и конкретных констант. В случае одной переменной такое уравнение имеет вид

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0,$$

где под степенью x^k мы, как обычно, понимаем k -кратное умножение x на само себя. Позже мы еще вернемся к изучению таких уравнений произвольной степени n . А пока заметим, что *линейное уравнение — это алгебраическое уравнение первой степени*. Поле \mathbb{Q} разрешает любое такое уравнение.

Посмотрим, как оно справится с уравнениями более высокой степени! Пусть дано уравнение $x^2 - 2 = 0$. Это — уравнение второй степени с целыми коэффициентами (1 и 2), т. е. алгебраическое уравнение, заданное над \mathbb{Z} (и над \mathbb{Q}). Разрешимо ли оно в \mathbb{Z} или хотя бы в \mathbb{Q} ?

Ответ: нет! Предположим, что $x = n/m$ разрешает такое уравнение, т. е. $(n/m)^2 = 2$. Предположим сразу же, что $n \perp m$, т. е. дробь n/m несократимая. Далее имеем

$$n^2 = 2m^2.$$

Отсюда видно, что n^2 делится на 2, а значит, 2 входит в разложение числа n^2 по степеням простых. Проблема в том, что если бы 2 не вошла в разложение числа n , то ее не было бы и в разложении числа n^2 , т. к. n^2 есть произведение степеней тех же самых простых, что и n , только в удвоенной степени. А значит, n делится на 2, откуда следует, что n^2 делится на 4. Но тогда m^2 делится на 2 и, аналогично рассуждая, получаем, что и m делится на 2. А это уже противоречит тому, что дробь n/m несократимая — ее как минимум можно сократить на 2.

Следовательно, корень уравнения $x^2 - 2 = 0$ не может быть рациональным числом.

Тем не менее положительный корень такого уравнения можно оценивать сверху и снизу сколь угодно точно при помощи рациональных чисел. Например, корень извлекается из числа 2.25 и равен 1.5, при этом $x^2 = 2 < 2.25$, так что $x < 1.5$. В то же время $2 > 1.96 = 1.4^2$, так что $x > 1.4$. Можно еще усилить оценку: $1.41 < x < 1.42$. И так далее. Это позволяет нам думать, что на самом деле число такое есть, просто оно сидит где-то между рациональными числами. Обоснование его существования мы отложим на потом, а пока просто обозначим его символом $\sqrt{2}$.

Есть еще один способ удостовериться в том, что $\sqrt{2}$ не является рациональным числом. И тут снова нам на выручку приходят цепные дроби.

Воспроизведем алгоритм Евклида для дроби $\alpha = r_0/r_1$, считая, что $r_0 > r_1$ (если это не так, то приведем дробь к виду $1/(r_1/r_0)$ и будем работать дальше только со знаменателем). Как и раньше, будем выделять остаток r_{s+1} от деления r_{s-1} на r_s и сохранять неполное частное k_s . Только запишем весь алгоритм не в несколько строк, а в виде многоэтажной дроби.

$$\begin{aligned} \frac{r_0}{r_1} &= \frac{k_1 r_1 + r_2}{r_1} = \boxed{k_1} + \frac{1}{\frac{r_1}{r_2}} = \boxed{k_1} + \frac{1}{\frac{k_2 r_2 + r_3}{r_2}} = \\ &= \boxed{k_1} + \frac{1}{\boxed{k_2} + \frac{1}{r_3/r_2}} = \boxed{k_1} + \frac{1}{\boxed{k_2} + \frac{1}{\boxed{k_3} + \dots + \frac{1}{\boxed{k_n} + r_{n+1}/r_n}}} \end{aligned}$$

где $r_0 > r_1 > r_2 > \dots > r_n > r_{n-1}$.

Поскольку остатки всегда являются натуральными числами, рано или поздно этот алгоритм прервется. Пусть это случится на шаге с номером n , так что мы полагаем $r_{n+1} = 0$, и цепная дробь закончится на числе k_n .

В таком случае цепную дробь принято записывать последовательно выделенных на каждом шаге целых частей:

$$\frac{r_0}{r_1} = [k_1, k_2, \dots, k_n].$$

Отсюда следует, что всякая рациональная дробь представима в виде конечной цепной дроби. Обратное, очевидно, также верно, ибо каждую конечную цепную дробь можно свернуть по правилам арифметики в обычную рациональную дробь.

Заметим также, что любое целое число представляется в виде тривиальной цепной дроби, в которой есть только k_1 .

Алгоритм Евклида можно применять к любым числам, лишь бы можно было выделять остаток от деления. Например, его можно применить к паре чисел $\pi/2$ и $\pi/3$ и получить конечную цепную дробь. А все потому, что отношение этих чисел является рациональным числом $3/2$. Поэтому, если отношение двух чисел a/b рационально, их принято называть **соизмеримыми**.

Соизмеримые числа хорошо иллюстрируются рис. 8.7. Видим, что прямоугольник $a \times b$ мы делим на квадраты, каждый раз выбирая максимальный квадрат, который вписывается в оставшуюся область. Если a и b соизмеримы, то процесс разрезания прямоугольника на квадраты

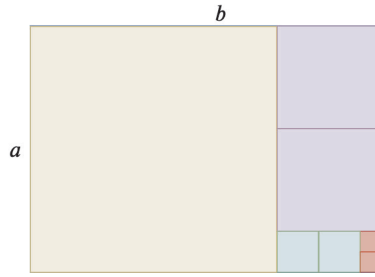


Рис. 8.7. Соизмеримые числа.

закончится за конечное число шагов, причем количество одинаковых квадратов, посчитанное в порядке их убывания, есть как раз те самые числа k_1, k_2, \dots, k_n , появляющиеся в записи цепной дроби (поскольку вырезание максимального квадрата — это не что иное, как процесс выделения единица за единицей целой части из остатка, т.е. алгоритм Евклида).

То, что сами числа a и b при этом могут не быть целыми или рациональными — не важно. Важно, что их отношение рационально. Также легко видеть, что всякое рациональное число соизмеримо с 1 и, наоборот, всякое число, соизмеримое с 1, рационально.

Посмотрим теперь, что происходит при попытке записать цепную дробь для $\sqrt{2}$.

Мы уже знаем, что $1 < \sqrt{2} < 2$, кроме того, $(\sqrt{2} + 1) = 1/(\sqrt{2} - 1)$ (здесь работает формула $(x - y)(x + y) = x^2 - y^2$, которую мы получили в разделе 1.3 с помощью рис. 1.12) так что

$$\begin{aligned} \sqrt{2} &= \boxed{1} + (\sqrt{2} - 1) = \boxed{1} + \frac{1}{1/(\sqrt{2} - 1)} = \boxed{1} + \frac{1}{\sqrt{2} + 1} = \\ &= \boxed{1} + \frac{1}{\boxed{2} + (\sqrt{2} - 1)} = \boxed{1} + \frac{1}{\boxed{2} + \frac{1}{\sqrt{2} + 1}} = \boxed{1} + \frac{1}{\boxed{2} + \frac{1}{\boxed{2} + \frac{1}{\dots}}} \end{aligned}$$

Как мы видим, остатком после выделения целой части всегда является одно и то же число $\sqrt{2} - 1$, и процесс алгоритма Евклида никогда не остановится. При этом цепная дробь характеризуется последовательностью одинаковых целых частей, равных 2. То есть представление для корня из 2 в виде цепной дроби будет бесконечным:

$$\sqrt{2} = [1, 2, 2, 2, 2, \dots],$$

и, следовательно, $\sqrt{2}$ не является рациональным числом.

Геометрический алгоритм Евклида здесь тоже заикливается. Действительно, возьмем прямоугольник со сторонами $1 + \sqrt{2}$ и 1. Следуя

алгоритму, вырежем из него два квадрата 1×1 . Посмотрим, какой прямоугольник остался: его сторонами будут 1 и $\sqrt{2} - 1$. А в каком соотношении друг к другу они находятся?

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

Таким образом, перед нами уменьшенная копия исходного прямоугольника. И если мы продолжим вырезать квадраты, мы будем вновь и вновь получать один и тот же прямоугольник, только все меньшего размера.

Числа, не являющиеся рациональными, называются **иррациональными**.

Наличие иррационального числа $\sqrt{2}$ позволяет нам рассмотреть числа вида $r + q\sqrt{2}$, где $r, q \in \mathbb{Q}$.

Множество таких чисел, полученных «присоединением» к полю \mathbb{Q} положительного корня уравнения $x^2 = 2$ и любых выражений вида $r + q\sqrt{2}$ с рациональными r и q , принято обозначать $\mathbb{Q}[\sqrt{2}]$ и называть **расширением поля** \mathbb{Q} , порожденным числом $\sqrt{2}$.

На числах множества $\mathbb{Q}[\sqrt{2}]$ операции сложения и умножения определяются естественным образом:

$$(r_1 + q_1\sqrt{2}) + (r_2 + q_2\sqrt{2}) = (r_1 + r_2) + (q_1 + q_2)\sqrt{2},$$

$$(r_1 + q_1\sqrt{2})(r_2 + q_2\sqrt{2}) = (r_1r_2 + 2q_1q_2) + (r_1q_2 + r_2q_1)\sqrt{2}.$$

Проверим, что для $\mathbb{Q}[\sqrt{2}]$ выполняются все аксиомы поля F1–F8, приведенные на стр. 130.

F1. Замкнутость операций выполнена в силу их определения.

F2. Ассоциативность операций

$$\begin{aligned} (r_1 + q_1\sqrt{2} + r_2 + q_2\sqrt{2}) + (r_3 + q_3\sqrt{2}) &= (r_1 + r_2 + r_3) + (q_1 + q_2 + q_3)\sqrt{2} = \\ &= (r_1 + q_1\sqrt{2}) + (r_2 + q_2\sqrt{2} + r_3 + q_3\sqrt{2}), \end{aligned}$$

$$\begin{aligned} [(r_1 + q_1\sqrt{2})(r_2 + q_2\sqrt{2})](r_3 + q_3\sqrt{2}) &= \\ = (r_1r_2r_3 + 2q_1q_2r_3 + 2r_1q_2q_3 + 2q_1r_2q_3) + (r_1r_2q_3 + 2q_1q_2q_3 + r_1q_2r_3 + q_1r_2r_3)\sqrt{2} &= \\ = (r_1 + q_1\sqrt{2})[(r_2 + q_2\sqrt{2})(r_3 + q_3\sqrt{2})]. \end{aligned}$$

F3. Коммутативность операций

$$\begin{aligned} (r_1 + q_1\sqrt{2}) + (r_2 + q_2\sqrt{2}) &= (r_1 + r_2) + (q_1 + q_2)\sqrt{2} = \\ &= (r_2 + r_1) + (q_2 + q_1)\sqrt{2} = (r_2 + q_2\sqrt{2}) + (r_1 + q_1\sqrt{2}), \end{aligned}$$

$$\begin{aligned}(r_1 + q_1\sqrt{2})(r_2 + q_2\sqrt{2}) &= (r_1r_2 + 2q_1q_2) + (r_1q_2 + r_2q_1)\sqrt{2} = \\ &= (r_2r_1 + 2q_2q_1) + (r_2q_1 + r_1q_2)\sqrt{2} = (r_2 + q_2\sqrt{2})(r_1 + q_1\sqrt{2}).\end{aligned}$$

F4. Нулем в $\mathbb{Q}[\sqrt{2}]$ является число $0 + 0\sqrt{2}$.

F5. Противоположным для $r + q\sqrt{2}$ является число $-r - q\sqrt{2}$.

F6. Единицей в $\mathbb{Q}[\sqrt{2}]$ является число $1 + 0\sqrt{2}$.

F7. Вычисление обратного элемента:

$$\frac{1}{r + q\sqrt{2}} = \frac{r - q\sqrt{2}}{(r + q\sqrt{2})(r - q\sqrt{2})} = \frac{r - q\sqrt{2}}{r^2 - 2q^2}.$$

При этом нужно показать, что $r^2 - 2q^2 \neq 0$, если r и q одновременно не обращаются в ноль. Предположим, что это не так, т.е. $r^2 = 2q^2$, причем $q \neq 0$ (ясно, что тогда и $r \neq 0$), тогда $2 = (r/q)^2$, но тогда $\sqrt{2}$ — рациональное число. Противоречие. Следовательно, если $r + q\sqrt{2} \neq 0$, то оно обратимо.

F8. Дистрибутивность:

$$\begin{aligned}[(r_1 + q_1\sqrt{2}) + (r_2 + q_2\sqrt{2})](r_3 + q_3\sqrt{2}) &= \\ &= (r_1r_3 + r_2r_3 + 2q_1q_3 + 2q_2q_3) + (r_1q_3 + r_2q_3 + r_3q_1 + r_3q_2)\sqrt{2} = \\ &= (r_1 + q_1\sqrt{2})(r_3 + q_3\sqrt{2}) + (r_2 + q_2\sqrt{2})(r_3 + q_3\sqrt{2}).\end{aligned}$$

Итак, $\mathbb{Q}[\sqrt{2}]$ является полем.

В поле $\mathbb{Q}[\sqrt{2}]$ уравнение $x^2 - 2 = 0$ разрешимо. Причем в нем лежат оба корня данного уравнения: $\sqrt{2}$ и $-\sqrt{2}$.

Отметим еще один важный факт. В поле $\mathbb{Q}[\sqrt{2}]$ выражение $x^2 - 2$ можно записать в виде произведения линейных членов $(x - \sqrt{2})(x + \sqrt{2})$, поскольку $\sqrt{2}$ здесь стал «разрешенным» числом. Точно так же мы ранее сначала не могли записывать уравнения $0.5x - 1 = 0$, т.к. работали только с целыми числами (но могли заменить его эквивалентным уравнением $x - 2 = 0$), а после выхода в поле \mathbb{Q} у нас появилась возможность использовать дробные коэффициенты.

Возникает резонный вопрос: а если уравнение какое-то более сложное? Например, $x^5 + 3x^3 - 5 = 0$. Всегда ли его можно разложить на линейные множители в поле $\mathbb{Q}[\sqrt{2}]$? Или понадобится какое-то новое расширение \mathbb{Q} ? Иначе говоря, всегда ли будут корни такого уравнения лежать в построенных нами полях?

Ответ: нет. Но существует такое всеобъемлющее поле, в котором это действительно возможно.

Упражнения

Обязательные упражнения

8.1° Разложить в цепную дробь отношения: а) $36/25$; б) $111/34$; в) $12/8$; д) $1024/333$.

8.2° Решить уравнение в целых числах методом цепных дробей: а) $100x + 77y = 1$; б) $355x + 113y = 1$; в) $271x - 100y = 7$; д) $707x + 500y = 10$.

8.3° Маша продавала на школьной ярмарке плетеные мандалы по 135 рублей, а потом купила несколько фенечек по 40 рублей, после чего у нее осталось 5 рублей. Пользуясь методом цепных дробей, найдите, сколько фенечек купила Маша.

8.4° а) В фирме 28 служащих с большим стажем и 37 — с маленьким. Хозяин фирмы выделил некую сумму для подарков служащим на Новый год. Бухгалтер подсчитал, что есть только один способ разделить деньги так, чтобы все служащие с большим стажем получили поровну и все с маленьким — тоже поровну (все получают целое число рублей, большее 0). Какую наименьшую и какую наибольшую сумму мог выделить хозяин на подарки? б*) А если еще требуется, чтобы служащий с большим стажем получил больше денег, чем служащий с маленьким стажем?

8.5° При каком c прямая $ax + (\sqrt{3})y + c = 0$ пройдет через рациональную точку (x, y) ?

8.6° Решить уравнение $(\sqrt{3})x - (\sqrt{12})y = \sqrt{75}$ в целых числах.

8.7° Имеет ли решения в целых числах следующее уравнения: $x\sqrt{6} + y\sqrt{24} = \sqrt{12}$?

8.8° Сколько решений в зависимости от c может иметь уравнение $x + y\sqrt{3} = c$?

8.9° Методом цепных дробей найти наилучшее приближение с точностью до 0.001 следующих иррациональных чисел: $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}$.

8.10° Английский ярд составляет 0.914383 метра. Найти приближенное отношение метра к ярду.

8.11° Год равен 365.2422 суткам. Разложить эту дробь в цепную и найти первые четыре подходящие дроби.

8.12° Разность между последней и предпоследней подходящими дробями равна $1/42$. Подберите два-три набора пар чисел, которые могли бы

быть, соответственно, числителями и знаменателями этих подходящих дробей.

8.13° Разложите в цепную дробь число $43/40$. Найдите все ее подходящие дроби. Чему равна разность между последней и предпоследней дробями?

8.14° Решите уравнения в целых числах:

- a) $12x = 42y$;
- b) $ax + by = 0$, где $\text{НОД}(a, b) = d$;
- c) $2x + 3y = 1$;
- d) $4x + 6y = 2$;
- e) $4x + 6y = 5$;
- f) $20x + 21y = 2021$.

Дополнительные упражнения

8.15* Натуральные числа a и b взаимно просты. Докажите, что уравнение $ax + by = c$:

- a) при любом целом c имеет такое решение в целых числах x и y , что $0 \leq x < b$;
- b) имеет решение в *целых неотрицательных* числах x и y , если c целое, большее $ab - a - b$;
- c) при целых c от 0 до $ab - a - b$ ровно в половине случаев имеет целое неотрицательное решение, причем если для $c = c_0$ такое решение есть, то для $c = ab - a - b - c_0$ таких решений нет.

8.16* Слонопотам типа $\langle p, q \rangle$ ходит по бесконечной клетчатой доске, сдвигаясь за ход на p клеток по любому направлению «горизонталь-вертикаль» и на q клеток по перпендикулярному. (Шахматный конь — слонopotам типа $\langle 1, 2 \rangle$.) Какие слонopotамы могут попасть на соседнее с собой поле?

8.17* Натуральные числа m и n взаимно просты. Известно, что дробь

$$\frac{m + 179n}{179m + n}$$

можно сократить на число k . Каково наибольшее такое k ?

8.18* Есть шоколадка в форме равностороннего треугольника со стороной n , разделенная бороздками на равносторонние треугольники со стороной 1. Играют двое. За ход можно отломить от шоколадки треугольный кусок вдоль бороздки, съесть его, а остаток передать противнику. Тот, кто получит последний кусок — треугольник со стороной 1, — победитель. Тот, кто не может сделать ход, досрочно проигрывает. Кто выигрывает при правильной игре?

Арифметика остатков

Аннотация

Исчисление остатков дает богатый фактологический материал для изучения свойств простых чисел, а также позволяет по-новому взглянуть на операции Минковского с числовыми множествами и выйти на такие важные вехи теории множеств, как виды отношений и фактормножества.

9.1. Арифметика остатков

Рассмотрим бытовую задачу. Вам нужно выключить печку через 40 минут, но у вас нет таймера, зато есть будильник, на котором можно выставить время звонка. Сейчас 12:30, на какое время требуется поставить будильник? Ответ: 13:10. Почему так? Дело в том, что в часе 60 минут, и если к 30 минутам прибавить 40, то получается 70 минут, что больше часа. Поэтому добавляем 1 час и остаток — 10 минут.

Еще пример: сколько часов будет через 20 часов, если сейчас 8 утра? Можно решать аналогично: $8 + 20 = 28$, затем убираем полные сутки, т. е. 24 часа, остается 4 часа утра.

Можно решать иначе. 20 часов — это -4 часа от суток. Следовательно, нужно просто вычесть из 8 утра 4 часа и получим те же 4 часа утра.

Во всех случаях мы решаем задачу нахождения остатка от деления на некоторое число. В случае минут это 60, в случае часов это 24.

Когда вас просят отметить в анкете количество полных лет, то вам, по сути, нужно найти неполное частное от деления вашего возраста на 1 год. Конечно, в данном случае нам это просто сделать, т. к. каждый год мы запоминаем именно количество прожитых лет, а не дней или недель.

Но например, во многих сферах деятельности планирование календаря происходит неделями (и даже у себя в компьютере в настройках календаря вы можете вывести номер текущей недели в году). А сколько полных недель в году? Для этого нужно найти неполное частное от деления 365 (или 366) на 7, оно составляет 52.

Остаток от деления на неделю есть число от 0 до 6, которое определяет сдвиг вперед относительно текущего дня недели. Например, если сегодня четверг, то какой день недели будет через 30 дней? Мы выбрасываем 4 полных недели, что составляет 28 дней, и находим остаток, который равен 2. Это значит, что через 30 дней будет четверг плюс 2 дня, т. е. суббота.

Точно так же можно легко заметить, что каждый год происходит смещение дат на один или два дня вперед относительно дней недели. Так, если в этом году 1 января было средой, то в следующем оно будет или четвергом (если мы не переходим через 29 февраля), или пятницей (если текущий год — високосный, т. е. содержит 366 дней), как на картинке 9.1.

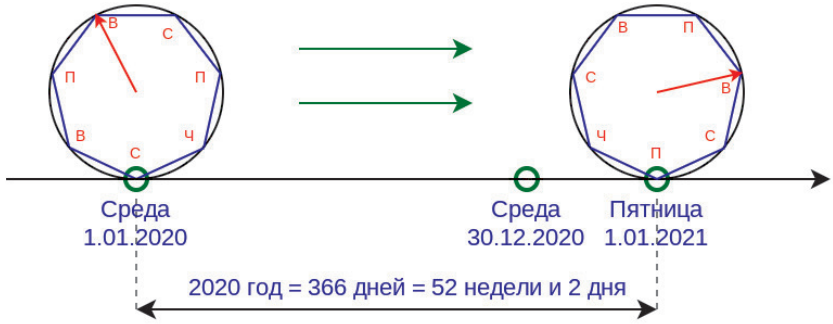


Рис. 9.1. Смещение дат в днях недели.

Каждые 28 лет (а 28 — это наименьшее общее кратное 7 и 4) соответствие дат и дней недели повторяется.

Если выписать последовательно сдвиг дней недели год за годом, то легко увидеть, что 28-летний цикл разбивается на несколько подциклов, через которые даты тоже повторяются. Рис. 9.2 дает представление о том, как это происходит: сначала даты повторяются через 6 лет (если текущая дата находится в интервале от 1 марта високосного года до 28 февраля следующего года), затем — через 11 лет, затем снова через 6 лет, затем — через 5 лет.



Рис. 9.2. Периоды повторения дат и дней недели.

При расчетах на более длительные периоды, а именно при переходе через 1900 год или 2100 год, нужно учитывать также, что три раза за

400 лет не происходит добавление лишнего дня (29 февраля) для более точного соответствия календаря астрономическому году, т. е. 1900, 1800, 1700 годы не являются високосными, как и 2100, 2200 и 2300.

Иными словами, часто в жизни встречается задача вычисления дня недели, и здесь нам на помощь приходит исчисление остатков по модулю 7. Например, сегодня 21 марта 2020 года, суббота, а нам нужно знать, какой день недели будет 31 августа 2020 года. Сначала мы находим день недели 21 августа, т. к. до этой даты целое число месяцев. При этом мы три раза переходим через 31 число (март, май, июль) и два раза не переходим (апрель, июнь). Следовательно, три раза прибавляется остаток 3, и два раза — остаток 2, итого сумма остатков составляет 13. Но это больше 7, причем очень близко к 14, поэтому сумму остатков мы запишем как -1 . Наконец, остается добавить 10 дней (от 21 августа до 31 августа). Итого получается 9, а по модулю 7 — всего 2. Таким образом, 31 августа 2020 года есть понедельник!

Из приведенной выше картинке с семиугольником на окружности, совмещенной с прямой линией, мы можем ясно представить себе, как работает исчисление остатков по модулю 7, т. е. исчисление дней недели. Мы катим окружность по прямой времени, пока не достигнем нужной нам даты. При этом неважно, сколько целых оборотов совершит семиугольник, т. е. сколько недель мы проедем, а вот последний неполный виток как раз и дает нам ответ на вопрос о дне недели. Так что, если мы пронумеруем дни недели цифрами от 0 до 6, то любое расстояние между датами можно представить как какое-то целое количество недель плюс остаток, лежащие в диапазоне от 0 до 6 (включительно).

Эта картинка легко обобщается на случай произвольного основания. Представим, что в неделе у нас не 7 дней, а, например, 28 (лунный месяц), и тогда любое расстояние между датами выражается как целое число 28-дневных циклов плюс некоторый остаток от 0 до 27. И так далее.

Следующим шагом обобщения является выход в целые числа. Ведь рассчитывать день недели можно не только в будущее, но и в прошлое с соответствующей сменой знака перед величиной сдвига. Так, если сегодня воскресенье, то 10 дней назад (-10) день недели был на 3 дня меньше (-3) или же на 4 дня больше, т. е. четверг.

Таким образом, мы приходим к тому, что всякое целое число a можно представить в виде $a = km + r$, где k — неполное частное от деления a на положительное m , r — остаток от деления, который находится в промежутке от 0 (включая) до m (не включая). Если число b имеет такой же точно остаток от деления на m , т. е. имеет место равенство $b = lm + r$ при некотором целом l , то говорят, что числа a и b **сравнимы**

по модулю m . Это записывается следующим образом:

$$a \equiv b \pmod{m}.$$

Читается: a **сравнимо с b** (по модулю m).

Причем если модуль m известен из контекста и не меняется при вычислениях, то его можно опускать, записывая просто $a \equiv b$.

Остаток от деления числа a на модуль m обозначается как $a \bmod m$ (без скобок у модуля). Таким образом, запись $a \equiv b \pmod{m}$ — это отношение на паре чисел a и b , его значениями могут быть только истина или ложь, а выражение $a \bmod m$ — это запись функции от аргументов a и m , ее значениями могут быть только числа от 0 до $m - 1$.

На картинке, приведенной выше, даты 01.01.2020 и 30.12.2020 сравнимы по модулю 7, т.е. по дням недели, что выражается фразой «это один и тот же день недели». А про интервал 366 дней мы запишем $366 \equiv 2 \pmod{7}$. Такая запись никак не информирует нас о коэффициенте k (количестве целых недель), но показывает самое главное — сколько дней надо прибавить к среде (см. рис. 9.1).

Остатками можно оперировать так же, как обычными числами, сбрасывая всякий раз накопленные при сложении целые «обороты» модулей. Иначе говоря, если мы хотим, например, к текущей среде прибавить 6 дней, то мы совмещаем наш семиугольник вершиной «среда» с прямой времени, а затем прокатываем его вперед на 6 делений (что чуть меньше полного оборота), и в точке касания с прямой получаем вторник. Заметим, что ровно тот же результат мы получим, если прокатим семиугольник назад на 1 деление. Это значит, что числа 6 и -1 сравнимы по модулю 7. И на практике можно также пользоваться отрицательными числами для исчисления остатков.

Ранее мы много времени уделяли таблицам композиций движений многоугольников. И, как мы помним, композиция вращений многоугольника соответствовала сложению углов этих вращений. При этом мы также отбрасывали 360 градусов (или 2π), если сумма углов переваливала за полный оборот. При описании конечных подгрупп движений правильных многоугольников мы выяснили, что каждый поворот является степенью некоторого минимального поворота на угол $2\pi/n$ (для n -угольника), т.е. все повороты выражаются углами $k(2\pi/n)$, где $k = 0, \dots, n - 1$ (а это остатки по модулю n).

Понаблюдаем теперь за степенями этих поворотов при композициях, т.е. при сложении углов. Для примера рассмотрим случаи $n = 7$ и $n = 8$, и выпишем таблицу композиций, которая представляет собой таблицу сложения остатков по модулям 7 и 8, соответственно (см. табл. 9.1).

Таблица сложения получается последовательными циклическими сдвигами верхней строки влево.

Таблица 9.1. Таблицы сложения остатков по модулям 7 и 8

	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7	0
2	2	3	4	5	6	7	0	1
3	3	4	5	6	7	0	1	2
4	4	5	6	7	0	1	2	3
5	5	6	7	0	1	2	3	4
6	6	7	0	1	2	3	4	5
7	7	0	1	2	3	4	5	6

Нетрудно видеть, что сложение по модулю удовлетворяет аксиомам абелевой группы, т. е. сложение ассоциативно, коммутативно, существует ноль, и для всякого элемента существует противоположный. Эти свойства наследуются из кольца \mathbb{Z} .

Помимо сложения остатков, мы можем их умножать (в терминологии вращений многоугольника умножение соответствует многократной композиции одинаковых поворотов, так что первое число произведения отвечает за величину поворота, а второе — за его кратность, либо наоборот). Таблицы умножения остатков по модулям 7 и 8 (отметим важную особенность этих таблиц: они имеют центральную симметрию, если вычеркнуть нулевые строку и столбец) приведены в табл. 9.2.

Таблица 9.2. Таблицы умножения остатков по модулям 7 и 8

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	0	2	4	6
3	0	3	6	1	4	7	2	5
4	0	4	0	4	0	4	0	4
5	0	5	2	7	4	1	6	3
6	0	6	4	2	0	6	4	2
7	0	7	6	5	4	3	2	1

Заметим, что умножение по модулю удовлетворяет аксиомам коммутативного моноида, т. е. оно ассоциативно, коммутативно и имеет нейтральный элемент — единицу. Эти свойства умножения наследуются из кольца \mathbb{Z} .

Таким образом, множество $\{0, 1, 2, \dots, m-1\}$ с операциями сложения и умножения по модулю m является конечным коммутативным кольцом с единицей. Оно называется **кольцом вычетов по модулю m** и обозначается так: \mathbb{Z}_m , либо $\mathbb{Z}/m\mathbb{Z}$. Природу второго обозначения мы поймем в последующих главах.

Отметим еще одно свойство умножения: строка или столбец, номер которого НЕ взаимно прост с модулем, содержит нули. Это легко доказать. Пусть номер строки равен k , и $s = \text{НОД}(k, m) > 1$. При этом ясно, что $s < m$, т.к. s является делителем m . Пусть также $t = m/s$. Рассмотрим тогда строку k и столбец t . Произведение их номеров равно $kt = km/s$. Поскольку k/s также целое, получаем, что kt кратно m , а значит, $kt \equiv 0 \pmod{m}$. Отметим, что $s = 1$ здесь не проходит ровно потому, что в этом случае t не будет номером столбца таблицы умножения.

На самом деле верно и обратное: если строка таблицы умножения содержит нули, то номер строки не взаимно прост с модулем. Для этого мы докажем эквивалентное утверждение.

Теорема 9.1. Пусть $k > 0$ и $k \perp m$, тогда все остатки

$$k \pmod{m}, \quad 2k \pmod{m}, \quad 3k \pmod{m}, \quad \dots, \quad (m-1)k \pmod{m}$$

попарно различны и отличны от нуля.

Доказательство. Предположим, что один из остатков равен нулю: $kl \equiv 0 \pmod{m}$, где $l \in \{1, 2, \dots, m-1\}$. Тогда $kl = mt$ при некотором t . Но поскольку $k \perp m$, в силу ОТА число k делит t , а значит, $k \leq t$. Однако $l < m$, следовательно, $kl < mt$. Противоречие.

Далее, если среди остатков есть равные, например, $kl \equiv kt$, то найдется и остаток $k(l-t)$ (или $k(t-l)$, если $t > l$), который равен 0. А это невозможно по доказанному.

Таким образом, эти остатки попарно различны и положительны. \square

Появление нулей в таблице умножения по непростому модулю означает, что в кольце вычетов по непростому модулю существуют **делители нуля**, т.е. такие ненулевые элементы, произведение которых равно нулю. Это — существенное отличие числовой системы \mathbb{Z}_m от тех, что мы встречали ранее.

Множество \mathbb{Z}_m^* , состоящее только из взаимно простых с модулем m элементов \mathbb{Z}_m , образует коммутативную группу по умножению по модулю m . Действительно, умножение замкнуто на множестве \mathbb{Z}_m^* , т.е. если $k \perp m$ и $l \perp m$, то $kl \perp m$ (это следует из основной теоремы арифметики: просто делитель m и kl будет и простым делителем одного из чисел k или l).

Ассоциативность и коммутативность умножения остатков наследуется от умножения в кольце целых чисел.

Кроме того, 1 является элементом \mathbb{Z}_m^* .

Чтобы доказать, что у произвольного элемента $k \in \mathbb{Z}_m^*$ имеется обратный, вспомним алгоритм Евклида, а точнее, его следствие: поскольку $k \perp m$, существуют такие целые числа a и b , что $ka + mb = 1$. Тогда число $ka + mb \equiv 1 \pmod{m}$, а значит, $ka \equiv 1 \pmod{m}$. Наконец, если $a = lm + r$, то верно равенство $kr \equiv 1 \pmod{m}$, т. е. остаток r является обратным к остатку k при умножении по модулю.

Остается показать, что $r \in \mathbb{Z}_m^*$, т. е. что $r \perp m$. Пусть $q|r$ и $q|m$. Тогда $q|a$ и, следовательно, $q|(ka + mb)$, т. е. $q|1$, откуда следует, что $q = \pm 1$. А это и означает, что r и m взаимно просты, т. е. $r \in \mathbb{Z}_m^*$. Тем самым, обратный элемент для произвольного остатка $k \in \mathbb{Z}_m^*$ существует и принадлежит \mathbb{Z}_m^* .

Следовательно, \mathbb{Z}_m^* — группа по умножению по модулю m .

Пусть p — простое число. В этом случае все числа $1, 2, \dots, p-1$ взаимно просты с p , так что $\mathbb{Z}_p^* = \{1, 2, \dots, p-1\}$. И поскольку, как мы только что выяснили, \mathbb{Z}_p^* является абелевой группой по умножению по модулю p , получаем, что \mathbb{Z}_p с операциями сложения и умножения по модулю p удовлетворяет аксиомам поля.

Итак, \mathbb{Z}_p при простом p — это пример *конечного поля*. Отметим, что конечные поля не исчерпываются одними лишь полями вычетов по простому модулю.¹

Рассмотрим таблицы умножения для групп \mathbb{Z}_5^* и \mathbb{Z}_8^* :

\mathbb{Z}_5^*	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

\mathbb{Z}_8^*	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

И тут мы снова видим знакомую ситуацию: если группа \mathbb{Z}_5^* циклическая (все ее элементы могут быть получены как степени двойки), и ее можно изоморфно сопоставить с группой \mathbb{Z}_4 с операцией сложения, а также с группой вращений квадрата, то группа \mathbb{Z}_8^* уже не является циклической, хотя остается коммутативной. И это — еще одна ипостась четверной группы Клейна. Чуть позже мы дадим сравнение нескольких вариантов групп 4-го порядка.

¹ Подробнее см. [13], Часть I.

9.2. Свойства арифметики остатков

Перечислим некоторые свойства отношения сравнимости:

- 1) $a \equiv b \pmod{m}$ тогда и только тогда, когда $a - b$ кратно m ;
- 2) если $a \equiv b$, $c \equiv d$, то $a + c \equiv b + d$, $a - c \equiv b - d$ и $ac \equiv bd$;
- 3) для $n \geq 0$ если $a \equiv b$, то $a^n \equiv b^n$;
- 4) признаки делимости на 3 и на 9: $a_0 + a_1 10 + a_2 10^2 + \dots + a_n 10^n \equiv a_0 + \dots + a_n$ по модулю 3 и по модулю 9;
- 5) если $m > 0$ и $d \perp m$, то

$$ad \equiv bd \pmod{m} \iff a \equiv b \pmod{m};$$

- 6) если $m, d > 0$, то

$$ad \equiv bd \pmod{md} \iff a \equiv b \pmod{m};$$

- 7) если $m > 0$, то для любого d

$$ad \equiv bd \pmod{m} \iff a \equiv b \pmod{m/\text{НОД}(m, d)};$$

- 8) если $m, d > 0$, $a \equiv b \pmod{md}$, то $a \equiv b \pmod{m}$;

- 9) если $m, n > 0$, то

$$a \equiv b \pmod{m}, \quad a \equiv b \pmod{n} \iff a \equiv b \pmod{\text{НОК}(m, n)};$$

- 10) если $m, n > 0$ и $m \perp n$, то

$$a \equiv b \pmod{m}, \quad a \equiv b \pmod{n} \iff a \equiv b \pmod{mn};$$

- 11) пусть n_p — степень простого числа p в разложении n по степеням простых (ОТА), тогда

$$a \equiv b \pmod{n} \iff \forall p \quad a \equiv b \pmod{p^{n_p}} \quad (p - \text{простое}).$$

Предположим, что у нас задана система из нескольких сравнений и требуется знать, разрешима ли она? Ответ на этот вопрос дает следующее утверждение.

Теорема 9.2 (Китайская теорема об остатках). *Если натуральные числа $m_1, m_2, \dots, m_n > 0$ попарно взаимно просты, то для любых целых r_1, r_2, \dots, r_n таких, что $0 \leq r_i < m_i$ при всех $i \in \{1, 2, \dots, n\}$, найдется число N , которое при делении на m_i дает остаток r_i при всех $i \in \{1, 2, \dots, n\}$.*

Более того, если найдутся два таких числа N_1 и N_2 , то $N_1 \equiv N_2 \pmod{m_1 \cdot m_2 \cdot \dots \cdot m_n}$.

Доказательство. Рассмотрим систему сравнений:

$$\begin{cases} x \equiv r_1 \pmod{m_1}, \\ x \equiv r_2 \pmod{m_2}, \\ \dots\dots\dots \\ x \equiv r_n \pmod{m_n}. \end{cases} \quad (9.1)$$

Обозначим $m = m_1 m_2 \dots m_n$ и рассмотрим x , вычисленный по формуле:

$$x = \sum_{j=1}^n r_j M_j M_j^{-1}, \quad (9.2)$$

где $M_j = \frac{m}{m_j}$, а M_j^{-1} — мультипликативно обратный к M_j элемент в группе $\mathbb{Z}_{m_j}^*$. Отметим, что обратный M_j^{-1} существует, т.к. $M_j \not\equiv 0 \pmod{m_j}$ в силу условий теоремы.

Проверим, что для такого x выполняется i -е сравнение в системе (9.1).

Если $j \neq i$, то j -е слагаемое в (9.2) сравнимо с нулем по модулю m_i , поскольку M_j кратно m_i , поэтому сумма (9.2) превращается лишь в одно слагаемое при $j = i$ (по модулю m_i):

$$x \equiv r_i M_i M_i^{-1} \equiv r_i \pmod{m_i},$$

последнее сравнение верно, т.к. M_i^{-1} является обратным к M_i именно по модулю m_i . Следовательно, x является решением системы (9.1).

Число x нами вычислено в обычной арифметике целых чисел. Покажем, что его выбор можно осуществлять с точностью до вычисления остатка по модулю m . Пусть $y \equiv x \pmod{m}$. Тогда $y = mk + x$ при некотором целом k . Так как m делится на m_i при всех $i = \overline{1, n}$, то $y \equiv x \pmod{m_i}$, т.е. также является решением системы (9.1).

Осталось показать, что если $y \not\equiv x \pmod{m}$, то y не является решением данной системы (единственность решения по модулю). Действительно, если y и x одновременно являются решениями системы (9.1), то $y - x \equiv 0 \pmod{m_i}$ при всех $i = 1, \dots, n$, т.е. $y - x$ делится на все m_i , а значит, делится и на их произведение (т.к. все они попарно взаимно просты). Тогда $y - x = mk$ при некотором целом k , т.е. $y \equiv x \pmod{m}$.

Итак, решениями системы (9.1) являются те и только те целые числа, которые сравнимы с x , заданным формулой (9.2), по модулю $m = m_1 m_2 \dots m_n$. □

Теорема 9.3 (Малая теорема Ферма). $n^{p-1} \equiv 1 \pmod{p}$, где p — простое и n не кратно p .

Доказательство. Согласно теореме 9.1 все остатки

$$n, 2n, 3n, \dots, (p-1)n \pmod{p}$$

различны и составляют множество $\{1, 2, \dots, p-1\}$. Тогда по свойствам сравнений будем иметь

$$n \cdot 2n \cdot 3n \dots (p-1)n \equiv 1 \cdot 2 \cdot 3 \dots (p-1) \pmod{p},$$

откуда $n^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$. Последнее тождество можно сократить на $(p-1)!$, поскольку все его множители взаимно просты с p . Откуда получаем

$$n^{p-1} \equiv 1 \pmod{p}.$$

□

Малая теорема Ферма позволяет вычислять обратный элемент по умножению в поле \mathbb{Z}_p . Достаточно n умножить на n^{p-2} , и мы получим 1. Но такой способ нахождения обратного элемента довольно трудоемкий. Проще всего воспользоваться алгоритмом Евклида, а точнее, его представлением в виде цепной дроби.

Пусть, например, $p = 101$. Это — простое число. И пусть требуется найти 77^{-1} по модулю 101. Можно воспользоваться малой теоремой Ферма и вычислить $77^{99} \pmod{101}$. Число

$$\begin{aligned} 77^{99} = & 578870998719483192604006423068948796479127581676622586145657719 \\ & 488743965292466241833275096332117841891352380031908081644221279 \\ & 6787687881706201273681438984199647603820668204476189223565013, \end{aligned}$$

как и его остаток от деления на 101, без компьютера не вычислить.

Теперь выпишем цепную дробь:

$$\frac{101}{77} = 1 + \frac{1}{3 + \frac{1}{5 - \frac{1}{5}}},$$

откуда, отсекая «хвост» $1/5$, сворачиваем дробь обратно и получаем

$$\frac{101}{77} \approx \frac{21}{16},$$

откуда видим, что $-101 \cdot 16 + 77 \cdot 21 = 1$. Поэтому $77 \cdot 21 \equiv 1 \pmod{101}$, т. е. $77^{-1} = 21$ в поле \mathbb{Z}_{101} .

Отметим, что существует обобщение малой теоремы Ферма, называемое **теоремой Эйлера**:

$$\text{если } n \perp m, \text{ то } n^{\varphi(m)} - 1 \equiv 1 \pmod{m},$$

где $\varphi(m)$ — функция Эйлера, определяемая по формуле (4.3).

Упражнения

Обязательные упражнения

9.1° Отметить на числовой оси целые числа, которые при делении на 7 дают остаток 2 (на рисунке должны поместиться числа, лежащие между -20 и 20).

9.2° Книги на столе пытались связывать в пачки по 2, по 3, по 4 и по 5 книг, и каждый раз оставалась одна лишняя. Сколько книг было на столе? (Известно, что их было не больше 100.)

9.3° Одному брату 6 лет, другому — 10. Значит, сумма возрастов — четная. Какой она будет в следующем году?

9.4° Если сегодня понедельник, то какой день недели будет через 10 дней, через 90 дней, через 2 года (рассмотреть случай без високосных лет и с високосным годом)?

9.5° Найти день недели через месяц, квартал, полгода и год, отправляясь от текущей даты.

9.6° Поезд Москва — Владивосток отправляется из Москвы в 7:00 и находится в пути 166 часов. Определите время прибытия (московское) поезда во Владивосток.

9.7° Построить таблицы сложения для модулей: 2, 3, 4, 5, 6, 10, 11.

9.8° Найти число, которое при делении на 2 дает остаток 1, при делении на 3 остаток 2, при делении на 4 остаток 3, при делении на 5 остаток 4, при делении на 6 остаток 5 и при делении на 7 дает остаток 6.

9.9° Докажите свойства сравнений 1)–11), перечисленные в разделе 9.2.

9.10° Верно ли, что а) если $n:k$ и $k:n$, то $n = \pm k$; б) если $a|b$ и $b|c$, то $a|c$; в) если $b:a$ и $c:a$, но $d \not\mid a$, то $(b+c):a$, но $(b+d) \not\mid a$; г) если a и b не делятся на c , то ab не делится на c^2 ?

9.11° Что означает запись $a \equiv b \pmod{0}$?

9.12° Обозначим за \oplus сложение по модулю 2, т. е. $a \oplus b = (a + b) \pmod{2}$, если $a, b \in \{0, 1\}$. Для битовых последовательностей эта операция применяется попозиционно (например, $110 \oplus 101 = 011$).

Алиса и Боб придумали следующий алгоритм шифрования. Каждый из них определил случайную последовательность длины n : A и B соответственно. Алиса передает Бобу сообщение t длиной в n битов следую-

щим способом: она отправляет ему сообщение $m_1 = m \oplus A$, в ответ Боб отправляет ей $m_2 = m_1 \oplus B$, затем Алиса отправляет Бобу $m_3 = m_2 \oplus A$. Как Боб сможет прочесть сообщение m , зная алгоритм и сообщение m_3 ? Как Ева, перехватившая сообщения m_1, m_2, m_3 , сможет прочесть исходное сообщение m ?

9.13° Целое положительное число увеличили на 1. Могла ли сумма его цифр: а) возрасти на 8; б) уменьшиться на 8; в) уменьшиться на 10?

9.14° Какие остатки может давать точный квадрат при делении на 4?

9.15° Последняя цифра точного квадрата равна 6. Доказать, что его предпоследняя цифра нечетна.

9.16° Остаток от деления простого числа на 30 — простое число или 1. Почему?

9.17° Какое наибольшее число различных целых чисел можно выбрать, если требуется, чтобы сумма и разность любых двух из них не делились на 15?

9.18° На какую цифру оканчивается число $33^{77} + 77^{33}$?

9.19° Могут ли среди m последовательных целых чисел какие-то два иметь равные остатки от деления на m ?

9.20° Пусть $5x \equiv 6 \pmod{8}$. Найти x .

9.21° Найти последнюю цифру: а) 7^{100} ; б) 7^{1942} .

9.22° Найдите остаток от деления: а) числа $1 + 31 + 331 + \dots + 3333333331$ на 3; б) 6100 на 7.

9.23° Найдите остаток от деления числа $1 - 11 + 111 - 1111 + \dots - 1111111111$ на 9.

9.24° Найдите остаток от деления: а) $10!$ на 11; б) $11!$ на 12.

9.25° а) Какой цифрой оканчивается 8^{18} ? б) При каких натуральных k число $2^k - 1$ кратно 7?

9.26° Найдите три последние цифры числа 1999^{2000} .

9.27° Найти: а) $3^{31} \pmod{7}$; б) $2^{35} \pmod{7}$; в) $128^{129} \pmod{17}$.

9.28° Докажите, что: а) $30^{99} + 61^{100}$ делится на 31; б) $43^{95} + 57^{95}$ делится на 100.

9.29° Докажите, что $1^n + 2^n + \dots + (n-1)^n$ делится на n при нечетном n .

9.30° Числа x и y целые, причем $x^2 + y^2$ делится на 3. Докажите, что x и y делятся на 3.

9.31° Какие целые числа дают при делении на 3 остаток 2, а при делении на 5 — остаток 3?

9.32° Докажите, что остаток от деления простого числа на 30 есть или простое число или 1.

9.33° а) Квадрат целого положительного числа оканчивается на ту же цифру, что и само число. Что это за цифра? (Указать все возможности.)

б) Квадрат целого положительного числа оканчивается на те же две цифры, что и само число. Что это за цифры? (Указать все возможности.)

с) Пятая степень числа оканчивается на ту же цифру, что и само число. Почему? Для каких еще степеней это верно?

9.34° Доказать, что для любого целого a число $10a$ дает при делении на 9 тот же остаток, что и само a .

9.35° Доказать, что число и его сумма цифр дают одинаковые остатки при делении на 3 и 9.

9.36° Найти обратные остатки к 5, 9, 12, 25, 51, 88, 99, 100 по модулю 101.

9.37° Найти (или доказать, что их не существует) обратные остатки к 10, 20, 30, 27, 51, 86 по модулю 2021. А по модулю 2022?

9.38° Докажите, что из любых n целых чисел всегда можно выбрать несколько, сумма которых делится на n (или одно число, делящееся на n).

9.39° Число a дает остаток 5 при делении на 9, число b дает остаток 7 при делении на 9. Можно ли по этим данным определить, какой остаток дают числа $a + b$ и ab при делении на 9?

9.40° Число p — простое. Докажите, что $\binom{p}{k}$ делится на p , если $0 < k < p$.

9.41° (Второе доказательство малой теоремы Ферма). Пусть p — простое, n — целое. а) Докажите индукцией по n , что $n^p - n$ делится на p .

б) Докажите, что если $n \perp p$, то $n^{p-1} - 1$ делится на p .

9.42° Докажите, что $2222^{5555} + 5555^{2222}$ делится на 7.

9.43° а) Числа p и q — простые, $2p - 1$ делится на q . Докажите, что $q - 1$ делится на p . б) Простое ли $2^{13} - 1$? с) Простое ли число $257^{60} + 60$?

9.44° Пусть p не равно 5. Докажите, что среди чисел, записываемых только единицами, есть число, которое делится на p .

9.45° Пусть p – простое число вида $4k + 1$, и пусть $x = (2k)!$. Докажите, что $x^2 \equiv -1 \pmod{p}$.

9.46° Пусть p – простое число вида $4k + 1$, и пусть x удовлетворяет сравнению $x^2 \equiv -1 \pmod{p}$. Докажите, что:

- $(a + xb)(a - xb) \equiv a^2 + b^2 \pmod{p}$ при $a, b \in \mathbb{Z}$;
- среди значений выражения $m + xn$, где $m, n \in \mathbb{Z}$, $0 \leq m, n \leq \lfloor \sqrt{p} \rfloor$, найдутся два различных с равными остатками от деления на p ;
- найдется ненулевое число $a + bx$, делящееся на p , где $a, b \in \mathbb{Z}$, причем $|a| < \sqrt{p}$ и $|b| < \sqrt{p}$;
- p представимо в виде суммы двух квадратов целых чисел.

9.47° а) Пусть p простое и имеет вид $4k + 3$. Найдется ли такое целое x , что $x^2 \equiv -1 \pmod{p}$? б) Докажите, что если $x^2 + 1$ делится на нечетное простое число p , то p имеет вид $4k + 1$. в) Докажите, что простых чисел вида $4k + 1$ бесконечно много. г) Пусть p простое и имеет вид $4k + 1$. Найдите такое целое x , что $x^2 \equiv -1 \pmod{p}$.

Дополнительные упражнения

9.48* Сколько есть способов записать 2018 как сумму натуральных слагаемых, любые два из которых равны или различаются на 1? (Способы лишь с разным порядком слагаемых считаем равными.)

9.49* Докажите, что из любых n целых чисел всегда можно выбрать несколько, сумма которых делится на n (или одно число, делящееся на n).

9.50* Боб хочет послать Алисе сообщение, выраженное числом m . На этот раз они используют алгоритм шифрования RSA.

RSA устроен следующим образом.

- Берем некоторое большое число N (все сообщения должны быть остатками по модулю N), которое плохо раскладывается на простые множители (например, полупростое, т.е. $N = pq$, где p, q – большие простые числа, обычно 1024 или 2048-битные).
- Берем также некоторое число $e < \varphi(N)$, взаимно простое с $\varphi(N)$.²
- Находим $d = e^{-1}$ по модулю $\varphi(N)$, т.е. такое, что $e \cdot d \equiv 1 \pmod{\varphi(N)}$.
- Пара $\langle e, N \rangle$ называется **открытым ключом**, пара $\langle d, N \rangle$ – **закрытым ключом**.
- Сообщение m , которое должно быть взаимно просто с N , кодируем числом $m_1 = m^e \pmod{N}$.

² Функция Эйлера $\varphi(N)$ задается формулой (4.3)

6) Чтобы расшифровать сообщение, пользуемся закрытым ключом d :

$$m_1^d = m^{e \cdot d} = m^{\varphi(N)k+1} \equiv m \pmod{N}$$

в силу теоремы Эйлера.

Алиса и Боб заранее обмениваются закрытым ключом d . Сообщение m пересылается в зашифрованном виде Алисе, а вместе с ним открытый ключ $e = 53$, $N = 299$. Зашифрованное сообщение $m^e \pmod{N}$ равно числу 171. Эти данные (открытый ключ и кодированное сообщение) перехватывает Ева.

Опишите действия Евы по расшифровке сообщения m и найдите число m .

Многочлены

Пусть дано какое-то коммутативное кольцо K с единицей (например, \mathbb{Z} или \mathbb{Z}_m). Тогда **многочленом степени** $n \geq 0$ **над** K называется всякое выражение вида

$$\sum_{s=0}^n k_s x^s = k_n x^n + k_{n-1} x^{n-1} + \dots + k_1 x + k_0,$$

где $k_n \neq 0$. Заметим, что многочлен степени $n = 0$ — это константа k_0 , отличная от нуля. Тожественный ноль принято называть многочленом степени $-\infty$. Степень многочлена P принято обозначать $\deg P$. Например, $\deg(x^2 - 1) = 2$.

Множество всех многочленов от переменной x с коэффициентами из K обозначается за $K[x]$. Например, $\mathbb{Z}[x]$ — многочлены с целыми коэффициентами, $\mathbb{Q}[x]$ — многочлены с рациональными коэффициентами.

Многочлены **равны**, если равны коэффициенты при соответствующих степенях, т. е. если $P(x) = \sum p_k x^k$ и $Q(x) = \sum q_k x^k$, то

$$P = Q \Leftrightarrow \forall k \ p_k = q_k.$$

Многочлены можно складывать, вычитать и умножать. Операции сложения и умножения вводятся следующим образом:

$$(P + Q)(x) = \sum_k (p_k + q_k) x^k, \quad (PQ)(x) = \sum_k \sum_{i+j=k} (p_i q_j) x^k. \quad (10.1)$$

Множество $K[x]$ с такими операциями называется **кольцом многочленов** и действительно является кольцом, причем коммутативным кольцом с единицей в силу таковых же свойств кольца K .

Проверим это, следуя списку аксиом кольца, приведенному на стр. 66–67.

R1: сумма и произведение многочленов являются многочленами — это легко видеть из формул (10.1), поскольку в обоих случаях получается конечная сумма степеней переменной x с коэффициентами, принадлежащими K (уже в силу свойств кольца K).

R2: ассоциативность операций сложения и умножения также следует из свойств кольца K . Покажем это на примере.

$$\begin{aligned} [(5x^2 + 3x + 1) + (x^2 + 2)] + (-3x - 1) &= \\ &= 6x^2 + 0x + 2 = \\ &= (5x^2 + 3x + 1) + [(x^2 + 2) + (-3x - 1)], \end{aligned}$$

здесь работает ассоциативность сложения в кольце K .

$$\begin{aligned} [(5x^2 + 3x + 1)(x^2 + 2)](-3x - 1) &= \\ &= -15x^5 - 14x^4 - 36x^3 - 29x^2 - 12x - 2 = \\ &= (5x^2 + 3x + 1)[(x^2 + 2)(-3x - 1)] \end{aligned}$$

R3: нулевой многочлен — это константа 0.

R4: противоположный многочлен: $-P(x) = (-k_n)x^n + (-k_{n-1})x^{n-1} + \dots + (-k_0)$.

R5: дистрибутивность: $[P(x) + Q(x)]R(x) = P(x)R(x) + Q(x)R(x)$ — следует из свойств кольца K . Предлагаем читателю проверить это самостоятельно.

R6: коммутативность сложения — прямо следует из коммутативности сложения в кольце K .

R7: единица — это константа 1, которая существует в кольце K .

R8: коммутативность умножения прямо следует из коммутативности умножения в кольце K .

Теорема 10.1 (Безу́). Пусть P — многочлен над коммутативным кольцом K с единицей. Тогда для любого $c \in K$ существует многочлен $Q \in K[x]$ такой, что

$$P(x) = (x - c)Q(x) + P(c).$$

Доказательство. Пусть $P(x) = p_0 + p_1x + \dots + p_nx^n$, $p_n \neq 0$, и $Q(x) = q_0 + q_1x + \dots + q_nx^n + \dots + q_{n+m}x^{n+m}$ (ниже мы увидим, что достаточно брать $m = 0$, т.к. вышестоящие коэффициенты обращаются в ноль). При этом мы не требуем, чтобы старший коэффициент не обращался в ноль, т.к. пока не знаем степени данного многочлена.

Решим уравнение

$$P(x) = (x - c)Q(x) + h$$

относительно коэффициентов q_k . Раскрывая скобки и приравнивая ко-

эффиценты при одинаковых степенях, получаем систему уравнений

$$\begin{aligned}
 p_0 &= h - cq_0 \\
 p_1 &= q_0 - cq_1 \\
 &\dots\dots \\
 p_{n-1} &= q_{n-2} - cq_{n-1} \\
 p_n &= q_{n-1} - cq_n \\
 0 &= q_n - cq_{n+1} \\
 &\dots\dots \\
 0 &= q_{n+m-1} - cq_{n+m} \\
 0 &= q_{n+m}
 \end{aligned}$$

Решая эту систему снизу вверх, находим, что

$$\begin{aligned}
 q_n &= q_{n+1} = \dots = q_{n+m} = 0 \\
 q_{n-1} &= p_n \\
 q_{n-2} &= p_{n-1} + cp_n \\
 &\dots\dots \\
 q_0 &= p_1 + cp_2 + \dots + c^{n-1}p_n \\
 h &= p_0 + p_1c + \dots + p_nc^n = P(c)
 \end{aligned}$$

Как видим, система однозначно разрешается в кольце K , и остаток h действительно равен $P(c)$. Кроме того, видим, что степень $Q(x)$ в точности равна $n - 1$. \square

Теорема Безу хороша тем, что работает в кольце многочленов над любым коммутативным кольцом с единицей!

Корнями многочлена называются числа, зануляющие его, т.е. это такие числа, которые, будучи подставленными вместо переменной x , обращают значение многочлена в ноль. Например, числа $\sqrt{2}$ и $-\sqrt{2}$ являются корнями многочлена $x^2 - 2$. Корни многочлена не всегда лежат в том же кольце, где и его коэффициенты. Это делает возможным расширять кольца и поля с помощью присоединения корней многочленов, заданных над этими кольцами и полями.

Из теоремы Безу следует, что α — корень $P(x)$ тогда и только тогда, когда P есть произведение двучлена $(x - \alpha)$ и другого многочлена меньшей степени, т.е. когда P делится на $(x - \alpha)$ в кольце многочленов. Например, многочлен $x^k - 1$ делится на $x - 1$, потому что 1 — корень $x^k - 1$. В самом деле,

$$x^k - 1 = (x - 1)(x^{k-1} + x^{k-2} + \dots + x + 1).$$

Заметим, что в привычной нам арифметике степень произведения многочленов равна сумме степеней сомножителей:

$$\deg(PQ) = \deg(P) + \deg(Q).$$

Однако это верно не всегда. Рассмотрим, к примеру, кольцо вычетов по модулю 8. В таком кольце:

$$(2x^2 + 3x + 7)(4x + 4) = (8x^3 + 20x^2 + 40x + 28) \equiv 4x^2 + 4 \pmod{8},$$

т. е. правило сложения степеней нарушилось, т. к. коэффициент перед старшей степенью оказался сравним с нулем по модулю 8.

Это — не единственная проблема многочленов над произвольными кольцами. Рассмотрим многочлен $x^2 - 1$ над тем же самым кольцом \mathbb{Z}_8 . Попробуем его разложить на множители. Школьная формула разности квадратов сразу дает ответ:

$$x^2 - 1 = (x - 1)(x + 1),$$

но, поскольку $1 \equiv -7 \pmod{8}$, правильнее будет записать так:

$$x^2 - 1 = (x - 1)(x - 7).$$

Однако легко проверить, что числа 3 и 5 также являются корнями многочлена $x^2 - 1$ в кольце \mathbb{Z}_8 . Стало быть, он делится также на $(x - 3)$ и $(x - 5)$.

Получается, с одной стороны, линейные многочлены $(x - 1)$, $(x - 3)$, $(x - 5)$, $(x - 7)$ — простые, т. е. не раскладываются в произведение других многочленов, а с другой стороны — все они являются делителями многочлена второй степени $x^2 - 1$. Интуиция, связанная с основной теоремой арифметики, подсказывает нам, что тут что-то не так: или произведение четырех многочленов первой степени дает многочлен второй степени (что мы уже видели выше, но там обнуление старших степеней случилось из-за «удачного» перемножения коэффициентов, здесь же такой довод не работает, т. к. все коэффициенты равны 1), или многочлен $x^2 - 1$ представляется разными способами в виде произведения «простых» многочленов, что уже напрямую противоречит ОТА.

На самом деле, оба вывода косвенно связаны между собой, поскольку в их основе лежит такое «нехорошее» свойство некоторых колец, как наличие *делителей нуля*, т. е. таких ненулевых элементов, произведение которых равно нулю (например, $4 \cdot 2 \equiv 0 \pmod{8}$).

В таких кольцах ОТА не то, чтобы не работает, там ее в принципе невозможно сформулировать, поскольку отсутствует понятие простого элемента (аналога простого числа).

Если же K является кольцом без делителей нуля (в науке алгебре коммутативное кольцо без делителей нуля еще называется *областью целостности*; к таковым, например, относится кольцо целых чисел), а еще лучше — полем, то такой проблемы нет, и кольцо многочленов $K[x]$ становится намного более привлекательным, а его арифметика — похожей на арифметику целых чисел.

Многочлены, заданные над полем, можно делить друг на друга с остатком так же, как это делается с обычными целыми числами. Например, произведем деление столбиком многочлена $x^3 - 4x^2 - x - 3$ на многочлен $x^2 + x + 1$:

$$\begin{array}{r|l} x^3 - 4x^2 - x - 3 & x^2 + x + 1 \\ \hline -x^3 + x^2 + x & \\ \hline -5x^2 - 2x - 3 & \\ -5x^2 - 5x - 5 & \\ \hline 3x + 2 & \end{array}$$

Отсюда видим, что

$$x^3 - 4x^2 - x - 3 = (x - 5)(x^2 + x + 1) + (3x + 2),$$

где красным цветом выделено неполное частное, а синим — остаток от деления.

Многочлен может разделиться нацело, т.е. без остатка, на другой многочлен, например, $x^3 - 1$ делится на $x - 1$ без остатка, т.к.

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

Что при этом происходит с коэффициентами многочленов — не так важно, поскольку мы работаем над полем, где деление разрешено на любой ненулевой элемент. Поэтому коэффициенты вполне могут быть нецелыми, хотя деление многочленов получилось без остатка:

$$(x^2 - 1)/(2x - 2) = 0.5x + 0.5.$$

На этом же примере легко увидеть, что разложение многочлена на множители определяется с точностью до коэффициентов:

$$x^2 - 1 = (x - 1)(x + 1) = (2x - 2)(0.5x + 0.5) = (100x - 100)(0.01x + 0.01)$$

и т.д.

Теория делимости многочленов над полем во многом повторяет теорию делимости целых чисел. Здесь также есть простые, или **неприводимые**, многочлены, которые невозможно разложить в произведение

многочленов меньшей положительной степени над тем же полем, есть обратимые элементы — это все многочлены-константы, кроме нуля. Наконец, есть алгоритм Евклида и аналог основной теоремы арифметики о единственности разложения многочлена в произведение неприводимых (с точностью до коэффициентов-констант).

Так, при делении многочленов в остатке всегда получается многочлен степени, меньшей, чем делитель. Точнее, пусть P_n — многочлен степени n , а Q_m — многочлен степени $m < n$, тогда справедливо представление

$$P_n = Q_m G + H,$$

где степень многочлена H меньше m . При этом степень неполного частного G будет равна $n - m$.

В процессе выполнения алгоритма Евклида степень остатка все время падает (так же, как падает модуль остатка при делении целых чисел). Такое снижение степени остатка позволяет провести алгоритм Евклида за конечное число шагов, т. к. в конце концов остаток будет иметь степень 0 или $-\infty$, т. е. будет каким-то числом, не зависящим от переменной x .

Если остаток оказался константой, отличной от нуля, то на следующем шаге предыдущий остаток разделится на эту константу без остатка. В любом случае алгоритм заканчивается нулевым остатком, т. е. многочленом степени $-\infty$. Тогда предыдущий остаток (будь то ненулевая константа или многочлен положительной степени) будет представлять собой НОД исходных многочленов.

Например, найдем $\text{НОД}(x^3 - 6x^2 + 11x - 6, x^2 - 1)$:

$$\begin{aligned} x^3 - 6x^2 + 11x - 6 &= (x^2 - 1)(x - 6) + (12x - 12), \\ x^2 - 1 &= (12x - 12) \left(\frac{x}{12} + \frac{1}{12} \right) + 0. \end{aligned}$$

Красным выделены коэффициенты алгоритма, т. е. неполные частные. Поскольку алгоритм закончился нулевым остатком, последний остаток положительной степени $12(x - 1)$ является искомым НОД.

Заметим, что НОД определяется с точностью до коэффициента-константы, поскольку умножение многочленов на ненулевую константу никак не влияет на их делимость или неприводимость. Поэтому проще записать найденный НОД в виде $x - 1$.

Сравните этот факт с разложением исходных многочленов: $x^3 - 6x^2 + 11x - 6 = (x - 1)(x - 2)(x - 3)$, $x^2 - 1 = (x - 1)(x + 1)$. Как видим, у них есть общий делитель $x - 1$.

Еще пример: пусть даны два многочлена $x^2 - 3x + 2$ и $x^2 - 2x - 3$, тогда

$$\begin{aligned}x^2 - 3x + 2 &= (x^2 - 2x - 3) \cdot 1 - (x - 5), \\x^2 - 2x - 3 &= (x - 5)(x + 3) + 12.\end{aligned}$$

В данном случае алгоритм Евклида в конце дает остаток 12, т. е. ненулевую константу. На следующем шаге остаток будет равен нулю, поэтому константа 12 (и любая отличная от нуля константа) и есть НОД многочленов $x^2 - 3x + 2$ и $x^2 - 2x - 3$. Как и в предыдущем случае по модулю умножения на константу мы можем считать, что НОД = 1.

Такие многочлены, по аналогии с целыми числами, называются **взаимно простыми**.

Сравните этот факт с разложением исходных многочленов: $x^2 - 3x + 2 = (x - 1)(x - 2)$, $x^2 - 2x - 3 = (x - 3)(x + 1)$. Как видим, у них нет общих делителей.

Разложение многочлена на линейные множители сразу же дает нам список корней этого многочлена. Но, как мы видели выше, этот список не всегда полный. Покажем, что для многочленов над полем (а если точнее — над целостным кольцом) такой проблемы не существует.

Теорема 10.2 (о корнях многочлена над полем). *Если K — поле, то количество различных корней многочлена из $K[x]$ не превышает его степени.*

Доказательство. Воспользуемся индукцией. Очевидно, что для линейного многочлена $k_0 + k_1x$ корень определяется однозначно: он равен числу $-k_0/k_1$.

Предположим, что для всех степеней ниже n теорема верна, и рассмотрим многочлен $P(x)$ степени n (т. е. $k_n \neq 0$).

Предположим, что $P(x)$ имеет более чем n различных корней. Пусть α — один из его корней. Тогда по теореме Безу

$$P(x) = (x - \alpha)Q(x), \tag{10.2}$$

где $Q(x)$ — многочлен степени $n - 1$. Но у $P(x)$ есть еще как минимум n различных корней, кроме α . Пусть β — один из таких корней $P(x)$, тогда $\beta \neq \alpha$ и

$$0 = P(\beta) = (\beta - \alpha)Q(\beta),$$

откуда $Q(\beta) = 0$ (поскольку в поле нет делителей нуля, а $(\beta - \alpha) \neq 0$), т. е. β оказался корнем многочлена $Q(x)$. И так — для всех корней $P(x)$, отличных от α . Следовательно, если таковых будет не меньше n , то многочлен $Q(x)$ имеет как минимум n различных корней, в то время как его степень равна $n - 1$. А это противоречит предположению индукции. Следовательно, $P(x)$ не может иметь более чем n различных корней. \square

Эту теорему можно уточнить, учитывая кратности корней. Корень α многочлена $P(x)$ имеет кратность k , если P делится на $(x - \alpha)^k$ и не делится на $(x - \alpha)^{k+1}$. Пусть многочлен P имеет степень n и корни α_1 — кратности k_1, \dots, α_m — кратности k_m . Тогда $k_1 + \dots + k_m \leq n$. Для доказательства этого факта нужно в разложении (10.2) делить сразу на максимальную степень двучлена $(x - \alpha)$.

Неравенство $k_1 + \dots + k_m \leq n$ превращается в равенство, если мы имеем дело с многочленами над полем комплексных чисел, поскольку над этим полем любой многочлен раскладывается в произведение линейных многочленов! И этот замечательный факт называется **Основной теоремой алгебры** (доказательство которой мы не рассматриваем в данном курсе в виду его немалой сложности).

Поскольку мы договорились о том, что рассматриваем многочлены над некоторым полем, мы всегда можем произвольный многочлен степени n привести к виду $x^n + k_{n-1}x^{n-1} + \dots + k_0$, т.е. разделить его на коэффициент при старшей степени x^n (т.к. он не нулевой и деление в поле на него разрешается). Такие многочлены называются **приведенными** (или нормированными). Легко понять, что произведение приведенных многочленов есть приведенный многочлен. Соответственно, если P, Q — приведенные многочлены и $P:Q$, то их частное P/Q — тоже приведенный многочлен. Как уже отмечалось ранее, НОД многочленов мы определяем с точностью до коэффициента-константы, а это значит, что можно всякий раз в качестве НОД выбирать приведенный многочлен.

Таким образом, когда мы говорим о делимости многочленов, мы вполне можем оперировать только приведенными многочленами (хотя остатки от деления уже могут оказаться не приведенными).

Можно кое-что сказать о соотношении между его корнями x_1, \dots, x_n и его коэффициентами k_0, \dots, k_{n-1} .

Теорема 10.3 (Виет). *Если имеет место разложение*

$$k_0 + k_1x + \dots + k_{n-1}x^{n-1} + x^n = (x - x_1)(x - x_2) \dots (x - x_n),$$

то

$$k_0 = (-1)^n x_1 \dots x_n \text{ (произведение всех корней)}$$

$$k_1 = (-1)^{n-1} x_1 \dots x_n / x_1 + \dots + (-1)^{n-1} x_1 \dots x_n / x_n$$

$$\text{(сумма всех произведений по } n - 1 \text{ корней)}$$

.....

$$k_{n-2} = (x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n) \text{ (сумма произведений всех пар корней)}$$

$$k_{n-1} = -(x_1 + \dots + x_n) \text{ (сумма всех корней)}$$

Доказательство. Указанные тождества можно получить, сравнивая коэффициенты при x^m у произведения $(x - x_1)(x - x_2) \dots (x - x_n)$. Действительно, перед x^0 мы получаем произведение всех корней, взятых со знаком «-», перед x^1 — сумму произведений всех корней со знаком «-» за исключением одного корня (вместо которого в произведение попадает переменная x), причем таких слагаемых столько, сколько корней, т. е. n . И так далее. Наконец, степень x^{n-1} собирает все произведения, в которых $n - 1$ раз входит переменная x и 1 раз — какой-то из корней со знаком «-». Коэффициент же при степени x^n у произведения $(x - x_1)(x - x_2) \dots (x - x_n)$, очевидно, равен 1 и от корней никак не зависит. \square

Теорема о корнях многочлена дает некоторые полезные следствия. Рассмотрим поле \mathbb{Z}_p вычетов по простому модулю p .

Теорема 10.4 (Вильсон). *Если p — простое, то $(p - 1)! + 1$ делится на p .*

Доказательство. Рассмотрим многочлен $x^{p-1} - 1$ над полем \mathbb{Z}_p . В силу малой теоремы Ферма все числа $1, \dots, p - 1$ являются его корнями, причем других корней нет (хотя это и так ясно, т. к. ноль, очевидно, не является корнем). Тогда в силу теоремы Безу данный многочлен раскладывается в произведение линейных членов:

$$x^{p-1} - 1 \equiv (x - 1)(x - 2) \dots (x - (p - 1)).$$

Внимание! Все операции в поле \mathbb{Z}_p производятся по модулю p . В обычном числовом поле, например, в \mathbb{Q} , такое разложение не будет выполняться, т. к. $x^m - 1 = (x - 1)(x^{m-1} + x^{m-2} + \dots + 1)$ и при любом положительном x , кроме 1, он не обращается в ноль.

Осталось применить теорему Виета для свободного члена $k_0 = -1$. При нечетном простом p мы имеем $(-1)^{p-1} = 1$, так что:

$$-1 \equiv (-1)^{p-1} 1 \cdot 2 \dots (p - 1) \equiv (p - 1)! \pmod{p},$$

что и доказывает теорему Вильсона для случая $p > 2$. А в случае $p = 2$ это легко проверить непосредственно. \square

Упражнения

Обязательные упражнения

10.1^o Выполните действия с многочленами:

a) $(1 + x)(1 + x^2)(1 + x^4)(1 + x^8)(1 + x^{16});$

- b) $(1 + x + x^2 + x^3 + \dots + x^9)^2$;
 c) $(x^4 - 9x^3 + 16x^2 + 36x - 80)/(x^2 - 4)$;
 d) $(x^{56} + x^{55} + x^{54} + \dots + x^2 + x + 1)/(x^{18} + x^{17} + \dots + x + 1)$.

10.2° Многочлены $P(x)$ и $Q(x)$ имеют целые коэффициенты, причем каждый из них имеет хотя бы один нечетный коэффициент. Докажите, что у произведения $P(x)Q(x)$ также есть хотя бы один нечетный коэффициент.

- 10.3° Делится ли: a) $x^{1000} + x^{999} + \dots + x + 1$ на $x^5 + x^4 + x^3 + x^2 + x + 1$;
 b) $x^4 + x - 2$ на $x + 2$; c) 57 на $x - 2$?

10.4° Пусть $P(x)$ — многочлен степени n , и пусть a — некоторое число. Докажите, что $P(x)$ можно записать в виде $c_0 + c_1(x - a) + \dots + c_n(x - a)^n$, подобрав подходящие числа c_0, \dots, c_n .

10.5° Доказать, что в любом кольце (даже некоммутативном и без единицы) выполняется аксиома нуля: $a \cdot 0 = 0 \cdot a = 0$.

10.6° Если K — поле, то будет ли $K[x]$ полем?

10.7° Решите уравнение $x^2 = 1$ в $\mathbb{Z}/m\mathbb{Z}$ для $2 \leq m \leq 10$.

10.8° Разделите многочлен $5x^5 - x^3 + 3x - 2$ на $x - 3$ в $\mathbb{Z}/m\mathbb{Z}$ для $m = 7, 11, 13, 17, 19$.

10.9° Доказать, что если $P:Q$ над полем, то либо $P = 0$, либо $\deg Q \leq \deg P$.

10.10° Докажите, что $(x^2 - 1)$ не делится на x^3 . Перечислите все делители многочлена $(x^2 - 1)$.

10.11° Пусть $P, Q, R \in \mathbb{Q}[x]$. Докажите, что:

- a) если $P:R$ и $Q:R$, то $(P \pm Q):R$;
 b) если $P:R$ и Q — произвольный многочлен, то $PQ:R$;
 c) если $P:Q$ и $Q:R$, то $P:R$.

10.12° Верно ли, что любые многочлены обладают следующими свойствами:

- a) если P делится на R , а Q не делится на R , то $(P + Q)$ не делится на R ;
 b) если P делится на Q , а Q не делится на R , то P не делится на R ;
 c) если PQ делится на R^2 , то P делится на R или Q делится на R ?

10.13° а) Остаток от деления $A(x)$ на $x - 1$ равен 5, а на $x - 3$ равен 6. Найдите остаток от деления $A(x)$ на $(x - 1)(x - 3)$. б) Найдите остаток от деления x^{1000} на $x^2 + x - 1$.

10.14° Найдите неполные частные и остатки от деления: а) 10305 на 102; б) $x^4 + 3x^2 + 5$ на $x^2 + 2$; в) $x^4 + 3x^2 + 5$ на $x^2 + x + 2$; д) $x^4 + 3x^2 + 5$ на $x - 1$; е) $x^4 + 3x^2 + 5$ на $x - c$, где c — некоторое число.

10.15° Найдите НОД многочленов над полем: а) $x(x - 1)^3(x + 2)$ и $(x - 1)^2(x + 2)^2(x + 5)$; б) $3x^3 - 2x^2 + x + 2$ и $x^2 - x + 1$; в) $x^m - 1$ и $x^n - 1$; д) $x^m + 1$ и $x^n + 1$.

Здесь и далее под символом \mathbb{R} понимается множество вещественных чисел, которое, в частности, включает всевозможные корни различных степеней из рациональных чисел.

10.16° Разложите на неприводимые множители над \mathbb{R} и на неприводимые множители над \mathbb{Q} : а) $5x + 7$; б) $x^2 - 2$; в) $x^3 + x^2 + x + 1$; д) $x^3 - 6x^2 + 11x - 6$; е) $x^3 + 3$; ф) $x^4 + 4$.

10.17° Разложите $x^4 - 1$ на неприводимые множители над кольцом $\mathbb{Z}/m\mathbb{Z}$ при $m = 2, 3, 4, 5, 6, 7, 8, 9, 10$.

10.18° Пусть K — поле и $P, Q \in K[x]$, причем $P \neq 0$ или $Q \neq 0$. Обозначим через \mathcal{M} множество всех многочленов, представимых в виде $PU + QV$, где $U, V \in K[x]$. Пусть D — один из многочленов наименьшей степени в \mathcal{M} .

а) Докажите, что каждый многочлен из \mathcal{M} делится на любой общий делитель многочленов P и Q .

б) Докажите, что каждый многочлен из \mathcal{M} делится на D .

в) Докажите, что НОД(P, Q) всегда существует, причем определен лишь с точностью до множителя-константы.

д) Сформулируйте и докажите алгоритм Евклида для многочленов.

е) Докажите, что НОД(P, Q) делится на любой другой общий делитель этих многочленов.

ф) Докажите, что существуют $U, V \in K[x]$ такие, что $\text{НОД}(P, Q) = PU + QV$.

10.19° Найдите НОД(P, Q), если: а) $P(x) = x^2 - 4x + 3$ и $Q(x) = 2x^2 + 4x - 6$; б) $P(x) = \frac{2}{3}x^4 + x^3 - \frac{1}{4}x^2 + 1$ и $Q(x) = x^2 - x + \frac{1}{2}$.

10.20° Найдите НОД(P, Q) и его линейное выражение через $P, Q \in \mathbb{R}[x]$, если а) $P(x) = x^4 + x^3 - 3x^2 - 4x - 1$ и $Q(x) = x^3 + x^2 - x - 1$; б) $P(x) = x^4 + 2x^3 - x^2 - 4x - 2$ и $Q(x) = x^4 + x^3 - x^2 - 2x - 2$.

10.21° Пусть $P, Q, R \in \mathbb{R}[x]$, причем $R \neq 0$. Докажите, что $P \equiv Q \pmod{R}$ тогда и только тогда, когда многочлены P и Q имеют одинаковые остатки при делении на R .

10.22° Пусть $P_1, Q_1, P_2, Q_2, R, S \in \mathbb{R}[x]$, причем $R, S \neq 0$. Докажите следующие свойства сравнений:

- a) если $P_1 \equiv Q_1 \pmod{R}$ и $P_2 \equiv Q_2 \pmod{R}$, то $P_1 \pm P_2 \equiv Q_1 \pm Q_2 \pmod{R}$;
- b) если $P_1 \equiv Q_1 \pmod{R}$ и $P_2 \equiv Q_2 \pmod{R}$, то $P_1 P_2 \equiv Q_1 Q_2 \pmod{R}$;
- c) если $P_1 \equiv Q_1 \pmod{R}$, а $k \in \mathbb{N}$, то $P_1^k \equiv Q_1^k \pmod{R}$;
- d) если $P_1 \equiv Q_1 \pmod{RS}$, то $P_1 \equiv Q_1 \pmod{R}$.

10.23° Известно, что $P(x) \equiv x^2 \pmod{x^4 - x^3}$. Какой остаток даст x при делении на $(x - 1)$, x , $(x^2 - x)$ и x^3 ?

10.24° Какой остаток может дать $P(x)$ при делении на $(x^4 - x^3)$, если известно, что a) $P(x) \equiv x^2 \pmod{x^3}$ и $P(x) \equiv 1 \pmod{x - 1}$; b) $P(x) \equiv x \pmod{x^2 - x}$ и $P(x) \equiv 0 \pmod{x^2}$?

10.25° Пусть $P \in \mathbb{Z}[x]$ и $m, n \in \mathbb{Z}$. Докажите, что число $P(m) - P(n)$ делится на число $m - n$.

10.26° В выражении $(x^5 - 6x^4 + 5x^2 + 1)^{2021}$ раскрыли скобки и привели подобные. Найдите: a) коэффициент при x^0 (свободный член); b) сумму всех коэффициентов; c) сумму коэффициентов при четных степенях получившегося выражения.

10.27° Сформулируйте и докажите признак делимости многочлена: a) на $(x^3 - 1)$; b) на $(x^2 + x + 1)$.

Дополнительные упражнения

10.28 Докажите, что для любого целого положительного m множество $m\mathbb{Z} = \{mk \mid k \in \mathbb{Z}\}$ является подкольцом кольца \mathbb{Z} .

10.29 Многочлен P таков, что $P(x^n)$ делится на $x - 1$. Докажите, что $P(x^n)$ делится на $x^n - 1$.

10.30 Даны многочлены положительной степени $P(x)$ и $Q(x)$, причем выполнены тождества $P(P(x)) = Q(Q(x))$ и $P(P(P(x))) = Q(Q(Q(x)))$. Обязательно ли $P(x)$ и $Q(x)$ совпадают?

10.31 Барон Мюнхгаузен попросил задумать непостоянный многочлен $P(x)$ с целыми неотрицательными коэффициентами и сообщить ему только значения $P(2)$ и $P(P(2))$. Барон утверждает, что лишь по этим данным всегда может восстановить задуманный многочлен. Не ошибается ли барон?

10.32 Существует ли такой многочлен $P(x)$, что у него есть отрицательный коэффициент, а у каждой его степени $(P(x))^n$, где $n > 1$, все коэффициенты положительны?

10.33 Существуют ли такие многочлены $P(x)$ и $Q(x)$ из $\mathbb{R}[x]$, что каждое рациональное число r представимо в виде $r = P(k)/Q(k)$ для некоторого целого числа k ?

10.34 Коэффициенты многочленов P и Q целые. Коэффициенты их произведения делятся на 5. Докажите, что либо коэффициенты P , либо коэффициенты Q делятся на 5.

10.35 На графике многочлена из $\mathbb{Z}[x]$ отмечены две точки с целыми координатами. Докажите, что если расстояние между ними — целое число, то у них одинаковые ординаты.

10.36 Пусть $P(x)$ — многочлен с целыми коэффициентами. **a)** Докажите, что $a - b$ делит $P(a) - P(b)$ при любых различных целых числах a и b . **b)** Пусть уравнения $P(x) = 1$ и $P(x) = 3$ имеют целое решение. Может ли уравнение $P(x) = 2$ иметь два различных целых решения?

10.37 Пусть $P(x)$ — непостоянный многочлен с целыми коэффициентами. **a)** Докажите, что при любом целом числе n либо $P(n)$ делит $P(n + P(n))$, либо $P(n) = P(n + P(n)) = 0$. **b)** Могут ли все числа $P(0), P(1), P(2), \dots$ быть простыми?

10.38 Квадратный трехчлен $ax^2 + bx + c$ при всех целых x принимает целые значения. Верно ли, что среди его коэффициентов **a)** хотя бы один — целое число; **b)** все — целые числа?

10.39 Докажите, что для любого многочлена $P(x)$ степени n , принимающего при всех целых x целые значения, существуют такие целые числа b_0, b_1, \dots, b_n , что

$$P(x) = b_n \binom{x}{n} + b_{n-1} \binom{x}{n-1} + \dots + b_1 \binom{x}{1} + b_0.$$

Указание: $P(x+1) - P(x)$ тоже многочлен, принимающий целые значения при целых x , но он меньшей степени.

10.40 Многочлен $P(x)$ степени $n - 1$ принимает целые значения при n последовательных целых значениях x . Докажите, что $P(x) \in \mathbb{Q}[x]$ и $P(k) \in \mathbb{Z}$ при всех $k \in \mathbb{Z}$.

10.41 Докажите, что многочлен $(x - a_1) \dots (x - a_n) - 1$ не раскладывается в произведение двух многочленов меньшей степени из $\mathbb{Z}[x]$ при любых попарно различных целых числах a_1, \dots, a_n .

10.42 а) Пусть многочлен $P(x) = x^3 + ax^2 + bx + c$ раскладывается на линейные множители (то есть многочлены первой степени): $P(x) = (x - a_1)(x - a_2)(x - a_3)$. Докажите, что справедливы формулы Виета: $a_1 + a_2 + a_3 = -a$, $a_1a_2 + a_2a_3 + a_3a_1 = b$, $a_1a_2a_3 = -c$. б) Доказать теорему Виета.

10.43 а) Пусть $a + b + c > 0$, $ab + bc + ac > 0$, $abc > 0$. Докажите, что a, b, c положительны. б) Пусть $a + b + c < 0$, $ab + bc + ac < 0$, $abc < 0$. Какие знаки могут иметь числа a, b, c ?

10.44 а) Пусть число $c \neq 0$. Докажите, что многочлен $x^5 + ax^2 + bx + c$ не может раскладываться на пять линейных множителей. б) Та же задача для многочлена $x^5 + ax^4 + bx^3 + c$.

10.45 а) Коэффициенты многочлена $(x - a)(x - b)$ целые. Докажите, что $a^n + b^n$ целое при $n \in \mathbb{N}$. б) Найдите первые n цифр после запятой в десятичной записи числа $(\sqrt{26} + 5)^n$.

10.46 Целые числа a, b, c таковы, что числа $\frac{a}{b} + \frac{b}{c} + \frac{c}{a}$ и $\frac{a}{c} + \frac{c}{b} + \frac{a}{b}$ целые. Докажите, что $|a| = |b| = |c|$.

10.47 Пусть p — простое число. а) Докажите, что для каждого ненулевого остатка a от деления на p найдется такой остаток b от деления на p , что $ab \equiv 1 \pmod{p}$. б) Для каких a из предыдущего пункта $b = a$? с) Докажите, что сравнение $x^2 \equiv a \pmod{p}$ имеет не больше двух корней. д) Что будет с этим сравнением в случае не простого модуля m ? е) Решите сравнение $x^2 \equiv 1 \pmod{p}$. ф) [Второе доказательство критерия Вильсона] Докажите, что $(p - 1)! + 1$ делится на p тогда и только тогда, когда p — простое.

10.48 Пусть p — простое вида $4k + 1$, и пусть $x = (2k)!$. Докажите, что $x^2 \equiv -1 \pmod{p}$.

10.49* Установить коммутативность произвольного кольца, в котором каждый элемент x удовлетворяет уравнению $x^2 = x$.

10.50* а) Докажите, что если P и Q — произвольные многочлены, а неприводимый многочлен R таков, что $PQ \div R$, то либо $P \div R$, либо $Q \div R$. б) Сформулируйте и докажите основную теорему арифметики для многочленов.

Перестановки

Аннотация

В этой главе мы в основном изучаем конечные группы на примере перестановок. Кроме того, изучаются некоторые свойства групп и вводятся необходимые теоретико-множественные определения.

11.1. Теория множеств: отношения и функции

Под **упорядоченной парой** точек a и b мы понимаем запись вида $\langle a, b \rangle$, которая обладает свойством позиционного равенства:

$$\langle a, b \rangle = \langle c, d \rangle \iff (a = c) \wedge (b = d). \quad (11.1)$$

Упорядоченную пару обычно обозначают с помощью угловых скобок, либо же круглых: $\langle a, b \rangle$ или (a, b) . Второй вариант, как правило, используется при обозначении векторов и точек на координатной плоскости.

Пусть заданы два множества A и B . Под их **прямым произведением** $A \times B$ мы понимаем множество всех упорядоченных пар $\langle a, b \rangle$, где $a \in A$, $b \in B$, т. е.

$$A \times B = \{\langle a, b \rangle \mid a \in A, b \in B\}.$$

В качестве примера можно рассмотреть множество пар целых чисел на плоскости. Мы с ними уже сталкивались, когда искали все решения линейных уравнений в целых числах. Решением такого уравнения является подмножество прямого произведения $\mathbb{Z} \times \mathbb{Z}$. Обычно это пары целых чисел, лежащие на одной прямой, но в некоторых случаях это может быть пустое множество (когда решений нет) или же все множество $\mathbb{Z} \times \mathbb{Z}$ (когда любая пара целых чисел удовлетворяет уравнению).

Отношением¹ между множествами A и B называется всякое подмножество $R \subseteq A \times B$. Обычно вместо $\langle a, b \rangle \in R$ принято записывать aRb . В случае, когда $A = B$, говорят, что R есть **отношение на множестве** A .

¹ Можно встретить также термин *соответствие между* A и B .

Рассмотрим несколько простых примеров отношений.

- Отношение «родитель — дитя» (a есть родитель b).
- Отношение прямого братства: a есть родной брат/сестра b (имеется в виду, что у них общие родители).
- Отношение «предок — потомок».
- Отношение $a < b$ на целых числах.
- Отношение сравнения по модулю: $a \equiv b \pmod{m}$.

На данных примерах рассмотрим несколько основных типов отношений. Обозначим за F отношение «родитель — дитя», т. е. aFb , если a есть родитель b . Легко понять, что если выполняется aFb , то не может выполняться bFa , а также не верно и aFa . Но есть одна интересная логическая конструкция с отношением F .

Введем новое отношение S по правилу:

$$aSb \iff \exists c (cFa) \wedge (cFb), \quad (11.2)$$

т. е. a состоит в отношении S с b в том и только том случае, если у a и b существует общий родитель (хотя бы один). Проще говоря, aSb означает, что a и b есть братья/сестры/брат и сестра по одному из родителей (в английском языке для этого есть удобное слово *siblings*).

Таким образом, мы сумели одно отношение логически выразить через другое.

Обозначим за B отношение прямого братства. То есть aBb значит, что a и b имеют общего отца и мать. Ясно, что из aBb следует aSb . В символах множеств (а всякое отношение есть множество) это записывается так: $B \subseteq S$.

Заметим, что оба эти отношения обладают свойством симметрии, т. е. $aSb \iff bSa$ (это можно доказать из (11.2)) и $aBb \iff bBa$.

Говорят, что отношение R **симметрично**, если $aRb \iff bRa$ для любых a и b .

Другим хорошо известным примером симметричного отношения является равенство: $a = b \iff b = a$. Равенство обладает также свойством рефлексивности. Отношение R называется **рефлексивным**, если aRa для всех элементов $a \in A$ (понятно, что R задается на одном множестве A). В случае равенства имеем $a = a$, так что равенство рефлексивно.

Что можно сказать про отношение братства? Если посмотреть на определения отношений S и B , то становится понятно, что они рефлексивны по определению, как бы странно это ни выглядело. Действительно, мы говорим, что a и b связаны отношением братства, если у них общие родители (или хотя бы один общий родитель как в случае

S), но ведь если $a = b$, то родители у a и b тем более являются общими! Следовательно, отношения S и B также рефлексивны.

А чтобы быть ближе к действительности, нам следовало бы переопределить S следующим образом:

$$aSb \iff (a \neq b) \wedge (\exists c cFa \wedge cFb).$$

Отношения, которые мы рассматривали до сих пор, называются **бинарными**, т.к. это отношения двух индивидов. На самом деле, существует более общее определение отношения. Через $\langle x_1, \dots, x_n \rangle$ будем обозначать упорядоченный набор (или кортеж) длины n , который удовлетворяет аксиоме позиционного равенства:

$$\langle x_1, \dots, x_n \rangle = \langle y_1, \dots, y_n \rangle \iff \forall k \in \{1..n\} : (x_k = y_k).$$

Отличие кортежа из n элементов от n -элементного множества в том, что, во-первых, некоторые или все элементы кортежа могут совпадать, а во-вторых, они чувствительны к перестановкам, т.е. если поменять местами два разных элемента кортежа, то получится другой кортеж. Кортежи можно интерпретировать как символичные строки.

Пусть заданы множества A_1, \dots, A_n . Тогда их прямым произведением (в порядке указанных номеров) называется множество

$$\prod_{k=1}^n A_k = A_1 \times \dots \times A_n = \{ \langle x_1, \dots, x_n \rangle \mid (x_1 \in A_1) \wedge \dots \wedge (x_n \in A_n) \}.$$

Соответственно, подмножество $R \subseteq \prod_{k=1}^n A_k$ называется **n -арным отношением** между множествами A_1, \dots, A_n . В частности, при $n = 1$ подмножество R есть просто подмножество A_1 и называется также **унарным отношением** на A_1 , при $n = 2$ кортеж становится упорядоченной парой, а отношение — бинарным отношением. Типичным примером **тернарного** ($n = 3$) отношения является геометрическое отношение $B(a, b, c)$, выражающее тот факт, что точка b лежит между точками a и c (на одной прямой).

Перейдем к отношению «предок — потомок». Для начала мы скажем, что a и b связаны отношением P , т.е. aPb , если a есть родитель (т.е. отец или мать) b . Далее, положим

$$aTb \iff \exists n \exists \langle a_0, \dots, a_n \rangle (a_0 = a) \wedge (a_n = b) \wedge \forall k \in \{1..n\} : (a_{k-1}Pa_k).$$

Иначе говоря, если существует цепочка людей таких, что каждый предыдущий по номеру есть родитель следующего, а началом цепочки

является a , концом цепочки является b , то a и b связаны отношением T . Отношение T и есть отношение «предок — потомок».

Это отношение обладает еще одним характерным свойством: оно транзитивное. По определению, отношение R называется **транзитивным**, если всякий раз из $aRb \wedge bRc$ следует aRc . Отношение T транзитивно, т.к. если существует цепочка людей, связывающая a и b , и цепочка людей, связывающая b и c , то объединение этих цепочек дает цепочку людей, связывающую a и c .

Отношение равенства также является транзитивным. То есть равенство — это рефлексивное, симметричное и транзитивное отношение! По определению, если отношение R рефлексивно, симметрично и транзитивно, то оно называется **отношением эквивалентности**.

Сюрпризом, наверное, будет то, что отношение братства, определенное в форме (11.2), также является отношением эквивалентности. Но если подумать, оно обязано этим свойство опять-таки равенству, а именно: равенству родителей братьев/сестер.

Не всякое транзитивное отношение является отношением эквивалентности. Например, отношение «предок — потомок» таковым не является, т.к. оно не рефлексивно и не симметрично. Есть очень похожее на него арифметическое отношение $a < b$. Действительно, не верно, что $a < a$ (отношение антирефлексивно), не верно также и $b < a$ при условии $a < b$ (отношение асимметрично), но зато верно $a < c$ всякий раз, когда $a < b$ и $b < c$ (транзитивно).

К отношениям, похожим на порядок чисел, мы еще вернемся, а пока рассмотрим отношение эквивалентности.

Прежде всего вспомним про сравнимость целых чисел по модулю m :

$$a \equiv b \pmod{m} \iff \exists k \in \mathbb{Z} : a - b = km.$$

Равенство двух чисел по модулю есть отношение эквивалентности (предлагаем в этом удостовериться самостоятельно, используя ранее изученные свойства сравнения, или просто по определению).

Отношение эквивалентности разбивает множество, на котором оно задано, на непересекающиеся классы эквивалентности:

$$A = A_1 \sqcup A_2 \sqcup \dots$$

Символ \sqcup обозначает объединение множеств и одновременно говорит нам, что это — объединение попарно непересекающихся множеств. При этом внутри каждого класса сидят эквивалентные друг другу элементы. Например, всех людей можно разделить на классы эквивалентности, в каждом из которых находятся родные братья и сестры (класс эквивалентности задает пара родителей).

Докажем это строго. Пусть R — отношение эквивалентности на множестве A . Для всякого элемента $a \in A$ обозначим за $[a]_R$ его класс эквивалентности, т. е.

$$[a]_R = \{x \in A \mid aRx\}.$$

Ясно, что A есть объединение всех классов эквивалентности:

$$A = \bigcup_{a \in A} [a]_R,$$

поскольку, с одной стороны, для любого $a \in A$ $a \in [a]_R$, т. к. отношение R рефлексивно, а с другой стороны, $[a]_R \subseteq A$.

Предположим, что $[a]_R \cap [b]_R \neq \emptyset$, т. е. существует общий элемент c , принадлежащий этим классам. Но тогда по определению класса эквивалентности имеем: aRc и bRc , откуда в силу симметричности отношения R получаем, что cRb , а в силу транзитивности — aRb , откуда следует, что $b \in [a]_R$. Теперь, если $x \in [b]_R$, то bRx и aRb дают aRx , т. е. $x \in [a]_R$, откуда следует вложение $[b]_R \subseteq [a]_R$. Аналогично доказывается обратное вложение. Следовательно, $[a]_R = [b]_R$.

Итак, любые два класса эквивалентности либо не пересекаются, либо совпадают, а объединение всех классов дает все множество A . Что в общем виде записывается как

$$A = \bigsqcup_{a \in A} [a]_R.$$

В частности, отношение сравнимости по модулю m на множестве \mathbb{Z} определяет классы эквивалентности. Посмотрим, что это за классы. Класс числа 0 — это все сравнимые с нулем целые числа (по модулю m), т. е. числа вида km , где $k \in \mathbb{Z}$. Класс числа 1 — это все сравнимые с 1 числа, т. е. числа вида $1 + km$, где $k \in \mathbb{Z}$. И так далее...

Пусть K — некоторое кольцо. **Сложением по Минковскому** элемента $x \in K$ с множеством $M \subseteq K$ называется операция, результатом которой является множество

$$x + M = \{x + y \mid y \in M\},$$

а **умножением по Минковскому** элемента $x \in K$ на множество $M \subseteq K$ называется операция, результатом которой является множество

$$xM = \{xy \mid y \in M\}.$$

Отсюда легко видеть, что классом эквивалентности нуля по отношению сравнимости по модулю m является множество $m\mathbb{Z}$, классом эквивалентности 1 является множество $1 + m\mathbb{Z}$ и так далее.

Однако, $m\mathbb{Z} = m + m\mathbb{Z}$, поскольку числа вида km и $m + km = (1 + k)m$ принадлежат по определению одному и тому же классу $m\mathbb{Z}$. Аналогично, $1 + m\mathbb{Z} = (m + 1) + m\mathbb{Z}$ и так далее.

И вообще, если $k \equiv s \pmod{m}$, то $k + m\mathbb{Z} = s + m\mathbb{Z}$. Таким образом, множество \mathbb{Z} разбивается на m классов эквивалентности:

$$\mathbb{Z} = m\mathbb{Z} \sqcup (1 + m\mathbb{Z}) \sqcup \dots \sqcup ((m - 1) + m\mathbb{Z}).$$

Множество, состоящее из элементов–классов эквивалентности некоторого множества A по отношению эквивалентности R , заданному на этом множестве, называется **фактормножеством** множества A по данному отношению R и обозначается за A/R , т. е.

$$A/R = \{[a]_R \mid a \in A\}.$$

В частности, если R — это отношение сравнимости по модулю m на целых числах, то

$$\mathbb{Z}/R = \{m\mathbb{Z}, 1 + m\mathbb{Z}, \dots, (m - 1)\mathbb{Z}\}.$$

Заметим, что отношение сравнимости по модулю m еще можно определить и так:

$$a \equiv b \pmod{m} \iff a - b \in m\mathbb{Z},$$

т. е. множество $m\mathbb{Z}$ является определяющим классом для данного отношения (остальные классы получаются его сдвигом на некоторое целое число).

Произвольное множество $M \subseteq \mathbb{Z}$ не может таким способом определить отношение эквивалентности. Действительно, во-первых, чтобы полученное отношение было рефлексивным, нужно, чтобы $0 \in M$ ($a - a \in M$). Чтобы отношение было симметричным, нужно также, чтобы $-a \in M$ при $a \in M$. А чтобы отношение было транзитивным, необходимо, чтобы $a + b \in M$ при $a, b \in M$.

Иначе говоря, множество $M \subseteq \mathbb{Z}$ задает отношение эквивалентности R по правилу

$$aRb \iff a - b \in M,$$

если $0 \in M$, а также M замкнуто относительно операции сложения и взятия противоположного элемента. То есть M должно быть подгруппой \mathbb{Z} по сложению.

В тех случаях, когда фактормножество множества A задается при помощи некоторого специального подмножества M «с хорошими свойствами», оно записывается как A/M , т. е. в обозначении подчеркивается, что факторизация произведена по отношению эквивалентности, которое задано с помощью множества M .

В частности, множество классов сравнимых по модулю m целых чисел записывается так: $\mathbb{Z}/m\mathbb{Z}$.

Введем понятие функции. Пусть у нас имеется отношение F между множествами X и Y . Отношение F называется

- **всюду значным**, если для каждого $y \in Y$ найдется $x \in X$ такое, что xFu ;
- **всюду определенным**, если для каждого $x \in X$ найдется $y \in Y$ такое, что xFu ;
- **однозначным**, если всякий раз из одновременного выполнения xFu и xFu' следует, что $y = y'$, т. е. каждому x соответствует не более одного y ;
- **обратно однозначным**, если всякий раз из одновременного выполнения xFu и $x'Fu$ следует, что $x = x'$, т. е. каждому y соответствует не более одного x ;
- **функцией из X в Y** , если оно всюду определенное и однозначное;
- **сюръекцией из X на Y** , если это всюду значная функция;
- **инъекцией из X в Y** , если это обратно однозначная функция;
- **биекцией из X в Y** , если это инъекция и сюръекция одновременно.

Взаимосвязь перечисленных терминов схематично представлена на рис. 11.1.

Обычно функция из X в Y обозначается $F : X \rightarrow Y$, а если $\langle x, y \rangle \in F$, то пишут $y = F(x)$. Можно также явно определить, что такое $F(x)$, следующим способом:

$$F(x) = \bigcup \{y \mid \langle x, y \rangle \in F\}, \quad x \in X.$$

Поясним. Поскольку функция является однозначным отношением, множество $\{y \mid \langle x, y \rangle \in F\}$ состоит ровно из одной точки y , соответствующей точке x в данном отношении, т. е. оно равно *синглетону* $\{y\}$. А его объединение (т. е. объединение всех его элементов) — это и есть сам элемент y .

Если имеет место равенство $y = F(x)$, то x называют *аргументом*, а y — *значением функции F* , соответствующим данному аргументу x .

Для обозначения биекции часто используется символ $F : X \leftrightarrow Y$.

Часто используется термин *частичная функция*, означающий, что функция может быть задана не на всем множестве X . Например, функция, которая определена только на простых числах, является частичной функцией, заданной на \mathbb{N} . Такая терминология бывает полезной, если используется одно базовое множество, а функции на нем задаются формулами, которые могут быть определены не для всех точек данного

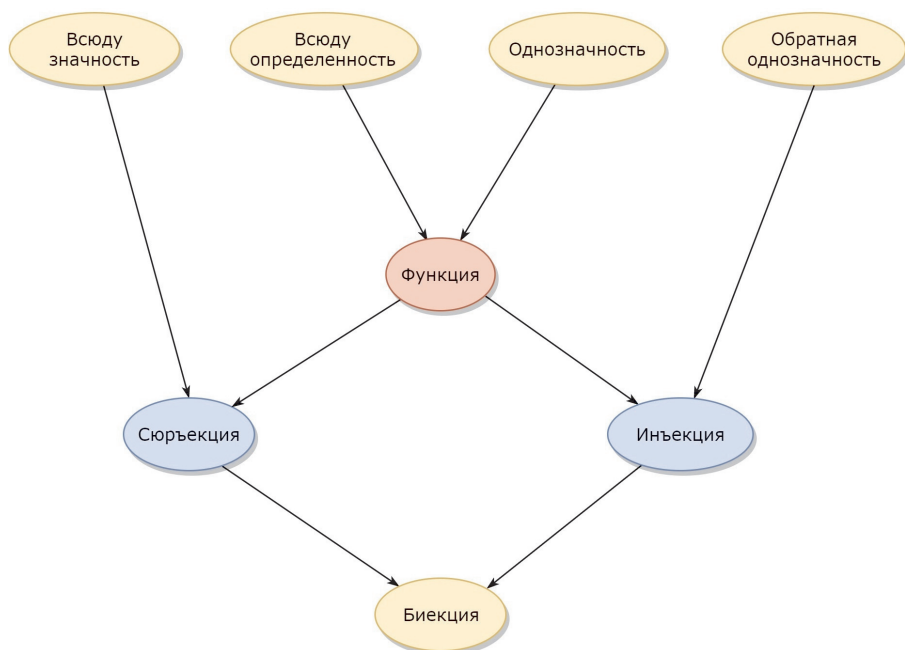


Рис. 11.1. Виды функций.

множества. Например, на \mathbb{Q} можно рассматривать частичные функции вида $F(x) = 1/(x - r)$, где $r \in \mathbb{Q}$. Эти функции будут определены всюду на \mathbb{Q} , кроме точки r .

Для частичных функций содержательный смысл имеет термин *область определения*, т.е. такое подмножество X , где данная функция определена. В случае функции (не частичной) $F : X \rightarrow Y$ ее областью определения называется само множество X . Область определения функции F принято обозначать $\text{dom}(F)$.

Обратным к R отношением называется отношение

$$R^{-1} = \{\langle y, x \rangle \mid xRy\}.$$

Если обратное отношение есть функция, то она называется **обратной функцией**. Легко видеть, что в том случае, когда существует обратная функция к функции F , выполняются равенства $F(F^{-1}(y)) = y$ и $F^{-1}(F(x)) = x$, если только $F^{-1}(y)$ определено.

Последняя оговорка связана с тем, что обратная функция является частичной функцией из Y в X , т.е. она может быть не определена всюду на Y , а значит, $F^{-1}(y)$ может не существовать.

Областью значений функции F называется подмножество Y , обозна-

чаемое $\text{ran}(F)$, для элементов которого F определена, т. е.

$$\text{ran}(F) = \{y \in Y \mid \exists x \in X : y = F(x)\}.$$

Образом множества $A \subseteq X$ относительно функции F называется множество

$$F[A] = \{y \mid \exists x \in A y = F(x)\},$$

т. е. образ множества — это результат действия на него функции F .

Соответственно, **прообразом** множества $B \subseteq Y$ называется множество

$$F^{-1}[B] = \{x \mid F(x) \in B\},$$

т. е. прообраз множества — это такое максимальное подмножество множества X , образ которого лежит в данном множестве B .

Заметим, что для существования прообраза не требуется существования обратной функции. Кроме того, заметим, что $F[F^{-1}[B]]$ не всегда равен B .

Внимание! Мы намеренно отличаем символы $F(x)$ и $F[A]$, а также $F^{-1}(y)$ и $F^{-1}[B]$, чтобы подчеркнуть, что в случае круглых скобок речь идет о точках (аргумент и значение), а в случае квадратных — о подмножествах.

Итак, функция — это есть однозначное соответствие элементов одного множества элементам другого (или того же самого). Функции обычно задаются формулами, предписывающими некоторой алгоритм вычисления y через x . Но иногда такие формулы не указаны явно или же их указать вовсе невозможно, хотя существование функции строго доказывается (такие теоремы называют теоремами существования).

11.2. Обозначения перестановок

Рассмотрим множество $X_n = \{x_1, \dots, x_n\}$, состоящее из n попарно различных элементов, и все возможные биекции множества X_n в себя. Обозначим множество всех биекций за

$$\mathbf{S}(X_n) = \{f \mid f : X_n \leftrightarrow X_n\}.$$

Проще говоря, $\mathbf{S}(X_n)$ — это множество всех возможных **перестановок** элементов множества X_n .

Для функций, заданных на множестве, естественной операцией является операция композиции, т. е. последовательное применение функций. Обычно композиция функций записывается символом \circ или вовсе пропускается. Мы будем пользоваться первым вариантом, так что:

$$(f \circ g)(x) = f(g(x))$$

по определению.

Итак, мы имеем множество биекций (перестановок) с операцией композиции. Свойства композиции биекций таковы:

- 1) композиция биекций есть снова биекция;
- 2) $f \circ (g \circ h) = (f \circ g) \circ h$, поскольку это последовательное вычисление $f(g(h(x)))$ при любом $x \in X_n$;
- 3) существует функция, которая ничего не меняет: $\text{id}(x) = x$, она также является биекцией;
- 4) для всякой биекции существует обратная функция, которая также является биекцией: $f \circ f^{-1} = f^{-1} \circ f = \text{id}$.

Таким образом, множество $\mathbf{S}(X_n)$ (и вообще, $\mathbf{S}(X)$ для любого конечного или бесконечного множества X) с операцией композиции является группой. Эта группа называется **группой перестановок** множества X_n , а ее элементы–биекции называются **перестановками** (или подстановками).

Пусть $\langle G_1, \circ \rangle$ и $\langle G_2, \star \rangle$ – некоторые группы.

Говорят, что функция $f : G_1 \rightarrow G_2$ является **гомоморфизмом** групп $\langle G_1, \circ \rangle$ и $\langle G_2, \star \rangle$, если она сохраняет групповую операцию:

$$\forall g, g' \in G_1 : f(g \circ g') = f(g) \star f(g').$$

Если функция $f : G_1 \rightarrow G_2$ является биекцией и гомоморфизмом, то она называется **изоморфизмом** групп $\langle G_1, \circ \rangle$ и $\langle G_2, \star \rangle$. При этом группы $\langle G_1, \circ \rangle$ и $\langle G_2, \star \rangle$ называются **изоморфными**. Тот факт, что группы $\langle G_1, \circ \rangle$ и $\langle G_2, \star \rangle$ изоморфны, обозначается так:

$$\langle G_1, \circ \rangle \cong \langle G_2, \star \rangle,$$

либо в упрощенной форме пишут просто $G_1 \cong G_2$.

Например, группа вращений правильного n -угольника изоморфна группе вычетов по модулю n с операцией сложения.

Понятие изоморфизма (не только групп, но и более сложных математических структур) является одним из фундаментальных понятий математики. Это – аналог равенства множеств, но на более высоком уровне, поскольку отвечает за тождество операций и отношений, но пренебрегает тождеством элементов. В Алгебре изоморфные структуры часто просто считаются равными, хотя природа их элементов может кардинально отличаться.

Один простой пример группы перестановок мы уже встречали, когда рассматривали все возможные симметрии правильного треугольника. Именно в этом случае все перестановки вершин треугольника соответствовали движениям треугольника, и только они.

Сколько всего перестановок в группе $\mathbf{S}(X_n)$?

Для ответа на этот вопрос посмотрим, сколько существует вариантов перехода одних элементов в другие. Очевидно, что первый элемент может перейти в любой, в том числе в самого себя, так что для него существует n вариантов. Второй элемент может перейти куда угодно, кроме того места, которое занял первый, так что для него существует $n - 1$ вариант. Третьему остается $n - 2$ варианта. И т.д. Последнему элементу остается выбор из одного оставшегося места. Таким образом, всего вариантов перестановок на n элементах существует ровно

$$n(n - 1)(n - 2) \dots 2 \cdot 1 = n!.$$

Иначе говоря, группа $\mathbf{S}(X_n)$ имеет порядок $n!$.

Группа перестановок на трех элементах имеет порядок $3! = 6$, что соответствует количеству симметрий правильного треугольника. Однако уже для квадрата число перестановок равно 24, в то время как число всех движений составляет всего лишь 8, а для ромба так и вовсе 4. Вообще, как мы помним, количество движений правильного n -угольника равно $2n$. С ростом n это число становится во много раз меньше, чем $n!$ (а точнее, в $3 \cdot 4 \dots (n - 1) = (n - 1)!/2$ раз).

Чтобы не заострять внимание на происхождении элементов множества X_n , обычно они обозначаются числами-символами от 1 до n , а соответствующая группа биекций — $\mathbf{S}_n = \mathbf{S}(\{1, 2, \dots, n\})$.

Заметим, что группы \mathbf{S}_n и $\mathbf{S}(X_n)$ изоморфны, т.е. между ними существует биекция, сохраняющая операцию композиции.

Для этого достаточно взять произвольную биекцию $H : \{1, 2, \dots, n\} \leftrightarrow X_n$ и всякой биекции $f \in \mathbf{S}_n$ поставить в соответствие функцию \bar{f} , заданную для всякого $x_i \in X_n$ по правилу

$$\bar{f}(x_i) = H(f(H^{-1}(x_i))) \quad (\text{т.е. } \bar{f} = H \circ f \circ H^{-1}).$$

Изобразим то же самое при помощи диаграммы:

$$\begin{array}{ccc} \{1, \dots, n\} & \xrightarrow{f} & \{1, \dots, n\} \\ \uparrow H^{-1} & & \downarrow H \\ X_n & \xrightarrow{\bar{f} = H \circ f \circ H^{-1}} & X_n \end{array}$$

Нетрудно видеть, что $\bar{f} : X_n \rightarrow X_n$ и что это биекция, поскольку \bar{f} является композицией трех биекций. Таким образом, $\bar{f} \in \mathbf{S}(X_n)$.

Покажем, что соответствие $f \mapsto \bar{f}$ является биекцией между \mathbf{S}_n и $\mathbf{S}(X_n)$, сохраняющей операцию композиции. Пусть $f, g \in \mathbf{S}_n$ и $f \neq g$. Это значит, что в некоторой точке j получим $f(j) \neq g(j)$. Но тогда и

$\bar{f}(H(j)) \neq \bar{g}(H(j))$, поскольку $\bar{f}(H(j)) = H(f(j))$ и $\bar{g}(H(j)) = H(g(j))$, а функция H в различных точках принимает различные значения, т. к. является биекцией. То есть соответствие $f \mapsto \bar{f}$ разные функции переводит в разные функции. Кроме того, какова бы ни была функция $g \in \mathbf{S}(X_n)$, существует $f \in \mathbf{S}_n$ такая, что $g = \bar{f}$. Действительно, пусть $f = H^{-1} \circ g \circ H$, тогда

$$\bar{f} = H \circ f \circ H^{-1} = H \circ H^{-1} \circ g \circ H \circ H^{-1} = g.$$

Итак, соответствие $f \mapsto \bar{f}$ биективно.

Осталось показать, что оно сохраняет композицию:

$$\overline{f \circ g} = H \circ (f \circ g) \circ H^{-1} = (H \circ f \circ H^{-1}) \circ (H \circ g \circ H^{-1}) = \bar{f} \circ \bar{g}.$$

Поэтому в дальнейшем, говоря о группе перестановок, мы будем иметь в виду группу \mathbf{S}_n , заданную на множестве $\{1, \dots, n\}$, не акцентируя внимание на природе переставляемых элементов. В связи с этим, для удобства дальнейшего изложения положим $X_n = \{1, \dots, n\}$.

Теория групп в XIX в. начиналась именно с изучения групп перестановок, и лишь позже понятие группы было обобщено Артуром Кэли. Он же сделал первый важный шаг на пути классификации групп.

Теорема 11.1 (Кэли). *Любая конечная группа порядка n изоморфна некоторой подгруппе группы \mathbf{S}_n .*

Доказательство. Пусть $G = \{g_1, \dots, g_n\}$. Для каждого элемента g_k определим функцию $h_k : G \rightarrow G$ по правилу:

$$h_k(g) = g_k g, \quad g \in G. \quad (11.3)$$

Это — т.н. «правые» биекции, т. к. домножение на аргумент происходит справа. Аналогично можно задать «левые» биекции. Легко проверить, что это и в самом деле биекции. Действительно, если $g \neq g'$, то $h_k(g) \neq h_k(g')$, иначе мы бы получили равенство $g_k g = g_k g'$, и после сокращения на g_k слева получилось бы, что $g = g'$. Таким образом, h_k — инъекция. Кроме того, для любого $g \in G$ имеем следующее:

$$h_k(g_k^{-1} g) = g_k g_k^{-1} g = g,$$

т. е. h_k — сюръекция и, следовательно, она есть биекция.

Покажем, что множество биекций $H = \{h_1, \dots, h_n\}$ образует группу с операцией композиции. Действительно, для любого $g \in G$ имеем $(h_k \circ h_j)(g) = (g_k g_j)g$. А так как $g_k g_j \in G$, то $h_k \circ h_j \in H$. Ассоциативность композиции функций выполняется автоматически. Единицей группы H является функция h_k при таком k , что $g_k = \mathbf{e}$, т. е. нейтральный элемент

группы G . Действительно, $h_k(g) = eg = g$ для всех $g \in G$. Наконец, обратной функцией к h_k является функция $h_l(g) = g_l g$, где $g_l = g_k^{-1}$.

Группа $\mathbf{S}(G)$, как мы уже установили, изоморфна группе \mathbf{S}_n , а значит, группа H изоморфна некоторой подгруппе \mathbf{S}_n .

Осталось заметить, что $H \cong G$. Это следует из определения элементов H . Соответствие $g_k \mapsto h_k$ по формуле (11.3) устанавливает изоморфизм этих групп, т.к. произведению $g_k g_j$ соответствует композиция $h_k \circ h_j$. \square

В группе \mathbf{S}_n , как и в любой другой конечной группе, можно построить циклическую подгруппу, отправляясь от произвольно взятого элемента, т.е. биекции на X_n . Например, пусть $s \in \mathbf{S}_n$, тогда можно рассмотреть циклическую подгруппу $G(s) = \{s, s^2, s^3, \dots\}$, где под степенью понимается многократная композиция биекции s с самой собою. Ясно, что эта подгруппа не может быть бесконечной, т.к. она входит в конечную группу, поэтому при некотором k имеем $s^k = \text{id}$.

Рассмотрим некоторую перестановку $s \in \mathbf{S}_n$. Ее можно записать в виде таблицы аргумент–значение:

$$s = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ s_1 & s_2 & \dots & s_{n-1} & s_n \end{pmatrix},$$

где $s_i = s(i)$. При этом $\{1, 2, \dots, n\} = \{s_1, s_2, \dots, s_n\}$ в полном соответствии с определением равенства множеств.

Возьмем теперь символ 1 и начнем «раскрывать» его:

$$1 \mapsto s(1) \mapsto s(s(1)) \mapsto \dots \mapsto s^k(1).$$

Мы получим то, что называется **орбитой** элемента 1 при действии группы $G(s)$ на множестве X_n . Действительно, все элементы данной цепочки составляют множество $G(s)1 = \{g(1) \mid g \in G(s)\}$.

Заметим, что все степени перестановки s являются перестановками, т.е. элементами конечного множества \mathbf{S}_n . А значит, с ростом k они начнут повторяться, т.е. найдутся такие k и j , $k > j$, что $s^k = s^j$. Умножая данное равенство на s^{-1} (обратную к s функцию) достаточное количество раз, мы придем к равенству $s^{k-j} = \text{id}$. То есть всегда существует такое k , что $s^k = \text{id}$.

В этом случае мы получим равенство $s^k(1) = 1$, и орбита элемента 1 превратится в **цикл**:

$$(1 \ s(1) \ s(s(1)) \ \dots \ s^{k-1}(1))$$

(единицу в конце мы не пишем, подразумевая, что последний элемент цикла переходит в первый).

Отметим, что длина цикла в данном случае может оказаться меньше k . Например, если перестановка s переводит 1 в 2 и 2 в 1, но при этом 3 — в 4, 4 — в 5, и 5 — в 3. В таком случае группа $G(s)$ содержит как минимум 6 элементов, в то время как цикл, начинающийся с 1, — всего лишь два.

Действие группы $G(s)$ на множестве X_n позволяет разбить это множество на несколько попарно непересекающихся орбит (или циклов). Покажем это.

Предположим, что орбиты $G(s)1$ и $G(s)2$ имеют общий элемент j . Тогда при некоторых натуральных $a, b \leq k$ будем иметь $s^a(1) = j = s^b(2)$. Далее, любую степень $s^t(2)$ можно записать в виде:

$$s^t(2) = (s^{t-b \bmod k} \circ s^b)(2) = (s^{t-b \bmod k} \circ s^a)(1) = s^{t-b+a \bmod k}(1),$$

где мы использовали вычисления степени по модулю k , т. к. $s^k = \text{id}$ и в случае $t < b$ можно добавить k к степени, чтобы выйти на положительную степень.

Таким образом, каждый элемент орбиты $G(s)2$ является элементом орбиты $G(s)1$. Симметрично рассуждая, получаем и обратное вложение, так что $G(s)1 = G(s)2$. Аналогично доказывается равенство для любой пары орбит $G(s)j$ и $G(s)l$ в предположении, что у них есть общий элемент. Следовательно, орбиты либо не пересекаются, либо совпадают.

То, что объединение всех орбит равно множеству X_n , следует из простого факта, что орбита $G(s)j$ всегда содержит элемент j . Итак,

$$X_n = \bigsqcup_{j=1}^n G(s)j.$$

Отсюда мы получаем представление самой перестановки s как набора независимых циклов (орбит). Поэтому перестановки принято записывать в виде последовательности циклов. Например, пусть

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix}.$$

В этой перестановке мы наблюдаем два цикла: (1243) и тривиальный (5). Тогда

$$s = (1243)(5),$$

причем тривиальные циклы принято пропускать в такой «циклической» записи, т. к. они однозначно восстанавливаются по всем остальным циклам и по параметру n (в нашем случае $n = 5$).

Рассмотрим более сложный пример:

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 3 & 1 & 7 & 5 & 6 \end{pmatrix} = (124)(3)(576) = (124)(576).$$

Предположим теперь, что у нас имеется три перестановки:

$$s_1 = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 1 \end{pmatrix}; \quad s_2 = \begin{pmatrix} 5 & 6 & 7 \\ 7 & 5 & 6 \end{pmatrix}; \quad s_3 = \text{id}.$$

Тогда исходная перестановка s получается как последовательное применение этих новых перестановок:

$$s = s_1 s_2 s_3,$$

причем порядок перестановок в этой композиции неважен, т. к. две из них «работают» на разных орбитах, а третья тождественна и коммутирует с любой перестановкой.

Таким образом, каждую перестановку из S_n можно единственным образом (с точностью до порядка) представить как композицию циклов, и, таким образом, запись перестановки в виде набора ее циклов является не только удобным соглашением, но еще и функционально полезной.

11.3. Пара слов о конечных группах

Пусть G — конечная группа. Приведем некоторые общие свойства групп.

В группе существует только одна единица. Действительно, если их две e и e' , то в силу их же свойств получим

$$e' = ee' = e$$

(в первом равенстве мы рассматривали e как единицу, а во втором e').

Обратный элемент для каждого $a \in G$ определен однозначно. Предположим, что для элемента a нашлось два обратных элемента b, c , т. е. $ab = ba = e$ и $ac = ca = e$. Тогда

$$b = be = b(ac) = (ba)c = ec = c.$$

Степень элемента $\underbrace{a \cdots a}_{k \text{ раз}}$ корректно определяется в силу закона ассоциативности и обозначается a^k , где $k \in \mathbb{N}$. Кроме того, по определению, $a^0 = e$.

Отрицательная степень элемента по определению: $a^{-k} = (a^{-1})^k$, $k \in \mathbb{N}$.

Операции со степенями:

$$(a^k)(a^m) = a^{k+m},$$

где $k, m \in \mathbb{Z}$. Если k и m одного знака, то это очевидно, а если разного, то пусть $k > 0, m < 0$, тогда

$$(a^k)(a^m) = \underbrace{a \cdots a}_{k \text{ раз}} \underbrace{a^{-1} \cdots a^{-1}}_{|m| \text{ раз}}.$$

Пользуясь ассоциативностью, начинаем сворачивать пары aa^{-1} , стоящие в середине, заменяя их на e , а затем выбрасывая e . В итоге либо ничего не останется (когда $k = -m$), либо останутся только a в количестве $k+m$ (если $k > -m$), либо останутся только a^{-1} в количестве $-m-k$ (когда $k < -m$). В любом случае это записывается как $(a^{-1})^{-m-k} = a^{m+k}$ по определению.

Как уже отмечалось, в конечной группе каждый элемент в некоторой конечной степени обращается в e , поэтому корректно определение: число $\min\{k \mid k > 0 \wedge a^k = e\}$ называется **порядком элемента** a .

Очевидно также, что всякую отрицательную степень элемента в конечной группе можно записать как положительную, поскольку

$$a^k = a^{k \bmod p},$$

где p — порядок элемента a .

Подмножество $T \subseteq G$, все возможные произведения степеней элементов которого, т.е. выражения вида $t_1^{k_1} \cdots t_m^{k_m}$, где $t_j \in T, k_j \in \mathbb{N}$, образуют всю группу G , называется **системой образующих** или **порождающим множеством** группы G . При этом пишут $G = \langle T \rangle$ или $G = \langle t_1, \dots, t_m \rangle$. Элементы системы образующих называются **образующими** группы.

Если система образующих состоит из одного нетривиального элемента, то группа называется **циклической**. При этом ее можно записать так: $G = \langle g \rangle$, где $T = \{g\}$. Иначе говоря, циклическая группа состоит из степеней одного своего элемента.

Например, группа $\mathbb{Z}/n\mathbb{Z}$ вычетов по модулю n с операцией сложения является циклической: $\mathbb{Z}/n\mathbb{Z} = \langle 1 \rangle$, поскольку все ее элементы — это конечные суммы единиц (от одной до n штук). Группа вращений правильного n -угольника является циклической, где образующим элементом является поворот на угол $2\pi/n$. Группа $(\mathbb{Z}/5\mathbb{Z})^*$, состоящая из элементов $1, 2, 3, 4$, с операцией умножения по модулю 5 является циклической, $(\mathbb{Z}/5\mathbb{Z})^* = \langle 2 \rangle = \langle 3 \rangle$.

Циклические группы коммутативны. Действительно, любые два элемента такой группы — это некоторые степени образующего элемента, поэтому $(a^k)(a^m) = a^{k+m} = a^{m+k} = (a^m)(a^k)$. Здесь коммутативность наследуется от сложения в группе \mathbb{Z} , где $k + m = m + k$.

Напомним, что подмножество $H \subseteq G$ называется **подгруппой** группы G , если H само является группой с той же операцией, которая определена в G . Например, $\{0, 2\}$ образует подгруппу группы \mathbb{Z}_4 . Тривиальная подгруппа $\{e\}$ является подгруппой любой группы.

Операция Минковского умножения элемента группы на ее подмножество порождает «смежные классы»:

$$gH = \{gh \mid h \in H\}, \quad Hg = \{hg \mid h \in H\},$$

где gH называется левым, а Hg — правым **смежным классом** (или классом смежности), порожденным элементом $g \in G$.

Лемма 11.1. *Левые (правые) классы смежности по данной подгруппе H содержат одинаковое количество элементов.*

Для доказательства этого факта построим взаимно-однозначное соответствие между произвольным смежным классом и множеством H .

Докажем, что $f : H \rightarrow gH$, определенная по правилу $f(h) = gh$, является биекцией. Сюръективность очевидна по определению. Покажем, что f — инъекция.

Действительно, пусть $h_1 \neq h_2$, где $h_1, h_2 \in H$. Предположим, что $f(h_1) = f(h_2)$, то есть $gh_1 = gh_2$. Домножая слева на g^{-1} , находим, что $h_1 = h_2$. Противоречие. Следовательно, функция f различные элементы переводит в различные.

Аналогично рассуждаем в случае, когда $f(h) = hg$. Таким образом, множества gH , Hg и H для любой подгруппы $H \subseteq G$ и любого элемента $g \in G$ имеют одинаковое количество элементов. \square

Лемма 11.2. *Классы смежности подгруппы H либо совпадают, либо не пересекаются, а их объединение равно G .*

Доказательство. Пусть классы g_1H и g_2H имеют общий элемент g . Этот элемент будет иметь два представления: $g = g_1h_1 = g_2h_2$, где $h_1, h_2 \in H$, откуда $g_1 = g_2h_2h_1^{-1}$. Возьмем любой элемент g_1h из первого класса, тогда

$$g_1h = g_2h_2h_1^{-1}h,$$

где $h_2h_1^{-1}h \in H$, т. к. H — подгруппа. Следовательно, $g_1h \in g_2H$, и $g_1H \subseteq g_2H$. Аналогично рассуждая, находим, что $g_2H \subseteq g_1H$, т. е. $g_1H = g_2H$.

Тот факт, что любой элемент G находится в каком-то классе смежности, следует из того, что $e \in H$, так что для любого $g \in G$ имеем $g \in gH$. И аналогично для правых классов. \square

Иными словами, классы смежности образуют разбиение множества G . Такую ситуацию мы уже наблюдали в связи с подгруппами $m\mathbb{Z}$ и их сдвигами внутри \mathbb{Z} и получали там m классов смежности.

Итак, множество G есть объединение непересекающихся классов одного размера, причем размер классов равен порядку подгруппы H . Следовательно, *порядок подгруппы делит порядок группы*. Это утверждение называется **теоремой Лагранжа**.

Индексом подгруппы H в группе G называется количество (в общем случае — мощность множества) классов смежности подгруппы H в группе G , которое обозначается символом $|G : H|$. Теорема Лагранжа позволяет вычислить индекс подгруппы по формуле $|G|/|H|$ для конечной группы G .

Подгруппа H группы G называется **нормальной**, если для любого $g \in G$ верно равенство $gH = Hg$, т. е. левые и правые классы смежности не различаются. Обозначение: $H \triangleleft G$.

В коммутативных группах любая подгруппа будет нормальной. В частности, $m\mathbb{Z}$ — нормальная подгруппа в \mathbb{Z} .

Следующее вложение является как критерием нормальности подгруппы, так и зачастую используется в качестве ее определения:

$$\forall g \in G : g^{-1}Hg \subseteq H \quad (\forall g \in G : gHg^{-1} \subseteq H). \quad (11.4)$$

Утверждение в скобках равносильно первому, поскольку g — произвольный элемент G , а значит вместо него можно подставить ему обратный. Отсюда же следует и то, что данное неравенство можно на самом деле заменить равенством. Умножим неравенство $g^{-1}Hg \subseteq H$ на g слева, а затем результат — на g^{-1} справа, и получим, что $H \subseteq gHg^{-1}$, поскольку для любой функции верно, что если $A \subseteq B$, то $f[A] \subseteq f[B]$. В частности, это верно для функции $f(x) = gxg^{-1}$.

Тогда получается, что если выполнено условие (11.4), то для любого $g \in G$ верно $H \subseteq gHg^{-1}$, что вместе с равносильным условием $gHg^{-1} \subseteq H$ дает равенство $gHg^{-1} = H$.

Проверим, что условие (11.4) равносильно нашему определению нормальной подгруппы, т. е. что для любого $g \in G$ имеем равенство $Hg = gH$. Для этого достаточно рассмотреть функцию $f(x) = g^{-1}x$, которая множество gH взаимно однозначно переводит в множество H , а множество Hg — в множество $g^{-1}Hg$, откуда следует, что $Hg = gH$ тогда и только тогда, когда $H = g^{-1}Hg$.

Классы смежности нормальной подгруппы можно умножать так же, как сами элементы группы G . Определим произведение подмножеств группы по Минковскому: если $A, B \subseteq G$, то положим

$$AB = \{ab \mid a \in A \wedge b \in B\}.$$

Тогда при $H \triangleleft G$ имеет место тождество для любых $g_1, g_2 \in G$:

$$(g_1H)(g_2H) = (g_1g_2)H. \quad (11.5)$$

Действительно,

$$(g_1H)(g_2H) = g_1(g_2Hg_2^{-1})(g_2H) = (g_1g_2)HH = (g_1g_2)H.$$

Такое свойство умножения классов смежности по нормальной подгруппе позволяет задать групповую операцию на множестве всех классов смежности. Действительно, как мы только что установили, произведение классов смежности, заданное по формуле (11.5), не выводит из множества классов смежности. Далее, ассоциативность операции напрямую наследуется от такого же свойства операции в группе G :

$$\begin{aligned} ((g_1H)(g_2H))(g_3H) &= (g_1g_2)H(g_3H) = ((g_1g_2)g_3)H = \\ &= (g_1(g_2g_3))H = (g_1H)(g_2g_3)H = (g_1H)((g_2H)(g_3H)). \end{aligned}$$

Нейтральным элементом является сама подгруппа H , т. к. для любого $g \in G$ имеем $(gH)H = g(HH) = gH$ и $H(gH) = H(Hg) = Hg = gH$. Наконец, обратный элемент к gH — это $g^{-1}H$.

Таким образом, множество классов смежности группы G по нормальной подгруппе H образует группу с операцией (11.5). Такая группа обозначается

$$G/H = \{gH \mid g \in G\}$$

и называется **факторгруппой** группы G по нормальной подгруппе H . Опять же, мы уже сталкивались с примером факторгруппы $\mathbb{Z}/m\mathbb{Z}$ при изучении группы вычетов (см. раздел 9.1).

Факторизацию группы можно воспринимать как делимость групп, и в этом смысле группы становятся подобны числам. Есть **простые группы** — они ни на что не делятся, т. е. не имеют собственных нормальных подгрупп, кроме тривиальной (делятся только на себя и на подгруппу $\{e\}$), а есть составные — они делятся на нетривиальные нормальные подгруппы.

Естественно ввести и умножение групп. Пусть даны две группы G_1 и G_2 с операциями \circ и \star , соответственно. Тогда на прямом произведении $G_1 \times G_2$ определим операцию умножения по правилу

$$\langle g_1, g_2 \rangle \langle g'_1, g'_2 \rangle = \langle g_1 \circ g'_1, g_2 \star g'_2 \rangle,$$

т. е. будем покомпонентно перемножать все пары элементов прямого произведения. Легко проверить, что это — групповая операция, т. е. она ассоциативна, имеет единицу, а для каждой пары есть обратная. Все эти свойства наследуются от исходных групп напрямую. Кроме того, если исходные группы коммутативны, то и произведение групп будет коммутативной группой. Такая конструкция называется **прямым произведением групп** G_1 и G_2 .

Если в исходных группах операция интерпретируется как сложение или группа является коммутативной (например, если речь идет об операции сложения в кольце), то прямое произведение называют **прямой суммой групп**. Обозначение: $G_1 \oplus G_2$.

Рассмотрим простой пример прямой суммы групп: $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. Вот ее таблица умножения:

$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	00	01	10	11
00	00	01	10	11
01	01	00	11	10
10	10	11	00	01
11	11	10	01	00

где вместо пары $\langle i, j \rangle$ мы просто пишем ij для краткости.

Видим, что эта группа абелева, но не циклическая. В этой группе есть три подгруппы: $\{00, 01\}$, $\{00, 10\}$ и $\{00, 11\}$.

Если сравнить ее с группой \mathbb{Z}_4 по сложению, то мы увидим существенную разницу. Во-первых, в \mathbb{Z}_4 только одна подгруппа $\{0, 2\}$, а во-вторых, \mathbb{Z}_4 является циклической группой.

Этот пример показывает нам, что порядок группы не определяют однозначно ее структуру (как нам того бы хотелось, памятуя об основной теореме арифметики).

Однако нам уже хорошо знакома группа, которая имеет такую же таблицу умножения, как и $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. Это все та же группа симметрий ромба, т. е. четверная группа Клейна. Чуть позже мы к ней вернемся.

Подгруппа группы G позволяет «разрезать» группу G на непересекающиеся классы одинакового размера, которые в случае нормальности подгруппы образуют некоторую новую группу, абстрактную по отношению к группе G . Существуют и другие способы «разрезания» группы с помощью чисто групповых приемов. Один из таких способов мы сейчас изучим.

Пусть $g, x \in G$, где G – группа. Элемент $x^{-1}gx$ называется **сопряженным** к элементу g при помощи x . Если сопряженный элемент обозначить как g^x , то у нас получаются знакомые свойства:

- 1) $g^{xy} = (g^x)^y$;
- 2) $(g_1g_2)^x = g_1^xg_2^x$;
- 3) $(g^{-1})^x = (g^x)^{-1}$.

Действительно,

$$\begin{aligned}
 g^{xy} &= (xy)^{-1}g(xy) = y^{-1}(x^{-1}gx)y = y^{-1}(g^x)y = (g^x)^y, \\
 (g_1g_2)^x &= x^{-1}(g_1g_2)x = (x^{-1}g_1x)(x^{-1}g_2x) = g_1^xg_2^x, \\
 (g^{-1})^x &= x^{-1}g^{-1}x = (gx)^{-1}(x^{-1})^{-1} = (x^{-1}gx)^{-1} = (g^x)^{-1}.
 \end{aligned}$$

Лемма 11.3. Скажем, что $g_1 \sim_G g_2$, если при некотором x имеет место равенство $g_1^x = g_2$. Отношение \sim_G является отношением эквивалентности.

Доказательство. Рефлексивность: $g = g^e$. Симметричность: если $g_1^x = g_2$, то $g_2^{x^{-1}} = g_1$. Транзитивность: если $g_1^x = g_2$ и $g_2^y = g_3$, то $g_1^{xy} = g_3$. \square

Таким образом, группа G разбивается на классы эквивалентности по отношению \sim_G , которые называются **классами сопряженности**. Класс сопряженности элемента g обозначается g^G . Это — «сопряжение по Минковскому», т. е.

$$g^G = \{g^x \mid x \in G\}.$$

На этом, правда, аналогия с классами смежности и заканчивается. Как мы увидим в дальнейшем, классы сопряженности имеют различный размер и могут по-разному соотноситься с классами смежности.

Лемма 11.4. Функция $g \mapsto g^x$ при фиксированном x является изоморфизмом группы G на себя.

Доказательство. Во-первых, $f(g) = g^x$ сохраняет операцию группы G , поскольку $f(g_1g_2) = f(g_1)f(g_2)$. Во-вторых, это инъекция, т. к. если $x^{-1}g_1x = x^{-1}g_2x$, то $g_1 = g_2$. В-третьих, это сюръекция, т. к. для любого $g \in G$ найдется g' такой, что $g = f(g')$, а именно, $g' = xgx^{-1}$. \square

Итак, при каждом x функция $f(g) = g^x$ задает изоморфизм группы G на себя. Изоморфизм группы на себя называется **автоморфизмом** группы G . Множество всех автоморфизмов группы G принято обозначать $\text{Aut}(G)$.

Классы сопряженности напрямую связаны с нормальными подгруппами.

Теорема 11.2. Подгруппа H группы G является нормальной тогда и только тогда, когда она есть объединение классов сопряженности.

Доказательство. Пусть $H \triangleleft G$, тогда $H = x^{-1}Hx$ для любого $x \in G$. Тогда вместе с любым элементом $h \in H$ имеем вхождение всего класса $h^G \subseteq H$. То есть если класс сопряженности пересекается с нормальной подгруппой, то он входит в нее целиком. Отсюда следует, что нормальная подгруппа является объединением классов сопряженности (не обязательно всех).

Обратно, пусть H — подгруппа группы G и $H = g_1^G \cup \dots \cup g_n^G$ при некотором n . Операция сопряжения сохраняет классы сопряженности:

$$x^{-1}g^Gx = \{x^{-1}g^yx \mid y \in G\} = \{g^{yx} \mid y \in G\} = g^G.$$

Последнее равенство справедливо, поскольку элемент yx — произвольный элемент G , т. е. для любого $z \in G$ имеем $z = yx$ при $y = zx^{-1}$.

Тогда получаем, что

$$x^{-1}Hx = x^{-1}g_1^Gx \cup \dots \cup x^{-1}g_n^Gx = g_1^G \cup \dots \cup g_n^G = H,$$

откуда следует, что подгруппа H является нормальной. \square

Отметим, что произвольное объединение классов сопряженности не обязано быть подгруппой, а значит, и нормальной подгруппой. Однако нормальные подгруппы следует искать только среди таких объединений, что сильно упрощает их поиск. Важно, чтобы эти объединения оказались замкнутыми относительно групповой операции! Поэтому такое объединение должно включать как минимум тривиальный класс сопряженности $e^G = \{e\}$.

Существует и еще один способ выделить нормальную подгруппу, а именно — воспользоваться гомоморфизмом групп.

Пусть $h : G \rightarrow G'$ является гомоморфизмом групп $\langle G, \cdot \rangle$ и $\langle G', \times \rangle$.

Ядром гомоморфизма h называется прообраз единицы группы G' :

$$\ker(h) = h^{-1}\{e'\} = \{g \in G \mid h(g) = e'\},$$

где e' — единица группы G' .

Гомоморфизм обладает следующими свойствами.

1) Единицу переводит в единицу. Действительно, пусть $h(e) = g'$. Тогда

$$g' = h(e) = h(ee) = h(e) \times h(e) = g' \times g',$$

откуда, используя сокращение в группе G' , получаем, что $e' = g'$.

2) Обратный элемент переводит в обратный. Пусть $g' = h(g)$ и $g'' = h(g^{-1})$, тогда

$$g' \times g'' = h(g) \times h(g^{-1}) = h(gg^{-1}) = h(e) = e',$$

т. е. элемент g'' — обратный к g' , или $h(g^{-1}) = h(g)^{-1}$.

3) Ядро гомоморфизма есть нормальная подгруппа: $\ker(h) \triangleleft G$.

Проверим аксиомы группы. Пусть $g_1, g_2 \in \ker(h)$, тогда $h(g_1g_2) = h(g_1) \times h(g_2) = e'$, откуда $g_1g_2 \in \ker(h)$, т. е. ядро замкнуто относительно групповой операции в G . Ассоциативность наследуется из G . Единица находится в ядре, согласно свойству $\text{Hom}1$.

Пусть $g \in \ker(h)$, тогда $h(g^{-1}) = h(g)^{-1} = (e')^{-1} = e'$, откуда $g^{-1} \in \ker(h)$. Таким образом, все требования группы выполнены, и $\ker(h)$ является подгруппой в G . Проверим ее нормальность.

Поскольку h сохраняет групповую операцию, для любого $g \in G$ легко вычислить образ

$$h[g^{-1} \ker(h)g] = h(g)^{-1} \times h[\ker(h)] \times h(g) = \{e'\},$$

откуда следует, что $g^{-1} \ker(h)g \subseteq \ker(h)$. А это по доказанному ранее критерию нормальности (11.4) означает, что $\ker(h) \triangleleft G$.

О тесной связи гомоморфизмов и нормальных подгрупп говорит также следующее утверждение.

Теорема 11.3 (Основная теорема о гомоморфизмах групп).

Пусть $h : G \rightarrow G'$ — гомоморфизм группы G в группу G' . Тогда факторгруппа $G/\ker(h)$ изоморфна области значений h в группе G' .

Наоборот, если $H \triangleleft G$, то существуют такие группа G' и гомоморфизм $h : G \rightarrow G'$, что $H = \ker(h)$.

Доказательство. 1) Обозначим за K ядро $\ker(h)$, а область значений h в группе G' обозначим за H' . Необходимо построить изоморфизм между G/K и H' .

Для произвольного смежного класса gK вычислим образ $h[gK]$. Он равен $\{h(g)\}$, т. е. множеству с единственной точкой $h(g)$. Отсюда следует, что корректным будет следующее определение функции:

$$f(gK) = h(g),$$

поскольку $f(gK) = \bigcup h[gK] = \bigcup \{h(g)\} = h(g)$. Проще говоря, значение f на классе из G/K можно вычислить, взяв значение h на любом элементе данного класса.

Далее, f является сюръекцией на множество H' , поскольку для всякого $h' \in H'$ существует $g \in G$ такой, что $h(g) = h'$, а значит, $f(gK) = h(g) = h'$.

Кроме того, f является инъекцией в H' . Действительно, пусть $f(g_1K) = f(g_2K)$ при некоторых $g_1, g_2 \in G$. Тогда $h(g_1) = h(g_2)$, откуда $g_1g_2^{-1} \in K$, а значит, $g_1 \in Kg_2$ и, следовательно, $g_1 \in g_2K$ (поскольку $Kg_2 = g_2K$ в силу $K \triangleleft G$). Тогда $g_1K = g_2K$ (т. к. классы смежности либо не пересекаются, либо совпадают).

Таким образом, f — биекция между G/K и H' . Осталось показать, что она сохраняет групповую операцию:

$$f(g_1Kg_2K) = f(g_1g_2K) = h(g_1g_2) = h(g_1)h(g_2) = f(g_1K)f(g_2K).$$

Итак, f — изоморфизм между G/K и $H' = \text{ran}(h)$.

2) Пусть $H \triangleleft G$. Обозначим за G' группу G/H . Определим гомоморфизм из G в G' следующим образом:

$$h(g) = gH, \quad g \in G.$$

Функция h сохраняет групповую операцию, т. к. $h(g_1g_2) = (g_1g_2)H = (g_1H)(g_2H) = h(g_1)h(g_2)$. Поскольку единицей группы G/H является H и функция h переводит в H все элементы H , и только их, то очевидно, что $\ker(h) = H$, т. е. h — гомоморфизм с ядром H . \square

11.4. Знакопеременная группа

Мы введем теперь такое важное понятие, как **транспозиция**. Это — цикл, состоящий из двух элементов, например, (12) или (59) . Транспозиция меняет местами два элемента X_n , а остальные оставляет на месте. Любой цикл длины k можно представить как композицию $(k-1)$ транспозиций. Например,

$$(1234) = (14)(13)(12),$$

причем это представление неоднозначное, поскольку

$$(1234) = (2341) = (21)(24)(23).$$

Тем не менее любая перестановка (не только цикл) имеет *инвариант* по разложению в транспозиции.

Теорема 11.4. Если перестановка $s \in \mathbf{S}_n$ имеет два представления транспозициями

$$s = t_1 \dots t_k = \tau_1 \dots \tau_m,$$

то $k \equiv m \pmod{2}$.

Доказательство. Пусть дана некоторая перестановка s на n элементах. Запишем ее в виде матрицы

$$s = \begin{pmatrix} 1 & 2 & \dots & n \\ s_1 & s_2 & \dots & s_n \end{pmatrix},$$

где $s_1, \dots, s_n \in \{1, \dots, n\}$ попарно различны. Для такой перестановки s рассмотрим функцию

$$\Delta(s) = \prod_{i < j} (s_j - s_i), \quad (11.6)$$

т. е. находим все «подъемы» и «спуски» нашей перестановки. Например, $\Delta((123)) = (3-2)(1-2)(1-3)$. Иными словами, мы перемножаем все возможные разности пар элементов перестановки, вычисляя разность в обратном порядке следования индексов: если функция s растет на паре индексов ij , то разность будет положительной.

Посмотрим, что происходит с функцией Δ , если перестановку s умножить на произвольную транспозицию (ij) , где $i < j$. Во-первых,

$$s' = s \circ (ij) = \begin{pmatrix} 1 & 2 & \dots & i & \dots & j & \dots & n \\ s_1 & s_2 & \dots & s_j & \dots & s_i & \dots & s_n \end{pmatrix},$$

т. е. s_i поменялось местами с s_j : $s'_i = s_j$, $s'_j = s_i$, а остальные элементы остались на своих местах: $s'_k = s_k$ для всех $k \notin \{i, j\}$.

Во-вторых, рассмотрим отношение $\Delta(s)/\Delta(s')$. В числителе и знаменателе стоят одни и те же разности, только местами отличающиеся знаком. Поэтому для нахождения данного отношения нужно лишь найти количество таких перемен знаков в этих разностях.

Рассмотрим все возможные пары индексов и соответствующие им разности, входящие в функции $\Delta(s)$ и $\Delta(s')$, включающие и не включающие индексы i и j :

- $(s'_j - s'_i) = (s_i - s_j) = (-1)(s_j - s_i)$, т. е. в данном случае происходит смена знака сомножителя для пары индексов ij ;
- при $k > j$ получаем, что $(s'_k - s'_j)(s'_k - s'_i) = (s_k - s_i)(s_k - s_j)$, т. е. для двух пар индексов ik и jk произведение соответствующих разностей не меняется (отличие только в порядке множителей), и смены знака нет;
- при $k < i$ получаем, что $(s'_j - s'_k)(s'_i - s'_k) = (s_i - s_k)(s_j - s_k)$, т. е. для двух пар индексов ki и kj произведение соответствующих разностей тоже не меняется;
- при $i < k < j$ получаем, что $(s'_k - s'_i)(s'_j - s'_k) = (s_k - s_j)(s_i - s_k) = (-1)(s_j - s_k)(-1)(s_k - s_i)$, и снова произведение разностей для пар индексов ik и kj не меняется (хотя обе разности сменили знак);
- для всех остальных пар индексов kl , где $k < l$ и $k, l \notin \{i, j\}$, разность $s_l - s_k$ сохраняется.

Таким образом, функции $\Delta(s)$ и $\Delta(s')$ содержат одинаковые разности, некоторых из которых отличаются знаком, причем во всех случаях, кроме одного, знак либо не меняется, либо смена знака одной разности компенсируется сменой знака другой разности. А значит, отношение $\Delta(s)/\Delta(s')$ равно -1 .

Пусть теперь $s = \text{id}$, т. е. $s_k = k$ при всех $k \in \{1, \dots, n\}$. Тогда все разности $(s_j - s_i) = j - i$ положительны, т. е. $\Delta(\text{id}) > 0$.

Отсюда следует, во-первых, что композиция тождественной перестановки и транспозиции не может быть тождественной перестановкой, поскольку $\Delta(\text{id} \circ (ij)) = -\Delta(\text{id})$. А во-вторых, знак функции $\Delta(s)$ определяется выражением $(-1)^k$, где k — количество транспозиций в разложении перестановки s .

Отсюда следует, что если имеет место представление

$$s = t_1 \dots t_k = \tau_1 \dots \tau_m,$$

то $(-1)^k = (-1)^m$, а это возможно только в том случае, когда $k \equiv m \pmod{2}$. Теорема доказана. \square

Данная теорема позволяет корректно определить понятие четности (или знака) перестановки. А именно, знак функции $\Delta(s)$ называется **четностью перестановки** s и обозначается $\text{sgn}(s)$, причем если этот знак положительный, то перестановка s называется **четной**, а в противном случае — **нечетной**. И, как мы выяснили в ходе доказательства теоремы, четность перестановки s равна $(-1)^k$, где k — количество транспозиций в разложении s , и не зависит от конкретного представления s в виде композиции транспозиций. В частности, $\text{sgn}(\text{id}) = +1$, $\text{sgn}(ij) = -1$.

Кроме того, из определения функции $\Delta(s)$ по формуле (11.6) видно, что **четность перестановки определяется количеством инверсий данной перестановки**. **Инверсией** называется такая пара индексов ij , что $s_j < s_i$, т. е. когда перестановка меняет отношение порядка. Если k — количество инверсий перестановки s , то $\text{sgn}(s) = (-1)^k$. Таким образом, **четность перестановки совпадает с четностью количества ее инверсий**.

Отметим одно важно свойство $\text{sgn}(s)$. Пусть заданы две перестановки s и g . Запишем их в виде композиции транспозиций

$$s = t_1 \dots t_k, \quad g = \tau_1 \dots \tau_m.$$

Тогда четность композиции sg будет равна $(-1)^{k+m}$, поскольку $sg = t_1 \dots t_k \tau_1 \dots \tau_m$. Таким образом,

$$\text{sgn}(sg) = \text{sgn}(s) \text{sgn}(g). \quad (11.7)$$

Функция sgn , определенная на элементах группы \mathbf{S}_n и принимающая значения из множества $B = \{-1, 1\}$, является гомоморфизмом группы \mathbf{S}_n в группу $\langle B, \cdot \rangle$ с операцией умножения, поскольку в силу (11.7) функция sgn композиции перестановок ставит в соответствие произведение их знаков.

Поскольку функция sgn на группе \mathbf{S}_n действует как гомоморфизм в группу $\langle B, \cdot \rangle$, прообраз 1 в группе \mathbf{S}_n относительно данного гомоморфизма, а именно, **все четные перестановки**, образуют нормальную подгруппу в группе \mathbf{S}_n . Эта нормальная подгруппа обозначается A_n и называется **знакопеременной группой** порядка n . Не следует путать употребленное здесь слово «порядок» с порядком группы, означающим количество ее элементов, поскольку в A_n находится ровно половина элементов группы \mathbf{S}_n , т. е. $n!/2$, что значительно больше n при $n > 3$.

11.5. Структура группы перестановок

Посмотрим, как работает отношение сопряженности на группе перестановок.

Пусть $G = \mathbf{S}_3 = \{\text{id}, (12), (13), (23), (123), (132)\}$. Будем записывать действие g^x в виде таблицы, похожей на таблицу умножения, предполагая, что на верхнюю строку мы действуем функцией g^x , где x берется из левого столбца:

$x \backslash g$	id	(12)	(13)	(23)	(123)	(132)
id	id	(12)	(13)	(23)	(123)	(132)
(12)	id	(12)	(23)	(13)	(132)	(123)
(13)	id	(23)	(13)	(12)	(132)	(123)
(23)	id	(13)	(12)	(23)	(132)	(123)
(123)	id	(23)	(12)	(13)	(123)	(132)
(132)	id	(13)	(23)	(12)	(123)	(132)

Мы видим, что когда x есть транспозиция, то сопряжение к g при помощи x — это либо само g (а именно, при $g = x$, т.к. $x^x = x$), либо другая перестановка того же вида: сопряжением к транспозиции является транспозиция, а сопряжением к 3-циклу — 3-цикл. Если же x — это 3-цикл, то сопряжением к транспозиции является обязательно другая транспозиция, а сопряжением к 3-циклу является он сам.

Кроме того, каждая строка таблицы — это перестановка над элементами группы \mathbf{S}_3 (в соответствии с леммой 11.4 это не просто перестановка, а автоморфизм группы \mathbf{S}_3). Эту перестановку можно также рассматривать как элемент группы $\mathbf{S}(\mathbf{S}_3)$, изоморфной группе \mathbf{S}_6 .

Перестановки над элементами группы \mathbf{S}_3 , получаемые их сопряжением, подчиняются строгим ограничениям. Мы специально разделили таблицу на три вертикальных блока, включив в первый блок id, во второй — транспозиции, а в третий — 3-циклы. Такое разделение соответствует найденному нами ранее разбиению на классы сопряженности по отношению эквивалентности \sim_G (в нашем случае это $\sim_{\mathbf{S}_3}$). И действительно, если смотреть теперь на столбцы таблицы, то внутри каждого столбца мы видим только элементы, сопряженные друг другу.

Мы можем построить аналогичную таблицу для \mathbf{S}_4 и убедиться в том, что **(a)** каждая строка таблицы есть перестановка над элементами \mathbf{S}_4 и **(b)** каждый столбец — выборка с повторениями из соответствующих классов сопряженности. Таковыми классами для \mathbf{S}_4 являются классы одинакового цикленного состава:

()	id	1
(**)	(12),(13),(14),(23),(24),(34)	6
(***)	(123),(132),(124),(142),(134),(143),(234),(243)	8
(**)(**)	(12)(34), (13)(24), (14)(23)	3
(****)	(1234),(1243),(1324),(1342),(1423),(1432)	6

Эти классы можно обозначить как $\text{id}^{\mathbb{S}_4}$, $(12)^{\mathbb{S}_4}$, $(123)^{\mathbb{S}_4}$, $[(12)(34)]^{\mathbb{S}_4}$, $(1234)^{\mathbb{S}_4}$.

На самом деле, для любой группы \mathbb{S}_n классы сопряженности состоят из перестановок одинакового цикленного состава (т. е. совпадает количество циклов одинаковой длины).

Как мы отмечали ранее, теорема 11.2 упрощает поиск нормальных подгрупп тем, что такие подгруппы следует искать только среди всевозможных объединений классов сопряженности, включая тривиальный класс.

В группе \mathbb{S}_3 есть только две возможности создать нетривиальную нормальную подгруппу: на основе класса $(12)^{\mathbb{S}_3}$ и на основе класса $(123)^{\mathbb{S}_3}$. Но в первом случае подгруппа не получается, поскольку $(12)(23) = (123)$ выходит из класса транспозиций. Зато второй случай позволяет создать уже знакомую нам знакопеременную подгруппу, состоящую из всех четных перестановок на 3-х символах:

$$A_3 = \{\text{id}, (123), (132)\}.$$

По теореме 11.2 она является нормальной.

В группе \mathbb{S}_4 возможностей сильно больше, но далеко не все они приводят к подгруппе. Вспомним, что порядок подгруппы делит порядок группы (теорема Лагранжа). Следовательно, собственные нетривиальные подгруппы \mathbb{S}_4 должны иметь порядок из множества $\{2, 4, 6, 8, 12\}$. А перечисленные выше классы сопряженности имеют размеры 1,6,8,3,6. И так как подгруппа должна включать класс $\{\text{id}\}$, ее порядок может складываться только следующими двумя способами: $1 + 3$, $1 + 3 + 8$. И действительно, в группе \mathbb{S}_4 существует ровно две нетривиальные собственные нормальные подгруппы:

$$V_4 = \{\text{id}, (12)(34), (13)(24), (14)(23)\},$$

$$A_4 = \{\text{id}, (12)(34), (13)(24), (14)(23),$$

$$(123), (132), (124), (142), (134), (143), (234), (243)\}.$$

Первая — это уже известная нам четверная группа Клейна, которая изоморфна группе симметрий ромба, группе \mathbb{Z}_8^* и группе $\mathbb{Z}_2 \times \mathbb{Z}_2$. Можно построить гомоморфизм из \mathbb{S}_4 в \mathbb{S}_3 с ядром V_4 . У этого факта существует простое геометрическое доказательство. Дело в том, что группа \mathbb{S}_4 является группой симметрий тетраэдра (включая отражения), тогда,

если в тетраэдре соединить середины противоположных ребер, получится три отрезка. При движениях тетраэдра движутся и эти отрезки, переходя друг в друга, и эти движения образуют группу, изоморфную S_3 . При этом, как легко проверить, движения вершин тетраэдра, соответствующие перестановкам из группы Клейна, не меняют положение этих отрезков (они могут переворачиваться, но переходят каждый сам в себя), т. е. группа Клейна является ядром такого гомоморфизма. Из основной теоремы о гомоморфизмах 11.3 также следует, что S_4/V_4 изоморфна S_3 , поскольку S_3 — образ гомоморфизма с ядром V_4 .

Вторая — подгруппа A_4 — это знакопеременная подгруппа, содержащая все четные перестановки на 4-х символах, и только их.

Соберем известные нам группы 4-го порядка в одну таблицу 11.1 для сравнения, а также добавим к ним группу комплексных корней 4-ой степени из 1, речь о которой пойдет в разделе 12.2.

Отметим, что с точностью до изоморфизма существует всего две группы 4-го порядка: группа Клейна и циклическая группа. Несмотря на все разнообразие их представителей!

Игра «Пятнадцать»

Четность перестановки является инвариантом на подгруппе A_n (т. е. принимает одно и то же значение на всех ее элементах), а также на ее смежном классе в S_n (принимает другое постоянное значение). Поиск инвариантов является одним из мощных математических инструментов при доказательстве невозможности некоторых объектов или действий.

С конца XIX века известна игра «пятнадцать», суть которой в следующем. Имеем поле 4×4 , в котором расставлены одинаковые фишки размером 1×1 . Всего фишек 15, и они пронумерованы числами от 1 до 15. Одно место на поле пустое, что позволяет производить следующие простые манипуляции: занимать данное место фишкой с любого смежного с пустым места, т. е. передвигать ее на это место, освобождая исходное. При этом нельзя совершать никакие другие действия, например, вынимать фишки с поля и расставлять их произвольным образом.

Стандартная позиция фишек на поле приведена на рис. 11.2.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

Рис. 11.2. Стандартная позиция игры «пятнадцать»

Глядя на этот квадрат с числами, мы можем закодировать его, пере-

Таблица 11.1. Группы 4-го порядка

Четверная группа Клейна					Циклическая 4-го порядка				
\diamond	id	R	S_1	S_2	\square	id	$R_{\pi/2}$	R_π	$R_{3\pi/2}$
id	id	R	S_1	S_2	id	id	$R_{\pi/2}$	R_π	$R_{3\pi/2}$
R	R	id	S_2	S_1	$R_{\pi/2}$	$R_{\pi/2}$	R_π	$R_{3\pi/2}$	id
S_1	S_1	S_2	id	R	R_π	R_π	$R_{3\pi/2}$	id	$R_{\pi/2}$
S_2	S_2	S_1	R	id	$R_{3\pi/2}$	$R_{3\pi/2}$	id	$R_{\pi/2}$	R_π
\mathbb{Z}_8^*	1	3	5	7	\mathbb{Z}_4	0	1	2	3
1	1	3	5	7	0	0	1	2	3
3	3	1	7	5	1	1	2	3	0
5	5	7	1	3	2	2	3	0	1
7	7	5	3	1	3	3	0	1	2
$\mathbb{Z}_2 \oplus \mathbb{Z}_2$	00	01	10	11	\mathbb{Z}_5^*	1	2	4	3
00	00	01	10	11	1	1	2	4	3
01	01	00	11	10	2	2	4	3	1
10	10	11	00	01	4	4	3	1	2
11	11	10	01	00	3	3	1	2	4
id	(12)(34)	(13)(24)	(14)(23)		$\sqrt[4]{1}$	1	i	-1	$-i$
(12)(34)	id	(14)(23)	(13)(24)		1	1	i	-1	$-i$
(13)(24)	(14)(23)	id	(12)(34)		i	i	-1	$-i$	1
(14)(23)	(13)(24)	(12)(34)	id		-1	-1	$-i$	1	i
					$-i$	$-i$	1	i	-1

числя все фишки в порядке «змейки»

$$\langle 1, 2, 3, 4, 8, 7, 6, 5, 9, 10, 11, 12, 15, 14, 13 \rangle, \tag{11.8}$$

пропуская пустое место при таком обходе.

Этот набор из 15 чисел кодирует позицию игры с точностью до положения пустого места на игровом поле (действительно, пустое место можно передвигать вдоль «змейки», не нарушая правил игры, так что все 16 положений пустого поля с одним и тем же порядком чисел в «змейке» задаются одним и тем же кодом). Разные коды всегда означают разные позиции игры, которые невозможно перевести друг в друга перемещением пустого поля «вдоль змейки». Если нам дали такую по-

зицию игры, где пустое поле находится где-то посередине поля, то мы легко можем переместить его в правый нижний угол, не меняя последовательности фишек в нашем коде. Таким образом, мы всегда можем считать, что код определяет именно такую позицию, где пустое поле находится в правом нижнем углу поля.

Заметим, что набор чисел (11.8) есть не что иное, как перестановка из группы S_{15} . Иначе говоря, все позиции игры «пятнадцать» (с точностью до положения пустого поля) задаются всеми перестановками группы S_{15} .

«Фишка» этой игры в том, что все разрешенные действия по перемещению фишек на игровом поле не меняют четности перестановки, соответствующей расположению фишек. А это значит, что никакую расстановку, закодированную нечетной перестановкой невозможно привести (разрешенными действиями) к расстановке, закодированной четной перестановкой, и наоборот. Например, две расстановки фишек, у которых кодирующие перестановки отличаются лишь одной транспозицией (обменом двух соседних фишек местами), не могут быть переведены одна в другую по правилам игры.

Создатель игры даже обещал большой приз тому, кто приведет расстановку

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

к виду

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

(они отличаются транспозицией (15, 14)).

С тех пор прошло больше 100 лет, и до сих пор многие пытаются это сделать, но алгебра дает нам беспощадный ответ: этого сделать невозможно! Потому что четность перестановки инвариантна относительно действий с фишками!

Нормальный ряд

Выше мы отыскиали все нормальные подгруппы S_3 и S_4 . В таблице 11.2 помещена полная таблица умножения группы S_4 с использованием кратких обозначений перестановок как произведений циклов. Там же выделены две подтаблицы, отвечающие группам A_4 и V_4 , а также отмечены (желтым) элементы (и их произведения) некоторой подгруппы 8-го порядка, которая не является ни нормальной, ни коммутативной (и поэтому она не является объединением классов сопряженности).

Выделенная подгруппа 8-го порядка:

$$\{e, (12)(34), (13)(24), (14)(23), (12), (34), (1324), (1423)\}.$$

Все подгруппы 8-го порядка группы \mathbf{S}_4 изоморфны. Аналогичная ситуация с подгруппами 6-го порядка, вот одна из них: $\{e, (123), (132), (12), (13), (23)\}$, которая совпадает с \mathbf{S}_3 .

Как уже отмечалось, в группе \mathbf{S}_4 существует только 2 нетривиальные нормальные подгруппы: A_4 и V_4 .

Говорят, что группа G имеет **субнормальный ряд** (называемый также **субнормальной башней**, **субинвариантным рядом**, **субнормальной матрешкой** или просто **рядом**) длины n , если имеют место вложения:

$$\{e\} = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_{n-1} \triangleleft G_n = G,$$

где G_i — собственная нормальная подгруппа в G_{i+1} . Ряд называется **нормальным**, если все G_i нормальны также в исходной группе G . Факторгруппы G_{i+1}/G_i называются **факторами** (факторгруппами) **ряда**.

Для простых групп (например, \mathbb{Z}_p) тривиальный субнормальный ряд длины 1 является единственно возможным: $\{e\} \triangleleft G$.

Для группы \mathbf{S}_4 имеем:

$$\{e\} \triangleleft V_4 \triangleleft A_4 \triangleleft \mathbf{S}_4, \quad \{e\} \triangleleft V_4 \triangleleft \mathbf{S}_4.$$

Эти утверждения можно извлечь непосредственно из таблицы 11.2. Например, нормальность V_4 в A_4 следует из того, что симметричные столбец и строка в зеленой области напротив и под группой V_4 совпадают с точностью до перестановки элементов (т. е. выполняется условие $gH = Hg$).

Если для группы G существует такой субнормальный ряд, что все его факторы — коммутативные группы, то группа G называется **разрешимой**.

Так как

- 1) $\mathbf{S}_4/A_4 \cong \mathbb{Z}_2$, т. е. является циклической группой,
- 2) $A_4/V_4 \cong \mathbb{Z}_3$, т. е. является циклической группой,
- 3) $V_4/\{e\}$ — коммутативная группа (см. таблицу 11.2),

то \mathbf{S}_4 разрешима. Заметим, что ряд $\{e\} \triangleleft V_4 \triangleleft \mathbf{S}_4$ не годится для установления разрешимости, поскольку фактор $\mathbf{S}_4/V_4 \cong \mathbf{S}_3$ не является коммутативной группой.

Для $n = 3$ имеем $\{e\} \triangleleft A_3 \cong \mathbb{Z}_3 \triangleleft \mathbf{S}_3$ и, таким образом, \mathbf{S}_3 также разрешима. Тем более разрешима и $\mathbf{S}_2 \cong \mathbb{Z}_2$.

Известно, что все \mathbf{S}_n порядка $n \geq 5$ неразрешимы. Именно на этом замечательном факте построено доказательство знаменитой теоремы Абеля — Руффини о неразрешимости в радикалах уравнений степени 5 и выше. Теорема Абеля — Руффини не заявляет о том, что уравнение

5 и более высокой степени вообще не имеет решения (на комплексной плоскости всякое уравнение n -й степени разрешимо при $n \geq 1$). Суть теоремы Абеля — Руффини сводится к тому, что для произвольных уравнений степени выше четвертой невозможно указать *явную общую формулу*, содержащую только арифметические операции и знаки корня, которая позволила бы вычислить хотя бы один корень данного уравнения. В некоторых частных случаях это возможно, а в общем случае — нет.

Образующие группы перестановок

Покажем, что любая перестановка в S_n может быть получена с помощью только лишь двух циклов (возможно, многократными композициями этих циклов и обратных к ним циклов):

$$(12) \quad \text{и} \quad (12 \dots n).$$

Ранее мы уже видели, что всякая перестановка есть композиция циклов, а всякий цикл — композиция транспозиций:

$$(a_1 a_2 \dots a_{n-1} a_n) = (a_1 a_2)(a_2 a_3) \dots (a_{n-1} a_n). \quad (11.9)$$

Стало быть, нужно научиться получать только транспозиции. Заметим, что если $a < b$, то

$$(ab) = \sigma^{-1}(b-1 \ b)\sigma,$$

где цикл $\sigma = (a \ a+1 \ \dots \ b-1) = (a \ a+1)(a+1 \ a+2) \dots (b-2 \ b-1)$, так что любая транспозиция (а значит, и вообще любая перестановка) сводится к транспозициям вида $(k \ k+1)$, т. е. к транспозициям соседних символов.

Например,

$$(25) = (23)(34)(45)(43)(32).$$

Предположим, что мы уже умеем задавать транспозиции $(12), (23), \dots, (k-1 \ k)$. Как получить транспозицию $(k \ k+1)$? И вот тут понадобится самый длинный цикл $(12 \dots n)$. Действительно,

$$(12 \dots n-1 \ n)(k-1 \ k)(n \ n-1 \ \dots \ 21) = (k \ k+1).$$

Например,

$$(12345)(34)(54321) = (45).$$

Таким образом, имея на старте транспозицию (12) и полный цикл $(12 \dots n)$, мы можем получить все транспозиции из соседних элементов, из них — вообще все транспозиции, а из последних — вообще все перестановки группы S_n .

Покажем, что любая четная перестановка (т. е. элемент группы A_n) может быть получена с помощью композиции 3-циклов. Для этого снова представим перестановку в виде композиции транспозиций, как в (11.9). Поскольку наша перестановка четная, этих транспозиций четное число. Разобьем их по парам и докажем, что любую комбинацию вида $(ij)(kl)$ можно заменить композицией не более чем двух 3-циклов.

Если $(ij) = (kl)$, то $(ij)(kl) = \text{id}$, и такую пару можно сразу исключить из представления четной перестановки.

Если k — один из символов i, j , а l — какой-то третий, то случай сводится к паре $(ij)(jl) = (ijl)$, т. е. 3-циклу.

Если же все 4 символа разные, то $(ij)(kl) = (ijk)(jkl)$, т. е. тоже композиция 3-циклов.

Таблица 11.2. Таблица умножения симметрической группы S_4

Знакопеременная группа A_4											
Четверная группа Клейна											
e	(12)(34)	(13)(24)	(14)(23)	(123)	(132)	(124)	(142)	(134)	(143)	(234)	(243)
(12)(34)	e	(14)(23)	(13)(24)	(243)	(143)	(234)	(134)	(142)	(132)	(124)	(123)
(13)(24)	(14)(23)	e	(12)(34)	(142)	(234)	(143)	(123)	(243)	(124)	(132)	(134)
(14)(23)	(13)(24)	(12)(34)	e	(134)	(124)	(132)	(243)	(123)	(234)	(143)	(142)
(123)	(134)	(243)	(142)	(132)	e	(13)(24)	(143)	(234)	(14)(23)	(12)(34)	(124)
(132)	(234)	(124)	(143)	e	(123)	(243)	(14)(23)	(12)(34)	(142)	(134)	(13)(24)
(124)	(143)	(132)	(234)	(14)(23)	(134)	(142)	e	(13)(24)	(243)	(123)	(12)(34)
(142)	(243)	(134)	(123)	(234)	(13)(24)	e	(124)	(132)	(12)(34)	(14)(23)	(143)
(134)	(123)	(142)	(243)	(124)	(14)(23)	(12)(34)	(234)	(143)	e	(13)(24)	(132)
(143)	(124)	(234)	(132)	(12)(34)	(243)	(123)	(13)(24)	e	(134)	(142)	(14)(23)
(234)	(132)	(143)	(124)	(13)(24)	(142)	(134)	(12)(34)	(14)(23)	(123)	(243)	e
(243)	(142)	(123)	(134)	(143)	(12)(34)	(14)(23)	(132)	(124)	(13)(24)	e	(234)
(12)	(34)	(1324)	(1423)	(23)	(13)	(24)	(14)	(1342)	(1432)	(1234)	(1243)
(13)	(1234)	(24)	(1432)	(12)	(23)	(1243)	(1423)	(34)	(14)	(1342)	(1324)
(14)	(1243)	(1342)	(23)	(1234)	(1324)	(12)	(24)	(13)	(34)	(1423)	(1432)
(23)	(1342)	(1243)	(14)	(13)	(12)	(1324)	(1432)	(1234)	(1423)	(34)	(24)
(24)	(1432)	(13)	(1234)	(1423)	(1342)	(14)	(12)	(1324)	(1243)	(23)	(34)
(34)	(12)	(1423)	(1324)	(1243)	(1432)	(1234)	(1342)	(14)	(13)	(24)	(23)
(1234)	(13)	(1432)	(24)	(1324)	(14)	(1342)	(34)	(1423)	(23)	(1243)	(12)
(1243)	(14)	(23)	(1342)	(1432)	(34)	(1423)	(13)	(24)	(1324)	(12)	(1234)
(1324)	(1423)	(12)	(34)	(14)	(1234)	(1432)	(23)	(1243)	(24)	(13)	(1342)
(1342)	(23)	(14)	(1243)	(24)	(1423)	(34)	(1234)	(1432)	(12)	(1324)	(13)
(1423)	(1324)	(34)	(12)	(1342)	(24)	(13)	(1243)	(23)	(1234)	(1432)	(14)
(1432)	(24)	(1234)	(13)	(34)	(1243)	(23)	(1324)	(12)	(1342)	(14)	(1423)

*Здесь **особым фоном** выделены элементы, образующие группу, изоморфную \mathbb{Z}_3 , поскольку их 3-я степень равна e.

Таблица 11.3. Продолжение таблицы 11.2

(12)	(13)	(14)	(23)	(24)	(34)	(1234)	(1243)	(1324)	(1342)	(1423)	(1432)
(34)	(1432)	(1342)	(1243)	(1234)	(12)	(24)	(23)	(1423)	(14)	(1324)	(13)
(1423)	(24)	(1243)	(1342)	(13)	(1324)	(1432)	(14)	(34)	(23)	(12)	(1234)
(1324)	(1234)	(23)	(14)	(1432)	(1423)	(13)	(1342)	(12)	(1243)	(34)	(24)
(13)	(23)	(1423)	(12)	(1243)	(1234)	(1342)	(1324)	(24)	(34)	(1432)	(14)
(23)	(12)	(1432)	(13)	(1324)	(1342)	(34)	(24)	(1243)	(1234)	(14)	(1423)
(14)	(1324)	(24)	(1234)	(12)	(1243)	(1423)	(1432)	(1342)	(13)	(23)	(34)
(24)	(1342)	(12)	(1423)	(14)	(1432)	(23)	(34)	(13)	(1324)	(1234)	(1243)
(1234)	(14)	(34)	(1324)	(1342)	(13)	(1243)	(12)	(1432)	(1423)	(24)	(23)
(1243)	(34)	(13)	(1432)	(1423)	(14)	(12)	(1234)	(23)	(24)	(1342)	(1324)
(1342)	(1423)	(1234)	(24)	(34)	(23)	(1324)	(13)	(14)	(1432)	(1243)	(12)
(1432)	(1243)	(1324)	(34)	(23)	(24)	(14)	(1423)	(1234)	(12)	(13)	(1342)
e	(132)	(142)	(123)	(124)	(12)(34)	(234)	(243)	(13)(24)	(134)	(14)(23)	(143)
(123)	e	(143)	(132)	(13)(24)	(134)	(12)(34)	(124)	(243)	(234)	(142)	(14)(23)
(124)	(134)	e	(14)(23)	(142)	(143)	(123)	(12)(34)	(132)	(13)(24)	(234)	(243)
(132)	(123)	(14)(23)	e	(243)	(234)	(134)	(13)(24)	(124)	(12)(34)	(143)	(142)
(142)	(13)(24)	(124)	(234)	e	(243)	(14)(23)	(143)	(134)	(132)	(123)	(12)(34)
(12)(34)	(143)	(134)	(243)	(234)	e	(124)	(123)	(14)(23)	(142)	(13)(24)	(132)
(134)	(14)(23)	(234)	(124)	(12)(34)	(123)	(13)(24)	(132)	(142)	(143)	(243)	e
(143)	(243)	(13)(24)	(12)(34)	(123)	(124)	(142)	(14)(23)	(234)	e	(132)	(134)
(14)(23)	(124)	(243)	(134)	(132)	(13)(24)	(143)	(142)	(12)(34)	(123)	e	(234)
(234)	(142)	(12)(34)	(13)(24)	(134)	(132)	(243)	e	(143)	(14)(23)	(124)	(123)
(13)(24)	(234)	(123)	(142)	(143)	(14)(23)	(132)	(134)	e	(243)	(12)(34)	(124)
(243)	(12)(34)	(132)	(143)	(14)(23)	(142)	e	(234)	(123)	(124)	(134)	(13)(24)

Желтым фоном выделена таблица подгруппы 8 порядка. Данная подгруппа некоммутативна.

Упражнения

Обязательные упражнения

11.1° Какие перестановки из S_4 не являются циклами? Разложите их в произведение независимых циклов.

11.2° Сколько всего различных циклов длины k в S_n ?

11.3° Докажите, что любая перестановка из S_n однозначно, с точностью до порядка множителей, разлагается в произведение независимых циклов (циклы длины 1 обычно пропускают).

11.4° а) Докажите, что если циклы независимы, то они коммутируют. б) Верно ли обратное?

11.5° Текст на русском языке зашифрован программой, заменяющей взаимно однозначно каждую букву на некоторую другую. а) Докажите, что существует такое число k , что текст расшифровывается применением k раз шифрующей программы. б) Найдите хотя бы одно такое k .

11.6° Найдите порядки: а) перестановок из S_3 ; б) цикла длины k .

11.7° Найдите все α из S_n , для которых $\alpha = \alpha^{-1}$.

11.8° Пусть $\alpha = (1\ 2\ \dots\ n)^k$. На сколько независимых циклов раскладывается α , каковы их длины?

11.9° а) Докажите, что произвольный цикл в некоторой степени даст тождественную перестановку. б) Докажите, что любая перестановка в некоторой степени даст тождественную. в) Найдите порядок цикла длины m . г) Найдите все возможные порядки перестановок множества из 7 и 8 элементов.

11.10° Найдите максимальный возможный порядок перестановки а) из S_5 ; б) из S_{13} ; в) из S_n .

11.11° Докажите, что порядок перестановки из S_n делит $n!$. Может ли он быть равен $(n!)$?

11.12° Упростите (представьте в виде цикла или произведения независимых циклов):

а) $\left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 1 & 3 \end{array} \right)^{100}$;

б) $(1\ k)(1\ k-1)\dots(1\ 3)(1\ 2)$; в) $(i+1\ i+2)(i\ i+1)(i+1\ i+2)$;

г) $(1\ 2\ \dots\ n)^{n-1}$; д) $(1\ 2\ \dots\ n)(1\ 2)(1\ 2\ \dots\ n)^{n-1}$.

11.13° Вычислите, чему равны следующие перестановки:

a) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}^{100}$; b) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}^{1000}$; c) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 1 & 4 \end{pmatrix}^{-1000}$;

d) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 1 & 3 \end{pmatrix}^{500}$; e) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 2 & 6 & 3 & 1 \end{pmatrix}^{-127}$;

f) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 5 & 1 & 2 & 3 & 4 \end{pmatrix}^{1001}$; g) $\begin{pmatrix} 1 & 2 & \dots & n \\ n & n-1 & \dots & 1 \end{pmatrix}^n$.

11.14° a) Пусть порядок перестановки равен двум. Разложим ее в произведение независимых циклов. Какими могут быть длины этих циклов? b) Пусть σ — это k -я степень цикла $(1, 2, \dots, n)$. На сколько независимых циклов раскладывается σ ? Каковы длины этих циклов?

11.15° Пусть $n \geq 2$. Какие перестановки из S_n получаются композициями перестановок, каждая из которых — транспозиция $(1\ 2)$ или цикл $(1\ 2\ \dots\ n)$?

11.16° Сколько различных перестановок встречается среди степеней перестановки $(123)(4567)$?

11.17° Докажите, что множество S_n содержит $n!$ элементов.

11.18° a) Сколько существует перестановок чисел $1, 2, \dots, 5$? Сколько из них оставляют число 1 на месте? b) Сколько из них переводят 1 в 5? c) Для скольких из них $\sigma(1) < \sigma(2)$? d) Для скольких из них $\sigma(1) < \sigma(2) < \sigma(3)$?

11.19° Перед Петей на столе лежат в ряд n шариков, пронумерованные по порядку числами от 1 до n . Петя переставил местами шарики. Пусть α сопоставляет числу k число $\alpha(k)$ — номер места в ряду, на котором оказался шарик под номером k . a) Покажите, что α — перестановка из S_n . b) Затем Петя повторил движения рук (опять переставил шарики, даже не глядя на них). На этот раз шарик под номером k оказался на месте под номером $\beta(k)$. Выразите перестановку β через перестановку α .

11.20° Вычислить следующие перестановки:

a) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$; b) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$;

c) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}$; d) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}^2$; e) $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}^3$;

f) $\begin{pmatrix} 3 & 5 & 6 & 1 & 2 & 7 & 9 & 4 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 5 & 6 & 1 & 2 & 7 & 9 & 4 & 8 \end{pmatrix}$.

11.21°

а) Всегда ли $\sigma\tau = \tau\sigma$? б) Пусть $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$, $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$. Найдите $\sigma\tau$ и $\tau\sigma$.

11.22° Докажите, что:

а) для любых перестановок σ, τ, η выполнено равенство $(\sigma\tau)\eta = \sigma(\tau\eta)$;

б) для любой перестановки σ справедливо равенство $(\sigma^{-1})^{-1} = \sigma$.

11.23° Найдите такую перестановку e , что $e\alpha = \alpha e = \alpha$ при всех α (она называется *тождественной*) и обозначается id . Докажите ее единственность.

11.24° Для всякой перестановки α найдите такую перестановку α^{-1} , что $\alpha\alpha^{-1} = \alpha^{-1}\alpha = \text{id}$. Такая перестановка называется *обратной* к перестановке α . Докажите ее единственность.

11.25° Найдите α^{-1} для каждой α из S_3 .

11.26° Пусть $\sigma = (123)$, $\tau = (34)$. Чему равно $\tau\sigma\tau^{-1}$?

11.27° Какой шарик стоит на месте k после применения перестановки α из задачи 11.19°?

11.28° Пусть p — простое число, $\mathbb{Z}/p\mathbb{Z}$ — классы вычетов по модулю p . Докажите, что умножение на ненулевой остаток $a \in \mathbb{Z}/p\mathbb{Z}$ является перестановкой ненулевых остатков $\{1, 2, \dots, p-1\}$, причем $a = 1$ соответствует тождественной перестановке, обратный элемент — обратной, а произведение — композиции.

11.29° Для скольких перестановок чисел 1, 2, 3, 4 выполнено равенство:

а) $\sigma^2 = \text{id}$; б) $\sigma = \sigma^{-1}$; в) $\sigma^2 = \sigma^{-1}$?

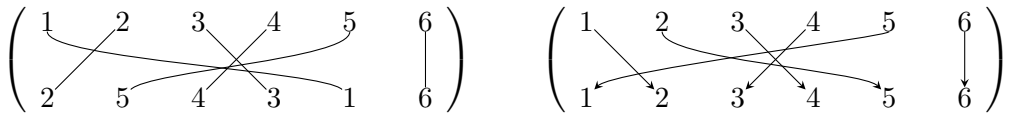
11.30° Во дворе стоят а) 17 б) 18 мальчиков. У каждого в руках мяч. Вдруг они одновременно кинули свои мячи друг другу. Петя и Вася наблюдали за ними. Петя утверждает, что может мысленно расположить мальчиков в круг так, что каждый кинул стоящему через одного по часовой стрелке. Аналогично Вася, но в кругу Васи каждый кидает стоящему через двух по часовой стрелке. Не врут ли Петя и Вася?

Дополнительные упражнения

11.31 Докажите, что четность перестановки при домножении на транспозицию меняется.

11.32 Чтобы увидеть число инверсий геометрически, на картинке, можно поступить двумя способами. Первый: в таблице, отвечающей перестановке α , соединим нитями одинаковые элементы (картинка слева).

Второй: нарисуем таблицу с двумя одинаковыми верхними строками — $1, 2, \dots, n$, — и каждый элемент i верхней строки соединим нитью с элементом $\alpha(i)$ во второй строке (картинка справа).



- a) Как увидеть количество инверсий на этой картинке (можно дать ответ для одного способа)?
- b) Сделайте это для $(2\ 3\ 4)$ и $(14)(23)$ из S_4 .
- c) Изменится ли четность числа инверсий, если в нижней строке таблицы поменять два элемента местами?

11.33 a) Какие перестановки в S_3 четные? b) Какие из перестановок задачи 11.20° четные?

11.34 Сколько инверсий у перестановки

$$\begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n & n-1 & \dots & 2 & 1 \end{pmatrix}?$$

11.35 Можно ли сказать, сколько инверсий у перестановки σ^{-1} , зная лишь число инверсий у σ ?

11.36 Для каких k в S_n существует перестановка, у которой ровно k инверсий?

11.37 Доказать, что четных перестановок столько же, сколько нечетных.

11.38 Докажите, что четность цикла зависит только от его длины. Как?

11.39 a) Как выражается четность $\sigma\tau$ через четности σ и τ ? b) Как выражается четность σ^n через четности σ и n ?

11.40 В таблице n строк и m столбцов. *Горизонтальный ход* — это любая перестановка элементов таблицы, при которой каждый элемент остается в той же строке, что и до перестановки. Аналогично определяется *вертикальный ход*. За какое наименьшее число горизонтальных и вертикальных ходов всегда удастся получить любую перестановку элементов таблицы?

11.41 У отца было 7 дочерей. Всякий раз, когда одна выходила замуж, каждая ее старшая сестра, оставшаяся в невестах, жаловалась отцу, что

нарушен обычай выходить замуж по старшинству. После того как вышла замуж последняя дочь, оказалось, что отец услышал всего 7 жалоб. В каком порядке дочери могли выходить замуж (приведите пример)? Сколько всего таких порядков?

11.42 В каждой клетке таблицы $2 \times n$ стоит одно из целых чисел от 1 до n , причем в каждой строке стоят разные числа, и в каждом столбце стоят разные числа. Сколько таких таблиц?

11.43 Докажите, что если в игре в «пятнашки» поменять местами фишки с номерами 14 и 15, то, следуя правилам, невозможно получить первоначальное расположение фишек.

11.44 Каждому из n мудрецов написали на лбу число и выдали две варежки: черную и белую. По сигналу все мудрецы одновременно надевают варежки. После этого их строят в шеренгу в порядке возрастания написанных на их лбах чисел и просят соседей взяться за руки. Как мудрецам надевать варежки, чтобы в результате каждая белая варежка взялась за белую, а каждая черная — за черную? (Мудрец видит все числа, кроме своего; все написанные на лбах числа различны.)

11.45 Пусть $n > 3$. Докажите, что A_n — это в точности множество перестановок из S_n , которые можно разложить в произведение циклов длины 3 (повторения разрешаются).

11.46 Пусть s_l — количество перестановок с числом инверсий l . Покажите, что

$$1 + s_1x + s_2x^2 + s_3x^3 + \dots = (1+x)(1+x+x^2)\dots(1+x+\dots+x^{n-1}).$$

11.47 Ассоциативна ли операция \circ на множестве M , если: **a)** $M = \mathbb{N}$, $x \circ y = x^y$; **b)** $M = \mathbb{Z}$, $x \circ y = x^2 + y^2$; **c)** $M = \mathbb{R}$, $x \circ y = \sin x \cdot \sin y$

11.48 Пусть на множестве X задана операция $x \circ y$ и для нее выполняются тождества: $\forall x, y \in X \quad (x \circ y) \circ y = x$ и $y \circ (y \circ x) = x$. Доказать, что данная операция коммутативна, т. е. $x \circ y = y \circ x$.

11.49 Пусть на \mathbb{Z} задана операция $n \circ m = n + m + nm = (1+n)(1+m) - 1$. Каким аксиомам группы она удовлетворяет? Существует ли единица этой операции и как она выглядит? Существуют ли обратные элементы и для каких элементов?

11.50 Пусть G — группа и непустое множество $H \subseteq G$. Доказать, что если множество H замкнуто относительно групповой операции и конечно, то оно является подгруппой группы G .

11.51 Доказать, что пересечение произвольного множества подгрупп группы G является подгруппой группы G .

11.52 Какое отношение эквивалентности соответствует разбиению группы G на: а) правые; б) левые классы смежности?

11.53 Доказать, что $m\mathbb{Z}$ является подгруппой в \mathbb{Z} с операцией сложения. Найти все смежные классы этой подгруппы. Найти индекс подгруппы $m\mathbb{Z}$ в группе \mathbb{Z} , т. е. вычислить $|\mathbb{Z} : m\mathbb{Z}|$.

11.54 Доказать, что \mathbb{Q} не содержит собственных подгрупп конечного индекса.

11.55 Доказать следующее обобщение теоремы Лагранжа. Пусть A и B — подгруппы группы G , и $A \subseteq B$. Индексы $|G : B|$ и $|B : A|$ конечны тогда и только тогда, когда $|G : A|$ конечен и $|G : A| = |G : B| \cdot |B : A|$.

11.56 Проверить, что $\{\text{id}, (12)(34), (13)(24), (14)(23)\}$ является подгруппой группы S_4 перестановок на 4 символах. Каков ее индекс в S_4 ? Найдите ее правые и левые смежные классы в группе S_4 .

11.57 В группе \mathbb{Z}_8^* найти обратные элементы: $3^{-1}, 5^{-1}, 7^{-1}$.

11.58 Доказать свойства степеней в группе:

а) $g^m g^n = g^{m+n}$;

б) $(g^m)^n = g^{mn}$.

11.59 Доказать, что если G — конечная группа, то $g^n = g^{n \bmod m}$, где m — порядок элемента g в группе G .

11.60 Доказать, что если группа $G = \langle S \rangle$ конечная, то любой элемент $g \in G$ имеет представление $s_1^{m_1} \dots s_n^{m_n}$ для некоторых $s_k \in S$, где все m_k — неотрицательные целые числа.

11.61 Доказать, что в конечной группе порядок всякого элемента делит порядок группы.

11.62 Доказать, что группа G четного порядка $|G| = 2n$ обязательно содержит элемент $g \neq e$ порядка 2. Указание: представить G как объединение пар g, g^{-1} .

11.63 Является ли циклической группа: а) $\langle \mathbb{Z}, + \rangle$; б) $\langle \mathbb{Z}, \cdot \rangle$?

11.64 Найдите порядок элемента R_φ в группе вращений окружности, если: а) $\varphi = \pi$; б) $\varphi = \pi/3$; в) $\varphi = \pi/6$; г) $\varphi = 32\pi/63$; д) $\varphi = \pi\sqrt{2}$; е) $\varphi = 0$.

11.65 Существует ли конечное множество образующих у следующих групп: а) группа движений прямой; б) группа движений окружности; в) группа движений правильного многогранника.

11.66 Доказать, что перестановочные элементы g, h произвольной группы G , имеющие взаимно простые порядки m, n , порождают в G циклическую подгруппу H порядка mn , где $H = \langle g, h \rangle = \langle gh \rangle$. Указание: использовать равенство $\text{НОД}(m, n) = km + ln$.

11.67 Показать, что $\mathbf{S}_n = \langle (12), (13), \dots, (1n) \rangle$.

11.68 Показать, что $\mathbf{S}_n = \langle (12), (123 \dots n) \rangle$.

11.69 Показать, что знакопеременная группа A_n , $n \geq 3$, порождается циклами длины 3, причем

$$A_n = \langle (123), (124), \dots, (12n) \rangle.$$

11.70 Показать, что порядок перестановки $\sigma \in \mathbf{S}_n$ (порядок циклической группы $\langle \sigma \rangle$) равен НОК длин независимых циклов, входящих в разложение σ .

11.71 Какую систему образующих можно предложить для мультипликативной группы $\langle \mathbb{Q}^+, \cdot \rangle$ положительных рациональных чисел? Указание: использовать ОТА.

11.72 Показать, что все группы порядка 4 абелевы. Указание: если для всех $x \in G$ имеем $x^2 = e$, то $abab = e$ и, следовательно, $ab = ba$.

11.73 Докажите, что для любых множеств A, B и C : а) $A \subseteq A$; б) если $A \subseteq B$ и $B \subseteq C$, то $A \subseteq C$; в) $A = B$ тогда и только тогда, когда $A \subseteq B$ и $B \subseteq A$.

11.74 Докажите, что: а) пустое множество является подмножеством любого множества; б) пустое множество единственно.

11.75 Для каждого из множеств $\{0\}, \emptyset, \{\emptyset\}, \{1, 2\}, \{1, 2, 3\}, \{\{1, 2, 3\}\}, \{\{1, 2\}, 3\}$ найдите количество его: а) элементов; б) подмножеств.

11.76 Может ли у множества быть ровно: а) 0; б) 7; в) 16 подмножеств?

11.77 Сколько подмножеств у множества, содержащего ровно n элементов?

11.78 Может ли у множества A быть ровно на 2022 подмножеств больше, чем у множества B ?

11.79 Пусть $a = \{b, c\}$, $b = \{\}$, $d = \{c, e\}$, $e = \{b\}$. Определите истинность высказываний: а) $a \in e$; б) $e \in b$; в) $b \in d$; д) $(a \in c \vee c \in d)$; е) $(b \in c \rightarrow a \in a)$; ф) $(e \in d \Leftrightarrow d \in e)$; г) $(e \in c \wedge a \in c)$; х) $\forall x x \notin b$; и) $\forall q (q \in c \rightarrow q \in a)$; ж) $\exists k k \in d$; к) $\forall s \exists t (s \in t)$; л) $\forall s \exists t (t \in s)$.

11.80 Пусть во всей вселенной есть только множества: $a = \{\}$, $b = \{d, a\}$, $c = \{a, e\}$, $d = \{a, e, c\}$, $e = \{a\}$, $f = \{a, b, c, d, e\}$. Чему может быть равно множество x , удовлетворяющее условию: **a)** $x \in e$; **b)** $x \notin f$; **c)** $d \subseteq x$; **d)** $f \subseteq x$; **e)** $\forall y y \notin x$; **f)** $x = \{\{\}\}$; **g)** $b = \{a, x\}$; **h)** $(x \in b \wedge x \notin d)$; **i)** $(x \in c \not\rightarrow x \in b)$; **j)** $\forall n n \notin x$?

Здесь символ \subset обозначает *собственное вложение*, т.е. $A \subset B$, если $A \subseteq B$ и $A \neq B$.

11.81 Пусть заданы следующие множества:

$$A = \{57, 91, 179, 239\}, \quad B = \{91, 239, 2014\}, \\ C = \{2, 57, 239, 2014\}, \quad D = \{2, 91, 2014, 2017\}.$$

Найдите:

a) $A \cup B$; **b)** $A \cap B$; **c)** $(A \cap B) \cup D$; **d)** $C \cap (D \cap B)$; **e)** $(A \cup B) \cap (C \cup D)$; **f)** $(A \cap B) \cup (C \cap D)$; **g)** $(D \cup A) \cap (C \cup B)$; **h)** $(A \cap (B \cap C)) \cap D$; **i)** $(A \cup (B \cap C)) \cap D$; **j)** $(C \cap A) \cup ((A \cup (C \cap D)) \cap B)$; **k)** $(A \cup B) \setminus (C \cap D)$; **l)** $(A \cup D) \setminus (B \cup C)$; **m)** $A \setminus (B \setminus (C \setminus D))$; **n)** $D \setminus ((B \cup A) \setminus C)$; **o)** $((A \setminus (B \cup D)) \setminus C) \cup B$.

11.82 Докажите, что для любых множеств A, B, C выполнены равенства: **a)** $A \cup A = A$, $A \cap A = A$; **b)** $A \cup B = B \cup A$, $A \cap B = B \cap A$; **c)** $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$; **d)** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

11.83 Верно ли, что для любых множеств A, B, C : **a)** $(A \setminus B) \cup B = A$; **b)** $A \setminus (A \setminus B) = A \cap B$; **c)** $A \setminus (B \setminus C) = (A \setminus B) \cup (A \cap C)$; **d)** $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$; **e)** $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$; **f)** $(A \setminus B) \cup (B \setminus A) = A \cup B$?

11.84 Приведите пример такого множества из четырех элементов, что для каждого двух его элементов один из них является элементом другого.

11.85 Сколько различных множеств можно получить из множеств A, B, C, D задачи 11.81 при помощи операций **a)** \cup, \cap, \setminus ; **b)** \cup, \cap ; **c)** \cup, \setminus ; **d)** \cap, \setminus ?

11.86 Из каких элементов состоят следующие множества: **a)** $\{0, 1\} \times \{9\}$; **b)** $\{0, 1\} \times \{0, 1\}$; **c)** $\emptyset \times \emptyset$; **d)** $\{5, 7\} \times \{1, 3, 17\}$; **e)** $\{16, 41\} \times \emptyset$?

11.87 Верно ли, что для всех множеств A, B, C, D выполняются равенства **a)** $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$; **b)** $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$?

11.88 Когда $A \times B = B \times A$?

11.89 Покажите, что $|A \times B| = |A| \cdot |B|$ для конечных множеств A и B .

11.90 (Принцип включения-исключения) а) Докажите, что

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|.$$

б) Сформулируйте и докажите аналогичное утверждение для n множеств.

11.91 Упорядоченной парой a и b по Куратовскому называется множество $\{\{a\}, \{a, b\}\}$. Докажите, что а) $\{\{a\}, \{a, b\}\} \neq \{\{b\}, \{b, a\}\}$, если $a \neq b$; б) $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$ тогда и только тогда, когда $a = c$ и $b = d$, т.е. такая конструкция реализует упорядоченную пару.

11.92 Постройте все возможные отношения на множестве: а) \emptyset ; б) $\{0\}$; в) $\{1, 2\}$. К какому виду относится каждое из полученных отношений?

11.93 Какова мощность: а) множества всех отношений, заданных на множестве мощности n ; б) множества всех отношений между множествами мощности m и n ?

11.94 Изучить рис. 11.3 с примерами отношений. Какой цвет и почему соответствует указанным отношениям? Функция $[x]$ обозначает целую часть числа.

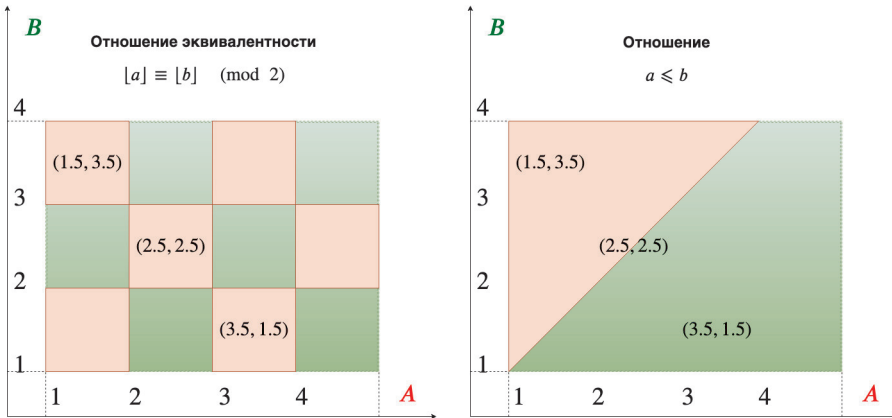


Рис. 11.3. Примеры бинарных отношений.

11.95 Докажите, что отношение сравнимости по любому заданному модулю является отношением эквивалентности на множестве целых чисел. Что представляют собой классы эквивалентности в этом случае?

11.96 Постройте фактормножество $\mathbb{Z}_9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ по отношению сравнимости по модулю 3.

11.97 Что представляет собой фактормножество произвольного множества X по отношению равенства? Рассмотрите также случай $X = \emptyset$.

11.98 Докажите что два класса эквивалентности либо не пересекаются, либо совпадают. Что можно сказать о сумме мощностей классов эквивалентности конечного множества?

11.99 **Разбиением множества** A называется такое множество непустых множеств $\{A_n\}$, что $A_i \cap A_j = \emptyset$ для всех i, j ($i \neq j$) и, кроме того, A есть объединение всех A_n . Докажите, что а) фактормножество является разбиением; б) всякому разбиению соответствует единственное отношение эквивалентности, фактормножество по которому совпадает с данным разбиением.

11.100 **Числом Бэлла** B_n для натурального n называется количество всех разбиений множества, состоящего из n элементов. Найдите B_5 и B_7 .

11.101 Найти факторгруппы: а) $\mathbb{Z}/n\mathbb{Z}$; б) $4\mathbb{Z}/12\mathbb{Z}$.

11.102 Доказать, что любая подгруппа абелевой группы нормальна.

11.103 Доказать, что $A_n \triangleleft S_n$.

11.104 Доказать, что $\mathbb{Z}/p\mathbb{Z}$ — простая группа, если p — простое число.

11.105 Доказать, что а) если $|G : H| = 2$, то $H \triangleleft G$; б) если $|G : H| = 2$, то $g^2 \in H$ для любого $g \in G$.

11.106 Доказать, что в группе \mathbb{Q}/\mathbb{Z} : а) каждый элемент имеет конечный порядок; б) для каждого натурального n имеется в точности одна подгруппа порядка n .

11.107 Пусть $H \triangleleft G$ и $f : G \rightarrow G/H$ — естественный гомоморфизм, т. е. $f(g) = gH$. Доказать, что f — эпиморфизм (т. е. сюръективный гомоморфизм), а также, что $\ker f = H$.

11.108 Найти все нетривиальные подгруппы S_3 . Какие из них являются нормальными? Каким группам \mathbb{Z}_m изоморфны факторы по данным нормальным подгруппам?

11.109 Показать, что группа Клейна $V_4 = \{\text{id}, (12)(34), (13)(24), (14)(23)\}$ является нормальной в S_4 .

11.110 Рассмотрим на прямой точки X, A, B, C, D , предполагая, что все они попарно различны. Обозначим за a разность точек $A - X$, которая равна длине отрезка XA , взятой со знаком направления (если A справа от X , то с плюсом, иначе — с минусом), а также $b = B - X$, $c = C - X$, $d = D - X$. Запишем *двойное отношение*

$$\lambda = [A, B; C, D] = \frac{(c - a)(d - b)}{(c - b)(d - a)}.$$

Докажите, что двойное отношение не зависит от выбора точки X .

11.111 Пусть точки A, B, C, D пронумерованы цифрами 1, 2, 3, 4. Через $\lambda(\sigma)$ обозначим двойное отношение $[\sigma(A), \sigma(B); \sigma(C), \sigma(D)]$, где $\sigma \in \mathbf{S}_4$, т. е. является перестановкой на 4 символах. Покажите, что

$$\lambda(\sigma) \in \Lambda = \left\{ \lambda, \frac{1}{\lambda}, 1 - \lambda, \frac{1}{1 - \lambda}, 1 - \frac{1}{\lambda}, \frac{\lambda}{1 - \lambda} \right\}$$

при любом σ .

11.112 Показать, что если на множестве Λ ввести операцию композиции (т. е. рассматривать его элементы как функции от аргумента λ и подставлять одну в другую), то получится группа, изоморфная \mathbf{S}_3 . Что в этой группе является единицей?

11.113 Покажите, что перестановки \mathbf{S}_4 , сохраняющие двойное отношение, — это перестановки группы Клейна V_4 .

11.114 Какой группе изоморфна группа \mathbf{S}_4/V_4 ? Как это связать со свойствами двойного отношения? Какой эпиморфизм $\mathbf{S}_4 \rightarrow \mathbf{S}_3$ можно задать с помощью двойного отношения?

11.115 Рассмотрим группу движений правильного n -угольника. Два движения будем называть эквивалентными, если их композиция является поворотом (или id). Докажите, что это и в самом деле отношение эквивалентности, постройте классы эквивалентности, постройте факторгруппу на этих классах. Какова ее таблица умножения?

11.116* Доказать, что если $f : G \rightarrow G'$ гомоморфизм и $g \in \ker f$ и $g \sim h$, то $h \in \ker f$.

11.117* Опишите перестановки, сопряженные перестановке (13)(24) в группе \mathbf{S}_4 .

11.118* Доказать, что при сопряжении при помощи а) перестановки (12); б) любой транспозиции; в) любой перестановки цикловая структура перестановки (т. е. количество циклов определенной длины) не меняется. д) Обратное: если перестановки имеют одинаковую цикловую структуру, то они сопряжены.

11.119* Опишите классы сопряженности в группе: а) \mathbf{S}_3 ; б) \mathbf{S}_4

11.120* Пусть $\sigma \in \mathbf{S}_n$. И пусть $\tilde{\sigma}(\tau) = \tau^\sigma = \sigma^{-1}\tau\sigma$. Поскольку $\tilde{\sigma}$ — автоморфизм \mathbf{S}_n , он также является биекцией на \mathbf{S}_n , т. е. $\tilde{\sigma} \in \mathbf{S}(\mathbf{S}_n)$.

а) Что представляет собой автоморфизм id?

б) Доказать, что $\tilde{\sigma}(\alpha \circ \beta) = \tilde{\sigma}(\alpha)\tilde{\sigma}(\beta)$.

с) Занумеруем элементы S_n числами от 1 до $n!$ в каком-либо порядке. Тогда автоморфизму $\tilde{\sigma}$ можно взаимно однозначно сопоставить перестановку из $S_{n!}$, которую мы обозначим $\varphi_n(\sigma)$. Ясно, что, вообще говоря, φ_n зависит от выбранного способа нумерации элементов S_n .

Доказать, что соответствие $\sigma \mapsto \varphi(\sigma)$ является гомоморфизмом из S_n в $S_{n!}$ при любом выборе нумерации элементов S_n .

11.121* Гомоморфизм $\varphi(\sigma) : S_n \rightarrow S_{n!}$, определенный в предыдущей задаче, можно рассматривать не на всем S_n , а только на перестановках, входящих в один класс сопряженности K . Доказать, что сужение $\varphi|_K$ является гомоморфизмом из S_n в $S_{|K|}$, где $|K|$ — количество элементов множества K .

11.122* Пусть $K = (12)^{S_3}$. Доказать, что $\varphi|_K : S_3 \rightarrow S_3$ — изоморфизм.

11.123* Пусть $K = (123)^{S_3}$. Доказать, что $\varphi|_K : S_3 \rightarrow S_2$ — отображение знака, т. е. четным перестановкам сопоставляется id , а нечетным — транспозиция.

11.124* Найти ядра гомоморфизмов из задач 11.122* и 11.123*.

11.125* Найти ядро гомоморфизма $\varphi|_K : S_4 \rightarrow S_{24}$, если $K = ((12)(34))^{S_4}$.

11.126* Доказать, что если $f : S_n \rightarrow S_m$ — гомоморфизм и $(12) \in \ker f$, то $\ker f = S_n$.

11.127* Несколько жителей города N хотят обменяться квартирами. У каждого есть по квартире, но каждый хочет переехать в другую (разные люди хотят переехать в разные квартиры). По законам города разрешены только парные обмены: если два человека обмениваются квартирами, то в тот же день они не участвуют в других обменах. Докажите, что можно устроить парные обмены так, что уже через два дня каждый будет жить в той квартире, куда хотел переехать.

11.128* Дана произвольная текстовая строка длины n . Как определить, что с помощью перестановки символов из нее можно сделать *палиндром* (строку, которая одинаково читается слева направо и справа налево)? Использовать перебор по всем перестановкам запрещается. Скорость алгоритма должна быть $O(n)$.

Комплексная арифметика и алгебра

Аннотация

В этой главе мы начинаем строить поле комплексных чисел, пока еще без участия вещественных. По сути, мы здесь работаем только с комплексными рациональностями, что, однако, не мешает показать тесную геометрическую связь комплексных чисел и движений плоскости, а также изучить гауссовы числа.

12.1. Алгебра комплексных чисел

Когда мы строили поле $\mathbb{Q}[\sqrt{2}]$, мы ввели в обращение новое число, которое позволяло решать уравнение $x^2 = 2$. Это число не является рациональным, но лежит где-то между рациональными числами. Тогда же мы задались вопросом, как быть с поиском корней других уравнений с целыми коэффициентами, неразрешимых в \mathbb{Q} .

Рассмотрим еще один пример уравнения: $x^2 = -1$. Вроде бы все коэффициенты — целые числа, и степень всего лишь вторая. Однако при детальном рассмотрении становится ясно, что у него нет решений не только в рациональных числах, но и где-то между ними, поскольку никакое известное нам число, возведенное в квадрат, и близко не подходит к -1 . Действительно, когда мы ищем, например, решение уравнения $x^2 = 2$, мы можем найти среди рациональных чисел такие дроби m/n , что их квадрат m^2/n^2 будет сколь угодно близок к числу 2, причем как сверху, так и снизу. Но какие бы рациональные числа мы не выбрали в качестве x , величина x^2 будет удалена от числа -1 как минимум на 1. Из чего мы можем сделать вывод, что корень уравнения $x^2 = -1$ не может находиться где-то между рациональными числами.

Стало быть, если мы хотим ввести в обращение корень такого уравнения, то его необходимо поместить где-то вне числовой оси, «подвесить в воздухе».

Сделаем это из чисто эстетико-геометрических соображений. Как геометрически проявляют себя числа на прямой? Они обеспечивают сдвиг вдоль прямой: положительные — вправо, отрицательные — влево. Причем у всех этих сдвигов есть единица измерения — число 1, которое заодно выступает и в роли мультипликативной единицы, когда мы определяем умножение чисел. Кроме того, сдвиг на 1 вправо и затем влево (или в обратном порядке) приводит нас обратно, т.е. является сдвигом на 0, т.е. совпадает с id .

Новое же число мы хотим поместить так, чтобы оно обеспечивало сдвиг на плоскости, аналогичный сдвигу вдоль прямой.

Поскольку мы привыкли считать направление «вверх» положительным, поместим это число над числовой осью.

Заложим в этом числе сразу и единицу измерения: пусть оно отстоит от нуля на расстояние 1, поскольку его квадрат равен 1, тем самым мы согласуем масштаб сдвигов на плоскости со сдвигами на прямой. Наконец, сдвиг в направлении и на величину этой новой единицы не должен содержать в себе горизонтальных сдвигов, их проще добавить потом, взяв от сдвигов прямой, которые нам уже известны. Иначе говоря, числовая прямая при сдвиге на эту новую единицу должна сдвинуться вверх на расстояние 1 и таким образом, чтобы ее числовая разметка никуда не сдвинулась вправо или влево.

Так мы приходим к тому, что новую единицу сдвига следует отложить от нуля строго вверх на расстояние 1.

На координатной сетке она окажется в точке $(0, 1)$.

Назовем это новое число-вектор буквой i , которую принято называть **мнимой единицей** (от фр. *imaginaire*).

Теперь всякий сдвиг плоскости мы можем записать как композицию сдвига, выраженного в единицах (горизонтальный сдвиг), и сдвига, выраженного в мнимых единицах (вертикальный сдвиг).

Иначе говоря, сдвиг на произвольный вектор z мы распишем как сдвиг на сумму векторов $x \cdot \vec{1} + y \cdot \vec{i}$. См. рис. 12.1.

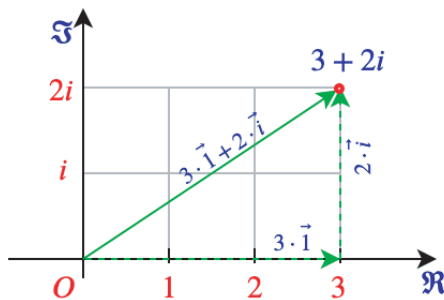


Рис. 12.1. Сдвиг как комплексное число.

Как и прежде, мы умеем различать на плоскости векторы и точки. Векторы — это направленные отрезки, которые можно откладывать от точек. Сложение векторов означает их последовательное откладывание. В результате таких откладываний мы уходим от некоторой стартовой точки и приходим в какую-то финишную точку. Результирующий вектор соединяет стартовую и финишную точки. Договоримся для удобства считать стартовой точкой начало координат O , а финишную точку обозначать почти так же, как вектор, который в нее входит, только без векторной символики.

Итак, если вектор равен $x \cdot \vec{1} + y \cdot \vec{i}$, то его финишная точка обозначается $x + yi$.

Пока все, что мы сделали, — это построили обычную арифметику векторов на плоскости. При чем же тут алгебраическая ипостась мнимой единицы, вытекающая из уравнения $x^2 = -1$?

Алгебраическая ипостась i нам нужна как раз для того, чтобы построить алгебру точек плоскости, т. е. научиться их не только складывать и умножать на число, но еще и умножать и делить друг на друга.

Примем за аксиому, что с числами вида $x + iy$ мы будем обращаться как с обычными числами, пользуясь аксиомами поля, и при этом пользоваться тем самым свойством мнимой единицы, которое ее определяет, т. е. равенством $i^2 = -1$.

Положим

$$(a + bi)(x + yi) = ax + ayi + bxi + byi^2 = (ax - by) + (ay + bx)i. \quad (12.1)$$

Числа вида $z = x + iy$ с заданными операциями сложения и умножения (сложение — покоординатное, а умножение определено выше) называются **комплексными числами**. При этом x называется **действительной** (или вещественной) частью комплексного числа z и обозначается за $\operatorname{Re} z$, а y называется **мнимой** частью числа z и обозначается за $\operatorname{Im} z$.

Координатная ось Ox на комплексной плоскости называется действительной осью, а координатная ось Oy — мнимой.

Дадим следующие определения. Число $\bar{z} = x - yi$ называется **комплексно сопряженным** к числу $z = x + iy$. Комплексное сопряжение — это отражение относительно действительной оси.

Модулем комплексного числа $z = x + yi$ называется число

$$|z| = \sqrt{x^2 + y^2}.$$

Нетрудно видеть, что модуль комплексного числа — это длина соответствующего ему вектора (по теореме Пифагора). Кроме того, из геометрических соображений понятно, что $|z_1 - z_2|$ — это расстояние между точками z_1 и z_2 на плоскости (см. рис. 12.2).

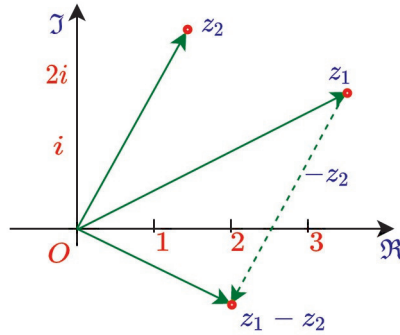


Рис. 12.2. Расстояние между комплексными числами.

Посмотрим, какие арифметические свойства комплексных чисел можно вывести из этих определений.

- 1) Умножение ассоциативно: $z_1(z_2 z_3) = (z_1 z_2)z_3$, что проверяется непосредственно по формуле (12.1).
- 2) Умножение коммутативно: $z_1 z_2 = z_2 z_1$, что следует из определения умножения по формуле (12.1), а также из коммутативности умножения действительных чисел.
- 3) Выполняется дистрибутивный закон: $(z_1 + z_2)z_3 = z_1 z_3 + z_2 z_3$, что также легко проверить, пользуясь (12.1) и тем, что сложение определено как сложение векторов, т. е. по координатам.
- 4) $z\bar{z} = |z|^2$. Действительно, $(x + yi)(x - yi) = x^2 + y^2$.
- 5) $z = 0$ (т. е. $z = 0 + 0i$) тогда и только тогда, когда $|z| = 0$. Это очевидно, ибо $x^2 + y^2 = 0 \Leftrightarrow x = y = 0$.
- 6) Обратное по умножению число для $z \neq 0$ существует и равно

$$z^{-1} = \frac{1}{x + yi} = \frac{x - yi}{(x + yi)(x - yi)} = \frac{\bar{z}}{|z|^2}.$$

Это можно получить и напрямую из свойства 4).

- 7) Мультипликативное свойство сопряжения: $\overline{z\bar{w}} = \bar{z} \cdot w$. Действительно,

$$\overline{(x + yi)(a + bi)} = \overline{(ax - by) + (ay + bx)i} = (ax - by) - (ay + bx)i$$

и

$$\overline{(x + yi)} \cdot \overline{(a + bi)} = (x - yi)(a - bi) = (ax - by) - (ay + bx)i.$$

- 8) Мультипликативное свойство модуля: $|zw| = |z||w|$. Действительно,

$$|zw|^2 = zw\bar{z}\bar{w} = zw\bar{z}\bar{w} = z\bar{z}w\bar{w} = |z|^2|w|^2.$$

Сложение с числом $z = x + iy$ — это параллельный перенос T_z на вектор $z = x\vec{1} + y\vec{i}$. Это следует из геометрических свойств комплексных чисел, о которых мы говорили выше.

Кроме того, это легко проверить арифметически. Пусть даны две точки z_1 и z_2 . Добавим к ним вектор z , получим новые точки $z'_1 = z_1 + z$ и $z'_2 = z_2 + z$. Во-первых, заметим, что расстояние сохранилось:

$$|z'_1 - z'_2| = |(z_1 + z) - (z_2 + z)| = |z_1 - z_2|,$$

т.е. прибавление z — это движение. Во-вторых, если $z \neq 0$, то у этого движения нет неподвижных точек, иначе мы бы получили равенство $z_1 + z = z_1$, откуда $z = 0$. Следовательно, в силу теоремы Шаля прибавление z есть параллельный перенос. Прибавление $z = 0$ есть id.

Умножение на комплексное число, по модулю равное 1, есть поворот с центром в нуле.

Пусть $|z| = 1$. Возьмем точки $w_1 = a_1 + b_1i$ и $w_2 = a_2 + b_2i$, умножим их на z , получим точки $w'_1 = w_1z$ и $w'_2 = w_2z$.

Найдем расстояние между w'_1 и w'_2 :

$$|w'_1 - w'_2| = |(w_1 - w_2)z| = |w_1 - w_2| \cdot |z| = |w_1 - w_2|,$$

т.е. умножение на z сохраняет расстояние. В то же время, очевидно, что при $z \neq 1$ единственной неподвижной точкой при умножении будет $w = 0$, иначе мы бы получили $wz = w$, т.е. $z = w/w = 1$. Умножение на $z = 1$ есть id.

Итак, умножение на число z , по модулю равное 1, является поворотом с центром в нуле. *Каков при этом угол поворота?*

Чтобы ответить на данный вопрос, рассмотрим для начала случай $|w| = 1$, т.е. точку с единичной окружности будем умножать на другую точку с единичной окружности. По свойствам модуля имеем $|zw| = |z||w| = 1$, т.е. в результате умножения мы вновь получим точку на единичной окружности! Иначе говоря, единичная окружность с операцией умножения комплексных чисел образует группу.

Теперь заметим, что на окружности радиуса 1 хорда однозначно определяет опирающийся на нее угол. Рассмотрим углы, которые опираются на хорду $[z; 1]$ и на хорду $[zw; w]$. На рисунке 12.3 они выделены красным цветом.

Легко видеть, что длины хорд равны: $|zw - w| = |z - 1||w| = |z - 1|$, так что и углы равны. Следовательно, точка zw получается из точки w поворотом на угол, соответствующий углу наклона вектора z относительно положительного направления действительной оси.

Что происходит в случае, когда w не лежит на единичной окружности и отлична от нуля? Для этого представим произведение zw следую-

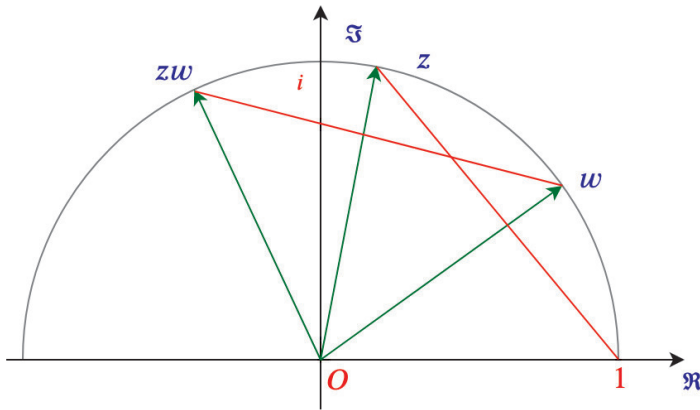


Рис. 12.3. Поворот как умножение комплексных чисел.

щим образом:

$$zw = z \frac{w}{|w|} |w|,$$

где отношение $w/|w|$ уже является комплексным числом единичной длины. Следовательно, число zw также является результатом поворота точки w на угол наклона вектора z относительно положительного направления действительной оси.

Угол, заданный числом z , а в общем случае, числом $z/|z|$ (если z — произвольное ненулевое комплексное число), называется **аргументом числа** z и обозначается $\arg z$.

Основные тригонометрические функции определяются с помощью комплексного числа с единичной окружности так: пусть задан угол φ . Повернем вектор $(1, 0)$ на этот угол и найдем число z на единичной окружности такое, что $\arg z = \varphi$, тогда

$$\cos \varphi = \operatorname{Re} z, \quad \sin \varphi = \operatorname{Im} z.$$

Как уже отмечалось выше, операция комплексного сопряжения есть не что иное как отражение относительно действительной оси. Таким образом, рассмотренные нами операции с комплексными числами обеспечивают следующие виды движений плоскости: сдвиг на произвольный вектор, поворот относительно начала координат, отражение относительно действительной оси.

Учитывая, что поворот с произвольным центром можно представить как композицию сдвига, поворота с центром в нуле и обратного сдвига, а отражение относительно произвольной оси — как композицию поворота или сдвига, отражения относительно действительной оси и обратного поворота или сдвига, мы приходим к тому, что все движения плоскости можно выразить через три изученных нами действия

с комплексными числами: сложение (произвольный сдвиг), умножение на число единичной длины (поворот с центром в нуле) и сопряжение (отражение относительно действительной оси).

На будущее у нас остается вопрос: *какое преобразование плоскости осуществляет умножение на произвольное ненулевое комплексное число?*

Поскольку мы пока знакомы только с рациональными дробями, комплексные числа у нас также являются рациональными, т. е. имеют вид $\frac{a}{b} + \frac{c}{d}i$, где $a, b, c, d \in \mathbb{Z}$ и $b, d \neq 0$. Но даже при таком существенном ограничении мы уже имеем дело с еще одним полем — **полем комплексных рациональностей**, поскольку сложение, вычитание, умножение и деление не выводит нас за пределы этого множества (правда, модуль числа может выпасть из \mathbb{Q} , в отличие от обычных рациональных чисел). Такое поле обозначается $\mathbb{Q}[i]$ и является расширением поля \mathbb{Q} , аналогично полю $\mathbb{Q}[\sqrt{2}]$, рассмотренному ранее.

12.2. Гауссовы целые числа

В этом разделе мы ограничимся рассмотрением комплексных чисел с целыми координатами, т. е. чисел вида

$$a + bi, \quad a, b \in \mathbb{Z}.$$

Легко видеть, что такие числа образуют коммутативное кольцо с единицей. Данное кольцо обозначается $\mathbb{Z}[i]$ и называется кольцом **гауссовых целых чисел**. На координатной плоскости точки $\mathbb{Z}[i]$ сосредоточены в узлах целочисленной решетки.

Число $a + bi$ **делится на** $c + di$, если существует число $a' + b'i$ такое, что $a + bi = (c + di)(a' + b'i)$. Обозначение аналогично обычному в натуральных числах: $(c + di) \mid (a + bi)$. Например, число 2 делится на $(1 + i)$, т. к. $2 = (1 + i)(1 - i)$.

Нормой гауссова числа $a + bi$ называется величина

$$N(a + bi) = (a + bi)(a - bi) = a^2 + b^2,$$

т. е. норма числа z равна $z\bar{z}$.

Несколько свойств нормы:

Norm1 $N(a + bi) = 0$ тогда и только тогда, когда $a = b = 0$.

Norm2 Нормы комплексно сопряженных чисел совпадают.

Norm3 Если норма нечетна, то она имеет вид $4k + 1$. Никакая норма не может быть равна $4n + 3$.

Так как $N(a + bi) = a^2 + b^2$, то легко видеть, что норма является нечетным числом только в том случае, когда либо a четное, b нечетное, либо наоборот. Пусть $a = 2k$, $b = 2j + 1$, тогда $a^2 + b^2 = 4k^2 + 4j^2 + 4j + 1 = 1 \pmod{4}$. Аналогично и наоборот.

Norm4 $N(zw) = N(z)N(w)$, где z, w — гауссовы числа.

Пусть $z = a + bi$, $w = c + di$, тогда

$$N(zw) = N(ac - bd + (ad + bc)i) = a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2,$$

$$N(z)N(w) = (a^2 + b^2)(c^2 + d^2) = a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2.$$

Последнее свойство означает, что обратимыми элементами могут быть только числа с нормой 1, т.е. ± 1 и $\pm i$. Других обратимых нет. Геометрически это можно охарактеризовать так: обратимыми элементами являются те и только те гауссовы числа, которые лежат на единичной окружности. Отметим также, что множество всех обратимых элементов совпадает с множеством всех корней четвертой степени из 1, т.е. корней уравнения $x^4 = 1$, т.е. с множеством $\{\pm 1, \pm i\}$.

Наконец, множество $\{\pm 1, \pm i\}$ является группой по умножению, причем уже хорошо знакомой нам группой, если не обращать внимание на символ операции и символы элементов группы. Сравните таблицу умножения этой группы с таблицей сложения группы вычетов по модулю 4:

*	1	i	-1	-i
1	1	i	-1	-i
i	i	-1	-i	1
-1	-1	-i	1	i
-i	-i	1	i	-1

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Если произвести соответствие $1 \mapsto 0$, $i \mapsto 1$, $-1 \mapsto 2$, $-i \mapsto 3$, а операции умножения поставить в соответствие операции сложения по модулю 4, то мы получим полное соответствие между результатами умножения в первой группе и сложения во второй: $i(-1) \mapsto 1 + 2$ и т.д.

В том случае, когда можно предъявить взаимно однозначное соответствие элементов двух групп так, чтобы операция в первой группе соответствовала операции во второй, говорят о том, что эти две группы **изоморфны**. Очень часто такие группы даже считают равными, хотя природа у них разная. Итак, группа по умножению обратимых гауссовых чисел изоморфна группе $\mathbb{Z}/4\mathbb{Z}$.

Все гауссовы числа делятся на обратимые. Это легко понять из групповых свойств делителей единицы. Действительно, разделить на 1 означает умножить на нее, т.к. 1 сама себе обратна по умножению. Разделить на i означает умножить на $-i$, т.к. эти числа взаимно обратны по

умножению. Аналогично, разделить на -1 означает умножить на -1 , и разделить на $-i$ означает умножить на i .

Обратимые элементы обладают еще одним замечательным свойством: умножение на них — это поворот относительно начала координат, причем умножение на i есть поворот на угол $\pi/2$, умножение на -1 — поворот на угол π (т.е. центральная симметрия), умножение на $-i$ — поворот на угол $3\pi/2$ или $-\pi/2$. То есть обратимые элементы образуют еще и группу вращений квадрата.

Два гауссовых числа называют **ассоциированными**, если одно получается из другого умножением на обратимый элемент. Ассоциированность является отношением эквивалентности, причем каждый класс эквивалентности включает ровно 4 числа, расположенных в углах квадрата с центром в 0. Например, $1 + 2i$, $-2 + i$, $-1 - 2i$ и $2 - i$ ассоциированы (см. рис. 12.4).

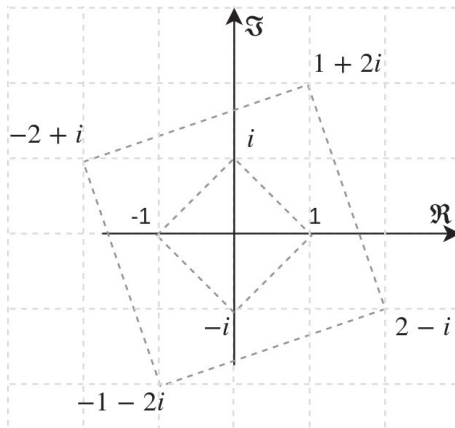


Рис. 12.4. Ассоциированные гауссовы числа.

Свойства делимости гауссовых чисел очень похожи на таковые свойства в арифметике натуральных чисел, но есть и отличия. Приведем несколько свойств делимости гауссовых чисел.

Div1 Если гауссово число $a + bi$ делится на обычное целое число $c + i0$, то $c | a$ и $c | b$ в целых числах. Верно и обратное.

Это легко видеть из равенства $a + bi = (x + yi)(c + i0) = xc + yci$, откуда $a = xc$ и $b = yc$.

Div2 Если $z | w$ и $w | z$, то z и w ассоциированы.

Пусть $w = z'z$ и $z = w'w$, откуда $N(w) = N(z')N(z)$ и $N(z) = N(w')N(w)$, откуда $N(z')N(w') = 1$ и, следовательно, $N(z') = N(w') = 1$, поскольку в натуральных числах это единственное решение. Стало быть, z' и w' — обратимые элементы.

Div3 Ассоциированность сохраняет делимость: если z и w ассоциированы, u и v ассоциированы, то $(z|u) \Leftrightarrow (w|v)$.

Действительно, пусть $z = z'w$ и $u = u'v$, где z', u' — обратимые элементы. Тогда

$$\frac{u}{z} = \frac{u'v}{z'w},$$

так что отношение u/z является гауссовым числом тогда и только тогда, когда отношение v/w является гауссовым числом.

Div4 z ($N(z) > 1$) имеет как минимум 8 делителей: своих ассоциированных и ассоциированных с 1.

Div5 Делители z являются делителями $N(z)$, если $N(z)$ рассматривать как гауссово число.

Пусть $w|z$, т.е. $z = uw$. Поскольку $N(z) = z\bar{z} = w(u\bar{z})$, очевидно, что $N(z)$ делится на w .

Div6 Норма $z = a + bi$ четна тогда и только тогда, когда $(1 + i)|z$. В частности, если a и b имеют разную четность, то z не делится на $1 + i$.

Для начала заметим, что норма $z = a + bi$ четна тогда и только тогда, когда a и b имеют одинаковую четность, т.е. сравнимы по модулю 2. Далее, поскольку $(1 + i)|z$, существует $c + di$ такое, что

$$a + bi = (1 + i)(c + di) = c - d + i(c + d),$$

т.е. $a = c - d$, $b = c + d$, что равносильно $a - b = -2d$, $a + b = 2c$ при некоторых $c, d \in \mathbb{Z}$, а это равносильно тому, что $a \equiv b \pmod{2}$.

В кольце $\mathbb{Z}[i]$ можно любое число u разделить на любое число $v \neq 0$ с остатком, так что получится

$$u = qv + r, \quad N(r) < N(v). \quad (12.2)$$

При этом выбор чисел q и r можно строго ограничить, выбирая q как ближайшее гауссово число к комплексному $u/v \in \mathbb{Q}[i]$, а r как разность между u и qv . В случае, когда выбор q неоднозначен (может быть максимум 4 числа), можно договориться выбирать такое число, которое на координатной сетке находится левее и/или ниже.

Приведем один из вариантов вычисления r . Пусть $u = a + bi$ и $v = c + di$. Далее оперируем в поле $\mathbb{Q}[i]$:

$$\frac{a + bi}{c + di} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} = q_1 + \frac{r_1}{c^2 + d^2} + q_2i + \frac{r_2}{c^2 + d^2}i,$$

где $ac + bd = q_1(c^2 + d^2) + r_1$ и $bc - ad = q_2(c^2 + d^2) + r_2$. Здесь мы воспользовались делением с остатком в кольце \mathbb{Z} . При этом мы выбираем знаки r_1 и r_2 так, чтобы выполнялись неравенства:

$$|r_1|, |r_2| \leq (c^2 + d^2)/2.$$

Это всегда возможно, поскольку остатки от деления можно выбирать не только из ряда $0, 1, 2, \dots, c^2 + d^2 - 1$, но также из ряда $0, \pm 1, \pm 2, \dots, \pm k$, где k — целая часть от деления $c^2 + d^2$ на 2, так что всегда $k \leq (c^2 + d^2)/2$. Здесь как раз и может оказаться вплоть до 4-х вариантов выбора.

Тогда

$$u = (q_1 + q_2i)v + \frac{(r_1 + r_2i)(c + di)}{c^2 + d^2} = (r_1 + r_2i) \frac{v}{N(v)}.$$

Эту последнюю дробь мы и выберем в качестве остатка r .

При этом заметим, что поскольку разность $u - (q_1 + q_2i)v$ является гауссовым числом, то таковым же будет и число $(r_1/N(v) + r_2i/N(v))v$, хоть оно и выглядит нецелым.

Далее,

$$N\left(\frac{r_1}{N(v)} + \frac{r_2}{N(v)}i\right) = \frac{r_1^2 + r_2^2}{(c^2 + d^2)^2} \leq \frac{1}{2},$$

откуда $N(r) \leq (1/2)N(v) < N(v)$.

На основе деления с остатком нетрудно вывести выполнимость алгоритма Евклида (см. раздел 4.2) для гауссовых чисел:

$$\begin{aligned} u &= q_1v + r_1, & N(r_1) &< N(v), \\ v &= q_2r_1 + r_2, & N(r_2) &< N(r_1), \\ r_1 &= q_3r_2 + r_3, & N(r_3) &< N(r_2), \\ &\dots\dots\dots & & \\ r_{n-1} &= q_{n+1}r_n + r_{n+1}, & N(r_{n+1}) &< N(r_n), \\ r_n &= q_{n+2}r_{n+1}, & & \end{aligned} \tag{12.3}$$

поскольку норма является натуральным числом и не может убывать бесконечно.

Отсюда — так же, как для целых чисел, — выводится и представление $\text{НОД}(u, v)$ в виде линейной комбинации исходных чисел u, v . Но мы докажем этот факт иным способом.

Лемма 12.1. *Для любых гауссовых чисел $u, v \neq 0$ существует гауссово число r такое, что:*

- 1) $r \mid u$ и $r \mid v$ (общий делитель);
- 2) если $(q \mid u) \wedge (q \mid v)$, то $q \mid r$ (наибольший общий делитель);
- 3) существуют гауссовы x, y такие, что $r = xu + yv$.

Кроме того,

4) число r , удовлетворяющее 1)–3), единственное с точностью до ассоциированности.

Доказательство. Пусть $R(u, v) = \{xu + yv \mid x, y \in \mathbb{Z}[i] \setminus \{0\}\}$. В множестве норм $\{N(z) \mid z \in R(u, v)\}$ существует наименьшее положительное число (т.к. ноль мы сразу исключили). Пусть $r \in R(u, v)$ есть такое число, у которого норма минимальная (оно может быть не единственным, тогда выберем одно из них). Остается показать, что r — искомое число.

Во-первых, r имеет вид $xu + yv$ по построению, т.е. выполняется пункт 3). Во-вторых, если $q \mid u$ и $q \mid v$, то очевидно, что $q \mid r$ также по построению r , т.е. выполняется пункт 2).

Докажем пункт 1). Из (12.2) имеем: $u = rt + s$, где $N(s) < N(r)$. Подставляя представление r , имеем: $u = xut + yvt + s$, откуда $s = (1 - xt)u + (-yt)v$. Если $s \neq 0$, то $s \in R(u, v)$ как линейная комбинация u и v , но тогда $N(s) \geq N(r)$ в силу выбора r , а это не так в силу (12.2). Следовательно, $s = 0$, откуда $r \mid u$. Аналогично, $r \mid v$.

Докажем пункт 4). Пусть $r' = x'u + y'v$ также удовлетворяет свойствам 1)–3). Тогда $r \mid r'$ и $r' \mid r$. Как мы видели выше (свойство Div2), это ведет к ассоциированности r и r' . \square

Доказанная лемма позволяет определить понятие НОД для гауссовых чисел с точностью до ассоциированности. В качестве НОД мы будем выбирать какое-то одно из четырех чисел (наиболее удобного вида).

Гауссово число называется **простым**, если оно не имеет никаких делителей, кроме тривиальных (ассоциированных с 1 и самим собой), и не является делителем 1, т.е. простое гауссово число имеет ровно 8 делителей. Два гауссовых числа называются **взаимно простыми** (обозначается $u \perp v$), если их НОД — обратимое число, т.е. 1 и ассоциированные с ней числа $i, -1, -i$.

Верны следующие свойства простых гауссовых чисел.

Prim1 Если $a + bi$ простое, то $a - bi$ также простое.

Действительно, если $a - bi = uv$, где u, v не ассоциированы с 1, то $a + bi = \bar{u}\bar{v}$, где \bar{u} и \bar{v} не ассоциированы с 1 (т.к. для делителей единицы сопряжение не нарушает ассоциированности), но тогда $a + bi$ не является простым.

Prim2 Если z простое, то его ассоциированные также простые. (Очевидно).

Prim3 Если z простое и $z \mid uv$, то $(z \mid u) \vee (z \mid v)$.

Действительно, пусть простое $z \mid uv$. Предположим, что $\neg(z \mid u)$, тогда $z \perp u$, откуда по лемме 12.1 получаем, что $1 = xz + yu$. Умножаем на v : $v = xzv + yuv$. Справа оба слагаемых делятся на z , следовательно, $z \mid v$. Аналогично, если $\neg(z \mid v)$, то $z \mid u$.

Prim4 Норма простого числа, неассоциированного с $1 + i$, всегда нечетна, т.е. имеет вид $4k + 1$.

Это следует из свойства Div6. Если простое число не ассоциировано с $1 + i$, то оно и не делится на него, а значит, по свойству Div6 его норма нечетная. То, что она имеет вид $4k + 1$, следует из свойства норм Norm3.

Prim5 Натуральное простое не всегда есть гауссово простое: $5 = (2 + i)(2 - i)$.

Prim6 Простое натуральное $4k + 1$ можно представить как сумму квадратов $a^2 + b^2$ (**рождественская теорема Ферма**).

Доказательство. Рассмотрим факториал $(p - 1)!$ в арифметике по модулю p . Поскольку $-1 \equiv p - 1$, $-2 \equiv p - 2$ и т.д., а всего множителей $p - 1 = 4k$, то все они разбиваются на пары вида $1, -1, 2, -2, \dots, (p - 1)/2, -(p - 1)/2$, откуда

$$(p - 1)! \equiv 1(-1)2(-2) \dots \frac{p - 1}{2} \frac{-p + 1}{2} \equiv (-1)^{(p-1)/2} \left(\frac{p - 1}{2}! \right)^2 \equiv \left(\frac{p - 1}{2}! \right)^2 \pmod{p},$$

поскольку $(p - 1)/2 = 2k$ — четное число. С другой стороны, по теореме Вильсона 10.4 $(p - 1)! \equiv p - 1 \pmod{p}$, так что

$$p - 1 \equiv \left(\frac{p - 1}{2}! \right)^2 \pmod{p},$$

то есть $p - 1$ сравнимо с квадратом некоторого числа c , откуда следует, что $c^2 + 1$ делится на p .

Теперь переходим в числа Гаусса: $c^2 + 1 = (c + i)(c - i)$. Если число p — простое в гауссовых числах, то в силу ОТА либо $c + i$ делится на p , либо $c - i$ делится на p , тогда по свойству Div1 число 1 делится на p , что невозможно. Следовательно, p — не простое гауссово число, а значит,

$$p = (a + bi)(x + yi),$$

где оба множителя нетривиальны. В то же время p — не вещественное число, т. е. $p = \bar{p}$, т. е.

$$p = (a - bi)(x - yi).$$

Наконец, норма p будет равна

$$p\bar{p} = p^2 = (a^2 + b^2)(x^2 + y^2).$$

А теперь возвращаемся в обычные натуральные числа, поскольку слева и справа именно они. Число p — простое, стало быть, его квадрат

в силу ОТА раскладывается единственным образом — произведение p и p , откуда

$$p = (a^2 + b^2) = (x^2 + y^2),$$

что и завершает доказательство. \square

Prim7 Рождественская теорема Ферма также выводится из критерия Гаусса того, что целое комплексное число является простым гауссовым числом. Доказательство этого критерия мы оставим за рамками курса.

Теорема 12.1 (Критерий Гаусса). $a + bi$ простое тогда и только тогда, когда

- 1) либо одно из чисел a, b нулевое, а второе — простое целое число вида $\pm(4k + 3)$ ($k > 0$),
- 2) либо a, b ненулевые и норма $N(a + bi) = a^2 + b^2$ — простое натуральное число.

Следствие 12.1. Простое натуральное вида $4k + 1$ не может быть простым гауссовым, простые натуральные вида $4k + 3$ являются простыми гауссовыми.

Prim8 Если $N(z) \perp N(w)$ в натуральных числах, то $z \perp w$ в гауссовых числах.

Действительно, пусть $u = \text{НОД}(z, w)$ в гауссовых числах. Тогда $z = ut$, $w = ut'$ и $N(z) = N(u)N(t)$, $N(w) = N(u)N(t')$. Откуда $N(u) \mid N(z)$ и $N(u) \mid N(w)$. Тогда из условия $N(z) \perp N(w)$ следует, что $N(u) = 1$, т. е. u — ассоциированное с 1 гауссово число, т. е. $z \perp w$.

Обратное, конечно, не верно.

Примеры простых гауссовых чисел: $\pm 3, \pm 7, \pm 3i, 1 \pm i, 1 \pm 2i, 1 \pm 4i$.

Для гауссовых чисел существует аналог основной теоремы арифметики (см. раздел 4.3):

Теорема 12.2 (Основная теорема арифметики для гауссовых чисел).

Если z — гауссово число с нормой $N(z) > 1$, то существуют простые гауссовы числа $\alpha_1, \dots, \alpha_n$ такие, что

$$z = \alpha_1 \dots \alpha_n,$$

и для любого другого представления z в виде произведения простых гауссовых чисел β_1, \dots, β_m справедливы утверждения:

- (i) $m = n$;
- (ii) существует перестановка σ такая, что α_k ассоциировано с $\beta_{\sigma(k)}$, $k = 1, \dots, n$.

Иначе говоря, разложение z в произведение простых гауссовых чисел единственно с точностью до отношения ассоциированности и порядка множителей.

Доказательство теоремы прямо следует из свойства Prim3.

Пример: $5 = (2 + i)(2 - i) = (1 + 2i)(1 - 2i)$ (множители переводятся друг в друга умножением на i и на $-i$).

Из ОТА легко выводится следующее утверждение.

Лемма 12.2. Если $u \perp v$ и $uv = c^n$, то существуют $a \perp b$ такие, что $u = a^n \delta_1$, $v = b^n \delta_2$, $c = ab$, где δ_1, δ_2 — обратимые гауссовы числа.

Заметим, что и в обычной арифметике целых чисел верна такая же лемма (с той разницей, что ассоциированные числа в этом случае — это числа, отличающиеся знаком). Более того, как основная теорема арифметики 12.2, так и лемма 12.2 верны в любом евклидовом кольце. Коммутативное кольцо R без делителей нуля называется **евклидовым**, если определена функция (евклидова норма) $N : R \setminus \{0\} \rightarrow \mathbb{N}$ такая, что для любых $a, b \in R$, $b \neq 0$, имеет место представление

$$a = bq + r, \text{ где } N(r) < N(b) \text{ или } r = 0. \quad (12.4)$$

Здесь число r называется *остатком от деления a на b* . Поскольку функция N принимает только целые неотрицательные значения, требование (12.4) приводит к возможности реализации в кольце R алгоритма Евклида в виде (12.3).

В частности, если мы рассмотрим кольцо многочленов $K[x]$ над произвольным полем K , то в качестве евклидовой нормы можем выбрать степень многочлена. В результате мы получим деление с остатком и алгоритм Евклида в кольце $K[x]$ (см. главу 10). Два многочлена $P, Q \in K[x]$ будут ассоциированными, если $Q = kP$ при некотором $k \in K \setminus \{0\}$, поскольку многочлены нулевой степени, и только они, являются обратимыми в кольце $K[x]$.

Приведем пару примеров, где числа Гаусса дают заметный выигрыш по скорости и простоте решения задач, связанных с уравнениями в целых числах, т. е. с **диофантовыми уравнениями**.

Рассмотрим уравнение

$$x^2 + 1 = y^3, \quad x, y \in \mathbb{Z}.$$

В гауссовых числах оно эквивалентно уравнению

$$(x + i)(x - i) = y^3.$$

Покажем, что $x + i \perp x - i$. Действительно, если это не так, т. е. $z \mid x + i$ и $z \mid x - i$, то $z \mid (x + i) - (x - i) = 2i$, откуда $z = 1 + i$ или ему

ассоциированное. Кроме того, $z \mid y^3$, причем, поскольку $1 + i$ — простое, оно должно входить в разложение y^3 трижды, т. е. $z^3 \mid y^3$, но тогда в разложение $x + i$ или $x - i$ входит $z^2 = 2i$, чего быть не может, т. к. $x \pm i$ не делится на 2 (см. свойство Div1). Следовательно, $x + i \perp x - i$.

Из предыдущего и леммы 12.2 следует, что существует число $a + bi$ такое, что $x + i = \delta(a + bi)^3$, где δ — обратимое, т. е. $\delta = \pm 1, \pm i$. Возводя в куб, находим, что

$$x + i = \delta(a^3 - 3ab^2) + \delta(3a^2b - b^3)i,$$

откуда, сравнивая коэффициенты при i , получаем, что либо $3a^2b - b^3 = \pm 1$ (при $\delta = \pm 1$), либо $a^3 - 3ab^2 = \pm 1$ (при $\delta = \pm i$).

Пусть $\delta = \pm 1$ и $3a^2b - b^3 = \pm 1$. Ясно, что $b = \pm 1$, т. к. это уравнение в целых числах. Сокращая на b , получим: $3a^2 - 1 = \pm 1$, откуда $3a^2 = 0$ или $3a^2 = 2$. Второе невозможно, так что получаем $a = 0$, а отсюда следует, что $x = 0$. Аналогично, в случае $\delta = \pm i$ и $a^3 - 3ab^2 = \pm 1$ находим, что $b = 0$, и тогда снова имеем $x = 0$.

Таким образом, единственно возможное решение в целых числах у исходного уравнения $x^2 + 1 = y^3$ — это $x = 0, y = 1$.

Рассмотрим **Теорему Ферма** при $n = 2$: $a^2 + b^2 = c^2$ (в натуральных числах). Ясно, что можно сразу считать, что все числа a, b, c попарно взаимно простые натуральные числа (иначе можно было бы сократить уравнение на общий множитель). Отсюда также следует, что a и b имеют разную четность. Действительно, если a и b четные, то таково же и c , а значит, они не взаимно простые. Если a и b нечетные, то $a^2 + b^2$ имеет остаток 2 при делении на 4, но c^2 может иметь остаток либо 0 (при четном c), либо 1 (при нечетном c). Таким образом, допускается только случай, когда a и b имеют различную четность. Тогда по свойству Div6 число $a + bi$ не делится на $1 + i$.

Заметим, что $(a + bi)(a - bi) = a^2 + b^2 = c^2$. Предположим, что НОД чисел $a + bi$ и $a - bi$ равен r и отличен от делителя 1. Тогда $r \mid 2a$ и $r \mid 2bi$. Но $a \perp b$ в натуральных числах, тогда $N(a) \perp N(b)$, откуда по свойству Prim8 $a \perp b$ в гауссовых числах. Это значит, что r есть НОД 2 и $2i$, т. е. $r = 1 + i$ или его ассоциированным. Но такое число не может быть делителем $a + bi$ и $a - bi$ по доказанному выше. Следовательно, $(a + bi) \perp (a - bi)$.

Теперь, по лемме 12.2 существуют такие z, w , что $a + bi = z^2\delta_1$, $a - bi = w^2\delta_2$, $c = zw$, где $\delta_1, \delta_2 \in \{\pm 1, \pm i\}$. Пусть $z = n + mi$, тогда $a + bi = (n^2 - m^2 + 2nmi)\delta_1$, откуда $a - bi = (n^2 - m^2 - 2nmi)\delta_1$, откуда $w = n - mi$, $\delta_2 = \overline{\delta_1}$ и $c = n^2 + m^2$.

При $\delta_1 = \pm 1$ получаем, что $a = \pm(n^2 - m^2)$, $b = \pm 2nm$, $c^2 = n^2 + m^2$.

При $\delta_1 = \pm i$ получаем, что $a = \pm 2nm$, $b = \pm(n^2 - m^2)$, $c^2 = n^2 + m^2$.

Таким образом, поскольку мы ищем решение в натуральных числах, не теряя общности, запишем формулу **пифагоровых троек** для случая четного b :

$$a = n^2 - m^2, \quad b = 2nm, \quad c = n^2 + m^2,$$

где натуральные $n, m > 0$ и $n > m$.

Рассмотрим теперь уравнение $x^4 + y^4 = z^4$, неразрешимость которого доказал еще сам Ферма методом, который мы покажем ниже.

Докажем более сильное утверждение: $x^4 + y^4 = z^2$ неразрешимо в целых положительных числах.

Как и прежде, считаем сразу же, что $x \perp y$. Посмотрим на это уравнение как на уравнение второй степени: $(x^2)^2 + (y^2)^2 = z^2$. Если оно разрешимо, то существуют ненулевые взаимно простые n, m такие, что

$$x^2 = n^2 - m^2, \quad y^2 = 2nm, \quad z = n^2 + m^2,$$

откуда вновь получаем уравнение второй степени $x^2 + m^2 = n^2$, а значит, его решение имеет вид:

$$x = a^2 + b^2, \quad m = 2ab, \quad n = a^2 + b^2,$$

где ненулевые $a \perp b$. Тогда для y имеет место равенство: $y^2 = 4nab$ и, поскольку число 2 простое (в обычных целых числах), $y = 2y'$.

Тогда $(y')^2 = nab$. Так как n, a, b попарно взаимно просты (это следует из того, что $a \perp b$ и $n = a^2 + b^2$), в силу леммы 12.2 (для обычных целых чисел) существуют такие s, t, k , что $n = s^2$, $a = t^2$, $b = k^2$. Подставляем это в равенство $n = a^2 + b^2$, получаем:

$$t^4 + k^4 = s^2,$$

где $t \perp k$ и $z > s > 0$ (это следует из того, что $s = \sqrt{n}$, $n^2 < z$).

Таким образом, имея одно решение $\langle x, y, z \rangle$ исходного уравнения, мы построили еще одно такое $\langle t, k, s \rangle$, где $s < z$. Продолжая применять эти построения далее, мы получим бесконечную последовательность решений $\langle t_j, k_j, s_j \rangle$ такую, что $z > s > s_1 > s_2 > \dots$. Но это невозможно, т. к. в натуральном ряде не существует бесконечной строго убывающей последовательности.

Полученное противоречие доказывает неразрешимость уравнения $x^4 + y^4 = z^2$ в целых положительных числах, а значит, и неразрешимость уравнения $x^4 + y^4 = z^4$. Заметим, что отсюда сразу же следует справедливость теоремы Ферма для всех степеней n , кратных 4.

Предъявленный здесь метод доказательства называется **методом бесконечного спуска**. Он напоминает индукцию, только не доказывающую, а опровергающую, поскольку приводит к противоречию.

Упражнения

Обязательные упражнения

- 12.1° Доказать, что $z + w = w + z$ и что $zw = wz$.
- 12.2° Докажите, что если $zw = 0$, то либо $z = 0$, либо $w = 0$.
- 12.3° Докажите, что если и сумма, и произведение двух комплексных чисел вещественны, то либо оба этих числа вещественны, либо сопряжены.
- 12.4° Вычислить: а) $(-i)^2$; б) i^{10} ; в) $(1 + i)^{10}$; г) $(1 - i)^{10}$; д) $(1 + i)^{101}$.
- 12.5° Вычислить: а) $(2 + 3i) + (7 - i)$; б) $(2 + 3i)(7 - i)$; в) $(1 + i)(1 - i)$, $(2 - 3i)(3 + 2i)$; г) $2(4 + 3i) - 3(2 - i)$.
- 12.6° Найти два комплексных числа, сумма и произведение которых равны 2.
- 12.7° Найти комплексное число z , для которого а) $z(1 + i) = 1$; б) $z(2 + 3i) = 3 - 2i$; в) $z(1 + i) = 3 + 4i$; г) $z(2 + 3i) = 5 + 4i$.
- 12.8° Вывести формулу для расстояния между числами z и w .
- 12.9° Найти сумму $1 + i + i^2 + i^3 + \dots + i^{100}$.
- 12.10° Найти все z , для которых а) $z^2 = 2$; б) $z^2 = -2$; в) $z^2 = 2i$; г) $z^2 = 1 + i$; д) $z^2 + 2z + 2 = 0$; е) $z^3 = -1$.
- 12.11° а) Каков геометрический смысл суммы комплексных чисел? б) Сравните $|z + w|$ и $|z| + |w|$ для $z, w \in \mathbb{C}$.
- 12.12° Найдите модуль и аргумент чисел: а) -4 ; б) $1 + i$; в) $1 - i\sqrt{3}$; г) $\sin \alpha + i \cos \alpha$; д) $1 + \cos \alpha + i \sin \alpha$.
- 12.13° Рассмотрим умножение точек комплексной плоскости на $\cos \varphi + i \sin \varphi$ как преобразование f этой плоскости, переводящее z в $(\cos \varphi + i \sin \varphi)z$. Куда при этом преобразовании перейдут
- а) точки действительной оси;
- б) точки мнимой оси?
- в) Докажите, что f — поворот против часовой стрелки на угол φ вокруг начала координат.
- г) Пусть $z, w \in \mathbb{C}$. Выразите $|zw|$ и $\arg(zw)$ через $|z|$, $|w|$, $\arg(z)$, $\arg(w)$.
- д) Выведите из предыдущего пункта формулы для косинуса суммы и синуса суммы.

12.14° а) Из любого ли комплексного числа можно извлечь квадратный корень?

б) Решите уравнение $z^2 = i$. в) Найдите ошибку: $1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i \cdot i = i^2 = -1$.

12.15° Докажите, что если m и n — суммы двух квадратов целых чисел, то и mn — тоже.

12.16° (Формула Муавра) Пусть $z = r(\cos \varphi + i \sin \varphi)$, $n \in \mathbb{N}$. Докажите: $z^n = r^n(\cos n\varphi + i \sin n\varphi)$.

12.17° Вычислите а) $(\sqrt{3} + i)^{30}$; б) $(1 + i)^{333}$; в) $(1 + i\sqrt{3})^{150}$.

12.18° Выразите $\cos nx$ и $\sin nx$ через $\cos x$ и $\sin x$ ($n \in \mathbb{N}$).

12.19° Выразите $|\bar{z}|$, $\arg(\bar{z})$ через $|z|$, $\arg(z)$.

12.20° Докажите, что если $P(x)$ — многочлен с вещественными коэффициентами и $P(z) = 0$, то $P(\bar{z}) = 0$.

12.21° Вычислить действительную и мнимую части суммы и произведения чисел $a + bi$ и $c + di$.

12.22° Доказать, что для любого z сумма $z + \bar{z}$ и произведение $z\bar{z}$ действительны, то есть имеют нулевую мнимую часть.

12.23° Вычислить $(1 + i)/(1 - i)$ и $(8 + i)/(1 + 2i)$.

12.24° Найти общую формулу для частного $(a + bi)/(c + di)$.

12.25° Вычислите а) $\frac{(5+i)(7-6i)}{3+i}$; б) $\frac{(1+i)^5}{(1-i)^3}$.

12.26° Для каких z найдется такое w , что $zw = 1$?

12.27° Доказать, что если $z^2 = w^2$, то $z = w$ или $z = -w$.

12.28° Найти все комплексные числа, для которых $z^2 = 2i$.

12.29° Доказать, что $\overline{z/w} = \bar{z}/\bar{w}$.

12.30° Как найти модуль и аргумент частного z/w , зная модули и аргументы комплексных чисел z и w ?

12.31° Доказать тождество $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2$.

12.32° Доказать, что если два целых числа представимы в виде суммы двух квадратов (например, $2 = 1^2 + 1^2$ и $13 = 2^2 + 3^2$), то их произведение обладает этим же свойством.

12.33° Доказать, что число $2a$ представимо в виде суммы двух квадратов тогда и только тогда, когда в этом виде представимо число a .

12.34° Доказать, что если сумма и произведение двух комплексных чисел вещественны, то эти числа либо оба вещественны, либо сопряжены.

12.35° Найти все комплексные числа z , для которых $z^2 + z + 1 = 0$.

12.36° Найти все комплексные числа z , для которых $z^3 = 1$.

12.37° Нарисуйте: а) $\{z \in \mathbb{C} \mid z^n + 1 = 0\}$; б) $\{z \in \mathbb{C} \mid 2 > |z - i|\}$; в) $\{z \in \mathbb{C} \mid \operatorname{Re} \frac{1}{z} = 1\}$; д) $\{\frac{1+ti}{1-ti} \mid t \in \mathbb{R}\}$.

12.38° Доказать, что для любого комплексного α преобразование $z \mapsto \alpha z$ увеличивает все расстояния в одно и то же число раз, и найти это число.

12.39° Что можно сказать о преобразовании $z \mapsto \alpha z$, если число α действительное?

12.40° Что можно сказать о преобразовании $z \mapsto \alpha z$, если $|\alpha| = 1$?

12.41° При каком числе α преобразование $z \mapsto \alpha z$ будет поворотом на 30° вокруг начала координат?

12.42° Доказать, что преобразование $z \mapsto \alpha z$ есть композиция гомотетии с коэффициентом $|\alpha|$ и поворота на угол $\arg \alpha$.

12.43° Доказать, что преобразование поворота на угол φ вокруг начала координат задается формулой $z \mapsto z(\cos \varphi + i \sin \varphi)$.

12.44° Куда переходит точка $z = 1$ при повороте на угол φ , а затем на угол ψ в том же направлении? Вывести формулы для $\cos(\varphi + \psi)$ и $\sin(\varphi + \psi)$.

12.45° а) Докажите, что $e^{i\varphi}e^{i\psi} = e^{i(\varphi+\psi)}$. б) Можно ли найти $e^{i\varphi} + e^{i2\varphi} + \dots + e^{in\varphi}$ по формуле суммы геометрической прогрессии?

12.46° Найдите: а) $\sin \varphi + \sin 2\varphi + \dots + \sin n\varphi$; б) $\binom{n}{1} - \binom{n}{3} + \binom{n}{5} - \binom{n}{7} + \dots$; в) $\binom{n}{0} + \binom{n}{4} + \binom{n}{8} + \binom{n}{12} + \dots$.

12.47° Какое преобразование плоскости переводит а) z в $2z$; б) z в $z + 1$; в) z в \bar{z} ; д) z в $-z$; е) z в $-\bar{z}$; ф) z в iz ?

12.48° Опишите геометрически преобразование плоскости, заданное формулой: а) $z \mapsto z + w$, где w — комплексное число; б) $z \mapsto kz$, где k — вещественное число; в) $z \mapsto 2z + 1$; д) $z \mapsto \bar{z}$; е) $z \mapsto e^{i\alpha}z$, где $0 \leq \alpha < 2\pi$; ф) $z \mapsto wz$, где w — произвольно.

12.49° Где находятся точки z , для которых $z + \bar{z} = 1$?

12.50° Где находятся точки z , для которых $z \cdot \bar{z} = 1$?

12.51° Где находятся числа z , для которых а) $|z| = 1$; б) $|z - 1| = 1$; в) $|z| = |z + 1|$?

12.52° Найти действительную и мнимую части числа z , если $|z| = 2$ и $\arg z = 60^\circ$. Найти действительную и мнимую части чисел z^2 и z^3 .

12.53° Вывести формулы для действительной и мнимой частей числа с модулем r и аргументом φ .

12.54° Для данного z найти w , при котором точки $0, z, iz$ и w лежат в вершинах квадрата.

12.55° Решите уравнения: а) $z^2 = e^{i\pi/3}$; б) $z^2 = 5 - 12i$; в) $z^2 + (2i - 7)z + 13 - i = 0$; д) $\bar{z} = z^2$; е) $\bar{z} = z^3$; ф) $z^3 + z^2 + z + 1 = 0$.

12.56° Запишите как функцию комплексного переменного (можно использовать переменную z , комплексные числа и операции сложения, вычитания, умножения, деления, сопряжения):

- а) симметрию относительно оси y ;
- б) ортогональную проекцию на ось x ;
- в) центральную симметрию с центром A ;
- г) поворот на угол φ относительно точки A ;
- д) гомотегию с коэффициентом k и центром A ;
- е) скользящую симметрию относительно прямой $y = 3$ со сдвигом на 1 влево;
- ж) поворот, переводящий ось Ox в прямую $y = 2x + 1$;
- з) симметрию относительно прямой $y = 2x + 1$.

12.57° Доказать, что точки $0, 1/z$ и \bar{z} лежат на одной прямой.

12.58° Доказать, что точки z , для которых $z + \bar{z} = z \cdot \bar{z}$, лежат на одной окружности, и найти ее центр и радиус.

12.59° Доказать, что точки, для которых число $z/(z - 1)$ является чисто мнимым, лежат на одной окружности, и найти ее центр и радиус.

12.60° Найти все z , для которых $|z - 3| \leq 2$ и $|z + 4i| \leq 3$.

12.61° Найти комплексное число α , при котором преобразование $z \mapsto \alpha z$ есть поворот на 45° . Чему равен квадрат этого числа?

12.62° Доказать, что преобразование $z \mapsto (1 + i)z$ увеличивает все расстояния в одно и то же число раз, и найти это число.

12.63° Как найти четвертую вершину параллелограмма, если три его вершины совпадают с точками u, v и w комплексной плоскости? Указать все возможности.

12.64° Где находится точка пересечения медиан треугольника, вершинами которого являются точки u , v и w комплексной плоскости?

12.65° Вывести формулу для преобразования комплексной плоскости, являющегося симметрией относительно прямой $\operatorname{Re} z = \operatorname{Im} z$.

12.66° Пусть z и w — ненулевые различные комплексные числа. Верно ли, что точки z , w , $1/z$, $1/w$ лежат на одной окружности?

12.67° Даны два комплексных числа a и b . Опишите множество таких $z \in \mathbb{C}$, что $(z-a)/(z-b)$ — а) вещественное число; б) чисто мнимое число. в) Каков геометрический смысл аргумента этого числа? г) Докажите, что треугольники $\triangle z_1 z_2 z_3$ и $\triangle w_1 w_2 w_3$ подобны тогда и только тогда, когда простые отношения точек z_1, z_2, z_3 и w_1, w_2, w_3 равны или сопряжены.

12.68° Рассмотрим двойное отношение:

$$\lambda = [z, w; z_1, w_1] = \frac{z_1 - z}{z_1 - w} \cdot \frac{w_1 - w}{w_1 - z}.$$

Найти все значения двойного отношения $[z, w; z_1, w_1]$ при перестановках символов внутри обозначения $[z, w; z_1, w_1]$, выразить через данное λ .

12.69° Каким будет число λ , если все точки двойного отношения лежат а) на одной прямой; б) на одной окружности?

12.70° Докажите с помощью комплексных чисел, что а) композиция двух поворотов является поворотом или параллельным переносом; б) композиция поворота на ненулевой угол и параллельного переноса является поворотом.

12.71° (Эйлер) Докажите, что сумма квадратов длин сторон четырехугольника отличается от суммы квадратов диагоналей на учетверенный квадрат длины отрезка, соединяющего середины диагоналей.

12.72° Пусть M — точка на плоскости, S — окружность, AB — ее диаметр. Докажите, что величина $|MA|^2 + |MB|^2$ не зависит от выбора диаметра AB окружности S .

12.73° Докажите теорему косинусов $|BC|^2 = |AB|^2 + |AC|^2 - 2 \cdot |AB| \cdot |AC| \cos \alpha$, расположив вершины треугольника ABC в точках $0, z$ и w соответственно, где w — вещественное число.

12.74° На плоскости даны точки A, B, C . Пусть A_1 — образ точки C при повороте вокруг A на 90° против часовой стрелки, B_1 — образ точки C при повороте вокруг B на 90° по часовой стрелке, K — середина $A_1 B_1$,

M — середина AB . Докажите, что отрезки M и AB перпендикулярны. Как соотносятся их длины?

12.75° На сторонах треугольника $A_1A_2A_3$ во внешнюю сторону построены квадраты с центрами B_1, B_2, B_3 . Докажите, что отрезки B_1B_2 и A_3B_3 равны по длине и перпендикулярны.

12.76° Пусть $A_1A_2A_3$ и $B_1B_2B_3$ — правильные треугольники, и их вершины занумерованы против часовой стрелки. Докажите, что середины отрезков A_1B_1, A_2B_2 и A_3B_3 — вершины правильного треугольника.

12.77° Доказать, что комплексные числа z , для которых $z^3 = 1$, являются вершинами правильного треугольника.

12.78° Доказать, что для любого целого $n > 2$ и любого комплексного α корни уравнения $z^n = \alpha$ являются вершинами правильного n -угольника.

12.79° Докажите, что три точки z_1, z_2, z_3 являются вершинами правильного треугольника тогда и только тогда, когда $z_1^2 + z_2^2 + z_3^2 = z_1z_2 + z_1z_3 + z_2z_3$.

12.80° Докажите, что **a)** три различные точки z_1, z_2, z_3 лежат на одной прямой тогда и только тогда, когда их простое отношение вещественно; **b)** прямая, проходящей через точки z_1 и z_2 , задается уравнением $(z_1 - z)(\bar{z}_2 - \bar{z}) = (\bar{z}_1 - \bar{z})(z_2 - z)$; **c)** каково уравнение перпендикуляра к этой прямой, проходящего через w ?

12.81° Докажите, что **a)** $(z_1 - z_2)(z_4 - z_3) + (z_2 - z_3)(z_4 - z_1) = (z_2 - z_4)(z_3 - z_1)$; **b)** в любом четырехугольнике произведение длин диагоналей не превосходит сумму произведений длин противоположных сторон; **c)** (теорема Птолемея) для четырехугольника, вписанного в окружность, достигается равенство. **d)** Верно ли, что если равенство достигается, то четырехугольник вписанный?

12.82° Докажите, что прямая, проходящая через точки a и b единичной окружности $z\bar{z} = 1$, имеет уравнение $z + ab\bar{z} = a + b$, а касательная в точке p этой окружности имеет уравнение $p\bar{z} + p\bar{z} = 2$.

12.83° **a)** Пусть z_1 и z_2 — точки на единичной окружности $z\bar{z} = 1$. Докажите, что точка пересечения касательных к этой окружности, проходящих через z_1 и z_2 , — это точка $2z_1z_2/(z_1 + z_2)$ **b)** (Задача Ньютона) В описанном около окружности четырехугольнике середины диагоналей и центр окружности лежат на одной прямой.

12.84° Пусть a, b, c, d — различные точки на единичной окружности $z\bar{z} = 1$. Докажите, что секущая, проходящая через a и b , и секущая, проходящая через c и d , пересекаются в точке, сопряженной к $\frac{(a+b)-(c+d)}{ab-cd}$.

12.85° При каких a и b преобразование $z \mapsto az + b$ является поворотом на 45° вокруг точки $1 = 1 + 0i$?

12.86° Числа 0 и z являются вершинами правильного треугольника. Где может находиться третья его вершина?

12.87° Числа 0 и z являются вершинами квадрата. Где могут находиться две другие его вершины?

12.88° Найти все корни уравнения $z^5 = 1$. Указание: в ответе могут остаться квадратные корни, но не должно быть синусов и косинусов.

12.89° Доказать, что сумма всех n корней уравнения $z^n = 1$ равна нулю.

12.90° Найти произведение всех n корней уравнения $z^n = 1$.

12.91° Доказать, что все корни уравнения $z^n = 1$ являются степенями некоторого из них. Замечание: корень с таким свойством называется **первообразным корнем**.

12.92° Проверить, что корни уравнения $z^n = 1$ образуют группу по умножению. Какой известной вам группе данная группа изоморфна? Чем является первообразный корень в теоретико-групповой терминологии?

12.93° Сколько существует первообразных корней степени 12 из единицы?

12.94° Сколько существует первообразных корней степени 1001 из единицы?

12.95° При каких a и b преобразование $z \mapsto az + b$ является а) поворотом; б) параллельным переносом; в) осевой симметрией?

12.96° При каких a и b преобразование $z \mapsto a\bar{z} + b$ является осевой симметрией?

12.97° Найти все значения корня: а) $\sqrt[3]{-i}$; б) $\sqrt[4]{-16}$; в) $\sqrt[5]{1+i}$.

12.98° Что с точки зрения движений плоскости представляют собой ассоциированные числа?

12.99° Покажите, что в $\mathbb{Z}[i]$ сложение и умножение не выходит за рамки $\mathbb{Z}[i]$.

12.100° Нарисуйте на комплексной плоскости: а) числа $\mathbb{Z}[i]$; б) числа, на которые делится z и с) числа, которые делятся на z для $z = 1 + i, 2 + i, 3 + i, 3 + 2i$.

12.101° Докажите свойства нормы:

- а) $N(z) = 0$ тогда и только тогда, когда $z = 0$;
- б) $N(z) = N(\bar{z})$;
- с) z делит $N(z)$;
- д) если z делит w , то $N(z)$ делит $N(w)$.

12.102° Для следующих пар чисел выясните, делится ли какое-либо из них на другое, и найдите частное: $1 + i$ и 8 ; $2 + i$ и $3 + i$; $4 - 3i$ и $3 + 4i$.

12.103° Докажите, что для гауссовых чисел x и y следующие свойства эквивалентны: (1) множество делителей x совпадает с множеством делителей y ; (2) x делит y и y делит x ; (3) $x = ry$, где $N(r) = 1$.

12.104° Показать, что норма $z = a + bi$ четна тогда и только тогда, когда $(1+i) \mid z$, в частности, если a и b имеют разную четность, то z не делится на $1 + i$.

12.105° а) Докажите, что обратимые числа в точности числа с нормой 1. б) Докажите, что гауссово число $z \neq 0$ является простым тогда и только тогда, когда для любого разложения $z = wr$ какое-то из чисел w, r обратимо.

12.106° Являются ли простыми следующие гауссовы числа: $-i, 2, 3, 1 + i, 2 + i, 1 + 2i$?

12.107° Докажите, что гауссово число с простой нормой является простым.

12.108° Докажите, что простые натуральные числа разбиваются на два непересекающихся множества: простые гауссовы числа и числа, которые являются нормой простых гауссовых чисел.

12.109° а) Докажите, что простое натуральное число p является нормой гауссового числа тогда и только тогда, когда $p = a^2 + b^2$ для натуральных a, b . б) Какие простые числа $p \leq 29$ являются простыми гауссовыми? с) Сформулируйте гипотезу об этих числах в общем виде и докажите ее (Указание: используйте задачу 9.47°).

Дополнительные упражнения

12.110 Докажите, что многочлен степени n с коэффициентами из \mathbb{C} имеет не более n корней из \mathbb{C} .

12.111 а) Найдите и нарисуйте все корни из 1 степеней 2, 3, 4, 5 и 6.
б) Сколько всего корней из 1 степени n ? Найдите их произведение и сумму их s -х степеней для каждого $s \in \mathbb{N}$.

12.112 Пусть P — многочлен степени k с коэффициентами из \mathbb{C} . Докажите, что среднее арифметическое значений P в вершинах правильного n -угольника равно значению P в центре многоугольника, если $n > k$.

12.113 Вершины правильного n -угольника покрашены в несколько цветов так, что точки одного цвета — вершины правильного многоугольника. Докажите: среди этих многоугольников есть равные.

12.114 а) Пусть $z = (3 + 4i)/5$. Найдется ли такое $n \in \mathbb{N}$, что $z^n = 1$?
б) Докажите, что $\frac{1}{\pi} \arctg \frac{4}{3} \notin \mathbb{Q}$.

12.115 Можно ли сравнивать комплексные числа так, чтобы сохранились основные свойства неравенств (умножение на число, большее 0, не меняет знак неравенства и т. п.)? Верно ли, что $i > 0$, или что $i < 0$?

12.116 Что можно сказать о расположении точек z, z_1, w, w_1 на плоскости, если их двойное отношение — вещественное?

12.117 Каждую сторону n -угольника продолжили на ее длину (обходя по часовой стрелке). Пусть концы построенных отрезков образуют правильный n -угольник. Докажите, что и исходный n -угольник правильный.

12.118 Пусть вписанная окружность треугольника ABC задается уравнением $z\bar{z} = 1$ и касается его сторон в точках p, q, r . Докажите, что: а) $\frac{2pqr(p+q+r)}{(p+q)(p+r)(q+r)}$ — центр описанной окружности треугольника ABC ; б) $\frac{(pq+pr+qr)^2}{(p+q)(p+r)(q+r)}$ — центр окружности Эйлера треугольника ABC ; в) точка $\frac{pq+pr+qr}{p+q+r}$ лежит и на вписанной окружности, и на окружности Эйлера (окружность 9 точек) треугольника ABC ; д) (теорема Фейербаха) вписанная окружность и окружность Эйлера треугольника ABC касаются друг друга.

12.119 Пусть $z, w \in \mathbb{Z}[i]$, причем $w \neq 0$. а) Докажите, что существуют $q, r \in \mathbb{Z}[i]$ такие, что $z = wq + r$ и $N(r) < N(w)$. б) Сколькими способами можно выбрать такую пару гауссовых чисел?

12.120 а) Докажите, что НОД определен однозначно с точностью до умножения на обратимые. б) Докажите, что в гауссовых числах алгоритм Евклида для чисел z, w останавливается, в конце получается общий делитель d чисел z, w , который линейно выражается через изначальные два числа.

12.121 Найдите: **a)** НОД($7 + 9i, 10 + 2i$); **b)** НОД($7 - i, -4 + 7i$); **c)** НОД($5 + 3i, 6 - 4i$); **d)** НОД($7, 3 + i$); **e)** НОД($10 + i, 3 + 4i$).

12.122 Найдите НОД(z, \bar{z}) для произвольного $z \in \mathbb{Z}[i]$.

12.123 Разложите следующие гауссовы числа в произведение простых: **a)** $7 + i$; **b)** $11 + 2i$.

12.124 Пусть a, b, c — такие взаимно простые целые числа, что $a^2 + b^2 = c^2$. Докажите, что $c = |z|^2$ для некоторого $z \in \mathbb{Z}[i]$. (Указание: воспользуйтесь задачей 9.46°)

12.125 Укажите все тройки целых чисел $a, b, c \in \mathbb{Z}$ таких, что $a^2 + b^2 = c^2$. (То есть напишите формулу, которая дает все такие тройки при подстановке в нее целых чисел)

12.126 **a)** Верно ли, что целые числа a и b взаимно просты (как целые), если они взаимно просты как гауссовы? **b)** Верно ли обратное? **c)** Верно ли, что $z, w \in \mathbb{Z}[i]$ — гауссовы взаимно простые числа, если их нормы $N(z)$ и $N(w)$ — взаимно просты (как натуральные)? **d)** Верно ли обратное?

12.127 Докажите, что если гауссово простое делит произведение zw , то оно делит либо z , либо w .

12.128 Сформулируйте и докажите основную теорему арифметики для гауссовых чисел.

12.129* (Теорема Паскаля) Докажите, что точки пересечения прямых, содержащих противоположные стороны вписанного шестиугольника, лежат на одной прямой.

Введение в линейную алгебру

Аннотация

Глава посвящена в первую очередь анализу преобразований подобия прямой и плоскости. В то же время она дает начальные сведения о линейных преобразованиях и матрицах, тем самым открывая дверь в линейную алгебру.

13.1. Преобразования

Ранее в главе 11 мы ввели несколько определений понятия функции. Прежде всего, функция есть однозначное соответствие элементов одного множества элементам другого (или того же самого) множества. Во многих разделах математики функции принято называть **отображениями**.

Например, рассмотрим множество всех треугольников и будем ставить им в соответствие меру наибольшего угла. Ясно, что это будет отображение из множества всех треугольников в множество чисел. Ясно также, что не все числа могут быть мерой наибольшего угла в треугольнике, поэтому данное отображение не является сюръективным. Однако если сузить область значений такого отображения до интервала $[60^\circ; 180^\circ)$, то оно становится сюръективным, или *отображением на* данный числовой интервал.

Данное отображение не является инъективным, поскольку разным треугольникам могут соответствовать одинаковые значения наибольших углов. И только если сузить область определения данного отображения до множества всех равнобедренных треугольников, а подобные треугольники не различать, то максимальный угол однозначно определит треугольник, и отображение станет инъективным.

Таким образом, отображение, которое равнобедренному треугольнику ставит в соответствие его наибольший угол (в градусах), является

взаимно однозначным, т. е. биекцией (с точностью до отношения подобия).

Биекция может быть не только между фигурами и числами, но и вообще между любыми видами объектов. В частности, биекция может быть установлена между объектами одного рода, например, между треугольниками или просто точками на плоскости или в пространстве.

В геометрии под словом **преобразование** понимается биекция, которая отображает какое-то множество объектов в себя. Например, все изученные нами движения прямой, окружности, плоскости являются их преобразованиями, т. к. устанавливают взаимно однозначное соответствие между точками, соответственно, прямой, окружности и плоскости.

Если заданы отображения $f : X \rightarrow Y$ и $g : Y \rightarrow Z$, то можно составить из них композицию $g \circ f$, которая также является отображением и действует из X в Z по правилу

$$(g \circ f)(x) = g(f(x)),$$

т. е. сначала применяется правый компонент, а затем — левый.

Композиция двух отображений легко обобщается на любое конечное число отображений, важно только следить за тем, чтобы всякий раз каждое следующее отображение содержало в своей области определения все те значения, которые были получены на предыдущей цепочке композиции отображений. Таким образом, если мы хотим составить композицию

$$f_n \circ (f_{n-1} \dots (f_2 \circ f_1) \dots), \text{ где } f_k : X_k \rightarrow X_{k+1},$$

то должны выполняться следующие вложения образов:

$$f_1[X_1] \subseteq X_2, (f_2 \circ f_1)[X_1] \subseteq X_3, \dots, (f_{n-1} \dots f_2 \circ f_1)[X_1] \subseteq X_n.$$

Если, более того, для каждого $k \in \{1, \dots, n-1\}$ имеет место вложение $f_k[X_k] \subseteq X_{k+1}$, то все композиции вида $f_{k+1} \circ f_k$ будут определены корректно. В этом случае композиция подчиняется аксиоме ассоциативности, т. е. в этой цепочке можно как угодно расставлять скобки, например:

$$f_4 \circ (f_3 \circ (f_2 \circ f_1)) = (f_4 \circ f_3) \circ (f_2 \circ f_1) = (f_4 \circ (f_3 \circ f_2)) \circ f_1$$

и т. д. Действительно, достаточно проследить, куда перейдет при всех этих отображениях произвольный элемент $x \in X_1$.

Заметим, что ассоциативность операции — ключевое требование для определения группы. Если мы рассматриваем преобразования некоторого множества X в себя, то мы автоматически получаем группу, поскольку:

- 1) операция композиции ассоциативна (требования о вложении выполняются, т. к. $X_1 = \dots = X_n = X$);
- 2) существует нейтральный элемент — это тождественное преобразование id , которое «ничего не делает»;
- 3) преобразования по определению являются биекциями, а значит, обратимы: для всякого F есть F^{-1} такой, что $F \circ F^{-1} = F^{-1} \circ F = \text{id}$.

Например, все движения плоскости с операцией композиции образуют группу, поскольку они биективны.

А вот с требованием коммутативности у преобразований не все так хорошо. И мы помним примеры движений, которые не коммутируют, например сдвиг и отражение прямой.

Пусть у нас задано расстояние между точками множества X , обозначаемое за $\rho(x, y)$. Например, это может быть обычная длина вектора, соединяющего две точки на плоскости. Если $P : X \rightarrow X$ такое преобразование, что выполняется тождество

$$\rho(P(x), P(y)) = k\rho(x, y),$$

т. е. когда расстояние между образами точек в k раз больше исходного расстояния, то такое преобразование называется **подобием**. Число k при этом называется коэффициентом подобия.

Заметим, во-первых, что $k \geq 0$, т. к. это число связывает два неотрицательных числа, поскольку расстояние всегда есть неотрицательное число. Кроме того, при $k = 0$ отображение P не будет биекцией, т. к. оно схлопывает все точки в одну. Действительно, пусть $y = P(x)$ при некотором x . Тогда для любого другого $x' \in X$ имеем $\rho(y, P(x')) = k\rho(x, x') = 0$, следовательно, отображение P все точки X переводит в одну точку y . В дальнейшем мы всегда предполагаем, что $k > 0$.

Еще один особый случай: $k = 1$. При таком коэффициенте подобия мы получаем сохранение расстояний между точками при действии преобразования P , а значит, такое преобразование является движением.

Композиция подобий с коэффициентами k и s есть подобие с коэффициентом ks , что прямо следует из определения подобия.

Отсюда же следует, что если есть подобия с коэффициентами k и $1/k$, то их композиция окажется движением. Это свойство поможет нам в дальнейшем свести изучение подобий к движениям, о которых мы уже почти все знаем.

Пусть теперь X — прямая, плоскость или пространство. В таком случае мы можем задать на нем следующее отображение в себя. Зафиксируем точку O и для каждой точки A построим точку H_O^k по правилу: на прямой OA в направлении вектора \overrightarrow{OA} отложим отрезок длины $k|OA|$,

полученную точку обозначим за $H_O^k(A)$. Это имеет смысл и для $k < 0$, тогда в общем виде мы получаем, что

$$H_O^k(A) = O + k\overrightarrow{OA},$$

где под суммой точки и вектора понимается точка, полученная откладыванием данного вектора от данной точки.

Отображение H_O^k называется **гомотетией** (растяжением) с центром O и коэффициентом k . Число k предполагается любым, однако, как и в случае подобия, при $k = 0$ мы получим схлопывание всех точек в одну точку O , и такое отображение не только не будет преобразованием, но и вовсе неинтересно для изучения. Поэтому в дальнейшем мы также положим, что для гомотетий $|k| > 0$, т. е. $k \neq 0$.

Если мы рассмотрим гомотетии с общим центром O , то для них легко увидеть мультипликативное тождество:

$$H_O^k \circ H_O^s = H_O^{ks},$$

т. е. композиция концентрических гомотетий — это не что иное, как произведение чисел, являющихся коэффициентами этих гомотетий. Говоря языком алгебры, множество всех гомотетий с общим центром с операцией композиции изоморфно числовой оси (без нуля) с операцией умножения.

Это напоминает ситуацию со сдвигами на прямой, когда композиция $T_a \circ T_b = T_{a+b}$ соответствует операции сложения на числовой оси.

Гомотетия является преобразованием подобия. Действительно, какова бы ни была точка A , отличная от O , можно применить к ней гомотетию $H_O^{1/k}$ и получить точку A_0 , которая под действием гомотетии H_O^k перейдет в точку A , поскольку $H_O^k(H_O^{1/k}(A)) = A$. Следовательно, гомотетия является сюръекцией. Также легко проверить, что разные точки под действием гомотетии переходят в разные точки: у них либо сразу же разные направления, либо, если направление общее, разное расстояние от центра. Следовательно, гомотетия есть биекция.

Наконец, пусть $A' = H_O^k(A)$ и $B' = H_O^k(B)$. Рассмотрим треугольники AOB и $A'O'B'$. Эти треугольники подобны, т. к. у них общий угол и одинаковые пропорции сторон (см. рис. 13.1):

$$\frac{|OA'|}{|OA|} = \frac{|OB'|}{|OB|} = |k|.$$

Отсюда следует, что расстояния $\rho' = |A'B'|$ и $\rho = |AB|$ находятся в той же пропорции $|k| : 1$. Так что гомотетия есть частный случай подобия.

При $k = 1$ гомотетия, очевидно, является преобразованием id .

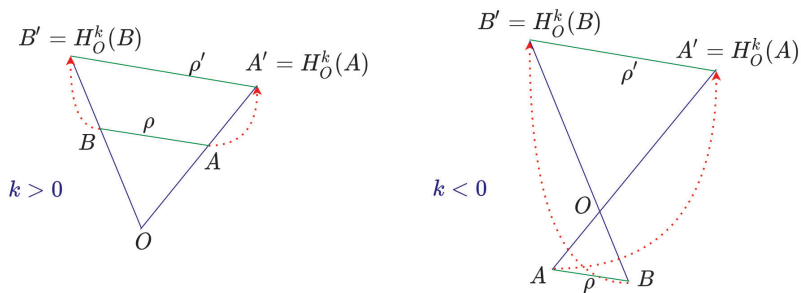


Рис. 13.1. Действие гомотетии на плоскости.

При $k = -1$ гомотетия с центром O является центральной симметрией, т.е. в случае прямой это — отражение относительно центра гомотетии, в случае плоскости — поворот на 180° относительно центра гомотетии, а в пространстве центральную симметрию можно представить как композицию отражения относительно плоскости, проходящей через центр гомотетии, с поворотом на 180° относительно прямой, проходящей через центр гомотетии и перпендикулярной данной плоскости отражения.

Таким образом, мы вновь замечаем, что числа — это не только меры длин и площадей, но это еще и преобразования! Ранее мы уже видели, что целые числа отвечали за кратности и направления преобразований, затем мы наблюдали композицию сдвигов и поворотов, соответствующую сложению произвольных чисел, теперь мы видим, что еще один вид преобразований эквивалентен умножению чисел.

13.2. Подобия прямой и плоскости

Пусть P — некоторое подобие прямой с коэффициентом подобия $k > 0$. Основная идея анализа видов подобий заключается в том, чтобы разложить подобие в композицию гомотетии и сдвига, поскольку то и другое есть частный случай подобия.

Рассмотрим некоторую гомотетию $H_A^{1/k}$, где точку A выберем произвольно. Тогда композиция

$$G = P \circ H_A^{1/k} \tag{13.1}$$

является движением прямой, т.е. либо сдвигом, либо отражением. Следовательно, исходное подобие P является композицией либо сдвига и гомотетии, либо отражения и гомотетии. Но отражение — частный случай гомотетии, так что все подобия прямой описываются таблицей умножения сдвигов и гомотетий 13.1.

Таблица 13.1. Композиция сдвигов и гомотетий

id	T_u	H_A^k
T_v	T_{u+v}	$H_{A+v/(1-k)}^k$ ($k \neq 1$)
H_B^s	$H_{B+su/(1-s)}^s$ ($s \neq 1$)	H_A^{ks} ($A = B$) $T_{(1-s)\overrightarrow{AB}}$ ($ks = 1$) $H_{A+\overrightarrow{AB}\frac{1-s}{1-ks}}^{ks}$ ($ks \neq 1$)

Итак, из формулы (13.1) мы получаем, что $P = G \circ H_A^k$, где G — это либо T_v , либо H_B^{-1} , т.е. отражение с центром в точке B . Тогда в первом случае получаем, что P есть либо T_v в случае $k = 1$, либо H_C^k , т.е. гомотетия с некоторым центром C . Во втором случае получаем, что $P = H_D^{-k}$, т.е. гомотетия с некоторым центром D .

Теорема 13.1. *Подобие на прямой является гомотетией, если коэффициент подобия отличен от 1. Если коэффициент равен 1, то это — сдвиг.*

Пусть P^k — некоторое подобие плоскости с коэффициентом подобия $k > 0$. Домножим его справа на гомотетию $H_O^{1/k}$ с центром в точке O . Снова получим какое-то движение, которое может быть либо параллельным переносом T_u , либо поворотом R_α^A с центром в точке A , либо скользящей симметрией W_u^l относительно оси l со сдвигом на вектор u . Следовательно, подобие P^k есть композиция одного из этих движений с гомотетией H_O^k . Посмотрим, что получается в каждом из трех случаев.

Мы упростим себе задачу, положив все преобразования на комплексную плоскость.

I. Чтобы найти $T_u \circ H_O^k$, поместим ноль комплексной плоскости в центр гомотетии, а вектор сдвига u будем считать сонаправленным действительной оси Re . Попробуем найти неподвижную точку такого преобразования. Проще всего ее искать на действительной оси, т.к. это сводит случай плоскости к прямой. Действительная точка x переходит сначала в kx под действием гомотетии, а затем сдвигается на число u , так что $P^k(x) = kx + u$. Эта точка будет неподвижной, если $kx + u = x$, или $x = u/(1 - k)$ (см. рис. 13.2).

Итак, при $k \neq 1$ у подобия $P^k = T_u \circ H_O^k$ существует неподвижная точка $O + u/(1 - k)$. При этом коэффициент подобия остается равным k , так что в итоге получаем, что

$$P^k = T_u \circ H_O^k = \begin{cases} H_{O+u/(1-k)}^k, & k \neq 1, \\ T_u, & k = 1. \end{cases}$$

II. Чтобы найти $R_\alpha^A \circ H_O^k$, поместим ноль комплексной плоскости в центр гомотетии O , а вещественную ось проведем через центр поворота A . Точка z переходит сначала в точку kz , а затем ее нужно повернуть относительно A . Для этого находим разность $kz - A$, временно перемещая центр координат в точку A , поворачиваем эту разность под действием R_α^A , домножая на единичный вектор w_α , соответствующий повороту на угол α , затем возвращаем полученное в начало координат смещением на $-A$, получаем $(kz - A)w_\alpha + A$. Находим неподвижную точку:

$$z_0 = \frac{A(1 - w_\alpha)}{1 - kw_\alpha}. \quad (13.2)$$

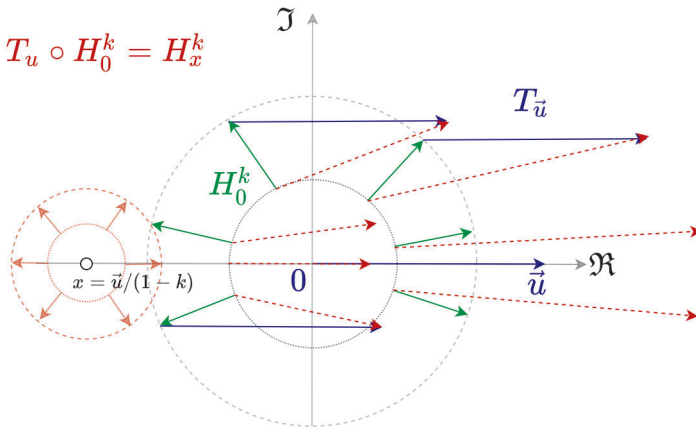


Рис. 13.2. Композиция сдвига и гомотетии.

Эта точка будет центром нового вида преобразований плоскости — **поворотной гомотетии**, которая является композицией концентрических поворота и гомотетии. Обозначим за $RH_O^{k,\alpha}$ поворотную гомотетию с центром в точке O , коэффициентом растяжения k и углом поворота α . Поворотная гомотетия иногда называется также *центрально-подобным вращением*. Результат композиции поворота и гомотетии в общем случае представлен на рис. 13.3.

Мы видим, что радиус-вектор, соединяющий центр поворотной гомотетии x и произвольную точку z , в результате действия такого подобия поворачивается на угол α и удлиняется в k раз.

Заметим, что если поместить ноль комплексной плоскости в центр поворотной гомотетии $RH_O^{k,\alpha}$, то ее действие можно записать как умножение на комплексное число $ke^{i\alpha}$. Отсюда, в частности, видно, что компоненты поворотной гомотетии — поворот и гомотетия с общим центром — коммутируют, т. е. $R_O^\alpha \circ H_O^k = H_O^k \circ R_O^\alpha$.

В общем случае при нахождении неподвижной точки z_0 преобразова-

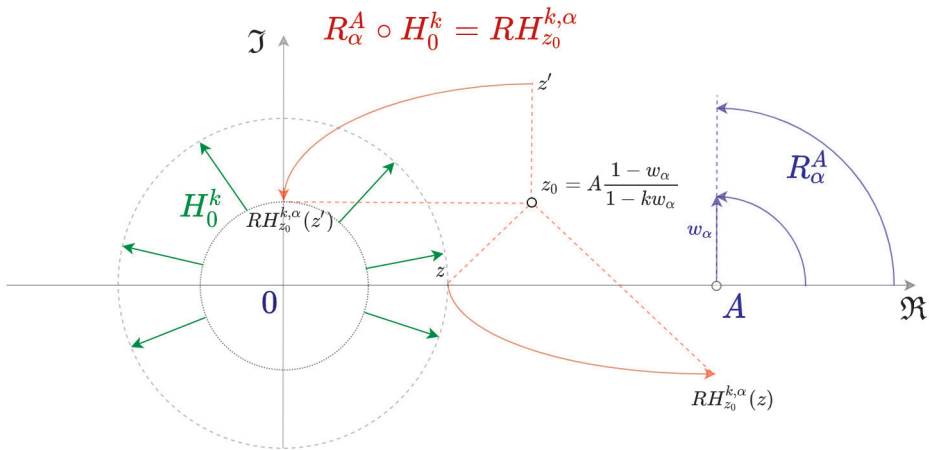


Рис. 13.3. Композиция поворота и гомотетии.

ния $R_\alpha^A \circ H_0^k$ по формуле (13.2) мы сталкиваемся с ограничением $kw_\alpha \neq 1$, которое нарушается лишь в двух случаях: $k = 1, \alpha = 0$ и $k = -1, \alpha = \pi$. Если ни одно из этих условий не выполняется, то неподвижная точка существует и вычисляется по указанной формуле. Если $k = 1, \alpha = 0$, то итоговая композиция есть id , т. е. изучаемое подобие P^k является тривиальным. Вариант $k = -1$ мы пропускаем, т. к. у рассматриваемого подобия P^k коэффициент k всегда положительный.

В итоге получаем

$$P^k = R_\alpha^A \circ H_0^k = \begin{cases} \text{id}, & k = 1, \alpha = 0, \\ RH_{A(1-w_\alpha)/(1-kw_\alpha)}^{k, \alpha}, & \text{иначе.} \end{cases}$$

III. Чтобы найти $W_u^l \circ H_0^k$, совместим действительную ось с осью симметрии l так, чтобы направление сдвига u совпадало с положительным направлением, а ноль совместим с проекцией точки O на ось l , так что O в данных координатах будет чисто мнимым числом, т. е. $O = xi, x \in \mathbb{R}$.

Снова ищем неподвижную точку. Произвольная точка комплексной плоскости может быть записана как $xi + z$. Под действием гомотетии она переходит в $xi + kz$, далее применяем отражение относительно действительной оси: $xi + kz = -xi + k\bar{z}$, далее сдвигаем результат на действительное число u и получаем точку $-xi + k\bar{z} + u$. Находим неподвижную точку: $xi + z = -xi + k\bar{z} + u$. Запишем действительную и мнимую часть этого уравнения по отдельности:

$$\begin{cases} \text{Re } z = k \text{Re } z + u, \\ x + \text{Im } z = -x - k \text{Im } z, \end{cases}$$

откуда

$$\begin{cases} \operatorname{Re} z = u/(1-k), \\ \operatorname{Im} z = -2x/(1+k), \end{cases}$$

и неподвижная точка имеет вид

$$z_0 = \frac{u}{1-k} + x \frac{k-1}{k+1} i.$$

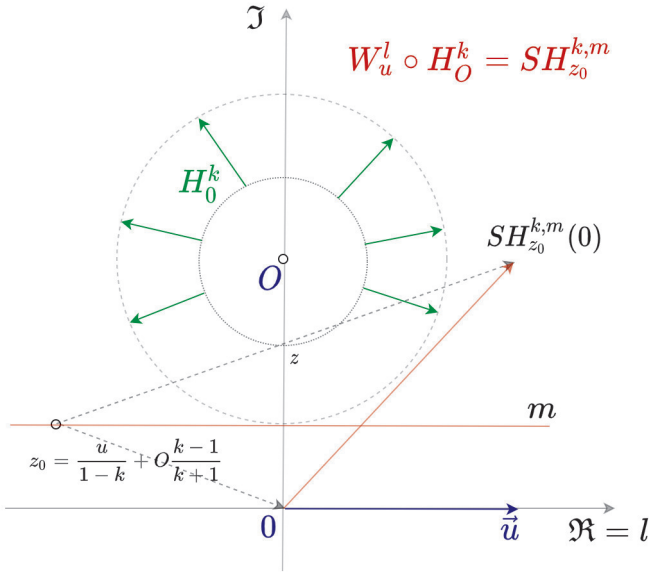


Рис. 13.4. Композиция скользящей симметрии и гомотетии.

Снова видим, что есть особые значения для k : $k = \pm 1$. При $k = 1$ гомотетия $H_O^k = \text{id}$, и в этом случае исходное подобие P^k совпадает с движением W_u^l . Вариант $k = -1$ мы пропускаем, т. к. у рассматриваемого подобия P^k коэффициент k всегда положительный.

Пусть теперь $k \neq 1$, $k > 0$. Рассмотрим прямые вида $z_0 + w_\alpha y$, где w_α — единичный вектор, имеющий угол наклона α к оси l , $y \in \mathbb{R}$ — параметр прямой. Найдем образ точки $z_0 + w_\alpha y$ при действии подобия $P^k = W_u^l \circ H_O^k$. Легко видеть, что

$$P^k(z_0 + w_\alpha y) = -xi + k(\overline{z_0 + w_\alpha y}) + u = z_0 + \bar{w}_\alpha(ky).$$

Таким образом, при действии рассматриваемого подобия прямая, проходящая через точку z_0 , переходит в прямую, проходящую через точку z_0 , симметричную относительно горизонтальной оси m , т. е. прямой, проходящей через z_0 и параллельной l . При этом точки на такой прямой становятся в k раз дальше от z_0 , чем на исходной прямой. См. рис. 13.4.

Это значит, что подобие P^k есть не что иное, как композиция гомотетии $H_{z_0}^k$ и отражения S_m , где прямая m параллельна l и проходит через точку z_0 . Мы получаем новый вид подобия плоскости — **отраженная гомотетия**. В некоторых источниках этот вид преобразований плоскости называется *центрально-подобной симметрией*. Заметим, что компоненты отраженной гомотетии, т. е. гомотетия и отражение, коммутируют. Более того, отраженную гомотетию легко описать в комплексных числах, если центр гомотетии совместить с нулем, а ось отражения — с действительной прямой. Тогда отраженная гомотетия есть не что иное, как умножение на число k с последующим сопряжением. Отсюда, в частности, легко увидеть свойство коммутативности компонентов отраженной гомотетии.

Обозначим за $SH_{z_0}^{k,m}$ отраженную гомотетию, центром которой является точка z_0 , коэффициентом гомотетии — k , а осью отражения — прямая m ($z_0 \in m$).

Итого, в третьем случае имеем:

$$P^k = W_u^l \circ H_O^k = \begin{cases} W_u^l, & k = 1, \\ SH_{\frac{u}{1-k} + O \frac{k-1}{k+1}}^{k,m}, & \text{иначе.} \end{cases}$$

Для наглядности на рис. 13.5 приведены схемы действий двух подобий плоскости, не являющихся движениями, — поворотной гомотетии и отраженной гомотетии.

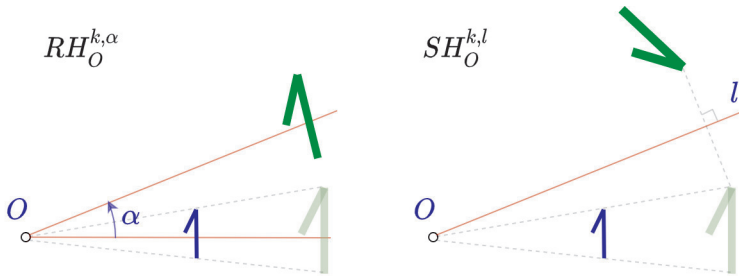


Рис. 13.5. Поворотная и отраженная гомотетия.

Подытожим.

Теорема 13.2. *Всякое подобие плоскости — это один из следующих видов преобразования:*

- параллельный перенос (в частности, id);
- поворотная гомотетия (в частности, id, поворот или гомотетия);
- скользящая симметрия (в частности, отражение);

- *отраженная гомотетия (в частности, отражение).*

Кроме того, можно добавить, что любое подобие плоскости является композицией не более чем трех симметрий и не более чем одной гомотетии.

Заметим также, что гомотетию с отрицательным коэффициентом всегда можно заменить на поворотную гомотетию с положительным коэффициентом и поворотом на угол π .

13.3. Линейное пространство

Пространства, с которыми мы до сих пор работали, — прямая и плоскость — это геометрические пространства точек. Несмотря на то что при определении движений мы предполагали сдвиг всех точек на какой-то вектор, этот вектор был для нас некоторым внешним объектом, под которым мы подразумевали само действие переноса всех точек на одно и то же расстояние в одном и том же направлении.

На самом деле векторы — это такие же полноценные математические сущности, как точки, прямые, отрезки, плоскости и фигуры. Наша локальная задача — выстроить понимание того, что такое вектор, как с ним работать, и какие новые возможности у нас при этом появятся.

Итак, первое, самое простое понятие вектора таково: вектор есть пара точек (A, B) , где первая точка называется началом вектора, вторая называется концом. Считается, что вектор соединяет эти две точки и имеет направление от первой ко второй. Заметим, что мы здесь не оговариваем специально, в каком пространстве мы находимся. В общем случае это может быть n -мерное пространство, но для простоты восприятия можно считать, что мы находимся на прямой или на плоскости. Обычно вектор, заданный парой точек, обозначается через \overrightarrow{AB} .

Такой вектор принято называть **фиксированным**, поскольку он жестко прибит двумя гвоздями—точками начала и конца к плоскости в определенном месте.

Если мы теперь обратимся к координатному методу, определим на плоскости начало координат O и две перпендикулярные оси Ox и Oy , на каждой из которых существует числовая сетка с нулем в начале координат, то мы можем рассматривать не пары точек, а пары чисел (x, y) , подразумевая, что откладываем их на соответствующих осях, чтобы получить конкретную точку на плоскости. В этом случае пару (x, y) можно рассматривать как точку, заданную координатами (с учетом расположения O и осей Ox, Oy), а можно — как вектор, отложенный от начала координат и заканчивающийся в точке (x, y) . Тогда x — это величина проекции данного вектора на ось Ox , а y — величина проекции данного

вектора на ось Oy . Мы используем термин «величина», поскольку она может быть отрицательной.

Этот второй подход к определению вектора, с одной стороны, также фиксирует его в определенном месте плоскости, а с другой стороны, любой параллельный перенос начала координат приведет нас к другому фиксированному вектору на плоскости, который, однако, будет задан все той же парой чисел. Поэтому, если отвлечься от фиксации начала координат, то под парой (x, y) можно понимать «алгоритм» построения вектора в любом месте плоскости: для этого достаточно выбрать его начало, а затем отложить проекции x и y вдоль соответствующих осей, чтобы вычислить конец вектора.

Эта идея алгоритма наводит нас на мысль о том, что мы можем ввести понятие вектора, не связанного с конкретными точками, а такого, который существует сразу во всех точках и может быть реализован в виде фиксированного вектора, если указать место его приложения. Такой вектор называется **свободным**.

Наконец, заметим, что свободный вектор, заданный алгоритмом (x, y) , будучи применен сразу во всех точках пространства, укажет единое направление и расстояние сдвига всех точек. А что это такое, если не параллельный перенос? Таким образом, под термином «свободный вектор», на самом деле, можно понимать параллельный перенос или сдвиг.

А теперь вспомним, что параллельные переносы образуют группу с операцией композиции, причем композиция переносов соответствует сумме векторов, на которые производится перенос всех точек. Следовательно, вектор v в записи переноса T_v есть не что иное, как свободный вектор, не привязанный к конкретным точкам плоскости. Опять-таки мы видим, что свободный вектор и параллельный перенос — суть одно и то же.

Существует еще один способ определить свободный вектор. С таким подходом мы уже сталкивались, когда говорили о классах вычетов и строили фактормножество $\mathbb{Z}/m\mathbb{Z}$. В том случае у нас числами новой арифметики стали классы точек, выбранных с равным шагом m на оси \mathbb{Z} . Точно так же мы поступим и с векторами. Скажем, что два вектора \vec{AB} и \vec{CD} эквивалентны ($\vec{AB} \sim \vec{CD}$), если четырехугольник $ABDC$ есть параллелограмм (в том числе вырожденный, когда все точки лежат на одной прямой, но при этом отрезок CD получается сдвигом отрезка AB вдоль этой прямой). По сути, это и означает, что данные векторы получаются один из другого некоторым параллельным переносом, поскольку если g — перенос, то четырехугольник $ABg(B)g(A)$ — параллелограмм, и, следовательно, $g(A)g(B) \sim AB$. И обратно, если даны два

эквивалентных вектора $\overrightarrow{AB} \sim \overrightarrow{CD}$, то перенос на вектор AC связывает их.

Далее, обозначим за $[\overrightarrow{AB}]$ класс фиксированных векторов, эквивалентных вектору \overrightarrow{AB} . И рассмотрим фактормножество, состоящее из всех таких классов. Поскольку все векторы в одном классе получают друг из друга переносами, каждый такой класс и есть свободный вектор.

На свободных векторах мы можем определить операцию сложения так же, как мы это делали с классами эквивалентности, когда работали с $\mathbb{Z}/m\mathbb{Z}$ или построением множества \mathbb{Q} , а именно, будем складывать их представителей:

$$[\overrightarrow{AB}] + [\overrightarrow{CD}] = [\overrightarrow{AB} + \overrightarrow{BB'}], \quad (13.3)$$

где вектор $\overrightarrow{BB'}$ эквивалентен вектору \overrightarrow{CD} , но стартует в точке B . Пользуясь все теми же параллелограммами, несложно доказать, что данное определение корректно, т. е. при смене представителей классов на эквивалентные получаем в результате вектор из того же самого класса.

В дальнейшем договоримся свободные векторы обозначать просто латинской буквой, без использования символики классов и отношения эквивалентности, понимая, что за этим стоит.

Итак, мы теперь имеем два сорта объектов: точки и свободные векторы. Векторы мы умеем складывать и даже знаем, что они образуют абелеву группу по сложению (по сути, это группа переносов). Как быть с точками?

Скажем, что суммой точки и вектора $A + v$ является точка, полученная откладыванием вектора v от точки A , т. е. мы находим вектор AB из класса v , и в качестве результата берем точку B . Обратная операция: $B - v$ — это такая точка A , что $A + v = B$. Кроме того, точки можно вычитать: $B - A$ — это такой вектор v , что \overrightarrow{AB} является элементом класса v . Отметим, что $B - A$ — это не фиксированный вектор, соединяющий точки A и B , а свободный вектор!

Тем самым мы подвели формальную базу под те записи движений, которые ранее использовали при изучении движений прямой и плоскости.

Если ранее мы рассматривали движения и подобия применительно к точкам, то теперь посмотрим, как они работают на свободных векторах. Пусть g — некоторое подобие прямой с коэффициентом $k > 0$, в частности, при $k = 1$ это движение. Пусть, кроме того, имеются точки A, B на прямой и вектор $v = B - A$. Каким будет вектор $g(B) - g(A)$? Очевидно, это будет вектор длины $k|v|$, при этом он будет либо сонаправлен с вектором v , и тогда это будет вектор kv , либо он будет ему противоположен по направлению, и тогда это будет вектор $-kv$.

Из строения подобий прямой мы знаем, что подобие есть либо сдвиг, либо гомотетия, а последняя может быть двух видов: с положительным коэффициентом k или с отрицательным $-k$, переворачивающая прямую в обратную сторону. Отсюда легко заключить, что все эквивалентные фиксированные векторы под действием подобия g переходят в эквивалентные же векторы, т. е. подобие сохраняет отношение эквивалентности.

Но тогда корректным будет следующее определение: $g(v) = \pm kv$, где знак перед k зависит от знака соответствующей гомотетии. При этом если g является сдвигом, то оно не меняет класс v вектора, т. е. в случае свободных векторов сдвигов, можно сказать, вовсе не существует. Иначе говоря, сдвиги прямой на свободных векторах действуют как id . Совершенно точно так же рассуждаем в случае плоскости и любого многомерного пространства.

В случае плоскости мы знаем также, что любое подобие g есть либо параллельный перенос, либо поворотная гомотетия на угол α с коэффициентом k , либо скользящая симметрия, либо отраженная гомотетия.

В случае, когда g есть параллельный перенос, очевидно, что $g(v) = v$.

В случае, когда g есть поворотная гомотетия $RH_O^{k,\alpha}$ ($k > 0$), получим, что $g(v) = kv'$, где v' есть вектор, повернутый на угол α относительно вектора v . См. рис. 13.5.

В случае, когда g есть скользящая симметрия W_u^l , ее действие эквивалентно обычной симметрии S^l , поскольку сдвиг T_u не меняет класс v . Поэтому $g(v) = v'$, где вектор v' получается из вектора v поворотом на угол -2α , где α — угол между прямой l и вектором v . Заметим, что на свободных векторах скользящая симметрия становится частным случаем отраженной гомотетии (при $k = 1$).

Наконец, в случае, когда g является отраженной гомотетией $SH_O^{k,l}$, получаем, что $g(v) = kv'$, где вектор v' получается из вектора v поворотом на угол -2α , где α — угол между прямой l и вектором v .

Отметим также, что во всех четырех случаях подобие сохраняет класс эквивалентности. Действительно, как видно из вышеприведенных рассуждений, подобие только умножает вектор на число и/или поворачивает его на какой-то угол, а эти действия сохраняют эквивалентность векторов.

Подводя итог, скажем, что подобия на свободных векторах определены корректно и делятся на три вида: id , поворотная гомотетия, отраженная гомотетия (в частности, отражение).

Ранее мы определили сложение свободных векторов по формуле (13.3). Гомотетия позволяет определить умножение свободного вектора на число аналогичным способом.

Положим по определению:

$$k[\overrightarrow{AB}] = [H_O^k(\overrightarrow{AB})]. \quad (13.4)$$

Выбор центра гомотетии — точки O — здесь не существенен, что легко видеть из рис. 13.6. Независимо от выбранного центра в результате получаются эквивалентные векторы. Таким образом, умножение свободного вектора на число с помощью гомотетии корректно.

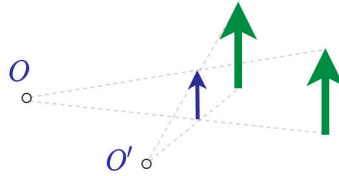


Рис. 13.6. Гомотетия на свободных векторах.

В дальнейшем, если не оговорено иное, под термином «вектор» мы будем понимать свободный вектор.

Рассмотрим, какими свойствами обладают операции над векторами, определенные по формулам (13.3) и (13.4).

- 1) $1 * v = v$, поскольку гомотетия с коэффициентом 1 есть id на векторах;
- 2) $k(u + v) = ku + kv$, поскольку гомотетия переводит сумму векторов в сумму;
- 3) $k_1(k_2v) = (k_1k_2)v$, поскольку произведение k_1k_2 соответствует композиции гомотетий;
- 4) $(k_1 + k_2)v = k_1v + k_2v$ — это следует из того, что вектор $(k_1 + k_2)v$ определяется как вектор, имеющий длину $|k_1 + k_2||v|$ и сонаправленный с v , если $k_1 + k_2 > 0$, и противоположно направленный, если $k_1 + k_2 < 0$.

Отметим также особый случай — гомотетия с коэффициентом 0, которая схлопывает все векторы в точку, или, в нулевой вектор. Нулевой вектор в арифметике векторов играет такую же роль, что и обычный ноль в кольце чисел: $k\vec{0} = \vec{0}$ и $v + \vec{0} = v$.

Итак, мы определили понятие (свободного) вектора, научились их складывать и умножать на числа, получили основные свойства этих операций.

Такая структура, составленная из векторов и чисел, подчиняющаяся ниже перечисленным свойствам, в алгебре называется **модулем** (над соответствующей системой чисел, например, \mathbb{Z} или \mathbb{Q} , или \mathbb{R}). Приведем все требования к модулю:

- Mod1** Векторы образуют коммутативную группу по сложению;
- Mod2** Числа образуют коммутативное кольцо с единицей;

Mod3 Умножение $*$ числа на вектор и сложение векторов согласованы в соответствии с правилами, указанными на рис. 13.7.

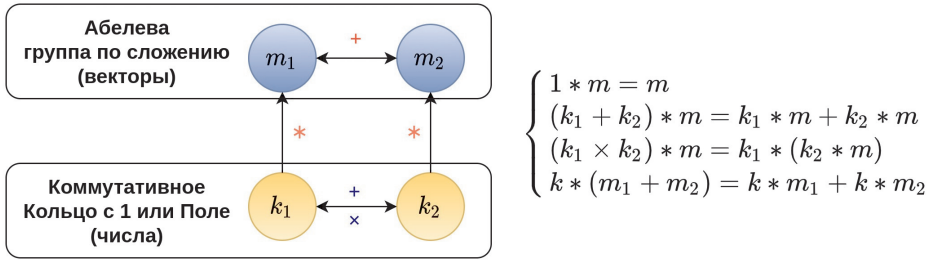


Рис. 13.7. Модуль над кольцом с единицей или полем.

В том случае, когда числовая структура является полем, т.е. числа можно делить друг на друга, модуль называется **векторным** (или **линейным**) **пространством** и является объектом изучения в разделе математики, называемом линейной алгеброй.

Рассмотрим произвольный модуль V над кольцом K . Элементы модуля, как и прежде, будем называть векторами, а элементы кольца — числами.

Линейной комбинацией векторов $e_1, \dots, e_n \in V$ называется всякое выражение вида

$$k_1 e_1 + \dots + k_n e_n,$$

где коэффициенты $k_1, \dots, k_n \in K$. Возьмем все возможные такие комбинации и соберем в множество V' :

$$V' = \{k_1 e_1 + \dots + k_n e_n \mid k_1, \dots, k_n \in K\}.$$

Тогда V' называется **линейной оболочкой** системы векторов $\{e_1, \dots, e_n\}$.

Нетрудно доказать, что линейная оболочка любой системы векторов, даже одного нулевого вектора, является сама по себе модулем (а если K — поле, то линейным пространством). Понятно, что $V' \subseteq V$. Если модуль V_1 является подмножеством другого модуля V_2 (над тем же кольцом, с теми же операциями), то его называют подмодулем модуля V_2 . Аналогично, если линейное пространство V_1 является подмножеством линейного пространства V_2 (над тем же полем, с теми же операциями), то оно называется подпространством пространства V_2 .

В том случае, когда $V' = V$, говорят, что система векторов $\{e_1, \dots, e_n\}$ порождает пространство V .

Например, на плоскости можно выбрать два перпендикулярных вектора, и тогда все возможные линейные комбинации этих векторов закрывают всю плоскость. Если к ним добавить третий вектор из той же плоскости, то ничего не изменится — они по-прежнему будут закрывать

всю плоскость, и не более того. Однако, если взять только один вектор, то его линейные комбинации будут покрывать всего лишь ту прямую, на которой лежит данный вектор, и эта прямая будет собственным подпространством плоскости. Поэтому системы векторов могут быть избыточными или недостаточными для того, чтобы получить равенство $V' = V$.

Система векторов $\{e_1, \dots, e_n\}$ называется **линейно независимой**, если никакая нетривиальная линейная комбинация этих векторов не равна нулю. Для примера, на плоскости никакая линейная комбинация, кроме как когда оба коэффициента — нули, не обратит в ноль систему из двух перпендикулярных векторов. В то же время, если векторы коллинеарны, то можно так подобрать не равные нулю коэффициенты, что их линейная комбинация обратится в ноль.

В том случае, когда мы имеем дело с линейным пространством (т.е. K — поле) и системой линейно зависимых векторов, то один из этих векторов можно выразить через остальные в виде линейной комбинации. Поэтому в линейном пространстве зависимость векторов равносильна возможности выразить один из них через остальные в виде линейной комбинации.

Система векторов $\{e_1, \dots, e_n\}$ называется **базисом** V , если она линейно независима, а ее линейная оболочка равна V , т.е. любой вектор пространства можно представить в виде линейной комбинации базисных векторов.

Лемма 13.1. *В ненулевом коммутативном кольце линейная система уравнений*

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \dots\dots\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$$

имеет нетривиальное решение x_1^, \dots, x_n^* , если $m < n$.*

Доказательство. Проведем доказательство индукцией по n . Для $n = 2$ имеем единственный вариант системы: $a_{11}x_1 + a_{12}x_2 = 0$.

Случай 1. $a_{11} = a_{12} = 0$. Тогда решение $x_1^* = 1, x_2^* = 0$ является нетривиальным решением (как и вообще любое другое).

Случай 2. $a_{11} \neq 0$ или $a_{12} \neq 0$. Тогда решение $x_1^* = a_{12}, x_2^* = -a_{11}$ является нетривиальным решением (мы воспользовались коммутативностью кольца K).

Предположим, что для n лемма доказана, и рассмотрим случай $n + 1$

Доказательство. Пусть $\mathbf{e}_1, \dots, \mathbf{e}_m$ и $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ — базисы и $m < n$. Тогда в силу определения базиса имеем:

$$\begin{aligned} \mathbf{e}'_1 &= a_{11}\mathbf{e}_1 + \dots + a_{m1}\mathbf{e}_m \\ &\dots\dots\dots \\ \mathbf{e}'_n &= a_{1n}\mathbf{e}_1 + \dots + a_{mn}\mathbf{e}_m \end{aligned}$$

и рассмотрим линейное уравнение $x_1\mathbf{e}'_1 + \dots + x_n\mathbf{e}'_n = 0$. Если $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ — базис, то это уравнение может иметь только тривиальное решение $x_1 = \dots = x_n = 0$. Подставим сюда разложения базисных векторов и получим:

$$\begin{aligned} 0 &= (a_{11}x_1 + \dots + a_{1n}x_n)\mathbf{e}_1 + \\ &\dots\dots\dots \\ &(a_{m1}x_1 + \dots + a_{mn}x_n)\mathbf{e}_m, \end{aligned}$$

откуда в силу того, что $\mathbf{e}_1, \dots, \mathbf{e}_m$ — базис, все коэффициенты должны быть равны нулю, т. е.

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \dots\dots\dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$$

Но в силу предыдущей леммы эта система имеет нетривиальное решение x_1^*, \dots, x_n^* , а это означает, что и уравнение $x_1\mathbf{e}'_1 + \dots + x_n\mathbf{e}'_n = 0$ имеет нетривиальное решение. Следовательно, $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ не может быть базисом при $n > m$.

Таким образом, все конечные базисы модуля (над коммутативным кольцом) равнозначны, если они существуют. \square

Следствие 13.1. *Если в линейном пространстве существует конечный базис, то все базисы этого линейного пространства имеют одинаковое количество элементов.*

Размерность линейного пространства называется количество векторов его базиса. Размерность пространства V обозначается через $\dim(V)$.

13.4. Линейные операторы

Пусть V и U — линейные пространства над полем K . Пока мы не будем уточнять, совпадают ли эти пространства, имеют ли они одинаковую размерность и какова их природа. Нам важно только то, что оба они заданы над одним и тем же полем K .

Отображение $L : V \rightarrow U$ называется **линейным оператором** (действующим из пространства V в пространство U), если выполнены условия:

L1 Аддитивность: $L(u + v) = L(u) + L(v)$.

L2 Однородность: $L(kv) = kL(v)$.

Если $V = U$, то говорят, что линейный оператор L **действует на** пространстве V .

Теорема 13.4. *Всякое подобие является линейным оператором на множестве (свободных) векторов.*

Эта теорема следует из наших предыдущих построений.

Какие еще бывают линейные операторы, действующие на линейном пространстве? Изучению этого вопроса мы посвятим оставшуюся часть главы.

Существенным признаком, разделяющим линейные операторы, действующие на одном пространстве, на два больших класса, является биективность, т. е. свойство быть преобразованием пространства векторов. Если линейный оператор является биекцией, то он называется **обратимым линейным оператором** или линейным преобразованием.

Обратимые линейные операторы с операцией композиции образуют группу (ведь мы всегда можем взять обратный оператор, который также является линейным, а кроме того, у нас имеется оператор id , ну а требование ассоциативности отображений выполняется автоматически). Эта группа, которая обозначается $\text{GL}(V)$, называется **полной линейной группой** пространства V .

Для линейных операторов можно естественным образом задать операции сложения и умножения на число из того же самого поля, над которым заданы векторы пространств V и U . Пусть L и M — линейные операторы, действующие из V в U , тогда положим

$$(L + M)(v) = L(v) + M(v), \quad (kL)(v) = kL(v),$$

т. е. мы переносим операции с векторов на операторы. Нетрудно видеть, что при такой арифметике сами операторы ведут себя ровно так же, как и векторы, т. е. заданные на них операции подчиняются аксиомам модуля. Это значит, что множество всех линейных операторов над линейным пространством V само по себе образует новое линейное пространство.

Но и это еще не все. На линейных операторах задана операция композиции, которую в случае операторов мы будем называть умножением и обозначать соответствующим образом. Умножение линейных операторов подчиняется правилу:

$$k(LM) = (kL)M = L(kM). \quad (13.6)$$

Действительно, $k(LM) = L(kM)$ в силу однородности L . В то же время, $(kL)M(v) = kL(M(v)) = k(LM)(v)$. То есть число k можно безнаказанно

проносить сквозь символ линейного оператора и сквозь любые композиции линейных операторов.

Если модуль над кольцом/полем (т. е. структура с аксиомами Mod1–Mod3) подчиняется еще и свойству (13.6), то он называется **алгеброй над кольцом/полем**. На рис. 13.8 схематично представлены компоненты алгебры и операции, которые в ней задействованы: сложение и умножение векторов, сложение и умножение чисел, умножение числа на вектор. На схеме даны только аксиомы операции $*$ умножения числа на вектор. Они в точности повторяют соответствующие аксиомы для модуля, к которым присоединилась аксиома, связывающая умножение векторов с умножением на число.

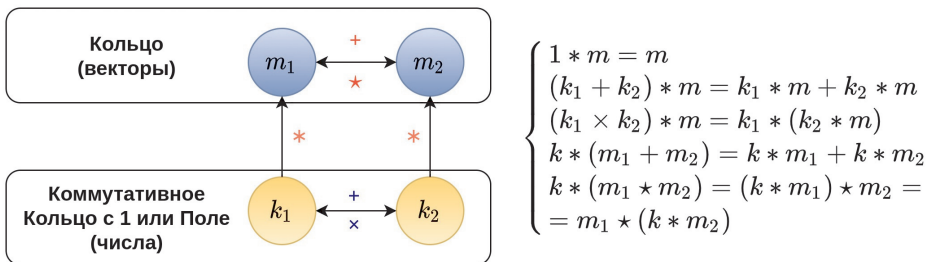


Рис. 13.8. Алгебра над кольцом с единицей или полем.

Как видим, множество всех линейных операторов, действующих из одного векторного пространства в другое (или то же самое), является алгеброй с операциями сложения и умножения (композиции).

Отметим, что при записи линейных операторов часто опускаются скобки, так что выражение $kL(v)$ приобретает вид kLv и может рассматриваться как умножение соответствующих сущностей. Мы также будем следовать этой традиции в тех случаях, когда это не вызовет двойного толкования записи.

Рассмотрим на плоскости два свободных вектора e_1 и e_2 . Мы будем предполагать, что эти векторы не лежат на одной прямой, т. е. неколлинеарны. В этом случае любой другой вектор v можно спроецировать на оси векторов e_1 и e_2 , и данные две проекции выразить числами, считая векторы e_1 и e_2 условными единицами каждый на своей оси. Таким образом, вектору v будет сопоставлена пара чисел (x, y) , где $x = \text{Pr}_{e_1} v$ и $y = \text{Pr}_{e_2} v$. Поскольку векторы e_1 и e_2 не лежат на одной прямой, т. е. линейно независимы, такое сопоставление $v \leftrightarrow (x, y)$ вектора и пары чисел является взаимно однозначным: вектор v легко восстанавливается с помощью этих чисел по правилу параллелограмма

$$v = xe_1 + ye_2.$$

Набор векторов $\{e_1, e_2\}$ на плоскости образует базис векторного про-

пространства, которое мы обозначим за V , поскольку любой вектор однозначно представляется в виде такой линейной комбинации векторов \mathbf{e}_1 и \mathbf{e}_2 , и в то же время векторы \mathbf{e}_1 и \mathbf{e}_2 независимы.

Если мы теперь вспомним про один из способов задания вектора, а именно, через пару координат на координатной плоскости, то поймем, что это частный случай разложения вектора по базису. Именно, пусть вектор задан парой (x, y) , где числа x и y суть его координаты, отложенные по осям Ox и Oy . Тогда

$$(x, y) = x(1, 0) + y(0, 1),$$

т. е. в данном случае векторы $(1, 0)$ и $(0, 1)$ являются базисом, в котором наш вектор имеет представление (x, y) .

Рассмотрим произвольный линейный оператор $L : V \rightarrow V$. В силу свойств линейности для произвольного вектора $v = x\mathbf{e}_1 + y\mathbf{e}_2$ имеем

$$L(v) = xw_1 + yw_2, \text{ где } w_1 = L(\mathbf{e}_1), \quad w_2 = L(\mathbf{e}_2),$$

т. е. образ вектора v точно так же раскладывается по векторам w_1 и w_2 , которые являются образами базисных векторов \mathbf{e}_1 и \mathbf{e}_2 .

Заметим, что векторы w_1 и w_2 могут оказаться коллинеарными, а значит, и все линейные комбинации $xw_1 + yw_2$ будут коллинеарны этим двум векторам. В этом случае отображение L «схлопнет» исходную плоскость в одну прямую, на которой лежат эти векторы, и окажется, что отображение L не является биективным, т. е. не будет преобразованием плоскости. Собственно, в этом и кроется отличие линейных преобразований от произвольных линейных отображений.

Поскольку векторы w_1 и w_2 лежат в том же самом пространстве V , их тоже можно разложить по базису $\mathbf{e}_1, \mathbf{e}_2$. Пусть

$$w_1 = w_{11}\mathbf{e}_1 + w_{21}\mathbf{e}_2, \quad w_2 = w_{12}\mathbf{e}_1 + w_{22}\mathbf{e}_2.$$

Составим из этих векторов квадратную матрицу 2×2 :

$$W = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}.$$

Матрица W называется **матрицей линейного оператора** L . Эта матрица составлена из векторов-образов базиса путем выписывания в столбик координат каждого вектора, в которые перешли базисные векторы под действием оператора L .

Посмотрим, как будут выглядеть координаты вектора $L(v)$ в исходном базисе:

$$\begin{aligned} L(v) &= xw_1 + yw_2 = x(w_{11}\mathbf{e}_1 + w_{21}\mathbf{e}_2) + y(w_{12}\mathbf{e}_1 + w_{22}\mathbf{e}_2) = \\ &= (xw_{11} + yw_{12})\mathbf{e}_1 + (xw_{21} + yw_{22})\mathbf{e}_2. \end{aligned}$$

Эти равенства задают умножение матрицы на вектор:

$$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xw_{11} + yw_{12} \\ xw_{21} + yw_{22} \end{pmatrix}. \quad (13.7)$$

Приведем пример. Пусть e_1 и e_2 — стандартный базис, т. е. векторы $(1, 0)$ и $(0, 1)$ на ортогональной координатной сетке. Пусть нам нужно сделать поворот плоскости на угол α относительно начала координат. В этом случае векторы w_1 и w_2 будут образами базисных векторов при таком повороте. Но тогда проекцией w_1 на e_1 будет $\cos \alpha$, проекцией w_1 на e_2 будет $\sin \alpha$, проекцией w_2 на e_1 будет уже $\sin(\pi/2 - \alpha) = -\sin \alpha$, проекцией w_2 на e_2 будет $\cos \alpha$ (см. рис. 13.9).

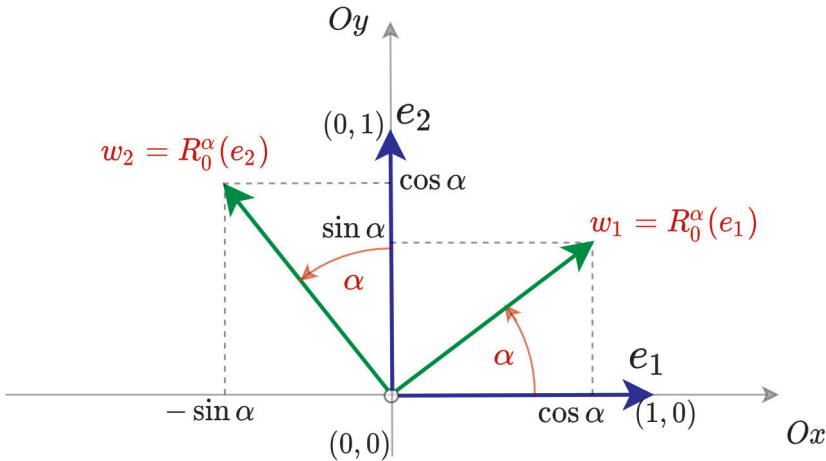


Рис. 13.9. Получение матрицы поворота на угол α .

Таким образом, матрицей поворота R_0^α будет матрица

$$R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Матрица $\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$ задает, как легко видеть, растяжение вдоль оси Ox , поскольку

$$\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ y \end{pmatrix}.$$

Внимание! Это линейное преобразование, которое не является ни движением, ни гомотетией и представляет собой пример такого преобразования плоскости, каких мы до сих пор не встречали!

Немного модифицируем предыдущую матрицу: пусть

$$H_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

Здесь мы уже видим, что H_λ переводит вектор (x, y) в вектор $(\lambda x, \lambda y)$, т. е. осуществляет гомотетию с центром в начале координат и коэффициентом λ .

Наконец, еще один пример:

$$S_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Данная матрица соответствует оператору, который переводит вектор (x, y) в вектор $(x, -y)$, т. е. осуществляет отражение относительно оси Ox .

13.5. Арифметика матриц

Выше было установлено, как умножать матрицу на вектор. На самом деле, вектор — это тоже матрица, только с одним столбцом, поэтому отчасти мы уже знаем, как умножать матрицу на матрицу. Выведем свойства арифметики матриц из свойств соответствующих им линейных операторов. Для простоты мы рассмотрим случай, когда базисные векторы \mathbf{e}_1 и \mathbf{e}_2 перпендикулярны и задают обычную координатную сетку, т. е. $\mathbf{e}_1 = (1, 0)$ и $\mathbf{e}_2 = (0, 1)$, а также всюду далее будем предполагать, что линейные операторы действуют на одном и том же пространстве. Но на самом деле, арифметика матриц основана исключительно на арифметике операторов, и потому не зависит от того, в каком базисе мы их изучаем.

Сложение матриц

Сложение матриц производится покомпонентно (так же, как векторов). Пусть даны два оператора L и M , и соответствующие им матрицы W и U . Как мы уже выяснили, эти матрицы составлены из векторов-столбиков, причем эти векторы есть представление образов базисных векторов в исходном базисе. То есть если $L\mathbf{e}_1 = w_1$ и $L\mathbf{e}_2 = w_2$, то матрица $W = [w_1; w_2]$. Такая запись означает, что мы берем столбики $w_1 = \begin{pmatrix} w_{11} \\ w_{21} \end{pmatrix}$ и $w_2 = \begin{pmatrix} w_{12} \\ w_{22} \end{pmatrix}$ и ставим их слева направо, образуя квадратную матрицу W .

Пусть также $M\mathbf{e}_1 = u_1$ и $M\mathbf{e}_2 = u_2$, тогда $U = [u_1; u_2]$.

Пользуясь тем, что сумма операторов $L + M$ определяется как сумма образов $(L + M)v = Lv + Mv$, заключаем, что

$$(L + M)\mathbf{e}_1 = w_1 + u_1, \quad (L + M)\mathbf{e}_2 = w_2 + u_2,$$

откуда следует, что матрицей оператора $L + M$ будет

$$W + U = [w_1 + u_1; w_2 + u_2],$$

т. е. сложение матриц, в самом деле, производится покомпонентно, как и векторов.

Заметим, что если мы перейдем в трехмерное пространство, то вся механика останется ровно такой же. Просто вместо двух базовых векторов будет 3, а матрица будет иметь размер 3×3 . То же самое относится и к пространствам более высоких размерностей.

Умножение матриц

Посмотрим теперь на матрицу композиции операторов LM . В этом случае мы должны разложить по исходному базису векторы $L(M\mathbf{e}_1)$ и $L(M\mathbf{e}_2)$ и составить из них матрицу. Но также мы можем выразить это разложение через промежуточные векторы u_1 и u_2 :

$$\begin{aligned} L(M\mathbf{e}_1) &= L(u_{11}\mathbf{e}_1 + u_{21}\mathbf{e}_2) = u_{11}L(\mathbf{e}_1) + u_{21}L(\mathbf{e}_2) = \\ &= u_{11}(w_{11}\mathbf{e}_1 + w_{21}\mathbf{e}_2) + u_{21}(w_{12}\mathbf{e}_1 + w_{22}\mathbf{e}_2) = \\ &= (u_{11}w_{11} + u_{21}w_{12})\mathbf{e}_1 + (u_{11}w_{21} + u_{21}w_{22})\mathbf{e}_2, \end{aligned}$$

$$\begin{aligned} L(M\mathbf{e}_2) &= L(u_{12}\mathbf{e}_1 + u_{22}\mathbf{e}_2) = u_{12}L(\mathbf{e}_1) + u_{22}L(\mathbf{e}_2) = \\ &= u_{12}(w_{11}\mathbf{e}_1 + w_{21}\mathbf{e}_2) + u_{22}(w_{12}\mathbf{e}_1 + w_{22}\mathbf{e}_2) = \\ &= (u_{12}w_{11} + u_{22}w_{12})\mathbf{e}_1 + (u_{12}w_{21} + u_{22}w_{22})\mathbf{e}_2, \end{aligned}$$

откуда

$$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \begin{pmatrix} u_{11}w_{11} + u_{21}w_{12} & u_{12}w_{11} + u_{22}w_{12} \\ u_{11}w_{21} + u_{21}w_{22} & u_{12}w_{21} + u_{22}w_{22} \end{pmatrix}.$$

Отметим, что при умножении матрицы располагаются в том же порядке, в каком действуют соответствующие им линейные операторы в композиции операторов: композиции $L \circ M$ соответствует произведение матриц $W \cdot U$. В общем случае произведение матриц некоммутативно.

Произведение векторов

Определим скалярное произведение векторов $x = (x_1, x_2)$ и $y = (y_1, y_2)$, заданных координатами в базисе $\mathbf{e}_1, \mathbf{e}_2$ по правилу

$$\langle x, y \rangle = x_1y_1 + x_2y_2. \quad (13.8)$$

Подчеркнем, что такое представление скалярного произведения векторов справедливо только в стандартном базисе $(1, 0)$, $(0, 1)$ (или базисе, который получается из стандартного движением).

В общем случае для *вещественного* линейного пространства V **скалярное произведение** вводится как функция $V \times V \rightarrow \mathbb{R}$ от двух векторов со следующими свойствами:

$$\text{SP1 } \langle x, y \rangle = \langle y, x \rangle;$$

$$\text{SP2 } \langle (x + y), z \rangle = \langle x, z \rangle + \langle y, z \rangle;$$

$$\text{SP3 } \langle (\lambda x), y \rangle = \lambda \langle x, y \rangle;$$

$$\text{SP4 } \langle x, x \rangle \geq 0, \langle x, x \rangle = 0 \iff x = 0.$$

Несложно проверить, что скалярное произведение, заданное по формуле (13.8) через координаты векторов в стандартном базисе, удовлетворяет данным требованиям.

Скалярное произведение позволяет в общем случае определить понятие ортонормированного базиса. Так, мы говорим, что векторы x и y **ортгоналичны**, если $\langle x, y \rangle = 0$ (обозначение: $x \perp y$). Кроме того, можно ввести понятие **нормы** (или *длины*) вектора: $\|x\| = \sqrt{\langle x, x \rangle}$. В таком случае **ортонормированным базисом** (ОНБ) называется такой набор векторов $\{x_i\}$, что все они имеют норму, равную 1, и все они попарно ортогональны ($x_i \perp x_j$ при $i \neq j$). Легко видеть, что если скалярное произведение задано по формуле (13.8), то стандартные базисные векторы $(1, 0)$, $(0, 1)$ образуют ОНБ.

Для наших целей координатная запись скалярного произведения понадобится, чтобы упростить запись произведения матриц: ровно по такому правилу получают элементы произведения, если представить, что мы скалярно умножаем строки первой матрицы на столбцы второй

$$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \begin{pmatrix} \langle w_{1\bullet}, u_{\bullet 1} \rangle & \langle w_{1\bullet}, u_{\bullet 2} \rangle \\ \langle w_{2\bullet}, u_{\bullet 1} \rangle & \langle w_{2\bullet}, u_{\bullet 2} \rangle \end{pmatrix}, \quad (13.9)$$

где значок \bullet вместо индекса означает, что мы используем вектор, который составлен из координат, получаемых заменой знака \bullet на допустимые индексы.

Произведение матриц, заданное в виде (13.9), легко обобщается на случай трехмерного пространства и пространств более высоких размерностей.

Отметим одну особенность обозначения скалярного произведения, которая пришла к нам из физики. Представим себе, что мы работаем в стандартном ОНБ и что векторы x и y записываем как векторы-столбцы через их координаты в этом базисе:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Тогда введем по определению **бра- и кет-символы Дирака**:

$$\langle x | = (x_1, x_2), \quad |x \rangle = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Иначе говоря, кет-символ — это вектор-столбец, а бра-символ — это вектор-строка.

Выше мы научились умножать квадратную матрицу на вектор по формуле (13.7), а также две квадратные матрицы по правилу (13.9). На самом деле это легко обобщить на случай матриц произвольной размерности и не обязательно квадратных, главное — чтобы в матрице, стоящей слева в произведении матриц, было столько же столбцов, сколько строк во второй матрице. В частности, это позволяет рассмотреть произведение двух векторов как произведение матриц:

$$\langle x | \cdot | y \rangle = (x_1, x_2) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = x_1 y_1 + x_2 y_2,$$

а ведь это в точности соответствует формуле (13.8). Отсюда происходит второе обозначение скалярного произведения: $\langle x | y \rangle$, которое по сути заменяет собой более сложное $\langle x | \cdot | y \rangle$.

Существует и еще более интересное обозначение, включающее три параметра: $\langle x | A | y \rangle$. По определению это скалярное произведение $\langle x, Ay \rangle$, где A — квадратная матрица линейного оператора. Нетрудно проверить, что отображение $\langle x | A | y \rangle$, как функция от двух векторов x, y из одного и того же линейного пространства, в силу свойств скалярного произведения является линейным по каждому из векторов x и y . Отображение $\langle x | A | y \rangle$ называется *сверткой оператора* A с бра-вектором $\langle x |$ и кет-вектором $| y \rangle$.

Можно задаться вопросом: почему указанная запись свертки оператора раскрывается именно таким способом: $\langle x | A | y \rangle = \langle x, Ay \rangle$, а не наоборот: $\langle xA, y \rangle$? Во-первых, заметим, что произведение xA не определено корректно для квадратной матрицы A , а во-вторых, с некоторыми поправками мы все-таки можем объединить x и A в одной компоненте скалярного произведения. Для этого мы должны от матрицы A перейти к транспонированной матрице A^T , которая получается из исходной матрицы отражением элементов относительно главной диагонали. Например,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad \langle x |^T = | x \rangle.$$

Транспонирование матриц согласовано с их произведением по следующему правилу:

$$(AB)^T = B^T A^T,$$

что достаточно просто проверить, пользуясь формулой (13.9). Тогда мы получаем, что

$$\langle x | A | y \rangle = \langle x, Ay \rangle = | x \rangle^T (A^T)^T | y \rangle = | A^T x \rangle^T | y \rangle = \langle A^T x, y \rangle.$$

Отметим также, что существует класс операторов, для которых это равенство становится еще проще. Скажем, что оператор A называется **эрмитовым**, если выполняется тождество $\langle x, Ay \rangle = \langle Ax, y \rangle$ для всех векторов x и y . В случае вещественного линейного пространства матрица эрмитова оператора обладает свойством *самосопряженности*: $A = A^T$. Таким образом, для эрмитова оператора мы получаем тождество

$$\langle x|A|y \rangle = \langle x, Ay \rangle = \langle Ax, y \rangle.$$

Эрмитовы операторы играют важную роль в квантовой механике.

Умножение числа на матрицу

Умножение числа на матрицу соответствует умножению этого же числа на оператор. В самом деле, если оператору L соответствует матрица W , то оператору kL , очевидно, соответствует матрица kW , где все элементы матрицы необходимо умножить на число k , и тогда результирующий вектор kLv будет в k раз больше вектора Lv по установленным нами правилам умножения матрицы на вектор.

Вычисление матриц сложных подобий

Ранее мы нашли матрицы поворота, гомотетии и отражения. Однако существуют еще два вида подобий плоскости: поворотная гомотетия и отраженная гомотетия. Поскольку они являются композицией более простых подобий, матрицы которых нам уже известны, для получения их матриц нужно воспользоваться произведением матриц:

$$RH_\alpha^k = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = \begin{pmatrix} k \cos \alpha & -k \sin \alpha \\ k \sin \alpha & k \cos \alpha \end{pmatrix},$$

$$SH_{Ox}^k = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} = \begin{pmatrix} k & 0 \\ 0 & -k \end{pmatrix},$$

где RH_α^k — поворотная гомотетия с центром в точке O , SH_{Ox}^k — отраженная гомотетия с центром O и осью отражения Ox .

Обратимые матрицы и операторы

Итак, мы видим, что на множестве всех квадратных матриц можно задать точно такие же операции, как и на операторах, причем между матрицами и операторами существует взаимно однозначное соответствие, если мы зафиксировали некоторый базис линейного пространства. В другом базисе матрицы операторов будут, вообще говоря, иметь

другой набор чисел в своих ячейках. Но такое соответствие между операторами и матрицами говорит нам о том, что вся арифметика операторов полностью воспроизводится в арифметике матриц, что мы и видим в предыдущих вычислениях. Более того, поскольку композиция операторов ассоциативна, а композиции соответствует произведение матриц, то и произведение матриц также ассоциативно.

Кроме того, существует единичная матрица, соответствующая оператору id . В двумерном случае она выглядит так:

$$E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

На самом деле, для любого числа измерений пространства матрица оператора id будет иметь такой же вид: на главной диагонали матрицы стоят единицы (того поля или кольца, над которым заданы векторы), а на всех остальных местах — нули (того же поля или кольца). Причем вид матрицы оператора id не зависит от выбора базиса!

Таким образом, квадратные матрицы образуют моноид по умножению. Существуют ли обратные элементы к матрицам? Ответ напрямую связан с обратимостью соответствующих линейных операторов.

Если задан обратимый линейный оператор L , то существует обратная функция L^{-1} , которая, как нетрудно проверить, также является линейным оператором. А значит, существует матрица обратного линейного оператора, которая при умножении на матрицу исходного оператора даст единичную матрицу.

Пусть дана матрица W , найдем ей обратную. Обратная матрица U должна удовлетворять уравнению $WU = UW = E$:

$$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \begin{pmatrix} u_{11}w_{11} + u_{21}w_{12} & u_{12}w_{11} + u_{22}w_{12} \\ u_{11}w_{21} + u_{21}w_{22} & u_{12}w_{21} + u_{22}w_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Можно решать эти 4 уравнения, а можно заметить, что

$$\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \begin{pmatrix} w_{22} & -w_{12} \\ -w_{21} & w_{11} \end{pmatrix} = \begin{pmatrix} w_{11}w_{22} - w_{12}w_{21} & 0 \\ 0 & w_{11}w_{22} - w_{12}w_{21} \end{pmatrix}.$$

Следовательно, для получения единичной матрицы требуется разделить вторую матрицу на число $w_{11}w_{22} - w_{12}w_{21}$, и мы получим, что

$$W^{-1} = \begin{pmatrix} \frac{w_{22}}{w_{11}w_{22} - w_{12}w_{21}} & \frac{-w_{12}}{w_{11}w_{22} - w_{12}w_{21}} \\ \frac{-w_{21}}{w_{11}w_{22} - w_{12}w_{21}} & \frac{w_{11}}{w_{11}w_{22} - w_{12}w_{21}} \end{pmatrix}.$$

Отсюда видно, что если число $w_{11}w_{22} - w_{12}w_{21}$ отлично от нуля, то матрица W обратима. На самом деле верно и обратное. Действительно,

предположим, что $w_{11}w_{22} = w_{12}w_{21}$. Тогда либо $w_{21}/w_{11} = w_{22}/w_{12} = \lambda$, т. е. вторая строка матрицы W пропорциональна первой, либо некоторая строка или некоторый столбец матрицы W содержат только нули.

Рассмотрим первый вариант:

$$W = \begin{pmatrix} w_{11} & w_{12} \\ \lambda w_{11} & \lambda w_{12} \end{pmatrix}.$$

Тогда

$$WU = \begin{pmatrix} u_{11}w_{11} + u_{21}w_{12} & u_{12}w_{11} + u_{22}w_{12} \\ \lambda(u_{11}w_{11} + u_{21}w_{12}) & \lambda(u_{12}w_{11} + u_{22}w_{12}) \end{pmatrix},$$

т. е. мы получим матрицу, у которой также строки пропорциональны с коэффициентом λ . Но нам нужно получить равенство $WU = E_2$, строки которой не пропорциональны. Противоречие.

Во втором случае сначала предположим, что в матрице W одна строка состоит из нулей, например, первая:

$$\begin{pmatrix} 0 & 0 \\ w_{21} & \lambda w_{22} \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} = \begin{pmatrix} 0 & \dots \\ \dots & \dots \end{pmatrix}.$$

Как видим, мы получаем ноль там, где должна стоять единица. Аналогично и случай второй нулевой строки сводится к противоречию.

Наконец, если в матрице W один из столбцов состоит из нулей, то легко показать, что тогда в матрице WU строки будут пропорциональными, а значит, WU не может быть единичной матрицей.

С числом $w_{11}w_{22} - w_{12}w_{21}$ очень многое связано в линейной алгебре! Это число называется **определителем матрицы W** и обозначается обычно $|W|$ или $\det W$.

Определитель матрицы и перестановки

Посмотрим на определитель квадратной матрицы 2×2 с точки зрения перестановок. Пусть

$$W = [w^{(1)}; w^{(2)}] = \begin{pmatrix} w_1^{(1)} & w_1^{(2)} \\ w_2^{(1)} & w_2^{(2)} \end{pmatrix}.$$

Тогда

$$\det W = w_1^{(1)}w_2^{(2)} - w_2^{(1)}w_1^{(2)}.$$

Пусть номера векторов — это элементы множества $\{1, 2\}$, на котором выполняется перестановка, а номера координат, выбранных у этих векторов, — это перестановка номеров векторов. Тогда произведение

$w_1^{(1)}w_2^{(2)}$ можно соотнести с тождественной перестановкой $\text{id} = (1)(2)$, а произведение $w_2^{(1)}w_1^{(2)}$ — с инверсной перестановкой (12) . Заметим также, что в случае переставленных номеров перед произведением появляется минус. В то же время перестановка (12) является нечетной и, таким образом, ее знак тоже равен -1 .

С этой точки зрения определитель матрицы 2×2 задается формулой:

$$\det W = \sum_{\sigma \in S_2} \text{sgn}(\sigma) w_{\sigma(1)}^{(1)} w_{\sigma(2)}^{(2)},$$

где суммирование ведется по всем перестановкам из указанной группы.

Таким же способом вводится определение для определителя квадратной матрицы произвольной размерности $n \times n$. Пусть матрица W составлена из векторов-столбцов $w^{(1)}, w^{(2)}, \dots, w^{(n)}$, каждый из которых состоит из n координат, тогда ее определитель будет равен

$$\det W = \begin{vmatrix} w_1^{(1)} & w_1^{(2)} & \dots & w_1^{(n)} \\ w_2^{(1)} & w_2^{(2)} & \dots & w_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ w_n^{(1)} & w_n^{(2)} & \dots & w_n^{(n)} \end{vmatrix} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) w_{\sigma(1)}^{(1)} w_{\sigma(2)}^{(2)} \dots w_{\sigma(n)}^{(n)}. \quad (13.10)$$

Таким образом, в определитель входит $n!$ слагаемых.

Следующие графические схемы помогают запомнить правило вычисления определителя третьего порядка:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = + \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

или

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} + & + & + \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ + & + & + \end{matrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Здесь подразумевается сумма тернарных произведений элементов, связанных одной фигурой (линией или треугольником), с соответствующим знаком, выбор которого указан цветом и значками $+$ и $-$.

Обозначим за E_n единичную матрицу размерности $n \times n$, т. е.

$$E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Пользуясь определением (13.10), видим, что для матрицы E_n ненулевым будет только такое слагаемое, где $\sigma = \text{id}$, откуда $\det E_n = 1$.

В определении (13.10) мы использовали перестановки, которые номерам столбцов ставили в соответствие номера строк матрицы W . Но как мы знаем, перестановки на n символах образуют группу, так что у каждой перестановки существует обратная, принадлежащая той же самой группе. Пусть перестановка σ переводит номер столбца в номер строки, тогда перестановку σ^{-1} можно рассматривать как соответствие номеров строк номерам столбцов. Перепишем определитель матрицы W через обратные перестановки:

$$\det W = \sum_{\sigma \in \mathbf{S}_n} \text{sgn}(\sigma) w_1^{(\sigma^{-1}(1))} w_2^{(\sigma^{-1}(2))} \dots w_n^{(\sigma^{-1}(n))}.$$

Теперь посмотрим на это выражение, как на определитель матрицы с элементами $u_j^{(k)} = w_k^{(j)}$, и получим

$$\det W = \sum_{\sigma \in \mathbf{S}_n} \text{sgn}(\sigma) u_{\sigma^{-1}(1)}^{(1)} u_{\sigma^{-1}(2)}^{(2)} \dots u_{\sigma^{-1}(n)}^{(n)} = \sum_{\sigma \in \mathbf{S}_n} \text{sgn}(\sigma) u_{\sigma(1)}^{(1)} u_{\sigma(2)}^{(2)} \dots u_{\sigma(n)}^{(n)},$$

поскольку $\text{sgn}(\sigma) = \text{sgn}(\sigma^{-1})$ и произвольный выбор $\sigma \in \mathbf{S}_n$ эквивалентен произвольному выбору $\sigma^{-1} \in \mathbf{S}_n$. Отсюда мы видим, что

$$\det W = \begin{vmatrix} u_1^{(1)} & u_1^{(2)} & \dots & u_1^{(n)} \\ u_2^{(1)} & u_2^{(2)} & \dots & u_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ u_n^{(1)} & u_n^{(2)} & \dots & u_n^{(n)} \end{vmatrix} = \begin{vmatrix} w_1^{(1)} & w_2^{(1)} & \dots & w_n^{(1)} \\ w_1^{(2)} & w_2^{(2)} & \dots & w_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ w_1^{(n)} & w_2^{(n)} & \dots & w_n^{(n)} \end{vmatrix},$$

т. е. определитель матрицы W равен определителю матрицы, полученной из W транспонированием. Итак,

$$\det W = \det W^T.$$

С помощью перестановок легко получить следующее свойство определителей: если в матрице W поменять местами два столбца (или две строки) j и k , то определитель сменит знак на противоположный. Это легко объясняется тем, что перестановка двух столбцов матрицы соответствует умножению перестановки σ в формуле (13.10) на транспозицию (jk) , которая меняет четность.

Такое свойство определителя называется **кососимметричностью**.

Нетрудно видеть также, что если для некоторого столбца $w^{(k)}$ имеет место равенство $w^{(k)} = \alpha u + \beta v$, где векторы u, v имеют ту же размерность, что и вектор $w^{(k)}$, а α, β — произвольные числа, то

$$\det W = \alpha \det[w^{(1)}, \dots, w^{(k-1)}, u, w^{(k+1)}, \dots, w^{(n)}] + \beta \det[w^{(1)}, \dots, w^{(k-1)}, v, w^{(k+1)}, \dots, w^{(n)}]. \quad (13.11)$$

Действительно, для каждого слагаемого в сумме (13.10) получаем

$$\begin{aligned} w_{\sigma(1)}^{(1)} w_{\sigma(2)}^{(2)} \cdots w_{\sigma(n)}^{(n)} &= w_{\sigma(1)}^{(1)} \cdots w_{\sigma(k-1)}^{(k-1)} (\alpha u_{\sigma(k)} + \beta v_{\sigma(k)}) w_{\sigma(k+1)}^{(k+1)} \cdots w_{\sigma(n)}^{(n)} = \\ &= \alpha w_{\sigma(1)}^{(1)} \cdots w_{\sigma(k-1)}^{(k-1)} u_{\sigma(k)} w_{\sigma(k+1)}^{(k+1)} \cdots w_{\sigma(n)}^{(n)} + \\ &\quad + \beta w_{\sigma(1)}^{(1)} \cdots w_{\sigma(k-1)}^{(k-1)} v_{\sigma(k)} w_{\sigma(k+1)}^{(k+1)} \cdots w_{\sigma(n)}^{(n)}, \end{aligned}$$

откуда легко видеть равенство (13.11). Это свойство определителя называется *полилинейностью*. Отметим также, что полилинейность выполняется не только для столбцов, но и для строк матрицы W , поскольку определитель сохраняется при транспонировании, как мы уже показали выше.

Определитель и площадь

Рассмотрим на координатной плоскости параллелограмм, построенный на векторах u и v . Пусть эти векторы заданы координатами

$$u = (a, b), \quad v = (c, d).$$

Найдем площадь этого параллелограмма, глядя на картинку 13.10.

Легко видеть, что площадь S параллелограмма вычисляется по формуле:

$$S = |ad - bc| = \left| \det \begin{pmatrix} a & c \\ b & d \end{pmatrix} \right|.$$

Если мы повернем параллелограмм на рисунке 13.10 вокруг точки O на произвольный угол α , то координаты векторов u и v изменятся под действием оператора поворота R_α (см. раздел 13.4) следующим образом:

$$u' = R_\alpha(u), \quad v' = R_\alpha(v).$$

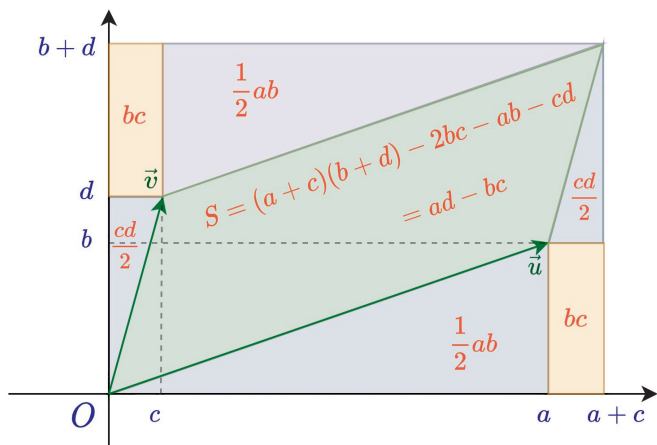


Рис. 13.10. Площадь параллелограмма и определитель.

Вычислим их:

$$u' = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \cos \alpha - b \sin \alpha \\ a \sin \alpha + b \cos \alpha \end{pmatrix}, \quad v' = \begin{pmatrix} c \cos \alpha - d \sin \alpha \\ c \sin \alpha + d \cos \alpha \end{pmatrix}.$$

Найдем определитель матрицы $[u'; v']$, он равен

$$\begin{aligned} u'_1 v'_2 - u'_2 v'_1 &= \\ &= (a \cos \alpha - b \sin \alpha)(c \sin \alpha + d \cos \alpha) - (a \sin \alpha + b \cos \alpha)(c \cos \alpha - d \sin \alpha) = \\ &= ad - bc. \end{aligned}$$

То есть при повороте на угол α определитель остался тем же самым, причем с сохранением знака. Кроме того, очевидно, что наш параллелограмм также не изменился при повороте, так что его площадь S вычисляется по формуле $|ad - bc|$ независимо от того, как на плоскости выбрана система координат.

Пусть даны три вектора в пространстве:

$$u = (a, b, c), \quad v = (d, e, f), \quad w = (g, h, j).$$

На этих трех векторах можно построить параллелепипед, объем V которого вычисляется через определитель матрицы, составленной из этих векторов:

$$V = \pm \begin{vmatrix} a & d & g \\ b & e & h \\ c & f & j \end{vmatrix},$$

где символ \pm указывает на то, что мы должны выбрать положительный знак, т. к. объем фигуры — величина положительная.

Если в качестве числовой характеристики параллелепипеда рассматривать именно определитель матрицы, составленной из векторов, которыми он задан, то такая характеристика называется **ориентированным объемом** (а в случае параллелограмма — **ориентированной площадью**).

На самом деле, в произвольном n -мерном евклидовом пространстве ориентированный объем n -мерного параллелепипеда в точности равен определителю матрицы, составленной из векторов, задающих данный параллелепипед.

Из этой связи между объемом (площадью) и определителем легко усмотреть случай, когда определитель равен нулю. Если векторы, задающие определитель, линейно зависимы, т. е. один из них есть линейная комбинация остальных, то объем фигуры равен нулю, т. к. параллелограмм схлопывается в отрезок, а параллелепипед — в многоугольник, т. е. плоскую фигуру, объем которой равен нулю.

Таким образом, из геометрических соображений легко понять, что в случае линейно зависимых столбцов матрицы ее определитель равен нулю. В силу свойства сохранения определителя при транспонировании матрицы верно также и то, что в случае линейно зависимых строк матрицы ее определитель равен нулю. Для случая матрицы 2×2 мы уже видели это, когда доказывали критерий существования обратной матрицы: при нулевом определителе мы получали пропорциональные или тождественно нулевые столбцы или строки.

Свойства определителя

Выше мы показали, что определитель как функция от квадратных матриц обладает свойствами кососимметричности и полилинейности. Для двумерного случая линейность по любому из векторов, задающих матрицу, можно продемонстрировать с помощью сложения и растяжения соответствующего параллелограмма.

На рис. 13.11 показано, что происходит с площадью параллелограмма, если одну из его сторон представить как сумму векторов $u_1 + u_2$, а также если одну из его сторон вытянуть в α раз. В первом случае площадь будет равна сумме исходных площадей, во втором — площадь увеличится в α раз. Именно так и ведет себя определитель.

На самом деле, все функции от квадратных матриц размерности $n \times n$, обладающие свойствами кососимметричности и полилинейности, отличаются друг от друга лишь коэффициентом, а если потребовать, чтобы такая функция на определенной матрице принимала какое-то определенное значение, то она и вовсе будет единственной!

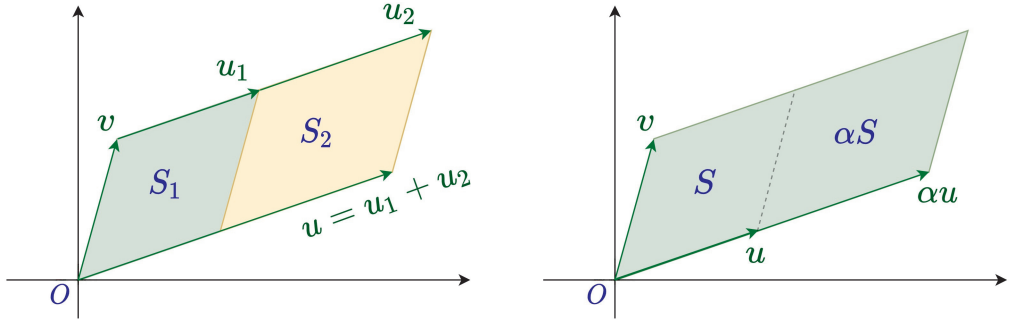


Рис. 13.11. Полилинейность определителя 2-го порядка.

Теорема 13.5 (об универсальности определителя).

Пусть $f(W)$ — вещественно-значная функция от квадратных вещественных матриц размерности $n \times n$, обладающая свойствами кососимметричности и полилинейности относительно строк и столбцов матрицы W . Тогда

$$f(W) = f(E_n) \det W.$$

Доказательство. Рассмотрим для простоты только двумерный случай. Для начала заметим, что в силу свойства полилинейности функции f относительно строк матрицы, легко получить равенства:

$$f \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = f \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = 0.$$

Далее, пусть $W = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, тогда заметим, что

$$\begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

поэтому, пользуясь свойством полилинейности f относительно столбцов матрицы W , получаем, что

$$f(W) = af \begin{pmatrix} 1 & c \\ 0 & d \end{pmatrix} + bf \begin{pmatrix} 0 & c \\ 1 & d \end{pmatrix} = af[\mathbf{e}_1; v] + bf[\mathbf{e}_2; v],$$

где

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} c \\ d \end{pmatrix}.$$

Воспользуемся свойством линейности f для вектора $v = c\mathbf{e}_1 + d\mathbf{e}_2$:

$$f(W) = a(cf[\mathbf{e}_1; \mathbf{e}_1] + df[\mathbf{e}_1; \mathbf{e}_2]) + b(cf[\mathbf{e}_2; \mathbf{e}_1] + df[\mathbf{e}_2; \mathbf{e}_2]).$$

Теперь, так как в матрицах $[e_1; e_1]$ и $[e_2; e_2]$ одна из строк состоит только из нулей, то, как уже отмечалось, $f[e_1; e_1] = f[e_2; e_2] = 0$. Кроме того, в силу кососимметричности функции f относительно столбцов матрицы, получаем, что $f[e_1; e_2] = -f[e_2; e_1]$.

Таким образом,

$$f(W) = a(c \cdot 0 + d \cdot f(E_2)) + b(c \cdot (-f(E_2)) + d \cdot 0) = f(E_2)(ad - bc) = f(E_2) \det W.$$

Аналогично теорема доказывается для случая произвольной размерности квадратной матрицы. \square

Следствие 13.2. $\det(AW) = \det A \det W$ для любых матриц размерности $n \times n$.

Доказательство. Рассмотрим функцию $f(W) = \det(AW)$.

Матрица AW состоит из столбцов $Aw^{(1)}, \dots, Aw^{(n)}$, где $W = [w^{(1)}; \dots; w^{(n)}]$. Каждый из столбцов $Aw^{(k)}$ является результатом действия на вектор $w^{(k)}$ линейного оператора, соответствующего матрице A , откуда следует, что если $w^{(k)} = \alpha u + \beta v$, то $Aw^{(k)} = \alpha Au + \beta Av$. Далее, поскольку определитель $\det(AW)$ обладает свойством линейности по столбцу $Aw^{(k)}$, мы получаем, что

$$\det(AW) = \alpha \det[Aw^{(1)}, \dots, Au, \dots, Aw^{(n)}] + \beta \det[Aw^{(1)}, \dots, Av, \dots, Aw^{(n)}],$$

откуда следует, что функция f полилинейна по столбцам матрицы W .

Аналогично можно показать, что f полилинейна также и по строкам матрицы W .

Далее, если в матрице W поменять местами два столбца, то и в матрице AW столбцы с теми же номерами поменяются местами, что приведет к смене знака определителя $\det(AW)$. Отсюда следует, что перестановка столбцов матрицы W приводит к смене знака функции $f(W)$. Аналогичное рассуждение применяем к строкам матрицы W .

Таким образом, функция f является кососимметричной и полилинейной относительно строк и столбцов матрицы W , а значит, по доказанной теореме получаем, что

$$f(W) = f(E_n) \det W = \det(AE_n) \det W = \det A \det W.$$

\square

Из свойства мультипликативности определителя следует также, что определитель обратной матрицы вычисляется по формуле $\det(W^{-1}) = 1/\det W$, если $\det W \neq 0$.

Суммируя полученные результаты, приведем следующие общие свойства определителя:

- 1) $\det \mathbf{E}_n = 1$;
- 2) $\det(WU) = \det W \det U$;
- 3) $\det(W^{-1}) = (\det W)^{-1}$, если $\det W \neq 0$;
- 4) $\det(W^T) = \det W$;
- 5) определитель равен нулю тогда и только тогда, когда столбцы матрицы линейно зависимы;
- 6) определитель равен нулю тогда и только тогда, когда строки матрицы линейно зависимы.

Подведем итог. При выборе базиса пространства (плоскости) каждому линейному оператору соответствует единственная квадратная матрица, и каждой квадратной матрице (при данном базисе) соответствует единственный линейный оператор. Операции на матрицах в точности соответствуют операциям на операторах, т. е. мы имеем изоморфизм между операторами и квадратными матрицами. И мы приходим к тому, что алгебра линейных операторов изоморфна алгебре квадратных матриц.

Более того, оператор обратим тогда и только тогда, когда его матрица имеет определитель, отличный от нуля. Оператору id соответствует единичная матрица, обратному оператору соответствует обратная матрица. Соответственно, полной линейной группе $\text{GL}(V)$ обратимых линейных операторов над пространством векторов V изоморфна группа обратимых квадратных матриц, размерность которых равна размерности пространства V . Такая группа матриц обозначается через $\text{GL}(n)$ для n -мерного пространства.

В группе $\text{GL}(n)$ можно выделить ряд подгрупп. Например, если мы рассматриваем матрицы с вещественными коэффициентами, то мы можем выделить подгруппу всех **ортогональных матриц**, т. е. таких матриц, что $A^T A = \mathbf{E}_n$. Такая подгруппа обозначается $\mathbb{O}(n)$ и называется **ортогональной группой**.

Следующий пример: **специальная линейная подгруппа** группы $\text{GL}(n)$, которая обозначается как $\text{SL}(n)$ и состоит из всех матриц, определитель которых по модулю равен 1.

Элементы ортогональной группы, модуль определителя которых в точности равен 1, образуют подгруппу ортогональной группы — **специальную ортогональную группу** $\text{SO}(n)$, которая также является подгруппой $\text{SL}(n)$.

Матрицы группы $\text{SL}(n)$ соответствуют операторам поворота относительно начала координат и отражения относительно оси, проходящей через начало координат. Матрицы группы $\text{SO}(n)$ соответствуют операторам поворота относительно начала координат.

13.6. Матрицы и комплексные числа

Ранее мы уже показали, что комплексные числа позволяют «оцифровать» движения плоскости. А именно, функция $T(z) = z + z_0$ есть параллельный перенос на плоскости \mathbb{C} , функция $R(z) = zz_0/|z_0|$ есть поворот относительно нуля, функция $S(z) = \bar{z}$ есть отражение относительно вещественной оси. За кадром остался вопрос о том, что же получится, если рассмотреть операцию умножения на произвольное комплексное число. Теперь уже должно быть очевидно, что это будет поворотная гомотетия с центром в нуле. Чуть ниже мы это докажем.

Мы также нашли, что тригонометрические функции появляются как компоненты комплексного числа с единичной окружности, а именно:

$$\operatorname{Re} z = \cos \varphi, \quad \operatorname{Im} z = \sin \varphi.$$

С другой стороны, мы знаем вид матрицы поворота:

$$R_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix},$$

что очень похоже на запись комплексного числа, олицетворяющего поворот плоскости относительно нуля.

На самом деле, умножение на всякое комплексное число z_0 можно интерпретировать как действие линейного оператора, поскольку

$$(\lambda z + \mu z')z_0 = \lambda z z_0 + \mu z' z_0, \quad \lambda, \mu \in \mathbb{R},$$

т. е. оно подчиняется аксиомам линейности.

Но если $z_0 = x + iy$ понимать как линейный оператор, то легко найти его матрицу в стандартном базисе, проследив за тем, куда переходят единичные базисные векторы: $\bar{1} = (1, 0)$ и $\bar{i} = (0, 1)$. Первый переходит в (x, y) , а второй, в соответствии с арифметикой комплексных чисел, — в вектор $(-y, x)$, ибо это $i(x + iy) = -y + ix$. Следовательно, матрицей линейного оператора, соответствующего домножению на число $z_0 = x + iy$, является двухпараметрическая матрица

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

Легко проверить, что алгебра таких матриц в точности соответствует алгебре комплексных чисел: сложение чисел переходит в сложение матриц, умножение чисел — в умножение матриц. Кроме того, комплексное сопряжение соответствует транспонированию матрицы:

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}^T = \begin{pmatrix} x & y \\ -y & x \end{pmatrix},$$

а определитель матрицы соответствует норме комплексного числа:

$$\det \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = x^2 + y^2.$$

Осталось вспомнить, что матрицей поворотной гомотетии с центром в нуле и коэффициентом λ является

$$RH_\varphi^\lambda = \begin{pmatrix} \lambda \cos \varphi & -\lambda \sin \varphi \\ \lambda \sin \varphi & \lambda \cos \varphi \end{pmatrix},$$

что соответствует домножению на комплексное число $\lambda \cos \varphi + i \lambda \sin \varphi$.

Наконец, матрицей отраженной гомотетии с центром в нуле и осью вещественных чисел в качестве оси отражения является

$$SH_{Ox}^\lambda = \begin{pmatrix} k & 0 \\ 0 & -k \end{pmatrix},$$

что соответствует домножению на действительное число λ с последующим сопряжением результата.

Итак, в терминах комплексных функций получается, что всякое подобие есть линейная функция от z , либо ее сопряжение. Обратное также верно.

Теорема 13.6. *Преобразование $z \mapsto w$ является преобразованием подобия тогда и только тогда, когда оно имеет вид:*

$$w = az + b, \text{ либо } w = a\bar{z} + b,$$

где a, b — произвольные комплексные числа, $|a| > 0$.

Доказательство. Учитывая наши предыдущие построения, нам осталось доказать только обратное утверждение.

Пусть $w = az + b$. Представим число a в виде матрицы

$$\begin{pmatrix} |a| \cos \varphi & -|a| \sin \varphi \\ |a| \sin \varphi & |a| \cos \varphi \end{pmatrix}.$$

Это — матрица поворотной гомотетии, как мы видели выше. Следовательно, преобразование $w = az + b$ является композицией сдвига на вектор b и поворотной гомотетии:

$$w = T_b \circ RH_\varphi^{|a|}(z).$$

Пусть $w = a\bar{z} + b$. В данном случае нам еще потребуется композиция с оператором отражения относительно действительной оси, и в итоге мы получим, что

$$w = T_b \circ RH_\varphi^{|a|} \circ S^{\text{Re}}(z).$$

Отметим, что требование $|a| > 0$ существенно, т. к. при $|a| = 0$ мы получим «схлопывание» всей комплексной плоскости в одну точку b , а это отображение уже не является преобразованием. \square

Стоит отметить также, что преобразование вида $w = az + b\bar{z} + c$, где $|a| \neq |b|$, представляет собой так называемое **аффинное преобразование**, при котором параллельные прямые переходят в параллельные прямые, а пересекающиеся — в пересекающиеся. Подобия являются частным случаем такого преобразования (либо $a = 0$, либо $b = 0$).

Наконец, преобразование вида $w = (az + b)/(cz + d)$, где $ad \neq bc$, представляет собой **дробно-линейное преобразование** — основной инструмент проективной геометрии. При дробно-линейном преобразовании окружности и прямые переходят в окружности и прямые (в том числе, окружность может перейти в прямую, и наоборот).

13.7. Действие линейных операторов

Поскольку линейный оператор сохраняет операцию сложения, он является гомоморфизмом векторного пространства V на себя. А поскольку V есть группа по сложению, то естественно вспомнить о том, что ядро гомоморфизма является (нормальной) подгруппой. Это значит, что для любого линейного оператора $L : V \rightarrow V$ множество

$$\ker L = \{v \in V \mid Lv = 0\}$$

замкнуто относительно операции сложения векторов, а кроме того, пространство V можно представить как сумму попарно непересекающихся множеств вида $\ker(L) + w$, где w — произвольный вектор пространства V .

Кроме того, ядро линейного оператора замкнуто и относительно операции умножения на число, т. к. если $Lv = 0$, то $L(kv) = 0$. Иначе говоря, ядро линейного оператора само по себе является векторным пространством внутри пространства V . Соответственно, у него есть базис, по количеству векторов не превосходящий базис всего пространства V . Таким образом, $\ker L$ является векторным подпространством пространства V , и его размерность не превосходит размерность самого V . Если размерность ядра равна размерности V , то $\ker L = V$.

Смещенное на вектор w подпространство называется **линейным мно-гообразием**.

Интересно, что в случае линейного оператора не только его ядро, но и образ $L[V]$ является линейным подпространством пространства V , т. к. если $u, v \in L[V]$, то любая их линейная комбинация также принадлежит $L[V]$. Действительно, если $u, v \in L[V]$, то существуют такие

$u', v' \in V$, что $u = L(u')$ и $v = L(v')$. Тогда для произвольных коэффициентов α, β получим: $\alpha u + \beta v = \alpha L(u') + \beta L(v') = L(\alpha u' + \beta v')$. А так как $\alpha u' + \beta v' \in V$, то $\alpha u + \beta v \in L[V]$.

Имеет место уравнение для размерностей (которое мы оставим без доказательства):

$$\dim \ker L + \dim L[V] = \dim V.$$

В частности, отсюда следует, что $L[V] = V$ тогда и только тогда, когда ядро линейного оператора тривиально, т. е. равно $\{0\}$, а это эквивалентно тому, что L является обратимым линейным оператором.

Например, если линейный оператор действует в трехмерном пространстве и переводит плоскость в ноль, то его образом будет прямая. Простейший пример: проекция на координатную ось $\text{Pr}_z(x, y, z) = z$ (см. рис. 13.12).

Рассмотрим данный пример подробнее. Наш оператор Pr_z определен для всех векторов пространства \mathbb{R}^3 и каждому вектору (x, y, z) ставит в соответствие вектор $(0, 0, z)$. Ядром данного оператора является плоскость $(x, y, 0)$. Очевидно, что размерность ядра при этом равна 2. Образ $\text{Pr}_z[\mathbb{R}^3]$ — это прямая, точнее, ось z . Размерность образа равна 1.

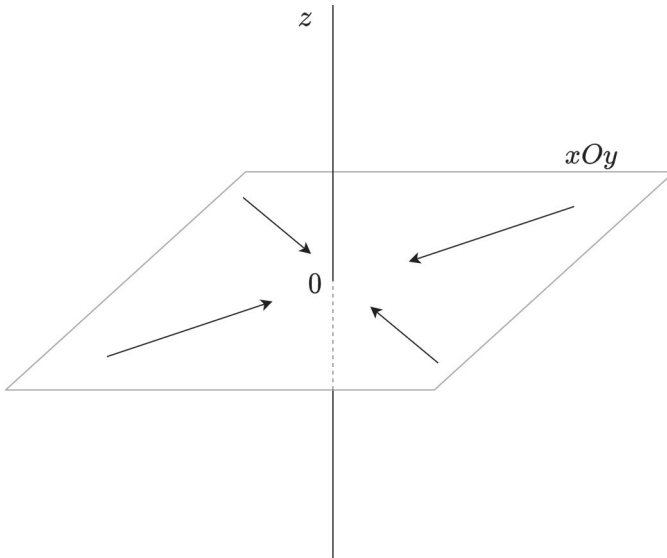


Рис. 13.12. Оператор Pr_z .

Запишем матрицу оператора Pr_z в стандартном базисе:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Здесь, как и раньше, столбцы — это образы базисных векторов относительно оператора Pr_z . Поскольку векторы, лежащие в плоскости xOy , принадлежат ядру оператора, их образом является нулевой вектор, что мы и видим. В то же время, третий базисный вектор ($z = 1$) переходит сам в себя, и его мы также видим в записи матрицы оператора Pr_z .

Заметим далее, что, поскольку матрица линейного оператора строится из векторов-столбцов, которые являются образами базисных векторов, то эти столбцы линейно независимы тогда и только тогда, когда размерность образа равна $\dim V$, т. е. когда оператор L обратим. С другой стороны, независимость столбцов матрицы эквивалентна тому, что ее определитель отличен от нуля.

Теорема 13.7. *Следующие утверждения о линейном операторе $L : V \rightarrow V$ эквивалентны:*

- (1) *линейный оператор L обратим;*
- (2) *матрица линейного оператора обратима;*
- (3) *определитель матрицы линейного оператора отличен от нуля;*
- (4) *размерность ядра L равна 0;*
- (5) *размерность образа $L[V]$ равна $\dim V$;*
- (6) $L[V] = V$.

Число линейно независимых столбцов матрицы линейного оператора (оно же — размерность образа соответствующего линейного оператора) называется **рангом матрицы**. Если W — матрица линейного оператора L , то $\text{rank}(W) = \dim L[V]$. Определитель матрицы равен нулю тогда и только тогда, когда ее ранг меньше ее размерности.

Если оператор нулевой, т. е. $Lv = 0$ для любого вектора v , то ранг матрицы такого оператора равен 0, а значит, любой столбец матрицы сам по себе не является линейно независимым, а такое возможно только в том случае, когда этот столбец нулевой. Следовательно, матрица такого оператора состоит из одних нулей.

Упражнения

Обязательные упражнения

- 13.1° В каких случаях гомотетия является движением?
- 13.2° Коммутирует ли гомотетия с поворотом, а) если их центры идентичны; б) если их центры различны? Коммутируют ли гомотетии с разными центрами?
- 13.3° Вычислить композиции: а) $H_O^2 \circ H_A^{1/2}$; б) $H_O^2 \circ R_O^{\pi/2} \circ H_O^{-1/2}$; в) $T_{OA} \circ H_A^{-1}$; д) $H_O^2 \circ H_A^3$ ($A \neq O$).
- 13.4° Доказать, что подобия прямой — это либо гомотетия, либо сдвиг (при $k = 1$).
- 13.5° Проверить, что подобия прямой и плоскости образуют группу с операцией композиции.
- 13.6° Найти все конечные подгруппы подобий прямой.
- 13.7° Доказать, что при подобии: а) прямая переходит в прямую; б) окружность — в окружность; в) отрезок — в отрезок (с сохранением внутренних точек); г) треугольник — в треугольник (вершины — в вершины).
- 13.8° Пусть дан треугольник $\triangle ABC$, и в нем проведены биссектрисы углов. Затем через середины его сторон провели прямые, параллельные этим биссектрисам (через середину BC провели прямую, параллельную биссектрисе $\angle A$, и т.д.). Доказать, что эти прямые пересекаются в одной точке. *Указание:* использовать гомотетию с центром в точке пересечения медиан $\triangle ABC$.
- 13.9° Взяли карту в одном масштабе, уменьшили ее (сделали масштаб более мелким) и положили на исходную карту так, что уменьшенная карта оказалась полностью внутри исходной. Доказать, что существует точка на карте, которая при этом совместилась сама с собою.
- 13.10° Вычислить, нарисовать на плоскости и указать модуль и аргумент следующих комплексных чисел:

$$i^2, i^3, i^4, 1/i, (1+2i)(2-i), (1+i)(1+2i)(1+3i), \frac{1}{1+i}, \frac{5}{2-i}.$$

- 13.11° Какие из следующих соответствий задают отображения между множествами X и Y ?

а) X — множество точек декартовой плоскости, Y — множество точек оси абсцисс, точке плоскости ставится в соответствие абсцисса этой точки.

б) $X = Y = \mathbb{N}$, числу $x \in X$ ставится в соответствие число x^2 .

с) $X = Y = \mathbb{Z}$, числу $y \in Y$ ставится в соответствие тем числам $x \in X$, для которых $|x| = |y|$.

д) $X = Y = \mathbb{R}$, числу $y \in Y$ ставится в соответствие тем числам $x \in X$, для которых $x^3 = y$.

е) $X = Y = \mathbb{R}$, числу $x \in X$ ставится в соответствие одно такое число $y \in Y$, что $x = y^2$.

13.12° Для каждого отображения из задачи 13.11° найдите прообраз каждого элемента $y \in Y$.¹

13.13° Выпишите все отображения множества $\{a, b, c\}$ в множество $\{0, 1\}$.

13.14° Какие из отображений задачи 13.11° являются биекциями?

13.15° Пусть X и Y — конечные множества. Определите в терминах отображений из X в Y и из Y в X следующие отношения между множествами X и Y :

а) в X не больше элементов, чем в Y ;

б) в X ровно столько же элементов, сколько в Y ;

с) в X строго меньше элементов, чем в Y .

13.16° Каких треугольников с целыми сторонами больше: а) тех, периметр которых равен 2002, или тех, периметр которых равен 2005? б) тех, периметр которых равен 2003, или тех, периметр которых равен 2006?

13.17° Докажите, что для произвольных отображений $f : X \rightarrow Y$, $g : Y \rightarrow Z$ и $h : Z \rightarrow W$ выполняется равенство $h \circ (g \circ f) = (h \circ g) \circ f$.

13.18° Пусть $f : X \rightarrow Y$, $g : Y \rightarrow Z$. Верно ли, что если f и g биекции, то и $g \circ f$ биекция?

13.19° Пусть $f : X \rightarrow Y$. Докажите, что существует f^{-1} тогда и только тогда, когда f — биекция. Докажите, что f^{-1} , если оно существует, — всегда биекция.

13.20° Найдите обратные к тем отображениям задачи 13.11°, которые являются биекциями.

¹ Имеется в виду $f^{-1}[\{y\}]$.

13.21° Докажите, что множество преобразований множества X , т.е. множество

$$\{f \mid (f : X \rightarrow X) \wedge (f - \text{биекция})\},$$

с операцией композиции образует группу.

13.22° Какое преобразование является композицией двух гомотетий с коэффициентами k_1 и k_2 , если а) $k_1 k_2 = 1$; б) $k_1 k_2 \neq 1$?

13.23° а) Даны два параллельных отрезка разной длины. Укажите все гомотетии, переводящие первый отрезок во второй. б) (Замечательное свойство трапеции) Докажите, что в любой трапеции точка пересечения диагоналей, точка пересечения продолжений боковых сторон и середины оснований лежат на одной прямой.

13.24° Какое преобразование является композицией гомотетии и параллельного переноса?

13.25° а) Даны две окружности. Укажите все гомотетии, переводящие первую во вторую. б) Даны три окружности различных радиусов. Для каждой пары окружностей нашли точку пересечения их общих внешних касательных. Докажите, что эти три точки лежат на одной прямой.

13.26° В окружности проведены два непараллельных радиуса. Постройте хорду, которая делится этими радиусами на три равные части.

13.27° Докажите, что любое преобразование подобия есть композиция гомотетии и движения.

13.28° Можно ли перевести а) любую параболу в любую другую параболу преобразованием подобия; б) график функции $y = \sin x$ в график функции $y = \sin 2x$ преобразованием подобия? А гомотетией?

13.29° Докажите, что всякое преобразование подобия с коэффициентом, не равным 1, а) имеет неподвижную точку; б) является композицией гомотетии и поворота с общим центром или композицией гомотетии и симметрии относительно оси, проходящей через центр гомотетии.

13.30° На стене висят двое часов, одни побольше, другие поменьше. Докажите, что прямые, соединяющие концы минутных стрелок в разные моменты времени, проходят через одну точку.

13.31° Какой формулой в комплексных числах можно записать: а) гомотетию с коэффициентом k с центром в нуле; б) поворотную гомотетию с коэффициентом k , углом α и центром в нуле? Для записи отображения воспользуйтесь формулой Эйлера.

13.32° Являются ли линейными пространствами (при подходящем определении операций):

- a) \mathbb{R} над \mathbb{Q} ;
- b) \mathbb{Q} над \mathbb{R} ;
- c) $\mathbb{Z}[i]$ над \mathbb{Z} ;
- d) множество всех векторов на плоскости над \mathbb{R} ;
- e) многочлены с коэффициентами из поля K (обозначение: $K[x]$) над K ;
- f) многочлены степени не выше n над \mathbb{R} ; ровно степени n и нулевой многочлен, над \mathbb{R} ;
- g) многочлены над \mathbb{R} , равные в точке $x = 1$ нулю; единице;
- h) строки (или столбцы) из n элементов поля K (обозначение: K^n);
- i) бесконечные последовательности действительных чисел;
- j) последовательности Фибоначчи (последовательности, удовлетворяющие условию $x_{n+1} = x_{n-1} + x_n$);
- k) множество решений однородной системы линейных уравнений; неоднородной.

13.33° Докажите, что в линейном пространстве векторы e_1, \dots, e_n ($n \geq 2$) независимы тогда и только тогда, когда один из них линейно выражается через остальные.

13.34° Являются ли линейно независимыми векторы следующих множеств: a) $\{(1, -1, 0), (-1, 0, 1), (0, 1, -1)\} \subset \mathbb{R}^3$; b) $\{(1, 1, 0), (1, 0, 1), (0, 1, 1)\} \subset \mathbb{R}^3$?

13.35° Доказать, что если система векторов $\{e_1, \dots, e_n\}$ линейно независима, а система векторов $\{e_1, \dots, e_n, b\}$ линейно зависима, то вектор b линейно выражается через векторы e_1, \dots, e_n .

13.36° Доказать, что если $\{e_1, \dots, e_n\}$ — базис пространства V , то для всякого $x \in V$ его координаты в данном базисе определяются однозначно.

13.37° Даны векторы e_1, e_2, e_3, e_4 и a в стандартном базисе пространства \mathbb{R}^4 . a) Показать, что векторы e_1, e_2, e_3, e_4 образуют базис пространства \mathbb{R}^4 ; b) найти разложение вектора a по этому базису.

$$e_1 = (1, 0, -2, 3); e_2 = (0, 1, 3, 2); e_3 = (1, 0, 0, 1); e_4 = (2, 3, 12, 2);$$
$$a = (9, 12, 5, 8).$$

13.38° Докажите, что в конечномерном векторном пространстве: a) всякая линейно независимая система векторов может быть дополнена до базиса (в частности, существует хотя бы один базис); b) из всякой порождающей системы векторов можно выбрать базис.

13.39° Докажите, что все базисы конечномерного векторного пространства V содержат одно и то же число векторов.

13.40° Обозначим за \mathbf{P}_n пространство многочленов степени не выше n над \mathbb{R} . Показать, что многочлены $p_1(x) = 2x^2 + x - 1$, $p_2(x) = x^2 - 2$ и $p_3(x) = -x^2 + 2x + 3$ образуют базис \mathbf{P}_2 . Найти координаты $f(x) = 3x - 1$ в этом базисе.

13.41° Образуют ли многочлены $p_1(x) = x^3 + x^2 - 1$, $p_2(x) = x^2 - 2x$, $p_3(x) = x^3 + x$, $p_4(x) = x^2 - 3$ базис в пространстве \mathbf{P}_3 ?

13.42° Доказать, что $L(V, V')$ является линейным пространством над тем же полем F , если сложение операторов и умножение на число определять по формулам

$$(A + B)(v) = A(v) + B(v), \quad (kA)(v) = kA(v), \quad v \in V, \quad k \in F.$$

13.43° Пусть $x = (x_1, x_2, \dots, x_n)^T$ — произвольный вектор-столбец n -мерного арифметического пространства \mathbb{R}^n . Исследовать линейность отображения f , если:

- a) $f(x) = (x_2, x_1 - x_2)^T$ ($n = 2$);
- b) $f(x) = (x_2, x_1 x_2)^T$ ($n = 2$);
- c) $f(x) = (x_2, x_1 - 3, x_3)^T$ ($n = 3$);
- d) $f(x) = (2x_3 + x_1, 2x_3 x_1, x_1 - x_2)^T$ ($n = 3$);
- e) $f(x) = (0, 0, \dots, 0)^T$;
- f) $f(x) = (0, x_1 + 3x_2, x_2^2)^T$ ($n = 3$);
- g) $f(x) = (0, 0, \dots, 0, 1)^T$;
- h) $f(x) = (\sin x_1, \cos x_2, x_3)^T$ ($n = 3$);
- i) $f(x) = (x_n, x_{n-1}, \dots, x_1)^T$;
- j) $f(x) = (2x_1, 2|x_2|, 2x_3)^T$ ($n = 3$).

13.44° Пусть f, g, h — линейные отображения пространств вида $\mathbb{R}^n \rightarrow \mathbb{R}^m$ (числа n и m для каждого отображения свои), k — некоторое вещественное число. При каких условиях на размерности пространств, в которых действуют эти отображения, справедливо каждое из следующих равенств?

- a) $f(gh) = (fg)h$;
- b) $f(g + h) = fg + fh$;
- c) $(f + g)h = fh + gh$;
- d) $k(f + g) = kf + kg$.

13.45° Является ли скалярное произведение $\langle x | y \rangle$ билинейной функцией?

13.46° Доказать, что композиция линейных отображений есть линейное отображение. Проверить свойства ассоциативности и дистрибутивности.

13.47° Доказать, что ядро и образ линейного отображения являются линейными пространствами.

13.48° Являются ли линейными следующие отображения $A : V \rightarrow V'$:

- a) $Ax = 0$;
- b) $V = V'$, $Ax = x$ (тождественное отображение; обозначение: id или E);
- c) $V = \mathbb{R}^4$, $V' = \mathbb{R}^3$, $A(x, y, z, t) = (x + y, y + z, z + t)$;
- d) $V = V' = \mathbb{R}^3$, $A(x, y, z) = (x + 1, y + 1, z + 1)$;
- e) $V = V' = F[x]$, $(Ap)(x) = p(ux^2 + v)$, где u, v — фиксированные многочлены из $F[x]$;
- f) $V = V' = F[x]$, $(Ap)(x) = q(x)p(x)$, где q — фиксированный многочлен из $F[x]$.

13.49° Найти ядра и образы линейных отображений предыдущей задачи.

13.50° Доказать линейность $\langle x \mid A \mid y \rangle$ по x и по y , где A — матрица линейного оператора.

13.51° Какие движения плоскости \mathbb{R}^2 являются линейными операторами?

13.52° Пусть A — отображение пространства многочленов степени не выше n с действительными коэффициентами в пространство функций на $M \subseteq \mathbb{R}$, которое переводит многочлен в его ограничение на M .

a) Доказать, что A линейно. b) При каких M ядро $\ker(A) = \{0\}$?

13.53° Какие из перечисленных в задаче 13.32° структур являются алгеброй над полем или кольцом? Умножение на векторах предполагается естественным.

13.54° Будет ли алгеброй (над полем F) a) множество всех линейных операторов $\mathcal{L}(V, V)$ (где V — линейное пространство над полем F), если в качестве умножения операторов выбрать их композицию как функций; b) множество всех линейных функционалов из $\mathcal{L}(V, F)$?

13.55° Проверить следующие свойства матриц 3-го порядка:

- a) $X(YZ) = (XY)Z$;
- b) $X(Y + Z) = XY + XZ$;
- c) $(X + Y)Z = XZ + YZ$;
- d) $k(X + Y) = kX + kY$;
- e) $(k_1 + k_2)X = k_1X + k_2X$;
- f) $(k_1k_2)X = k_1(k_2X)$;
- g) $k(XY) = (kX)Y = X(kY)$.

Найдите нулевую и единичную матрицы. Сделайте вывод о том, какую алгебраическую структуру образуют квадратные матрицы.

13.56° Построить матрицы линейных операторов из задачи 13.43° (тех, которые являются линейными) в стандартном базисе.

13.57° Показать, что каждое из следующих отображений, действующих в линейном пространстве \mathbb{R}^3 , является линейным оператором, найти его матрицу в базисе $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$. Для любого вектора $x = (x_1, x_2, x_3) \in \mathbb{R}^3$:

- a) $Ax = (-3x_1 + 3x_2 - 2x_3, x_1 + 2x_2 - x_3, -x_1 - 3x_2 + 2x_3)$;
- b) $Ax = (-2x_2 - x_3, 3x_1 + 2x_2 + 3x_3, x_1 + 2x_2 + 2x_3)$;
- c) $Ax = (-3x_1 + 3x_2 - 2x_3, x_1 + 2x_2 - x_3, -x_1 - 3x_2 + 2x_3)$;
- d) $Ax = (-3x_1 + 3x_2 - 2x_3, x_1 + 2x_2 - x_3, -x_1 - 3x_2 + 2x_3)$;
- e) $Ax = (-3x_1 + 3x_2 - 2x_3, x_1 + 2x_2 - x_3, -x_1 - 3x_2 + 2x_3)$;
- f) $Ax = (-3x_1 + 3x_2 - 2x_3, x_1 + 2x_2 - x_3, -x_1 - 3x_2 + 2x_3)$.

13.58° Для операторов, действующих в \mathbb{R}^2 найти: a) $A - B$; b) $2A + 3B$; c) AB ; d) BA ; e) A^2 ; f) B^3 , если $Ax = \begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} x$, $Bx = \begin{pmatrix} -1 & 0 \\ 2 & 1 \end{pmatrix} x$, где $x = (x_1, x_2)$.

13.59° Вычислить размерности ядра и образа линейных отображений, заданных матрицами:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 3 & 5 \\ 1 & 1 & 3 \\ 1 & 4 & 6 \\ 1 & 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 0 & 3 & 0 \\ 3 & 1 & 0 & 2 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 4 & 1 \end{pmatrix}.$$

13.60° Какой числовой структуре изоморфно множество матриц вида $aE + bI$, где $a, b \in \mathbb{R}$ и $E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ и $I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

13.61° Найдите матрицу оператора поворота \mathbb{R}^2 на угол α относительно точки $(0, 0)$ и матрицу оператора отражения \mathbb{R}^2 относительно оси, проходящей через точку $(0, 0)$ под углом φ к оси Ox . Базис — стандартный.

13.62° Найти $3A - B - 4C$, если:

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

13.63° Найти $2A - B$, если:

$$A = \begin{pmatrix} 1 & 2 & 7 & -15 \\ 1 & -5 & -6 & 11 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 24 & -1 & -1 \\ -1 & -2 & 7 & 3 \end{pmatrix}.$$

13.64° Найти $AB - BA$, если:

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & -2 \\ 3 & -2 & 4 \\ -3 & 5 & -1 \end{pmatrix}.$$

13.65° Найти $(A - B) \cdot A + 2B$, если:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

13.66° Показать, что матрицы вида $\begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix}$, где $q \in \mathbb{Q} \setminus \{0\}$, образуют поле, изоморфное \mathbb{Q}^* .

13.67° Найти A^3 для следующих матриц:

$$\text{a) } A = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}; \quad \text{b) } A = \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix}; \quad \text{c) } A = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

13.68° Найти A^n для матрицы $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

13.69° Найти общий вид матриц A второго порядка, квадрат которых равен нулевой матрице, т. е. $A^2 = O$.

13.70° Найти все матрицы A второго порядка, квадрат которых равен диагональной матрице $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, где $a \neq b$.

13.71° Найти условие, при котором матрица A второго порядка перестановочна со всеми матрицами второго порядка. При этом матрицы A и B называются **перестановочными**, если их **коммутатор** $[A, B]$ равен нулевой матрице, где коммутатор определяется по формуле $[A, B] = AB - BA$.

13.72° Матрица A называется **инволютивной**, если $A^2 = E$, и **идемпотентной**, если $A^2 = A$. Найти общий вид инволютивной и идемпотентной матрицы второго порядка.

13.73° Каким условиям должны удовлетворять элементы матрицы A второго порядка, чтобы она была перестановочна со всеми диагональными матрицами того же порядка?

13.74° Найти все степени матрицы

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

13.75° Показать на примере матриц второго порядка, что равенство $AB - BA = E$ невозможно.

13.76° Найти общий вид матрицы A третьего порядка, для которой

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot A = O.$$

13.77° Найти все матрицы, перестановочные с данными:

$$\text{a) } \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \text{ b) } \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}; \text{ c) } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

13.78° Показать, что матрицы A и B перестановочны:

$$A = \begin{pmatrix} 3 & 1 & -2 \\ 3 & -2 & 4 \\ -3 & 5 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix}.$$

Дополнительные упражнения

13.79 Для матриц второго и третьего порядка проверьте свойство определителя:

$$\det(AB) = \det(A) \det(B).$$

Выведите отсюда критерий обратимости матрицы.

13.80 Вычислить определители:

$$\text{a) } \begin{vmatrix} 4 & -5 \\ 3 & 8 \end{vmatrix}; \text{ b) } \begin{vmatrix} 5 & -2 \\ -9 & 6 \end{vmatrix}; \text{ c) } \begin{vmatrix} -10 & 2 \\ -6 & 7 \end{vmatrix}; \text{ d) } \begin{vmatrix} 1.5 & -0.2 \\ 0.3 & -4 \end{vmatrix}; \text{ e) } \begin{vmatrix} 7 & -4 \\ 5 & 2 \end{vmatrix};$$

$$\text{f) } \begin{vmatrix} 3 & -7 & 1 \\ -2 & 4 & -3 \\ 9 & -1 & -5 \end{vmatrix}; \text{ g) } \begin{vmatrix} 3 & -5 & 2 \\ 4 & 1 & 5 \\ 2 & 7 & -3 \end{vmatrix}; \text{ h) } \begin{vmatrix} 4 & -3 & 5 & 6 \\ -7 & 2 & 1 & 8 \\ 0 & 5 & 4 & 0 \\ 9 & 7 & 0 & -2 \end{vmatrix}.$$

13.81 Вычислить определитель Вандермонда:

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix}.$$

Выяснить, при каких значениях x, y, z определитель равен нулю.

13.82 Вычислить определители:

$$\text{a) } \begin{vmatrix} \sin^2 \alpha & 1 & \cos^2 \alpha \\ \sin^2 \beta & 1 & \cos^2 \beta \\ \sin^2 \gamma & 1 & \cos^2 \gamma \end{vmatrix}; \text{ b) } \begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^3 & y^3 & z^3 \end{vmatrix}; \text{ c) } \begin{vmatrix} \cos 2\alpha & \cos^2 \alpha & \sin^2 \alpha \\ \cos 2\beta & \cos^2 \beta & \sin^2 \beta \\ \cos 2\gamma & \cos^2 \gamma & \sin^2 \gamma \end{vmatrix}.$$

13.83 Вычислить определители:

$$\text{a) } \begin{vmatrix} 2 & 1 & 0 & 2 \\ 3 & 2 & 1 & 0 \\ -1 & 0 & 1 & 3 \\ -1 & 2 & 1 & 3 \end{vmatrix}; \text{ b) } \begin{vmatrix} a & b & b & a \\ 1 & 2 & 1 & 0 \\ 3 & 2 & 0 & 1 \\ 2 & 0 & 1 & 2 \end{vmatrix}; \text{ c) } \begin{vmatrix} 0 & a & b & c \\ 1 & x & 0 & 0 \\ 1 & 0 & y & 0 \\ 1 & 0 & 0 & z \end{vmatrix}.$$

13.84 Найти матрицы, обратные для следующих:

$$\text{a) } \begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix}; \text{ b) } \begin{pmatrix} \sin x & \cos x \\ -\cos x & \sin x \end{pmatrix}; \text{ c) } \begin{pmatrix} 1 & 3 & 4 \\ 2 & 0 & 3 \\ -2 & 1 & -3 \end{pmatrix}; \text{ d) } \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}.$$

13.85 Доказать, что определитель ортогональной матрицы (над \mathbb{R}) равен ± 1 .

13.86 Проверить, что $\mathbb{O}(n)$ и $\text{SO}(n)$ — группы по умножению матриц.

13.87 Доказать, что линейный оператор над \mathbb{R}^2 является оператором поворота плоскости с центром в точке $(0, 0)$ тогда и только тогда, когда его матрица в стандартном базисе принадлежит группе $\text{SO}(2)$.

13.88 Доказать, что линейный оператор над \mathbb{R}^2 является оператором отражения плоскости относительно оси, проходящей через точку $(0, 0)$, тогда и только тогда, когда его матрица в стандартном базисе принадлежит группе $\text{SL}(2)$ и имеет определитель, равный -1 .

13.89 Диагональной называется квадратная матрица, у которой все элементы, кроме, быть может, a_{ii} , равны 0. Доказать, что определитель диагональной матрицы равен произведению ее диагональных элементов. Как выглядит обратная к диагональной матрица?

13.90 Треугольной называется квадратная матрица, у которой все элементы, расположенные выше (или ниже) главной диагонали, равны 0. Доказать, что определитель треугольной матрицы равен произведению ее диагональных элементов.

13.91 Как изменится определитель матрицы, если все элементы матрицы заменить комплексно-сопряженными числами?

13.92 Матрица A коммутирует с B . Доказать, что тогда A^{-1} коммутирует с B^{-1} (предполагается, что матрицы обратимы).

13.93 Обозначим через $D(2, \mathbb{Q})$ множество всех диагональных матриц порядка 2 над полем \mathbb{Q} . Доказать, что $D(2, \mathbb{Q}) \cap SL(2, \mathbb{Q}) \cong \mathbb{Q}^*$.

13.94 Матрицы второго порядка над полем \mathbb{C}

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

называются **матрицами Паули**.

Найти: **a)** определители матриц Паули; **b)** обратные к ним матрицы; **c)** квадраты матриц Паули; **d)** все их попарные произведения; **e)** коммутаторы всех пар матриц Паули.

13.95 Показать, что вещественные линейные комбинации матриц Паули и единичной матрицы, т. е. выражения вида

$$\alpha E_2 + \beta \sigma_x + \gamma \sigma_y + \delta \sigma_z, \quad \alpha, \beta, \gamma, \delta \in \mathbb{R}, \quad (13.12)$$

образуют четырехмерное линейное пространство над полем \mathbb{R} .

13.96* Докажите, что если $\{e_1, \dots, e_n\}$ — ОНБ, то скалярное произведение векторов выражается по формуле

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n,$$

где x_i, y_i — координаты векторов x и y в данном базисе.

13.97* Докажите **неравенство Коши — Буняковского** для вещественного пространства:

$$\langle x, y \rangle^2 \leq \|x\|^2 \|y\|^2.$$

Указание: рассмотрите скалярное произведение $\langle \lambda x + y, \lambda x + y \rangle$.

13.98* Пусть $\hat{A}, \hat{B} : V \rightarrow V$ — вещественные эрмитовы линейные операторы с матрицами A и B . Докажите неравенства:

$$\langle x | AB | x \rangle^2 \leq \|Ax\|^2 \|Bx\|^2, \quad \langle x | AB | x \rangle \langle x | BA | x \rangle \leq \|Ax\|^2 \|Bx\|^2. \quad (13.13)$$

Указание: воспользуйтесь неравенством Коши — Буняковского.

13.99* Из неравенств (13.13) докажите, что для вещественных эрмитовых операторов выполняется **соотношение Робертсона – Шрёдингера** для коммутатора:

$$\langle x|[A, B]|x\rangle^2 \leq 4\|Ax\|^2\|Bx\|^2, \quad (13.14)$$

где $[A, B] = AB - BA$ – коммутатор матриц A и B .

13.100* Из неравенства для коммутатора (13.14) выведите **соотношение неопределенности Гейзенберга** для вещественных эрмитовых операторов при условии, что $\|\psi\| = 1$:

$$\Delta_\psi A \cdot \Delta_\psi B \geq \frac{1}{2} |\langle [A, B] \rangle_\psi|,$$

где

$$\langle X \rangle_\psi = \langle \psi | X | \psi \rangle, \quad \Delta_\psi X = \sqrt{\langle X^2 \rangle_\psi - \langle X \rangle_\psi^2}.$$

Указание: докажите, что $[A, B] = [\alpha, \beta]$, где $\alpha = A - \langle A \rangle_\psi$ и $\beta = B - \langle B \rangle_\psi$, а также, что $\|\alpha\psi\|^2 = (\Delta_\psi A)^2$ и $\|\beta\psi\|^2 = (\Delta_\psi B)^2$.

Алгебраические числа

Аннотация

В этой главе мы заглянем за пределы поля рациональных чисел при помощи многочленов с рациональными коэффициентами, построим поле алгебраических чисел, разберем некоторые теоремы о многочленах над кольцами и полями.

14.1. Плотные линейно упорядоченные множества

Вернемся к анализу поля рациональных чисел \mathbb{Q} . Данное поле интересно тем, что какие бы два различных числа мы ни взяли, между ними всегда найдется третье рациональное число. Действительно, пусть есть две дроби $r = n/m$ и $q = t/s$, тогда их среднее арифметическое $(r + q)/2$ является рациональным числом и лежит строго между ними. Множества с таким свойством принято называть **плотными**.

Данное определение стоит уточнить, поскольку мы еще не дали строгого определения понятию «лежать между», которое, с одной стороны, геометрически очевидно, а с другой — требует некоторой формализации, если мы хотим рассуждать не только о точках на прямой. Для этого вспомним, что *отношением на множестве X* называется всякое подмножество $R \subseteq X \times X$. Ранее мы уже рассмотрели такие два важных типа отношений как *отношение эквивалентности* и *функция* (см. раздел 11.1). Здесь мы определим отношение порядка.

Скажем, что отношение $R \subseteq X \times X$ является

- **антирефлексивным**, если $\forall a \in X \neg(aRa)$;
- **транзитивным**, если $\forall a, b, c \in X (aRb) \wedge (bRc) \rightarrow (aRc)$;
- **связным**, если $\forall a, b \in X (aRb) \vee (bRa) \vee (a = b)$.

Пример антирефлексивного отношения: « x является родителем y ». Ясно, что не бывает ситуации, когда x является родителем x , так что данное отношение — антирефлексивное.

Пример транзитивного отношения: x содержит y . Действительно, если x содержит y , а y содержит z , то x содержит z (можно держать в голове образ матрешки или вложенных фигур).

Если отношение R является одновременно антирефлексивным и транзитивным, то оно называется отношением (строгого) **предпорядка** или (строгого) **частичного порядка**. Например, отношение предок–потомок является предпорядком.

Заметим, что отношение предпорядка асимметрично, т. е., в противоположность симметричному отношению, для него выполняется условие $\forall a, b \in X : (aRb) \rightarrow \neg(bRa)$. Если бы это было не так, то мы бы пришли к противоречию с тем, что отношение порядка антирефлексивно: если существуют такие $a, b \in X$, что aRb и bRa , то в силу транзитивности получаем, что aRa .

Наконец, связное отношение предпорядка называется (строгим) **линейным порядком**. Чаще всего линейный порядок обозначается символом $<$. Например, естественное упорядочение целых чисел является линейным порядком. А вот отношение «предок — потомок» — нет. Потому что у одного и того же предка могут существовать потомки, которые не состоят в таком отношении (например, два брата или дядя с племянником).

Часто линейным порядком называется нестрогий линейный порядок, который допускает равенство: если $<$ — линейный порядок в том смысле, как мы определили выше, то \leq — нестрогий линейный порядок. В этом курсе под термином «линейный порядок» мы будем понимать именно строгий линейный порядок, т. е. антирефлексивное транзитивное связное отношение.

Теперь, если на некотором множестве X задан линейный порядок $<$, то пара $\langle X, < \rangle$ называется **линейно упорядоченным множеством** (сокращенно л.у.м.). Например, таковы $\langle \mathbb{Z}, < \rangle$ и $\langle \mathbb{Q}, < \rangle$.

Для л.у.м. мы можем определить понятие *лежать между* следующим образом: точка a лежит между точками b и c , если выполнены неравенства $b < a$ и $a < c$, либо $c < a$ и $a < b$. Отсюда очевидным образом мы выходим и на понятие *интервала*: $(a; b) = \{x \in X \mid (a < x) \wedge (x < b)\}$.

Пусть $A \subseteq B$, где B — линейно упорядоченное множество. Говорят, что множество A **плотно в B** , если между любыми двумя точками множества B найдется точка множества A . Иначе говоря, A плотно в B , если для любых $a, b \in B$ из того, что $a < b$, следует, что существует $x \in A$ такой, что $x \in (a; b)$. Например, множество \mathbb{Q} плотно в себе (или просто плотно) в силу приведенного ранее рассуждения.¹

¹ На самом деле, когда мы находим среднее значение $(r + q)/2$ двух рациональных чисел $r, q \in \mathbb{Q}$, и на этом основании делаем вывод о том, что оно лежит между числами r и q , мы прячем под ковер еще одно требование к порядку на числах, а именно

Можно взять, например, множество $\mathbb{B} \subseteq \mathbb{Q}$ всех двоичных дробей, т.е. дробей вида $k/2^n$, где $k \in \mathbb{Z}$, $n \in \mathbb{N}$. Такое множество, во-первых, является плотным в себе, поскольку среднее арифметическое любых двух его представителей

$$\left(\frac{k}{2^n} + \frac{l}{2^m}\right) / 2 = \frac{k \cdot 2^m + l \cdot 2^n}{2^{m+n+1}}$$

является двоичной дробью.

А во-вторых, множество \mathbb{B} плотно в \mathbb{Q} , поскольку каковы бы ни были два рациональных числа $r \neq q$, между ними найдется двоичная дробь. Докажем это.

Для начала заметим, что мы можем рассматривать только случай неотрицательных r, q . Действительно, если $r, q \leq 0$, то мы рассмотрим пару $-r, -q$, найдем между ними число $k/2^n$, следовательно, число $-k/2^n$ будет лежать между исходными r и q . Если же r и q имеют разные знаки, то число 0 лежит между ними.

Далее, для любого натурального x справедливо неравенство $x < 2^x$, которое легко доказать по индукции.

Наконец, пусть нам даны две дроби $0 \leq a/b < c/d$ ($b, d > 0$). Тогда рассмотрим число $bc \cdot 2^{bd} - 1$, которое мы разделим на bd с остатком (очевидно, что $bd > 0$):

$$bc \cdot 2^{bd} - 1 = k(bd) + r, \quad 0 \leq r < bd.$$

Теперь заметим, во-первых, что

$$\frac{k}{2^{bd}} = \frac{k(bd)}{bd \cdot 2^{bd}} < \frac{c}{d}.$$

Во-вторых, так как $ad \leq bc - 1$ и $1 + r \leq bd < 2^{bd}$, то справедлива такая оценка:

$$ad \cdot 2^{bd} \leq bc \cdot 2^{bd} - 2^{bd} < bc \cdot 2^{bd} - 1 - r = k(bd),$$

откуда

$$\frac{a}{b} < \frac{k}{2^{bd}}.$$

Таким образом, между дробями a/b и c/d находится двоичная дробь $k/2^{bd}$. То есть множество двоичных дробей \mathbb{B} плотно в \mathbb{Q} .

На самом деле это легко понять, если представить себе, как нужно наносить на числовую ось двоичные дроби. Сначала мы берем все

согласованность порядка с операциями сложения и умножения. Это значит, что если $x < y$, то для любого z : $x + z < y + z$, и для любого $z > 0$: $xz < yz$. Эти свойства для геометрической прямой мы разбирали в разделе 1.1.

целые числа, затем ровно между ними ставим все полуцелые (с шагом $1/2$), затем в оставшихся полуцелых интервалах отмечаем середины (получаем числа с шагом $1/4$), затем снова делим эти интервалы пополам (получаем шаг $1/8$) и т. д. Ясно, что чем больше шагов мы пройдем, тем мельче будет сетка двоичных дробей.

Множества, подобные \mathbb{B} , т. е. такие, что их точки можно обнаружить в любом интервале числовой прямой, называются **всюду плотными** множествами. Ясно, что и множество \mathbb{Q} всюду плотно, поскольку оно содержит в себе \mathbb{B} .

Более того, всякое множество, плотное в \mathbb{Q} , всюду плотно на прямой. И вообще, отношение « X плотно в Y » транзитивно на подмножествах числовой прямой \mathbb{R} .

14.2. Зазоры между рациональными числами

Ранее мы уже приводили пример числа, которое определяется уравнением в целых числах, но не является рациональным. Это число $\sqrt{2}$, которое разрешает уравнение $x^2 - 2 = 0$.

Такое число как бы вставляет клин между рациональными числами, рассекая их на две части. Действительно, если мы посмотрим на два множества

$$X = \{r \in \mathbb{Q} \mid r^2 < 2 \text{ или } r < 0\}, \quad Y = \{r \in \mathbb{Q} \mid r^2 \geq 2 \text{ и } r \geq 0\},$$

то можно заметить, что, во-первых, их объединение $X \cup Y$ равно \mathbb{Q} . Это следует из того, что определяющие их условия

- (1) $r^2 < 2$ или $r < 0$;
- (2) $r^2 \geq 2$ и $r \geq 0$;

являются взаимоисключающими и дополняющими.

Во-вторых, эти множества не пересекаются: $X \cap Y = \emptyset$. Это также следует из того, что (1) является отрицанием (2).

Наконец, в-третьих, $X < Y$ в смысле сравнения множеств. Действительно, пусть $r \in X, q \in Y$, тогда или $r < 0$, или $r \geq 0, r^2 < 2$. В первом случае $r < q$, т. к. $q \geq 0$. Во втором случае мы сравниваем два неотрицательных числа, связанных условиями $r^2 < 2$ и $q^2 \geq 2$. Но для неотрицательных чисел неравенство $r < q$ равносильно неравенству $r^2 < q^2$ (поскольку $r^2 \leq rq < q^2$), поэтому из неравенств $r^2 < 2 \leq q^2$ следует $r < q$.

Такие пары множеств называются **дедекиндовыми сечениями**, и к ним мы еще вернемся чуть позже.

Особенностью данного разбиения $\langle X, Y \rangle$ является то, что новое число $\sqrt{2}$ должно быть размещено строго между множествами X и Y , т. к.

оно не меньше никакого числа из X и не больше никакого числа из Y , и в то же время не попадает ни в X , ни в Y , т.к. не является рациональным числом. Получается, что $X < \sqrt{2} < Y$. В таком случае говорят, что сечение $\langle X, Y \rangle$ определяет число $\sqrt{2}$.

Стало быть, между точками множества рациональных чисел есть дырки. Оказывается, что их довольно много.

Теорема 14.1. *Уравнение $x^k = r$, где r — положительное рациональное число, k — ненулевое натуральное число, имеет рациональный корень x тогда и только тогда, когда все степени простых чисел, входящих во взаимно простые числитель и знаменатель r , кратны k .*

Доказательство. (\Leftarrow) Очевидно, поскольку при $r = a^k/b^k$, $a, b \in \mathbb{N} \setminus \{0\}$, существует решение данного уравнения: $x = a/b$.

(\Rightarrow) Предположим, что $x = a/b$, где $a, b > 0$ и $a \perp b$. Пусть также $r = c/d$, где $c, d > 0$ и $c \perp d$. Пользуясь ОТА, напишем разложения:

$$c = p_1^{\alpha_1} \dots p_n^{\alpha_n}, \quad d = q_1^{\beta_1} \dots q_m^{\beta_m},$$

где $p_1, \dots, p_n, q_1, \dots, q_m$ — попарно различные простые числа, $\alpha_i, \beta_j > 0$. Тогда из уравнения $x^k = r$ получаем, что

$$a^k q_1^{\beta_1} \dots q_m^{\beta_m} = b^k p_1^{\alpha_1} \dots p_n^{\alpha_n}.$$

Поскольку q_1 не совпадает ни с одним из p_i , число b^k имеет в своем разложении по степеням простых множитель $q_1^{\beta_1}$, откуда следует, что β_1 делится на k . Аналогично для всех остальных q_i и всех p_j получаем, что их степени β_i и α_j также делятся на k . □

Таким образом, мы имеем рецепт изготовления иррациональных чисел: взяв произвольное натуральное $k > 1$ и произвольное рациональное число $r > 0$, в котором числитель и знаменатель взаимно просты, причем в разложении числителя или знаменателя по степеням простых чисел есть хотя бы одна степень простого, не кратная k (например, от 1 до $k - 1$), мы получаем, что корень уравнения $x^k = r$ не является рациональным числом.

Например, пусть $r = p_1^{\alpha_1} \dots p_n^{\alpha_n} / q_1^{\beta_1} \dots q_m^{\beta_m}$ (где $p_i \neq q_j$ для всех i, j), тогда если $k = 1 + \max\{p_i, q_j\}$, то корнем $x^k = r$ не может быть рациональное число.

Имея такую пару чисел k и r , мы снова можем построить сечение $\langle X, Y \rangle$, где

$$X = \{q \in \mathbb{Q} \mid q < 0 \text{ или } q^k < r\}, \quad Y = \{q \in \mathbb{Q} \mid q \geq 0 \text{ и } q^k \geq r\}.$$

Такое сечение будет определять единственно возможную точку на числовой прямой, расположенную строго между множествами X и Y . Эту точку мы обозначим за $\sqrt[k]{r}$.

Как и в случае $\sqrt{2}$, мы можем сколь угодно близко подбираться к числу $\sqrt[k]{r}$ при помощи рациональных чисел. Это прямо следует из наших рассуждений о всюду плотности множества \mathbb{W} .

Итак, мы видим, что «дырки» между рациональными числами — явление нередкое. И точно так же, как мы расширяли \mathbb{Q} до поля $\mathbb{Q}[\sqrt{2}]$, мы можем строить любые расширения $\mathbb{Q}[\sqrt[k]{r}]$, получая новые поля всякий раз, когда рациональное число r не является k -ой степенью какого-то рационального числа.

14.3. О построениях циркулем и линейкой

Одной из классических задач геометрии является изучение вопроса о том, какие геометрические построения можно произвести, имея только циркуль и линейку. Первый позволяет строить окружности заданного радиуса, вторая — соединять любые две заданные точки и неограниченно продлевать отрезок за его границы. Все в точности с первыми тремя постулатами Евклида. При этом предполагается, что у нас есть некий мерный отрезок, задающий масштаб (единицу длины), а следовательно, вопрос о построениях циркулем и линейкой сводится к умению строить отрезки различной длины или, проще говоря, строить числа.

Числа при этом получаются как длины отрезков, соединяющих получаемые при построениях точки пересечения линий — прямых и окружностей. Ясно, что такие пересечения могут давать только числа, являющиеся решениями линейных и квадратных уравнений. Иначе говоря, имея единицу, мы можем строить все рациональные числа (по теореме Фалеса), затем все рациональные комбинации различных корней из рациональных чисел, затем корней из этих корней и т. д. Речь идет конечно же о квадратных корнях.

В качестве упражнения предлагаем читателю самостоятельно построить циркулем и линейкой отрезки длины $\sqrt{2}$ и $\sqrt{3}$. А чтобы задача не казалась сложной, скажем, что юный Гаусс в начале XIX века построил правильный 17-угольник впервые в истории математики. Это построение любопытно посмотреть в динамике, например, [тут](https://en.wikipedia.org/wiki/Constructible_polygon) (https://en.wikipedia.org/wiki/Constructible_polygon). Там же смотрите построение 15-, 257- и 65537-угольников.

Возникает вопрос: можно ли построить с помощью циркуля и линейки ребро куба, объем которого равен 2, т. е. число $\sqrt[3]{2}$ (это античная

задача об удвоении куба, известная наравне с задачей о квадратуре круга)?

Чтобы ответить на него, вспомним о линейных пространствах. Дело в том, что всякое расширение поля \mathbb{Q} (и не только этого поля) с помощью присоединения каких-либо новых чисел, являющихся корнями алгебраических уравнений с рациональными коэффициентами, можно рассматривать как конечномерное линейное пространство над \mathbb{Q} .

Действительно, пусть нам дано уравнение вида

$$x^n + a_1x^{n-1} + \dots + a_n = 0, \quad a_1, \dots, a_n \in \mathbb{Q}, \quad n > 1, \quad (14.1)$$

и число γ — такой иррациональный корень данного уравнения, который не является корнем никакого многочлена степени меньше n с рациональными коэффициентами. Теперь мы хотим получить такую систему чисел K , которая бы содержала \mathbb{Q} и число γ , при этом чтобы в ней выполнялись аксиомы поля, и она была бы в некотором смысле минимальной (т.е. чтобы никакое собственное подмножество этой системы не обладало бы перечисленными свойствами).

Для начала мы построим кольцо чисел K . Поскольку $\mathbb{Q} \cup \{\gamma\} \subseteq K$, в это кольцо должны входить и все комбинации вида $r + q\gamma$, где коэффициенты $r, q \in \mathbb{Q}$. Но как только мы начинаем умножать такие комбинации друг на друга, немедленно выясняем, что нам требуется присоединить и γ^2 , т.е. включить в систему K комбинации вида $r_0 + r_1\gamma + r_2\gamma^2$. После чего снова возникает потребность в увеличении степени γ , принадлежащей K , и так далее.

Однако, как только мы достигнем степени γ^n , мы сможем ее выразить через более низкие степени, пользуясь уравнением (14.1):

$$\gamma^n = -a_1\gamma^{n-1} - \dots - a_n.$$

То же самое можно применить и к более высоким степеням γ . Таким образом, K будет исчерпываться комбинациями вида $r_0 + r_1\gamma + \dots + r_n\gamma^n$, где $r_i \in \mathbb{Q}$. Операции сложения, вычитания и умножения чисел такого вида приводят к числам такого же вида, т.е. K является кольцом.

Введем следующие обозначения:

$$\mathbf{e}_0 = 1, \quad \mathbf{e}_1 = \gamma, \quad \mathbf{e}_2 = \gamma^2, \dots, \mathbf{e}_{n-1} = \gamma^{n-1}.$$

Тогда любое число из K можно представить в виде линейной комбинации $\mathbf{e}_0, \dots, \mathbf{e}_{n-1}$. Иначе говоря, K является линейной оболочкой $\{\mathbf{e}_k\}$ и линейным пространством над полем \mathbb{Q} .

Кроме того, нетрудно показать, что набор чисел $\mathbf{e}_0, \dots, \mathbf{e}_{n-1}$ является линейно независимым, поскольку в противном случае нашелся бы

многочлен степени меньше n , корнем которого является γ , что противоречит требованию минимальности степени многочлена (14.1).

Таким образом, набор чисел e_0, \dots, e_{n-1} является базисом пространства K , которое также обозначается через $\mathbb{Q}[\gamma]$.

Например, $\mathbb{Q}[\sqrt{2}]$ является пространством размерности 2 над полем \mathbb{Q} , базисными векторами в нем являются числа 1 и $\sqrt{2}$ (они линейно независимы, т.к. $\sqrt{2}$ — иррациональное число), и $\mathbb{Q}[\sqrt{2}] = \{x + y\sqrt{2} \mid x, y \in \mathbb{Q}\}$.

Можно доказать также, что $\mathbb{Q}[\gamma]$ является полем, т.е. отношение чисел вида $r_0 + r_1\gamma + \dots + r_{n-1}\gamma^{n-1}$ является числом такого же вида. Это легко видеть на примере $\mathbb{Q}[\sqrt{2}]$:

$$\frac{x + y\sqrt{2}}{a + b\sqrt{2}} = \frac{(x + y\sqrt{2})(a - b\sqrt{2})}{(a + b\sqrt{2})(a - b\sqrt{2})} = \frac{(ax - 2by) + (ya - xb)\sqrt{2}}{a^2 - 2b^2}.$$

Для корней более высокой степени доказательство этого факта является более сложным и в данном курсе не рассматривается.

Повторяя описанный выше процесс построения «нового» поля чисел с помощью присоединения некоторого числа к «старому» полю, мы можем строить все более широкие поля, которые также будут линейными пространствами над \mathbb{Q} . Так, имея поле $K = \mathbb{Q}[\sqrt{2}]$, мы можем построить его расширение $K[\sqrt{3}] = \mathbb{Q}[\sqrt{2}][\sqrt{3}]$, в котором все числа будут иметь вид $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$, где $a, b, c, d \in \mathbb{Q}$. И так далее.

Присоединяя к \mathbb{Q} квадратные корни натуральных чисел, мы либо не увеличиваем поле, либо удваиваем его размерность. Как показано в примере выше, $\mathbb{Q}[\sqrt{2}][\sqrt{3}]$ имеет размерность 4 над \mathbb{Q} . Доказательство независимости базисных векторов $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$ мы оставим за рамками курса. Главное для нас состоит в том, что присоединение квадратных корней к полю \mathbb{Q} дает только такие размерности расширений поля \mathbb{Q} , которые являются степенями двойки.

Следующим замечательным фактом из высшей алгебры является такое утверждение: если какое-то число α содержится в каком-то поле $\mathbb{Q}[\beta_1] \dots [\beta_n]$, то это поле является конечным расширением над $\mathbb{Q}[\alpha]$, а значит, размерность поля $\mathbb{Q}[\beta_1] \dots [\beta_n]$ делится на размерность поля $\mathbb{Q}[\alpha]$. Напоминает делимость порядка группы на порядок подгруппы, не правда ли? Доказательство этого факта также выходит за рамки нашего курса.

Теперь о $\sqrt[3]{2}$. Можно доказать, что числа $1, \sqrt[3]{2}, \sqrt[3]{4}$ линейно независимы над полем \mathbb{Q} , поэтому поле $\mathbb{Q}[\sqrt[3]{2}] = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} \mid a, b, c \in \mathbb{Q}\}$ имеет размерность 3 над \mathbb{Q} . Проблема в том, что никакое число 2^n не делится на 3 в силу основной теоремы арифметики. Это значит, что как бы мы ни расширяли \mathbb{Q} с помощью квадратных корней, мы никогда не

сможем построить число $\sqrt[3]{2}$. Следовательно, удвоить куб циркулем и линейкой невозможно!

14.4. Многочлены и алгебраические числа

Корни многочленов над \mathbb{Q} называются **алгебраическими числами**. Множество всех алгебраических чисел обозначается через \mathbb{A} . Заметим, что алгебраические числа, вообще говоря, лежат в комплексной плоскости, т. е. могут иметь мнимую часть.

Все корни многочленов из $\mathbb{Q}[x]$ и все корни многочленов из $\mathbb{Z}[x]$ — это одно и то же множество \mathbb{A} . Действительно, $\mathbb{Z}[x] \subseteq \mathbb{Q}[x]$, поскольку $\mathbb{Z} \subseteq \mathbb{Q}$, так что корни многочленов с целыми коэффициентами принадлежат \mathbb{A} . С другой стороны, если какое-то число x зануляет многочлен с рациональными коэффициентами, то оно же зануляет и многочлен с целыми коэффициентами. Этот многочлен получается из исходного домножением на все знаменатели всех его коэффициентов, так что вместо дробей мы получаем целые числа, а равенство нулю при этом сохраняется.

Поэтому часто при анализе корней многочленов используется один из следующих подходов: либо рассматриваются многочлены с произвольными целыми коэффициентами, либо рассматриваются многочлены с рациональными коэффициентами, у которых старший коэффициент равен 1:

$$k_n x^n + k_{n-1} x^{n-1} + \dots + k_1 x + k_0, \quad k_s \in \mathbb{Z}, \quad \text{либо} \quad x^n + q_{n-1} x^{n-1} + \dots + q_1 x + q_0, \quad q_s \in \mathbb{Q}.$$

Пусть $f \in \mathbb{Z}[x]$. **Содержанием многочлена** f называется наибольший общий делитель всех его коэффициентов:

$$\text{cont}(k_0 + k_1 x + \dots + k_n x^n) = \text{НОД}(k_0, k_1, \dots, k_n)$$

Лемма 14.1 (Гаусса). Пусть $f, g \in \mathbb{Z}[x]$, тогда

$$\text{cont}(fg) = \text{cont}(f) \text{cont}(g).$$

Доказательство. Ясно, что достаточно рассмотреть случай $\text{cont}(f) = \text{cont}(g) = 1$. Остальные случаи сводятся к нему делением коэффициентов многочленов на их содержание. При этом, очевидно, они не перестают быть многочленами над \mathbb{Z} .

Итак, предполагая $\text{cont}(f) = \text{cont}(g) = 1$, покажем, что $\text{cont}(fg) = 1$. Пусть, кроме того,

$$f(x) = f_0 + f_1 x + \dots + f_n x^n, \quad g(x) = g_0 + g_1 x + \dots + g_m x^m.$$

Предположим, что $\text{cont}(fg) = d > 1$. Пусть p — простое число, делящее d . Так как $\text{cont}(f) = 1$, то существует хотя бы один коэффициент многочлена f , который не делится на p . Пусть f_k — коэффициент с минимальным номером, не делящийся на p . Аналогично обозначим за g_s коэффициент многочлена g с минимальным номером, не делящийся на p .

Найдем коэффициент многочлена fg при степени x^{k+s} . Он равен

$$f_0g_{k+s} + f_1g_{k+s-1} + \cdots + f_kg_s + f_{k+1}g_{s-1} + \cdots + f_{k+s}g_0 \equiv f_kg_s \pmod{p},$$

поскольку $f_0, \dots, f_{k-1} \equiv 0 \pmod{p}$ и $g_{s-1}, \dots, g_0 \equiv 0 \pmod{p}$. Для упрощения записи мы предполагаем, что $f_j = 0$, если $j > n$, и $g_l = 0$, если $l > m$.

Но $f_kg_s \not\equiv 0 \pmod{p}$ в силу их выбора, а значит, и коэффициент многочлена fg при степени x^{k+s} не делится на p , откуда следует, что p не может быть общим делителем коэффициентов fg , а значит, не делит и наибольший общий делитель $\text{cont}(fg) = d$. Противоречие. \square

Следствие 14.1. *Многочлен с целыми коэффициентами неприводим над \mathbb{Z} тогда и только тогда, когда он неприводим над \mathbb{Q} .*

Доказательство. Ясно, что если многочлен неприводим над \mathbb{Q} , то он неприводим и над \mathbb{Z} , поэтому докажем обратное. А точнее, покажем, что если многочлен из кольца $\mathbb{Z}[x]$ раскладывается в произведение над \mathbb{Q} , то его можно разложить и над \mathbb{Z} .

Пусть $f \in \mathbb{Z}[x]$ и $f = gh$, где $g, h \in \mathbb{Q}[x]$. Можно считать также, что $\text{cont}(f) = 1$ (если это не так, то разделим f и g на $\text{cont}(f)$ и получим то, что требуется, f при этом не выпадет из $\mathbb{Z}[x]$, а g — из $\mathbb{Q}[x]$).

Найдем такое натуральное число $n > 0$, что $ng \in \mathbb{Z}[x]$ (например, произведение всех знаменателей коэффициентов g). И пусть $m = \text{cont}(ng)$. Тогда рациональное число $r = n/m$ таково, что $rg \in \mathbb{Z}[x]$ и $\text{cont}(rg) = 1$. Аналогично найдем рациональное s такое, что $sh \in \mathbb{Z}[x]$ и $\text{cont}(sh) = 1$.

По лемме Гаусса получаем

$$1 = \text{cont}(rg) \text{cont}(sh) = \text{cont}(rsgh) = \text{cont}(rsf) = rs \text{cont}(f) = rs.$$

Отсюда получаем следующее разложение

$$f = gh = rsgh = (rg)(sh),$$

где $rg, sh \in \mathbb{Z}[x]$, т. е. f разложим над \mathbb{Z} . \square

Если α — алгебраическое число, то минимальная степень многочлена, корнем которого является α , называется **степенью алгебраического числа** α .

Например, мы ранее показали, что $\sqrt{2}$ не является рациональным числом, значит, никакой многочлен первой степени, т.е. $x + q$, не обращается в ноль при $x = \sqrt{2}$, значит, $\sqrt{2}$ не является алгебраическим числом первой степени.

На самом деле все рациональные числа, и только они, являются алгебраическими числами первой степени. То есть $\mathbb{Q} \subset \mathbb{A}$ (вложение собственное!)

С другой стороны, $x^2 - 2$ зануляется числом $\sqrt{2}$, так что это число является алгебраическим числом степени 2. То же самое можно сказать про квадратный корень из любого рационального числа, не являющегося квадратом какого-либо рационального числа (теорема 14.1).

Далее мы установим, что число $\sqrt[3]{2}$ является алгебраическим числом степени 3. Для этого нам нужно показать, что никакой многочлен степени 2 не может его занулить.

Лемма 14.2. Пусть $f(x)$ — многочлен минимальной степени, зануляющий число $\alpha \in \mathbb{A}$, и пусть $g(\alpha) = 0$ для некоторого многочлена $g(x)$. Тогда существует многочлен $h(x)$ такой, что

$$g = fh.$$

Доказательство. Пусть степень f равна m , а степень g равна n . Ясно, что $n \geq m$ в силу минимальности f . Разделим g на f с остатком: $g = fh + r$. Здесь степень h равна $n - m$, а степень r строго меньше m .

Но так как $g(\alpha) = 0$ и $f(\alpha) = 0$, то мы получаем, что и $r(\alpha) = 0$. Таким образом, r является многочленом, зануляющим α , и притом его степень меньше m в противоречии с определением числа m . Следовательно, r есть тождественный ноль, следовательно, g делится на f без остатка. \square

Перейдем к $\sqrt[3]{2}$, точнее, к определяющему его многочлену.

Лемма 14.3. Многочлен $x^3 - 2$ неразложим на множители с целыми коэффициентами степени ≥ 1 .

Доказательство. Предположим, что это не так, тогда $x^3 - 2$ делится на линейный многочлен вида $(ax + b)$ (если он делится на многочлен второй степени, то делится и на многочлен первой степени):

$$x^3 - 2 = (ax + b)(kx^2 + tx + n),$$

где $a, b, k, t, n \in \mathbb{Z}$ и $a, k \neq 0$. Сравним коэффициенты при одинаковых степенях:

$$\begin{aligned}
1 &= ak \\
0 &= am + bk \\
0 &= an + bm \\
-2 &= bn
\end{aligned}$$

Из первого равенства следует, что либо $a = k = 1$, либо $a = k = -1$. Учитывая это, из второго равенства получаем, что $m = -b$, откуда с помощью третьего равенства получаем, что $an = b^2$. Наконец, четвертое равенство предлагает варианты:

$$b = 2, n = -1 \text{ или } b = -2, n = 1 \text{ или } b = 1, n = -2 \text{ или } b = -1, n = 2.$$

В первом и втором случае получаем, что $an = 4$, но это невозможно, поскольку $a, n \in \{1, -1\}$.

В третьем и четвертом случае $an = 1$, но и это невозможно при $a \in \{1, -1\}, n \in \{2, -2\}$. \square

В силу следствия из леммы Гаусса и леммы о неразложимости $x^3 - 2$ получаем, что $x^3 - 2$ неразложим на множители с рациональными коэффициентами. Теперь, если бы минимальный многочлен для $\sqrt[3]{2}$ имел степень 1 или 2, то в силу леммы 14.2 ($x^3 - 2$) делился бы на него. А это невозможно в силу неразложимости $x^3 - 2$. Следовательно, минимальным многочленом для $\sqrt[3]{2}$ является многочлен третьей степени, а значит, $\sqrt[3]{2}$ — алгебраическое число степени 3.

Существует подробно разработанная теория алгебраических чисел, из которой, в частности, следует, что число $\sqrt[k]{p}$ является алгебраическим числом степени k . Откуда, в частности, следует, что для любого натурального $k \geq 1$ существует алгебраическое число степени k .

Про множество алгебраических чисел известна следующая теорема, доказательство которой опирается на теорию расширений полей.

Теорема 14.2. (1) \mathbb{A} является полем; (2) \mathbb{A} алгебраически замкнуто.

Первое утверждение говорит нам о том, что алгебраические числа можно складывать, вычитать, умножать и делить, а результат все равно останется алгебраическим числом, т.е. корнем некоторого многочлена с целыми коэффициентами. Второе — о том, что если даже мы рассмотрим кольцо многочленов $\mathbb{A}[x]$, т.е. всех многочленов с коэффициентами из поля \mathbb{A} , то корни таких многочленов все равно будут алгебраическими числами. Иначе говоря, замкнутость \mathbb{A} означает, что его невозможно расширить алгебраическими методами, как это мы проделывали с полем \mathbb{Q} . Для дальнейшего расширения \mathbb{A} понадобятся многочлены бесконечной степени, т.е. ряды.

Отметим, что если число α — алгебраическое степени $n > 1$, то числа $\alpha, \alpha^2, \dots, \alpha^{n-1}$ также являются алгебраическими степени не выше n и притом иррациональными (если бы какое-то из них было рациональным, то степень α оказалась бы ниже n). Например, рассмотрим число $(\sqrt[3]{2})^2 = \sqrt[3]{4}$. Очевидно, что это алгебраическое число степени не выше 3. Докажем, что его степень не может быть меньше 3.

Действительно, если бы оно было алгебраическим числом меньшей степени, то некоторый многочлен вида $ax^2 + bx + c$ ($a, b, c \in \mathbb{Q}$) занулял бы его, т. е. мы бы получили, что

$$a(\sqrt[3]{4})^2 + b\sqrt[3]{4} + c = 0,$$

а это эквивалентно равенству

$$2a(\sqrt[3]{2}) + b(\sqrt[3]{2})^2 + c = 0,$$

откуда следует, что многочлен $bx^2 + 2ax + c$ зануляет число $\sqrt[3]{2}$, а это неверно по доказанному ранее.

Более того, если α — алгебраическое число степени выше 1, то любая комбинация $k + at$ с целыми коэффициентами k, t ($t \neq 0$) также будет алгебраическим числом той же степени. И еще более того, любая комбинация вида

$$k_0 + k_1\alpha + k_2\alpha^2 + \dots + k_m\alpha^m,$$

где α — алгебраическое степени n , k_m — целое ненулевое число, $m < n$, также будет алгебраическим числом степени выше 1 (иначе $k_0 + k_1x + k_2x^2 + \dots + k_mx^m - (k_0 + k_1\alpha + k_2\alpha^2 + \dots + k_m\alpha^m)$ был бы многочленом степени $< n$, зануляющий α), причем все такие комбинации различны (если бы нашлись две равные друг другу комбинации с разными наборами коэффициентов, то их разность задавала бы многочлен $(k_0 - k'_0) + (k_1 - k'_1)x + \dots + (k_m - k'_m)x^m$ степени $< n$, зануляющий α).

Таким образом, мы уже существенно пополнили множество иррациональных чисел всего лишь с помощью алгебраических чисел и целочисленных линейных комбинаций их степеней, которые также оказались алгебраическими числами. Точнее, мы надстроили над множеством \mathbb{Q} бесконечный слоеный пирог, в каждом слое которого сидят алгебраические числа какой-то одной степени (а первый слой пирога — это само \mathbb{Q}), причем каждый слой содержит бесконечно много чисел. И все эти слои каким-то образом умещаются в «дырках» между рациональными числами, не смотря на то, что \mathbb{Q} всюду плотно на прямой.

Насколько же плотны алгебраические числа заданной степени на числовой прямой?

Возьмем произвольное алгебраическое число α (например, $\sqrt{2}$) степени > 1 . Это число иррациональное, т. е. не соизмеримо с целыми

числами. И пусть у нас кузнечик одной ногой прыгает на 1, а второй — на α . *Вопрос*: какими свойствами обладает множество всех тех точек, куда может допрыгнуть кузнечик?

Ясно, что все точки, в которые попадает кузнечик, описываются в общем виде формулой

$$k + \alpha t, \quad k, t \in \mathbb{Z},$$

и все они — числа-собратья исходного числа α по степени своей алгебраичности (хотя они далеко не исчерпывают множество всех алгебраических чисел такой же степени, что и степень α). То есть, выбрав число определенной алгебраической степени, кузнечик прыгает только по числам такой же степени.

Вспомним наш пример с наматыванием прямой на окружность (колесо на дороге), и выберем радиус окружности так, чтобы один полный виток по ней составлял как раз 1 единицу длины ($R = 1/2\pi$). Тогда однократное наматывание прямой на окружность будет соответствовать прыжку кузнечика на 1, а значит, с точки зрения жителя окружности кузнечик будет топтаться на месте всякий раз, когда он прыгает на любое целое число.

В то же время, если кузнечик начинает прыгать с шагом α , то он, очевидно, начинает встречаться с жителем окружности в каких-то других точках, отличных от нуля. Посмотрим, как расположены эти точки на окружности.

Для начала заметим, что кузнечик при этом никогда не повторяется, т. к. точки $k\alpha$ и $l\alpha$ при $k \neq l$, намотанные на окружность, соответствуют длинам дуг за вычетом каких-то целых оборотов:

$$k\alpha = n + \beta, \quad l\alpha = m + \gamma,$$

и если бы мы получили совпадение дуг β и γ , то получилось бы уравнение

$$k\alpha - n = l\alpha - m,$$

откуда видно, что $\alpha = (n - m)/(k - l)$ — рациональное число, а это противоречит иррациональности α .

Итак, все шаги кузнечика вида $0, \pm\alpha, \pm 2\alpha, \dots$ дают попарно различные точки на окружности.

Далее. Покажем, что какое бы маленькое ε мы ни выбрали, найдется такое целое t , что число αt будет отстоять от некоторого целого числа на расстояние меньше этого ε .

Действительно, выберем целое число N заведомо большее, чем $1/\varepsilon$, и поделим окружность на N равных секторов. В каждом секторе длина дуги будет равна $1/N$, что меньше ε . Но всех различных точек, кратных α , как мы доказали чуть выше, бесконечно много, а значит, хотя бы на

одной дуге из этих N штук окажется хотя бы 2 точки, и мы получим, что

$$k\alpha = n + \beta, \quad l\alpha = m + \gamma, \quad |\beta - \gamma| < \varepsilon.$$

Для удобства можем считать, что $\beta > \gamma$ и, следовательно, $\gamma < \beta < \gamma + \varepsilon$. Тогда

$$l\alpha - m < k\alpha - n < l\alpha - m + \varepsilon,$$

откуда

$$n - m < (k - l)\alpha < n - m + \varepsilon,$$

т. е. при $t = k - l$ число at отстоит от целого $n - m$ меньше, чем на ε .

А это значит, что кузнечик попадает в точки, сколь угодно близкие к нулю (ведь до нуля он может допрыгать своей целочисленной ногой, сделав $m - n$ шагов).

Но, умея сдвигаться на какое-то число δ от нуля хоть в какую-то сторону, кузнечик может повторять этот алгоритм раз за разом, и уходить от 0 на расстояние $\pm\delta, \pm 2\delta, \pm 3\delta$ и т. д. Таким образом, числа вида $k + at$, достигаемые кузнечиком, могут быть найдены сколь угодно близко к любой точке на прямой!

Но в таком случае множество $\{k + at \mid k, t \in \mathbb{Z}\}$ всюду плотно на прямой, как и множество \mathbb{Q} . Хотя, строго говоря, оно не содержит в себе \mathbb{Q} (ведь коэффициенты k, t — целые, а при $t \neq 0$ комбинация $k + at$ и вовсе иррациональна). Получается, что множество алгебраических чисел одной выбранной степени всюду плотно, причем не за счет \mathbb{Q} , а за счет точек, лежащих вне \mathbb{Q} !

По сути, мы получаем, что между рациональными числами столь много «дырок», что там умещается бесконечно много всюду плотных попарно непересекающихся множеств.

Вопрос: а все ли «дырки» могут быть ими заполнены? Сколько вообще существует алгебраических чисел и сколько существует «дыр» на рациональной прямой?

Для ответа на этот вопрос нам снова следует обратиться к теории множеств.

Упражнения

Обязательные упражнения

14.1^o Показать, что $\sqrt{2}$ является алгебраическим числом 2 степени.

14.2^o Показать, что $\sqrt[3]{2}$ является алгебраическим числом 3 степени.

14.3° Пусть m/n — несократимая дробь. Если m или n не являются k -ой степенью натурального числа, то $\sqrt[k]{m/n}$ является иррациональным числом.

14.4° Показать, что множество \mathbb{A} счетно.

14.5° Показать, что множества \mathbb{Q} и \mathbb{A} всюду плотны на прямой \mathbb{R} , а также, что \mathbb{A} всюду плотно на плоскости \mathbb{C} .

Для произвольных чисел x, y рассмотрим множество $Z_{(x,y)}$ точек вида $mx + ny$, где m, n — произвольные целые числа, через $Q_{(x,y)}$ — множество точек вида $px + qy$, где p, q — произвольные рациональные числа.

14.6° Если $\alpha \in Z_{(x,y)}$, то и $k\alpha \in Z_{(x,y)}$ для любого целого k .

14.7° Если x, y — целые взаимно простые числа, то $Z_{(x,y)} = \mathbb{Z}$.

14.8° Докажите, что $mx + ny = 0$ при $m, n \neq 0, m, n \in \mathbb{Z}$ тогда и только тогда, когда x/y — рациональное число.

Будем говорить, что $Z_{(x,y)}$ отделено от нуля, если найдется $C > 0$, для которого все положительные элементы $Z_{(x,y)}$ больше C .

14.9° Докажите, что α — рациональное число тогда и только тогда, когда:

- a) множество $Z_{(1,\alpha)}$ отделено от нуля;
- b) множество $Z_{(1,\alpha)}$ состоит из попарно различных элементов;
- c) множество $Q_{(1,\alpha)}$ состоит из попарно различных элементов.

14.10° По окружности длины 1 по часовой стрелке прыгает кузнечик, все прыжки имеют иррациональную длину α . Пусть M — множество точек, куда может попасть кузнечик. Докажите, что a) кузнечик никогда не попадет дважды в одну и ту же точку; b) любая дуга, содержащая начало, пересекается с M по бесконечному числу точек; c) M всюду плотно на окружности.

14.11° Доказать, что отношение « X плотно в Y », определенное для всех пар подмножеств числовой прямой \mathbb{R} , транзитивно.

14.12° Пусть α иррационально. Рассмотрим множество дробных частей чисел вида $n\alpha$, где $n \in \mathbb{N}$. Докажите, что это множество всюду плотно на отрезке $[0; 1]$ (кстати, а что это значит?).

14.13° Внутри круга запускается точечный бильярдный шар и отражается от границы по закону «угол падения равен углу отражения». Докажите, что траектория шара либо зацикливается, либо всюду плотно заполняет a) граничную окружность; b) некоторое кольцо.

14.14° Точечный конь прыгает скачками $\sqrt{2}$ и $\sqrt{3}$ по плоскости, где в каждой целой точке растёт кукуруза (круг с центром в точке). Докажите, что он обязательно сшибет хотя бы один росток (конь сшибает росток только в том случае, если приземляется на него; в прыжках конь ростки не задевает).

Дополнительные упражнения

14.15 Пусть даны точки 0 и 1 на прямой. Тогда при помощи циркуля и линейки можно построить отрезок, длина которого а) произвольное целое число; б) произвольное рациональное число; в) $\sqrt{2}$.

14.16 Пусть даны точки $(0, 0)$ и $(1, 0)$ на плоскости. Тогда при помощи циркуля и линейки можно построить точку, обе координаты которой а) произвольные целые числа; б) произвольные рациональные числа.

с) Предположим, что мы построили отрезок длины a . Докажите, что мы можем построить отрезок длины $1/a$.

д) Предположим, что мы построили отрезки длин a, b . Докажите, что мы можем построить отрезок длины ab .

е) Предположим, что мы построили отрезок длины a . Докажите, что мы можем построить отрезок длины \sqrt{a} .

14.17 Имея 0 и 1, на комплексной плоскости постройте при помощи циркуля и линейки все корни а) 3; б) 4; в) 6; г) 5-ой степени из 1.

14.18 Имея 0, 1, z, w на комплексной плоскости постройте а) \bar{z} ; б) zw ; в) z/w при $w \neq 0$.

14.19 Покажите, что построение правильного n -угольника равносильно построению числа $\sin(2\pi/n)$ или $\cos(2\pi/n)$.

14.20 Покажите, что: а) построение правильного 5-угольника равносильно построению корня уравнения $x^2 + x + 1 = 0$; б) при помощи циркуля и линейки можно отложить угол в 3 градуса; в) построение правильного 9-угольника при помощи циркуля и линейки равносильно построению угла в 10 градусов.

14.21 а) Покажите, что сумма всех корней n -ой степени из 1 равна 0. б) Найдите сумму квадратов всех корней n -ой степени из 1.

14.22 Пусть ξ_n — корень n -ой степени из 1, $\xi = \xi_7$. Покажите, что а) $\xi + \xi^{-1} = 2 \cos(2\pi/7)$.

б) Пусть $x = \xi + \xi^{-1}$. Выразите x^2, x^3 через ξ, ξ^{-1} и покажите, что $x^3 + x^2 - 2x - 1 = 0$.

в) Покажите, что многочлен $x^3 + x^2 - 2x - 1$ неприводим над \mathbb{Z} .

14.23* Докажите, что определения, введенные выше, корректны (то есть существует минимальное поле/кольцо, содержащее поле/кольцо и данное α).

14.24* Докажите, что: **a)** $\mathbb{Z}[\alpha] = \{f(\alpha) \mid f(x) \in \mathbb{Z}[x]\}$; **b)** $\mathbb{Q}(\alpha) = \{f(\alpha)/g(\alpha) \mid f(x), g(x) \in \mathbb{Z}[x], g(x) \neq 0\}$.

14.25* Докажите, что: **a)** $\mathbb{Z}[\sqrt{2}] = \mathbb{Z}_{(1, \sqrt{2})}$, **b)** $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}_{(1, \sqrt{2})} = \mathbb{Q}[\sqrt{2}]$.

14.26* Образуют ли поле числа вида: **a)** $a + b\sqrt[3]{2} + c\sqrt[3]{4}$, где $a, b, c \in \mathbb{Q}$; **b)** $a + b\sqrt{p} + c\sqrt{q} + d\sqrt{pq}$, где $a, b, c, d \in \mathbb{Q}$, а p, q — фиксированные различные простые; **c)** $\mathbb{Z}[i]/(3 + 4i)\mathbb{Z}[i]$; **d)** $\mathbb{Z}[i]/(2 + i)\mathbb{Z}[i]$?

14.27* Избавьтесь от иррациональности в знаменателе: $\frac{1}{5 + \sqrt[3]{2} + \sqrt[3]{4}}$.

14.28* Какую размерность имеют $\mathbb{Q}[\sqrt{2}]$ и $\mathbb{Q}[\sqrt[3]{2}]$ как линейные пространства над \mathbb{Q} ?

14.29* Пусть $R(x) = \left\{ \frac{P(x)}{Q(x)} \mid P(x), Q(x) \in \mathbb{R}[x], Q(x) \neq 0 \right\}$. Докажите, что $R(x)$ — поле (с обычным сложением и умножением).

Континуум

Аннотация

В этой главе мы полностью завершим наполнение геометрической прямой числами, обсудим разные версии понятия полноты вещественной прямой, построим до конца комплексную плоскость.

15.1. Мощности множеств

Множество X называется **конечным**, если его элементы можно перенумеровать натуральными числами от 1 до некоторого $n \in \mathbb{N}$, иначе говоря, X — конечно, если существует биекция $f: X \leftrightarrow \{k \in \mathbb{N} \mid 1 \leq k \leq n\}$ при некотором $n \in \mathbb{N}$. При этом число n называется **мощностью множества** X и обозначается символом $\#X$ или $\text{Card}(X)$ (встречаются также обозначения $|X|$, $\|X\|$, \overline{X}).

Отдельно оговорим случай *пустого множества* — оно тоже относится к конечным. Если в определении, приведенном выше, положить $n = 0$, то мы получим, что $\{k \in \mathbb{N} \mid 1 \leq k \leq n\} = \emptyset$, а единственная функция из X в \emptyset — это $f = \emptyset$, которая и является биекцией в случае $X = \emptyset$. Из данного выше определения следует, что мощностью пустого множества является число 0.

Конечное множество можно задать некоторой нумерованной последовательностью: $A = \{a_1, \dots, a_n\}$, предполагая при этом, что все a_k попарно различны — в этом случае $\#A = n$ (если допускать повторы среди значений a_k , то $\#A < n$).

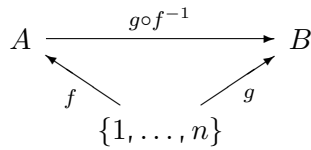
Множество, не являющееся конечным, называется **бесконечным**.

Теперь, если задано некоторое множество $B = \{b_1, \dots, b_n\}$, у которого также все b_k попарно различны, то мы вновь имеем дело со взаимно однозначным соответствием между B и множеством-эталонном $\{1, \dots, n\}$. Иначе говоря, мы имеем биекции

$$f: \{1, \dots, n\} \leftrightarrow A \text{ и } g: \{1, \dots, n\} \leftrightarrow B.$$

Но тогда сложная функция $h(a) = g(f^{-1}(a))$ устанавливает взаимно однозначное соответствие между A и B . Это можно проиллюстрировать

на диаграмме:



Таким образом, мы видим, что конечные множества A и B обладают одинаковой мощностью тогда и только тогда, когда между ними существует биекция.

Именно такой подход и выбирается для определения равномощности произвольных множеств!

Множества X и Y **равномощны** (будем это записывать так: $X \leftrightarrow Y$), если существует биекция $f : X \leftrightarrow Y$. Г. Кантор в своих ранних трудах по теории множеств определял мощность множества как «то общее, что есть у всех равномощных множеств».

Отношение \leftrightarrow рефлексивно (всякое множество само себе равномощно), симметрично (если $X \leftrightarrow Y$, то $Y \leftrightarrow X$) и транзитивно (если $X \leftrightarrow Y$ и $Y \leftrightarrow Z$, то $X \leftrightarrow Z$), т. е. является отношением эквивалентности. Это значит, что, вообще говоря, все множества можно разделить на непересекающиеся классы так, что внутри каждого класса будут находиться равномощные множества. Такой класс и принято называть **мощностью множества**.

Заметим, что при этом мы не требуем наличия какого-то эталонного множества, хотя для конечных множеств, как мы уже видели, такими могут считаться множества $\{1, \dots, n\}$ или, как принято в теории множеств, $\{0, \dots, n - 1\}$. Тем не менее, если углубиться в формальную теорию множеств, то в ней мы обнаружим кардинальные числа, которые и являются эталонными представителями классов равномощных множеств, правда, для существования кардинального числа у произвольного множества требуется *аксиома выбора*.

Ниже мы ограничимся рассмотрением всего лишь двух разных бесконечностей — *счетной* и *континуальной*, оставив в стороне упомянутые тонкости оснований математики, а эталонные множества выберем наиболее удобным способом.

Итак, описав все конечные множества как равномощные начальным отрезкам натурального ряда, возникает естественное желание посмотреть на множества, которые равномощны всему натуральному ряду, т. е. множеству \mathbb{N} . Такие множества называются **счетными**.

Например, счетными являются следующие множества:

$$2\mathbb{N}, \mathbb{Z}, 2\mathbb{Z}, \{p \in \mathbb{N} \mid p - \text{простое}\}, \mathbb{Z}[i], \mathbb{Q}, \mathbb{A}.$$

Для доказательства этого факта, очевидно, нужно предъявить взаимно однозначное соответствие между \mathbb{N} и каждым из перечисленных мно-

жеств, т. е. задать их нумерацию всеми натуральными числами (без повторов).

Нумерацию $2\mathbb{N}$ представить очень просто: каждому номеру k поставим в соответствие элемент $2k$, тем самым мы перенумеруем все четные числа, а значит, множество четных чисел счетное. Здесь мы впервые сталкиваемся с тем, что бесконечное множество равномощно какой-то своей части (которая может казаться очень маленькой, ведь точно так же устанавливается биекция между \mathbb{N} и $10^9\mathbb{N}$ и т. п.). Иногда такое свойство бесконечных множеств берется за их определение: *множество бесконечное, если оно равномощно некоторому своему собственному подмножеству*.

Биекцию $\mathbb{N} \leftrightarrow \mathbb{Z}$ установить также относительно просто: будем нумеровать целые числа по мере их удаления от нуля: $0, 1, -1, 2, -2, 3, -3$ и т. д. Ясно, что каждое целое число будет пронумеровано, и притом только один раз. Следовательно, \mathbb{Z} — счетное множество.

Для нумерации $\mathbb{Z}[i]$ нужно придумать алгоритм нумерации по «квадратным орбитам» вокруг нуля. Например, это можно сделать способом, показанным на рис. 15.1.

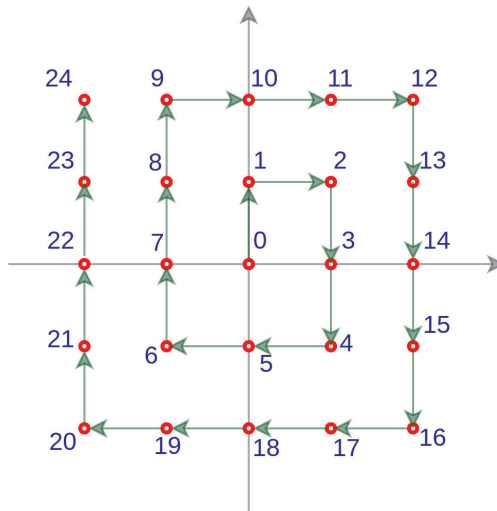


Рис. 15.1. Нумерация гауссовых чисел натуральными.

Такая нумерация дает взаимно однозначное соответствие между $\mathbb{Z} \times \mathbb{Z}$ и \mathbb{N} , и мы получаем замечательный факт, который называется **теоремой о квадрате**. То есть счетное множество равномощно своему квадрату. Конечным множествам такое и не снилось!

Можно развить это достижение и дальше. Имея нумерацию $\mathbb{Z} \times \mathbb{Z}$ и \mathbb{Z} , мы можем пронумеровать $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$, т. е. куб множества \mathbb{Z} и так далее. Любая конечная степень \mathbb{Z} равномощна \mathbb{Z} .

Рациональные числа можно рассматривать как некоторые специальные гауссовы числа, а именно, пусть дано рациональное число p/q , где $p, q \in \mathbb{Z}$ и $q > 0$, причем $p \perp q$. Тогда ему в соответствие можно поставить гауссово число $p + qi$. Ясно, что такое соответствие чисел из \mathbb{Q} и чисел из $\mathbb{Z}[i]$ является инъективным, но не взаимно однозначным, т. к. при данном соответствии мы пропускаем все числа $p + qi$ при $q \leq 0$, а также все числа вида $kp + kqi$, т. е. кратные тем, которые соответствуют рациональным. Таким образом, мы имеем инъективное вложение множества \mathbb{Q} в множество $\mathbb{Z} \times \mathbb{Z}$.

Таким образом, если мы смогли перенумеровать натуральными числами все множество $\mathbb{Z} \times \mathbb{Z}$, то мы тем самым перенумеровали и \mathbb{Q} , только с пропусками — очень много натуральных чисел выпало из такой нумерации. Интуитивно понятно, что можно скорректировать эту нумерацию так, чтобы получить взаимно однозначное соответствие между \mathbb{Q} и \mathbb{N} . То есть множество \mathbb{Q} также счетно.

Другое интуитивное соображение заключается в следующем. Нумерация $\mathbb{Z}[i]$ показывает нам, что \mathbb{Q} инъективно вкладывается в \mathbb{N} (как часть), но, с другой стороны, множество \mathbb{N} также инъективно вкладывается в \mathbb{Q} . Действительно, каждому натуральному числу n соответствует рациональная дробь вида $n/1$. Получается, что мы имеем цепочку вложений

$$\mathbb{N} \rightarrow \mathbb{Q} \rightarrow \mathbb{N},$$

т. е. множества \mathbb{N} и \mathbb{Q} взаимно друг в друга инъективно вкладываются. Отсюда напрашивается вывод, что ни одно из них не может быть «больше» другого в смысле мощности. И действительно, ниже мы докажем общую теорему об этом (см. теорему 15.2).

Для нумерации \mathbb{A} воспользуемся следующим приемом. Вспомним, что \mathbb{A} — это все корни всех многочленов с целыми коэффициентами. Возьмем тогда некоторое целое положительное число L и соберем в множество \mathbb{A}_L те и только те алгебраические числа, которые определяются многочленами вида $P(x) = k_0 + k_1x + \dots + k_nx^n$ ($k_n \neq 0$), удовлетворяющими условию

$$|k_0| + |k_1| + \dots + |k_n| + n = L.$$

Ясно, что таких многочленов существует лишь конечный набор, т. к. выбор коэффициентов k_s и степени n ограничен по модулю числом L . Но и корней у многочлена — тоже конечное количество, не превышающее его степень (см. выше теорему о корнях над полем). Таким образом, множество \mathbb{A}_L конечно.

С другой стороны, множество \mathbb{A} есть объединение всех множеств \mathbb{A}_L при $L = 1, 2, \dots$. Поэтому, нумеруя последовательно, сперва числа

из \mathbb{A}_1 , затем числа из \mathbb{A}_2 и т. д., мы пронумеруем все множество \mathbb{A} , а значит, это множество счетное!

Что же получается в итоге? Множество алгебраических чисел, которыми мы старались заткнуть все дыры между рациональными числами, равномощно множеству \mathbb{Q} ! С точки зрения мощности множества мы так ничего и не добавили к рациональным числам, хотя и позатыкали много дыр.

Возникает вопрос: *а бывают ли вообще какие-то другие мощности, кроме счетной?* Ответ на этот вопрос дает следующая теорема.

Теорема 15.1 (Кантора). *Никакое множество не равномощно множеству всех своих подмножеств.*

Доказательство. Пусть имеется множество X . Мы можем сразу считать, что оно непустое, т. к. для пустого множества теорема, очевидно, верна (в пустом множестве 0 элементов, а в множестве $\{\emptyset\}$ — один элемент). Через $\mathcal{P}(X)$ обозначим множество всех подмножеств множества X (оно называется **булеаном** множества X).

Предположим, что существует биекция $f : X \leftrightarrow \mathcal{P}(X)$. Ясно, что поскольку для всякого $x \in X$ значение $f(x)$ есть какое-то подмножество X , то возможны две ситуации: либо $x \in f(x)$, либо $x \notin f(x)$. Соберем тогда в множество Y все такие элементы x , которые удовлетворяют второму соотношению:

$$Y = \{x \in X \mid x \notin f(x)\}.$$

Понятно, что $Y \subseteq X$, а значит, $Y \in \mathcal{P}(X)$, а значит, существует единственный элемент $y \in X$ такой, что $f(y) = Y$ (поскольку f — биекция по предположению).

Вопрос: $y \in Y$ или нет?

Если $y \in Y$, то по определению множества Y получаем, что $y \notin f(y)$, но тогда $y \notin Y$. Противоречие.

Если $y \notin Y$, то по определению множества Y получаем, что **неверно** $y \notin f(y)$, т. е. $y \in Y$. Противоречие.

Любой вариант приводит к противоречию, следовательно, предположение о существовании биекции $f : X \leftrightarrow \mathcal{P}(X)$ неверно, т. е. множество X и множество всех его подмножеств неравномощны. \square

Итак, как конечные, так и бесконечные множества бывают различных мощностей. Заметим, что благодаря теореме Кантора мы можем предъявить как минимум счетное множество бесконечных попарно неравномощных множеств:

$$\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}(\mathcal{P}(\mathbb{N})), \dots, \mathcal{P}^n(\mathbb{N}), \dots$$

Напрашивается определение, позволяющее сравнивать множества по мощности, т. е. говорить о том, что одно множество больше другого в смысле мощностей. Скажем, что множество A является менее мощным, чем множество B (и множество B мощнее множества A), если существует инъекция $A \rightarrow B$, но не существует биекции $A \leftrightarrow B$. В указанном выше ряду множеств каждое следующее множество мощнее предыдущего в силу теоремы Кантора. Действительно, в $\mathcal{P}(X)$ существует подмножество следующего вида:

$$S(X) = \{\{x\} \mid x \in X\}.$$

Это — множество всех синглетонов (т. е. одноточечных множеств), образованных точками множества X . Ясно, что $X \leftrightarrow S(X)$, следовательно, существует инъекция $X \rightarrow \mathcal{P}(X)$. Таким образом, $\mathcal{P}(X)$ получается более мощным множеством, чем X .

Приведем пример. Пусть $X = \{1, 2, 3\}$. Тогда

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Подмножество синглетонов здесь — это множество $\{\{1\}, \{2\}, \{3\}\}$.

Для конечного множества X мощности n мощность множества его подмножеств $\mathcal{P}(X)$ равна 2^n . Это легко проверить, поскольку каждое подмножество X задается состояниями его элементов: каждый из них может входить в данное подмножество или не входить. Поскольку у каждого элемента ровно 2 состояния, а всего элементов n , то общее количество состояний всех элементов равно $2 \cdot 2 \cdot \dots \cdot 2$ (n раз), т. е. 2^n .

Но если с конечными множествами все укладывается в рамки обычной арифметики, то с множеством $\mathcal{P}(\mathbb{N})$ возникают проблемы. Его мощность не только не выражается натуральным числом, но она также не равна и мощности \mathbb{N} , как мы только что доказали. Эта мощность называется **мощностью континуума** и обозначается как c . Как мы видели выше, континуум — далеко не самая большая из возможных мощностей.

Итак, мы теперь знаем, что бесконечные множества отличаются по мощности. Как же можно устанавливать их равномощность, если сложно или невозможно в явном виде указать биекцию между множествами? Ответ на этот вопрос дает

Теорема 15.2 (Кантора — Бернштейна). *Если существуют две инъекции $f : A \rightarrow B$ и $g : B \rightarrow A$, то множества A и B равномощны.*

Доказательство. Введем следующие обозначения:

$$A_0 = A, \quad B_0 = B, \quad A_{n+1} = g[B_n], \quad B_{n+1} = f[A_n], \quad n \in \mathbb{N}.$$

Напомним, что квадратные скобки отвечают за обозначение образа подмножества области определения функции. Посмотрим, какими свойствами обладают две последовательности множеств A_n и B_n .

Легко видеть, что они монотонны, т. е. $A_0 \supseteq A_1 \supseteq A_2 \supseteq \dots$ и $B_0 \supseteq B_1 \supseteq B_2 \supseteq \dots$. Действительно, так как $f[A] \subseteq B$, то $B_1 \subseteq B_0$. Так как $g[B] \subseteq A$, то $A_1 \subseteq A_0$. Далее применяем соответствующие функции к множествам в этих неравенствах и получаем:

$$g[B_1] = A_2 \subseteq g[B_0] = A_1, \quad f[A_1] = B_2 \subseteq f[A_0] = B_1,$$

и так далее.

Пусть $X = \bigcap_n A_n$ и $Y = \bigcap_n B_n$. Нетрудно видеть, что $f|_X : X \leftrightarrow Y$. Действительно,

$$f[X] = f \left[\bigcap_{n=0}^{\infty} A_n \right] = \bigcap_{n=0}^{\infty} f[A_n] = \bigcap_{n=1}^{\infty} B_n = Y,$$

поскольку f — инъекция. Далее мы опускаем обозначение $f|_X$ сужения функции на множество для удобочитаемости.

Теперь, в силу того, что f и g являются инъекциями, они устанавливают следующие биективные соответствия:

$$f : A_n \setminus A_{n+1} \leftrightarrow B_{n+1} \setminus B_{n+2}, \quad g : B_n \setminus B_{n+1} \leftrightarrow A_{n+1} \setminus A_{n+2}, \quad n \in \mathbb{N}.$$

Для наглядности эти соответствия изображены на рис. 15.2.

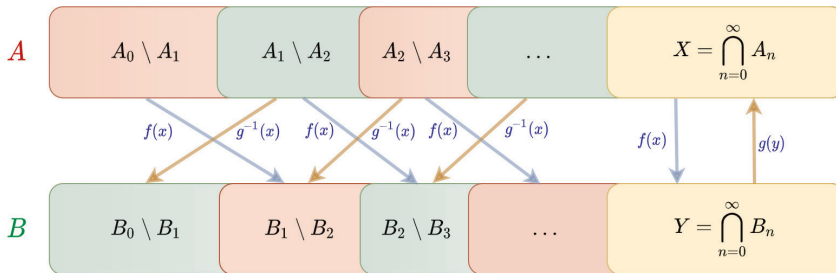


Рис. 15.2. A' и B'' — красные, A'' и B' — зеленые.

Пусть

$$A' = \bigcup_{n=0}^{\infty} A_{2n} \setminus A_{2n+1},$$

$$A'' = \bigcup_{n=0}^{\infty} A_{2n+1} \setminus A_{2n+2},$$

$$B' = \bigcup_{n=0}^{\infty} B_{2n} \setminus B_{2n+1},$$

$$B'' = \bigcup_{n=0}^{\infty} B_{2n+1} \setminus B_{2n+2}.$$

Тогда из предыдущего получаем, что $f : A' \leftrightarrow B''$ и $g : B' \leftrightarrow A''$. При этом $A' \cap A'' = \emptyset$ и $B' \cap B'' = \emptyset$.

Далее, $A = A' \sqcup A'' \sqcup X$ и $B = B' \sqcup B'' \sqcup Y$, причем $f : X \leftrightarrow Y$, как было показано выше. Теперь определим новую функцию $h : A \rightarrow B$ следующим образом:

$$h(x) = \begin{cases} f(x), & \text{если } x \in A', \\ g^{-1}(x), & \text{если } x \in A'', \\ f(x), & \text{если } x \in X. \end{cases}$$

Ясно, что $h[A] = B$ и h является инъекцией, следовательно, h — искомая биекция между A и B . \square

Теорема Кантора — Бернштейна дает простой инструмент сравнения мощностей. Достаточно показать, что первое множество равномощно какой-то части второго, а второе — какой-то части первого, т. е. что они взаимно друг в друга вкладываются. Это будет означать, что между ними существует биекция, т. е. что они равномощны.

Кроме того, эта теорема вносит ясность в терминологию о сравнении мощностей множеств. Введем следующее обозначение: $X \preceq Y$, если X менее мощное, чем Y , либо они равномощны. Нетрудно видеть, что это равносильно наличию инъекции из X в Y . Тогда теорему Кантора–Бернштейна можно сформулировать так:

$$(X \preceq Y) \wedge (Y \preceq X) \rightarrow (X \leftrightarrow Y).$$

Отношение \preceq на всех множествах является рефлексивным и транзитивным, т. е. частичным (нестрогим) порядком, поскольку $X \leftrightarrow X$ и если есть инъекции $f : X \rightarrow Y$ и $g : Y \rightarrow Z$, то $g \circ f$ — инъекция из X в Z .

Замечание для более глубокого изучения: линейным порядком отношения \preceq было бы только в том случае, если бы для любых двух множеств можно было бы их сравнить по мощности, т. е. если бы выполнялось условие $\forall X, Y (X \preceq Y) \vee (Y \preceq X)$. Это будет верно в том и только том случае, если мы принимаем аксиому выбора.

С помощью теоремы Кантора–Бернштейна легко показать равномощность \mathbb{Q} и натурального ряда. Действительно, \mathbb{Q} легко вкладывается в часть множества $\mathbb{Z} \times \mathbb{Z}$ (пары взаимно простых составляют его часть). Но $\mathbb{Z} \times \mathbb{Z}$ мы умеем явно нумеровать целыми числами, т. е. умеем построить инъекцию из $\mathbb{Z} \times \mathbb{Z}$ в \mathbb{Z} , а значит, мы имеем вложение \mathbb{Q} в \mathbb{Z} . Вложение в обратную сторону тривиально. Таким способом получается много результатов о равномощности.

Назовем множество **не более чем счетным**, если оно счетно или конечно.

Теорема 15.3. *Объединение не более чем счетного множества не более чем счетных множеств не более чем счетно.*

Доказательство. Поскольку мы имеем не более чем счетный набор множеств, они уже как-то пронумерованы натуральными числами либо от 0 до некоторого n , либо от 0 до бесконечности. Иначе говоря, задана функция $A(k)$, действующая из \mathbb{N} или его начального отрезка в некоторое множество множеств. Если эта функция задана только для $k = 0, \dots, n$, мы можем ее продолжить константой $A(k) = \emptyset$ при $k > n$. Тем самым при необходимости мы продолжим конечную последовательность множеств до счетной, не меняя результат их объединения. Поэтому мы сразу можем предполагать, что нам дан счетный набор не более чем счетных множеств.

Пусть $A_n = A(n)$. Тогда требуемое объединение равно

$$A = \bigcup_{n=0}^{\infty} A_n,$$

где все A_n — не более чем счетные множества.

Поскольку нам известно, что A_n — не более чем счетное, значит, существуют биекции f_n между A_n и либо \mathbb{N} , либо каким-то конечным отрезком \mathbb{N} . В любом случае f_n — это инъекция из A_n в \mathbb{N} .

Теперь построим инъекцию из A в $\mathbb{Z} \times \mathbb{Z}$.

Пусть $a \in A$. Тогда существует n такой, что $a \in A_n$. Может оказаться так, что a лежит сразу во многих множествах A_n , в таком случае выберем наименьший из их номеров и обозначим его за n_a . Поскольку $a \in A_{n_a}$, мы можем применить к нему функцию f_{n_a} . Положим далее

$$g(a) = \langle n_a, f_{n_a}(a) \rangle.$$

Ясно, что g — инъекция, т. к. для разных точек a и a' либо отличаются номера n_a и $n_{a'}$, и следовательно, $g(a) \neq g(a')$ по свойствам упорядоченной пары, либо у них общий номер n_a , но тогда отличаются значения $f_{n_a}(a)$ и $f_{n_a}(a')$, поскольку f_{n_a} является инъекцией, и, стало быть, $g(a) \neq g(a')$.

Итак, у нас есть инъекция $g : A \rightarrow \mathbb{Z} \times \mathbb{Z}$. К ней можно применить биекцию $f : \mathbb{Z} \times \mathbb{Z} \leftrightarrow \mathbb{N}$, которую мы ранее выписывали в явном виде (см. рис. 15.1). Тогда композиция $h(a) = f(g(a))$ будет инъекцией из A в \mathbb{N} .

Область значений h есть подмножество в \mathbb{N} . Она либо имеет максимум, и тогда это конечное множество, либо не имеет максимума, и тогда это бесконечное множество. В любом случае, область значений h можно перенумеровать так, чтобы номера шли подряд от нуля без пропусков. И тогда с помощью h^{-1} мы построим биекцию между A и либо конечным отрезком натурального ряда, либо самим \mathbb{N} . \square

15.2. Изоморфизмы

Установление биекции между множествами позволяет нам судить об их количественном сходстве, но ничего не говорит о том, насколько похожи могут быть структуры, заданные на них. Поэтому на базе понятия биекции строятся более сильные критерии «похожести» двух множеств.

Центральным термином здесь является «**изоморфизм**». Это — биекция, сохраняющая операции (например, сложение или умножение) и/или отношения (например, линейный порядок или отношение эквивалентности) и/или функционалы (например, норма или длина), заданные на этих двух множествах.

Поясним. Пусть даны группа \mathbb{Z}_4 со сложением по модулю 4 и группа корней четвертой степени из единицы с операцией умножения (обратимые гауссовы числа). В обоих множествах 4 элемента, следовательно, существует биекция между ними. Причем таких биекций ровно столько, сколько перестановок в группе \mathbf{S}_4 , т. е. 24 штуки. Однако, если дополнительно потребовать, чтобы результат сложения двух элементов в первой группе переходил в результат умножения образов этих элементов во второй группе, то таких биекций окажется всего две. Они-то и будут изоморфизмами этих групп.

Рассмотрим таблицы умножения группы $(\mathbb{Z}/9\mathbb{Z})^*$ и сложения группы $\mathbb{Z}/6\mathbb{Z}$:

$(\mathbb{Z}/9\mathbb{Z})^*$	1	2	4	5	7	8
1	1	2	4	5	7	8
2	2	4	8	1	5	7
4	4	8	7	2	1	5
5	5	1	2	7	8	4
7	7	5	1	8	4	2
8	8	7	5	4	2	1

$\mathbb{Z}/6\mathbb{Z}$	0	1	2	5	4	3
0	0	1	2	5	4	3
1	1	2	3	0	5	4
2	2	3	4	1	0	5
5	5	0	1	4	3	2
4	4	5	0	3	2	1
3	3	4	5	2	1	0

Во второй таблице мы специально перемешали порядок элементов таким образом, чтобы показать изоморфизм групп, при котором умножение в $(\mathbb{Z}/9\mathbb{Z})^*$ соответствует сложению в $\mathbb{Z}/6\mathbb{Z}$, а соответствие элементов можно установить по правилу: $a \mapsto 2^a \pmod{9}$, где $a \in \mathbb{Z}/6\mathbb{Z}$, поскольку $(\mathbb{Z}/9\mathbb{Z})^* = \langle 2 \rangle$. Аналогичное соответствие можно построить, используя число 5 в качестве основания степени. А вот соответствие $a \mapsto 7^a \pmod{9}$ не будет изоморфизмом, поскольку $\langle 7 \rangle = \{1, 4, 7\}$. Аналогично: $a \mapsto 8^a \pmod{9}$ не будет изоморфизмом, поскольку $\langle 8 \rangle = \{1, 8\}$.

Заметим, что не любая мультипликативная группа $(\mathbb{Z}/m\mathbb{Z})^*$ изоморфна некоторой аддитивной группе $\mathbb{Z}/n\mathbb{Z}$. Например, в группе $(\mathbb{Z}/8\mathbb{Z})^*$ содержится 4 элемента, но ни один из них не является образующим,

группа $(\mathbb{Z}/8\mathbb{Z})^*$ не является циклической, а значит, она не может быть изоморфна $\mathbb{Z}/4\mathbb{Z}$.

Еще пример. Рассмотрим множества \mathbb{Z} и $2\mathbb{Z}$ с обычными операциями сложения и умножения и обычным линейным порядком на них. Мы уже знаем, что эти множества равномощны. Но посмотрим повнимательнее на биекцию $f(n) = 2n$, действующую из \mathbb{Z} в $2\mathbb{Z}$. Оказывается, что:

$$f(n + m) = f(n) + f(m), \quad n < m \Leftrightarrow f(n) < f(m),$$

т.е. f сохраняет сложение и порядок. А это значит, что f является изоморфизмом упорядоченных групп $\langle \mathbb{Z}, +, < \rangle$ и $\langle 2\mathbb{Z}, +, < \rangle$.

Однако f не сохраняет умножение, поскольку $f(nm) = 2nm \neq f(n)f(m)$. Следовательно, f не является изоморфизмом колец $\langle \mathbb{Z}, +, \cdot \rangle$ и $\langle 2\mathbb{Z}, +, \cdot \rangle$. Более того, эти два кольца вовсе не изоморфны. Дело в том, что изоморфизм должен сохранять единицу, т.е. если какое-то число e является единицей по умножению в первом кольце, то $f(e)$ будет единицей во втором кольце. Просто потому, что $ne = n$ соответствует $f(n) = f(n)f(e)$. Но чему бы ни было равно $f(1)$ в кольце $2\mathbb{Z}$, оно не обладает свойствами единицы (в $2\mathbb{Z}$ просто нет единицы), а значит, эти кольца не изоморфны.

Бывает и еще хуже. Изоморфизм работает только по отношению порядка, но не работает по алгебраическим операциям. Для этого достаточно вспомнить два изученных нами поля: \mathbb{Q} и \mathbb{A} .

Ясно, что эти поля не могут быть изоморфны по операциям, т.к. иначе в обоих полях одинаково бы разрешалось или не разрешалось уравнение $x^2 = 2$. Но мы знаем, что оно разрешается в \mathbb{A} и не разрешается в \mathbb{Q} . Тем не менее с порядковым изоморфизмом у них все в порядке.

Теорема 15.4. *Все счетные неограниченные сверху и снизу плотные линейно упорядоченные множества порядково изоморфны друг другу.*

Иначе говоря, пусть у нас имеется два множества A и B , которые счетны (т.е. все их элементы можно перенумеровать натуральными числами), на них заданы линейные порядки $<_A$ и $<_B$ такие, что в обоих множествах нет ни наибольшего, ни наименьшего элемента, и эти множества плотны в своем порядке, тогда существует изоморфизм $f : A \leftrightarrow B$, сохраняющий порядок, т.е. такой, что $f(x) <_B f(y) \Leftrightarrow x <_A y$.

Доказательство. Будем строить соответствие пошагово. Пусть мы сделали некоторое соответствие для подмножеств $A_n \subset A$ и $B_n \subset B$ из n элементов. Возьмем любой элемент одного из множеств (для определенности A), который не вошел в A_n . Посмотрим, в каком отношении он находится со всеми элементами из A_n . Он оказался либо наибольшим

элементом, либо наименьшим элементом, либо стоящим между некоторыми элементами a_i и a_{i+1} . Найдем элемент в B , находящийся в таком же отношении со всеми элементами B_n . Мы можем это сделать, так как B — плотное множество без наибольшего и наименьшего элементов. Будем считать эти два элемента сопоставленными. Таким образом, мы научились получать из соответствия для n элементов соответствие для $n + 1$ элемента. Чтобы в пределе получить соответствие для всех элементов, воспользуемся счетностью множеств A и B . Пронумеруем все элементы и на каждом четном шаге будем выбирать еще не взятый элемент из множества A с наименьшим номером, а на нечетном — из B . \square

Из этой теоремы следует, например, что множество \mathbb{Q} с обычным линейным порядком и множество \mathbb{A} всех алгебраических чисел с обычным линейным порядком порядково изоморфны! Более того, рациональный интервал $(a; b)$ порядково изоморфен всему множеству \mathbb{Q} .

15.3. Действительные числа

Для обозначения множества всех точек числовой прямой мы используем символ \mathbb{R} . При этом мы предполагаем естественные включения $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. Основная задача данной главы заключается в том, чтобы дать более точное определение этому символу.

Вспомним снова про множество \mathbb{B} , которое состоит из рациональных чисел вида $k/2^n$, где $k \in \mathbb{Z}$, $n \in \mathbb{N}$.

Это множество есть собственное подмножество \mathbb{Q} . Оно счетно и всюду плотно. Но главное — оно очень удобно устроено. При $n = 0$ мы имеем целочисленную решетку на числовой прямой, при $n = 1$ мы получаем все целые и полуцелые числа, при $n = 2$ — все числа с шагом $1/4$. Обозначим

$$\mathbb{B}_n = \{k/2^n \mid k \in \mathbb{Z}\}.$$

Тем самым мы определяем некоторый слой в множестве \mathbb{B} с фиксированным шагом, равным $1/2^n$.

Множества \mathbb{B}_n хороши тем, что образуют возрастающую последовательность по вложению, которая стартует с \mathbb{Z} и заканчивается \mathbb{B} :

$$\mathbb{Z} = \mathbb{B}_0 \subset \mathbb{B}_1 \subset \mathbb{B}_2 \subset \dots \subset \mathbb{B}.$$

В таком случае принято называть множество \mathbb{B} **пределом возрастающей цепи множеств**.

Поскольку расстояние между соседними точками \mathbb{B}_n очень быстро сокращается с ростом n , то любая точка на прямой может быть сколь угодно точно приближена точками множества \mathbb{B} .

Здесь мы раз и навсегда должны договориться о том, что мы работаем на *архимедовой* числовой прямой. Это значит, что как бы мы ни выбрали на прямой две точки A и B , найдется такое натуральное число n , что длина отрезка AB не превосходит n . По-другому этот принцип формулируется так: каковы бы ни были два отрезка на прямой, взятые кратно, они могут превзойти друг друга.¹ Данный принцип, называемый **аксиомой Архимеда**, был впервые сформулирован Евдоксом Книдским и развит в сочинениях Архимеда. Аксиома Архимеда исключает на числовой прямой бесконечно малые и бесконечно большие величины, а также гарантирует, что множество целых чисел \mathbb{Z} не ограничено ни сверху, ни снизу.

Ниже мы отметим места в наших рассуждениях, где и как она помогает получать некоторые результаты.

Процесс последовательного приближения произвольной точки A на числовой прямой можно осуществить следующим образом (рис. 15.3).

Шаг 1. Находим целое число k такое, что точка A лежит в полуинтервале $[k; k + 1)$, т. е. либо между соседними целыми числами, либо совпадает с целым числом. Очевидно, что такое k определяется однозначно и всегда существует (благодаря аксиоме Архимеда). Обозначим отрезок $[k; k + 1]$ за Δ_0 . Ясно, что его границы k и $k + 1$ есть элементы множества \mathbb{B}_0 .

Шаг 2. Пусть далее n — это номер текущего интервала (начиная с нуля).

Шаг 3. Находясь в отрезке Δ_n , делим его на 2 части посередине так, чтобы получилось два одинаковых полуинтервала: если $\Delta_n = [a; b]$, то новые полуинтервалы будут $[a; (a + b)/2)$ и $[(a + b)/2; b)$. Точка A лежит в одном из этих полуинтервалов: либо в левом, либо в правом, третьего не дано.

Заметим, что и границы Δ_n , и его середина — это точки множества \mathbb{B} , причем границы Δ_n находятся в множестве \mathbb{B}_n , а его середина — в множестве \mathbb{B}_{n+1} . Таким образом, это подразбиение является переходом к следующему уровню разбиения в множестве \mathbb{B} .

Шаг 4. Тогда через Δ_{n+1} обозначим отрезок $[a; (a + b)/2]$, если $A \in [a; (a + b)/2)$, и отрезок $[(a + b)/2; b]$, если $A \in [(a + b)/2; b)$. После чего перейдем на предыдущий шаг, увеличивая номер n на 1.

В результате мы получим последовательность вложенных отрезков

$$[k; k + 1] = \Delta_0 \supset \Delta_1 \supset \Delta_2 \supset \dots$$

¹ Цитата из «Начал» Евклида.

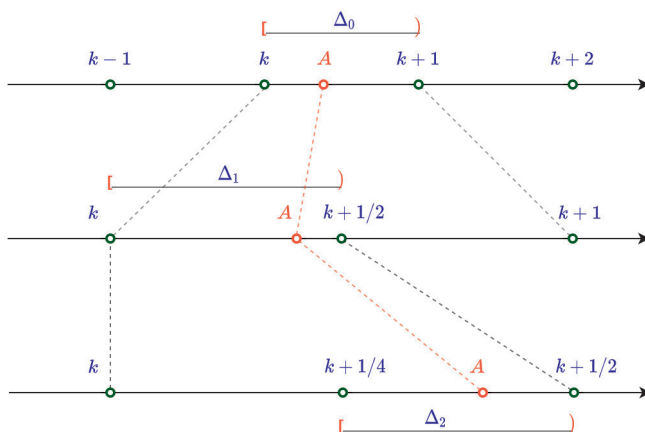


Рис. 15.3. Приближение методом деления отрезка пополам.

Эта последовательность монотонно убывает, причем на каждом шаге отрезок становится вдвое короче, а концы отрезков прыгают по точкам множества \mathbb{B} , постепенно переходя ко все более мелкой сетке — от \mathbb{B}_n к \mathbb{B}_{n+1} .

Где же в итоге окажется точка A ?

Поскольку $A \in \Delta_n$ для всех n , то она также лежит в пересечении всех этих отрезков:

$$A \in \bigcap_n \Delta_n.$$

Такое пересечение называется **пределом убывающей цепи множеств**.

Может ли в этом пересечении лежать еще какая-то точка? Ответ: нет. Если мы возьмем любую другую точку $B \neq A$, то, очевидно, что между ними есть какое-то расстояние $\varepsilon > 0$. Возьмем² тогда такое n , что $\varepsilon > 1/2^n$, и посмотрим на отрезок Δ_n . Его длина равна $1/2^n$, и он содержит точку A . Но тогда он не содержит точку B , а значит, и пересечение всех отрезков Δ_n не содержит точку B .

Итак, точка A — единственный представитель пересечения отрезков Δ_n :

$$\bigcap_n \Delta_n = \{A\}.$$

По сути, мы уже сформулировали принцип непрерывности (полноты) числовой прямой — *принцип вложенных отрезков*. Однако здесь нужно проявить осмотрительность. Дело в том, что мы не доказали существование точки A , а сразу же выбрали ее из существующих то-

² Чтобы определить такое n , заметим, что $n < 2^n$ при $n \in \mathbb{N}$ (что легко доказать индукцией), и выберем в качестве n целую часть $1 + 1/\varepsilon$.

чек прямой. Но в настоящий момент нам известны только рациональные числа и алгебраические, поэтому разумно ожидать, что точка A есть одно из таких чисел. Но рассмотренный принцип «ловли» произвольной точки на прямой с помощью стягивающейся сетки двоично-рациональных чисел сам по себе является мощным инструментом для определения новых чисел и, что самое главное, для ликвидации всех возможных дыр на числовой оси, состоящей из алгебраических чисел.

На самом деле вовсе не очевидно, что если мы выберем произвольную последовательность вложенных отрезков, длина которых стремится к нулю с ростом номера, то в пределе получим множество, состоящее из одной точки. Быть может, никакой точки там вовсе не окажется. Поэтому приходится вводить принцип непрерывности при помощи **аксиомы непрерывности** (она же — *аксиома полноты*).

У этой аксиомы существует несколько равносильных (с учетом аксиомы Архимеда) формулировок, и мы начнем с той, к которой нас подготовил сюжет по поимке точки A в сеть множества \mathbb{B} .

A1. Принцип вложенных отрезков

Формулировка A1. Пусть дана последовательность вложенных отрезков на прямой:

$$[a_0; b_0] \supseteq [a_1; b_1] \supseteq [a_2; b_2] \supseteq \dots [a_n; b_n] \supseteq \dots,$$

где $a_n < b_n$ для всех n . Тогда множество точек, принадлежащих всем отрезкам одновременно, не пусто.

В терминах, которые мы упоминали выше, принцип **A1** можно переформулировать так: любая убывающая по вложению бесконечная цепь отрезков имеет непустой предел.

Отметим, что в формулировке мы не требовали, чтобы концы отрезков были двоично-рациональными числами, а также не требовали, чтобы их длина стремилась к нулю. Очевидно, что случай с двоично-рациональными отрезками является частным случаем последовательности вложенных отрезков, а значит, в силу принципа **A1** имеет непустой предел.

На самом деле, если сформулировать принцип вложенных отрезков, используя только двоично-рациональные концы отрезков и деление отрезка пополам на каждом шаге, мы получим эквивалентную формулировку принципа вложенных отрезков. Но доказательство этого факта мы оставим за рамками курса.

Принцип вложенных отрезков уже позволяет нам доказать, что на числовой прямой существуют не только алгебраические числа, более того, что точек на прямой существует несчетное множество.

Предположим, что это не так, и пусть на прямой есть только счетный набор точек. В соответствии с определением счетности мы можем перенумеровать все эти точки натуральными числами, так что вся прямая — это множество $\{x_0, x_1, x_2, \dots\}$.

Построим цепь вложенных отрезков следующим способом. Выберем любой отрезок Δ_0 (можно считать, что он имеет концы в множестве \mathbb{B} , но это не обязательно) так, чтобы $x_0 \notin \Delta_0$. Точка x_1 может лежать или не лежать в отрезке Δ_0 , но в любом случае мы можем выбрать отрезок Δ_1 так, чтобы выполнялись условия: $\Delta_1 \subseteq \Delta_0$ и $x_1 \notin \Delta_1$. Заметим, что при этом также $x_0 \notin \Delta_1$. Далее, точка x_2 может лежать или не лежать в отрезке Δ_1 , но мы всегда можем выбрать отрезок $\Delta_2 \subseteq \Delta_1$ так, чтобы $x_2 \notin \Delta_2$. Продолжаем эту процедуру до бесконечности, поддерживая следующую ситуацию:

$$\Delta_{n+1} \subseteq \Delta_n, \quad x_0, \dots, x_{n+1} \notin \Delta_{n+1}.$$

Но тогда в пересечении $\bigcap \Delta_n$ нет ни одной точки x_n , т. е. нет вообще ни одной точки числовой прямой. Однако в силу принципа **A1** там должна быть хотя бы одна точка. Противоречие.

Таким образом, принцип вложенных отрезков гарантирует нам несчетность множества чисел на прямой. Насколько велико это множество, мы сможем оценить чуть позже.

Следующее, что можно отметить, опираясь на наш пример с поимкой точки A в сеть множества \mathbb{B} , это то, что последовательность левых границ вложенных отрезков не убывает. На каждом шаге левая граница Δ_n либо остается такой же, как у предыдущего отрезка, либо перескакивает в его середину. Но при этом все левые границы отрезков Δ_n остаются ограниченными сверху правой границей начального отрезка Δ_0 (а также правой границей любого следующего отрезка). То же самое мы видим и в ситуации с произвольной последовательностью вложенных отрезков, которую мы описывали при формулировке принципа **A1**.

Аналогичное наблюдение можно вывести и для правых концов вложенных отрезков. Мы имеем некую бесконечную монотонную последовательность точек, и притом ограниченную, т. е. находящуюся в некотором заранее известном отрезке.

Введем определения. Последовательность $\{x_n\}_{n=0}^{\infty}$ называется **невозрастающей (неубывающей)**, если для всех n выполняется неравенство $x_n \geq x_{n+1}$ (соответственно, $x_n \leq x_{n+1}$). Невозрастающие и неубывающие последовательности называются **монотонными**. Последовательность **строго** монотонна (строго убывающая или строго возрастающая), если указанное неравенство всегда строгое.

Множество X на прямой называется **ограниченным сверху**, если существует точка a такая,³ что $X \leq a$. Если ситуация противоположная, т. е. $X \geq a$, то множество X называется **ограниченным снизу**. Если множество ограничено сверху и снизу, то оно называется **ограниченным**.

Заметим, что в силу архимедовости числовой прямой определение ограниченного множества можно дать в эквивалентной форме: множество X ограничено сверху, если $X \leq n$ при некотором $n \in \mathbb{N}$, и множество X ограничено снизу, если $X \geq -n$ при некотором $n \in \mathbb{N}$.

Последовательность **ограничена** (сверху и/или снизу), если ограничено множество ее значений (сверху и/или снизу). Отметим, что последовательность мы рассматриваем не просто как множество точек на прямой, а как функцию из \mathbb{N} в множество точек прямой или любое другое множество. Это позволяет рассматривать, например, стационарные последовательности, когда $x_n = \text{const}$, или циклические последовательности, когда x_n принимает конечный набор значений, последовательно повторяя их. Например, $x_n = n \bmod m$ повторяет значения $0, 1, \dots, m-1$.

Имея любую монотонную ограниченную последовательность, мы можем сопоставить ей цепь вложенных отрезков. А именно, если эта последовательность неубывающая, то в качестве левых границ отрезков берем ее элементы, а правую границу зафиксируем в какой-то одной точке, которая точно больше всех членов последовательности (ее существование следует из ограниченности последовательности).

Введем одно из основных понятий математического анализа. Число a называется **пределом последовательности** $\{x_n\}$, что обозначается равенством

$$a = \lim_{n \rightarrow \infty} x_n,$$

если для любого $\varepsilon > 0$ существует такой номер N , что для всех $n > N$ имеет место неравенство $|x_n - a| < \varepsilon$. Если последовательность $\{x_n\}$ имеет предел, то она называется **сходящейся** (к данному пределу), в противном случае — **расходящейся**.

Назовем интервал $U_\varepsilon(a) = (a - \varepsilon; a + \varepsilon)$ ε -окрестностью точки a . Ясно, что утверждение $|x_n - a| < \varepsilon$ означает, что x_n лежит в ε -окрестности точки a . Кроме того, очень часто говорят «почти все члены последовательности», когда имеют в виду некоторый ее хвост, т. е. ту же последовательность, но за исключением, быть может, какого-то ее конечного начального отрезка. Поэтому тот факт, что a является пределом последовательности $\{x_n\}$, можно записать следующими словами: *в любой сколь угодно малой окрестности точки a лежат почти все члены последовательности $\{x_n\}$.*

³ Мы ранее уже вводили сравнение множеств и множества с точкой. $X \leq Y$, если для всех $x \in X, y \in Y$ имеем $x \leq y$. $X \leq a$, если $X \leq \{a\}$.

Стоит отметить, что символика пределов в обязательном порядке предписывает указывать, при каком именно изменении параметра совершается предельный переход. В нашем случае параметром последовательности является индекс (или номер) ее членов, т. е. число n . И если еще раз внимательно прочесть определение предела, а также смысл фразы «почти все члены последовательности», то мы увидим, что требование малого отклонения от предела выполняется при больших n , т. е. при $n > N$ при некотором номере N , который, вообще говоря, зависит от выбранного ε . И весь смысл данного предела в том, что при забегании индекса n в сторону бесконечности величина x_n становится близкой к a .

Кроме того, параметров может быть несколько, и поэтому всегда следует указывать, относительно какого из них осуществляется предельный переход.

Приведем пример. Пусть $x_{n,m} = m/(1 + (n - m)^2)$. Найдем ее предел при $n \rightarrow \infty$. Мы можем построить экспериментальный график (см. рис. 15.4) и убедиться в том, что предел равен нулю, при каком бы значении m он ни вычислялся.

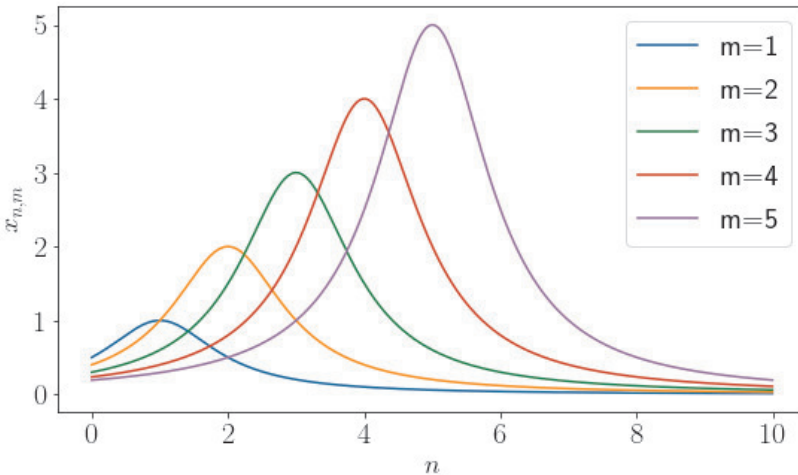


Рис. 15.4. График $x_{n,m} = \frac{m}{1 + (n - m)^2}$ для некоторых значений m .

Докажем, что так оно и есть на самом деле. Выберем произвольный достаточно малый $\varepsilon > 0$. Поскольку $x_{n,m} > 0$, нам достаточно установить, при каких n выполняется неравенство $x_{n,m} < \varepsilon$, что и будет означать близость к нулю. Заметим, что нас интересует поведение $x_{n,m}$ при больших n , поэтому мы можем сразу же предположить, что $n > m$, чтобы выполнялось равенство $\sqrt{(n - m)^2} = n - m$. Тогда, решая указанное

неравенство, находим, что

$$n > m + \sqrt{\frac{m}{\varepsilon} - 1},$$

откуда видно, что начиная с некоторого N (например, можно округлить вверх корень и добавить m) для всех последующих n требуемое неравенство выполняется. Стало быть, по определению получаем, что

$$\lim_{n \rightarrow \infty} \frac{m}{1 + (n - m)^2} = 0.$$

Однако даже по графику видно, что если мы будем менять параметр m вместе с n , оставаясь на гребне волны графика, то мы увидим рост величины $x_{n,m}$. Можно так подобрать зависимость параметров n и m (например, положить $m = kt^2$ и $n = kt^2 + t$), чтобы получить сходимость $x_{n,m}$ к любому наперед заданному положительному числу k ($x_{n,m} = kt^2/(1 + t^2) \rightarrow k$ при $t \rightarrow \infty$). Можно даже добиться стремления $x_{n,m}$ к бесконечности (например, полагая $n = m$).

Поэтому всегда очень важно указывать, при каком изменении какого параметра ищется предел.

A2. Предел монотонной последовательности

Формулировка A2. *Всякая монотонная ограниченная последовательность имеет предел.*⁴

Из рассуждений, проведенных выше относительно монотонной ограниченной последовательности, ясно, что она должна иметь предел в силу принципа **A1**. На самом деле, верно и обратное: если все монотонные ограниченные последовательности имеют предел, то выполняется принцип вложенных отрезков. Это совсем легко установить, поскольку границы вложенных отрезков образуют две монотонные ограниченные последовательности, а значит, имеют пределы. Причем как эти пределы, так и точки, лежащие между ними, будут принадлежать пределу цепи вложенных отрезков.

Таким образом, справедлива следующая теорема.

Теорема 15.5. *Принцип A2 равносильен принципу A1.*

Следующее, что мы можем заметить из сюжета с поимкой точки A сетью точек множества \mathbb{B} , это то, что границы отрезков не просто стремятся к точке A , но и как бы стремятся друг к другу, т. е. расстояния

⁴ Если быть точным, то еще требуется архимедовость системы чисел, в которой этот принцип постулируется, но для числовой прямой, в которой множество \mathbb{Z} не ограничено ни сверху, ни снизу, это выполняется автоматически.

между всеми членами последовательности, начиная с некоторого номера, становятся сколь угодно малыми. Поэтому, даже ничего не зная о существовании предела, мы можем сделать некоторые выводы о последовательности.

Последовательность $\{x_n\}$ называется **фундаментальной**, если для любого $\varepsilon > 0$ существует такой номер N , что для всех номеров $n, m > N$ выполняется $|x_n - x_m| < \varepsilon$.

А3. Предел фундаментальной последовательности

Формулировка А3. *Всякая фундаментальная последовательность имеет предел.*

Теорема 15.6. *Принцип А3 равносильен двум предыдущим принципам (при выполнении принципа Архимеда).*

Доказательство. Докажем, что **А3** следует из **А1**. Пусть дана фундаментальная последовательность $\{x_n\}$. Она, очевидно, ограничена, поскольку, выбирая, например, $\varepsilon = 1$, в силу свойства фундаментальности получаем, что существует номер N такой, что для всех $n, m > N$ имеет место неравенство $|x_n - x_m| < 1$. В частности, положим $n = N + 1$, тогда для произвольного $m > N$ будем иметь

$$|x_{N+1} - x_m| < 1 \Rightarrow x_{N+1} - 1 < x_m < x_{N+1} + 1,$$

т. е. $x_m \in [x_{N+1} - 1; x_{N+1} + 1]$ при $m > N$. Пусть далее $a = x_{N+1} - 1$ и $b = x_{N+1} + 1$. Тогда $x_m \in [a; b]$ для всех $m > N$. При этом длина отрезка $[a; b]$ равна 2.

Теперь разделим отрезок $[a; b]$ пополам и рассмотрим два новых отрезка $[a; (a+b)/2]$ и $[(a+b)/2; b]$. Как минимум в одном из них находится бесконечное количество членов последовательности $\{x_n\}$, его и обозначим за $[a_1; b_1]$. Пусть $x_{n_1} \in [a_1; b_1]$ (мы можем выбрать в качестве n_1 наименьший такой номер $n > N$, что $x_n \in [a_1; b_1]$). Разделим отрезок $[a_1; b_1]$ пополам точкой $(a_1 + b_1)/2$: снова один из получившихся отрезков содержит бесконечно много членов последовательности $\{x_n\}$, обозначим его через $[a_2; b_2]$. Снова выберем в нем $x_{n_2} \in [a_2; b_2]$, причем так, что $n_2 > n_1$. Продолжая этот процесс, получаем последовательность отрезков $[a_k; b_k]$, в которой каждый последующий является половиной предыдущего, и подпоследовательность элементов x_{n_k} таких, что $x_{n_k} \in [a_k; b_k]$ и $n_{k''} > n_{k'}$ при $k'' > k'$.

Последовательность отрезков $[a_k; b_k]$ является последовательностью вложенных отрезков по построению, а значит, ее предел $\bigcap_k [a_k; b_k]$ не пуст в силу принципа вложенных отрезков **А1**. В то же время, множество $\bigcap_k [a_k; b_k]$ не может содержать более одной точки, т. к. длина

отрезка $[a_k; b_k]$ равна $(b - a)/2^k = 2/2^k$, и если бы в множестве $\bigcap_k [a_k; b_k]$ было бы как минимум две точки на расстоянии $\delta > 0$, то мы нашли бы такое k ,⁵ что $\delta > 2/2^k$, т.е. некоторый отрезок $[a_k; b_k]$ оказался бы короче, чем расстояние между этими точками. Следовательно, в $\bigcap_k [a_k; b_k]$ ровно одна точка. Обозначим ее за a .

Покажем, что a является пределом $\{x_n\}$. Возьмем произвольное $\varepsilon > 0$. Тогда по построению отрезков $[a_k; b_k]$ найдется такое $K = K_\varepsilon$, что $b_K - a_K < \varepsilon/2$ (снова работает аксиома Архимеда). В то же время $x_{n_k}, a \in [a_K; b_K]$ при $k > K$, следовательно, $|x_{n_k} - a| < \varepsilon/2$ при $k > K$. Так как последовательность $\{x_n\}$ фундаментальна, то существует такое $M = M_\varepsilon$, что для всех $n, m > M_\varepsilon$ имеем $|x_n - x_m| < \varepsilon/2$.

Положим $N_\varepsilon = \max\{K_\varepsilon, M_\varepsilon\}$ и зафиксируем некоторое $m = n_k \geq N_\varepsilon$. Тогда для всех $n > N_\varepsilon$ получим

$$|x_n - a| \leq |x_n - x_{n_k}| + |x_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

а это и доказывает, что $\lim_{n \rightarrow \infty} x_n = a$.

Таким образом, **A1** \Rightarrow **A3**.

Докажем, что из **A3** следует **A2**. Пусть $\{x_n\}$ — неубывающая ограниченная сверху последовательность, и пусть при всех n $x_n \leq A$ при некотором A (в силу ограниченности $\{x_n\}$ оно существует). Предположим, что $\{x_n\}$ не является фундаментальной. Тогда существует $\varepsilon > 0$ такой, что для любого номера N существуют номера $n_N, m_N > N$ такие, что $|x_{n_N} - x_{m_N}| \geq \varepsilon$. Не ограничивая общности, мы можем считать, что $n_N > m_N$, так что в силу неубывания нашей последовательности мы сможем записать $x_{n_N} \geq x_{m_N} + \varepsilon$ и, следовательно, $x_{n_N} \geq x_N + \varepsilon$. Иначе говоря, для любого N найдется $n > N$ такой, что $x_n \geq x_N + \varepsilon$.

Пусть $N_1 = 1$, тогда выберем $n_1 > N_1$ такой, что $x_{n_1} \geq x_1 + \varepsilon$. Далее выберем $N_2 = n_1$ и для некоторого $n_2 > N_2$ получим неравенство $x_{n_2} \geq x_{N_2} + \varepsilon \geq x_1 + 2\varepsilon$. Далее выберем $N_3 = n_2$ и для некоторого $n_3 > N_3$ получим $x_{n_3} \geq x_1 + 3\varepsilon$. И так далее. На k -ом шаге мы найдем такое n_k , что $x_{n_k} \geq x_1 + k\varepsilon$. Остается выбрать k так, чтобы получилось неравенство $x_{n_k} \geq A + \varepsilon$. Для этого достаточно взять в качестве k целую часть числа $(A - x_1)/\varepsilon + 1$ (и тут мы снова используем принцип Архимеда).

Итак, предполагая, что последовательность $\{x_n\}$ не является фундаментальной, мы пришли к противоречию с ее ограниченностью. Аналогично рассматривается случай невозрастающей последовательности $\{x_n\}$, ограниченной снизу.

⁵ В данном месте используется принцип Архимеда: взяв число $1/\delta$, мы можем найти превосходящее его натуральное число $k-1$, которое (что уже доказывается индукцией) строго меньше, чем 2^{k-1} .

Теперь, поскольку $\{x_n\}$ фундаментальна, в силу принципа **A3** получаем, что она имеет предел. Следовательно, **A3** влечет **A2**.

Итак, мы доказали, что **A1** \Rightarrow **A3** \Rightarrow **A2**. Тот факт, что из **A2** следует **A1**, мы доказали ранее. Отсюда следует, что все три принципа равносильны. \square

Мы уже говорили о том, что множество может быть ограничено сверху или снизу. Пусть X ограничено сверху числом a . Тогда число a называется его **верхней гранью**. Ясно, что если a — верхняя грань X , то верхними гранями будут также $a + 1, a + 10000, a + 0.0001$ и т. д. Все числа, бóльшие a , будут верхними гранями X .

Вспомним уравнение $x^2 = 2$. Пусть $X = \{x \mid x^2 < 2 \wedge x > 0\}$. Это есть интервал $(0; \sqrt{2})$. Мы помним, что в \mathbb{Q} нет числа $\sqrt{2}$, поэтому в рамках множества \mathbb{Q} верхними гранями X будут положительные рациональные числа r такие, что $r^2 > 2$. И среди этих чисел нет наименьшего в \mathbb{Q} . Но стоит нам выйти в поле алгебраических чисел, как мы уже можем использовать $\sqrt{2}$, и он будет не просто верхней гранью X , а наименьшей из всех верхних граней.

Наименьшая из верхних граней X называется **супремумом** X или **точной верхней гранью** X . Обозначение: $\sup X$.

Предлагаем читателю самостоятельно определить понятие точной нижней грани.

К точной верхней грани X мы можем подбираться, находясь внутри X , выбирая каждый раз все большее число из X , находящееся как можно ближе к его верхней грани. Например, у цепи вложенных отрезков есть множество левых границ, образующее неубывающую последовательность. При этом все правые границы отрезков будут верхними гранями для этой последовательности. И если в пределе получится множество, состоящее из одной точки (как в сюжете с ловлей точки A точками множества \mathbb{B}), то эта точка и будет точной верхней гранью для последовательности левых границ вложенных отрезков.

Мы снова видим некоторую связь между существованием супремума и аксиомой непрерывности в ее трех предыдущих формулировках. На самом деле, эта связь прямая и двусторонняя.

A4. Принцип точных граней

Формулировка A4. *Всякое непустое ограниченное сверху подмножество числовой прямой имеет точную верхнюю грань.*

A4'. *Всякое непустое ограниченное снизу подмножество числовой прямой имеет точную нижнюю грань.*

Теорема 15.7. *A4 и A4' равносильны принципам A1, A2, A3 с присоединенной к ним аксиомой Архимеда.*

Доказательство. Покажем, что **A4** влечет аксиому Архимеда. Предположим, что это не так и что существует число x , которое больше любого натурального числа n . Тогда $\mathbb{N} < x$, т. е. множество натуральных чисел ограничено сверху. В силу **A4** найдется $\alpha = \sup \mathbb{N}$. Рассмотрим $\beta = \alpha - 1$. С одной стороны, $\beta < \alpha$ в силу свойств упорядоченного поля. С другой стороны, не существует такого n , что $\beta \leq n$, т. к. отсюда следовало бы, что $\alpha \leq n + 1$, следовательно, β — верхняя грань \mathbb{N} , но тогда $\alpha \leq \beta$. Противоречие.

Докажем, что **A4** влечет **A2**. Пусть последовательность $\{x_n\}$ ограничена и не убывает. Тогда множество $X = \{x_n \mid n \in \mathbb{N}\}$ не пусто и ограничено сверху, и по аксиоме **A4** имеет точную верхнюю грань. Пусть $a = \sup X$. Ясно, что $x_n \leq a$, т. к. a является верхней гранью X . Предположим, что a не является пределом последовательности $\{x_n\}$, тогда существует $\varepsilon > 0$ такой, что для любого N существует $n > N$ такой, что $|x_n - a| \geq \varepsilon$. Так как $x_n \leq a$, то последнее неравенство можно переписать так: $x_n \leq a - \varepsilon$. Таким образом, выбирая все большие N , мы построим подпоследовательность x_{n_N} такую, что $x_{n_N} \leq a - \varepsilon$. Но в силу неубывания $\{x_n\}$ получаем, что $x_n \leq a - \varepsilon$ для любого n . Тогда число $a - \varepsilon$ также будет верхней гранью X в противоречии с тем, что $a = \sup X$. Следовательно, $a = \lim_{n \rightarrow \infty} x_n$.

Докажем, что из аксиомы Архимеда и **A1** следует **A4**. Пусть X — непустое ограниченное сверху множество. Пусть $X < a$. Возьмем какую-нибудь точку $x \in X$ и рассмотрим отрезок $[x; a]$. Разделим его пополам на два отрезка $[x; (x+a)/2]$ и $[(x+a)/2; a]$. Если в правом отрезке нет точек X , то в качестве $[a_1; b_1]$ выберем левый отрезок (в нем точки X точно есть), иначе — правый. Затем к отрезку $[a_1; b_1]$ применим ту же процедуру: разделим его пополам и выберем в качестве $[a_2; b_2]$ правую половину, если в ней есть точки X , и левую — в противном случае. И так далее. На каждом шаге в получаемом отрезке всегда есть точки X , но справа от b_n точек X нет.

В итоге мы получаем последовательность вложенных отрезков, длина которых стремится к нулю. В силу принципа **A1** пересечение $\bigcap_n [a_n; b_n]$ не пусто и, более того, состоит из единственной точки c (для единственности нам требуется аксиома Архимеда).

Точка c является верхней гранью X . Действительно, если это не так, то существует $x' \in X$ такой, что $c < x'$. Тогда выберем отрезок $[a_n; b_n]$ с таким номером n , что его длина меньше $x' - c$. Получим: $c \in [a_n; b_n]$ и при этом $b_n < x'$, а это противоречит выбору отрезков при их построении (справа от них нет точек X).

Число $\sup X$ существует и равно c . Действительно, если это не так, то существует меньшая, чем c , верхняя грань X . Пусть $X \leq c' < c$, тогда снова выберем отрезок $[a_n; b_n]$ с таким номером n , что его длина меньше $c - c'$. Получим: $c \in [a_n; b_n]$ и $c' < a_n$, но на отрезке $[a_n; b_n]$ есть точки X по построению этих отрезков, следовательно, найдется точка $x' \in X$ такая, что $x' \geq a_n > c'$, а это противоречит тому, что c' — верхняя грань X .

Итого получаем: **A4** \Rightarrow **A2** и **A1** \Rightarrow **A4**. Поскольку **A1** равносильно **A2** и **A3**, мы получаем, что все четыре аксиомы **A1**, **A2**, **A3**, **A4** эквивалентны (в предположении аксиомы Архимеда).

Для **A4'** можно провести аналогичное доказательство, а можно заметить, что если множество X не пусто и ограничено снизу, то множество $-X = \{-x \mid x \in X\}$ не пусто и ограничено сверху. При этом $\sup X = -\inf(-X)$. Поэтому равносильность **A4'** и **A4** устанавливается простой сменой знака. \square

Заметим, что существует некоторая двойственность между верхними и нижними гранями. Так, если взять некоторое непустое множество X , ограниченное сверху, то оно само будет множеством нижних граней (не обязательно всех) для множества Y всех своих верхних граней. При этом окажется, что $\sup X = \inf Y$. Действительно, пусть $c = \sup X$. Тогда по определению $c = \min Y$ и, следовательно, c есть нижняя грань Y . В то же время не существует нижней грани Y , которая была бы больше $\min Y$, следовательно, c есть наибольшая нижняя грань Y , т. е. $c = \inf Y$.

Аналогичная ситуация и с ограниченным снизу множеством. Возникает желание разбить всю числовую прямую на два луча — левый и правый — так, чтобы левый был множеством нижних граней для правого, и наоборот.

Пусть $\langle L, < \rangle$ — линейно упорядоченное множество. Пара $\langle X, Y \rangle$ непустых подмножеств множества L таких, что $X \cup Y = L$, $X \cap Y = \emptyset$ и $X < Y$, называется **сечением**.⁶

Ранее мы уже видели такое разбиение. Оно представляло собой два интервала рациональных чисел: $(-\infty; \sqrt{2}) \cap \mathbb{Q}$ и $(\sqrt{2}; +\infty) \cap \mathbb{Q}$. Действительно, их объединение равно \mathbb{Q} , пересечение пусто, и левый интервал меньше правого.

В случае \mathbb{Q} сечение может состоять из двух интервалов, т. е. таких множеств, что верхнее не имеет минимума, а нижнее не имеет максимума, и при этом между ними ничего нет. И это говорит нам о том, что в \mathbb{Q} имеются дырки. В случае \mathbb{A} дырки найти сложнее, но тоже можно. Разбиение на интервалы $(-\infty; \pi) \cap \mathbb{A}$ и $(\pi; +\infty) \cap \mathbb{A}$ доставляет

⁶ Иногда определяется только нижний класс, и он называется сечением.

такой пример, поскольку число π не является алгебраическим числом, т. е. *трансцендентно*.⁷

Еще один подход к определению непрерывности числовой прямой заключается в том, чтобы исключить такие дырки.

А5. Принцип дедекиндовых сечений

Формулировка А5. Если $\langle X, Y \rangle$ — сечение числовой прямой, то существует точка z такая, что $X \leq z \leq Y$.

Заметим, что здесь речь идет не о сечении \mathbb{Q} , а о сечении числовой прямой вещественных чисел. При этом точка z с необходимостью попадает либо в верхний, либо в нижний класс разбиения (ей просто деваться некуда). Таким образом всякое сечение числовой прямой должно обладать свойством: либо у верхнего класса есть минимум, либо у нижнего класса есть максимум (но не одновременно). Сечение с таким свойством называется **дедекиндовым**, а принцип **А5**, таким образом, утверждает, что всякое сечение числовой прямой является дедекиндовым.

Как и прежде, покажем, что имеет место равносильность принципа дедекиндовых сечений предыдущим четырем принципам непрерывности.

Теорема 15.8. Принципы **А4** и **А5** эквивалентны.

Отметим, что в данной теореме не требуется присоединение аксиомы Архимеда.

Доказательство. Покажем, что из **А4** следует **А5**. Пусть два подмножества $X, Y \subseteq \mathbb{R}$ таковы, что: $X, Y \neq \emptyset$, $X \cup Y = \mathbb{R}$, $X \cap Y = \emptyset$ и $X < Y$. Тогда очевидно, что X ограничено сверху элементами Y , а значит, существует $z = \sup X$. Легко видеть, что $X \leq z \leq Y$, и **А5** выполняется.

Докажем, что **А5** \Rightarrow **А4**. Пусть X — непустое ограниченное сверху подмножество \mathbb{R} . Необходимо найти его \sup .

Обозначим за Y множество всех верхних граней X , а за X' — дополнение к Y , т. е. $X' = \mathbb{R} \setminus Y$. Нетрудно видеть, что $X' < Y$ (т. к. если $y < x$ и $y \in Y$, то $x \in Y$), $X' \neq \emptyset$ (т. к. в нем есть точки $x' < x$, где $x \in X$), $Y \neq \emptyset$ (т. к. X ограничено), следовательно, пара $\langle X', Y \rangle$ является сечением. Тогда в силу **А5** существует точка z такая, что $X' \leq z \leq Y$.

Ясно, что $z \geq X$, т. к. $X \subseteq X'$. Тогда z является верхней гранью X , т. е. $z \in Y$. Но тогда $z = \min Y$ и, тем самым, $z = \sup X$. \square

Итак, мы получили, что при выполнении аксиомы Архимеда все пять форм аксиомы непрерывности равносильны друг другу. При этом мы

⁷ Этот факт был доказан в 1882 году Фердинандом фон Линдеманном.

§ 4. Созидание иррациональных чисел

Последними словами уже достаточно ясно указывается, каким образом разрывная область R рациональных чисел должна быть дополнена до превращения ее в непрерывную. Как это поставлено было на вид в § 1 (III), каждое рациональное число a производит разложение системы R на два класса A_1 и A_2 такого рода, что каждое число a_1 первого класса меньше каждого числа a_2 второго класса. Число a представляет либо наибольшее число класса A_1 , либо наименьшее число класса A_2 . Если теперь дано какое-либо подразделение системы R на два класса A_1, A_2 , обладающее только тем характерным свойством, что каждое число a_1 из A_1 меньше каждого числа a_2 из A_2 , то для краткости мы будем называть такое подразделение *сечением* и будем его обозначать через (A_1, A_2) . Мы можем тогда сказать, что каждое число a производит одно или, собственно, два сечения, на которые мы, однако, не будем смотреть, как на существенно различные*); это сечение имеет *кроме того* то свойство, что либо между числами первого класса есть наибольшее, либо между числами второго класса существует наименьшее. И наоборот, если сечение обладает и этим свойством, то оно производится этим наибольшим или наименьшим числом.

Рис. 15.5. Цитата из работы О. Дедекинда «Непрерывность и иррациональные числа» в пер. Шатуновского, 1923.

установили также, что **A5** и **A4**, будучи эквивалентны друг другу, влекут аксиому Архимеда.

Замечание для более глубокого изучения: существуют числовые системы, в которых **a)** выполняются все аксиомы упорядоченного поля, **б)** содержатся вещественные числа, **в)** но не выполняется аксиома Архимеда. Это так называемые гипердействительные числа или сюрреальные числа Конвея. Можно доказать, что в гипердействительных числах всякая фундаментальная последовательность является стационарной (т. е. принимает одно и то же значение начиная с некоторого номера), а значит, имеет предел, т. е. выполняется принцип **A3**. Таким образом, гипердействительные числа предоставляют нам пример числовой системы, где принцип **A3** логически слабее принципов **A4** и **A5**. В такой системе вновь появляются «дырки» в линейном порядке (например, между бесконечно малыми числами и вещественными положительными).

Наконец, дадим самое главное определение данной главы (несмотря на то, что мы уже давно им пользуемся). Числовая прямая с операциями сложения, умножения и линейным порядком, согласованным с данными операциями, удовлетворяющая аксиоме непрерывности в форме **A4** (**A5** или **A1–A3** плюс аксиома Архимеда), называется **вещественной (действительной) прямой** и обозначается за \mathbb{R} .

15.4. Модели действительных чисел

В предыдущем разделе было сформулировано пять вариантов аксиомы непрерывности, которая необходима для того, чтобы узаконить действительные числа как непрерывную числовую структуру. Аксиома непрерывности не оставляет дыр на числовой прямой, поскольку вводит в обращение все числа, к которым можно обратиться с помощью счетной последовательности рациональных чисел.

Тем не менее одной лишь аксиомы недостаточно, чтобы действительные числа имели право на существование. Необходимо убедиться в том, что их можно непротиворечиво сконструировать. Поэтому здесь мы рассмотрим несколько подходов к построению действительных чисел.

Дедекиндовы сечения

Первый подход связан непосредственно с тем, чем мы закончили предыдущий раздел — с сечениями. А именно рассмотрим все сечения множества рациональных чисел, причем только такие, у которых нижний класс не содержит наибольшего элемента (т. е. если между классами существует граница, то она отнесена к верхнему классу). И соберем в множество R нижние классы всех таких сечений \mathbb{Q} . Таким образом, R состоит из подмножеств $X \subset \mathbb{Q}$ таких, что

$$X \neq \emptyset, \quad X \neq \mathbb{Q}, \quad \forall r, q \in \mathbb{Q} (q \in X) \wedge (r < q) \rightarrow (r \in X), \quad \nexists \max X.$$

Иначе говоря, R — это множество лучей из рациональных чисел, направленных в $-\infty$.

На множестве R введем операцию сложения: пусть $\alpha, \beta \in R$, тогда положим

$$\alpha + \beta = \{x + y \mid x \in \alpha, y \in \beta\},$$

т. е. просто сложим их по Минковскому (см. рис. 15.6).

Можно проверить, что R с такой операцией сложения является абелевой группой, т. е. сложение ассоциативно, коммутативно, имеется

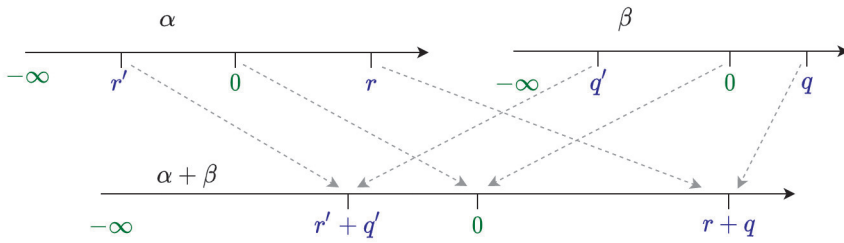


Рис. 15.6. Сложение сечений.

нейтральный элемент $0_R = \{x \in \mathbb{Q} \mid x < 0_{\mathbb{Q}}\}$, где $0_{\mathbb{Q}}$ — ноль в системе рациональных чисел, и для каждого α имеется противоположный элемент $-\alpha = \{x \in \mathbb{Q} \mid (-x > \alpha) \wedge (-x \neq \min(\mathbb{Q} \setminus \alpha))\}$.

Здесь появляется первая тонкость определения. По сути, в качестве $-\alpha$ мы берем соответствующий ему верхний класс сечения и умножаем на -1 (в поле рациональных чисел). Однако верхний класс может иметь минимум, а элемент R не должен иметь максимума, поэтому мы подстраховываемся и выбрасываем из верхнего класса $\mathbb{Q} \setminus \alpha$ его минимум (если такой существует).

Достаточно легко определяется и порядок на R . Скажем, что $\alpha < \beta$, если $\alpha \subset \beta$ (как собственное подмножество).

Отсюда же следует согласованность сложения и порядка, поскольку сдвиг вверх или вниз интервалов $\alpha < \beta$ не меняет их вложенности.

Сложнее дело обстоит с умножением. Если мы попытаемся умножить $\alpha \cdot \beta$ по Минковскому, то произведение будет содержать сколь угодно большие числа (при перемножении двух чисел, сильно меньших 0 в поле \mathbb{Q}). Поэтому сначала определяется произведение положительных чисел: пусть $\alpha > 0_R$ и $\beta > 0_R$, тогда (см. рис. 15.7)

$$\alpha \cdot \beta = \{xy \mid (x \in \alpha) \wedge (x \geq 0) \wedge (y \in \beta) \wedge (y \geq 0)\} \cup \mathbb{Q}^-,$$

где \mathbb{Q}^- — все отрицательные рациональные числа, т. е. $\mathbb{Q}^- = 0_R$.

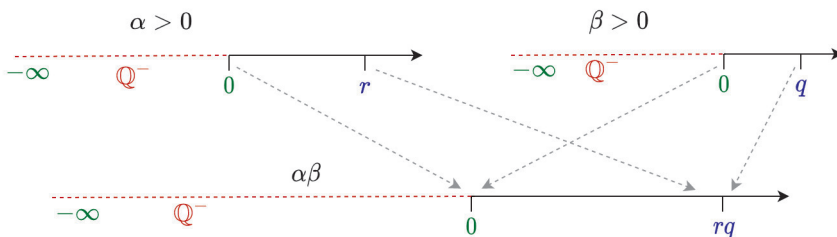


Рис. 15.7. Умножение сечений.

Далее просто полагаем, что

$$\alpha \cdot \beta = \begin{cases} 0, & \text{если } (\alpha = 0_R) \vee (\beta = 0_R), \\ -((-\alpha) \cdot \beta), & \text{если } (\alpha < 0_R) \wedge (\beta > 0_R), \\ -(\alpha \cdot (-\beta)), & \text{если } (\alpha > 0_R) \wedge (\beta < 0_R), \\ (-\alpha) \cdot (-\beta), & \text{если } (\alpha < 0_R) \wedge (\beta < 0_R). \end{cases}$$

Можно проверить, что такая операция умножения на R ассоциативна, коммутативна, имеет единицу $1_R = \{x \in \mathbb{Q} \mid x < 1_{\mathbb{Q}}\}$, где $1_{\mathbb{Q}}$ — единица в системе рациональных чисел, и для каждого положительного $\alpha \in R$ имеется обратный по умножению

$$1/\alpha = \{x \in \mathbb{Q} \mid \exists y \in \mathbb{Q} : y > \alpha : x < 1/y\},$$

а обратный к отрицательному числу определяется сменой знака: $1/\alpha = -(1/(-\alpha))$, если $\alpha < 0$.

Кроме того, можно доказать, что операции сложения и умножения удовлетворяют дистрибутивному закону, а также что умножение согласовано с порядком.

Таким образом, R с указанными операциями и порядком есть упорядоченное поле. В целях экономии места и времени мы не стали утомлять читателя суровыми выкладками по проверке этого факта, а только лишь привели само построение системы чисел. Тем не менее мы настоятельно рекомендуем проработать тщательные доказательства свойств определенных выше порядка и операций сложения и умножения, например, с помощью учебника У. Рудина [19].

Нам остается доказать то, ради чего мы строили это поле, — его непрерывность.

Для этого воспользуемся формулировкой аксиомы непрерывности в виде **A4**. Возьмем непустое ограниченное сверху множество $X \subset R$. Положим $\alpha = \bigcup X$, т. е. мы включаем в множество α все рациональные числа, входящие во все элементы множества X . Во-первых, $\alpha \in R$. Действительно, $\alpha \neq \emptyset$, т. к. не пусты все элементы X . Если $y < x$ и $x \in \alpha$, то $y \in \alpha$, поскольку для некоторого $\beta \in X$ имеем $x \in \beta$, откуда следует, что $y \in \beta$, а значит, $y \in \alpha$. Наконец, в α нет максимума, т. к. если бы существовал $r = \max \alpha$, то для некоторого $\beta \in X$ мы бы получили $r \in \beta$, но т. к. $r = \max \alpha$, он же будет максимумом β , а это противоречит определению элементов R . Итак, $\alpha \in R$.

Легко видеть, что для любого $\beta \in X$ имеет место неравенство $\beta \leq \alpha$, т. е. α является верхней гранью X . Действительно, ведь $\beta \subseteq \alpha$, а значит, либо $\beta < \alpha$, либо $\beta = \alpha$ по определению отношения $<$ на множестве R (см. выше).

В то же время любая верхняя грань X окажется не меньше α . Предположим, что это не так и что существует $\gamma \in R$ такой, что $\beta \leq \gamma < \alpha$ для всех $\beta \in X$. Поскольку $\gamma \in \alpha$, множество $\alpha \setminus \beta \neq \emptyset$, следовательно, существует $r \in \alpha$ такой, что $\gamma < r$, но тогда существует $\beta \in X$ такой, что $r \in \beta$ и $\gamma < r$. Противоречие с выбором γ .

Таким образом,

$$\sup X = \bigcup X.$$

Итак, множество R , построенное из подмножеств \mathbb{Q} специального вида, с заданными на нем операциями и порядком, является непрерывным упорядоченным полем, т. е. полем действительных чисел \mathbb{R} .

Двоичные дроби

Следующий подход к моделированию \mathbb{R} прямо связан с нашим множеством \mathbb{B} . Вспомним сюжет о поимке точки A в сеть точек множества \mathbb{B} , т. е. рациональных точек со знаменателями вида 2^n . Мы выстраивали убывающую цепь отрезков, концы которых находятся в множестве \mathbb{B}_n , т. е. имеют вид $[k/2^n; (k+1)/2^n]$. Длина этих отрезков быстро стремится к нулю, а по аксиоме непрерывности в форме **A1** существует непустой предел этой цепи отрезков, который состоит из единственной точки A .

Почему бы тогда не обращаться к точке A с помощью этой последовательности? Точнее, мы определим некоторый условный код, который позволит нам однозначно поймать точку A и никакую другую.

Вспомним алгоритм построения этих отрезков. Сначала мы выбираем целочисленный полуинтервал $[k; k+1)$, в котором лежит A . Запишем число k как стартовое число кода (в контексте дальнейшего построения можно считать, что число k записано в виде конечной двоичной последовательности цифр).

Затем мы делим этот интервал ровно пополам: $[k; k+1) = [k+1/2; k+1/2) \cup [k+1/2; k+1)$. Точка A лежит либо в левом полуинтервале, либо в правом. Следующим числом кода запишем 0, если A лежит в левом интервале, и 1 — если в правом. Перейдем к соответствующему интервалу.

Снова поделим его пополам и произведем аналогичную процедуру записи следующей цифры кода. И так будем продолжать до бесконечности.

В итоге у нас получится код, стартующий с некоторого целого числа (записанного в двоичной системе), после которого идет бесконечная (счетная) цепочка нулей и единиц. Этот код однозначно формируется по заданной точке A (поскольку мы всегда работаем с полуинтервалами, а они всякий раз выбираются единственным способом).

Особенностью данного кода является то, что в нем нет хвоста единиц, т. е. когда, начиная с некоторой позиции, все цифры равны 1.

Это объясняется очень просто: если есть хвост единиц, то, начиная с какого-то шага, алгоритм всегда выбирал правый интервал, в результате чего хвост последовательности вложенных отрезков имел бы вид

$$[r/2^m - 1/2; r/2^m] \supset [r/2^m - 1/4; r/2^m] \supset [r/2^m - 1/8; r/2^m] \supset \dots,$$

где r и m не зависят от n . Но пределом такой цепи будет, очевидно, множество $\{r/2^m\}$, т.е. такая цепь вложенных отрезков должна сходиться к числу $r/2^m$. Но тогда алгоритм еще на предыдущем $m - 1$ шаге выберет полуинтервал, лежащий справа от этой точки, в результате чего отрезками, сходящимися к точке $r/2^m$, будут такие

$$[r/2^m; r/2^m + 1/2] \supset [r/2^m; r/2^m + 1/4] \supset [r/2^m; r/2^m + 1/8] \supset \dots,$$

и мы увидим не хвост единиц, а хвост нулей! Правда, перед ним будет стоять единица.

То есть кодовая последовательность вида $k, d_1 \dots d_s 0111111 \dots$, где $d_i \in \{0, 1\}$, невозможна. Вместо нее будет последовательность вида $k, d_1 \dots d_s 1000000 \dots$.⁸

Верно и обратное. По такому коду (без хвоста единиц) можно однозначно восстановить закодированную им точку A .

Поэтому между точками вещественной прямой \mathbb{R} и кодами указанного вида существует взаимно однозначное соответствие (биекция), и это значит, что за модель \mathbb{R} может быть принято множество всех таких цепочек.

Точнее, положим

$$R = \{ \langle k, f \rangle \mid k \in \mathbb{Z}, f : \mathbb{N} \rightarrow \{0, 1\}, \forall n \exists m > n (f(m) = 0) \}.$$

Здесь условие $\forall n \exists m > n (f(m) = 0)$ как раз и означает, что в последовательности f нет хвоста единиц (ноль встречается бесконечно часто).

Сложности в такой модели \mathbb{R} начинаются, когда мы хотим определить операции сложения и умножения. Порядок же определяется предельно просто. Пусть даны две последовательности $\langle k, f \rangle$ и $\langle k', f' \rangle$. Отношение порядка между ними основано на сравнении первого расхождения кодов. Если $k < k'$, то $\langle k, f \rangle < \langle k', f' \rangle$. Если $k = k'$, смотрим $f(0)$ и $f'(0)$. Если $f(0) < f'(0)$, то $\langle k, f \rangle < \langle k', f' \rangle$. Если они равны, то переходим к следующей цифре кода и т.д. Если не нашлось ни одного расхождения в коде, то $\langle k, f \rangle$ и $\langle k', f' \rangle$ равны.

⁸ Вместо длинного выражения $d_1 \dots d_s$, где $d_i \in \{0, 1\}$, удобнее воспользоваться языком *регулярных выражений* и написать $[01]^*$, что и будет означать конечную последовательность произвольной длины (в том числе нулевой) из нулей и единиц.

Мы не будем здесь заниматься рекурсивным определением операций сложения и умножения. Скажем только, что его можно задать, используя арифметику двоично-рациональных чисел множества \mathbb{B} , и во многом он напоминает определение операций в следующей модели \mathbb{R} , основанной на классах эквивалентных последовательностей. Действительно, ведь двоичный код задает не только алгоритм вычисления вложенных отрезков, он задает последовательность их левых границ, которая сходится к адресуемому числу A . А это — последовательность двоично-рациональных чисел, которые мы умеем складывать и умножать, не выходя за рамки множества \mathbb{B} .

Более того, число, которое закодировано парой $\langle k, f \rangle$, можно записать в виде бесконечной суммы

$$A = k + \sum_{n=0}^{\infty} \frac{f(n)}{2^n},$$

поскольку переход к правому отрезку на n -ом шаге в описанном алгоритме означает добавление $1/2^n$ к левой границе предыдущего отрезка, а переход к левому отрезку означает добавление $0/2^n$. Так что любое действительное число можно записать в виде разложения по степеням 2, а это и есть не что иное как запись произвольного числа в двоичной системе счисления. Число k при этом можно тоже записать в двоичном коде, и тогда код произвольного числа имеет следующий вид: конечный набор нулей и единиц (начиная с 1), затем стоит точка, затем идет бесконечный набор нулей и единиц (без хвоста единиц).

Завершая описание двоичной модели, скажем, что в качестве основания можно выбрать любое натуральное число $d > 1$. Например, если мы хотим получить троичные последовательности, нам следует модифицировать алгоритм разбиения на отрезки следующим образом: интервал $[k; k + 1)$ делить на три части $[k; k + 1/3)$, $[k + 1/3; k + 2/3)$ и $[k + 2/3; k + 1)$, и далее к каждому следующему интервалу применять аналогичное деление на 3 части. В результате для записи кода мы будем выбирать 0, если A оказалась в левом интервале, 1 — если в среднем, 2 — если в правом. Получится код из цифр 0, 1, 2, причем без хвоста двоек. Все рассуждения здесь полностью аналогичны предыдущему.

Мы можем использовать число $d = 10$ в качестве основания, и каждое действительное число записывать кодом из цифр $0 \dots 9$ без хвоста девяток (хвост девяток всегда можно заменить хвостом нулей, увеличив стоящую перед девятками цифру на 1). Получится стандартная запись числа в виде десятичной дроби.

Двоичное представление вещественных чисел открывает нам возможность оценить мощность множества \mathbb{R} , а точнее, полуинтервала $[0; 1)$. Всякое число $\alpha \in [0; 1)$ имеет код, заданный функцией $f : \mathbb{N} \rightarrow$

$\{0, 1\}$, т. е. мощность множества вещественных чисел в полуинтервале $[0; 1)$ равна мощности множества таких функций без хвоста единиц.

С другой стороны, всякая функция вида $f : \mathbb{N} \rightarrow \{0, 1\}$ взаимно однозначно задает некоторое подмножество в \mathbb{N} . Нужно в этом подмножестве собрать только те элементы, на которых $f = 1$.

Это значит, что мы можем построить инъекцию $F : [0; 1) \rightarrow \mathcal{P}(\mathbb{N})$.

Теперь по произвольному подмножеству \mathbb{N} построим функцию $f : \mathbb{N} \rightarrow \{0, 1\}$. Такая функция может содержать хвост единиц. Но теперь мы этот код будем рассматривать не как двоичный, а как троичный! У нас гарантированно не будет хвоста двоек, а значит, мы инъективно построим какие-то числа в $[0; 1)$ (точнее, даже в $[0; 2/3)$, причем с очень многими дырами). Тем самым, мы имеем инъекцию $G : \mathcal{P}(\mathbb{N}) \rightarrow [0; 1)$.

Окончательно, согласно теореме Кантора – Бернштейна, мы получаем равномощность множеств $[0; 1)$ и $\mathcal{P}(\mathbb{N})$. То есть интервал $[0; 1)$ имеет мощность континуума!

Чтобы перейти к \mathbb{R} , нужно сначала научиться строить биекцию между интервалом и полуинтервалом.

Легко видеть, что функция (см. рис. 15.8)

$$f(x) = \begin{cases} 1/2, & x = 0, \\ x/2, & x = 1/2^n, n = 1, 2, \dots, \\ x, & \text{иначе,} \end{cases}$$

биективно переводит $[0; 1)$ в $(0; 1)$. Все точки вида $1/2^n$ сдвигаются вниз на 1 шаг, а ноль переходит в точку $1/2$.

Далее, функция $g(x) = 2x - 1$, очевидно, биективно переводит $(0; 1)$ в $(-1; 1)$.

Наконец, функция

$$h(x) = \begin{cases} \frac{x}{1-x}, & 0 \leq x < 1, \\ \frac{x}{1+x}, & -1 < x < 0 \end{cases}$$

биективно переводит интервал $(-1; 1)$ в \mathbb{R} . График функции $h(x)$ представлен на рисунке справа. Таким образом, композиция $h(g(f(x)))$ биективно переводит $[0; 1)$ в \mathbb{R} . Следовательно, множество вещественных чисел имеет мощность континуума.

Эквивалентные последовательности

Рассмотрим еще один способ конструирования множества \mathbb{R} .

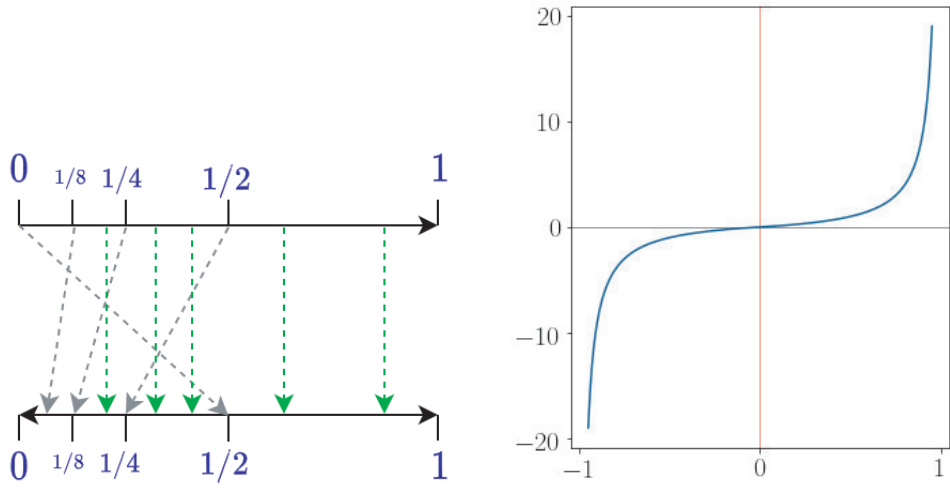


Рис. 15.8. Биекции $[0; 1) \leftrightarrow (0; 1)$ и $(-1; 1) \leftrightarrow \mathbb{R}$.

Обозначим за \mathcal{F} множество всех фундаментальных последовательностей рациональных чисел. На множестве \mathcal{F} введем отношение \sim следующим образом:

$$q \sim r \Leftrightarrow \lim_{n \rightarrow \infty} (q_n - r_n) = 0,$$

где $q = \{q_n\}$ и $r = \{r_n\}$ — произвольные элементы множества \mathcal{F} .

Вспоминая аксиому непрерывности в форме **A3**, мы понимаем, что если $q, r \in \mathcal{F}$ и $q \sim r$, то эти последовательности имеют пределы, причем эти пределы совпадают. Отсюда легко видеть, что отношение \sim является отношением эквивалентности, а значит, мы можем разбить \mathcal{F} на классы эквивалентных последовательностей, т.е. построить фактормножество $R = \mathcal{F} / \sim$, которое мы и объявим множеством действительных чисел.

После чего мы должны ввести соответствующие операции и отношение сравнения.

Сложение классов эквивалентности вводится с помощью их представителей по правилу:

$$[q] + [r] = [q + r].$$

Необходимо лишь доказать, что если $q \sim q'$ и $r \sim r'$, то $q + r \sim q' + r'$. Это легко заметить из следующего неравенства

$$|(q + r)_n - (q' + r')_n| \leq |q_n - q'_n| + |r_n - r'_n|,$$

поскольку два модуля справа стремятся к нулю.

Аналогично вводится умножение:

$$[q] \cdot [r] = [qr].$$

Для доказательства корректности определения заметим, что если $q \sim q'$ и $r \sim r'$, то

$$|q_n r_n - q'_n r'_n| \leq |q_n| |r_n - r'_n| + |r'_n| |q_n - q'_n|.$$

Здесь справа стоят слагаемые, в каждом из которых ограниченная величина (в силу фундаментальности) умножается на величину, стремящуюся к нулю. Так что и все вместе стремится к нулю.

Наконец, определим сравнение:

$$[q] < [r], \text{ если } \exists \varepsilon > 0 \exists N (\forall n > N)(q_n < r_n - \varepsilon),$$

т. е. для почти всех индексов разность $r_n - q_n$ отделена от нуля положительным числом ε (оно может быть очень маленьким, но не нулевым).

Такое определение корректно. Действительно, пусть $q' \sim q$ и $r' \sim r$. Пусть также $[q] < [r]$ в том смысле, как определено выше. Пусть $\varepsilon > 0$, тогда существует такой номер N (общий для всех трех случаев выбираем как максимум), что для всех $n > N$

$$q_n < r_n - \varepsilon, \quad |q_n - q'_n| < \varepsilon/3, \quad |r_n - r'_n| < \varepsilon/3.$$

Отсюда получаем, что

$$q'_n = q_n + (q'_n - q_n) < r_n - \varepsilon + \varepsilon/3 = r'_n + (r_n - r'_n) - 2\varepsilon/3 < r'_n - \varepsilon/3.$$

Откуда следует, что $[q'] < [r']$, т. е. порядок на классах эквивалентных последовательностей определен корректно.

Итак, мы видим, что при определении \mathbb{R} через классы эквивалентных фундаментальных последовательностей и операции, и упорядочение чисел просто переносятся с рациональных чисел, образующих соответствующие последовательности. Главная задача тут — показать корректность такого определения. Кроме того, здесь мы активно пользуемся понятием предела.

Выше мы рассмотрели три модели \mathbb{R} :

- Модель дедекиндовых сечений.
- Модель двоичных (в общем случае: d -ичных) дробей.
- Модель классов фундаментальных последовательностей.

Попутно мы установили раномощность \mathbb{R} и множества всех подмножеств \mathbb{N} .

Возникает вопрос: если \mathbb{R} такое большое множество, а \mathbb{Q} такое маленькое (по мощности), есть ли какие-то множества, имеющие промежуточные мощности между счетной и континуумом? Ответ на этот вопрос дали два человека: К. Гёдель (1936) и П. Коэн (1963). Первый доказал, что отсутствие промежуточных мощностей не противоречит аксиоматике теории множеств, второй — что существование таких мощностей также не противоречит аксиоматике теории множеств. Таким образом, мы оказываемся в ситуации пятого постулата Евклида, когда можем принимать или отвергать континуум-гипотезу Кантора (именно так называется утверждение о том, что между счетной мощностью и континуумом нет промежуточных мощностей), не опасаясь получить противоречие.

Упражнения

Обязательные упражнения

15.1° Рассмотрим соответствие множества людей и множества всех возрастов (целых лет). В каком направлении соответствие между ними является функцией?

15.2° Рассмотрим соответствие множества людей и банковских счетов. Является ли это соответствие функцией хоть в каком-то направлении?

15.3° Доказать, что функция $f(x) = 3x - 7$ биективно отображает \mathbb{R} в \mathbb{R} .

15.4° Доказать, что функция $g(x) = x^2 + 3x - 6$ действует инъективно при $x \in [-3/2, +\infty)$.

15.5° Доказать, что прообраз Y — это образ Y для обратного отношения.

15.6° Пусть $R \subseteq A \times B$ и $X_1, X_2 \subseteq A$, $Y_1, Y_2 \subseteq B$. Проверьте, что:

a) $R[X_1 \cup X_2] = R[X_1] \cup R[X_2]$; b) $R^{-1}[Y_1 \cup Y_2] = R^{-1}[Y_1] \cup R^{-1}[Y_2]$;

c) $R[X_1 \cap X_2] \subseteq R[X_1] \cap R[X_2]$; d) $R^{-1}[Y_1 \cap Y_2] \subseteq R^{-1}[Y_1] \cap R^{-1}[Y_2]$.

Что изменится, если R будет функцией? Инъекцией? Биекцией?

15.7° Привести примеры, когда $f[X_1 \cap X_2] \neq f[X_1] \cap f[X_2]$, если $f : A \rightarrow B$.

15.8° Что можно сказать о a) равенстве множеств X и $f^{-1}[f[X]]$; b) равенстве множеств Y и $f[f^{-1}[Y]]$, если такие множества определены?

15.9° Докажите, что равномощность является отношением эквивалентности, то есть a) $A \leftrightarrow A$; b) если $A \leftrightarrow B$, то $B \leftrightarrow A$; c) если $A \leftrightarrow B$ и $B \leftrightarrow C$, то $A \leftrightarrow C$.

15.10° Докажите, что для любых множеств X и Y выполнено $X \times Y \leftrightarrow Y \times X$.

15.11° Верно ли, что если $A \leftrightarrow B$ и $C \leftrightarrow D$, то а) $A \times C \leftrightarrow B \times D$; б) $A \cup C \leftrightarrow B \cup D$; в) $A \cap C \leftrightarrow B \cap D$?

15.12° Доказать, что **A2** равносильно следующим утверждениям: а) всякая возрастающая ограниченная сверху последовательность имеет предел; б) всякая убывающая ограниченная снизу последовательность имеет предел.

15.13° Доказать, что любая сходящаяся последовательность фундаментальна.

15.14° Доказать, что **A4** эквивалентно следующему утверждению: (A4') всякое непустое ограниченное снизу множество имеет точную нижнюю грань.

15.15° Доказать, что а) супремум множества, если он существует, есть минимум верхних граней этого множества; б) инфимум множества, если он существует, есть максимум нижних граней этого множества.

15.16° Число $x \in (0; 1)$ назовем **вычислимым**, если есть конечный алгоритм, который позволяет для каждого $n \in \mathbb{N}$ определить n -ый знак после запятой в десятичной записи числа x .

а) Докажите, что множество вычислимых чисел из $(0; 1)$ счетно.

б) Выпишем десятичные записи всех вычислимых чисел в таблицу и диагональным методом построим вычислимое число, не входящее в таблицу. (Это можно сделать, написав алгоритм, который перебирает алгоритмы, дающие вычислимые числа, и меняет у n -го числа n -ю цифру на какую-нибудь другую.) Объясните это противоречие.

15.17° Найдите ограниченную последовательность, у которой а) есть и наибольший, и наименьший члены; б) есть наибольший, но нет наименьшего; в) есть наименьший, но нет наибольшего; г) нет ни наименьшего, ни наибольшего.

15.18° Верно ли, что а) сумма; б) разность; в) произведение; г) отношение ограниченных последовательностей — тоже обязательно ограниченная последовательность?

15.19° Докажите, что элемент c есть $\sup M$ тогда и только тогда, когда выполнены два условия: а) для всех $x \in M$ верно, что $x \leq c$; б) для любого элемента $c_1 < c$ найдется такой $x \in M$, что $x > c_1$.

15.20° Найдите $\sup M$ и $\inf M$, если $M = \{a^2 + 2^a \mid -5 < a \leq 5\}$.

15.21° Может ли у множества быть несколько точных верхних (нижних) граней?

15.22° Пусть A и B — некоторые подмножества \mathbb{R} , и пусть известны $\sup A$ и $\sup B$. а) Найдите $\sup(A \cup B)$. б) Найдите $\sup(A + B)$. в) Найдите $\inf(A \cdot B)$, если A и B состоят из отрицательных элементов.

15.23° Пусть $-X = \{-x \mid x \in X\}$ и $X \neq \emptyset$. Доказать, что а) $\sup(-X) = -\inf X$; б) $\inf(-X) = -\sup X$.

15.24° Пусть $X, Y \neq \emptyset$. Доказать равенства: а) $\sup(X + Y) = \sup X + \sup Y$; б) $\inf(X + Y) = \inf X + \inf Y$.

15.25° Пусть $X, Y \neq \emptyset$ и $X \geq 0, Y \geq 0$. Доказать, что: а) $\sup(XY) = \sup X \sup Y$; б) $\inf(XY) = \inf X \inf Y$.

Дополнительные упражнения

15.26* Докажите, что $\mathcal{P}(A) \leftrightarrow \{0, 1\}^A$ (включая случай $A = \emptyset$).

15.27* Доказать, что если $\#A = n$, то множество B^A равномощно множеству $\underbrace{B \times \cdots \times B}_n$ раз.

15.28* Пусть $\#A = n$ и $\#B = k$. Сколько элементов в множествах A^B и B^A (включая случаи $n = 0$ и/или $k = 0$)?

15.29* Докажите, что следующие множества являются счетными: а) множество четных натуральных чисел; б) множество нечетных натуральных чисел; в) множество натуральных чисел без числа 2021; г) множество целых чисел \mathbb{Z} .

15.30* Докажите, что всякое подмножество счетного множества не более чем счетно.

15.31* Докажите, что если X счетно, а Y конечно и непусто, то $X \times Y$ счетно.

15.32* Докажите, что следующие множества являются счетными: а) объединение конечного числа счетных множеств; б) прямое произведение двух счетных множеств; в) объединение счетного числа различных конечных множеств; г) объединение счетного числа счетных множеств.

15.33* Докажите, что множество рациональных чисел \mathbb{Q} является счетным.

15.34* Докажите, что множество алгебраических чисел \mathbb{A} является счетным.

15.35* Докажите, что $\mathbb{N}^{\mathbb{N}}$ континуально. Указание: докажите сначала, что $\mathbb{N}^{\mathbb{N}} \approx 2^{\mathbb{N} \times \mathbb{N}}$, а затем примените биекцию $\mathbb{N} \leftrightarrow \mathbb{N} \times \mathbb{N}$.

15.36* Верно ли, что следующие множества являются счетными: а) множество всевозможных конечных последовательностей нулей и единиц; б) множество всевозможных русских «слов» (т.е. конечных последовательностей букв русского алфавита); в) множество конечных подмножеств множества \mathbb{N} ?

15.37* Докажите, что если $B \cap C = \emptyset$, то $A^B \times A^C \leftrightarrow A^{B \cup C}$.

15.38* Докажите, что если A бесконечно, а B не более чем счетно, то $A \cup B \leftrightarrow A$.

15.39* Рассмотрим таблицу из 0 и 1, бесконечную «вправо-вниз». Докажите, что найдется бесконечная последовательность из 0 и 1, которая не совпадает ни с одной из строк таблицы.

15.40* Докажите, что следующие множества попарно равномощны: а) отрезок $[0; 1]$; б) интервал $(0; 1)$; в) \mathbb{R} ; г) $\{0, 1\}^{\mathbb{N}}$ (т.е. множество бесконечных последовательностей 0 и 1).

15.41* Докажите счетность следующих множеств: а) $\mathbb{Z}[x]$ (множество многочленов от x с целыми коэффициентами); б) $\mathbb{Q}[x_1, x_2, \dots]$ (множество многочленов с рациональными коэффициентами от счетного числа переменных).

15.42* Пусть A , B и C — произвольные попарно непересекающиеся множества. Докажите, что:

а) $A^C \times B^C \leftrightarrow (A \times B)^C$; б) $(A^B)^C \leftrightarrow A^{B \times C}$.

15.43* Докажите, что следующие множества континуальны:

а) $(\{0, 1\}^{\mathbb{N}})^2$; б) $[0; 1]^2$; в) \mathbb{R}^2 ; г) $\mathbb{R}^{\mathbb{N}}$.

15.44* Докажите, что объединение континуального числа континуальных множеств континуально.

15.45* Докажите, что следующие множества не более чем счетны: а) произвольное множество попарно непересекающихся интервалов на прямой; б) произвольное множество попарно непересекающихся «восьмерок» на плоскости («восьмерка» — объединение двух касающихся внешним образом окружностей, одна восьмерка может находиться целиком внутри другой); в) произвольное множество попарно непересекающихся «букв Т» на плоскости (т.е. наборов из трех невырожденных отрезков, имеющих один общий конец и никаких других общих точек); г) множество точек разрыва произвольной монотонной функции.

Элементы математического анализа

Аннотация

Здесь мы познакомимся с базовой терминологией и методиками анализа функций вещественного переменного.

16.1. Оценки и пределы

Выше мы ввели общее понятие предела числовой последовательности: число a называется пределом последовательности $\{x_n\} \subset \mathbb{R}$, если для любого $\varepsilon > 0$ найдется такой номер N , что для всех $n > N$ имеем $|x_n - a| < \varepsilon$. Предел a в этом случае обозначается за

$$a = \lim_{n \rightarrow \infty} x_n.$$

Очень часто используется более простая нотация в виде $x_n \rightarrow a$, и если ясно из контекста, по какому параметру берется предел (в данном случае при $n \rightarrow \infty$), то больше ничего не дописывают. О подводных камнях такой нотации мы уже поговорили чуть выше.

Для отыскания пределов существует масса прямых и косвенных методов. Прямой, т. е. по определению, предполагает в явном виде выписать номер N , зависящий от переменной ε , либо же доказать существование такого номера, отправляясь от каких-то известных фактов о строении множества \mathbb{R} . Приведем простой пример. Как доказать, что $1/2^n \rightarrow 0$?

Возьмем произвольный $\varepsilon > 0$. Необходимо показать, что выбором достаточно большого n величину $1/2^n$ можно сделать меньше ε . Для начала заметим, что число $1/\varepsilon$ хоть и огромное (при маленьких ε), но все-таки конечное. А это значит, что существует натуральное $n > 1/\varepsilon$ (здесь снова работает аксиома Архимеда). А далее, как легко видеть, $2^n > n > 1/\varepsilon$ (то, что $2^n > n$, доказывается индукцией), откуда уже по арифметическим правилам следует, что $\varepsilon > 1/n > 1/2^n$. Следовательно, $1/2^n \rightarrow 0$. Попутно, кстати, мы доказали, что и $1/n \rightarrow 0$.

К косвенным методам можно отнести арифметику пределов.

1) Если $x_n \rightarrow a$ и $y_n \rightarrow b$, то $x_n + y_n \rightarrow a + b$.

Действительно, для любого $\varepsilon > 0$ найдутся номера N_1 и N_2 такие, что при $n > N_1$ имеем $|x_n - a| < \varepsilon/2$ (поскольку ε произвольный, почему бы не взять его половинку?) и при $n > N_2$ имеем $|y_n - b| < \varepsilon/2$. Далее, пользуясь неравенством треугольника для модуля, получаем, что

$$|x_n + y_n - (a + b)| \leq |x_n - a| + |y_n - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

при $n > N = \max(N_1, N_2)$.

2) Если $x_n \rightarrow a$ и $y_n \rightarrow b$, то $x_n y_n \rightarrow ab$.

Здесь — аналогичные рассуждения. Возьмем произвольное $\varepsilon > 0$ и найдем такое N , что для всех $n > N$ имеют место неравенства:

$$|x_n - a|, |y_n - b| < \frac{\varepsilon}{|a| + |b| + 1} \text{ и одновременно } |x_n - a| < 1,$$

где, опять же, деление ε на константу никак не ограничивает произвольности его выбора, а дальше

$$\begin{aligned} |x_n y_n - ab| &= |x_n y_n - x_n b + x_n b - ab| \leq |x_n| |y_n - b| + |b| |x_n - a| < \\ &< \frac{(|a| + 1)\varepsilon}{|a| + |b| + 1} + \frac{|b|\varepsilon}{|a| + |b| + 1} = \varepsilon. \end{aligned}$$

Тут весь фокус заключается в том, что ε для результирующей последовательности $x_n y_n$ выбирается произвольный, а ε из определения сходимости x_n и y_n выбирается уже на основе исходного ε так, чтобы потом в итоге получилось то, что требует определение предела.

3) Если $x_n \rightarrow a$ и $y_n \rightarrow b \neq 0$, то $x_n/y_n \rightarrow a/b$.

Действительно, пусть $\varepsilon > 0$, тогда найдем такое N , что при всех $n > N$ имеем $|x_n - a|, |y_n - b| < \varepsilon$ и, кроме того, $|y_n - b| < |b|/2$ при $n > N$. Тогда $y_n \neq 0$ при $n > N$ и, более того, $|y_n| > |b|/2$. Пользуясь этим, получаем, что

$$\begin{aligned} \left| \frac{x_n}{y_n} - \frac{a}{b} \right| &= \frac{1}{|y_n b|} |x_n b - ab + ab - y_n a| < \\ &< \frac{1}{|b|^2/2} (|x_n - a| \cdot |b| + |y_n - b| \cdot |a|) < \frac{2(|a| + |b|)}{|b|^2} \varepsilon. \end{aligned}$$

Последнее выражение будет сколь угодно малым выбором достаточно малого ε , так что $x_n/y_n \rightarrow a/b$.

4) Если $x_n \rightarrow a$, то $kx_n \rightarrow ka$ при любом фиксированном k .

Заметим, что слово «фиксированный» означает, что k не меняется при изменении параметра n , по которому берется предел.

Другие косвенные методы нахождения пределов связаны с различными оценками последовательности.

Лемма 16.1 (о двух милиционерах). *Если последовательности $\{x_n\}$, $\{a_n\}$ и $\{b_n\}$ таковы, что $a_n \rightarrow x_0$ и $b_n \rightarrow x_0$, и, кроме того, $a_n \leq x_n \leq b_n$, то $x_n \rightarrow x_0$.*

Это легко доказать, пользуясь неравенствами для модулей, но заметим, что этот факт уже вполне очевиден из тех построений, которые мы проводили при определении вещественных чисел и описании аксиомы непрерывности.

Таким образом, если нам удастся зажать последовательность между двумя сходящимися к одному и тому же числу последовательностями, то мы легко находим ее предел — она стремится к тому же числу.

Например, рассмотрим величину $x^n/n!$ при фиксированном $x > 0$. Очевидно, что снизу она оценивается последовательностью, тождественно равной нулю ($a_n = 0$), а для оценки сверху заметим, что, начиная с некоторого номера N , для всех $n > N$ будет верно неравенство $n > 2x$ (аксиома Архимеда), так что

$$\frac{x^n}{n!} = \frac{x \cdot x \dots x}{1 \cdot 2 \dots n} = \frac{x^N}{N!} \left(\frac{x}{N+1} \dots \frac{x}{n} \right) < \frac{x^N}{N!} \left(\frac{1}{2} \right)^{n-N} / (1/2)^N.$$

В итоге мы имеем некое постоянное число $k = x^N/N!/(1/2)^N$, а также последовательность $1/2^n$. Их произведение стремится к нулю, так что

$$b_n = \frac{x^N}{N!} \left(\frac{1}{2} \right)^{n-N} / (1/2)^N \rightarrow 0.$$

Но поскольку

$$0 < \frac{x^n}{n!} < b_n,$$

заключаем, что и $x^n/n! \rightarrow 0$. На графике 16.1 можно проследить, как быстро это происходит при различных значениях постоянной x .

Наконец, очень важен такой метод, как оценка порядка малости. Если мы складываем две положительные последовательности $x_n + y_n$, причем $x_n \rightarrow a$, $y_n/x_n \rightarrow 0$, то сумма $x_n + y_n$ ведет себя ровно так же, как x_n , поскольку y_n вносит бесконечно малый вклад в сумму с ростом n .

Действительно,

$$x_n + y_n = x_n(1 + y_n/x_n),$$

где выражение в скобках стремится к 1, т.к. $y_n/x_n \rightarrow 0$, а тогда по правилу умножения пределов получаем, что

$$\lim_{n \rightarrow \infty} (x_n + y_n) = \lim_{n \rightarrow \infty} x_n.$$

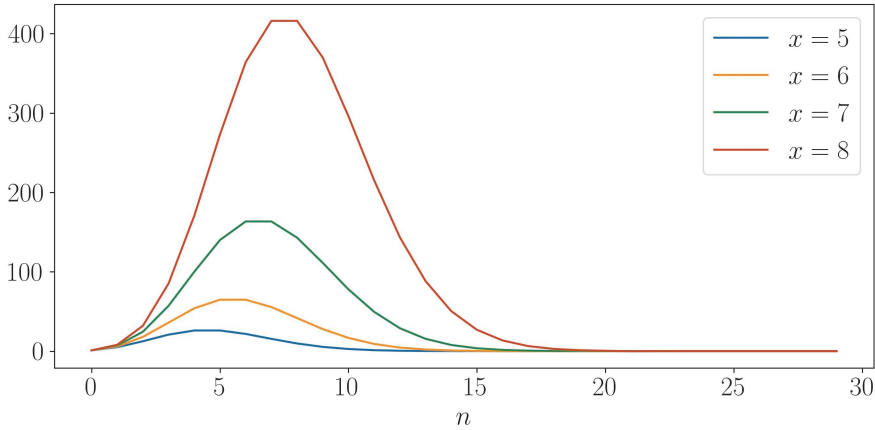


Рис. 16.1. График $y = x^n/n!$ при различных x .

Например, если у нас имеется конечная сумма вида $ax_n + bx_n^2 + cx_n^3$ и при этом $x_n \rightarrow 0$, то слагаемые, содержащие x_n^2 , x_n^3 и т. д., можно отбросить при нахождении предела, т. к. $x_n^k/x_n = x_n^{k-1} \rightarrow 0$ при $k \geq 2$, и получаем, что

$$\lim_{n \rightarrow \infty} (ax_n + bx_n^2 + cx_n^3) = \lim_{n \rightarrow \infty} (ax_n) = 0.$$

Этот прием (отбрасывание бесконечно малых слагаемых) характерен при вычислении пределов, связанных со сложными комбинаторными оценками, где количество слагаемых постоянно, а их значение поддается порядковой оценке.

Покажем, например, что $x_n = n^k/a^n \rightarrow 0$ при любом фиксированном $k \in \mathbb{N}$ и любом фиксированном $a > 1$.

Рассмотрим отношение следующего члена последовательности к предыдущему:

$$\frac{x_{n+1}}{x_n} = \frac{(n+1)^k/a^{n+1}}{n^k/a^n} = \frac{1}{a} \left(1 + \frac{1}{n}\right)^k. \quad (16.1)$$

Рассмотрим второй множитель. Если раскрыть скобки, то мы получим выражение (бином Ньютона)

$$\left(1 + \frac{1}{n}\right)^k = 1 + \frac{k}{n} + \frac{k(k-1)}{2n^2} + \dots + \frac{1}{n^k}.$$

Здесь, как мы видим, имеется конечное постоянное число слагаемых (их k штук), у которых порядок бесконечно малый по сравнению с первым слагаемым, т. е. с единицей: они все имеют вид α/n^j и стремятся к

нулю с ростом n . Следовательно,

$$\left(1 + \frac{1}{n}\right)^k \rightarrow 1.$$

Тогда по определению предела для любого $\varepsilon > 0$ найдется такой номер N , что для любого $n > N$ имеет место неравенство

$$\left| \left(1 + \frac{1}{n}\right)^k - 1 \right| < \varepsilon.$$

Положим $\varepsilon = (a - 1)/2$ (поскольку $a > 1$). Тогда при $n > N$ (N зависит от выбранного ε) получим, что

$$\left(1 + \frac{1}{n}\right)^k < 1 + \frac{a - 1}{2}.$$

Подставим эту оценку в (16.1), и тогда для $n > N$ будем иметь

$$\frac{x_{n+1}}{x_n} = \frac{1}{a} \left(1 + \frac{1}{n}\right)^k < \frac{a + 1}{2a} < 1.$$

Перемножая такие отношения, начиная с $n = N$ и заканчивая произвольным $n > N$, получаем

$$\frac{x_n}{x_N} = \frac{x_{N+1}}{x_N} \cdot \frac{x_{N+2}}{x_{N+1}} \cdots \frac{x_n}{x_{n-1}} < \left(\frac{a + 1}{2a}\right)^{n-N},$$

откуда следует, что

$$\frac{n^k}{a^n} < \frac{N^k}{a^N} \left(\frac{a + 1}{2a}\right)^{n-N}.$$

То есть мы получаем произведение константы на некоторое число < 1 в растущей степени. И по аналогии с $(1/2)^n \rightarrow 0$ заключаем, что $((a + 1)/2a)^n \rightarrow 0$, и, следовательно,

$$\frac{n^k}{a^n} \rightarrow 0.$$

На графике 16.2 представлено четыре варианта параметров a и k , а по оси Ox отложен параметр n . Как видим, несмотря на большой скачок в начале, эти последовательности довольно быстро уходят в ноль с ростом n .

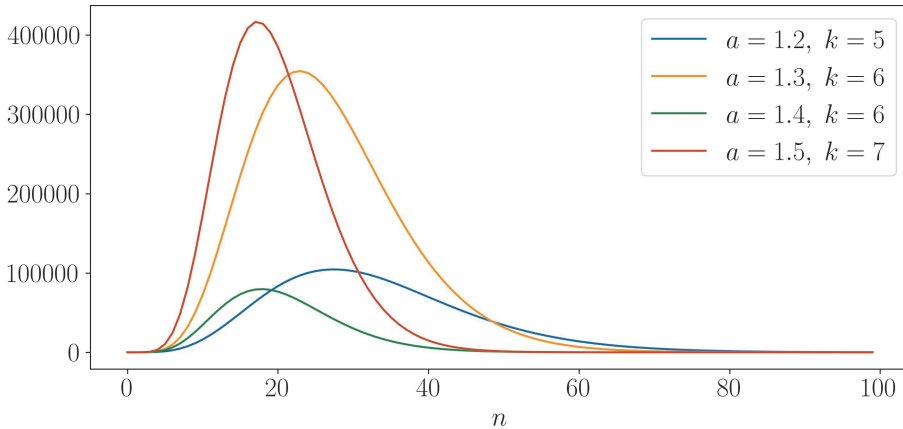


Рис. 16.2. График $y = n^k / a^n$ при различных параметрах k и a .

Рассмотрим теперь такую функцию, как корень натуральной степени $\sqrt[n]{x}$. Ее определение таково: $\sqrt[n]{x}$ — это такое положительное y , что $y^n = x$. Обоснование существования корня любого положительного x мы отложим до следующего раздела, а пока решим такую задачу: чему равен

$$\lim_{n \rightarrow \infty} \sqrt[n]{x}$$

при любом $x > 0$?

Пусть для начала $x > 1$. Докажем такое неравенство:

$$(1 + c)^n \geq 1 + cn \quad (c > -1).$$

Для этого воспользуемся индукцией. При $n = 1$ неравенство очевидно. Предположим, что оно верно при n , и выведем его при $n + 1$.

Действительно, по предположению индукции имеем $(1 + c)^n \geq 1 + cn$, тогда при умножении данного неравенства на положительное число $1 + c$ оно сохраняет знак, и мы получаем, что

$$(1 + c)^{n+1} = (1 + c)^n(1 + c) \geq (1 + cn)(1 + c) = 1 + c + cn + c^2n \geq 1 + c(n + 1),$$

что и требовалось доказать.

Из доказанного, в частности, следует, что $(1 + c)^n > cn$.

Подставим в полученное неравенство $c = \sqrt[n]{x} - 1$ и получим

$$x > (\sqrt[n]{x} - 1)n \quad \text{или} \quad \sqrt[n]{x} < 1 + \frac{x}{n}.$$

С другой стороны, $\sqrt[n]{x} > 1$. Тогда по лемме о двух милиционерах получаем, что

$$\lim_{n \rightarrow \infty} \sqrt[n]{x} = 1.$$

Если $x < 1$, то перейдем к величине $y = 1/x$. Для нее получим, что $\sqrt[n]{y} \rightarrow 1$. Но

$$\sqrt[n]{x} = 1 / \sqrt[n]{y},$$

так что и в этом случае получаем тот же самый предел.

На рис. 16.3 представлено несколько графиков корня различной степени. Хорошо видно, что с ростом n кривая графика все плотнее прижимается к прямой $y = 1$ как слева, так и справа от точки $x = 1$.

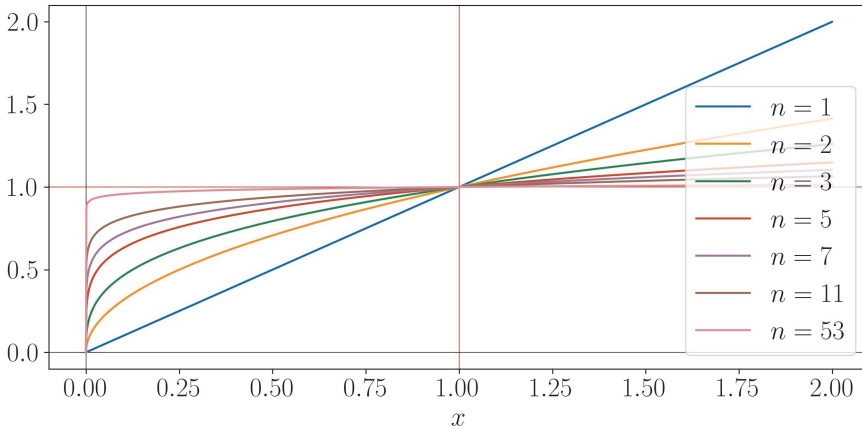


Рис. 16.3. График $y = \sqrt[n]{x}$ при различных значениях n .

Выше мы рассмотрели определение и несколько примеров предела последовательности. Заметим, что последовательность — это частный случай функции, а именно, это функция из \mathbb{N} в некоторое множество. Введем теперь определение для предела функции, заданной на подмножестве числовой оси \mathbb{R} . Напомним, что символом $U_\varepsilon(x)$ мы обозначаем ε -окрестность точки x , т.е. интервал $(x - \varepsilon; x + \varepsilon)$. Если из множества $U_\varepsilon(x)$ выбросить саму точку x , то получается множество $\dot{U}_\varepsilon(x) = (x - \varepsilon; x) \cup (x; x + \varepsilon)$, которое называется *проколотой* ε -окрестностью точки x .

Пусть $\dot{U}_r(\hat{x}) \subseteq \text{dom}(f)$, т.е. функция f определена как минимум на некоторой проколотой r -окрестности точки \hat{x} (область определения f может быть шире). Тогда число a называется **пределом функции f в точке \hat{x}** , если

$$\forall \varepsilon > 0 \exists \delta \in (0; r) : \forall x (0 < |x - \hat{x}| < \delta) \Rightarrow (|f(x) - a| < \varepsilon). \quad (16.2)$$

Иначе говоря, a есть предел функции f в точке \hat{x} , если значения $f(x)$ сколь угодно близки к числу a при выборе x , достаточно близких к (но при этом отличных от) \hat{x} . Этот факт записывается следующим образом:

$$a = \lim_{x \rightarrow \hat{x}} f(x).$$

Заметим, что в точке \hat{x} функция f может быть не определена, может принимать значение, отличное от a , может совпадать с a . Поэтому в определении предела функции существенным является требование $|x - \hat{x}| > 0$, т. е. $x \neq \hat{x}$.

Определение предела функции в форме (16.2) называется определением «по Коши». Существует эквивалентное определение предела функции «по Гейне» с использованием последовательностей. Число a называется пределом функции f в точке \hat{x} , если для любой последовательности $\{x_n\} \subseteq \dot{U}_r(\hat{x})$ такой, что $x_n \rightarrow \hat{x}$, предел последовательности $\{f(x_n)\}$ существует и равен a .

В этом определении предела функции существенным является требование того, что все значения x_n лежат в проколотой окрестности точки \hat{x} , т. е. отличны от \hat{x} .

Теорема 16.1. *Определения предела функции по Коши и по Гейне эквивалентны.*

Доказательство. Пусть a является пределом $f(x)$ в точке \hat{x} по Коши. Пусть $x_n \rightarrow \hat{x}$, причем $\{x_n\} \subseteq \dot{U}_r(\hat{x})$. Нужно показать, что $f(x_n) \rightarrow a$.

Пусть $\varepsilon > 0$. Тогда существует $\delta > 0$ такой, что при $0 < |x - \hat{x}| < \delta$ имеет место неравенство $|f(x) - a| < \varepsilon$. Поскольку $x_n \rightarrow \hat{x}$, для данного δ найдется такое N , что при всех $n > N$ имеем $|x_n - \hat{x}| < \delta$, откуда следует, что $|f(x_n) - a| < \varepsilon$. А это и означает, что $f(x_n) \rightarrow a$.

Обратно. Пусть a является пределом $f(x)$ в точке \hat{x} по Гейне. Предположим, что при этом a не является пределом $f(x)$ в точке \hat{x} по Коши. Тогда существует $\varepsilon > 0$ такой, что для любого $\delta > 0$ существует $x = x(\delta)$ такой, что

$$(0 < |x - \hat{x}| < \delta) \wedge (|f(x) - a| \geq \varepsilon).$$

Так как выбор δ произволен, будем считать, что $\delta = 1/n$. Для каждого n найдем соответствующий $x = x_n$ такой, что

$$(0 < |x_n - \hat{x}| < 1/n) \wedge (|f(x_n) - a| \geq \varepsilon).$$

Но тогда мы получаем последовательность $\{x_n\}$, сходящуюся к \hat{x} (поскольку $1/n \rightarrow 0$), $x_n \neq \hat{x}$, для которой $|f(x_n) - a| \geq \varepsilon$, т. е. $f(x_n) \not\rightarrow a$. А это противоречит пределу функции по Гейне. \square

Рассмотрим еще один важный предел, который называется **первым замечательным пределом**:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

Далее для простоты будем считать, что $x > 0$ (к отрицательным x перейти очень просто, т. к. $\sin(-x) = -\sin(x)$).

Вспомним, что угол, измеренный в радианах, на единичной окружности равен длине дуги, соответствующей данному углу. Так что возьмем единичную окружность с центром в точке O и отложим от направления Ox угол x .

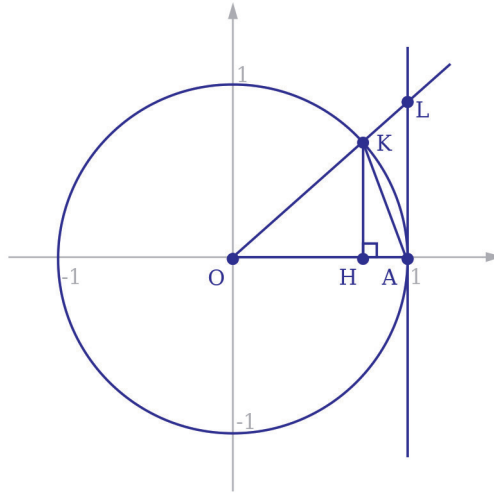


Рис. 16.4. Вычисление первого замечательного предела.

На картинке 16.4 этот угол называется $\angle AOK$, где $A = (1, 0)$. Кроме того, построим две нормали KH и LA .

Очевидно, что

$$S_{\triangle OAK} < S_{\text{sect}KOA} < S_{\triangle OAL}, \quad (16.3)$$

где $S_{\text{sect}KOA}$ — площадь сектора KOA .

Поскольку $|KH| = \sin x$, $|LA| = \operatorname{tg} x$:

$$\begin{aligned} S_{\triangle OAK} &= \frac{1}{2} \cdot |OA| \cdot |KH| = \frac{1}{2} \cdot 1 \cdot \sin x = \frac{\sin x}{2}, \\ S_{\text{sect}KOA} &= \frac{1}{2} \cdot |OA|^2 \cdot x = \frac{x}{2}, \\ S_{\triangle OAL} &= \frac{1}{2} \cdot |OA| \cdot |LA| = \frac{\operatorname{tg} x}{2}. \end{aligned}$$

Подставляя в (16.3), получим:

$$\frac{\sin x}{2} < \frac{x}{2} < \frac{\operatorname{tg} x}{2}.$$

Так как угол x близок к нулю и положителен, то можно считать, что он находится в первой четверти плоскости, поэтому $\sin x > 0$, $x > 0$, $\operatorname{tg} x > 0$, откуда

$$\frac{1}{\operatorname{tg} x} < \frac{1}{x} < \frac{1}{\sin x}.$$

Умножаем на $\sin x$:

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

Отсюда, поскольку $\cos x \rightarrow 1$, получаем требуемый предел.

Следствием первого замечательного предела является такой предел:

$$\lim_{x \rightarrow 0} \frac{1 - \cos(x)}{x^2} = \frac{1}{2}.$$

Действительно,

$$\frac{1 - \cos(x)}{x^2} = \frac{1 - \cos(x)}{x^2} \cdot \frac{1 + \cos(x)}{1 + \cos(x)} = \left(\frac{\sin x}{x}\right)^2 \cdot \frac{1}{1 + \cos x} \rightarrow \frac{1}{2},$$

поскольку $\sin(x)/x \rightarrow 1$ и $\cos x \rightarrow 1$.

16.2. Экспонента

Итак, \mathbb{R} , или вещественная прямая, — это непрерывное линейно упорядоченное поле. То есть в \mathbb{R} можно складывать, вычитать, умножать и делить, а также сравнивать, причем сравнение согласовано с операциями сложения и умножения. Кроме того, в нем «нет дыр», т. е. каждую точку прямой можно адресовать пределом счетной последовательности вложенных отрезков и, что самое главное, каждый адрес, заданный пределом цепи вложенных отрезков, заселен как минимум одной точкой. Кроме того, в \mathbb{R} выполняется аксиома Архимеда.

Мы также получили доказательство того, что вся прямая равномошна интервалу $(-1; 1)$, а значит, и любому открытому интервалу $(a; b)$, где $a < b$. Для этого достаточно было предъявить биекцию в явном виде. Проще всего в таких случаях строить строго монотонную биекцию, чтобы быть уверенным в ее инъективности. Кроме того, монотонная инъекция сохраняет порядок, т. е. устанавливает порядковый изоморфизм между всей прямой и ее частью.

В этом плане было бы интересно задаться вопросом — нет ли биекций, которые сохраняли бы хоть как-то еще и операции сложения и умножения, заданные на \mathbb{R} ? Как уже отмечалось выше, поле \mathbb{R} уникально по своей природе, т. е. не существует других полей (содержащихся в \mathbb{R} или содержащих \mathbb{R}), полностью изоморфных ему, т. е. с сохранением всех операций и согласованного с ними отношения линейного порядка. Однако при некотором ослаблении требований к изоморфизму кое-что интересное мы можем отыскать.

Посмотрим на операции сложения и умножения как на две операции из разных групп. Точнее, рассмотрим группу по сложению $(\mathbb{R}, +)$

и группу по умножению $\langle \mathbb{R}^+, \cdot \rangle$, где под \mathbb{R}^+ мы понимаем множество всех положительных действительных чисел. Эти группы, кроме того, линейно упорядочены стандартным отношением $<$.

Мы хотим найти взаимно однозначное соответствие $f : \mathbb{R} \leftrightarrow \mathbb{R}^+$ между этими группами такое, чтобы выполнялось функциональное тождество:

$$f(x + y) = f(x)f(y). \quad (16.4)$$

Иначе говоря, f должно переводить сложение в умножение. Кроме того, мы хотим, чтобы f сохраняло и порядок:

$$x < y \Rightarrow f(x) < f(y). \quad (16.5)$$

Требования (16.4) и (16.5) называются функциональными уравнением и неравенством, поскольку ограничивают не выбор переменной, а выбор функции. Переменные x и y предполагаются произвольными из области определения f .

Вместо условия (16.5) можно требовать сохранение обратного порядка, т. е. $f(x) > f(y)$ при $x < y$. К этому случаю мы вернемся чуть позже.

Итак, мы ищем изоморфизм структур $\langle \mathbb{R}, +, < \rangle$ и $\langle \mathbb{R}^+, \cdot, < \rangle$, сохраняющий операцию и порядок.

Для начала заметим, что, как и положено изоморфизму, f переводит нейтральный элемент в нейтральный: $f(0) = 1$. Действительно, $f(x) = f(x + 0) = f(x)f(0)$, откуда $f(0) = 1$ (сокращать на $f(x)$ мы можем, т. к. $f(x) > 0$ по определению).

Далее обозначим за a число $f(1)$. В силу требования (16.5) имеем $a > 1$.

Легко видеть, что

$$f(2) = f(1)f(1) = a^2, \quad f(3) = f(2)f(1) = a^3, \quad \dots, \quad f(n) = a^n,$$

причем $n \in \mathbb{Z}$ может быть и отрицательным числом, поскольку $f(-n) = 1/f(n)$. Таким образом, уже на целых числах мы видим, что f является **показательной функцией**.

Пусть теперь $x = p/q$. Сложим x сам с собой q раз, и получим

$$f\left(q \frac{p}{q}\right) = f(p/q)^q = f(p) = a^p.$$

Так что $f(p/q) = \sqrt[q]{a^p}$ или $f(p/q) = a^{p/q}$.

Отметим, что на рациональных точках наша функция монотонно возрастает, как того и требует условие (16.5). Действительно, пусть $p/q < t/s$. Сравним $a^{p/q}$ и $a^{t/s}$.

Поскольку порядок на \mathbb{R} согласован с операцией умножения, легко получить, что для положительных x, y и натурального m неравенство

$x^m < y^m$ верно тогда и только тогда, когда $x < y$. Поэтому, полагая $m = qs$, получаем, что

$$a^{p/q} < a^{t/s} \Leftrightarrow (a^{p/q})^{qs} < (a^{t/s})^{qs} \Leftrightarrow a^{ps} < a^{qt},$$

а это уже легко выводится из определения натуральной степени, поскольку при $ps < qt$ имеем

$$a^{qt} = a^{ps} \cdot \underbrace{a \cdot \dots \cdot a}_{qt-ps \text{ раз}},$$

и, так как $a > 1$, неравенство выполняется. Для отрицательных дробей все сводится к положительным, если рассмотреть обратные числа, а для разнознаковых дробей достаточно отметить, что $a^{-p/q}a^{p/q} = 1$, так что если одно число больше 1, то второе меньше, и монотонность снова имеет место.

Итак, мы научились рассчитывать функцию f для рациональных чисел, причем она достраивается однозначно в зависимости только от параметра $a = f(1)$. Кроме того, она оказалась строго возрастающей при $a > 1$. Как осуществить переход к иррациональным числам?

Для этого у нас есть аксиома непрерывности, которая позволяет осуществлять предельные переходы. Действительно, предположим, что x есть предел вложенных отрезков с двоично-рациональными концами:

$$[r_0; s_0] \supset [r_1; s_1] \supset [r_2; s_2] \supset \dots \supset \{x\},$$

где все $r_k, s_k \in \mathbb{B}_k$, и каждый следующий отрезок вдвое короче предыдущего.

Построим точки $R_k = f(r_k) = a^{r_k}$, $S_k = f(s_k) = a^{s_k}$. По доказанному ранее получаем, что если $r_k < s_k$, то $R_k < S_k$ и, кроме того, вложенность отрезков также сохранится, т. е.

$$[R_0; S_0] \supset [R_1; S_1] \supset [R_2; S_2] \supset \dots$$

По аксиоме полноты пределом такой цепи будет непустое множество

$$X = \bigcap_{k=0}^{\infty} [R_k; S_k].$$

Покажем, что это множество состоит из одного элемента. Предположим, что это не так, и в X есть хотя бы два элемента $c < d$. Но тогда $R_k \leq c < d \leq S_k$ для всех k , а значит, разность $S_k - R_k$ не может быть меньше, чем $d - c$.

Рассмотрим отношение $S_k/R_k = a^{s_k}/a^{r_k}$. По уже доказанным свойствам функции f легко получить, что это отношение равно $a^{s_k - r_k}$. В

силу того, что r_k, s_k выбирались как двоично-рациональные числа последовательным делением предыдущего отрезка пополам, очевидно, что $s_k - r_k = 1/2^k$.

Далее, в силу монотонности функции f для рациональных чисел, получаем, с одной стороны, что

$$1 < a^{s_k - r_k} = a^{\frac{1}{2^k}}. \quad (16.6)$$

С другой стороны, вспомним ранее доказанное неравенство $(1+b)^m > bm$ для всех $b > -1$ и натуральных m . Положим $b = a/m$, тогда $(1 + a/m)^m > a$, откуда следует: $(1 + a/m) > a^{1/m}$. Теперь положим $m = 2^k$ и получим:

$$a^{s_k - r_k} = a^{\frac{1}{2^k}} < 1 + \frac{a}{2^k}. \quad (16.7)$$

Итак, в силу неравенств (16.6) и (16.7) величина $a^{s_k - r_k}$ оказывается зажатой между 1 и $1 + a/2^k$, а какое бы большое a ни было, с ростом k отношение $a/2^k$ приближается к нулю. В частности, можно найти такое k , начиная с которого $a/2^k < (d - c)/2d$.

Тогда для тех же k получим

$$S_k - R_k = R_k \left(\frac{S_k}{R_k} - 1 \right) < d(a^{s_k - r_k} - 1) < d \frac{a}{2^k} < \frac{d - c}{2},$$

а это противоречит тому, что $S_k - R_k \geq d - c$.

Следовательно, в X нет двух различных точек, т.е. $X = \{y\}$. Более того, заметим, что так как $r_k \leq x \leq s_k$ для всех k , то и $R_k \leq f(x) \leq S_k$ для всех k в силу требования монотонности функции f . Но тогда $f(x)$ больше некуда деваться, кроме как быть равным числу y — единственному, удовлетворяющему таким же неравенствам, по доказанному выше.

Итак,

$$a^{r_k} \leq f(x) \leq a^{s_k}.$$

В этом случае вместо $f(x)$ мы также пишем a^x . Тем самым мы продлили определение $f(x)$, используя аксиому непрерывности, на все иррациональные числа. Можно показать, что данное определение корректно, т.е. не зависит от выбора рациональных последовательностей $\{r_k\}$ и $\{s_k\}$.

Чтобы убедиться в том, что построенная функция $f(x)$ является искомым изоморфизмом упорядоченных групп $\langle \mathbb{R}, +, < \rangle$ и $\langle \mathbb{R}^+, \cdot, < \rangle$, нужно проверить также, что она является биекцией.

Прежде всего, заметим, что это инъекция, т.к. из того, что $f(x) = f(y)$ следует $x = y$ в силу требования (16.5).

Чтобы показать сюръективность $f(x)$, рассмотрим ее на двух подмножествах \mathbb{R} — отрицательных x и положительных x , и покажем, что

образами этих подмножеств будут $(0; 1)$ и $(1; +\infty)$, т. е.

$$f[(-\infty; 0)] = (0; 1), \quad f[(0; \infty)] = (1; +\infty),$$

что вместе с равенством $f(0) = 1$ дает сюръективность f .

Пусть для начала $x > 0$. Вспомним о неравенстве $(1 + b)^m > bm$. Полагая $1 + b = a$, находим, что $a^m > (a - 1)m$, так что, выбирая достаточно большое m , мы можем получить сколь угодно большое число a^m , т. е. $f(x)$ не ограничена сверху.

На самом деле $f(x)$ при $x > 0$ принимает и все промежуточные значения между 1 и ∞ . Докажем это. Пусть $C > 1$ — некоторое число. Требуется найти такое $c > 0$, что $f(c) = C$. Так как f не ограничена сверху, то найдется такое $y > 0$, что $f(y) > C$. При этом мы сразу будем считать, что $y = 2^m$ (если это не так, то мы можем легко найти натуральное m такое, что $y < 2^m$, пользуясь аксиомой Архимеда). Тогда $C \in (f(0); f(y))$, поскольку $f(0) = 1$.

Мы построим последовательность вложенных отрезков, отправляясь от отрезка $[a_0; b_0] = [0; y]$, следуя той же логике по поимке произвольной точки в сеть точек множества \mathbb{B} , какой мы пользовались при описании принципа вложенных отрезков, а также при доопределении функции f в иррациональных точках. Только на этот раз будем использовать образ этой сети относительно функции f .

Итак, пусть $x_0 = (b_0 - a_0)/2$. Ясно, что $x_0 \in \mathbb{B}$ и $f(a_0) < f(x_0) < f(b_0)$. При этом $C \in [f(a_0); f(x_0)]$ или $C \in [f(x_0); f(b_0)]$ в силу монотонности f . Если выполняется первое утверждение, то положим $[a_1; b_1] = [a_0; x_0]$, иначе — $[a_1; b_1] = [x_0; b_0]$. Снова имеем: $C \in [f(a_1); f(b_1)]$. Затем полагаем $x_1 = (b_1 - a_1)/2$, и повторяем предыдущие рассуждения. И так далее. В итоге мы получим последовательность

$$[a_0; b_0] \supseteq [a_1; b_1] \supseteq \dots [a_n; b_n] \supseteq \dots$$

вложенных отрезков, длина которых стремится к нулю (т. к. $b_n - a_n = y/2^n$) и при этом $C \in [f(a_n); f(b_n)]$ при всех натуральных n . В силу принципа вложенных отрезков мы получим, что

$$\bigcap_{n=0}^{\infty} [a_n; b_n] = \{c\},$$

т. е. в пересечении всех этих отрезков лежит единственная точка c , для которой нам и нужно показать, что $C = f(c)$.

Мы можем также заметить, что и отрезки $[f(a_n); f(b_n)]$ образуют последовательность вложенных отрезков (в силу монотонности f) с длиной, стремящейся к нулю, т. к.

$$f(b_n) - f(a_n) = a^{b_n} - a^{a_n} = f(a_n)(a^{b_n - a_n} - 1) \leq f(y)(a^{\frac{2^m}{2^n}} - 1) \leq f(y) \frac{a}{2^{n-m}},$$

где мы воспользовались тем, что $a_n < y$, а также неравенством (16.7). Отсюда следует, что $\bigcap_n [f(a_n); f(b_n)] = \{C\}$.

Вспомним, как мы определяли f для произвольных точек прямой с помощью последовательностей $\{r_k\}$ и $\{s_k\}$ двоично-рациональных чисел, образующих последовательность вложенных отрезков. А именно, если $r_k \leq c \leq s_k$ и при этом r_k и s_k имеют общий предел, равный c , то мы доказывали, что и последовательности $f(r_k)$ и $f(s_k)$ обладают таким же свойством — имеют общий предел, который мы и определили в качестве значения $f(c)$.

Теперь положим $r_n = a_n$ и $s_n = b_n$. Это — двоично-рациональные последовательности, сходящиеся к точке c (которую мы определили выше с их помощью). Следовательно, единственная точка, лежащая в пересечении отрезков $[f(a_n); f(b_n)]$, т.е. точка C , и является по определению значением функции f в точке c , т.е. $C = f(c)$.

Итак, выбрав произвольную точку $C \in (1; +\infty)$, мы указали алгоритм поиска ее прообраза, т.е. такой точки $c \in (0; \infty)$, что $C = f(c)$. Следовательно, $f[(0; +\infty)] = (1; +\infty)$.

Для того, чтобы показать, что f взаимно однозначно отображает отрицательные числа в числа из интервала $(0; 1)$, достаточно знать, что $f(-x) = 1/f(x)$. Отсюда следует, что $f[(-\infty; 0)] = (0; 1)$. И на этом доказательство сюръективности f завершается.

Таким образом, требования (16.4) и (16.5) и условие $f(1) > 0$ приводят нас к построению взаимно однозначного соответствия между линейно упорядоченными группами $\langle \mathbb{R}, +, < \rangle$ и $\langle \mathbb{R}^+, \cdot, < \rangle$.

Функция $f(x)$, обозначаемая a^x и удовлетворяющая условиям (16.4) и (16.5), единственна с точностью до выбора числа a . Такая функция называется **показательной** с основанием a .

Если число a выбрать меньше 1, то мы можем повторить все те же рассуждения, заменяя всюду знак $<$ на $>$, т.е. мы построим изоморфизм между линейно упорядоченными группами $\langle \mathbb{R}, +, < \rangle$ и $\langle \mathbb{R}^+, \cdot, > \rangle$, инвертировав порядок в мультипликативной группе положительных вещественных чисел.

На самом деле можно поступить еще проще, и вместо a^x рассмотреть функцию $(1/a)^x$, которая будет изоморфизмом, сохраняющим прямой порядок на \mathbb{R} , а уже функция a^x будет выражаться через нее как $1/(1/a)^x$.

Единственный случай, который выбивается из требований сохранения порядка, но при этом сохраняет операции, — это случай $a = 1$. Поскольку тогда мы получим $a^x = 1$ для всех x . Очевидно, что требование (16.4) выполняется, а требование (16.5) — нет. Однако все три случая ($a > 1$, $a = 1$, $a < 1$) относятся к показательной функции без ограничения общности.

На графике 16.5 изображено несколько случаев показательной функции при различных a .

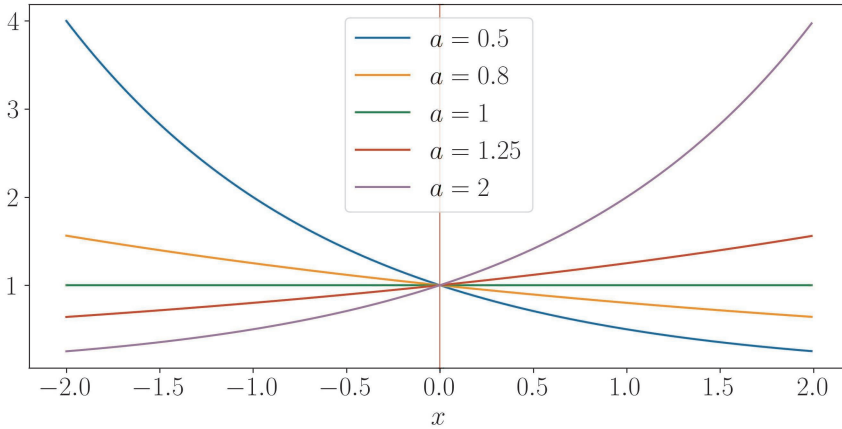


Рис. 16.5. График $y = a^x$ при некоторых значениях параметра a .

Выше мы упомянули о том, что функция a^x продлена по непрерывности на иррациональные числа. В данном случае мы, конечно, имели в виду непрерывность в смысле аксиомы непрерывности действительных чисел, поскольку именно ею мы пользовались для построения показательной функции. Однако термин **непрерывность** имеет намного более широкий спектр значений. И прежде всего он связан с непрерывностью функций. Грубо говоря, непрерывная функция действует таким образом, что сохраняет непрерывность своей области определения, переводит непрерывное в непрерывное (обратное не обязательно верно).

В математике существует несколько эквивалентных определений непрерывности функции. Приведем два из них. Пусть $f : X \rightarrow \mathbb{R}$, где X — непустое подмножество \mathbb{R} , содержащее некоторую окрестность точки x_0 . Функция f **непрерывна в точке** x_0 , если

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in X (|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon). \quad (16.8)$$

По-другому это записывается так:

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Иначе говоря, значения $f(x)$ становятся сколь угодно близки к значению $f(x_0)$ при x , достаточно близких к x_0 , т. е. у непрерывной функции не только существует предел в точке x_0 , но он еще и равен значению функции в данной точке. Для упрощения мы не рассматриваем случай изолированной точки, т. е. такой точки $x_0 \in X$, что $\dot{U}_r(x_0) \cap X = \emptyset$ при некотором $r > 0$ (формально говоря, для изолированной точки определение (16.8) выполняется).

Заметим, что стандартным словарем в математическом анализе являются выражения: «сколь угодно близкий» или «сколь угодно малый» (что означает произвольно малое отклонение от какой-то величины или от нуля и формально сопровождается квантором $\forall \varepsilon > 0$), а также «достаточно близкий» или «достаточно малый» (что означает возможность найти некоторую малость отклонения от какой-то величины или от нуля, достаточную для выполнения некоего условия, и формально сопровождается квантором $\exists \delta > 0$)

Второе определение связано с последовательностями. Функция $f : X \rightarrow \mathbb{R}$, где $X \subseteq \mathbb{R}$ и $U_r(x_0) \subseteq X$ при некотором $r > 0$, непрерывна в точке x_0 , если для любой последовательности $\{x_n\}$

$$\lim_{n \rightarrow \infty} x_n = x_0 \Rightarrow \lim_{n \rightarrow \infty} f(x_n) = f(x_0), \quad (16.9)$$

т. е. f непрерывна, когда она *сохраняет предельные переходы*.

Как и в случае с пределами, верна теорема об эквивалентности определений.

Теорема 16.2. *Определения (16.8) и (16.9) непрерывности функции в точке эквивалентны.*

Доказательство этой теоремы практически полностью повторяет доказательство теоремы 16.1.

Если функция f непрерывна в каждой точке некоторого множества $D \subseteq \text{dom}(f)$ (не обязательно всей области определения), то говорят, что f **непрерывна на D** .

Из построения функции a^x , в принципе, видно, что она непрерывна на \mathbb{R} , но проведем строгие рассуждения для доказательства этого факта. Для определенности положим $a > 1$. Требуется показать, что для любого x и для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что для любого y , если $|x - y| < \delta$, то $|a^x - a^y| < \varepsilon$.

Выберем произвольные x и ε , а в качестве δ возьмем число $1/m$ такое, что $\varepsilon > a^{1+x}/m$ (какие бы a и x ни были, такое число m найдется в силу аксиомы Архимеда). Пусть теперь $|x - y| < 1/m$, причем $y > x$. В силу монотонности показательной функции $a^x < a^y$. Заметим, что в этом случае $|a^x - a^y| = a^x(a^{y-x} - 1)$.

Далее, поскольку $y - x < 1/m$, оценим $a^{y-x} - 1 < a^{1/m} - 1 < a/m$ (это мы уже показывали ранее), откуда

$$|a^x - a^y| < a^{x+1}/m < \varepsilon.$$

Если же $y < x$, то модуль раскрывается иначе, поскольку $a^y < a^x$:

$$|a^x - a^y| = a^y(a^{x-y} - 1) < a^x(a^{x-y} - 1) < \varepsilon.$$

Таким образом, по определению (16.8) функция a^x при $a > 1$ непрерывна. Совершенно точно так же доказывается и непрерывность в случае $a < 1$, а при $a = 1$ доказательство непрерывности тривиально.

Приведем несколько свойств показательной функции:

- 1) $a^x a^y = a^{x+y}$;
- 2) $a^x / a^y = a^{x-y}$;
- 3) $(ab)^x = a^x b^x$;
- 4) $(a/b)^x = a^x / b^x$;
- 5) $\sqrt[n]{a^x} = a^{x/n}$;
- 6) $a^x < a^y$ тогда и только тогда, когда $x < y$, если $a > 1$;
- 7) $(a^x)^y = a^{xy}$.

Все эти свойства сначала доказываются для натуральных и рациональных чисел, а затем переносятся на иррациональные числа по непрерывности.

Следующее замечательное применение теории пределов находит в определении производной. Если существует предел

$$\lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x},$$

то такой предел называется **производной функции** f в точке x_0 и обозначается $f'(x_0)$. Геометрически производная в конкретной точке x_0 — это тангенс угла наклона касательной к графику функции в точке $(x_0, f(x_0))$.

Имея значение выражения $\alpha = (f(x_0 + \Delta x) - f(x_0))/\Delta x$ для данной конкретной функции $f(x)$, мы можем построить прямую l , имеющую тангенс угла наклона α и проходящую через точку $(x_0, f(x_0))$. Ясно, что l зависит как от функции f , так и от выбранной точки x_0 и приращения Δx . Уравнение l будет иметь вид:

$$y = \alpha(x - x_0) + f(x_0).$$

Если же в качестве α взять производную, то мы получим уравнение касательной к графику функции f в точке $(x_0, f(x_0))$. Это уравнение будет иметь вид: $y = f'(x_0)(x - x_0) + f(x_0)$.

Рассмотрим для примера $f(x) = e^x$ и $x_0 = 0$. Выберем несколько вариантов Δx и построим соответствующие варианты прямой l и касательную с помощью указанных уравнений. На рис. 16.6 жирная синяя линия соответствует самой функции $f(x) = e^x$, оранжевая линия — прямой l при $\Delta x = -2$, зеленая линия — прямой l при $\Delta x = -1$, красная — прямой l при $\Delta x = 1$, фиолетовая — прямой l при $\Delta x = 2$. Коричневая

прямая соответствует $\Delta x = 0$, т. е. является касательной к графику e^x в точке $x_0 = 0$.¹

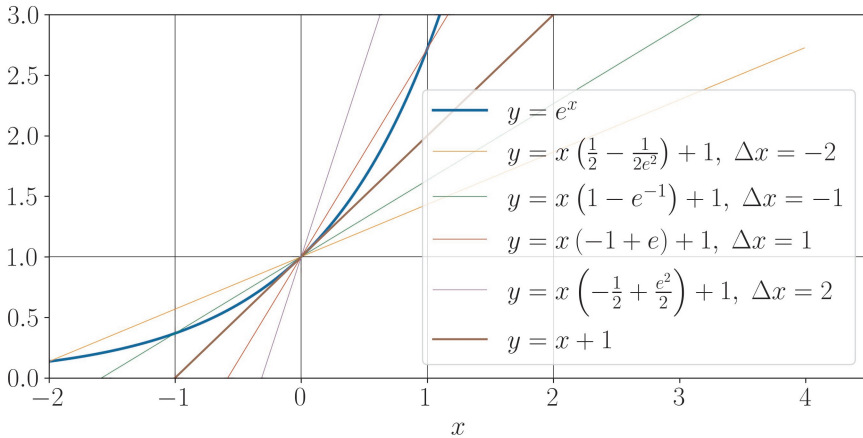


Рис. 16.6. Приближение касательных к графику $y = e^x$ в точке $x_0 = 0$.

Нетрудно видеть, что прямая l при $\Delta x \neq 0$ является хордой (или секущей) для графика функции $f(x)$, и только если устремить Δx к нулю, то есть перейти к производной в точке x_0 , то точка пересечения l с графиком $f(x)$ устремляется к точке $(x_0, f(x_0))$, а сама прямая l превращается в касательную.

Конечно, для того чтобы в пределе получить касательную, функция должна быть достаточно хорошей или, как говорят, **гладкой**. Собственно, функция и называется гладкой, если у нее есть производная.

Теорема 16.3. Если у функции $f(x)$ есть производная в точке x , то она непрерывна в этой точке.

Доказательство. В силу существования предела

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

получаем, что

$$\frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x) + \varepsilon(\Delta x),$$

где $\varepsilon(\Delta x) \rightarrow 0$ при $\Delta x \rightarrow 0$. Находим разность значений функции:

$$f(x + \Delta x) - f(x) = f'(x)\Delta x + \varepsilon(\Delta x)\Delta x.$$

Как видим, справа стоит сумма двух бесконечно малых, так что $f(x + \Delta x) \rightarrow f(x)$ при $\Delta x \rightarrow 0$. Следовательно, f непрерывна в точке x . \square

¹ Построение графиков и вычисление уравнений выполнено на языке Python с использованием библиотек `sympy` и `matplotlib`.

Каковы могут быть значения производной для показательной функции? Если посмотреть на график различных показательных функций с основанием $a = 2, 1.25, 1, 0.8, 0.5$ (рис. 16.5), то легко заметить, что даже в точке $x = 0, y = 1$ наклон касательной может быть каким угодно, кроме абсолютно вертикального. В частности, касательная может быть горизонтальной, если $a = 1$.

Действительно, в случае $a = 1$ получаем, что $y = a^x = 1$ для всех x , т. е. является константой, а для константы производная вычисляется очень легко:

$$\lim_{\Delta x \rightarrow 0} C - C\Delta x = \lim_{\Delta x \rightarrow 0} 0 = 0.$$

Среди всех показательных функций принято особо выделять одну, а именно, такую, у которой наклон касательной в точке $x = 0$ равен 45° или, иначе говоря, производная равна 1. Основание a в этом случае обозначается буквой e и называется **числом Эйлера**.

$$e \approx 2.718281828459045$$

Итак, по определению число e таково, что $(e^x)' = 1$ в точке $x = 0$, т. е.

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1, \text{ или } e^x = 1 + x + o(x),$$

где $o(x)$ — бесконечно малая в сравнении с x величина при $x \rightarrow 0$.

Функция e^x называется **экспонентой**. Иногда также используется обозначение $\exp(x)$.

Найдем производную экспоненты в произвольной точке x :

$$\frac{e^{x+\Delta x} - e^x}{\Delta x} = e^x \frac{e^{\Delta x} - 1}{\Delta x} \rightarrow e^x,$$

так что $(e^x)' = e^x$. Итак, производная экспоненты есть сама же экспонента! В терминах высшей математики это можно выразить так: *экспонента является неподвижной точкой оператора дифференцирования*.

Найдем теперь производную a^x в общем случае. Так как число a — некоторое положительное число, а функция e^x принимает все положительные значения, то существует такое число x_0 , что $e^{x_0} = a$. В таком случае,

$$a^x = (e^{x_0})^x = e^{xx_0},$$

и далее получаем, что

$$\frac{a^{x+\Delta x} - a^x}{\Delta x} = x_0 \frac{e^{xx_0+x_0\Delta x} - e^{xx_0}}{x_0\Delta x} \rightarrow x_0 e^{xx_0} = x_0 a^x.$$

Таким образом, число x_0 — это тангенс угла наклона касательной к графику функции a^x в точке $(0; 1)$. Число x_0 , определяемое равенством $e^{x_0} = a$, называется **натуральным логарифмом** числа a и обозначается как $\ln a$. Натуральный логарифм — это функция, обратная к e^x .

Во многих науках, в том числе в физике, часто встречается экспоненциальный (т. е. показательный) закон роста или убывания какой-либо величины. Например, такое понятие как период полураспада связано с показательной функцией. Точнее, для каждого радиоактивного вещества существует время T , за которое распадается половина ядер атомов этого вещества. Если мы обозначим количество оставшихся атомов за τ_n , то

$$\tau_{n+1} = \tau_n/2.$$

Отсюда легко получить, что $\tau_n = \tau_0/2^n$. Конечно, это закон статистический, и его точность тем хуже, чем меньше осталось атомов, но на гигантских количествах он работает очень точно.

И здесь мы видим показательную функцию с основанием $1/2$.

Во многих задачах встречается дифференциальное уравнение

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

которое говорит о том, что скорость роста величины $f(x)$ (то есть тангенс угла наклона ее графика) пропорциональна самой этой величине. Это — обобщение предыдущего уравнения (в конечных разностях) для непрерывной функции. Решением такого уравнения является показательная функция

$$f(x) = Ae^{kx},$$

где константа $A > 0$ обозначает начальное значение при $x = 0$. Например, рост банковского вклада с фиксированной процентной ставкой подчиняется такому закону (если считать, что проценты начисляются через один и тот же относительно короткий интервал времени).

Если производная функции положительна на некотором интервале, то сама функция на данном интервале строго возрастает, а если производная отрицательна, то сама функция строго убывает. Этот факт вполне очевиден из того же графика с касательными и секущими, который мы видели выше. Кроме того, из определения производной мы видим, что

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + o(x - x_0),$$

т. е. при малых отклонениях аргумента от x_0 разность значений $f(x) - f(x_0)$ имеет такой же знак, как производная в точке x_0 .

Проще всего, конечно, монотонность видна при интегрировании производной, но это мы оставим за рамками курса.

Покажем, что экспонента имеет следующие два представления.

Теорема 16.4. Для любого вещественного x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots, \quad e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Доказательство. Пусть

$$f(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Необходимо проверить, что $f(x)$ удовлетворяет условиям (16.4) и (16.5), а кроме того, $f'(0) = 1$.

Равенство $f(x + y) = f(x)f(y)$ проверяется следующим образом. Поскольку мы должны сравнить один бесконечный степенной ряд с произведением двух других рядов от двух переменных x и y , мы должны действовать так же, как при сравнении многочленов, а именно: сравнить коэффициенты при одинаковых степенях $x^n y^m$. Правда, при этом еще нужно обосновать возможность почленного перемножения рядов (для этого достаточно показать, что перемножаемые ряды сходятся абсолютно, но это мы оставим за рамками курса).

Ясно, что при перемножении $f(x)f(y)$ существует только один способ получить произведение степеней $x^n y^m$ — взять слагаемое с x^n из $f(x)$ и слагаемое с y^m из $f(y)$, поэтому в произведении $f(x)f(y)$ коэффициент перед $x^n y^m$ будет равен $\frac{1}{n!m!}$.

В случае $f(x + y)$ требуемая степень возникает только в слагаемом $(x + y)^{n+m} / (n+m)!$, которое по формуле бинома Ньютона раскладывается следующим образом:

$$\frac{(x + y)^{n+m}}{(n + m)!} = \frac{1}{(n + m)!} \sum_{k+j=n+m} x^k y^j \frac{(n + m)!}{k!j!},$$

где степень $x^n y^m$ можно получить единственным способом при $k = n$, $j = m$. Так что коэффициент перед $x^n y^m$ в $f(x + y)$ равен $\frac{1}{n!m!}$. Следовательно, $f(x + y) = f(x)f(y)$.

Покажем ее монотонность. Для случая $0 \leq x < y$ все очевидно, т. к. $x^n < y^n$, откуда и сумма ряда $f(x)$ меньше, чем сумма ряда $f(y)$. Кроме того, ясно, что $f(x) > 1$, когда $x > 0$.

Далее, из свойства сохранения операции мы видим, что $f(-x) = 1/f(x)$. Это значит, что для отрицательных x значение суммы ряда $f(x) < 1$, т. е. $f(x) < f(y)$, если $x < 0 < y$. Наконец, в случае $x < y < 0$ мы просто переходим к обратным величинам:

$$f(x) = \frac{1}{f(-x)} < \frac{1}{f(-y)} = f(y),$$

поскольку $f(-x) > f(-y)$ по доказанному выше (т. к. $0 < -y < -x$).

Итак, $f(x)$ монотонно возрастает. Следовательно, $f(x) = a^x$ при каком-то положительном a . Не вдаваясь в анализ степенных рядов, скажем, что ряд с факториальными коэффициентами настолько хорош, что с ним можно работать как с обычной суммой. В частности, $f(x) = 1 + x + o(x)$, т. к. сумма всех членов степени выше 1 имеет порядок малости сильнее, чем x . Но $y = 1 + x$ есть уравнение касательной к экспоненте, т. к. имеет наклон $\pi/4$. А это и означает, что $a = e$, т. е. $f(x) = e^x$.

Перейдем к доказательству соотношения, именуемого также **вторым замечательным пределом**:

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

Мы проведем его способом, типовым для получения предельных соотношений в математическом анализе. Он заключается в том, чтобы исследуемый ряд $f(x) = e^x$ и требуемый предел $l(x) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$ так удачно разделить на части, полагая $f(x) = K_n(x) + \varepsilon_n(x)$ и $l(x) = K_n(x) + \delta_n(x)$, что разность между «довесками» $\varepsilon_n(x) - \delta_n(x)$ будет стремиться к нулю. Тогда разность $f(x) - l(x)$ будет просто равна нулю, поскольку она не зависит от n . Итак, снова бином Ньютона:

$$\begin{aligned} \left(1 + \frac{x}{n}\right)^n &= 1 + x + \frac{n(n-1)}{2!} \frac{x^2}{n^2} + \dots + \frac{n!}{(n-k)!k!} \frac{x^k}{n^k} + \dots = \\ &= 1 + x + \frac{n-1}{n} \frac{x^2}{2!} + \dots + \frac{(n-1)\dots(n-k+1)}{n^{k-1}} \frac{x^k}{k!} + \dots = \\ &= 1 + x + \left(1 - \frac{1}{n}\right) \frac{x^2}{2!} + \dots + \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{x^k}{k!} + \dots \end{aligned}$$

Мы видим, что получается ряд, очень похожий на ряд экспоненты, который мы получили выше. Только, во-первых, это конечная сумма, а во-вторых, перед каждым слагаемым стоит коэффициент вида

$$\left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right).$$

Если мы зафиксируем k , то перед нами — обычное конечное произведение некоторых величин, которые с ростом n стремятся к 1. Следовательно, по свойствам пределов и все произведение стремится к 1. Это значит, что любой сколь угодно большой отрезок фиксированной длины k из разложения $\left(1 + \frac{x}{n}\right)^n$ стремится к такому же отрезку ряда экспоненты:

$$1 + x + \left(1 - \frac{1}{n}\right) \frac{x^2}{2!} + \dots + \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \frac{x^k}{k!} \rightarrow 1 + x + \frac{x^2}{2!} + \dots + \frac{x^k}{k!}$$

при $n \rightarrow \infty$. Сумма, стоящая слева, и есть наш $K_n(x)$, «основная» часть изучаемого ряда $f(x) = e^x$. Число k хоть мы и считаем фиксированным, но на самом деле его можно считать зависящим от n , только $k = k(n)$ растет так медленно, как нам нужно для получения данного предельного перехода. Остается показать, что все остальное можно сделать сколь угодно малым выбором достаточно больших n (и k).

Посмотрим на итоговую разность:

$$\left| \left(1 + \frac{x}{n}\right)^n - e^x \right| \leq \left| \sum_{j=0}^k \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{j-1}{n}\right) \frac{x^j}{j!} - \sum_{j=0}^k \frac{x^j}{j!} \right| + \left| \left(\frac{x^{k+1}}{(k+1)!} - \frac{(n-1) \dots (n-k)}{n^k} \frac{x^{k+1}}{(k+1)!} \right) + \dots + \left(\frac{x^n}{n!} - \frac{x^n}{n!} \right) + \frac{x^{n+1}}{(n+1)!} + \dots \right|.$$

Теперь заметим, что первый модуль можно сделать меньше произвольного $\varepsilon > 0$ (при любом сколь угодно большом k), а во втором модуле все слагаемые со знаком «минус» по модулю строго меньше слагаемых со знаком «плюс», поскольку

$$\frac{(n-1) \dots (n-k)}{n^k} \leq 1.$$

А это значит, что второй модуль можно оценить «хвостом» ряда экспоненты, т. е. выражением

$$\begin{aligned} \frac{|x|^{k+1}}{(k+1)!} + \dots + \frac{|x|^n}{n!} + \dots &= \\ &= |x|^{k+1} \left(\frac{|x|^0}{(k+1)!} + \dots + \frac{|x|^{n-k-1}}{(n-k-1)!(n-k) \dots n} + \dots \right) \leq \\ &\leq |x|^{k+1} \left(\frac{|x|^0}{(k+1)!} + \dots + \frac{|x|^{n-k-1}}{(n-k-1)!(k+1)!} + \dots \right), \end{aligned}$$

где мы воспользовались тем, что $(n-k) \dots n \geq (k+1)!$. Отсюда следует, что

$$\frac{|x|^{k+1}}{(k+1)!} + \dots + \frac{|x|^n}{n!} + \dots \leq \frac{|x|^{k+1}}{(k+1)!} e^{|x|}, \quad (16.10)$$

а эта величина, как мы уже ранее доказывали, стремится к нулю с ростом k . Следовательно, хвост экспоненты можно сделать меньше ε выбором достаточно большого k .

Таким образом, разность $\left| \left(1 + \frac{x}{n}\right)^n - e^x \right|$ оценивается величиной 2ε , причем сначала выбирается достаточно большое k , при котором хвост ряда экспоненты меньше ε , а затем уже для этого k выбирается достаточно большое $n \geq k$, при котором начальный отрезок ряда экспоненты

отличается от аналогичного отрезка ряда бинома Ньютона меньше, чем на ε . В итоге вся разность может быть сделана сколь угодно малой, а значит, имеет место предельный переход:

$$\left(1 + \frac{x}{n}\right)^n \rightarrow e^x$$

при $n \rightarrow \infty$. □

16.3. Комплексная экспонента

Комплéксные числа мы теперь рассматриваем в их полном объеме, т. е. как множество векторов $z = (x, y) = x + yi$ с вещественными координатами, подчиненных арифметическим операциям

$$(x+iy)+(x'+y'i) = (x+x')+(y+y')i, \quad (x+iy)(x'+y'i) = (xx'-yy')+(xy'+x'y)i.$$

Множество всех комплексных чисел обозначается символом \mathbb{C} .

Модулем комплексного числа $x + yi$ называется $|x + yi| = \sqrt{x^2 + y^2}$.

Так же, как в \mathbb{R} , мы можем рассматривать последовательности комплексных чисел, которые представляют собой две упакованные в одну последовательности чисел действительных:

$$\{z_n\} = \{x_n + y_n i\}.$$

Соответственно, точно так же, с использованием модуля комплексного числа, определяется предел последовательности:

$$z = \lim_{n \rightarrow \infty} z_n \Leftrightarrow \forall \varepsilon > 0 \exists N \forall n > N |z - z_n| < \varepsilon.$$

Термин *ε -окрестность* на плоскости приобретает уже привычное бытовое понимание. ε -окрестностью точки z_0 называется круг радиуса ε с центром в точке z_0 , т. е. множество

$$B_\varepsilon(z_0) = \{z \mid |z - z_0| < \varepsilon\}.$$

Таким образом, сходимость $z_n \rightarrow z$ означает, что в любой сколь угодно малой окрестности точки z находятся почти все члены последовательности $\{z_n\}$.

Сходимость последовательности $\{z_n = x_n + y_n i\}$ к какому-то пределу $z = x + yi$ равносильна сходимости последовательностей $\{x_n\}$ и $\{y_n\}$, соответственно, к точкам x и y .

Действительно, если $z_n \rightarrow z$, то для любого $\varepsilon > 0$ найдется номер N такой, что для $n > N$ выполняется неравенство $|z_n - z| < \varepsilon$, т. е.

$\sqrt{(x_n - x)^2 + (y_n - y)^2} < \varepsilon$. Откуда следует, что для тех же ε и n имеем $|x_n - x| < \varepsilon$ и $|y_n - y| < \varepsilon$, что и означает сходимость $x_n \rightarrow x$ и $y_n \rightarrow y$.

Если же, наоборот, мы имеем сходимость $x_n \rightarrow x$ и $y_n \rightarrow y$, то для любого $\varepsilon > 0$ найдутся такие номера N_1 и N_2 , что для всех $n_1 > N_1$ и всех $n_2 > N_2$ имеют место неравенства $|x_{n_1} - x| < \varepsilon$ и $|y_{n_2} - y| < \varepsilon$. Тогда для того же ε и для всех $n > \max\{N_1, N_2\}$ мы получим, что

$$|z_n - z| = \sqrt{(x_n - x)^2 + (y_n - y)^2} < \sqrt{2\varepsilon^2} = \varepsilon\sqrt{2}, \text{ т. е. } z_n \rightarrow z.$$

Поэтому все свойства полноты (непрерывности) \mathbb{R} наследуются множеством комплексных чисел за одним важным исключением, а именно, за исключением свойства упорядоченности.

Так, всякая фундаментальная последовательность, т. е. такая $\{z_n\}$, что

$$\forall \varepsilon > 0 \exists N \forall n, m > N |z_n - z_m| < \varepsilon,$$

имеет предел.

Полнота (непрерывность) \mathbb{C} означает, что в нем не появляется никаких новых «дыр», которые можно было бы задать уравнениями с целыми коэффициентами. Более того, даже если коэффициенты многочлена являются комплексными числами, никаких новых расширений с помощью корней таких многочленов мы не получим.

Более точно, для поля \mathbb{C} справедлива следующая теорема.

Теорема 16.5 (основная теорема алгебры). *Всякий многочлен степени $n \geq 1$ над \mathbb{C} раскладывается в произведение линейных множителей, и это разложение однозначно. То есть для всякого многочлена $f(z) = a_0 + a_1z + \dots + a_nz^n$ имеет место разложение*

$$f(z) = a_n(z - z_1)(z - z_2) \dots (z - z_n),$$

где корни z_1, \dots, z_n могут быть кратными, т. е. повторяться.

Из этой теоремы следует также, что любой многочлен с вещественными коэффициентами (который можно рассматривать как частный случай многочлена с комплексными коэффициентами) раскладывается в произведение многочленов первой и второй степени. Дело тут в том, что если $f(z) = 0$, то и $f(\bar{z}) = 0$, т. е. комплексные корни (не являющиеся вещественными) всегда идут парой вместе со своим сопряженным. Ну а произведение $(z - z_1)(z - \bar{z}_1)$ имеет уже строго вещественные коэффициенты. Поэтому такой квадратный трехчлен входит в разложение исходного многочлена. Если же корень z_1 сам себе сопряжен, то он является вещественным числом, и потому сам двучлен $(z - z_1)$ входит в разложение исходного многочлена.

Итак, \mathbb{C} — это алгебраически замкнутое полное поле. Оно не является упорядоченным полем ни при каком линейном порядке (т. к. в упорядоченном поле квадрат отрицательного числа всегда есть положительное число, т. е. невозможна мнимая единица). Поле \mathbb{C} есть двумерное расширение поля \mathbb{R} и его можно записать как $\mathbb{R}[i]$.

Мощность множества \mathbb{C} есть континуум.

Докажем этот факт в упрощенной форме, а именно, покажем, что квадрат $[0; 1) \times [0; 1)$ равномощен отрезку $[0; 1)$. Возьмем произвольную точку (x, y) из квадрата и запишем ее координаты двоичной последовательностью без хвоста единиц:

$$x = 0.x_1x_2x_3 \dots x_n \dots, \quad y = 0.y_1y_2y_3 \dots y_n \dots,$$

где $x_n, y_n \in \{0, 1\}$. Построим точку z , чередуя двоичные цифры из исходных представлений:

$$z = 0.x_1y_1x_2y_2x_3y_3 \dots x_ny_n \dots$$

Такое соответствие точек квадрата и точек отрезка является инъекцией, поскольку (x, y) однозначно восстанавливается по z , но не является биекцией, т. к. точки вида $z = 0.01010101010101$ не имеют соответствия в квадрате.

С другой стороны, очевидное вложение $z \mapsto (z, 0)$ точек отрезка в квадрат также является инъекцией. Следовательно, по теореме Кантора — Бернштейна квадрат и отрезок равномощны.

Расширить доказательство на случай \mathbb{C} и \mathbb{R} можно с помощью преобразований, которые мы ранее уже использовали при установлении равномощности $[0; 1)$ и \mathbb{R} . В случае \mathbb{C} их нужно применить к каждой координате по отдельности.

Введем некоторые вспомогательные определения, связанные с топологией. Ранее мы уже говорили, что круг без границы с центром z_0 и радиусом r , обозначаемый за $B_r(z_0) = \{z \mid |z - z_0| < r\}$, называется r -окрестностью точки z_0 . Обобщим это понятие.

Подмножество $D \subseteq \mathbb{C}$ называется **открытым**, если любая его точка содержится в нем вместе с некоторой своей окрестностью ($\forall z \in D \exists \varepsilon > 0: B_\varepsilon(z) \subseteq D$). В частности, всякая r -окрестность является открытым множеством, пустое множество и вся плоскость \mathbb{C} являются открытыми множествами. Совокупность всех открытых подмножеств \mathbb{C} называется **топологией** \mathbb{C} .

Подмножество $D \subseteq \mathbb{C}$ называется **связным**, если не верно, что оно содержится в объединении двух открытых непересекающихся множеств и его пересечение с каждым из них не пусто (см. рис. 16.7), т. е. не существует таких открытых множеств U_1 и U_2 , что

$$(U_1 \cap U_2 = \emptyset) \wedge (D \subseteq U_1 \cup U_2) \wedge (D \cap U_1 \neq \emptyset) \wedge (D \cap U_2 \neq \emptyset).$$

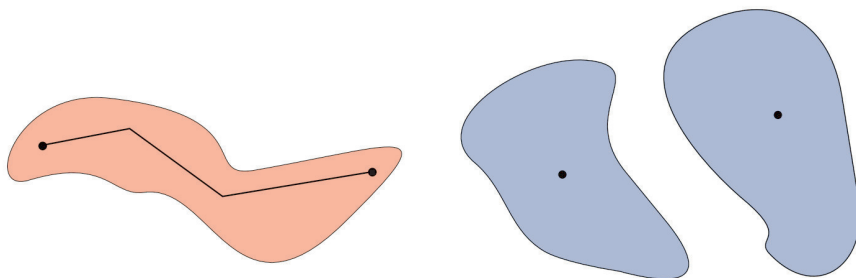


Рис. 16.7. Слева — связное, справа — несвязное множество.

В случае непустого открытого множества D его связность равносильна тому, что любые две точки D можно соединить ломаной, целиком лежащей в D .

Подмножество $D \subseteq \mathbb{C}$ называется **областью**, если оно открытое и связное.

Аналогично вещественному случаю вводится понятие предела функции комплексного переменного:

$$w = \lim_{z \rightarrow z_0} f(z), \text{ если } \forall \varepsilon > 0 \exists \delta > 0 : \forall z \in \dot{B}_\delta(z_0) \cap D : |f(z) - w| < \varepsilon,$$

где D — область определения функции f , $\dot{B}_\delta(z_0) = \{z \mid 0 < |z - z_0| < \delta\}$ — проколота δ -окрестность точки z_0 .

Функция $f : D \rightarrow \mathbb{C}$, где D — непустое подмножество \mathbb{C} , называется **непрерывной** в точке $z_0 \in D$, если

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall z \in D : (|z - z_0| < \delta \rightarrow |f(z) - f(z_0)| < \varepsilon).$$

В терминах пределов это означает, что $\lim_{z \rightarrow z_0} f(z)$ существует и равен значению $f(z_0)$.

Как и в случае вещественной прямой, можно доказать, что непрерывность f эквивалентна тому, что для любой последовательности $\{z_n\} \subseteq D$, если имеет место сходимость $z_n \rightarrow z \in D$, то $f(z_n) \rightarrow f(z)$.

Богатый инструментарий дает понятие комплексной производной. Определяется она точно так же:

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

только приращение Δz здесь — комплексное, и это приводит ко многим хорошим свойствам комплексного дифференцирования.

Функция $f(z)$ называется **регулярной** в области $D \subseteq \mathbb{C}$, если она дифференцируема в каждой ее точке. Регулярная на всей комплексной плоскости функция называется **целой**.

В отличие от вещественного случая, если комплексная функция дифференцируема в некоторой области, то у нее существуют все производные высших порядков в этой области. Это обеспечивает возможность представить регулярную функцию (в области регулярности) в виде степенного ряда:

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + f''(z_0)\frac{(z - z_0)^2}{2} + f'''(z_0)\frac{(z - z_0)^3}{3!} + \dots$$

Такой ряд называется **рядом Тейлора** (как в комплексном, так и вещественном случае).

Функцию, для которой мы знаем разложение в ряд Тейлора в вещественном случае на некотором непустом интервале, можно продлить до функции комплексного переменного, заданной на некоторой области, содержащей данный интервал, в которой она регулярна.

Теорема 16.6 (единственности). Пусть функции f и g регулярны в области D , и существует последовательность $\{z_n\} \subset D$, сходящаяся к точке $w \in D$, на которой эти функции совпадают: $f(z_n) = g(z_n)$. Тогда $f(z) = g(z)$ на всей области D .

Доказательство данной теоремы мы оставляем за рамками нашего курса (подробнее см. в [8]). У теоремы единственности есть ряд замечательных следствий, таких, например, как:

Следствие 16.1. Отличная от тождественного нуля регулярная в области D функция имеет лишь конечное число нулей в любом замкнутом круге $\{z \mid |z - z_0| \leq r\} \subset D$.

Следствие 16.2. Если две регулярные в области D функции совпадают на некоторой кривой, лежащей в замкнутом круге $\{z \mid |z - z_0| \leq r\} \subset D$, то они совпадают на всей области D .

Следствие 16.2 позволяет сделать следующий вывод. Пусть функции $f(z)$ и $g(z)$ регулярны в $B_r(0)$ и совпадают на вещественном интервале $(-r; r)$, где $r > 0$, тогда они совпадают на всей области $B_r(0)$. В частности, если f и g — целые функции, совпадающие на каком-то интервале $(-r; r)$, то они совпадают на всей плоскости \mathbb{C} .

Для нас это означает следующее. Возьмем функцию e^x , у которой ряд Тейлора в точке $x = 0$ имеет вид

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots$$

Затем определим функцию комплексного переменного с помощью точно такого же ряда:

$$e^z = 1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \dots, \quad (16.11)$$

которую будем называть **комплексной экспонентой**.

Важное замечание: хотя мы записываем комплексную экспоненту как показательную функцию e^z , на самом деле мы всего лишь задали некоторую комплексную функцию в каждой точке z , поэтому правильнее было бы обозначать ее как $\exp(z)$. Возведение же комплексного числа в комплексную степень является предметом более длительного и тщательного изучения.

Нетрудно видеть, что комплексная экспонента сходится в каждой точке $z \in \mathbb{C}$. Для этого достаточно оценить «хвост» ряда Тейлора (16.11) с помощью полученной ранее оценки (16.10), поскольку $|z|$ уже будет вещественным положительным числом. И тогда очень просто доказать, что последовательность частных сумм

$$z_n = 1 + z + \frac{z^2}{2!} + \dots + \frac{z^n}{n!}$$

является фундаментальной, т. к. мы будем иметь оценку при $n < m$

$$|z_n - z_m| \leq \frac{|z|^n}{n!} + \dots + \frac{|z|^{m-1}}{(m-1)!} \leq \frac{|z|^n}{n!} e^{|z|} \rightarrow 0$$

при $n \rightarrow \infty$. И далее в силу полноты \mathbb{C} мы заключаем, что существует предел $\{z_n\}$, который по определению и есть e^z .

На рис. 16.8 построено несколько шагов приближения величины $\exp(3 + 4i)$ с помощью частных сумм z_n ряда Тейлора.

Нетрудно проверить также, что производная комплексной экспоненты e^z существует и равна e^z в каждой точке $z \in \mathbb{C}$.

Таким образом функция e^z является целой и, кроме того, совпадает с e^x на всей вещественной оси. Следовательно, e^z является единственным возможным продолжением вещественной экспоненты до регулярной на \mathbb{C} функции комплексного переменного.

С помощью того же самого ряда Тейлора (16.11) экспонента может быть определена для многих числовых структур, однако для сохранения свойства $e^{x+y} = e^x e^y$ требуется коммутативность умножения, что не всегда выполняется. Например, матричная экспонента, когда вместо z мы подставляем квадратную матрицу, тоже определяется по формуле (16.11), но равенство $e^{x+y} = e^x e^y$ не всегда выполняется.

Для примера достаточно взять $X = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ и $Y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Обозначая за E единичную матрицу $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, получим:

$$e^X = E + (e - 1)X, \quad e^Y = E + Y, \\ e^X e^Y = E + (e - 1)X + eY, \quad e^{X+Y} = E + (e - 1)(X + Y).$$

Таким образом, $e^{i\varphi} = \lim_{n \rightarrow \infty} w_n^n$ и $z_n - w_n = \gamma_n/n$.

Остается показать, что на самом деле $e^{i\varphi} = \lim_{n \rightarrow \infty} z_n^n$. Для этого рассмотрим величину $\Delta_n = |w_n^n - z_n^n|$ и покажем, что она стремится к нулю.

Несложно доказать следующее неравенство:

$$|z^n - w^n| \leq n|z - w| \max\{|z|^{n-1}, |w|^{n-1}\}.$$

Действительно,

$$(z^n - w^n) = (z - w)(z^{n-1} + z^{n-2}w + \dots + zw^{n-2} + w^{n-1}),$$

откуда получаем требуемое неравенство, оценивая модуль суммы с помощью неравенства треугольника. Тогда

$$\Delta_n \leq n \frac{\gamma_n}{n} \max\{|z_n|^n, |w_n|^n\}.$$

Но $|z_n| = 1$, откуда $|z|^n = 1$. В то же время

$$|w_n|^n \leq \left| 1 + \frac{|\varphi|}{n} \right|^n \rightarrow e^{|\varphi|},$$

т. е. $|w_n|^n$ ограничено некоторой константой. Следовательно, $\Delta_n \rightarrow 0$ и

$$e^{i\varphi} = \lim_{n \rightarrow \infty} z_n^n = \lim_{n \rightarrow \infty} (\cos(\varphi/n) + i \sin(\varphi/n))^n.$$

Но последнее выражение под знаком предела — это произведение n одинаковых чисел, лежащих на единичной окружности. А как мы знаем, при умножении таких чисел их аргументы складываются, а модуль остается равным единице, поэтому $(\cos(\varphi/n) + i \sin(\varphi/n))^n = \cos(\varphi) + i \sin(\varphi)$, откуда окончательно получаем, что

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

Это равенство называется **формулой Эйлера**.

В частности, отсюда следует знаменитое тождество Эйлера, связывающее сразу 5 фундаментальных математических констант:

$$e^{i\pi} + 1 = 0.$$

Из формулы Эйлера легко получить разложения в ряд Тейлора для синуса и косинуса. Действительно, посмотрим на ряд

$$e^{ix} = 1 + ix - \frac{x^2}{2} - i \frac{x^3}{3!} + \frac{x^4}{4!} + i \frac{x^5}{5!} - \frac{x^6}{6!} - i \frac{x^7}{7!} + \dots$$

Поскольку $e^{ix} = \cos x + i \sin x$, собирая для косинуса все вещественные слагаемые ряда, а для синуса — мнимые, мы получим:

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots, \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots \end{aligned}$$

Упражнения

Обязательные упражнения

16.1° Может ли последовательность иметь более одного предела?

16.2° Найдите предел $\{x_n\}$, если он есть:

- a) $x_n = 1 + (-0.1)^n$;
b) $x_n = \frac{n}{n+1}$;
c) $x_n = (-1)^n$;
d) $x_n = \frac{2^n - 1}{2^n + 1}$;
e) $x_n = 1 + 0.1 + \dots + (0.1)^n$;
f) $x_n = \frac{1 + 3 + \dots + 3^n}{5^n}$;
g) $x_n = \sqrt{n+1} - \sqrt{n}$.

16.3° Запишите без отрицания: a) «число a — не предел $\{x_n\}$ »; b) « $\{x_n\}$ не имеет предела».

16.4° Пусть $\lim_{n \rightarrow \infty} x_n > 0$. Верно ли, что a) $x_n > 0$ при достаточно больших n ; b) $1/x_n$ ограничена (если определена)?

16.5° Найдите предел при $n \rightarrow \infty$:

- a) $1 + q + \dots + q^n$, где $|q| < 1$; b) $\frac{n^2 - n + 1}{n^2}$; c) $\sqrt[n]{2}$; d) $\frac{n^{50}}{2^n}$; e) $\sqrt[n]{n}$.

16.6° Может ли последовательность без наименьшего и наибольшего членов иметь предел?

16.7° a) Последовательность $\{x_n\}$ имеет предел. Докажите, что $\{x_{n+1} - x_n\}$ бесконечно малая (т. е. стремится к нулю). b) Верно ли обратное?

16.8° Последовательность $\{x_n\}$ положительна, а последовательность $\{x_{n+1}/x_n\}$ имеет пределом некоторое число, меньшее 1. Докажите, что $\{x_n\}$ бесконечно малая.

16.9° Найдите:

- a) $\lim_{n \rightarrow \infty} \frac{4n^2}{n^2 + n + 1}$; b) $\lim_{n \rightarrow \infty} \frac{n^2 + 2n - 2}{n^3 + n}$; c) $\lim_{n \rightarrow \infty} \frac{n^9 - n^4 + 1}{2n^9 + 7n - 5}$; d) $\lim_{n \rightarrow \infty} \frac{\binom{50}{n}}{n^{50}}$.

16.10° Найдите ошибку в следующем рассуждении. «Пусть $x_n = (n - 1)/n$. Тогда

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} (1 - 1/n) = 1.$$

С другой стороны,

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} (1/n) \lim_{n \rightarrow \infty} (n-1) = 0 \cdot \lim_{n \rightarrow \infty} (n-1) = 0.$$

Отсюда $0 = 1$ ».

16.11° Пусть $\lim_{n \rightarrow \infty} x_n = a$, $\lim_{n \rightarrow \infty} y_n = b$. Найдите: **а)** $\lim_{n \rightarrow \infty} (x_n \pm y_n)$; **б)** $\lim_{n \rightarrow \infty} (x_n y_n)$. **с)** Что можно сказать о $\lim_{n \rightarrow \infty} x_n / y_n$?

16.12° Пусть $\lim_{n \rightarrow \infty} x_n = a$, $\lim_{n \rightarrow \infty} y_n = b$ и $x_n > y_n$ при $n \in N$. Верно ли, что: **а)** $a > b$; **б)** $a \geq b$?

16.13° Пусть $\lim_{n \rightarrow \infty} x_n = 1$. Найдите: **а)** $\lim_{n \rightarrow \infty} \frac{x_n^2}{7}$; **б)** $\lim_{n \rightarrow \infty} \frac{x_n^2 + x_n - 2}{x_n - 1}$; **с)** $\lim_{n \rightarrow \infty} \sqrt{x_n}$.

16.14° В два сосуда разлили (не поровну) 1 л воды. Из 1-го сосуда перелили половину имеющейся в нем воды во 2-ой, затем из 2-го перелили половину оказавшейся в нем воды в 1-ый, снова из 1-го перелили половину во 2-ой и т. д. Сколько воды (с точностью до 1 мл) будет в 1-ом сосуде после 50 переливаний?

16.15° В банк кладут 1000 рублей. В каком случае спустя 10 лет получат больше денег: если банк начисляет 5% от имеющейся суммы раз в год или если он начисляет $(5/12)\%$ раз в месяц?

16.16° Докажите, что при всех натуральных n и при всех неотрицательных a выполнено неравенство $(1+a)^n \geq 1+an+a^2n(n-1)/2$.

16.17° Укажите такое целое $n > 1$, что **а)** $1.001n > 10^5$; **б)** $0.999^n < 10^{-5}$; **с)** $\sqrt[n]{n} < 1.001$.

16.18° **а)** Докажите: $n^n > 106 \cdot n!$ при $n \gg 0$. **б)** Можно ли заменить 10^6 на любое другое число?

16.19° **а)** Докажите, что $0.001n^2 > 100n + 57$ при $n \gg 0$. **б)** Число C — любое, n и m — натуральные, причем $n > m$. Докажите, что $x^n > Cx^m$ при $x \gg 0$. **с)** Дан многочлен $P(x) = p_k x^k + p_{k-1} x^{k-1} + \dots + p_1 x + p_0$, где $p_k > 0$. Верно ли, что $P(x) > 0$ при $x \gg 0$?

16.20° Докажите, что если $a > 1$ и C — любое, то **а)** $a^n > C$ при $n \gg 0$; **б)** $a^n > n$ при $n \gg 0$.

16.21° Верно ли, что **а)** $1,01^n > 100n$ при $n \gg 0$; **б)** если $a > 1$, $C > 0$, то $a^n > Cn$ при $n \gg 0$?

16.22° а) Пусть $q > 1$ и последовательность положительных чисел $\{x_n\}$ такова, что $x_{n+1}/x_n > q$ при $n \gg 0$. Докажите, что $x_n > 1$ при $n \gg 0$.
 б) Верно ли утверждение предыдущего пункта, если $q = 1$?

16.23° Докажите, что при натуральных $n \gg 0$: а) $2^n > n^{100}$; б) если $a > 1$ и $k \in \mathbb{N}$ фиксировано, то $a^n > n^k$.

16.24° Докажите, что для любого a неравенство $n! > a^n$ выполнено при $n \gg 0$.

16.25° а) Докажите, что $\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n} \geq \frac{1}{2}$ при любом $n \in \mathbb{N}$.
 б) Для любого ли числа C найдется такое $n \in \mathbb{N}$, что будет выполнено неравенство $1 + \frac{1}{2} + \dots + \frac{1}{n} \geq C$?
 в) Тот же вопрос для неравенства $1 + \frac{1}{2^2} + \dots + \frac{1}{n^2} \geq C$.

16.26° Докажите для $n \in \mathbb{N}$:

а) $a^{n+1}/b^n \geq (n+1)a - nb$, если $a, b > 0$;

б) $(1 + \frac{1}{n+1})^{n+1} \geq (1 + \frac{1}{n})^n$;

в) $(1 + \frac{1}{n-1})^n \geq (1 + \frac{1}{n})^{n+1}$;

г) $2 \leq (1 + \frac{1}{n})^n \leq 4$;

д) $(1 + \frac{1}{n})^n \leq 1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!}$;

е) [неравенство Коши] $(a_1 + \dots + a_n)/n \geq \sqrt[n]{a_1 \dots a_n}$, если числа a_1, \dots, a_n положительны;

ж) $(n/4)^n \leq n! \leq ((n+1)/2)^n$.

16.27° Для последовательности $\{x_n\}$ найдите по данному числу $\varepsilon > 0$ какой-нибудь номер $N = N(\varepsilon)$, начиная с которого верно, что $|x_n| < \varepsilon$, если:

а) $x_n = \frac{1}{n}$; б) $x_n = \frac{2}{n^3}$; в) $x_n = \frac{\sin n}{n}$; г) $x_n = \frac{1}{2n^2 + n}$.

16.28° Доказать неравенства:

а) $1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \dots + \frac{1}{\sqrt{n}} > \sqrt{n}$ ($n \geq 2$);

б) $n^{n+1} > (n+1)^n$ ($n \geq 3$);

в) $\left| \sin \left(\sum_{k=1}^n x_k \right) \right| \leq \sum_{k=1}^n \sin(x_k)$ ($0 \leq x_k \leq \pi$, $k = 1, 2, \dots, n$);

г) $(2n)! < 2^{2n}(n!)^2$.

16.29° Доказать существование предела последовательности $\{x_n\}$, если

а) $x_n = p_0 + \frac{p_1}{10} + \dots + \frac{p_n}{10^n}$, где p_1, p_2, \dots — целые неотрицательные числа, причем $p_k \leq 9$ при $k \geq 1$;

б) $x_n = \frac{10}{1} \cdot \frac{11}{3} \cdot \frac{n+9}{2n-1}$;

в) $x_n = \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{4}\right) \dots \left(1 - \frac{1}{2^n}\right)$;

г) $x_n = \left(1 + \frac{1}{2}\right) \left(1 + \frac{1}{4}\right) \dots \left(1 + \frac{1}{2^n}\right)$;

е) $x_1 = \sqrt{2}, x_2 = \sqrt{2 + \sqrt{2}}, \dots, x_n = \underbrace{\sqrt{2 + \sqrt{2 + \dots + \sqrt{2}}}}_{n \text{ корней}}$.

16.30° Какими арифметическими свойствами обладают символы $o()$ и $O()$?

16.31° Доказать, что если $x_n > 0$ для всех n , то $x_n = o(1)$ тогда и только тогда, когда $x_n^{-1} \rightarrow \infty$. Почему требование $x_n > 0$ здесь является существенным?

16.32° Последовательности $\{x_n\}$ и $\{y_n\}$ бесконечно малые, а последовательность $\{z_n\}$ такова, что $x_n \leq z_n \leq y_n$ при всех натуральных n . Докажите, что последовательность $\{z_n\}$ бесконечно малая.

16.33° Известно, что последовательности $\{x_n\}$ и $\{y_n\}$ бесконечно малые. Составим последовательность $x_1, y_1, x_2, y_2, x_3, y_3, \dots$. Будет ли эта последовательность бесконечно малой?

16.34° Является ли бесконечно малой последовательность:

а) $x_n = \frac{\cos n - 0.5^n}{n + 7}$; б) $y_n = \frac{3^n + 4^n}{2^n + 5^n}$?

16.35° Дана последовательность $\{x_n\}$ с положительными членами. Верно ли, что $\{x_n\}$ бесконечно малая тогда и только тогда, когда последовательность $\{\sqrt{x_n}\}$ бесконечно малая?

16.36° Даны две последовательности: $\{x_n\}$ — бесконечно малая, $\{y_n\}$ — ограниченная. Докажите, что: а) $\{x_n + y_n\}$ — ограниченная последовательность; б) $\{x_n y_n\}$ — бесконечно малая последовательность.

16.37° В бесконечно малой последовательности $\{x_n\}$ переставили члены (то есть взяли какое-то взаимно однозначное соответствие $f: \mathbb{N} \rightarrow \mathbb{N}$ и получили новую последовательность $\{y_n\}$, где $y_n = x_{f(n)}$ для всех $n \in \mathbb{N}$). Обязательно ли полученная последовательность будет бесконечно малой?

16.38° Последовательность состоит из положительных членов, причем сумма любого количества ее членов не превосходит 1. Докажите, что эта последовательность бесконечно малая.

16.39° Последовательность не является ограниченной. Обязательно ли она бесконечно большая?

16.40° Запишите без отрицания: $\{x_n\}$ не является а) бесконечно малой; б) бесконечно большой.

16.41° Какие из последовательностей ограничены, какие — бесконечно малые, а какие — бесконечно большие: а) $x_n = (1.1)^n$; б) $y_n = (0.9)^n$; в) $z_n = \sqrt{n^3 + n} - \sqrt{n^3}$; д) $t_n = \sqrt[n]{n!}$; е) $s_n = \frac{n^5 + 1}{n^4 + n^2}$?

16.42° Одна последовательность бесконечно большая, а другая бесконечно малая. Что можно сказать а) о сумме; б) об отношении; в) о произведении этих последовательностей?

16.43° Любую ли последовательность можно представить как отношение а) двух ограниченных; б) двух бесконечно малых последовательностей?

16.44° Пусть $\{x_n\}$ — последовательность, для которой $x_n/x_{n-1} < q$ при всех n . При каких q такая последовательность всегда является бесконечно малой?

16.45° Есть ли последовательность, члены которой найдутся в любом интервале числовой оси?

Дополнительные упражнения

16.46* Пусть $\{x_n\}$ — последовательность положительных чисел, стремящаяся к a . Докажите, что тогда для любого $k \in \mathbb{N}$ существует предел последовательности $\{\sqrt[k]{x_n}\}$, равный $\sqrt[k]{a}$.

16.47* Пусть $a, b > 0$ и $n \in \mathbb{N}$. Докажите, что выполнено неравенство $a^{n+1}/b^n \geq (n+1)a - nb$.

16.48* Докажите, что: а) $e_n = (1 + \frac{1}{n})^n$ монотонно возрастает; б) $E_n = (1 + \frac{1}{n})^{n+1}$ монотонно убывает; в) $\lim_{n \rightarrow \infty} e_n = \lim_{n \rightarrow \infty} E_n$; д) выполнено неравенство $2.25 < e < 3.375$.

Найдите такое n , $|e - e_n| < 10^{-6}$.

16.49* Докажите, что:

а) $\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^n = \frac{1}{e}$;

б) $\lim_{n \rightarrow \infty} \left(1 + \frac{k}{n}\right)^n = e^k$, если $k \in \mathbb{Z}$;

с) $\lim_{n \rightarrow \infty} \left(1 + \frac{r}{n}\right)^n = e^r$, если $r \in \mathbb{Q}$;

д) одна из последовательностей $\left(1 + \frac{r}{n}\right)^n$ и $\left(1 + \frac{r}{n}\right)^{n+r}$ монотонно возрастает, другая — монотонно убывает, и пределы обеих последовательностей равны e^r .

16.50* Обозначим сумму $1 + \frac{1}{1!} + \dots + \frac{1}{n!}$ через s_n , а число $\left(1 + \frac{1}{n}\right)^n$ — по-прежнему через e_n .

а) Докажите, что для любого натурального n выполнено неравенство $e_n \leq s_n$.

б) Зафиксируем натуральное число N и рассмотрим любое натуральное $n > N$. Раскроем скобки в выражении e_n по биному Ньютона и оставим лишь первые $N + 1$ слагаемых. Докажите, что предел при $n \rightarrow \infty$ полученной таким образом последовательности с фиксированным параметром N равен s_N .

с) Докажите, что $s_N \leq e$.

д) Докажите, что $\lim_{n \rightarrow \infty} s_n = e$.

е) Докажите, что $\sum_{k=m}^n \frac{1}{k!} \leq \frac{1}{m!} \frac{1}{1 - \frac{1}{m+1}}$.

ф) Найдется ли $n < 100$ такое, что $|e - s_n| < 10^{-6}$?

16.51* Докажите, что $\lim_{n \rightarrow \infty} n(e^n - 1) = 1$.

16.52* Пусть $\{x_n\}$ — последовательность положительных чисел, стремящаяся к a . Докажите, что тогда для любого $k \in \mathbb{N}$ существует предел последовательности $\{\sqrt[k]{x_n}\}$, равный $\sqrt[k]{a}$.

16.53* Доказать неравенства

$$\left(\frac{n}{e}\right)^n < n! < e \left(\frac{n}{2}\right)^n.$$

16.54* Доказать неравенства:

а) $\frac{1}{n+1} < \ln\left(1 + \frac{1}{n}\right) < \frac{1}{n}$, где $n > 0$ — целое число;

б) $1 + \alpha < e^\alpha$, где $\alpha \neq 0$ — вещественное число.

16.55* Доказать, что

$$\lim_{n \rightarrow \infty} n \left(a^{1/n} - 1\right) = \ln a, \quad (a > 0).$$

16.56* Проверить, будут ли непрерывны следующие функции на множестве M , если f и g непрерывны на том же множестве: а) $f+g$; б) $f-g$; в) kf (k — любое число); д) fg ; е) f/g ; ф) $|f|$; г) f^2 .

16.57* Продлить по непрерывности функции в точке $x = 0$, если это возможно:

а) $\frac{1 - \cos(x)}{x^2}$; б) $\frac{\sqrt{1+x}-1}{\sqrt[3]{1+x}-1}$; в) $\frac{\operatorname{tg} 2x}{2x}$; д) $\sin x \sin \frac{1}{x}$;
е) $(1+x)^{1/x}$; ф) $\frac{1}{x^2} e^{-1/x^2}$; г) x^x ; х) $x \ln^2 x$.

16.58* Доказать, что если функция $f(x)$ непрерывна, то функция

$$f_c(x) = \begin{cases} -c, & f(x) < -c, \\ f(x), & |f(x)| \leq c, \\ c, & f(x) > c, \end{cases}$$

где c — любое положительное число, тоже непрерывна.

16.59* Доказать, что если функция $f(x)$ непрерывна на отрезке $[a; b]$, то таковы же и функции:

$$m(x) = \inf_{a \leq t \leq x} f(t), \quad M(x) = \sup_{a \leq t \leq x} f(t).$$

16.60* Доказать, что если функции $f(x)$ и $g(x)$ непрерывны, то

$$m(x) = \min\{f(x), g(x)\}, \quad M(x) = \max\{f(x), g(x)\}$$

также непрерывны.

16.61* Доказать, что если функция $f(x)$ определена и монотонна на отрезке $[a; b]$ и в качестве своих значений принимает все значения между $f(a)$ и $f(b)$, то эта функция непрерывна на $[a; b]$.

16.62* Доказать, что если функция $f(x)$ непрерывна на интервале $(a; b)$ и x_1, \dots, x_n — любые значения из этого интервала, то между ними найдется число ξ такое, что

$$f(\xi) = \frac{1}{n}(f(x_1) + \dots + f(x_n)).$$

Задачи на индукцию

A.1 Докажите, что $1 \cdot 1! + 2 \cdot 2! + \dots + n \cdot n! = (n + 1)! - 1$ при любом натуральном n .

A.2 Докажите, что при любом натуральном n :

a) $2^n > n$; **b)** $\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{n^2} \leq 2 - \frac{1}{n}$.

A.3 Докажите: модуль суммы любого числа слагаемых не больше суммы модулей этих слагаемых.

A.4 Докажите, что $5^n - 4n + 15$ делится на 16 при всех $n \in \mathbb{N}$.

A.5 Докажите, что при любом натуральном n число $a_n = 2n^3 + 3n^2 + 7n$ делится на 6.

A.6 Докажите, что $4^{2n-1} + 1$ кратно 5 для всех $n \geq 1$.

A.7 Докажите, что для любого натурального числа: $6^{2n-2} + 3^{n+1} + 3^{n-1}$ кратно 11.

A.8 Докажите неравенство:

$$\frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n} > \frac{13}{24} \quad (n > 1).$$

A.9 Докажите неравенство:

$$2! \cdot 4! \cdot \dots \cdot (2n)! > [(n+1)!]^n, \quad (n > 2).$$

A.10 Найдите ошибку: Докажем, что в любой табуне все лошади одной масти, индукцией по числу лошадей. Если в табуне одна лошадь, все очевидно. Пусть в любой табуне из n лошадей все лошади одной масти. Возьмем любой табун из $n + 1$ лошади и построим в ряд. По предположению, первые n лошадей одной масти и последние n тоже, то есть все лошади той же масти, что и «средняя» лошадь. Поэтому все лошади в табуне одной масти, что и требовалось доказать.

A.11 Верна ли теорема: если треугольник разбит отрезками на треугольники, то хотя бы один из треугольников разбиения не остроугольный?

Вот ее доказательство (нет ли в нем ошибки?):

1. Если треугольник разбит отрезком на два треугольника, то один из них не остроугольный (ясно).

2. Пусть имеется треугольник, как-то разбитый на n треугольников. Проведем еще один отрезок, разбив один из маленьких треугольников на два. Получим разбиение на $n + 1$ треугольник, причем один из двух новых треугольников будет не остроугольный. По индукции теорема доказана.

A.12 Докажите, что любое натуральное число можно представить как сумму нескольких разных степеней двойки (возможно, включая и нулевую; сумма может состоять и из одного слагаемого).

A.13 Число $x + \frac{1}{x}$ — целое. Докажите, что $x^n + \frac{1}{x^n}$ — тоже целое при любом натуральном n .

A.14 (Ханойские башни) Есть детская пирамида с n кольцами и два пустых стержня той же высоты. Разрешается перекладывать верхнее кольцо с одного стержня на другой, но нельзя класть большее кольцо на меньшее. Докажите, что **a)** можно переложить все кольца на один из пустых стержней; **b)** можно сделать это за $2^n - 1$ перекладываний; **c)** меньшим числом перекладываний не обойтись; **d)** напишите рекурсивный алгоритм, который по заданному числу n выводит список из $2^n - 1$ инструкций вида «кольцо номер k переложить с башни α на башню β », позволяющий переместить башню с одного заданного стержня на другой заданный стержень; при этом кольца занумерованы числами от 1 до n в порядке возрастания диаметра, а стержни помечены буквами A, B, C .

A.15 На краю пустыни имеется неограниченный запас бензина и канистр, а также машина, которая при полной заправке может проехать 50 км. В канистры можно сливать бензин из бензобака машины и оставлять на хранение (в любой точке пустыни). Докажите, что машина может проехать любое расстояние. (Канистры с бензином возить нельзя, пустые можно возить в любом количестве.)

A.16 На какое максимальное число частей могут разбить плоскость **a)** n прямыми; **b)** n окружностей? **c)** На какое максимальное число частей могут разбить пространство n плоскостей?

A.17 При каких n гири весом $1, 2, \dots, n$ кг можно разложить на три равные по весу кучи?

A.18 Бизнесмен заключил с чёртом такую сделку: он может любую имеющуюся у него купюру обменять у чёрта на любой набор купюр любого меньшего достоинства (по своему выбору, без ограничения общей суммы). Он может также тратить деньги, но не может получать их в другом месте (кроме как у чёрта). При этом каждый день на еду ему нужен рубль. Сможет ли он так жить бесконечно долго?

A.19 Рассмотрим два определения отношения \leq на натуральных числах. Первое определение:

$$a \leq_1 b \iff \exists x \in \mathbb{N} \ a + x = b.$$

Второе (рекурсивное) определение:

$$a \leq_2 0 \iff a = 0, \quad a \leq_2 b + 1 \iff (a \leq_2 b) \vee (a = b + 1).$$

Докажите, что эти определения эквивалентны, т. е. $a \leq_1 b \iff a \leq_2 b$ для всех $a, b \in \mathbb{N}$.

Указание. Воспользуйтесь следующими свойствами сложения как аксиомами:

A1 $a + 0 = a$ для любого натурального a ;

A2 $0 \neq x + 1$ для любого натурального x ;

A3 ассоциативность и коммутативность сложения;

A4 если $a \neq 0$, то $a = x + 1$ для некоторого натурального x ;

A5 закон сокращения: если $a + x = b + x$, то $a = b$.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

А

автоморфизм группы	188
аксиомы	
Архимеда	328
группы	38
кольца	65
полноты	330
поля	129
алгебра над кольцом	263
алгебраические числа	306
алгоритм Евклида	27, 67, 118, 159, 230
ассоциативность	12, 18, 38

Б

базис пространства	259
--------------------	-----

В

вектор	253
свободный	254
фиксированный	253
вложенные отрезки	328

Г

гауссовы целые числа	222
гомоморфизм групп	177
группа	38
абелева	39
вычетов	144
движений окружности	56
движений плоскости	102
движений прямой	44
знакопеременная	193
Клейна	87
нормальная подгруппа	185
перестановок	177

полная линейная	262
простая	186
система образующих	183
факторгруппа	186
циклическая	39, 83, 183

Д

движения	34, 36
многоугольника	81
окружности	51
плоскости	89
поворот	52
пространства	105
прямой	34
ромба	86
скользящая симметрия	97
сферы	104
треугольника	79
делимость чисел	22, 222
делители нуля	18
дистрибутивность	18

И

изоморфизм	325
групп	87, 177, 223
порядковый	326
интервал	299
инфимум	337

К

кольцо	65
вычетов	144
гауссовых целых чисел	222
коммутативное	66
многочленов	154
с единицей	66

коммутативность	12, 18, 39
континуум	321
континуум-гипотеза	351
кратность чисел	22
критерий нормальности	188

Л

лемма

Гаусса о неприводимости 306

о двух милиционерах 357

линейная оболочка 258

линейное многообразие 283

линейный оператор 261

матрица оператора 264

обратимый 262

логарифм 375

М

матрица 264

Паули 296

мнимая единица 217

многочлен 154

неприводимый 158

множество

бесконечное 318

всюду плотное 301

конечное 316

континуальное 321

линейно упорядоченное 299

ограниченное 332

открытое 381

плотное 299

связное 381

счетное 317

модуль над кольцом 257

мощность множества 317

Н

натуральные числа 21

нейтральный элемент . 12, 18, 38

неравенство

Коши — Буняковского 296

норма

вектора 268

гауссова числа 222

О

область 382

значений функции 176

определения функции 175

образ множества 176

обратный элемент 12, 39

определитель матрицы 272

основная теорема

алгебры 380

арифметики 69

отношение 168

сравнимости по модулю . 142

эквивалентности 171

отображение 243

П

перестановка 177

подгруппа 39

движений окружности 83

нормальная 185

поле 129

аксиомы поля 129

алгебраических чисел 309

действительных чисел 342

комплексных чисел 379

конечное расширение 303

порядок

группы 83

линейный 299

частичный 299

элемента группы 84

последовательность 331

монотонная 331

сходящаяся 332

фундаментальная 335

предел

второй замечательный 377

первый замечательный 362

последовательности .	332, 355
функции в точке	361
цепи множеств	327
преобразование	244
гомотетия	246
отраженная	252
поворотная	249
подобия	245
принцип	
бесконечного спуска	232
вложенных отрезков	330
дедекиндовых сечений	340
Дирихле	30
индукции	394
точных граней	337
произведение матриц	268
производная	
вещественная	372
комплексная	382
прообраз множества	176
пространство	
векторное	258
линейное	258
прямое произведение	168
Р	
размерность пространства	261
ранг матрицы	285
расширение поля	135
рациональные числа	122, 128
ряд Тейлора	383
С	
сечение Дедекинда	340
скалярное произведение	267
соизмеримость	25
соотношение	
Гейзенберга	297
Робертсона — Шрёдингера	297
сравнение по модулю	142
супремум	337

Т	
таблица	
композиций	45, 98, 247
сложения по модулю	142
умножения по модулю	143
теорема	
Безу	155
Виета	161
Вильсона	162
Кантора	320
Кантора — Бернштейна	321
критерий обратимости	285
Кэли	179
Лагранжа	185
малая Ферма	147
о базисе ЛП	260
о гомоморфизмах	190
о классификации подобий	252
о корнях многочлена	160
о мощности квадрата	318
о подобиях в \mathbb{C}	282
о трех гвоздях	94
основная алгебры	161, 380
основная арифметики	69
ОТА гауссовых чисел	229
рождественская Ферма	228
Ферма	231
Шаля	47, 59, 98
Эйлера	148
точная верхняя грань	337
точная нижняя грань	337
транспозиция	191
У	
упорядоченная пара	168
уравнение	
в целых числах	116
диофантово	230
линейное однородное	116
прямой	115

Ф

формула Эйлера	386
функция	174
биекция	174
инъекция	174
непрерывная	370, 382
обратная	175
показательная	365
регулярная	382
сюръекция	174

Ц

целые числа	65
цепная дробь	119, 133
цепь множеств	327

Ч

числа	
алгебраические	306
ассоциированные гауссовы .	224

взаимно простые	68
гауссовы	222
двоично-рациональные ..	300
иррациональные	135
комплексные	218, 379
натуральные	21
простые	68
рациональные	122, 128
совершенные	77
соизмеримые	133
целые	65
Эйлера	374

Э

экспонента	374
матричная	384

Я

ядро гомоморфизма	189
-------------------------	-----

ОБОЗНАЧЕНИЯ

$\emptyset, \{\}$	пустое множество
$a \in b$	a является элементом множества b
$a \subseteq b$	множество a является частью (или равно) множества b
$\{a, b, \dots, c\}$	множество, состоящее из элементов a, b, \dots, c
$\{x \mid \varphi(x)\}$	множество, состоящее из всех x , для которых $\varphi(x)$
$\bigcup a$	объединение элементов множества a
$a \cap b$	пересечение множеств a и b
$a \cup b$	объединение множеств a и b
$a \setminus b$	разность множеств a и b
$\mathcal{P}(a)$	множество всех подмножеств множества a
$\langle a, b \rangle$	упорядоченная пара a и b
$\langle a, b, \dots, c \rangle$	упорядоченный набор a, b, \dots, c
$\langle g \rangle$	циклическая группа, порожденная элементом g
$\langle T \rangle$	группа, порожденная системой образующих T
$A \times B$	прямое произведение множеств A и B
$\neg \varphi$	отрицание формулы φ
$\varphi \wedge \psi$	одновременное выполнение формул φ и ψ
$\varphi \vee \psi$	верна хотя бы одна из формул φ или ψ
\iff	равносильно (логически)
\rightarrow, \Rightarrow	следует (логически)
$\exists x \varphi(x)$	для некоторого x истинна формула $\varphi(x)$
$\exists x \in a : \varphi(x)$	для некоторого x истинна формула $(x \in a) \wedge \varphi(x)$

$\forall x \varphi(x)$	для любого x истинна формула $\varphi(x)$
$\forall x \in a : \varphi(x)$	для любого x истинна формула $(x \in a) \rightarrow \varphi(x)$
$\text{dom}(f)$	область определения функции f
$\text{ran}(f)$	область значений функции f
$f : A \rightarrow B$	функция f определена на A и действует в B
$f \upharpoonright_D$	сужение функции f на множество D
B^A	множество всех функций из A в B
aRb	$\langle a, b \rangle \in R$ (a и b связаны отношением R)
$[a]_R$	класс эквивалентности элемента a по отношению R
A/B	фактормножество множества A по отношению эквивалентности B , а также факторгруппа группы A по нормальной подгруппе B
$H \triangleleft G$	H есть нормальная подгруппа группы G
$A \leftrightarrow B$	множества A и B равномоцны
$A \preceq B$	существует инъекция из A в B
$\#A, \text{Card}(A)$	количество элементов множества A
$f \circ g$	композиция функций f и g
$f^{-1}A, f^{-1}[A]$	прообраз множества A относительно функции f
$fA, f[A]$	образ множества A относительно функции f
\mathbb{N}	множество всех натуральных чисел, начиная с нуля
\mathbb{Z}	кольцо целых чисел
$\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}_m$	кольцо вычетов по модулю m
\mathbb{Z}_m^*	группа вычетов по умножению по модулю m
\mathbb{Q}	поле рациональных чисел
$\mathbb{Q}[\alpha, \beta]$	расширение поля \mathbb{Q} присоединением α и β
\mathbb{A}	поле алгебраических чисел

\mathbb{R}	поле действительных (вещественных) чисел
\mathbb{R}^n	конечномерное вещественное линейное пространство со скалярным произведением
\mathbb{C}	поле комплексных чисел
$k[x]$	кольцо многочленов над полем k
$a \mid b$ ($b : a$)	a делит b (b делится на a), т. е. $\exists k ak = b$
$a \perp b$	числа a и b взаимно просты
НОД	наибольший общий делитель
НОК	наименьшее (положительное) общее кратное
$a \equiv b \pmod{p}$	a сравнимо с b по модулю p , т. е. $(a - b) : p$
$[A, B]$	$AB - BA$, коммутатор операторов A и B
$[a; b]$	отрезок $\{x \mid a \leq x \leq b\}$ в линейно упорядоченном множестве
$[F : K]$	степень расширения поля F над полем K
$[k_0, k_2, \dots, k_n]$	цепная дробь $k_1 + (k_2 + (\dots + k_n^{-1} \dots)^{-1})^{-1}$, где $k_i \in \mathbb{N}$
$[v_1; \dots; w_n]$	запись матрицы через вектор-столбцы линейного пространства
$[z_1, z_2; z_3, z_4]$	двойное отношение $\frac{(z_3 - z_1)(z_4 - z_2)}{(z_3 - z_2)(z_4 - z_1)}$, где $z_i \in \mathbb{C}$
$\prod_{k=1}^n x_k$	то же, что $x_1 x_2 \cdot \dots \cdot x_n$
$\sum_{k=1}^n x_k$	то же, что $x_1 + x_2 + \dots + x_n$
$\prod_{k=1}^{\infty} x_k$	то же, что $\lim_{n \rightarrow \infty} \prod_{k=1}^n x_k$
$\sum_{k=1}^{\infty} x_k$	то же, что $\lim_{n \rightarrow \infty} \sum_{k=1}^n x_k$
$n!$	факториал натурального числа n , т. е. $1 \cdot 2 \cdot \dots \cdot n$

$\binom{n}{m}$	число сочетаний из n по m , т. е. $\frac{n!}{m!(n-m)!}$
id	тождественное преобразование, $\text{id}(x) \equiv x$
R_α	поворот окружности на угол α
R_α^O	поворот плоскости на угол α относительно центра O
S_A	отражение прямой относительно точки A
S_l	отражение плоскости относительно оси l
T_v	сдвиг на вектор v
W_v^l	скользящая симметрия с осью l и вектором сдвига v
H_O^k	гомотетия с центром O и коэффициентом k
$RH_O^{k,\alpha}$	поворотная гомотетия с центром O , коэффициентом k и поворотом на угол α
$SH_O^{k,m}$	отраженная гомотетия с центром O , коэффициентом k и осью отражения m
(123)(45)	перестановка, образованная двумя циклами (123) и (45)
$\mathfrak{A} \cong \mathfrak{B}$	изоморфизм структур \mathfrak{A} и \mathfrak{B}
$\dim(V)$	размерность пространства V
$\ker(h)$	ядро гомоморфизма h , $\{x \mid h(x) = \mathbf{e}\}$
$\mathbf{S}(X)$	группа всех биекций множества X
\mathbf{S}_n	группа перестановок на n символах
A_n	знакопеременная подгруппа группы \mathbf{S}_n
V_4	четверная группа Клейна
$\text{GL}(V)$	полная линейная группа пространства V
$\text{GL}(n)$	группа обратимых (вещественных) матриц порядка n
$\text{SL}(n)$	специальная линейная группа
$\text{SO}(n)$	специальная ортогональная группа
$\det A$	определитель квадратной матрицы A
A^T	транспонированная матрица A

СПИСОК ЛИТЕРАТУРЫ

- [1] Айерленд К., Роузен М. Классическое введение в современную теорию чисел. М.: Мир, 1987.
- [2] Арнольд В. И. Гюйгенс и Барроу, Ньютон и Гук. М.: Наука, 1989.
- [3] Берже М. Геометрия: Пер. с франц. М.: Мир, 1984.
- [4] Ван дер Варден Б. Л. Алгебра. М.: Наука, 1976.
- [5] Верещагин Н. К., Шень А. Начала теории множеств. М.: МЦНМО, 2012.
- [6] Гельфанд И. М. Лекции по линейной алгебре. М.: Наука, 1971.
- [7] Грэхем Р., Кнут Д., Паташник О. Конкретная математика. М.: Мир, 1998.
- [8] Домрин А. В., Сергеев А. Г. Лекции по комплексному анализу. В 2 частях. М.: МИАН, 2004.
- [9] Зарисский О., Самюэль П. Коммутативная алгебра. В двух томах. М.: ИИЛ, 1963.
- [10] Зорич В. А. Математический анализ. Часть I и Часть II. 6-е изд., дополн. М.: МЦНМО, 2012.
- [11] Клайн М. Математика. Утрата определенности. М.: Мир, 1984.
- [12] Курант Р., Роббинс Г. Что такое математика? Изд. 7-е., стереот. М.: МЦНМО, 2015.
- [13] Кострикин А. И. Введение в алгебру. М.: ФИЗМАТЛИТ, 2004.
- [14] Кострикин А. И., Манин Ю. И. Линейная алгебра и геометрия. Учебное пособие для вузов. 2-е изд., перераб. М.: Наука, 1986.
- [15] Математическая составляющая / Редакторы-составители Н. Н. Андреев, С. П. Коновалов, Н. М. Панюнин; Художник-оформитель Р. А. Кокшаров. 2-е изд., расш. и доп. М.: Фонд «Математические этюды», 2019.
- [16] Прасолов В. В. Многочлены. 4-е изд., испр. М.: МЦНМО, 2014.
- [17] Прасолов В. В. Наглядная топология. 3-е изд., стереотип. М.: МЦНМО, 2012.
- [18] Рид К. Гильберт. М.: Наука, 1977.
- [19] Рудин. У. Основы математического анализа. — М.: Мир, 1976.
- [20] Савватеев А. В. Математика для гуманитариев. Живые лекции. М.: Русский фонд содействия образованию и науке, 2019.

Научно-популярное издание

Казимиров Николай Игоревич, Савватеев Алексей Владимирович

ВВЕДЕНИЕ В НАСТОЯЩУЮ МАТЕМАТИКУ:
ПОСОБИЕ ДЛЯ УЧИТЕЛЕЙ МАТЕМАТИКИ ПО МОТИВАМ КУРСА
«100 УРОКОВ МАТЕМАТИКИ» АЛЕКСЕЯ САВВАТЕЕВА

Корректор *Е. В. Субботина*
Верстка *Н. И. Казимиров*
Оформление *А. А. Кузьмин*

Подписано в печать 20.08.2022. Формат 70 × 100 1/16.
Бумага офсетная.
Тираж 500 экз. Заказ

Русский фонд содействия образованию и науке.
Университет Дмитрия Пожарского
117246, Москва, Научный проезд, д. 8, стр. 1, каб. 430
Сайт: www.publisher.usdp.ru
Почта: knigiudp@gmail.com

Отпечатано: АО «Т 8 Издательские технологии».
109316, Москва, Волгоградский проспект, д. 42, корп. 5.