В. Г. Болтянский, А. П. Савин

БЕСЕДЫ О МАТЕМАТИКЕ

КНИГА 1
ДИСКРЕТНЫЕ ОБЪЕКТЫ

МОСКВА
«ФИМА» • МЦНМО
2002

УДК 51(07)
ББК 22.12
Б79

Болтянский В.Г., Савин А.П. **Беседы о математике. Книга 1.**
Б79 **Дискретные объекты.** – М.: ФИМА, МЦНМО, 2002. – 368 с.
ISBN 5-89492-011-6 («ФИМА»)
ISBN 5-94057-040-2 (МЦНМО)

Книга вводит читателя в круг идей современной математики. В популярной форме рассказывается о теории множеств, комбинаторике, теории графов, теории вероятностей и многом другом.

Издание будет интересно учителям математики. Специальная глава посвящена вопросам, связанным с поиском учащимися решений задач.

В то же время эта книга может служить основой курса математики для студентов гуманитарных специальностей. Такой курс был прочитан авторами для психологов.

Учащиеся и учителя математических школ, лицеев и гимназий могут использовать издание в качестве учебного пособия.

Научно-популярное издание
БОЛТЯНСКИЙ Владимир Григорьевич, *САВИН Анатолий Павлович*
БЕСЕДЫ О МАТЕМАТИКЕ. Книга 1. Дискретные объекты.
Редактор А.Н.Виленкин
Дизайн обложки А.Е.Шабельник

Подписано в печать с оригинал-макета 24.06.2002
Формат 60×90/16. Бумага тип. № 2. Гарнитура Times
Печать офсетная. Печ. л. 23. Тираж 2000 экз. Заказ № 87

ООО «ФИМА»
121248, г. Москва, Украинский бульв., д. 3/5, корп. 2
E-mail: pub-fima@narod.ru

Издательство Московского Центра
непрерывного математического образования
121002, г. Москва, Большой Власьевский пер., д. 11
Тел. 241-72-85
E-mail: adm@mccme.ru

Отпечатано с готовых диапозитивов в ГУП «Облиздат»
248640, г. Калуга, пл. Старый торг, 5

ISBN 5–89492–011–6 («ФИМА»)
ISBN 5–94057–040–2 (МЦНМО)

ОГЛАВЛЕНИЕ

Любимой подруге жизни Эрике
посвящаю

В. Болтянский

ПРЕДИСЛОВИЕ

Математика в системе наук занимает особое место. Ее методами пользуются все существующие науки. «Во всякой науке ровно столько науки, сколько в ней математики» — в справедливости этого высказывания теперь никто уже не сомневается.

Однако если достижения других наук — физики, химии, биологии, астрономии — систематически освещаются в средствах массовой информации и через короткое время попадают на страницы школьных учебников, то о новых математических открытиях прочесть в газете или журнале не удастся. Единственной математической теоремой, время от времени появлявшейся в широкой печати, была Великая теорема Ферма. Но и к ней исчез интерес у журналистов после появления ее доказательства, занимающего сотни страниц и использующего аппарат, недоступный среднему интеллигенту.

Причиной этому — дедуктивный характер математики. Для того, чтобы понять то или иное достижение в алгебре, геометрии или теории вероятностей, как правило, требуется знание огромного фундамента, на котором появилось это достижение. К тому же математика выработала свой специфический язык, овладение которым сродни овладению трудным иностранным языком.

В конце 40-х годов XX века американские математики Р. Курант и Г. Роббинс предприняли удачную попытку рассказать широкому кругу читателей, в первую очередь учителям математики в школе и школьникам, о содержании и методах современной математики в книге «Что такое математика». Эта книга не раз издавалась и на русском языке.

С тех пор прошло полвека и назрела необходимость еще раз попытаться ответить на вопрос «Что такое математика?». Эта книга написана именно с этой целью. Первый ее том «Дискретные объекты» вы держите в руках. Вторым томом предполагается «Непрерывность», а третьим «Экстремум».

Основной текст книги написан В. Г. Болтянским. А. П. Савину принадлежит составление задач и упражнений к материалу книги и написание решений этих задач. Такое построение книги делает возможным ее использование в качестве учебного пособия. Большую и существенную работу проделал редактор книги А. Н. Виленкин, которому авторы чрезвычайно благодарны.

Авторы

ВВЕДЕНИЕ

Беседа 1. Предмет математики

1. Мнения о пользе математики

Наш великий соотечественник Михаил Васильевич Ломоносов говорил: «Математику уже затем учить надо, что она ум в порядок приводит». Но, может быть, за те два с лишним столетия, которые прошли со времен Ломоносова, математика утратила свою актуальность? Приведем свидетельство наших дней.

Американский президент Рональд Рейган в последний год своего правления представил доклад «Нация в опасности». «...Кичась своей демократичностью, Америка позволила старшеклассникам называть, по своему выбору, три главных, с их точки зрения, учебных предмета, по которым они хотели бы быть аттестованными. Разумеется, лишь 10 – 12% молодых людей пожелали видеть среди этих трех предметов математику. И результат не замедлил сказаться: за десятилетие с лишним эта демократичность привела к тому, что молодые юристы стали хуже логически мыслить, молодые врачи стали лечить хуже, чем их более зрелые коллеги в бытность их в том же возрасте, вновь испеченные экономисты стали хуже понимать законы рынка». Президент поставил задачу срочно выбраться из этой интеллектуальной ямы и добиться, чтобы американская молодежь лучше всех знала математику. (Кстати, некоторые наши руководители образования чуть не затолкали нас в эту же яму, пытаясь на этом заработать популярность у сердобольных родителей.)

А известный польский математик и популяризатор Г. Штейнгауз выразил ту же мысль иначе, в виде шутливого принципа: «Математик это сделает лучше», добавив при этом, что его принцип универсален. Не в том, конечно, смысле, что медиков и юристов надо вербовать только среди математиков, но с назиданием, что представитель каждой специальности, владеющий стилем и методом мышления, почерпнутым при творческом изучении математики, будет и в своей области работать лучше.

Впрочем, некоторые ученые (особенно часто — физики) считают, что математика — лишь формалистическая одежда к важнейшим естественно-научным открытиям, тогда как сами эти открытия и не нуждаются в этой скучной, бесполезной с их точки зрения науке. Есть остроумный анекдот (наверное, придуманный физиками) на эту тему. Однажды Шерлок Холмс и доктор Ваттсон задумали распутать сложную криминальную историю. В селение, где случилось происшествие, не вела хорошая дорога, и они решили воспользоваться воздушным шаром. Однако, ввиду ненастной погоды, воздушный шар покрылся корочкой льда, и наши герои куда-то шлепнулись, к счастью, благополучно.

— Где мы находимся? — спросил Шерлок Холмс у стоящего невдалеке человека.

— В кабине вашего воздушного шара, — подумав, ответил тот.

— Он — математик! — воскликнул Шерлок Холмс. На вопрос доктора Ваттсона, откуда он это узнал, Холмс ответил, пожав плечами:

— Это очевидно: он ответил не сразу, а подумав; его ответ был абсолютно точен и совершенно бесполезен!

Конечно, физики, мягко говоря, не очень искренни в этом своем мнении. Великий Эйнштейн, создатель теории относительности, взял основную идею о пространственно-временном интервале у крупнейшего математика А. Пуанкаре, а также пользовался идеями и понятиями математика М. Гроссмана, чтобы облечь свои идеи по специальной теории относительности в точную форму. Что же касается общей теории относительности, то и здесь огромное значение имеют исследования выдающегося математика Давида Гильберта, который также внес большой вклад в создание этой теории.

Математический стиль мышления нужен всем, а в естественных науках (в том числе физике) также нужны фундаментальные математические результаты и теории.

Бытует и еще одно мнение. Нередко люди, далекие от математики и ее приложений, полагают, что наука эта сухая и неинтересная (хотя, может быть, и полезная), что все в ней давно уже открыто и ничего нового не придумаешь (все те же «дважды два — четыре», все те же квадратные уравнения и та же теорема Пифагора), а ученый-математик — это тот, кто умеет хорошо считать. Теперь же, поскольку появились быстродействующие вычислительные машины, сама профессия математика скоро будет ненужной.

Однако подобные рассуждения в корне ошибочны. Компьютеры (и даже карманные калькуляторы) действительно лучше справляются с утомительными и скучными арифметическими подсчетами, чем считающий вручную вычислитель. Но математик видит интерес и пользу своей науки вовсе не в вычислениях, а в богатстве ее идей и приложений. Ведь алгоритм вычислений должен дать компьютеру человек.

Ошибочно и мнение о том, что в математике не появляется ничего нового. Напротив, именно сейчас эта наука бурно развивается, движется семимильными шагами по непроторенным, неизведанным путям. Школьникам известны алгебра, геометрия, начала математического анализа. Инженер может добавить к этому еще несколько разделов так называемой высшей математики: интегральное исчисление, теорию дифференциальных уравнений, теорию вероятностей, математическую статистику... Современная же математика включает десятки новых направлений и областей, превосходящих по объему и идейному богатству названные разделы. Причем многие из них обязаны своим развитием нескольким последним десятилетиям и даже годам.

В нашей стране издается реферативный журнал «Математика», в котором кратко излагается содержание научных работ, публикуемых математиками всего мира. Аналогичные журналы (Mathematical Reviews, Zentral Blatt) выпускаются и за рубежом. За год эти журналы помещают десятки тысяч рефератов. И в каждой реферируемой работе есть что-то новое, какое-то продвижение вперед (хотя, может быть, не всегда крупное). Таких темпов развития математика не знала никогда.

Но если умение хорошо считать — далеко не самое важное в математике, то какие же проблемы волнуют представителей этой науки? Откуда эти проблемы возникают? И почему именно сейчас так велики темпы развития математики? Постараемся ответить на эти вопросы.

2. Понятия математики и их возникновение

Как и другие науки, математика использует много различных понятий. Все они почерпнуты из жизни и являются плодом многовековых наблюдений человечества. Сами математические термины убедительно свидетельствуют о возникновении математических понятий путем абстракции от реальных предметов и отношений.

Так, термин *трапеция* (рис. 1) происходит от древнегреческого слова τραπεζιον («трапезион»), означавшего *столик* (рис. 2) — вспомните древнерусское слово *трапеза*. Математическому термину *линия* основой возникновения послужило латинское слово linum, означающее *лен*, *льняная нить*. Английский термин line означает не только «линия», но также «прямая». Вообще, слово «линия» часто употребляется в смысле «прямая линия» (отсюда происходит термин «линейка»). Очевидно, понятие прямой линии является абстракцией от натянутой льняной нити.

От греческого слова σφαιρα (означавшего *мяч*), происходит математический термин *сфера* (рис. 3). От слова πρισμα (означавшего *опиленная*) происходит термин *призма* (рис. 4). От греческого слова κωνοζ (*сосновая шишка*) возник термин *конус* (рис. 5); заметим, что до сих пор в английском языке слово cone означает и конус, и шишку хвойного дерева (*pine cone* — сосновая шишка, рис. 6).

Слово *точка* означало в древнерусском языке отточенное перочинным ножом острие гусиного пера (рис. 7), которым писали (и ставили «точки») в древности. Математические термины круг, окружность имеют, очевидно, общее происхождение с общеязыковыми словами кружить, окружать, округа.

Понятия возникают в голове человека (в результате его практической деятельности) посредством процесса абстракции. Много тысячелетий люди наблюдали предметы, сгруппированные парами: две

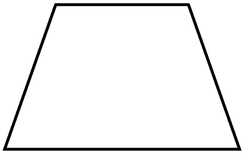


Рис. 1

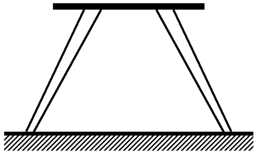


Рис. 2

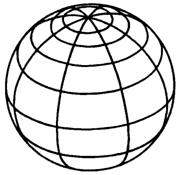


Рис. 3

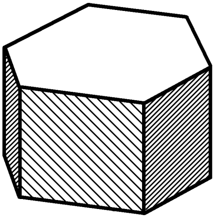


Рис. 4

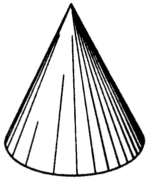


Рис. 5

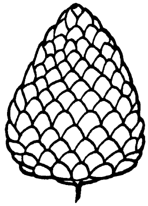


Рис. 6

Рис. 7

руки, два глаза, два уха у человека, два рога у коровы... Отвлекаясь, абстрагируясь от конкретных особенностей этих предметов и оставляя лишь существенное, общее свойство предметов, сгруппированных в пары (неважно, руки это, глаза или уши; существенно что имеется пара этих предметов), мы постепенно формируем в своем сознании абстрактное понятие *два*. А возникновение термина (в данном случае, «два») завершает, закрепляет формирование абстрактного понятия и является символом, знаком, словесным эквивалентом этого понятия.

Заметим, что ребенок, проходящий путь познания окружающего мира, гораздо легче усваивает абстрактные понятия в сравнении с теми трудностями, которые в процессе исторического развития преодолело человечество при первоначальной выработке этих абстрактных понятий. Объясняется это тем, что родители, уже владеющие абстрактными понятиями, сознательно акцентируют внимание ребенка на соответствующих предметах и обстоятельствах, используя термины, выражающие нужное понятие: «У тебя *две* ручки, *два* глазика», «Смотри: на лугу *две* коровки» и т. д. Повторение термина «два» помогает возникновению абстрактного понятия в сознании ребенка.

Другой причиной, облегчающей ребенку формирование, скажем, абстрактных геометрических понятий, является «овеществление» этих понятий в окружающих геометрических формах. Первобытный человек почти не имел перед глазами предметов, помогающих выработке абстрактных понятий *прямая линия*, *окружность*, *прямой угол* и т. д. Сегодня ребенок на каждом шагу видит прямолинейные углы зданий, ребра прямоугольных коробок, прямые углы в переплетах окон и страницах книги...

3. Некоторые виды абстракции

Современная теория познания различает несколько видов абстракции, и все они играют важную роль при образовании математических понятий. Самой простой является *абстракция отождествления*. Ребенок сотни раз держит в руках ложку и приучается пользоваться ею. Но это не одна и та же ложка. Он видит чайную и столовую ложку; деревянную, стальную или серебряную; свою ложку и мамину; ложку строгой формы или украшенную орнаментом, инкрустацией... Но эти индивидуальные отличия ложек являются второстепенными, несущественными. Главным же отличием ложек от других предметов является их функциональное назначение: столовый прибор, предназначенный для набора жидкой или вязкой пищи при еде. Мы объединяем предметы этого класса по их основному признаку (в данном случае по их функциональному назначению), отвлекаясь, абстрагируясь от индивидуальных, несущественных особенностей. Мы как бы *отождествляем* между собой все предметы этого класса: ложка «вообще». Абстракция отождествления, завершенная введением термина, и приводит к формированию абстрактного понятия.

Формирование отмеченного выше абстрактного понятия «два» также является, конечно, результатом абстракции отождествления. Выработка абстрактных понятий «один», «два», «три» и т. д. была многовековой и мучительной. Достаточно сказать, что еще в XIX столетии у некоторых племен, населявших Полинезию, были в употреблении три сорта чисел: одни для пересчета людей, другие для счета животных, третьи — для пересчета оружия, утвари и других неодушевленных предметов. Абстракция отождествления, позволяющая перейти к *одной* системе чисел, тогда еще не была проведена до конца у этих народов.

Другим видом абстракции, имеющим особенно важное значение для математики (в частности, для геометрии), является *абстракция идеализации*. Сущность ее состоит в том, что мы не просто отождествляем между собой предметы определенного класса, но также идеализируем их, наделяем такими свойствами, которыми реальные предметы не обладают. Так, понятие геометрической *точки* мы получаем, отождествляя между собой предметы очень малых, пренебрежимых размеров по сравнению с другими предметами, встречающимися в рассматриваемой ситуации. Однако мы не просто отождествляем эти предметы между собой, а наделяем их *идеальным*, воображаемым свойством. Мы считаем, что предметы совсем не имеют размеров, и такие идеализированные предметы считаем точками. Мы абстрагируемся не только от цвета, температуры, материала предмета, но и от его размеров. В реальном мире не существует предмета, о котором можно было бы сказать, что это — геометрическая точка. Понятие *линии*, имеющей протяженность только в одном направлении (имеющей длину, но не обладающей толщиной), также является результатом применения абстракции идеализации. То же относится и к любой геометрической фигуре. Коробка, шкаф, высотный дом являются реальными прототипами понятия *прямоугольный параллелепипед*, но ни об одном из этих предметов нельзя сказать, что это и есть прямоугольный параллелепипед. Мы не только должны отвлечься от второстепенных деталей (ручек, окон и т. п.), но и наделить предмет идеальными свойствами: абсолютно ровные грани, в точности прямолинейные ребра и т. д. Мы не только отождествляем предметы похожей формы, но и идеализируем их, наделяя свойствами, которыми реальные предметы обладают лишь с определенной степенью приближения.

Отметим еще *абстракцию потенциальной бесконечности*. Мы говорим, что прямая *неограниченно* продолжается в обе стороны, говорим, что натуральных чисел 1, 2, 3, ... *бесконечно много*. В действительном мире нет предмета, обладающего такой протяженностью, как бесконечная геометрическая прямая, нет такого набора предметов, для пересчета которого потребовалось бы «бесконечное» множество чисел. Понятие о натуральном числе мы получаем, пересчитывая *небольшое* количество предметов (десятки, в крайнем случае сотни). Но, составив свое мнение о свойствах этих не очень больших чисел, мы позволяем себе мысленно раздвинуть эти ограничительные рамки, вообразить бесконечное множество всех натуральных чисел.

4. Многоступенчатые абстракции

Математике свойственны *многоступенчатые абстракции*, т. е. абстракции над абстракциями. Разумеется, они свойственны и другим областям знания. Абстрактные понятия *идти*, *работать*, *говорить* приводят (с помощью абстракции отождествления) к понятию *глагола*. Это уже абстракция второй ступени, абстракция от абстракций. А еще один шаг, и мы объединяем понятия глагол, существительное, наречие и т. п. одним абстрактным синтаксическим понятием *часть речи*. Это уже абстракция следующей ступени. Таким образом, грамматические абстракции являются многоступенчатыми.

Для математики особенно свойственно образование многоступенчатых абстракций. Абстрактные понятия два, три, пять, ... приводят к следующей ступени абстракции — понятию натурального числа. Первоначально оно было связано с непосредственным пересчетом, исчислением предметов (откуда и происходит современный термин *число*). Еще одна ступень абстракции (на этот раз абстракции потенциальной бесконечности), и мы приходим к понятию натурального ряда. Оно как бы существует само по себе, со всеми его свойствами. К ним относятся *действия*: сложение, умножение, причем эти действия неограниченно выполнимы и обладают обобщенными свойствами. Таковы *коммутативность* (переместительность) сложения и умножения, выражаемая формулами

$$m + n = n + m; \; mn = nm;$$

ассоциативность (сочетательность) этих действий:

$$m + (n + p) = (m + n) + p; \quad m(np) = (mn)p;$$

дистрибутивность (распределительность):

$$m(n + p) = mn + mp.$$

И когда вам говорят «некоторое натуральное число», то вы в первую очередь имеете в виду именно весь натуральный ряд чисел и эти его свойства, а не вспоминаете пересчет конкретных предметов или, тем более, два уха и четыре ножки у стола.

Но, несмотря на очень высокую ступень абстракции, натуральное число — это далеко не последняя ступень обобщения. В процессе исторического развития математики (и, разумеется, в связи с практической деятельностью человека) возникли нуль и отрицательные числа, рациональные числа. Затем появились действительные числа, а впоследствии и комплексные числа. Теперь уже складывать и умножать можно не только натуральные, но также действительные (или комплексные) числа, причем и в этом случае сохраняются коммутативность, ассоциативность, дистрибутивность. Но и на этом ряд абстракций не кончается. Возникают *векторы*, которые также можно складывать (и умножать на числа), и вновь указанные свойства действий сохраняются.

Абстрагируясь от конкретных (теперь уже второстепенных) черт тех объектов, которые мы складываем, применяя снова абстракцию отождествления, мы приходим к понятию *группы* (точнее, *коммутативной группы*). Иными словами, коммутативная группа — это множество каких-то элементов (не обязательно чисел или векторов), для которых определена операция сложения, обладающая отмеченными выше свойствами. Точнее, сложение в коммутативной группе обладает свойствами коммутативности и ассоциативности:

$$a + b = b + a, \quad a + (b + c) = (a + b) + c. \tag{1}$$

Кроме того, в группе имеется элемент 0 (*нуль*), обладающий тем свойством, что

$$a + 0 = a \tag{2}$$

для любого элемента a; наконец, для каждого a имеется *противоположный элемент* $-a$, для которого выполняется соотношение

$$a + (-a) = 0. \tag{3}$$

К высшим ступеням абстракции относится также понятие *поля* (получающееся при рассмотрении множеств с двумя операциями — сложением и умножением). Ряд высших ступеней абстракции связан и с геометрическими понятиями (метрическое пространство, риманово пространство, топологическое пространство). В каком-то смысле каждая область современной математики является сложным переплетением алгебраических и топологических конструкций. Эту мысль образно выразил французский математик Андре Вейль, который говорил, что «за душу каждого математика борются ангел топологии и дьявол абстрактной алгебры».

Высокая абстрактность математических понятий и теорий привела к появлению у некоторых философов-математиков мнения о том, что математические теории — это не более чем «свободная игра разума». Это мнение можно пояснить следующим образом. Коммутативная группа — это собрание некоторых символов, не имеющих никакого конкретного смысла и связи с реальной действительностью, важно лишь, чтобы было указано, какие символы получаются при «сложении» двух других, и выполнялись указанные выше свойства сложения (1), (2), (3). Можно также свободным творением разума предписать формальным символам следовать не этим свойствам (аксиомам), а любым другим, создаваемым игрой воображения, и тогда получится не коммутативная группа, а иное математическое понятие. Делом чистой удачи является нахождение таких аксиом, которые приводят к «математически интересным» понятиям, допускающим достаточно длинную цепочку логических выводов, умозаключений, следствий... Именно такой игрой символов и занимается математика. А тот факт, что эта математическая игра символов иногда приводит к выводам, которые оказываются полезными и применимыми на практике, является удивительным и необъяснимым...

Современная теория познания смотрит на математику иначе. Окружающий материальный мир, воздействуя на наши органы чувств, вызывает ощущения, приводящие к образованию в нашем сознании представлений, образов, абстрактных понятий. Самая сложная идея, связанная с высокими ступенями абстракции, может быть с помощью возвратного самоанализа прослежена и сведена к все более и более простым понятиям, а в конце концов, к образам и представлениям, являющимся отражениями реальных предметов и процессов — подобно тому, как огромное мозаичное панно оказывается при более детальном рассмотрении сложенным из фрагментов и отдельных разноцветных камешков. Самая дерзкая и неожиданная идея ученого-математика, как бы она ни казалась «свободной» игрой разума, всегда имеет отдаленные аналогии и связи с представлениями, образами, понятиями, «подсмотренными» в окружающей действительности. Именно в том, что даже самые абстрактные математические понятия отражают определенные черты реальности, и заключается причина того, что из этих понятий могут быть построены математические модели, которые правильно, адекватно описывают (с той или иной точностью) определенные стороны реальных процессов и явлений.

5. Пространственные и пространственноподобные формы

Многие математические понятия имеют гораздо более абстрактный характер, чем понятия других естественных наук. Это означает, что математик, наблюдая окружающий мир, отвлекается от большего количества черт и признаков реальных процессов и явлений, чем его коллеги биологи, химики, физики. В связи с этим в поле зрения математика остаются наиболее глубинные черты предметов и явлений. Именно этим объясняется то, что математика находит применения во всех других науках. Мы говорим, что молекула воды содержит *два* атома водорода, отмечаем *векторный* характер силы и скорости, рассматриваем *группу* Лоренца в специальной теории относительности...

Как же охарактеризовать те глубинные черты реальной действительности, которые находят отражение в математических понятиях? Иными словами, *что* изучает математика, каков ее предмет?

Энгельс писал в своей «Диалектике природы», что математика изучает *количественные отношения* и *пространственные формы* реального мира. Например, числа и действия над ними, уравнения, функции и их свойства (в том числе такие относящиеся к ним понятия, как производная и интеграл) — все это математические понятия и закономерности, являющиеся отражениями *количественных отношений* реального мира. Геометрические фигуры и их свойства относятся к *пространственным формам*. Имеются и сложные переплетения количественных отношений и пространственных форм друг с другом. Так, измерение геометрических величин (длин, площадей, объемов) дает некоторые количественные характеристики геометрических фигур. Другим примером взаимопроникновения количественных отношений и пространственных форм является графическое изображение функциональных зависимостей, лежащее в основе аналитической геометрии и открытое великим французским математиком и философом Рене Декартом. Это соединение числа и фигуры, функции и линии (ее графика) является мощным математическим средством описания реальных процессов и познания мира, и это взаимопроникновение количественного и пространственного весьма характерно и существенно для современной математики.

Однако в связи с развитием современной математики ученые приходят сегодня к выводу, что энгельсово определение математики как науки о количественных отношениях и пространственных формах реального мира является несколько ограничительным и нуждающимся в расширении, уточнении. Академик А. Д. Александров считал, что следует говорить не только о пространственных, но и о *пространственноподобных* формах реального мира. Так, в математике (и не только в теоретической, абстрактной математике, но и в приложениях к физике, химии, экономике, биологии и другим наукам) систематически рассматриваются *многомерные пространства* (четырехмерное, n-мерное и даже бесконечномерное). Геометрические фигуры, изучаемые в элементарной геометрии, расположены в *трехмерном* пространстве, и именно это является отражением пространственных форм реального мира.

Вообразим себе двумерное существо, живущее в плоскости и могущее свободно перемещаться в ней. Для такого существа окружность будет замкнутой «тюремной камерой» (рис. 8). Но стоит только заметить, что плоскость лежит в трехмерном пространстве, как мы, трехмерные существа, сможем освободить узника, вынув его из плоскости и затем положив обратно, но вне «тюремной камеры» (рис. 9). Освобождение реального узника в нашем трехмерном пространстве

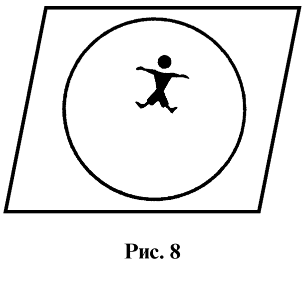


Рис. 8

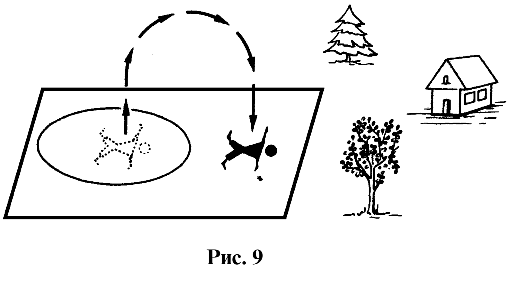


Рис. 9

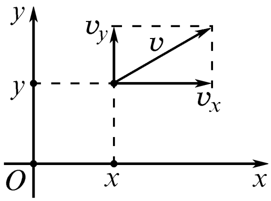


Рис. 10

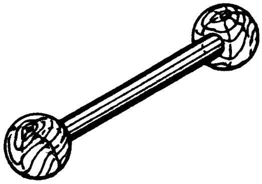


Рис. 11

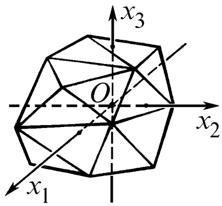


Рис. 12

можно было бы сделать, если бы оно было вложено в «реально существующее» четырехмерное пространство. Тогда можно было бы выйти из камеры заключения в четвертое измерение, а затем снова вернуться в наше трехмерное пространство, но уже вне пределов тюрьмы. Однако такой возможности реально нет: наше пространство трехмерно, и никакого четвертого измерения реально, в буквальном смысле, не существует. Никто, даже самый гениальный математик, не в состоянии *видеть*, зрительно ощущать четырехмерные (а тем более многомерные) фигуры. Но это вовсе не означает, что многомерные пространства являются заумной, надуманной и никому не нужной математической функцией. Состояния механических систем, имеющих более трех степеней свободы, адекватно описываются именно точками многомерных пространств, а расчеты таких систем, осуществляемые на основании формул многомерной геометрии, хорошо подтверждаются на практике. Например, механическое состояние материальной точки, движущейся в плоскости, определяется четырьмя координатами: двумя геометрическими координатами и двумя компонентами скорости (рис. 10). Иначе говоря, ее механическое состояние изображается точкой *четырехмерного* пространства. А механическое состояние системы, состоящей из двух материальных точек, соединенных жестким стержнем (рис. 11), описывается точкой *десятимерного* пространства. Вот еще пример. Для расчета наиболее выгодного плана вывоза сырья, скажем, с двух складов на четыре завода (при строго регламентированной потребности заводов в сырье) требуется ввести *четыре* параметра, характеризующие количество сырья, вывозимого на заводы с первого склада (тогда со второго склада надо будет довезти на заводы недостающее количество сырья). Поэтому каждый возможный план перевозок изображается точкой *четырехмерного* пространства. А *все* возможные планы перевозок, как можно доказать, образуют *выпуклый четырехмерный многогранник*. Не трехмерный (рис. 12), а четырехмерный. И самый выгодный план перевозок изображается одной из вершин этого многогранника. А при большем количестве складов и заводов для описания ситуации требуются пространства еще больших размерностей. Именно геометрия многомерных пространств позволяет рассчитывать оптимальные планы перевозок и решать многие другие задачи математической экономики. Российский ученый, академик Леонид Канторович, открывший эту новую область математики и ее приложения к экономике, за это открытие был удостоен Нобелевской премии.

Фигуры, формулы, закономерности многомерных пространств, изучаемые математикой, являются, таким образом, «слепками», моделями, имитациями реальных процессов, но относятся они не к пространственным, в буквальном понимании, а к пространственноподобным формам реального мира.

И хотя в буквальном смысле «видеть» многомерные фигуры никому не дано, но, переходя от двумерного случая (плоскости) к трехмерному пространству и рассуждая далее по аналогии (с учетом математических формул многомерной геометрии), можно выработать многомерные представления и интуицию, которые заменяют «видение» многомерных фигур обычным зрением. Именно в этом смысле математики охватывают мысленным взором геометрические фигуры и открывают закономерности в многомерных и даже бесконечномерных пространствах (находящих важные приложения, например, в квантовой физике). А аналитические расчеты позволяют точно подтвердить правильность этих наглядно подмеченных фактов многомерной геометрии.

Такое «многомерное зрение» настолько характерно для современной математики и ее приложений, что геометрия сегодня не только является наукой о непосредственно наблюдаемых пространственных формах реального мира, но и служит средством осмысления математических (в том числе количественных) соотношений и закономерностей с помощью образов многомерной геометрии, служит средством переработки и организации математической информации. Это — своеобразное «геометрическое мировоззрение» современного математика, физика, биолога.

Многомерная геометрия — это лишь один (хотя и очень важный) пример пространственноподобных отношений реального мира. Наиболее общие пространственноподобные отношения изучаются в топологии, одном из наиболее абстрактных и в то же время имеющих широкие приложения разделов современной математики.

6. Количественные отношения реального мира

Понимание *количественных отношений*, составляющих другую часть энгельсова определения математики, также нуждается в уточнении и расширении. При непосредственном, численном понимании количественных отношении мы имеем в виду: результат *счета*; результат *измерения* длин, площадей, объемов, скоростей, температур и других величин; *функциональную зависимость* (связывающую значения двух или нескольких переменных величин и позволяющих найти значение зависимой переменной, если известно значение независимой переменной, называемой аргументом), а также дальнейшие числовые характеристики, связанные с этими понятиями — производная, интеграл, дифференциальное уравнение и его решения и т. д. Разумеется, буквенные обозначения, идущие от работ Рене Декарта и Франсуа Виета, и связанное с ними дальнейшее развитие алгебры (тождества, уравнения, многочлены — вплоть до понятий группы, поля и дальнейших обобщений) также подпадают под категорию количественных отношений. Все это приводит к численным результатам (или предполагает возможность свести, в конце концов, все к численным результатам). Однако за последнее время весьма важное значение приобрело *качественное выражение* количественных закономерностей реального мира. Приведем несколько примеров для пояснения того, о чем здесь идет речь.

При работе паровой машины (или другого двигателя) угловая скорость вращения вала не является, конечно, строго постоянной. Некоторые флуктуации давления пара в котле, небольшие колебания нагрузки и другие «возмущения», как говорят математики и инженеры, приводят к тому, что угловая скорость вращения вала то и дело отклоняется от номинального значения. Однако возникшие отклонения гаснут, и угловая скорость вновь и вновь приближается к этому значению. В таком случае говорят, что работа паровой машины является *устойчивой*.

Мы специально остановили внимание на паровой машине, этом несколько старомодном двигателе, чтобы рассказать об исторических фактах, существенно связанных с развитием математики. В XIX столетии паровые машины становились все более совершенными. Увеличивались скорости вращения и, в связи с этим, уменьшались размеры и массы маховиков (массивный маховик большого размера может разорваться при быстром вращении), более массивными становились заслонки паропроводов и в связи с этим увеличивалась масса шаров центробежного регулятора Уатта (напомним, что этот великолепный по своему замыслу регулятор, схематически показанный на рис. 13, приводит к *расхождению* шаров при чрезмерном увеличении скорости вращения, благодаря чему прикрывается заслонка паропровода и скорость вращения начинает уменьшаться). Кроме того, совершенствование машин было связано с улучшением обработки поверхностей и уменьшением трения... И вдруг наступил кризис в машиностроении. Одна за другой вновь спроектированные паровые машины ломались, шли вразнос.

Русский инженер И. А. Вышнеградский изящным математическим исследованием вскрыл причину этого кризиса.

Он изучил дифференциальное уравнение, приближенно описывающее работу паровой машины, и обнаружил, что у прежних (устойчиво работавших) машин корни соответствующего «характеристического уравнения» были расположены в плоскости комплексных чисел *слева* от мнимой оси (рис. 14), чем и гарантировалась устойчивая работа машин. Уменьшение маховиков, увеличение массы шаров регулятора, уменьшение трения в сочленениях привело к постепенному перемещению корней характеристического уравнения вправо. В конце концов корни «перешагнули» через мнимую ось, вследствие чего работа машины стала неустойчивой: малейшее отклонение от положения равновесия приводило к увеличению, а не уменьшению амплитуд отклонений, происходила все увеличивающаяся раскачка,

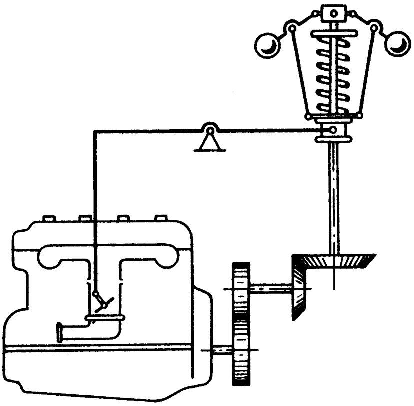


Рис. 13

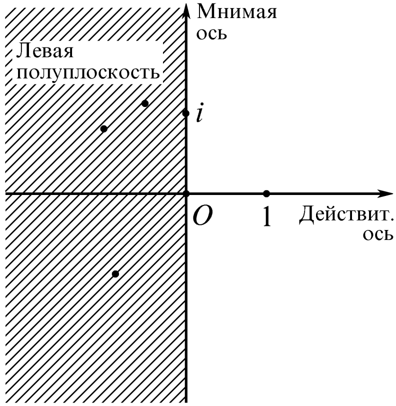


Рис. 14

машина «шла вразнос». И поскольку уменьшение маховиков и увеличение шаров регулятора было необходимым объективным условием увеличения мощности машин, Вышнеградский выдвинул тезис: надо искусственно вводить дополнительное «вязкое» трение при помощи специальных устройств, демпферов. Этот вывод — *качественный*, характеризующий глубокое понимание существа происходящих процессов, хотя выражался он через количественные расчеты и, в свою очередь, позволял видоизменять конструкции и расчеты в нужном направлении. Вскоре возникла новая ветвь математики, *теория устойчивости*, в создание и развитие которой наибольший вклад внесли академик А. А. Ляпунов (1857 – 1918) и другие русские математики.

Второй пример относится к области экологии. В скалистых фиордах Норвегии водятся белые куропатки, являющиеся важным объектом охотничьего промысла. Было замечено, что значительную часть куропаток уничтожают орланы. И тогда было решено истребить конкурентов-орланов, для чего за убитого орлана была назначена цена, несколько превышающая цену куропатки. Охотники переключились на этот более выгодный промысел, и вскоре орланы были выбиты почти полностью. Результат был обнадеживающим: поголовье куропаток стало расти из года в год. И вдруг, через 17 лет, наступила катастрофа. Без всяких видимых причин наступил страшный мор в стаях куропаток, и они почти исчезли.

Математическое исследование дает хорошо подтверждающуюся *качественную* картину этого явления. Куропатки болеют вирусным орнитозом. Больные птицы менее подвижны, и именно они в первую очередь становились добычей орланов. Подобно волкам в наших лесах, орланы были санитарами, отбраковывающими больных птиц. А в отсутствие орланов скученность растущего поголовья куропаток привела к эпидемии. Расчеты, связанные с составлением и решением дифференциальных уравнений «совместного существования видов», позволили дать полную качественную картину. Без учета орланов уравнение описывало «совместное существование двух видов» (куропатки — вирусы), решение которого обнаруживало устойчивый периодический процесс (возрастание поголовья куропаток — возникновение эпидемии — медленное восстановление небольшой популяции сохранившихся куропаток — новое возрастание поголовья и т. д.). И неважно, дает ли расчет точно 17 лет в качестве периода повторения процесса или 15, или 20. Важно, что он дает правильную качественную картину, позволяющую правильно вести охотничье хозяйство. И действительно, через много лет, когда восстановилось поголовье орланов-санитаров, все вошло в прежнее устойчивое состояние.

В рассмотренных примерах качественный характер проявления количественных закономерностей был совершенно различным. В первом примере изучалась устойчивость приближения к равновесному режиму, во втором — периодический, колебательный характер проявляющихся видоизменений. Количественные отношения реального мира проявляются здесь в их *качественном* выражении. Отражением этого является возникновение в современной математике теории устойчивости, качественной теории дифференциальных уравнений, теории катастроф и других разделов, в которых количественное исследование направлено на получение не *численных*, а *качественных* результатов.

Таким образом, несмотря на уточнения, которые должны быть сегодня сделаны в понимании пространственных форм и количественных отношений реального мира, энгельсово определение предмета математики (даже в применении к современному ее развитию) правильно отражает наличие в ней двух основополагающих начал. Пространственное и количественное порождают всю математику, взаимопроникая друг в друга и обогащаясь все новыми связями. Учитывая те уточнения в понимании пространственного и количественного, которые были отмечены выше, можно в итоге сформулировать следующее определение предмета математики: предметом математики являются, во-первых, пространственные (а также пространственноподобные) формы реального мира, и, во-вторых, количественные отношения (в их численном или качественном выражении) реального мира. Взаимосвязь и взаимопроникновение этих двух ее частей обусловливают все многообразие разделов и направлений современной математики.

Заметим в заключение, что мы здесь обсуждали только *предмет* математики, оставив вне рассмотрения ее методы. Они также весьма специфичны. Здесь и дедукция, и математический аспект индукции, анализ и синтез, аналогия и многое другое. Здесь также аксиоматический метод, интересный подход к пониманию непротиворечивости теории, глубокое обобщение аристотелевской логики. Но о методах математики мы поговорим позже, постепенно раскрывая содержание различных разделов этой науки.

Глава I. МНОЖЕСТВА И ОТОБРАЖЕНИЯ

Беседа 2. Конечные и бесконечные множества

7. Множество и его элементы

Мы часто говорим о стае птиц, наборе фломастеров, коллекции минералов, собрании картин и т. д. Обобщая, можно сказать, что речь идет о *множестве* некоторых предметов. Термин *множество* широко применяется в математике. Предметы, входящие в рассматриваемое множество, называются его *элементами*. Кроме основного термина множество, в математике иногда используются его синонимы: система, семейство, класс и др. Например, мы говорим о системе уравнений, о семействе всех прямых на плоскости, о классе всех непрерывных функций и т. п. Однако эти синонимы используются лишь в целях разнообразия языка, и каждый из них может быть заменен универсальным термином «множество».

Множество является первоначальным математическим понятием, т. е. никакого определения понятия «множество» не предусматривается. А апелляция к общеязыковым словам стая, набор, коллекция, собрание — это не определение (в точном математическом смысле), а лишь пояснение смысла этого понятия. Однако есть одно непременное требование к математическому понятию *множество*: о каждом предмете, каждом элементе должно быть возможно точное выяснение того, принадлежит ли он рассматриваемому множеству или нет.

Обозначим, например, через $\boldsymbol{F}$ множество всех цифр, используемых при десятичной записи чисел. Тогда 5 есть элемент этого множества, 8 тоже его элемент, а 12 и –1 не являются элементами множества $\boldsymbol{F}$. Это записывают следующим образом:

$$5 \in \boldsymbol{F}, \quad 8 \in \boldsymbol{F}, \quad 12 \notin \boldsymbol{F}, \quad -1 \notin \boldsymbol{F}.$$

Знак $\in$ говорит о том, что элемент принадлежит рассматриваемому множеству (или включается в данное множество). Он называется *знаком включения* и изображается одним из видов написания греческой буквы эпсилон. Тот же знак, но перечеркнутый ($\notin$), говорит о том, что рассматриваемый элемент не принадлежит рассматриваемому множеству.

Когда мы говорим о том, что для каждого предмета должно быть возможным точное выяснение того, принадлежит ли он рассматриваемому множеству, речь вовсе не идет о том, насколько просто это может быть выяснено. Речь идет лишь о *принципиальной* возможности выяснения этого. Например, в математике и ее приложениях часто используется число $\pi = 3{,}141592653...$. На 4-м и 8-м местах после запятой в этой записи стоит цифра 5. Обозначим через $\boldsymbol{P}$ множество всех «адресов» пятерок, т. е. таких натуральных чисел n, что на n-м месте после запятой в десятичной записи числа π стоит цифра 5. Как видно из записи нескольких первых цифр, множеству $\boldsymbol{P}$ принадлежат числа 4 и 8. А принадлежит ли число миллиард множеству $\boldsymbol{P}$? Сегодня этого никто не знает, но нельзя сказать, что этот вопрос *в принципе* не может быть решен. Существует некий способ вычисления, *алгоритм* вычисления знаков числа π. Математик Шенкс, живший в XIX столетии, потратил много лет жизни на вычисление десятичных знаков числа π. Он нашел 707 знаков после запятой (правда, вследствие ошибки, вкравшейся в его вычисления, верными из них были лишь первые четыре сотни цифр).

Современные компьютеры позволили найти несколько тысяч знаков числа π. Проведя огромное количество вычислений (что потребует огромного времени работы компьютера), можно найти, какая цифра стоит на миллиардном месте после запятой в записи числа π и, тем самым, какая из формул $10^9 \in \boldsymbol{P}$, $10^9 \notin \boldsymbol{P}$ верна.

Но, скажем, множество всех молекул воды, налитой в стакан, не является математически точно определенным вследствие непрерывно происходящих процессов испарения и конденсации.

Приведем пример, показывающий, что иногда «определение» множества, кажущееся вполне четким, может оказаться математически некорректным. В одном взводе был взводный парикмахер. Командир приказал ему брить *тех и только тех солдат, которые не бреются сами*. Обозначим множество этих солдат через $\boldsymbol{B}$. Казалось бы, множество это вполне четко определено. Но принадлежит ли этому множеству сам парикмахер? Если да, то, значит, он *не бреется сам*. Но тогда его должен брить парикмахер, т. е. он сам. Выходит, что он все же *бреется сам*, т. е. он не принадлежит множеству $\boldsymbol{B}$. Но в таком случае парикмахер не должен его брить, т. е. он *не бреется сам*. Поэтому он все же *принадлежит* множеству $\boldsymbol{B}$. Получающийся парадокс показывает, что множество $\boldsymbol{B}$ не определено корректно.

Вот еще пример. Зададим $\boldsymbol{A}$ как множество всех натуральных чисел, которые можно определить при помощи не более двадцати слов русского языка. Это множество конечно (поскольку слов в русском языке имеется лишь конечное множество). Пусть это будут числа $n_1, n_2, ..., n_k$. Возьмем теперь число $m = n_1 + n_2 + ... + n_k$. Иначе говоря, m есть *сумма всех натуральных чисел, каждое из которых определяется не более чем двадцатью словами русского языка*. Поскольку m описано фразой, содержащей менее двадцати слов, это число принадлежит множеству $\boldsymbol{A}$, т. е. совпадает с одним из чисел $n_1, n_2, ..., n_k$. Но это невозможно, поскольку число $m = n_1 + n_2 + ... + n_k$ *больше* каждого из чисел $n_1, n_2, ..., n_k$. Полученное противоречие показывает, что множество $\boldsymbol{A}$ не определено корректно.

В рассмотренном примере мы употребили термин «конечное множество», имея в виду множество, содержащее лишь конечное число элементов. Смысл этого термина совершенно ясен. Но в математике рассматриваются и бесконечные множества, т. е. такие, которые содержат *бесконечно много* элементов. Например, множество $\boldsymbol{N}$ всех натуральных чисел бесконечно. Также бесконечным является и множество всех рациональных чисел (т. е. чисел, представимых в виде отношения $\frac{p}{q}$ двух целых чисел, где знаменатель q отличен от нуля), множество всех действительных чисел, множество всех многочленов, множество всех окружностей на плоскости и т. д. Если множество конечно и притом содержит небольшое число элементов, то его можно задать явным перечислением всех его элементов. Для этого используют фигурные скобки. Например, рассмотренное выше множество $\boldsymbol{F}$ всех цифр можно задать следующим образом:

$$\boldsymbol{F} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$

Иногда приходится рассматривать *пустое множество*, т. е. такое, которое не содержит ни одного элемента. Его обозначают символом $\varnothing$. Например, множество людей, у которых жизнь продолжалась более тысячи лет, пусто (даже с учетом библейских сказаний). Необходимость рассмотрения пустого множества связана с тем, что иногда о вполне корректно определенном множестве мы заранее не знаем, содержит ли оно хотя бы один элемент. Интересный пример такого рода дает так называемая *Великая теорема Ферма*. Чтобы рассказать о ней, мы сначала рассмотрим более простой вопрос.

Известно, что существуют такие тройки натуральных чисел, в которых сумма квадратов двух из них равна квадрату третьего, т. е. натуральные числа x, y, z, удовлетворяющие уравнению

$$x^2 + y^2 = z^2. \qquad (1)$$

Согласно теореме Пифагора, такие числа позволяют построить прямоугольный треугольник с целочисленными длинами сторон (рис. 15). Поэтому такие тройки чисел называют *пифагоровыми*. Например, равенства $3^2 + 4^2 = 5^2$, $5^2 + 12^2 = 13^2$, $8^2 + 15^2 = 17^2$ дают пифагоровы тройки 3, 4, 5; 5, 12, 13; 8, 15, 17. В древнем Египте первый из этих треугольников (рис. 16) применялся для разметки прямых углов на местности, и поэтому его называют *египетским треугольником*. Пифагоровым треугольникам придавали в античном мире мистические свойства. Но сейчас все тройки пифагоровых чисел известны. Именно, любая такая тройка может быть получена по формулам

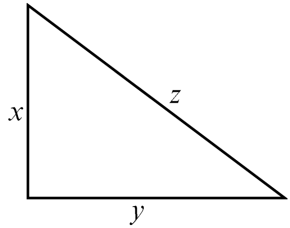


Рис. 15

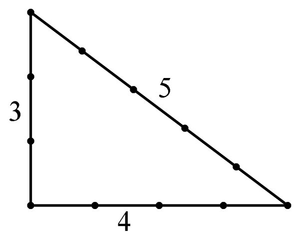


Рис. 16

$$x = k(m^2 - n^2),$$
$$y = 2kmn,$$
$$z = k(m^2 + n^2),$$

где k, m, n — натуральные числа (причем $m \neq n$).

Решение неопределенных уравнений (т. е. содержащих более одного неизвестного в одном уравнении) в натуральных числах (как, например, уравнения (1)) изучалось древнегреческим математиком Диофантом, и потому эти уравнения называют *диофантовыми уравнениями*. Великий математик Пьер Ферма, основатель теории чисел и один из создателей математического анализа, доказал ряд важных теорем о свойствах чисел. В частности, он рассмотрел диофантово уравнение

$$x^n + y^n = z^n \tag{2}$$

при $n > 2$, т. е. задачу о нахождении натуральных чисел x, y, z, которые удовлетворяют соотношению (2) при заданном натуральном $n > 2$. На полях книги Диофанта «Арифметика» Ферма сформулировал эту задачу и написал, что ни при каком натуральном $n > 2$ уравнение (2) *не имеет* натуральных решений. Иными словами, *множество таких натуральных* $n > 2$, *для которых уравнение* (2) *имеет хотя бы одно решение в натуральных числах*, *является пустым множеством*.

Ферма приписал к формулировке этого утверждения, что он нашел удивительное его доказательство, но поля книги слишком узки, чтобы его уместить. В других бумагах Ферма доказательство найдено не было. Долгие годы многие математики пытались доказать это утверждение, названное «Великой теоремой Ферма». Она была установлена для $n = 3$, затем для всех n до ста, затем еще для многих n, но полного доказательства найти не удавалось. Таким образом, было неизвестно заранее, является ли множество натуральных чисел $n > 2$, для которых уравнение (2) имеет хотя бы одно натуральное решение, пустым или непустым. И лишь в 1993 году решение получил американский математик Эндрю Уайлс: он установил, что это множество в самом деле пусто, т. е. теорема, которая была сформулирована Ферма, верна.

В заключение укажем еще одно бесконечное множество, которое имеет важное значение в математике. Напомним, что натуральное число $p > 1$ называется *простым*, если оно не имеет других натуральных делителей, кроме 1 и p. Вот несколько первых простых чисел:

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 37,\ 41,\ 43,\ 47,\ \ldots\ . \tag{3}$$

Получать простые числа можно при помощи следующего приема, называемого *решетом Эратосфена*. Мы выписываем какое-то количество натуральных чисел; скажем, числа от 1 до 50:

~~1~~, $\underline{2}$, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50.

Здесь первое число 1 вычеркнуто, поскольку оно не считается простым (простое число должно быть больше единицы). Следующее число, т. е. 2, простое — оно подчеркнуто. Теперь можно вычеркнуть все числа, делящиеся на 2 (т. е. четные числа 4, 6, 8, ...), — простыми они не являются, поскольку делятся не только на себя и единицу, но также на 2. Мы их вычеркиваем:

~~1~~, $\underline{2}$, $\underline{3}$, ~~4~~, 5, ~~6~~, 7, ~~8~~, 9, ~~10~~, 11, ~~12~~, 13, ~~14~~, 15, ~~16~~, 17, ~~18~~, 19, ~~20~~, 21, ~~22~~, 23, ~~24~~, 25, ~~26~~, 27, ~~28~~, 29, ~~30~~, 31, ~~32~~, 33, ~~34~~, 35, ~~36~~, 37, ~~38~~, 39, ~~40~~, 41, ~~42~~, 43, ~~44~~, 45, ~~46~~, 47, ~~48~~, 49, ~~50~~,

а первое из оставшихся чисел, т. е. 3, подчеркиваем — оно простое. Теперь вычеркиваем все числа, кратные числу 3, т. е. 6, 9, 12, ... Получается следующая картина:

~~1~~, $\underline{2}$, $\underline{3}$, ~~4~~, $\underline{5}$, ~~6~~, 7, ~~8~~, ~~9~~, ~~10~~, 11, ~~12~~, 13, ~~14~~, ~~15~~, ~~16~~, 17, ~~18~~, 19, ~~20~~, ~~21~~, ~~22~~, 23, ~~24~~, 25, ~~26~~, ~~27~~, ~~28~~, 29, ~~30~~, 31, ~~32~~, ~~33~~, ~~34~~, 35, ~~36~~, 37, ~~38~~, ~~39~~, ~~40~~, 41, ~~42~~, 43, ~~44~~, ~~45~~, ~~46~~, 47, ~~48~~, 49, ~~50~~,

в которой снова подчеркнуто первое оставшееся число, т. е. 5, — оно простое. Продолжая таким образом, мы постепенно вычеркиваем все числа, не являющиеся простыми. Остаются лишь простые числа — это и дает последовательность чисел (3), т. е. простых чисел, меньших 50.

Естественно возникает вопрос, *как много* имеется простых чисел? И прежде всего: конечное или бесконечное множество чисел остается после просеивания чисел решетом Эратосфена, т. е. конечно или бесконечно множество всех простых чисел? Замечательный результат — доказательство *бесконечности* множества простых чисел — был установлен еще в древней Греции; он содержится в книге Евклида «Начала», написанной свыше двух тысяч лет назад. Рассмотрим это изящное евклидово доказательство.

Допустим, что простых чисел существует лишь конечное число. Обозначим их в порядке возрастания через p_1, p_2, ..., p_k ($p_1 = 2$, $p_2 = 3$ мы уже нашли, дальше идут еще какие-то простые числа, а последнее p_k — наибольшее простое число).

Рассмотрим теперь число

$$n = p_1 p_2 \dots p_k + 1. \tag{4}$$

Как и всякое натуральное число, большее 1, оно либо само является простым, либо может быть разложено на простые множители, т. е. представлено в виде произведения нескольких сомножителей, каждый из которых есть одно из чисел p_1, p_2, ..., p_k. Но число *n не* *делится* ни на одно из чисел $p_1, p_2, ..., p_k$, что видно из его выражения (4) (поскольку $p_1 p_2 ... p_k$ делится на любое из чисел $p_1, p_2, ..., p_k$, а после прибавления единицы получается число, не делящееся ни на одно из них). Таким образом, остается единственная возможность: число n само является простым, т. е. *совпадает* с одним из чисел $p_1, p_2, ..., p_k$. Но это невозможно, поскольку, согласно (4), число n больше p_k, т. е. *больше* каждого из чисел $p_1, p_2, ..., p_k$. Итак, предположение о том, что имеется лишь конечное число простых чисел, приводит к противоречию, т. е. это предположение должно быть отброшено как ложное. Значит, множество простых чисел является бесконечным.

Задачи и упражнения

1. Мадемуазель Рембо любит домашних животных. Известно, что у нее не менее трех животных. Все ее животные, кроме двух — собаки; все кроме двух — кошки; все кроме двух — попугаи; все, кроме собак, кошек и попугаев — тараканы. Опишите множество животных у мадемуазель Рембо.

2. В люстре 5 лампочек. Переключатель имеет 6 положений, при которых горит разное количество лампочек: от 0 до 5. Однажды несколько лампочек перегорело. Может ли человек, не знающий схемы работы переключателя, определить множество перегоревших лампочек?

3. В комнате находятся 12 человек. Некоторые из них всегда говорят правду, а остальные всегда лгут. Каждый сказал по фразе. Первый сказал: «Здесь нет ни одного правдивого человека», второй: «Здесь не более одного правдивого человека», третий: «Здесь не более двух правдивых людей», и т. д., двенадцатый: «Здесь не более одиннадцати правдивых людей».

Определите, кто из них принадлежит множеству правдивых людей.

4. По полю проходит прямолинейная дорога. Человек, находящийся на дороге в точке A, может идти по полю со скоростью не больше 3 км/ч, а по дороге со скоростью не больше 6 км/ч. Определите множество точек, в которых он может оказаться через час.

5. На плоскости дан равносторонний треугольник ABC. Укажите множество точек M, для которых оба треугольника ABM и ACM — равнобедренные.

8. Взаимно однозначное соответствие

Студенты приехали на машинах копать картошку. Привезли и лопаты. Как узнать, хватит ли лопат? Конечно, можно пересчитать студентов и отдельно пересчитать лопаты. Но организаторы мероприятия решили вопрос иначе. Они предложили студентам подходить по одному и брать лопаты. Очень быстро все лопаты были разобраны и их, по счастливой случайности, как раз хватило. Этим было установлено, что студентов и лопат *одинаковое* количество. Как это было выяснено? Было установлено *взаимно однозначное соответствие* между множеством студентов и множеством лопат. Это означает, что каждому элементу первого множества (студенту) был поставлен в соответствие один определенный элемент второго множества (лопата, которую взял этот студент), причем каждый элемент второго множества (лопата) был поставлен в соответствие одному и только одному элементу первого множества.

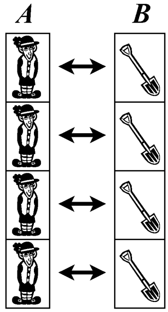


Рис. 17

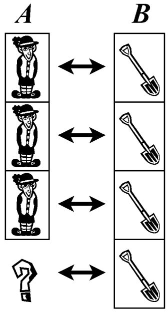


Рис. 18

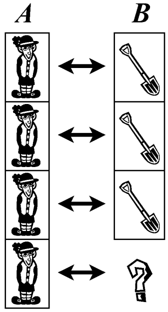


Рис. 19

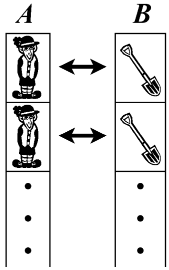


Рис. 20

На рис. 17 схематически показано это взаимно однозначное соответствие. Если бы был некоторый избыток лопат (рис. 18) или если бы некоторым студентам не хватило лопат (рис. 19), то взаимно однозначного соответствия не было бы. Возможность установления взаимно однозначного соответствия между элементами двух множеств означает, что в этих множествах «одинаково много» элементов. Такие множества называются *эквивалентными.*

Вот другой пример. На танцплощадке собрались юноши и девушки. Оркестр заиграл вальс, и юноши стали приглашать девушек на танец. Если окажется, что все пришедшие на танцплощадку распределились по парам, то этим будет установлено взаимно однозначное соответствие между множеством всех юношей и множеством всех девушек: каждому элементу первого множества (юноше) поставлен в соответствие в точности один элемент второго множества (девушка, которую он пригласил танцевать), причем каждый элемент второго множества поставлен в соответствие одному и только одному элементу первого множества. Это соответствие показывает, что оба множества эквивалентны, т. е. юношей присутствует столько же, сколько девушек.

Заметим, что каждый из нас многие тысячи раз устанавливает взаимно однозначное соответствие между множествами. Ребенок, трогая пальчиком один за другим разложенные перед ним предметы, производит пересчет, т. е. произносит «один», «два», «три», «четыре». Этим он устанавливает взаимно однозначное соответствие между множеством расположенных перед ним предметов и множеством $\{1, 2, 3, 4\}$, т. е. между заданным множеством предметов и *отрезком* натурального ряда, состоящим из нескольких первых натуральных чисел. Именно наличие универсального счетного множества, т. е. натурального ряда, позволяет сегодня ребенку легче постичь абстракцию числа. Число, названное последним (в данном случае «четыре»), отвечает на вопрос, *сколько* предметов (элементов) имеется в заданном множестве.

Натуральные числа позволяют сформировать первоначальное понятие «меньше» (или «больше»). Предположим, что мы пытаемся установить взаимно однозначное соответствие между двумя множествами предметов, постепенно группируя их парами: «элемент из ***A***» ↔ «соответствующий элемент из ***B***», «элемент из ***A***» ↔ «элемент из ***B***» и т. д. (рис. 20). Если при этом окажется, что мы исчерпали все элементы множества ***A***, но в множестве ***B*** остались «лишние» элементы (не поставленные в соответствие элементам из ***A***, как на рис. 18), то мы говорим, что в множестве ***A*** *меньше* элементов, чем в ***B*** (а в множестве ***B*** *больше* элементов, чем в ***A***). В этом случае установлено взаимно однозначное соответствие между множеством ***A*** и *частью* множества ***B***. «Часть меньше целого», гласит один из принципов, выдвинутых великим древнегреческим мыслителем Аристотелем, создателем логики. А универсальное счетное множество (натуральный ряд) позволяет с помощью пересчета установить, в каком множестве меньше элементов, не группируя предметы парами (что было бы неудобно, если предметы громоздкие или тяжелые). Если при пересчете оказалось, что в множестве ***A*** восемь элементов, а в множестве ***B*** — десять, то мы заключаем, что в ***A*** *меньше* элементов, чем в ***B***, поскольку число 8 стоит в натуральном ряду *раньше* числа 10. И мы убеждены, что в этом случае можно установить взаимно однозначное соответствие между всем множеством ***A*** и *частью* множества ***B***.

Все вышесказанное относится к *конечным* множествам. В древности рассматривали почти исключительно конечные множества. Построение *бесконечного* натурального ряда чисел и замечательное доказательство Евклида *бесконечности* множества простых чисел являются первыми прорывами в математику бесконечного.

Особенно характерно разграничение конечного и бесконечного в древнегреческой геометрии. Понятие «геометрического места точек», восходящее к Платону, является отголоском этого разграничения. Поскольку геометрическая точка имеет «нулевые» размеры, какое бы (конечное!) число точек мы ни «прикладывали» друг к другу, ничего, кроме точки, получиться не может. Поэтому, по мнению древнегреческих мыслителей, прямая (или окружность, или иная линия) не может «состоять» из точек, не может быть представлена как *множество* точек, а есть лишь «место», где могут располагаться точки.

Евклиду (или его предшественникам — скорее всего, Евдоксу) принадлежит хорошо известная теорема, которую можно сформулировать следующим образом: *множество всех точек, равноудаленных от двух заданных точек A и B, есть прямая, перпендикулярная отрезку AB и проходящая через его середину* (рис. 21).

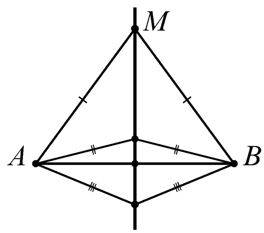


Рис. 21

Иными словами, точка M в том и только в том случае одинаково удалена от точек A и B, если она расположена на указанном перпендикуляре, и в этом смысле этот перпендикуляр есть «место» таких точек. Сегодня мы говорим, что прямая есть бесконечное *множество* точек (и вообще любая геометрическая фигура есть некоторое множество точек), но древнегреческим ученым такая идеология была чужда.

Впервые перешагнуть пропасть между конечным и бесконечным попытался в XIX столетии чешский ученый Бернард Больцано. Именно он пришел к понятию взаимно однозначного соответствия. Но трудности, к которым вело это понятие применительно к бесконечным множествам, остановили его. Несмотря на много интересных идей и результатов, содержавшихся в труде Б. Больцано (опубликованном лишь через много лет после его смерти), задача построения общей теории множеств им не была решена. Ее решение дал выдающийся немецкий математик Георг Кантор. О трудностях, связанных с рассмотрением бесконечных множеств, которые обнаружил Больцано и которые были преодолены Кантором, мы и поговорим в следующем пункте.

Задачи и упражнения

6. На балу каждый кавалер станцевал ровно с тремя дамами, а каждая дама — ровно с тремя кавалерами. Докажите, что дам и кавалеров было поровну.

7. У кондуктора автобуса две катушки билетов: в одной билеты с номерами от 573000 до 573999, в другой — с номерами от 426000 до 426999. В какой из катушек больше «счастливых» билетов, т. е. таких, что сумма первых трех цифр номера билета равняется сумме следующих трех цифр?

8. На окружности отмечены 2001 белая точка и одна красная. Рассмотрим всевозможные многоугольники с вершинами в этих точках. Каких многоугольников будет больше — с красной точкой или без нее?

9. Число 222122111121 получается, если в некотором слове заменить буквы на их номера в русском алфавите. Какое это слово?

10. Имеется 3 синих и 5 красных палочек разной длины. Сумма длин синих палочек равна сумме длин красных. Можно ли распилить эти палочки так, чтобы их потом можно было расположить парами палочек равной длины и разного цвета?

9. Счетные множества

В одном театре было бесконечно много мест, и они были занумерованы натуральными числами: первое место, второе, ..., n-е место, ... Соответственно и в гардеробе было бесконечно много крючков: № 1, № 2, № 3, ... Однажды на интересную премьеру все билеты были раскуплены, и в театр пришло бесконечно много зрителей: первый (т. е. обладатель билета на место № 1), второй, ... День был дождливый, все зрители пришли с зонтиками, которые они повесили в гардеробе (в порядке самообслуживания) на соответствующие крючки: первый — на крючок № 1, второй зритель — на крючок № 2 и т. д. После спектакля, впопыхах, первый зритель взял зонтик с крючка № 2, второй зритель взял зонтик с крючка № 3, n-й зритель взял зонтик с крючка с номером $(n+1)$, ... (рис. 22). В результате все зрители ушли с зонтиками, но (к удивлению администрации этого бесконечного театра) на крючке № 1 остался «лишний» зонтик. Таким образом, если мы обозначим зрителей символами s_1, s_2, ..., а

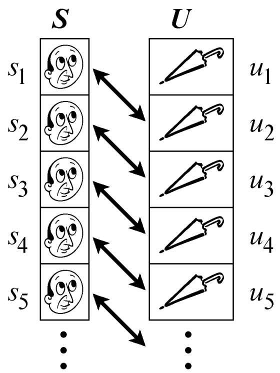


Рис. 22

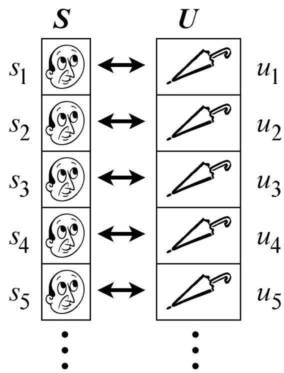


Рис. 23

зонтики — символами u_1, u_2, ..., то ситуацию можно описать следующим образом. Каждый зритель пришел со своим зонтиком, т. е. имелось взаимно однозначное соответствие между множеством $\boldsymbol{S} = \{s_1, s_2, ...\}$ всех зрителей и множеством $\boldsymbol{U} = \{u_1, u_2, ...\}$ всех зонтиков (рис. 23). Однако после спектакля было установлено другое соответствие (рис. 22), при котором взаимно однозначно соответствовали друг другу множество $\boldsymbol{S} = \{s_1, s_2, ...\}$ и *часть* $\{u_2, u_3, ...\}$ множества $\boldsymbol{U}$ всех зонтиков. Выходит, что часть $\{u_2, u_3, ...\}$ множества $\boldsymbol{U}$ всех зонтиков содержит «столько же» элементов, сколько и все множество $\boldsymbol{U}$ (поскольку каждое из этих множеств может быть приведено во взаимно однозначное соответствие с множеством $\boldsymbol{S}$ всех зрителей).

Это означает, что аристотелевский принцип «часть меньше целого», подмеченный при рассмотрении конечных множеств, рушится при переходе к бесконечным множествам. Вообще, в любом бесконечном множестве $\boldsymbol{M}$ можно выбрать такую его часть $\boldsymbol{M}_i$, (содержащую *не все* элементы множества $\boldsymbol{M}$), что существует взаимно однозначное соответствие между $\boldsymbol{M}$ и $\boldsymbol{M}_i$, т. е. между множеством и его частью. Это — характерное отличие бесконечных множеств от конечных. (Не захочет ли читатель попытаться доказать это утверждение?)

Прежде чем переходить к осмыслению этого факта, опишем события, которые происходили в этом «бесконечном» театре во время других спектаклей, чтобы лучше представить себе специфику бесконечных множеств.

Во время одного из спектаклей также была дождливая погода, и все зрители пришли с зонтиками. Пока первый зритель беседовал с режиссером после спектакля, остальные забрали зонтики и ушли. При этом второй зритель взял зонтик № 1, третий зритель взял второй зонтик и т. д. (рис. 24). В результате первый зритель остался без зонтика! В один из дней (также дождливый) после спектакля первый

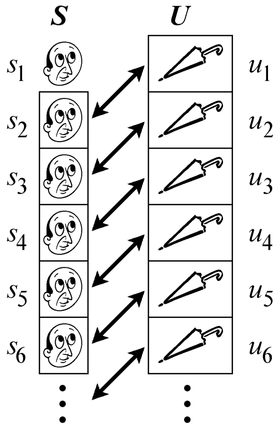


Рис. 24

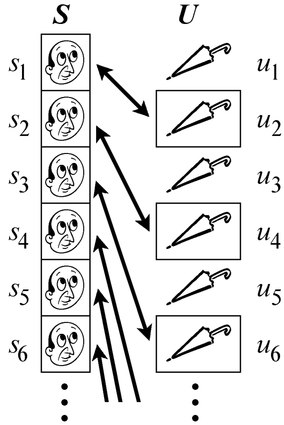


Рис. 25

зритель взял второй зонтик, второй взял четвертый зонтик, ..., n-й зритель взял зонтик с крючка с номером $2n$, ... (рис. 25). В результате все ушли с зонтиками, но в гардеробе осталось бесконечно много зонтиков (а именно, зонтики с *нечетными* номерами).

На следующий день все пришли без зонтиков по случаю солнечной погоды. Но за время спектакля погода испортилась, хлынул дождь. Однако администрация успокоила зрителей, сказав, что в гардеробе как раз имеется бесконечно много не востребованных вчера зонтиков. Значит, натуральный ряд как бы содержит «вдвое больше» элементов, чем всего в нем имеется...

Теперь читатель подготовлен к восприятию некоторых основных определений, связанных с бесконечными множествами. Как мы говорили, два множества называются *эквивалентными,* если между ними возможно установить взаимно однозначное соответствие. Если множества $\boldsymbol{P}$ и $\boldsymbol{Q}$ эквивалентны, это вовсе не означает, что при любой попытке сгруппировать их элементы парами (один элемент из $\boldsymbol{P}$ и один элемент из $\boldsymbol{Q}$, еще один элемент из $\boldsymbol{P}$ и еще один из $\boldsymbol{Q}$ и т. д.) мы непременно получим взаимно однозначное соответствие. Так, хотя в рассмотренном выше примере множества $\boldsymbol{S}$ и $\boldsymbol{U}$ эквивалентны (как показывает рис. 23), но изображенная на рис. 22 попытка сгруппировать элементы парами *не приводит* к установлению взаимно однозначного соответствия.

Однако наличие неудачных попыток ничего не означает. Если существует хотя бы один «удачный» способ, при котором *устанавливается* взаимно однозначное соответствие между рассматриваемыми множествами, то эти множества эквивалентны.

Каждое множество, эквивалентное натуральному ряду $\boldsymbol{N} = \{1, 2, 3, ...\}$, называется *счетным.* Иными словами, множество $\boldsymbol{M}$ является счетным, если оно бесконечно, и при этом все его элементы можно

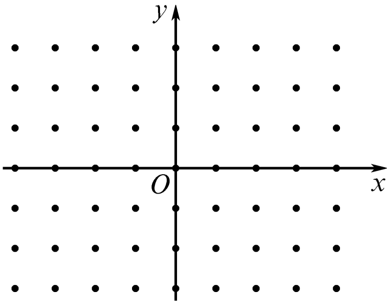


Рис. 26

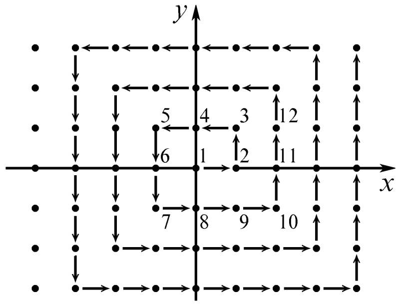


Рис. 27

перенумеровать натуральными числами, т. е. записать это множество в виде $\boldsymbol{M} = \{m_1, m_2, m_3, ...\}$.

Например, множество всех целых чисел (содержащее натуральные числа, нуль и отрицательные целые числа) является счетным. В самом деле, все целые числа встречаются в следующей последовательности:

$$0,\ 1,\ -1,\ 2,\ -2,\ 3,\ -3,\ ...$$

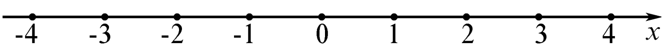
и теперь не представляет труда их все перенумеровать.

Несколько неожиданным является утверждение о том, что все *целочисленные точки* координатной плоскости (т. е. точки, у которых обе координаты являются целыми числами, рис. 26) также составляют *счетное* множество. Однако это становится понятным при рассмотрении того способа нумерации, который показан на рис. 27: идя вдоль спиралевидной ломаной, мы последовательно встретим (и перенумеруем) все целочисленные точки координатной плоскости.

Наконец, еще один пример. Целочисленные точки расположены на координатной прямой равномерно и изолированно (рис. 28). Рациональные же числа, т. е. числа, представимые в виде дробей $\frac{p}{q}$ с целыми p, q ($q \neq 0$), изображаются на координатной прямой всюду *плотно*, т. е. на любом отрезке, как бы мал он ни был, имеются рациональные точки.

Интуитивно кажется, что рациональных точек «существенно больше», чем целых. Во всяком случае целые числа составляют лишь часть множества всех рациональных чисел и, на первый взгляд, часть лишь «незначительную». Тем не менее, множество всех рациональных чисел *счетно*, т. е. их все можно перенумеровать натуральными числами, и в этом смысле рациональных чисел *столько же*, сколько и целых (или натуральных).

Рис. 28



В самом деле, назовем *весом* дроби $\frac{p}{q}$ натуральное число $|p| + |q|$, т. е. сумму модулей ее числителя и знаменателя. Например, обе дроби $\frac{-2}{3}$ и $\frac{1}{4}$ имеют вес 5. Каждое рациональное число может быть представлено (и притом однозначно) в виде несократимой дроби с натуральным знаменателем. Например, все дроби $\frac{2}{-3}, \frac{-2}{3}, \frac{4}{-6}, \frac{-6}{9}$ представляют одно и то же рациональное число, но только дробь $\frac{-2}{3}$ дает запись этого числа в виде несократимой дроби с натуральным знаменателем. Теперь будем последовательно нумеровать все несократимые дроби с натуральными знаменателями (т. е. рациональные числа), беря сначала дроби веса 1, затем дроби веса 2, веса 3 и т. д. Поскольку дробей каждого веса имеется лишь конечное число, мы постепенно перенумеруем все рациональные числа. Вот в какой последовательности они будут встречаться:

$$\underbrace{\frac{0}{1}}_{\text{вес 1}},\ \underbrace{\frac{1}{1}, \frac{-1}{1}}_{\text{вес 2}},\ \underbrace{\frac{1}{2}, \frac{-1}{2}, \frac{2}{1}, \frac{-2}{1}}_{\text{вес 3}},\ \underbrace{\frac{1}{3}, \frac{-1}{3}, \frac{3}{1}, \frac{-3}{1}}_{\text{вес 4}},\ \underbrace{\frac{1}{4}, \frac{-1}{4}, \frac{2}{3}, \frac{-2}{3}, \frac{3}{2}, \frac{-3}{2}, \frac{4}{1}, \frac{-4}{1}}_{\text{вес 5}}, \ldots$$

Этим и устанавливается, что множество всех рациональных чисел счетно.

Задачи и упражнения

11. Докажите, что множество всех рациональных точек на координатной плоскости (т. е. точек, у которых обе координаты рациональны) счетно.

12. Докажите, что множество окружностей, имеющих центры в рациональных точках координатной плоскости и рациональные радиусы, счетно.

13. Докажите, что множество всех многочленов

$$a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

данной степени n с рациональными коэффициентами счетно.

14. Фигуру на плоскости назовем «буквой Т», если она состоит из двух перпендикулярных отрезков, причем второй имеет один из концов в середине первого отрезка. Докажите, что любое бесконечное множество непересекающихся «букв Т» на плоскости счетно.

15. Число называется алгебраическим, если оно является корнем некоторого многочлена с целыми коэффициентами (например, числа 1, $\sqrt{2}$, $\sqrt[5]{3} + \sqrt[3]{2}$, $\sqrt[7]{6}$ являются алгебраическими). Докажите, что множество алгебраических чисел счетно.

10. Понятие мощности множества

Выше мы имели ряд примеров счетных множеств. Естественно возникает вопрос: все ли бесконечные множества счетны или же существуют несчетные множества? Ответ был дан Георгом Кантором. Он доказал, что *множество **R** всех действительных чисел несчетно.*

Приведем его весьма элегантное доказательство, известное под названием *канторовского диагонального процесса.*

Для простоты рассмотрим не все действительные числа, а лишь те, которые расположены между 0 и 1, и докажем, что это множество несчетно (следовательно, множество всех действительных чисел подавно несчетно). Допустим, напротив, что множество $\boldsymbol{D}$ всех действительных чисел x, удовлетворяющих условию $0 < x < 1$, счетно, и занумеруем все эти числа:

$$x_1,\ x_2,\ \dots,\ x_n,\ \dots \tag{5}$$

Каждое действительное число записывается в виде бесконечной десятичной дроби (если десятичная дробь конечна, то это означает, что, начиная с некоторого места, все цифры десятичной записи равны нулю; например, $\frac{2}{5} = 0{,}4 = 0{,}4000\dots$). Запишем числа (5) в виде десятичных дробей; пусть, скажем,

$$\begin{aligned} x_1 &= 0{,}\underline{4}83751\dots \\ x_2 &= 0{,}4\underline{7}9512\dots \\ x_3 &= 0{,}03\underline{8}179\dots \\ x_4 &= 0{,}576\underline{3}09\dots \\ &\dots\dots\dots\dots\dots \end{aligned} \tag{6}$$

В этой таблице подчеркнуты цифры, стоящие *на диагонали.* Создадим теперь бесконечную десятичную дробь (число z) по следующему правилу: если у числа x_n на диагональном (т. е. n-м) месте после запятой стоит цифра 7, то у числа z мы ставим на n-м месте цифру 3; во всех остальных случаях у числа z стоит на n-м месте цифра 7. Так, в рассматриваемой таблице, первые четыре строки которой выписаны выше, только у числа x_2 диагональная (подчеркнутая) цифра равна 7, а у чисел x_1, x_3, x_4 диагональные цифры отличны от 7. Поэтому в числе z вторая цифра равна 3, а остальные цифры (из первых четырех) равны 7:

$$z = 0{,}7377\dots$$

И дальше, если у очередного числа из (6) диагональная цифра равна 7, то вместо нее мы поставим на соответствующем месте в z цифру 3, а если диагональная цифра не 7, то на соответствующем месте ставим у числа z цифру 7. Таким образом, на первом месте у числа z стоит цифра, *отличная* от первой цифры числа x_1; на втором месте у числа z стоит цифра, *отличная* от второй цифры числа x_2, и т. д. Вообще, на n-м месте у числа z стоит цифра, *отличная* от n-й цифры числа x_n. Значит, число z *не совпадает* с x_1 (поскольку первые цифры у них различны). Число z *не совпадает* и с x_2 (у них вторые цифры различны). Вообще, z не совпадает *ни с одним числом* x_n (поскольку на n-м месте у числа z и у числа x_n стоят различные цифры). Таким образом, z не содержится среди чисел (5). Однако это противоречит тому, что среди чисел (5) имеются *все* действительные числа, заключенные между 0 и 1.

Мы видим, что предположение о счетности множества $\boldsymbol{D}$ приводит к противоречию. Значит, это предположение ложно, т. е. множество $\boldsymbol{D}$ *несчетно*. Поэтому и множество $\boldsymbol{R}$ всех действительных чисел несчетно.

Доказанная теорема послужила поводом для введения понятия *мощности* множества. Вспомним, что, рассматривая класс эквивалентных между собой конечных множеств, мы с помощью абстракции отождествления приходим к понятию *числа*. Например, рассматривая множество всех ножек обычного стола и рассматривая всевозможные эквивалентные (т. е. содержащие столько же элементов) множества, мы приходим, абстрагируясь от всех конкретных свойств множеств и элементов, к понятию числа 4. Подобно этому для эквивалентных бесконечных множеств мы говорим, что у них одинаковая *мощность* (аналог «числа элементов»). Все счетные множества эквивалентны между собой, т. е. имеют одну и ту же мощность. Она обозначается, согласно Кантору, символом $\aleph_0$; буква $\aleph$ («алеф») — первая в древнееврейском алфавите.

Мощность множества $\boldsymbol{R}$ всех действительных чисел обозначается готической буквой $\mathfrak{C}$ («цэ»); ее называют мощностью *континуума* (от латинского слова *continuum* — непрерывный). Согласно доказанной выше теореме Кантора, действительных чисел *больше*, чем элементов любого счетного множества, т. е. $\mathfrak{C} > \aleph_0$. В свою очередь, мощность $\aleph_0$ *больше* мощности любого конечного множества: $\aleph_0 > n$ для любого натурального n.

Заметим, что $\aleph_0$ — *наименьшая* бесконечная мощность. Это означает, что из любого бесконечного множества можно выбрать счетную часть, или, как говорят, счетное *подмножество*. В этом нетрудно убедиться. В самом деле, пусть $\boldsymbol{M}$ — бесконечное множество. Возьмем произвольный его элемент и обозначим его через a_1. Так как множество $\boldsymbol{M}$ бесконечно, то кроме a_1 в нем имеются еще элементы. Пусть a_2 — один из них. Кроме a_1 и a_2 в множестве $\boldsymbol{M}$ имеются еще элементы (поскольку $\boldsymbol{M}$ бесконечно); обозначим через a_3 какой-либо из них. Продолжая таким образом, мы построим счетное множество $\boldsymbol{M}_1 = \{a_1, a_2, a_3, \ldots\}$, содержащееся в множестве $\boldsymbol{M}$ (т. е. являющееся его *подмножеством*). Значит, множество $\boldsymbol{M}$ содержит *не меньше* элементов, чем счетное множество $\boldsymbol{M}_1$, т. е. мощность множества $\boldsymbol{M}$ не меньше $\aleph_0$.

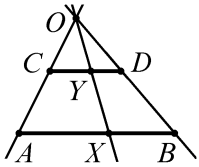


Рис. 29

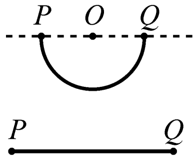


Рис. 30

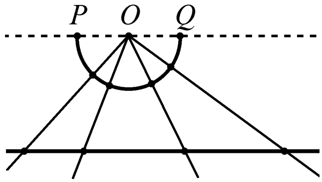


Рис. 31

Что же касается мощности континуума $\mathfrak{C}$, то она не является «наибольшей» мощностью. Канторовский диагональный процесс (в несколько усложненной форме) позволяет доказать, что *каково бы ни было бесконечное множество* M, *можно построить множество, которое имеет еще большую мощность, чем* M. Математики называют мощности также *кардинальными числами*. Это как бы «бесконечно большие» числа. Сказанное выше означает, что существует *бесконечно много* различных кардинальных чисел.

Весьма интересна связь мощностей с геометрическими понятиями. Для начала рассмотрим один *софизм*. Софизмом называют рассуждение, которое на первый взгляд выглядит вполне логичным и не вызывающим сомнения, тогда как вывод, к которому приводит это рассуждение, является явно абсурдным. *Раскрыть софизм* — значит указать ошибочное место в этом внешне логичном рассуждении и тем самым показать, почему это рассуждение приводит к абсурдному выводу.

Софизм, о котором идет речь, утверждает, что «все отрезки имеют одинаковую длину». В самом деле, возьмем два отрезка AB и CD, не равные между собой, и расположим их параллельно друг другу (рис. 29). Прямые AC и BD пересекаются в некоторой точке O. Теперь для любой точки X отрезка AB мы проведем прямую OX, которая пересечет второй отрезок в некоторой точке Y. Эту точку Y мы и поставим в соответствие точке $X \in AB$. Этим устанавливается *взаимно однозначное соответствие* между отрезками AB и CD, т. е. эти отрезки содержат *одинаковое* количество точек. Значит, они имеют *одинаковую длину*.

Раскрытие этого софизма несложно. Действительно, оба отрезка «содержат одинаковое количество точек», т. е. имеют одинаковую мощность. Однако мощность и длина — понятия разные, и из совпадения мощностей не вытекает равенство длин.

Заметим, что вся бесконечная прямая содержит «столько же» точек, сколько и любой отрезок. В самом деле, если отрезок изогнуть в виде дуги полуокружности с центром O (рис. 30), то центральное проектирование этой полуокружности из центра O на прямую, параллельную диаметру PQ полуокружности (рис. 31), устанавливает взаимно однозначное соответствие между точками первоначального отрезка и всеми точками прямой. Правда, концы отрезка (или концы P, Q полуокружности) не поставлены в соответствие никаким точкам прямой, как видно из рис. 31, т. е. отрезок содержит «на две точки

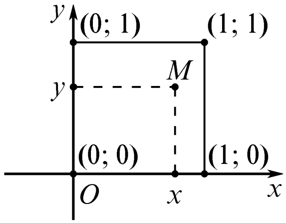


Рис. 32

больше», чем прямая. Однако это несущественно, так как добавление к бесконечному множеству конечного числа точек не изменяет его мощности. (Не может ли читатель доказать это?) Таким образом, любой отрезок имеет ту же мощность, что и прямая, т. е. мощность любого отрезка равна $\mathfrak{C}$. Вообще, любая ломаная, любая окружность, любая ее дуга имеют мощность $\mathfrak{C}$.

Обратимся теперь к плоским фигурам. Кажется наглядно «очевидным», что квадрат (или круг) содержит «больше» точек, чем отрезок. Однако это не так; здесь интуиция обманывает нас. Подобно тому, как рациональных точек на плоскости столько же, сколько рациональных точек на прямой (оба множества являются счетными), так и *всех* точек на плоскости «столько же», сколько и точек на прямой. Иными словами, мощность множества всех точек плоскости (или даже пространства) равна $\mathfrak{C}$, т. е. совпадает с мощностью всех точек прямой.

Чтобы примирить читателя с его интуицией (которая, возможно, протестует против такой «несправедливости»), мы приведем канторовское доказательство того факта, что в квадрате «столько же» точек, сколько и в его стороне.

Рассмотрим на координатной плоскости квадрат с вершинами (0; 0), (1; 0), (0; 1), (1; 1) (рис. 32). Каждая точка M этого квадрата имеет две координаты x, y, причем каждая из этих координат представляет собой действительное число, принадлежащее отрезку [0; 1], т. е. $0 \le x \le 1$, $0 \le y \le 1$. Это означает, что числа x, y могут быть записаны в виде бесконечных десятичных дробей:

$$x = 0{,}a_1a_2a_3\ \ldots,\quad y = 0{,}b_1b_2b_3\ \ldots \tag{7}$$

(заметим, что мы всегда можем считать целую часть равной нулю, т. е. если $x = 1$, то мы можем написать $x = 0{,}9999\ldots$). Теперь мы «перемешаем» десятичные цифры чисел x и y, выписывая их поочередно, т. е. рассмотрим число

$$z = 0{,}a_1b_1a_2b_2a_3b_3\ \ldots$$

Это число z мы и поставим в соответствие точке M квадрата (с координатами x, y). Получается *взаимно однозначное соответствие* между всеми точками квадрата и точками отрезка [0; 1]. Это показывает, что квадрат содержит «столько же» точек, сколько и отрезок, т. е. мощность квадрата равна $\mathfrak{C}$. (Может ли читатель доказать, что в кубе «столько же» точек, сколько и в отрезке?)

Рассмотрим теперь вопрос о связи мощностей с неравенствами. Прежде всего уточним понятие *подмножества.* Мы говорим, что некоторое множество $\boldsymbol{A}$ является подмножеством множества $\boldsymbol{M}$, если каждый элемент множества $\boldsymbol{A}$ является также и элементом множества $\boldsymbol{M}$. Это выражают записью $\boldsymbol{A} \subset \boldsymbol{M}$. Заметим, что, согласно этому определению, само множество $\boldsymbol{M}$ является своим подмножеством: $\boldsymbol{M} \subset \boldsymbol{M}$. Всякое же другое подмножество $\boldsymbol{A}$ множества $\boldsymbol{M}$ (т. е. такое, что в $\boldsymbol{M}$ найдется хотя бы один элемент, не принадлежащий $\boldsymbol{A}$), называется *собственным подмножеством* множества $\boldsymbol{M}$. В аристотелевском постулате «часть меньше целого» под частью понимается именно *собственное* подмножество. На языке теории множеств аристотелевский постулат означает, что если $\boldsymbol{A}$ есть собственное подмножество множества $\boldsymbol{M}$, то мощность множества $\boldsymbol{A}$ (т. е. число его элементов) *меньше* мощности множества $\boldsymbol{M}$. Обозначая мощности этих множеств готическими буквами $\mathfrak{A}$ («а») и $\mathfrak{M}$ («эм»), мы можем записать аристотелевский постулат неравенства:

$$\mathfrak{A} < \mathfrak{M}. \tag{8}$$

Однако мы теперь знаем, что этот постулат верен лишь для *конечных* множеств.

В случае же *бесконечных* множеств из того, что $\boldsymbol{A}$ есть *собственное* подмножество множества $\boldsymbol{M}$, для их мощностей $\mathfrak{A}$ и $\mathfrak{M}$ вытекает лишь нестрогое неравенство

$$\mathfrak{A} \leq \mathfrak{M}. \tag{9}$$

Примеры мы уже имели выше. Достаточно напомнить, что хотя множество $\{2, 4, 6, ...\}$ всех четных чисел является собственным подмножеством натурального ряда $\{1, 2, 3, ...\}$, но их мощности *равны* (оба множества имеют мощность $\aleph_0$). Точно так же если отрезок PQ является *собственным* подмножеством отрезка AB, то из этого не следует, что отрезок PQ имеет меньшую мощность: в самом деле, как мы знаем, оба отрезка имеют одну и ту же мощность $\mathfrak{C}$. Таким образом, в применении к бесконечным множествам аристотелевский принцип должен быть уточнен: часть (т. е. собственное подмножество) множества $\boldsymbol{M}$ имеет *либо меньшую*, *либо ту же самую мощность*, что и целое (т. е. все множество $\boldsymbol{M}$). Иными словами, при переходе от конечных множеств к бесконечным мы должны заменить аристотелевское *строгое* неравенство (8) *нестрогим* неравенством (9).

Заметим, впрочем, что аристотелевский принцип «часть меньше целого» относится не только к *числу элементов* множества и его собственного подмножества («части»), но также к длинам отрезков, площадям фигур, объемам. О смысле аристотелевского постулата применительно к этим случаям мы еще будем иметь случай поговорить в дальнейших беседах.

Задачи и упражнения

16. Докажите, что счастливых автобусных билетов (см. упр. 7) столько же, сколько билетов, у которых сумма цифр равна 27.

17. В языке племени Ододо всего два звука: «Д» и «О». Два слова обозначают одно и то же, если одно получается из другого при помощи следующих операций: исключения идущих подряд звуков ДО или ООДД и добавления в любое место сочетания ОД. Означают ли слова ОДД и ДОО одно и то же? Какова мощность множества различных слов этого языка?

18. На плоскости имеется некоторое количество непересекающихся треугольников. Докажите, что это множество конечно или счетно. (Треугольник вместе со своим контуром содержит и внутренность этого контура.)

19. Докажите, что множество точек куба имеет ту же мощность, что и множество точек отрезка.

20. Докажите, что множество всех подмножеств непустого множества $\boldsymbol{M}$ имеет мощность большую, чем мощность самого множества $\boldsymbol{M}$.

Беседа 3. Операции над множествами

11. Пересечение множеств

Прежде чем переходить к рассмотрению операций над множествами, мы рассмотрим один очень удобный и часто применяемый прием наглядного изображения взаимоотношений между множествами. Речь идет о так называемых диаграммах Эйлера–Венна, в которых множества изображаются фигурами на плоскости (чаще всего кругами), по расположению которых можно легко судить о рассматриваемых множествах.

Так на рис. 33 некоторые два множества $\boldsymbol{A}$ и $\boldsymbol{B}$ изображены кругами, причем видно, что $\boldsymbol{B}$ содержится в множестве $\boldsymbol{A}$ и, кроме того, $\boldsymbol{B}$ является собственным подмножеством множества $\boldsymbol{A}$, т. е. не совпадает с ним. Для множеств $\boldsymbol{A}$ и $\boldsymbol{B}$, изображенных на рис. 34, также справедливо включение $\boldsymbol{B} \subset \boldsymbol{A}$, но при этом множества $\boldsymbol{A}$ и $\boldsymbol{B}$ совпадают. Таким образом, запись $\boldsymbol{B} \subset \boldsymbol{A}$ означает, что либо $\boldsymbol{B}$ является собственным подмножеством множества $\boldsymbol{A}$, либо $\boldsymbol{A} = \boldsymbol{B}$. А на рис. 35 изображены два множества, ни одно из которых не является подмножеством другого.

Пусть, например, $\boldsymbol{A}$ есть множество всех натуральных чисел, не превосходящих 10, $\boldsymbol{B}$ — множество всех нечетных натуральных чисел, не превосходящих 10, а $\boldsymbol{C}$ — множество всех простых чисел, не превосходящих 10. Иначе говоря,

$$\boldsymbol{A} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}, \ \boldsymbol{B} = \{1, 3, 5, 7, 9\}, \ \boldsymbol{C} = \{2, 3, 5, 7\}.$$

Диаграмма Эйлера–Венна, показанная на рис. 36, наглядно изображает взаимоотношения между этими множествами. Каждое из

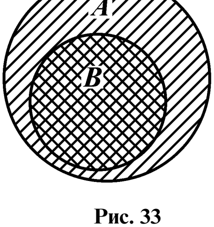


Рис. 33

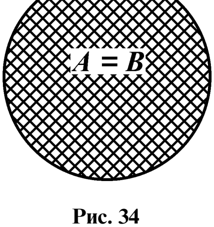


Рис. 34

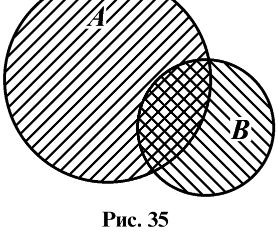


Рис. 35

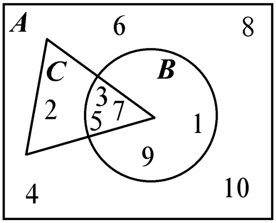


Рис. 36

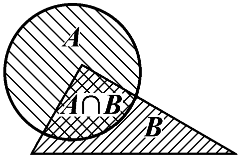


Рис. 37

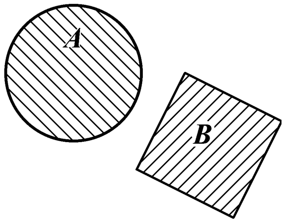


Рис. 38

множеств $\boldsymbol{B}$, $\boldsymbol{C}$ является подмножеством множества $\boldsymbol{A}$, т. е. $\boldsymbol{B} \subset \boldsymbol{A}$, $\boldsymbol{C} \subset \boldsymbol{A}$, причем $\boldsymbol{B}$ и $\boldsymbol{C}$ — собственные подмножества множества $\boldsymbol{A}$. Что же касается множеств $\boldsymbol{B}$ и $\boldsymbol{C}$, то ни одно из них не является подмножеством другого, и это также наглядно показано на рис. 36 — каждая из фигур, изображающих множества $\boldsymbol{B}$ и $\boldsymbol{C}$, выходит за пределы другой фигуры.

Перейдем теперь к рассмотрению операции пересечения множеств. Пусть $\boldsymbol{A}$ и $\boldsymbol{B}$ — два множества. Все те элементы, которые принадлежат как множеству $\boldsymbol{A}$, так и множеству $\boldsymbol{B}$, составляют, вместе взятые, некоторое множество, которое называется *пересечением* множеств $\boldsymbol{A}$ и $\boldsymbol{B}$; оно обозначается записью $\boldsymbol{A} \cap \boldsymbol{B}$ (рис. 37). Если множества $\boldsymbol{A}$ и $\boldsymbol{B}$ не имеют общих элементов, т. е. $\boldsymbol{A} \cap \boldsymbol{B} = \varnothing$, множества $\boldsymbol{A}$ и $\boldsymbol{B}$ называются *непересекающимися* (рис. 38). Например, для множеств $\boldsymbol{B}$ и $\boldsymbol{C}$, рассмотренных выше, справедливо соотношение $\boldsymbol{B} \cap \boldsymbol{C} \neq \varnothing$, т. е. пересечение этих множеств непусто:

$$\boldsymbol{B} \cap \boldsymbol{C} = \{3,\ 5,\ 7\}.$$

Проиллюстрируем понятие пересечения множеств на фигурах из школьного курса геометрии. Рассмотрим полосу, т. е. часть плоскости между двумя параллельными прямыми (рис. 39).

Любой параллелограмм может быть представлен как пересечение двух полос $\boldsymbol{A} = \boldsymbol{B} \cap \boldsymbol{C}$ (рис. 40). Треугольник можно рассматривать как пересечение трех полуплоскостей: $\boldsymbol{T} = \boldsymbol{A} \cap \boldsymbol{B} \cap \boldsymbol{C}$ (рис. 41). Выпуклый многоугольник с n вершинами является пересечением n полуплоскостей (случай $n = 5$ представлен на рис. 42). Сегмент круга (рис. 43) можно себе представлять как пересечение круга и полуплоскости.

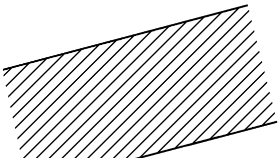


Рис. 39

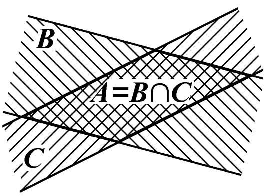


Рис. 40

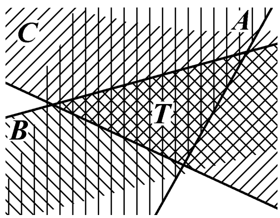


Рис. 41

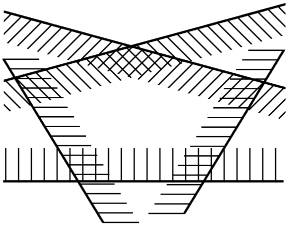


Рис. 42

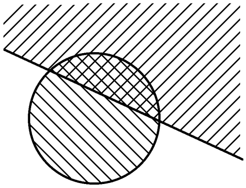


Рис. 43

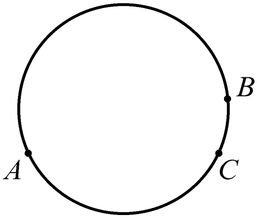


Рис. 44

Однако не только отдельные геометрические фигуры дают повод для использования понятия пересечения множеств. Существует важный метод решения геометрических задач (особенно задач на построение), который основан на применении пересечения множеств. Мы рассмотрим этот метод на примере двух задач.

В качестве первого примера возьмем следующую задачу. На плоскости даны три точки A, B, C, не лежащие на одной прямой; требуется построить окружность, проходящую через эти три точки (рис. 44). Чтобы решить эту задачу, мы разобьем требование задачи на две части: 1) искомая окружность должна проходить через точки A и B; 2) искомая окружность должна проходить через точки A и C.

Если нам удастся найти окружность, удовлетворяющую как условию 1, так и условию 2, то задача будет решена.

Если окружность удовлетворяет условию 1, то ее центр O находится от точек A и B на одинаковом расстоянии (равном радиусу, рис. 45). Следовательно, точка O находится на серединном перпендикуляре P_{AB} к отрезку AB. Верно и обратное: если O находится на этом серединном перпендикуляре, то окружность с центром O, проходящая через точку A, пройдет и через точку B. Иначе говоря, прямая P_{AB} есть множество всех центров окружностей, удовлетворяющих условию 1. Аналогично, серединный перпендикуляр P_{AC} к отрезку AC (рис. 46) есть множество всех центров окружностей, удовлетворяющих условию 2. Но искомая окружность должна удовлетворять обоим условиям 1 и 2. Значит, ее центр O должен принадлежать каждому из множеств P_{AB}, P_{AC}, т. е. $O \in P_{AB} \cap P_{AC}$. Так как пересе-

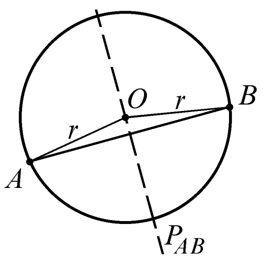


Рис. 45

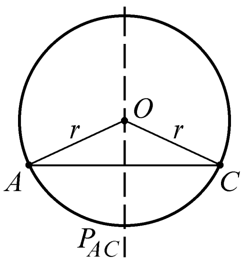


Рис. 46

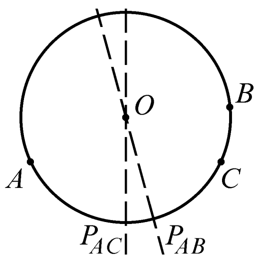


Рис. 47

чение $P_{AB} \cap P_{AC}$ состоит только из одной точки, то эта точка и является центром искомой окружности (рис. 47), т. е. поместив иголку циркуля в эту точку и проведя окружность, проходящую через точку A, мы и решим задачу — эта окружность пройдет также через точки B и C.

Решенную задачу можно интерпретировать следующим образом. В трех пунктах A, B, C удар грома услышали одновременно; указать место, где ударила молния. Известно, что звук распространяется с постоянной скоростью примерно 330 м/с. Поскольку в пунктах A, B и C удар грома был услышан одновременно, время от грозового разряда до момента его восприятия (на слух) было одинаковым для всех трех пунктов. Обозначим это время (в секундах) через t. За это время звук распространяется на расстояние $330t$. Таким образом, окружность с центром в месте грозового разряда и радиусом $330t$ проходит через все три точки A, B, C (рис. 48), и потому рассмотренная выше геометрическая задача дает возможность найти место, где ударила молния. Кстати, это решение позволяет также приближенно определить расстояние (от любого из пунктов A, B, C) до места грозового разряда, поскольку звук распространяется за 3 секунды примерно на 1 километр, а свет — практически мгновенно (около 300 000 км в секунду). Скажем, если время от зрительного восприятия молнии до слухового восприятия грома составляет 12 секунд, то грозовой разряд произошел на расстоянии 4 км.

Приведенный пример хорошо иллюстрирует приближенность всякого математического описания реального явления. Хотя в самом математическом решении (по формуле $s = vt$, относящейся к равномерному прямолинейному движению) вычисления осуществляются абсолютно точно, но *математическая модель* явления, так называемая его *идеализация*, дает лишь приближенное описание процесса.

Ведь скорость звука зависит от влажности воздуха, атмосферного давления, наличия препятствий, которые влияют на распространение звуковых волн, и т. п. Кроме того, хотя свет распространяется практически мгновенно (в данной задаче), но все же с некоторой скоростью, вспышка молнии происходит на некоторой высоте над землей. Имеется определенная (физиологическая) погрешность в фиксации тех моментов времени, когда обнаружены молния и гром. Пренебрегая всеми этими деталями (второстепенными для данного рассмотрения явления), мы и получаем некоторую модель явления, некоторую его идеализацию: пункты A, B, C и место разряда являются *геометрическими* точками, которые находятся в одной плоскости, моменты грозового удара фиксируются *точно*, звук распространяется прямолинейно и с фиксированной скоростью $\frac{1}{3}$ км в секунду. Вообще, для математического решения

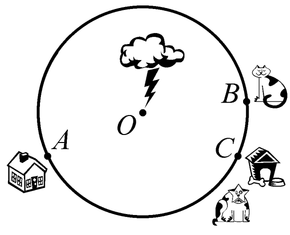


Рис. 48

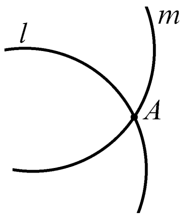


Рис. 49

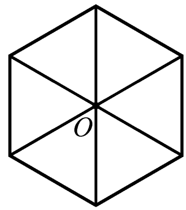


Рис. 50

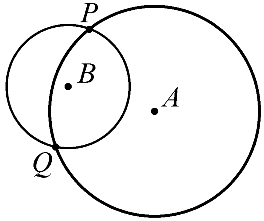


Рис. 51

любой жизненной задачи требуется, прежде всего, составить идеализированную математическую модель явления. И если результаты расчетов не соответствуют реальному ходу явления, то это лишь означает (если, конечно, нет ошибок в расчетах), что выбранная идеализация *неадекватна* действительному явлению и должна быть заменена более точной моделью. Следует также иметь в виду, что и сама математика не всегда дает точный численный ответ, а нередко использует приближенные методы расчета (не говоря уже о том, что в некоторых случаях нас интересует не точный численный расчет, а качественный прогноз).

Рассмотренный прием решения является типичным для многих геометрических задач, особенно для задач *на построение*. Геометрическая фигура задается обычно *конечным* числом характеристических точек. Например, отрезок задается двумя его концевыми точками; многоугольник (в частности, треугольник) задается его вершинами; окружность (или круг) задается двумя точками: центром и одной точкой на окружности (или тремя точками, лежащими на окружности) и т. п. Поэтому любая задача типа «построить геометрическую фигуру с такими-то свойствами» сводится, в конечном итоге, к построению нескольких точек. Но наши чертежные инструменты (карандаш, фломастер, рапидограф и т. д.) приспособлены к вычерчиванию *линий*. Конечно, они фактически воспроизводят узкие полоски, но, идеализируя, мы говорим о математических линиях, не имеющих ширины. Точка же может быть определена как *пересечение* двух линий на плоскости (рис. 49).

Разумеется, иногда три или большее число линий пересекаются в одной точке (например, три диагонали, соединяющие противоположные вершины правильного шестиугольника, пересекаясь, определяют его центр, рис. 50). Однако пересечения двух линий достаточно, чтобы задать точку. Именно поэтому мы обычно расчленяем требования, сформулированные в задаче на построение, на два отдельных условия (подобно условиям 1 и 2 рассмотренной выше задачи), каждое из которых определяет некоторое множество точек (чаще всего *линию*, или, как говорили в старину, «геометрическое место точек»), а *пересечение* этих двух множеств задает искомую точку (или несколько точек — в тех случаях, когда задача допускает несколько решений).

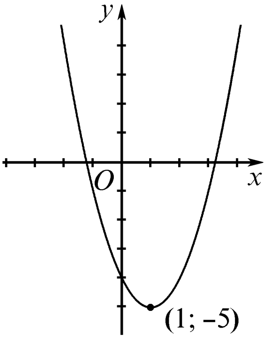


Рис. 52

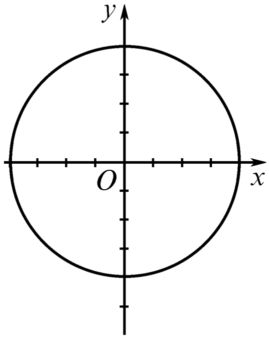


Рис. 53

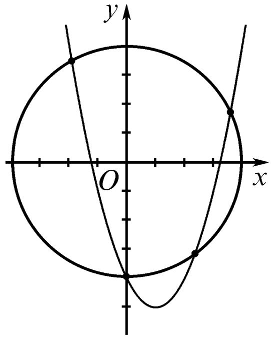


Рис. 54

Рассмотрим, например, следующую задачу. Звук грома достиг пункта A через 5 секунд после вспышки молнии, а пункта B — через 3 секунды; найти точку, где произошел грозовой разряд. Для решения достаточно заметить, что (в той же идеализации) за 5 с звук распространяется на $\frac{5}{3}$ км, а за 3 с — на 1 км. Поэтому множество всех точек, из которых звук разряда доходит до пункта A за 5 с, есть окружность с центром A и радиусом $\frac{5}{3}$ км, а множество всех точек, из которых он доходит до B за 3 с, есть окружность радиуса 1 км с центром B. Искомая точка принадлежит *пересечению* этих двух множеств (рис. 51), т. е. совпадает с одной из двух точек P, Q.

Геометрически обе точки равноправны и дают решение задачи. Практически же, если мы заметим, в каком направлении была видна вспышка молнии, то можем из этих двух точек отобрать одну; однако это означает привлечение *дополнительных* данных, не указанных в сформулированной вначале задаче.

Еще один яркий пример — *графическое решение уравнений*. Рассмотрим следующую систему уравнений:

$$\begin{cases} x^2 - 2x - y - 4 = 0, \\ x^2 + y^2 = 16. \end{cases}$$

Решить эту систему — значит найти пару чисел x, y (или все такие пары, если их несколько), которые удовлетворяют *обоим* этим уравнениям. Множество всех точек координатной плоскости, удовлетворяющих только первому уравнению, представляет собой *параболу*, вершина которой (т. е. нижняя точка) имеет координаты (1; –5) — это становится ясным, если записать первое уравнение в виде $y = (x - 1)^2 - 5$ (рис. 52). Второе уравнение задает *окружность* с центром в начале координат и радиусом 4 (рис. 53). Точки же $(x; y)$, удовлетворяющие обоим уравнениям, составляют *пересечение* этих двух линий (рис. 54). Эта графическая интерпретация показывает, что рассматриваемая система имеет *четыре* решения. Вот приближенное значение чисел, составляющих эти решения (их можно вычислить с помощью компьютера или даже микрокалькулятора):

$$1)\ x=-1{,}917\ldots,\ y=3{,}510\ldots;\ 2)\ x=0,\ y=-4;$$

$$3)\ x=2{,}319\ldots,\ y=-3{,}259\ldots;\ 4)\ x=3{,}597\ldots,\ y=1{,}748\ldots$$

В заключение приведем еще один практический пример применения понятия пересечения множеств. При отборе детей для участия в конкурсе решили использовать результаты обследования их по четырем параметрам u, v, w, x. Начало таблицы обследования приведено ниже (результаты указаны в процентах):

№	Фамилия, инициалы	u	v	w	x
1	Малинин И. И.	82%	47%	51%	43%
2	Куликова Е. А.	73%	61%	52%	42%
3	Комаров В. Г.	69%	39%	53%	63%
4	Закиров Н. И.	71%	63%	51%	41%
...	....	...	...	...	...

Было решено отобрать детей, имеющих по параметрам u, v, w, x результаты не ниже, чем соответственно 70%, 60%, 50%, 40%. Комиссия вначале стала последовательно исследовать результаты детей по списку. Однако это было признано неудобным, так как все время приходилось переключать внимание от критерия 70% к критерию 60% и т. д. Вместо этого решили применить другой метод. Комиссию разделили на 4 группы экспертов, каждой из которых выдали копию таблицы обследования. Первая группа экспертов вычеркивала фамилии детей, не прошедших по критерию u, вторая — по критерию v и т. д. Вот как выглядели результаты работы первой группы:

№	Фамилия, инициалы	u	v	w	x
1	Малинин И. И.	82%	47%	51%	43%
2	Куликова Е. А.	73%	61%	52%	42%
3	~~Комаров В. Г.~~	69%	39%	53%	63%
4	Закиров Н. И.	71%	63%	51%	41%
...	....	...	...	...	...

Аналогичную работу провели остальные группы. Обозначим теперь через $\boldsymbol{U}$ множество всех детей, пропущенных (т. е. невычеркнутых) первой группой, через $\boldsymbol{V}$ — второй, $\boldsymbol{W}$ — третьей, $\boldsymbol{X}$ — четвертой. Ясно, что множество детей, отобранных для участия в конкурсе, совпадает с *пересечением* $\boldsymbol{U} \cap \boldsymbol{V} \cap \boldsymbol{W} \cap \boldsymbol{X}$. Это множество было определено сразу: листки сложили вместе и на просвет стал виден список детей, не вычеркнутых ни одной из групп, т. е. вошедших в пересечение $\boldsymbol{U} \cap \boldsymbol{V} \cap \boldsymbol{W} \cap \boldsymbol{X}$ (хотя таких могло и не оказаться).

Задачи и упражнения

21. Опишите множество, являющееся пересечением множества четных чисел и множества чисел, делящихся на 5.

22. Опишите множество четырехугольников, являющихся одновременно прямоугольниками и ромбами.

23. В Монреале 80% жителей знают французский язык и 70% — английский. Сколько процентов жителей знают оба языка, если каждый житель знает хотя бы один из этих языков?

24. Дан куб. Проведите плоскость так, чтобы ее пересечение с кубом было правильным шестиугольником.

25. Сколько корней имеет уравнение: $10 \sin x = x$?

12. Объединение множеств

Все элементы множества $\boldsymbol{A}$ и все элементы множества $\boldsymbol{B}$, взятые вместе, составляют *объединение* множеств $\boldsymbol{A}$ и $\boldsymbol{B}$. Оно обозначается через $\boldsymbol{A} \cup \boldsymbol{B}$. Диаграмма Эйлера–Венна, приведенная на рис. 55, показывает взаимоотношение между множествами $\boldsymbol{A}$, $\boldsymbol{B}$, их пересечением и объединением. Двойной штриховкой обозначено пересечение множеств $\boldsymbol{A}$ и $\boldsymbol{B}$, а вся заштрихованная фигура (хотя бы одной штриховкой) — их объединение. Чтобы получить объединение $\boldsymbol{A} \cup \boldsymbol{B}$, можно взять все элементы множества $\boldsymbol{A}$ и добавить к ним те элементы множества $\boldsymbol{B}$, которые в $\boldsymbol{A}$ не входят (или наоборот: взять все элементы множества $\boldsymbol{B}$ и добавить к ним те элементы множества $\boldsymbol{A}$, которые в $\boldsymbol{B}$ не входят). Из сказанного ясно, что пересечение множеств $\boldsymbol{A}$ и $\boldsymbol{B}$ содержится в их объединении: $\boldsymbol{A} \cap \boldsymbol{B} \subset \boldsymbol{A} \cup \boldsymbol{B}$.

Для любого элемента множества $\boldsymbol{A} \cup \boldsymbol{B}$ имеются три возможности:

1) либо взятый элемент принадлежит $\boldsymbol{A}$, но не принадлежит $\boldsymbol{B}$ (элемент x на рис. 55);

2) либо же взятый элемент принадлежит $\boldsymbol{B}$, но не принадлежит $\boldsymbol{A}$ (элемент y на рис. 55);

3) либо же, наконец, взятый элемент принадлежит каждому из множеств $\boldsymbol{A}$ и $\boldsymbol{B}$, т. е. принадлежит их пересечению (элемент z на рис. 55).

Несомненно, представление об объединении двух непересекающихся множеств (скажем, соединение двух отдельных «кучек» предметов в одну общую «кучку») *предшествовало* возникновению понятия о сумме двух чисел. Если два конечные множества $\boldsymbol{A}$ и $\boldsymbol{B}$ не пересекаются (т. е. не имеют общих элементов), причем первое из них содержит a элементов, а второе содержит b элементов, то множество

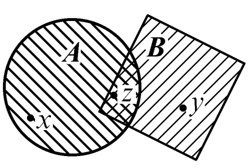


Рис. 55

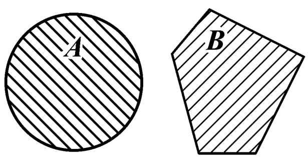


Рис. 56

$\boldsymbol{A} \cup \boldsymbol{B}$ содержит $a+b$ элементов (рис. 56). При этом «смешивание в одну кучку» не зависит от того, в каком порядке мы берем «кучки» $\boldsymbol{A}$ и $\boldsymbol{B}$, т. е.

$$\boldsymbol{A} \cup \boldsymbol{B} = \boldsymbol{B} \cup \boldsymbol{A}$$

и, в соответствии с этим, сумма чисел не зависит от порядка слагаемых:

$$a+b=b+a. \tag{1}$$

Разумеется, это не есть «доказательство» коммутативного свойства сложения, выражаемого равенством (1), а лишь пояснение. Подмеченное при смешивании «кучек», содержащих небольшое число реальных предметов, коммутативное свойство сложения («от перестановки слагаемых сумма не изменяется») было впоследствии сформулировано, как общее свойство сложения натуральных чисел. Формулировка этого свойства, выражаемая равенством (1), есть результат многовековой практики человечества и сложного процесса абстракции.

Таким образом, в теории натуральных чисел коммутативность сложения фактически применяется без доказательства, т. е. является *аксиомой*. Согласно довольно распространенному мнению, аксиомами называются истины, принимаемые в рассматриваемой теории без доказательства «ввиду их очевидности». Однако математики придерживаются несколько иного взгляда. Ведь то, что кажется «очевидным» одному человеку, может быть вовсе не «очевидным» другому. И даже для одного индивидуума ощущение «очевидности» может меняться в процессе его интеллектуального развития и приобретения знаний. Правильнее будет сказать, что аксиомы — это *первоначальные* (и именно поэтому не доказываемые) утверждения, лежащие в основе рассматриваемой теории и служащие для вывода (доказательства) дальнейших ее утверждений, называемых *теоремами*. Что же касается «очевидности» аксиом, их «простоты» и, может быть, «априорности», то эти вопросы в самой теории не рассматриваются; они связаны с историей развития науки (вскрывающей происхождение аксиом) и с пояснением наглядного смысла аксиом в процессе изучения или преподавания теории. Об аксиомах и о доказательствах теорем мы еще будем иметь случай поговорить в дальнейших беседах, а сейчас вернемся к операции объединения множеств.

Рассмотрим вопрос о *мощности* объединения множеств. Для удобства условимся обозначать мощность множества $\boldsymbol{A}$ символом $|\boldsymbol{A}|$. Как мы отмечали выше, для конечных множеств $\boldsymbol{A}$, $\boldsymbol{B}$, не имеющих общих элементов, справедливо равенство

$$|\boldsymbol{A} \cup \boldsymbol{B}| = |\boldsymbol{A}| + |\boldsymbol{B}|. \tag{2}$$

Посмотрим, как видоизменяется это соотношение в случае, когда пересечение $\boldsymbol{A} \cap \boldsymbol{B}$ непусто, а также в случае, когда хотя бы одно из множеств $\boldsymbol{A}$, $\boldsymbol{B}$ бесконечно. Нетрудно доказать, что для конечных множеств при $\boldsymbol{A} \cap \boldsymbol{B} \neq \varnothing$ соотношение (2) заменяется следующим:

$$|\boldsymbol{A} \cup \boldsymbol{B}| = |\boldsymbol{A}| + |\boldsymbol{B}| - |\boldsymbol{A} \cap \boldsymbol{B}|. \tag{3}$$

Иными словами, если множество $\boldsymbol{A}$ содержит a элементов, множество $\boldsymbol{B}$ содержит b элементов, а их пересечение $\boldsymbol{A} \cap \boldsymbol{B}$ содержит p элементов, то число элементов множества $\boldsymbol{A} \cup \boldsymbol{B}$ равно $a + b - p$. В самом деле, пересчитав все элементы множества $\boldsymbol{A}$, а затем все элементы множества $\boldsymbol{B}$, мы насчитаем $a + b$ элементов, но при этом каждый из p элементов, входящих в пересечение $\boldsymbol{A} \cap \boldsymbol{B}$, засчитаем *дважды*, тогда как в множестве $\boldsymbol{A} \cup \boldsymbol{B}$ каждый из них следует засчитывать лишь один раз. В качестве упражнения рекомендуем читателю написать формулу, аналогичную (3), для объединения $\boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{C}$ *трех* конечных множеств.

Обратимся теперь к случаю, когда хотя бы одно из множеств $\boldsymbol{A}$, $\boldsymbol{B}$ бесконечно. В этом случае формула (2) существенно изменяется. Именно, если мощность множества $\boldsymbol{A}$ *не меньше* мощности множества $\boldsymbol{B}$, т. е. $|\boldsymbol{A}| \geq |\boldsymbol{B}|$, причем множество $\boldsymbol{A}$ бесконечно, то независимо от того, имеют ли $\boldsymbol{A}$ и $\boldsymbol{B}$ общие элементы, справедливо равенство

$$|\boldsymbol{A} \cup \boldsymbol{B}| = |\boldsymbol{A}|. \tag{4}$$

В частности, по аналогии с равенством (2) можно написать следующие соотношения:

$$\aleph_0 + n = \aleph_0,$$

$$\aleph_0 + \aleph_0 = \aleph_0,$$

$$\mathfrak{C} + n = \mathfrak{C},$$

$$\mathfrak{C} + \aleph_0 = \mathfrak{C},$$

$$\mathfrak{C} + \mathfrak{C} = \mathfrak{C},$$

в которых через n обозначена любая конечная мощность (т. е. натуральное число). Например, второе из этих соотношений означает, что объединение двух счетных множеств также является счетным множеством. В самом деле, если

$$\boldsymbol{A} = \{a_1, a_2, a_3, \ldots\}, \ \boldsymbol{B} = \{b_1, b_2, b_3, \ldots\},$$

то, «перемешав» элементы этих множеств, мы запишем их объединение в виде

$$\boldsymbol{A} \cup \boldsymbol{B} = \{a_1, b_1, a_2, b_2, a_3, b_3, \ldots\}$$

(не обращая внимания на то, что в правой части могут встречаться повторяющиеся элементы), откуда видно, что элементы множества

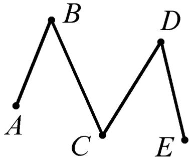


Рис. 57

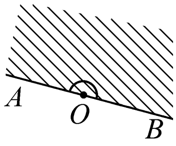


Рис. 58

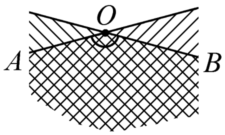


Рис. 59

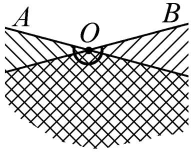


Рис. 60

$\boldsymbol{A} \cup \boldsymbol{B}$ можно перенумеровать натуральными числами, т. е. это множество счетно.

Рассмотрим теперь некоторые примеры применения понятия объединения множеств в школьной математике. Всякая *ломаная* есть объединение составляющих ее отрезков (называемых *звеньями* этой ломаной). Например, ломаная с последовательными вершинами A, B, C, D, E (рис. 57) представляет собой объединение четырех отрезков $AB \cup BC \cup CD \cup DE$.

Развернутый угол (содержащий 180°, рис. 58) представляет собой *полуплоскость*. Угол, *меньший развернутого*, можно представить в виде пересечения двух полуплоскостей (рис. 59). Угол, *больший развернутого*, представить в таком виде не удается; однако он является *объединением* двух полуплоскостей (рис. 60). Всякий четырехугольник можно разбить диагональю на два треугольника, причем это верно не только для выпуклых четырехугольников (рис. 61), но и для невыпуклых (рис. 62). Иными словами, всякий четырехугольник представляет собой объединение двух треугольников, которые не перекрывают друг друга, т. е. не имеют общих внутренних точек. Вообще, любой n-угольник можно разбить диагоналями на $n-2$ треугольника. Для этого в *выпуклом* многоугольнике достаточно провести все диагонали, выходящие из одной его вершины (рис. 63). Для невыпуклого многоугольника не всякая диагональ пригодна (рис. 64), однако и в этом случае можно подобрать такие диагонали, которые делят его на $n-2$ треугольника (рис. 65). Таким образом, всякий многоугольник представляется в виде *объединения* нескольких треугольников, не перекрывающих друг друга. Так как известны формулы, позволяющие вычислить площадь любого треугольника, то с помощью разбиения на треугольники можно найти площадь любого многоугольника.

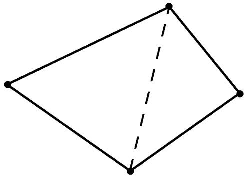


Рис. 61

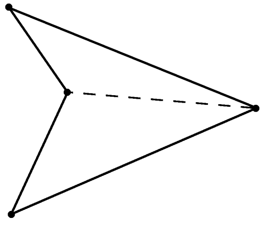


Рис. 62

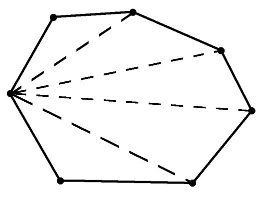


Рис. 63

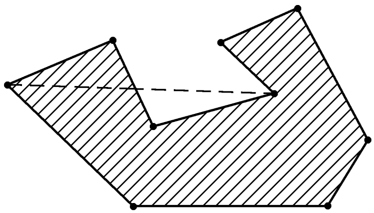


Рис. 64

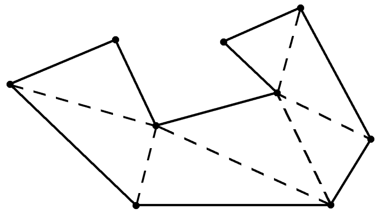


Рис. 65

Можно привести и ряд других геометрических примеров. Приведенные примеры показывают, что само понимание фигуры как *множества* точек и использование только терминов «объединение», «пересечение» создает удобства в осмыслении традиционного школьного материала и в выражении мыслей.

Приведем теперь некоторые примеры из области алгебры.

В левой части уравнения

$$(x^3 - 3x^2 + 4x - 12)(x^2 - 5x + 6) = 0 \tag{5}$$

стоит *произведение* двух многочленов

$$f(x) = x^3 - 3x^2 + 4x - 12,$$
$$g(x) = x^2 - 5x + 6.$$

Ясно, что каждый корень первого многочлена (т. е. такое число x_1, при подстановке которого этот многочлен обращается в нуль) является также и корнем уравнения (5), поскольку при умножении нуля на любое число снова получается нуль. Точно так же любой корень многочлена $g(x)$ является корнем уравнения (5). Если же число x не является корнем ни одного из многочленов $f(x)$, $g(x)$, то оно не является и корнем уравнения (5), поскольку произведение двух отличных от нуля чисел также отлично от нуля. Иначе говоря, множество всех корней уравнения (5) является *объединением* $\boldsymbol{F}_0 \cup \boldsymbol{G}_0$, где $\boldsymbol{F}_0$ — множество корней многочлена $f(x)$, а $\boldsymbol{G}_0$ — множество корней многочлена $g(x)$. Вообще, если левая часть некоторого уравнения $p(x) = 0$ представляет собой *произведение* многочленов или иных функций, то для нахождения всех корней этого уравнения следует взять *объединение* нескольких множеств, каждое из которых представляет собой множество всех корней одного из сомножителей (о методах решения уравнения и о так называемых «посторонних» корнях мы еще поговорим в п. 18).

В качестве второго примера рассмотрим неравенство

$$x^2 + x - 6 \geq 0. \tag{6}$$

График многочлена, записанного в левой части, показан на рис. 66. Из рассмотрения графика видно, что этот многочлен принимает *неотрицательные* значения при всех $x \le -3$, а также при $x \ge 2$ (рис. 67). Иначе говоря, множество $\boldsymbol{M}$ всех решений неравенства (6) представляет собой *объединение* двух лучей $(-\infty; -3]$ и $[2; \infty)$, т. е. это множество можно записать в виде

$$\boldsymbol{M} = (-\infty; -3] \cup [2; \infty). \tag{7}$$

Заметим попутно, что символы $-\infty$ и ∞ вовсе не означают какие-то «бесконечные» действительные числа, а лишь символически выражают изображенные на рис. 67 лучи. Например, второй из них условно записывается неравенствами $2 \le x < \infty$, которые показывают, что можно брать лишь числа, не меньше 2, но *как угодно большие*, т. е. этот луч простирается вправо «до бесконечности».

Наконец, еще один алгебраический пример. Рассмотрим решение неравенства, в левой части которого стоит тот же многочлен (5):

$$(x^3 - 3x^2 + 4x - 12)(x^2 - 5x + 6) > 0. \tag{8}$$

Число x в том и только в том случае удовлетворяет этому неравенству, если либо оба многочлена $f(x)$, $g(x)$ принимают в точке x положительные значения, либо оба принимают отрицательные значения. Обозначим через $\boldsymbol{F}_+$ множество тех действительных чисел, для которых многочлен $f(x)$ принимает положительное значение, $\boldsymbol{G}_+$ — где $g(x)$ положительно и аналогично через $\boldsymbol{F}_-$ и $\boldsymbol{G}_-$ — множества тех чисел, в которых многочлены отрицательны. Тогда $\boldsymbol{F}_+ \cap \boldsymbol{G}_+$ есть множество всех точек x, в которых оба многочлена $f(x)$, $g(x)$ имеют положительные значения, а $\boldsymbol{F}_- \cap \boldsymbol{G}_-$ — где оба принимают отрицательные значения. Следовательно, объединяя эти пересечения, мы и получаем множество всех решений неравенства (8): $(\boldsymbol{F}_+ \cap \boldsymbol{G}_+) \cup (\boldsymbol{F}_- \cap \boldsymbol{G}_-)$.

Эти алгебраические примеры (число которых можно было бы увеличить) показывают, что, как и в случае геометрии, речь идет об использовании терминов «множество», «пересечение», «объединение», и этим создаются удобства в осмыслении традиционного школьного материала и в выражении мыслей, т. е. обозначения, связанные с множествами, упрощают математический язык и *облегчают* изучение материа-

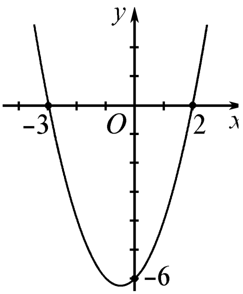


Рис. 66

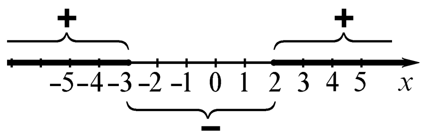


Рис. 67

ла. По существу, символы $\in$, $\notin$, $\subset$, $\bigcap$, $\bigcup$ являются лишь стенографическими знаками (и к тому же общепринятыми в математике) для сокращения и упрощения записи.

Более того, использование указанных терминов и знаков *роднит* геометрию и алгебру, показывает *единство* математики, *общность* ее методов и терминологии.

Задачи и упражнения

26. Докажите, что для конечных множеств справедливо равенство
$|\boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{C}| = |\boldsymbol{A}| + |\boldsymbol{B}| + |\boldsymbol{C}| - |\boldsymbol{A} \cap \boldsymbol{B}| - |\boldsymbol{A} \cap \boldsymbol{C}| - |\boldsymbol{B} \cap \boldsymbol{C}| + |\boldsymbol{A} \cap \boldsymbol{B} \cap \boldsymbol{C}|$.

27. На плоскости даны три точки A, B и C, не лежащие на одной прямой. Эти точки являются серединами трех сторон некоторого выпуклого четырехугольника. Укажите множество точек, в которых может лежать середина четвертой стороны. Укажите множество точек, в которых могут лежать вершины этого четырехугольника.

28. Даны три точки A, B и C. Через точку C проводятся всевозможные прямые. Опишите множество проекций отрезка AB на эти прямые.

29. Среди математиков каждый седьмой — философ, а среди философов каждый десятый — математик. Кого больше — философов или математиков?

30. Докажите, что если $\boldsymbol{A} \cap \boldsymbol{B} = \boldsymbol{C}$, то $(\boldsymbol{A} \cap \boldsymbol{B}) \cup \boldsymbol{C} = (\boldsymbol{A} \cup \boldsymbol{B}) \cap \boldsymbol{C}$.

13. Дополнение множеств

Прежде чем переходить к рассмотрению следующей операции, укажем один «парадокс» теории множеств; его обнаружил известный английский математик Бертран Рассел.

Как правило, множества не содержат себя в качестве элемента. Например, множество $\boldsymbol{A}$ всех целых чисел содержит в качестве элементов только целые числа; так как само $\boldsymbol{A}$ не есть целое число, а есть множество целых чисел, то $\boldsymbol{A}$ себя в качестве элемента не содержит. Условимся называть такие множества «ординарными». Но могут существовать и такие множества, которые содержат себя в качестве элемента. Рассмотрим, например, множество $\boldsymbol{S}$, определенное следующим образом: «$\boldsymbol{S}$ содержит в качестве элементов все множества, которые можно определить посредством предложения, содержащего меньше двадцати слов». Так как само множество $\boldsymbol{S}$ определяется предложением, содержащим менее двадцати слов, то выходит, что оно является элементом множества $\boldsymbol{S}$. Такие множества назовем «экстраординарными». Как бы то ни было, большинство множеств — ординарные; попробуем не иметь дела с дурно ведущими себя экстраординарными множествами и будем рассматривать только *множество всех ординарных множеств*. Обозначим его буквой $\boldsymbol{C}$. Каждый элемент множества $\boldsymbol{C}$ представляет собой множество, притом ординарное множество. Но вот возникает вопрос: а *само множество* $\boldsymbol{C}$ — ординарное или экстраординарное? Несомненно, оно должно быть или тем, или другим. Если $\boldsymbol{C}$ — ординарное множество, то оно содержит себя в качестве элемента, так как $\boldsymbol{C}$ определено как множество всех ординарных множеств. Раз дело обстоит так, значит $\boldsymbol{C}$ — экстраординарное множество, так как экстраординарными согласно определению названы множества, содержащие себя в качестве элемента. Получается противоречие. Значит, C должно быть экстраординарным множеством. Но тогда множество C содержит себя в качестве элемента; так как C есть экстраординарное множество, это противоречит определению C как множества всех ординарных множеств. Итак, мы видим, что уже одно только допущение существования множества C внутренне противоречиво.

Описанный выше парадокс Рассела — не единственный, обнаруженный в теории множеств. Эти парадоксы несколько обескуражили математиков, находившихся после успехов, достигнутых Г. Кантором, в состоянии некоторой эйфории. Как же избежать этих противоречий? Очевидно, нужно ограничиваться рассмотрением лишь таких множеств, которые определены четко и без противоречий (вспомните парикмахера в п. 7). Выход был найден в том, что в каждом случае, при каждом рассуждении рассматривается некоторое фиксированное корректно определенное множество U, называемое *универсальным множеством*, и изучению подлежат *только* элементы (и подмножества) этого универсального множества. Если рассуждение не выходит за эти пределы, оно не приводит к противоречию. В беседе 13 мы еще вернемся к вопросу о противоречивости и непротиворечивости, а пока ограничимся сказанным.

Возьмем, например, следующее рассуждение Кантора, которое произвело на математиков весьма сильное впечатление. Алгебраических чисел, т. е. чисел, являющихся корнями алгебраических уравнений с рациональными коэффициентами, имеется лишь *счетное* множество (см. задачу 15). В то же время множество *всех* действительных чисел несчетно. Значит, кроме алгебраических чисел существуют действительные числа, не являющиеся алгебраическими; их называют *трансцендентными*. Более того, трансцендентных чисел существенно *больше*, чем алгебраических: мощность множества всех трансцендентных чисел равна $\mathfrak{C}$. До того времени математики нашли лишь очень немного трансцендентных чисел, да и это было связано с существенными трудностями. Кантор же подарил математическому миру целый континуум трансцендентных чисел! Но корректно ли это канторовское рассуждение, нет ли в нем противоречия? Безусловно корректно, поскольку можно взять в качестве U множество всех действительных чисел, а все рассуждения Кантора о несчетности (и применение к трансцендентным числам) проводятся только *внутри* этого универсального множества, не выходя за его пределы.

Теперь можно перейти к рассмотрению следующей операции над множествами. Будем предполагать, что фиксировано некоторое универсальное множество U. В дальнейшей части этого пункта слово «множество» будет означать некоторое *подмножество* этого универсального множества U. Для любого множества A условимся через cA обозначать *дополнение* множества A, т. е. множество всех элементов $x \in U$, *не принадлежащих* множеству A (обозначение cA происходит от английского слова *complement* — дополнение).

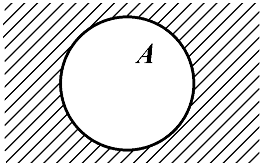


Рис. 68

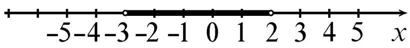


Рис. 69

В планиметрии (т. е. геометрии на плоскости) целесообразно взять всю плоскость $\boldsymbol{U}$ в качестве универсального множества, поскольку любая фигура, изучаемая в планиметрии, расположена в этой плоскости, т. е. является *подмножеством* универсального множества $\boldsymbol{U}$. Например, если $\boldsymbol{A}$ — круг, то $c\boldsymbol{A}$ есть *внешность* этого круга (рис. 68), т. е. то, что остается, если «вырезать» круг $\boldsymbol{A}$ из плоскости $\boldsymbol{U}$. При этом, если круг $\boldsymbol{A}$ рассматривается *замкнутым*, т. е. содержащим все точки ограничивающей его окружности, то его дополнение $c\boldsymbol{A}$ будет *открытым* множеством, т. е. ни одна точка этой окружности множеству $c\boldsymbol{A}$ не принадлежит. Далее, при решении уравнений (и при рассмотрении ряда других алгебраических вопросов) в качестве универсального множества целесообразно принять множество $\boldsymbol{R}$ всех действительных чисел. (Впрочем, в некоторых случаях, которые мы в этой беседе не рассматриваем, в качестве универсального множества удобно принять множество всех *комплексных* чисел.) Например, множество $\boldsymbol{M}$, дающее решение неравенства (6) (см. (7)), имеет своим дополнением $c\boldsymbol{M}$ множество всех решений *противоположного* неравенства $x^2 + x - 6 < 0$. Это множество (рис. 69) представляет собой интервал (–3; 2); концевые точки $x = -3$ и $x = 2$ ему не принадлежат (поскольку они принадлежат множеству $\boldsymbol{M}$). Ограничимся этими примерами, достаточно ясно показывающими, что такое дополнительное множество.

Теперь мы рассмотрим еще одну операцию, выражающуюся через рассмотренные ранее. Пусть $\boldsymbol{A}$ и $\boldsymbol{B}$ — два множества. Их *разность* $\boldsymbol{A} \setminus \boldsymbol{B}$ состоит из всех тех элементов множества $\boldsymbol{A}$, которые *не принадлежат* $\boldsymbol{B}$ (заштрихованное множество на рис. 70). Таким образом, разность $\boldsymbol{A} \setminus \boldsymbol{B}$ содержит все элементы x, для которых $x \in \boldsymbol{A}$, $x \notin \boldsymbol{B}$, или, что то же самое, все элементы x, для которых $x \in \boldsymbol{A}$ и, в то же время, $x \in c\boldsymbol{B}$. Иначе говоря, разность $\boldsymbol{A} \setminus \boldsymbol{B}$ есть пересечение множеств $\boldsymbol{A}$ и $c\boldsymbol{B}$: $\boldsymbol{A} \setminus \boldsymbol{B} = \boldsymbol{A} \cap c\boldsymbol{B}$.

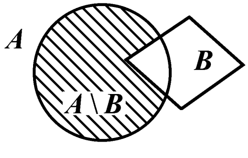


Рис. 70

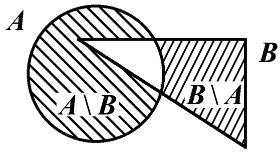


Рис. 71

Например, для любого множества $\boldsymbol{A}$ справедливо соотношение

$$\boldsymbol{U} \setminus \boldsymbol{A} = c\boldsymbol{A}.$$

А диаграмма Эйлера–Венна, приведенная на рис. 71, показывает, что справедливо соотношение

$$(\boldsymbol{A} \setminus \boldsymbol{B}) \bigcup (\boldsymbol{B} \setminus \boldsymbol{A}) = (\boldsymbol{A} \bigcup \boldsymbol{B}) \setminus (\boldsymbol{A} \bigcap \boldsymbol{B}).$$

Для любых трех множеств $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ справедливы соотношения *дистрибутивности*:

$$\boldsymbol{A} \bigcap (\boldsymbol{B} \bigcup \boldsymbol{C}) = (\boldsymbol{A} \bigcap \boldsymbol{B}) \bigcup (\boldsymbol{A} \bigcap \boldsymbol{C}), \quad \boldsymbol{A} \bigcup (\boldsymbol{B} \bigcap \boldsymbol{C}) = (\boldsymbol{A} \bigcup \boldsymbol{B}) \bigcap (\boldsymbol{A} \bigcup \boldsymbol{C}) \quad (9)$$

и соотношения *двойственности*:

$$c(\boldsymbol{A} \bigcap \boldsymbol{B}) = (c\boldsymbol{A}) \bigcup (c\boldsymbol{B}), \quad c(\boldsymbol{A} \bigcup \boldsymbol{B}) = (c\boldsymbol{A}) \bigcap (c\boldsymbol{B}). \quad (10)$$

Диаграммы, приведенные на рис. 72 – 74, иллюстрируют первое из соотношений дистрибутивности, а рис. 75 – 77 — первое соотношение двойственности. Не приводя аналогичные диаграммы для вторых соотношений (9) и (10) (читатель может изобразить их самостоятельно), мы дадим лишь жизненные примеры, иллюстрирующие их.

Во вторник в одном классе проходило занятие физического кружка. В нем принимали участие все мальчики этого класса, а также все интересующиеся физикой девочки. В среду в этом же классе проходило занятие химического кружка, где присутствовали все мальчики и интересующиеся химией девочки. Таким образом, обозначая через $\boldsymbol{A}$ множество всех мальчиков этого класса, через $\boldsymbol{B}$ — множество девочек, интересующихся физикой, а через $\boldsymbol{C}$ — множество девочек, интересующихся химией, мы заключаем, что во вторник присутствовало множество учащихся $\boldsymbol{A} \bigcup \boldsymbol{B}$, а в среду $\boldsymbol{A} \bigcup \boldsymbol{C}$.

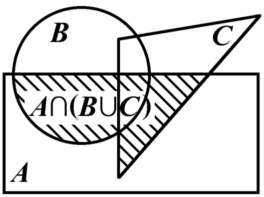


Рис. 72

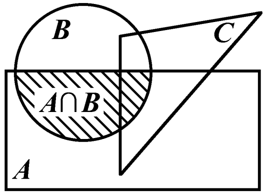


Рис. 73

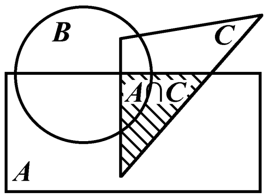


Рис. 74

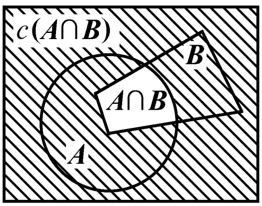


Рис. 75

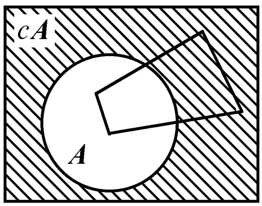


Рис. 76

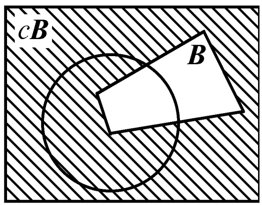


Рис. 77

Кто же из учеников присутствовал на заседаниях *обоих* кружков? Ясно, что это были все мальчики, а также все девочки, одновременно интересующиеся и физикой, и химией: $\boldsymbol{A} \bigcup (\boldsymbol{B} \bigcap \boldsymbol{C})$. С другой стороны, так как во вторник присутствовало множество учащихся $\boldsymbol{A} \bigcup \boldsymbol{B}$, а в среду $\boldsymbol{A} \bigcup \boldsymbol{C}$, то на заседания *обоих* кружков присутствовало множество учащихся $(\boldsymbol{A} \bigcup \boldsymbol{B}) \bigcap (\boldsymbol{A} \bigcup \boldsymbol{C})$. Таким образом, одно и то же множество учащихся может быть записано двумя способами — в полном соответствии со вторым соотношением дистрибутивности.

Второй пример также возьмем из области кружковой работы. В одном классе работали фотокружок и кружок рисования. Обозначим через $\boldsymbol{A}$ множество участников первого кружка, а через $\boldsymbol{B}$ — второго. Были учащиеся, которые принимали участие в работе обоих кружков, а были и такие, которые не посещали ни один кружок. Охарактеризуем множество учащихся, не работавших ни в одном из этих кружков. Для этого в качестве универсального множества $\boldsymbol{U}$ возьмем множество всех учеников класса. Тогда $\boldsymbol{A} \bigcup \boldsymbol{B}$ есть множество учеников, работавших *хотя бы в одном кружке*, а его дополнение $c(\boldsymbol{A} \bigcup \boldsymbol{B})$ как раз и есть множество учащихся, не посещавших ни один из кружков. С другой стороны, $c\boldsymbol{A}$ есть множество учащихся, не посещавших первый кружок, а $c\boldsymbol{B}$ — не посещавших второй. Значит, их пересечение $(c\boldsymbol{A}) \bigcap (c\boldsymbol{B})$ представляет собой множество учащихся, не посещавших ни один из кружков. Это дает два способа записи одного и того же множества — в соответствии со вторым соотношением (10).

Заметим в заключение, что соотношения дистрибутивности и двойственности справедливы и для большего числа множеств. Например, первое соотношение (9) обобщается следующим образом:

$$\boldsymbol{A} \bigcap (\boldsymbol{B}_1 \bigcup \boldsymbol{B}_2 \bigcup \ldots \bigcup \boldsymbol{B}_k) = (\boldsymbol{A} \bigcap \boldsymbol{B}_1) \bigcup (\boldsymbol{A} \bigcap \boldsymbol{B}_2) \bigcup \ldots \bigcup (\boldsymbol{A} \bigcap \boldsymbol{B}_k).$$

Читателю предлагается самостоятельно найти «кружковую» иллюстрацию этого соотношения и других обобщений равенств (9) и (10).

Задачи и упражнения

31. В классе 40% мальчиков. Математический кружок посещают 40% учеников, при этом 40% участников математического кружка составляют девочки. Какая часть мальчиков посещает математический кружок?

32. Докажите, что $(c\boldsymbol{A}) \bigcap (c\boldsymbol{B}) \bigcap (c\boldsymbol{C}) = c(\boldsymbol{A} \bigcup \boldsymbol{B} \bigcup \boldsymbol{C})$.

33. Учитель задал на уроке замысловатую задачу. В результате количество мальчиков, решивших эту задачу, оказалось равным количеству девочек, ее не решивших. Кого в классе больше — решивших задачу или девочек?

34. На шахматном турнире каждый участник выиграл белыми фигурами столько же партий, сколько все остальные — черными. Докажите, что все участники выиграли одинаковое количество партий.

35. Из 100 студентов колледжа 28 изучают английский язык, 42 — французский, 30 — немецкий, 8 — английский и немецкий, 10 — английский и французский, 5 — немецкий и французский, 3 — все три языка. Сколько студентов не изучает ни одного языка? Сколько студентов изучает только французский язык?

14. Произведение множеств

На билете для посещения кинотеатра написано: «ряд 4, место 5». Пара чисел 4 и 5 определяет положение места в зрительном зале. Можно было просто написать (4; 5), имея в виду, что первое число — номер ряда, а второе — номер места в этом ряду (рис. 78). Числа 4 и 5 — *координаты* места в зрительном зале. И если $\boldsymbol{A}$ — множество номеров всех рядов в зрительном зале, скажем, $\boldsymbol{A} = \{1, 2, ... 20\}$, а $\boldsymbol{B}$ — множество чисел, которыми нумеруются места в каждом ряду (скажем, $\boldsymbol{B} = \{1, 2, ..., 25\}$ — мы считаем, что все ряды одинаковые), то каждое место зрительного зала определяется парой чисел $(x; y)$, где $x \in \boldsymbol{A}$, $y \in \boldsymbol{B}$. Иными словами, множество мест в зрительном зале находится во взаимно однозначном соответствии с множеством всех пар $(x; y)$, где $x \in \boldsymbol{A}$, $y \in \boldsymbol{B}$.

Аналогично этому, если в плоскости выбрана координатная система, то положение каждой точки на плоскости определяется парой чисел $(x; y)$, где x — абсцисса рассматриваемой точки, а y — ее ордината (рис. 79). Иными словами, множество всех точек плоскости находится во взаимно однозначном соответствии с множеством всех пар $(x; y)$, где $x \in \boldsymbol{R}$, $y \in \boldsymbol{R}$ (т. е. x и y — действительные числа). Существенно, что (как и в первом примере) пары $(x; y)$ — *упорядоченные*, т. е. четко различается, какое число в паре является первым, а какое — вторым; пары $(x; y)$ и $(y; x)$ означают при $x \neq y$ разные точки плоскости (рис. 80). Принято говорить «точка $(x; y)$», т. е. отождествлять точку плоскости и пару чисел $(x; y)$ — ее координат.

На собрание кооператива пришло некоторое количество его членов; $\boldsymbol{A}$ — множество пришедших мужчин, $\boldsymbol{B}$ — множество женщин. Решено было избрать председателя и секретаря собрания, причем председателем должен быть мужчина, а секретарем — женщина. Сколькими способами можно их избрать? Здесь речь идет о выборе пары $(x; y)$, где $x \in \boldsymbol{A}$, $y \in \boldsymbol{B}$. Число способов выбора председателя равно $|\boldsymbol{A}|$, т. е. равно численности (мощности) множества $\boldsymbol{A}$, поскольку любой из пришедших мужчин может быть выбран председателем. И, независимо от этого (т. е. на каждый способ выбора председателя), имеется $|\boldsymbol{B}|$ способов выбрать секретаря. Таким образом, при указанных условиях всего имеется $|\boldsymbol{A}|\cdot|\boldsymbol{B}|$ способов выбрать руководство собрания, т. е. пару (председатель; секретарь).

В этих случаях (и ряде других) рассматриваются два множества $\boldsymbol{A}$, $\boldsymbol{B}$ и берется *множество всех пар* $(x; y)$, где $x \in \boldsymbol{A}$, $y \in \boldsymbol{B}$. Это множество пар называется *произведением* взятых множеств $\boldsymbol{A}$, $\boldsymbol{B}$ и обозначается через $\boldsymbol{A}\times\boldsymbol{B}$.

В кинотеатр пришло несколько учеников класса. Они купили оставшиеся разрозненные билеты и сели в разных местах зала. Отождествляя каждое место с соответствующей парой чисел $(x; y)$, мы можем сказать, что множество мест, на которых сидели пришедшие в кино ученики, есть некоторое *подмножество* произведения $\boldsymbol{A}\times\boldsymbol{B}$, где

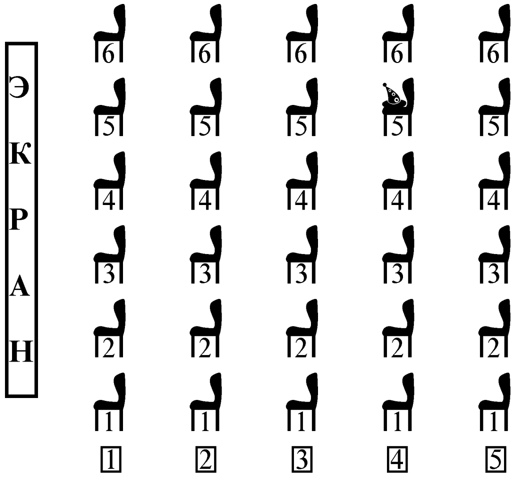


Рис. 78

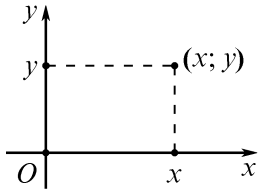


Рис. 79

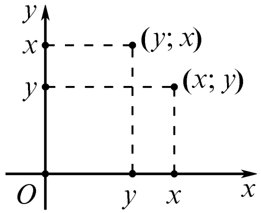


Рис. 80

A — множество номеров рядов, а ***B*** — множество номеров мест (мы здесь снова рассматриваем лишь случай, когда в каждом ряду содержится *одно и то же* число мест). Если, например, это подмножество имеет вид $\{(5;\ 13),\ (6;\ 11),\ (12;\ 9)\}$, то это означает, что в кино пришли три ученика, первый сидел в 5-м ряду на 13-м месте, второй — в 6-м ряду на 11-м месте и третий — в 12-м ряду на 9-м месте (рис. 81).

Так как координатная плоскость есть произведение $\boldsymbol{R}\times\boldsymbol{R}$, то *график* функции, например, многочлена $y = x^3 + 2x^2 - 5x - 6$, представляет собой некоторое подмножество этого произведения (рис. 82). Однако вовсе не любое подмножество ***G*** координатной плоскости $\boldsymbol{R}\times\boldsymbol{R}$ является графиком некоторой функции; необходимо (и достаточно), чтобы каждая прямая, параллельная оси ординат, пересекала ***G*** не более чем в одной точке (рис. 83). Подробнее мы поговорим об этом в беседе 4.

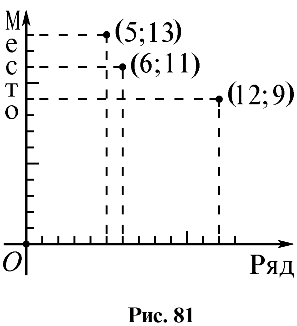


Рис. 81

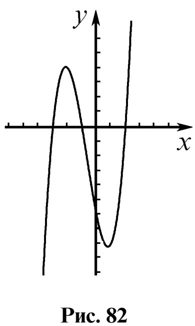


Рис. 82

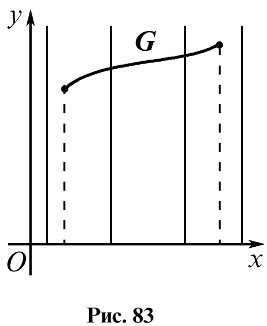


Рис. 83

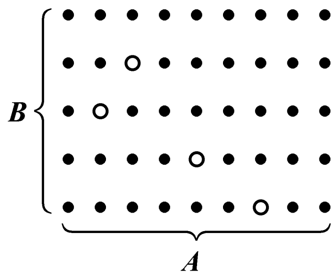


Рис. 84

	Иванов	Петров	Сидоров	Костин	Маркин	Ломов
Иванов			•			
Петров				•		
Сидоров	•					
Костин		•				
Маркин						•
Ломов					•	

Рис. 85

Вернемся к примеру о собрании членов кооператива. Выяснилось, что некоторые из присутствующих состоят в родственных отношениях — являются мужем и женой; в связи с этим было решено, что председатель и секретарь не должны быть из одной семьи. В этих условиях пару (председатель; секретарь) следует выбирать не из всего произведения $A\times B$, а из некоторого его подмножества, поскольку некоторые пары являются запрещенными (рис. 84, на котором кольца изображают семейные пары).

Допустим, что планируется шахматный турнир, A — множество всех его участников (будем считать, что в A *четное* число элементов, так что в каждом туре заняты все участники). При составлении расписания на один из туров следует в произведении $A\times A$ исключить диагональные клетки, поскольку никакой участник не может играть «с самим собой», и отметить в таблице $|A|$ клеток, попарно симметричных относительно диагонали (если третий играет с первым, то это означает также, что первый играет с третьим), причем так, что каждая строка (и каждый столбец) таблицы имеет ровно *одну* отмеченную клетку (рис. 85). Любое такое подмножество произведения $A\times A$ может служить графиком расписания для одного из дней турнира.

Наконец, рассмотрим еще вопрос о связи произведения множеств с произведением чисел. На рис. 86 изображена ситуация, в которой рассматривается набор из пяти предметов, взятый (повторенный) три раза. Объединяя все эти предметы, мы получаем случай, когда складываются три одинаковых слагаемых: 5 + 5 + 5. Или, как еще гово-

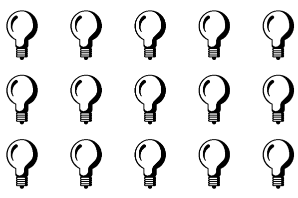


Рис. 86

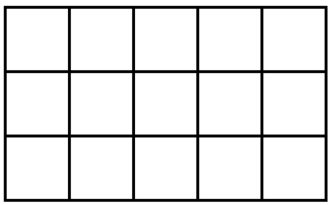


Рис. 87

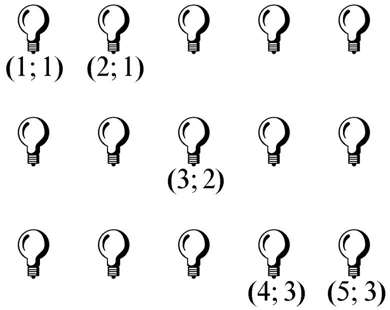


Рис. 88

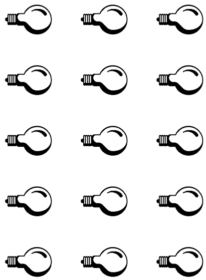


Рис. 89

рят, «надо пять взять три раза». Такое многократное сложение одинаковых слагаемых называется *умножением* (натуральных чисел). В данном случае речь идет о действии $5 \cdot 3$.

Таким же способом вычисляется *площадь прямоугольника* (с целыми длинами сторон, рис. 87), в связи с чем и возникает формула $S = ab$, выражающая площадь прямоугольника через длины его сторон (о единицах длины и площади, используемых при таком вычислении, нам еще предстоит поговорить впоследствии).

Связь с произведением множеств здесь непосредственно прослеживается. В самом деле, перенумеруем предметы в наборе (в случае, изображенном на рис. 86, нумерация осуществляется числами из множества $\boldsymbol{A} = \{1, 2, 3, 4, 5\}$) и сами наборы (нумерация осуществляется числами из множества $\boldsymbol{B} = \{1, 2, 3\}$). Тогда каждый предмет на рис. 86 обозначается парой чисел $(x; y)$, где $x \in \boldsymbol{A}$, $y \in \boldsymbol{B}$. Например, первый предмет из первого набора — пара (1; 1); третий предмет из второго набора — пара (3; 2) и т. д. (рис. 88). А всего предметов взято столько, сколько есть элементов $(x; y)$ в произведении $\boldsymbol{A} \times \boldsymbol{B}$.

В общем случае (для конечных множеств) произведение $\boldsymbol{A} \times \boldsymbol{B}$ имеет $|\boldsymbol{A}| \cdot |\boldsymbol{B}|$ элементов, т. е. произведению множеств соответствует произведение чисел.

Эта связь между произведением множеств и произведением чисел дает хорошую наглядную иллюстрацию свойства *коммутативности* умножения натуральных чисел: достаточно повернуть рисунок на 90°, и мы получаем модель произведения $3 \cdot 5$, хотя число предметов, очевидно, не изменилось (рис. 89). Таким образом, $5 \cdot 3 = 3 \cdot 5$, и вообще, $mn = nm$ для любых натуральных чисел m, n. Как и в случае сложения, это не есть «доказательство» свойства коммутативности, а лишь его наглядное пояснение.

В теории натуральных чисел коммутативность умножения является *аксиомой*, а соображения, подобные тем, которые связаны с рассмотренными рисунками, поясняют *происхождение* этой аксиомы.

Задачи и упражнения

36. Опишите геометрический образ произведения отрезка на окружность.

37. Каков будет геометрический образ произведения двух окружностей?

38. Огромный военный оркестр демонстрировал свое искусство на площади. Сначала музыканты построились в виде квадрата, а затем перестроились в прямоугольник, причем число шеренг увеличилось на 5. Сколько музыкантов в оркестре?

39. Сколько ладей можно поставить на шахматной доске так, чтобы они не били друг друга? Сколько способов такой расстановки?

40. Опишите множество точек координатной плоскости, координаты которых удовлетворяют соотношению $|x| + |y| = 1$.

Беседа 4. Отображения

15. Общее понятие отображения и школьная математика

Мы говорим, что задано *отображение* f множества $\boldsymbol{A}$ в множество $\boldsymbol{B}$, если указано правило, сопоставляющее *каждому* элементу $x \in \boldsymbol{A}$ *один* вполне определенный элемент множества $\boldsymbol{B}$, называемый *образом* элемента x и обозначаемый через $f(x)$. Задание отображения f множества $\boldsymbol{A}$ в множество $\boldsymbol{B}$ обозначается записью $f\colon \boldsymbol{A} \to \boldsymbol{B}$ или $\boldsymbol{A} \xrightarrow{f} \boldsymbol{B}$ (а иногда просто $\boldsymbol{A} \to \boldsymbol{B}$, если из текста понятно, о каком отображении идет речь). Для отображения $f\colon \boldsymbol{A} \to \boldsymbol{B}$ множество $\boldsymbol{A}$ называется его *областью определения*. На рис. 90 схематически показано некоторое отображение множества $\boldsymbol{A}$ в множество $\boldsymbol{B}$: от каждого элемента из $\boldsymbol{A}$ проведена стрелка к его образу.

Заметим, что один и тот же элемент множества $\boldsymbol{B}$ может оказаться образом нескольких элементов из $\boldsymbol{A}$. Так, на рис. 90 элемент $p \in \boldsymbol{B}$ является образом каждого из элементов b, с, т. е. $f(b) = p$ и $f(c) = p$. Может оказаться, далее, что некоторый элемент множества $\boldsymbol{B}$ не является образом никакого элемента из $\boldsymbol{A}$. Например, на рис. 90 нет в $\boldsymbol{A}$ элемента, образ которого совпадал бы с q. Но для любого отображения $\boldsymbol{A} \to \boldsymbol{B}$ непременно должно быть выполнено условие: каждому $x \in \boldsymbol{A}$ сопоставлен только один элемент множества $\boldsymbol{B}$ в качестве его образа. Соответствие, показанное на рис. 91, отображением не является, поскольку элементу $c \in \boldsymbol{A}$ не поставлен в соответствие никакой элемент множества $\boldsymbol{B}$ (т. е. от c не идет стрелка к какому-либо элементу множества $\boldsymbol{B}$). Соответствие на рис. 92 также

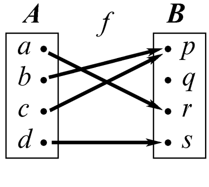


Рис. 90

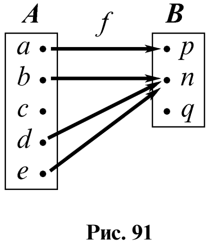


Рис. 91

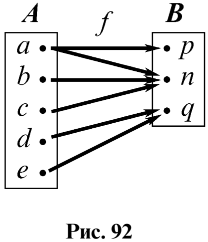


Рис. 92

Рис. 93

Рис. 94

отображением не является, так как от элемента $a \in \boldsymbol{A}$ ведут две стрелки к элементам множества $\boldsymbol{B}$, а не одна.

Отображения, рассматриваемые в алгебре и математическом анализе, называют *функциями*. Чаще всего функции задаются формулами. Например, запись

$$f(x) = x^2 - 3x + 5 \tag{1}$$

означает, что каждому действительному числу x сопоставляется в качестве его образа (или, как говорят, значения этой функции в точке x) число, определенное правой частью равенства (1). Например, $f(2) = 3$, $f(0) = 5$, $f(-2) = 15$. Поскольку в качестве x может быть взято любое действительное число, областью определения этой функции является все множество $\boldsymbol{R}$ действительных чисел. И образ $f(x)$ (т. е. значение этой функции в точке x) также является действительным числом, т. е. формула (1) задает функцию (отображение) $f\colon \boldsymbol{R} \to \boldsymbol{R}$. Функция $g(x) = \sqrt{x-5}$ имеет своей областью определения луч $[5; \infty)$, т. е. множество всех действительных чисел, удовлетворяющих неравенству $x \geq 5$ (поскольку при $x < 5$ выражение под корнем становится отрицательным), т. е. эта запись определяет функцию (отображение)

$$g\colon [5; \infty) \to \boldsymbol{R}.$$

А запись

$$h(x) = \frac{1}{(x-5)\sqrt{x}} \tag{2}$$

определяет функцию (отображение) $h\colon \boldsymbol{A} \to \boldsymbol{R}$, где множество $\boldsymbol{A}$, являющееся *областью определения* этой функции, имеет вид $A = (0; 5) \bigcup (5; \infty)$, т. е. представляет собой объединение интервала $(0; 5)$, определяемого неравенствами $0 < x < 5$, и луча $(5; \infty)$, состоящего из всех точек, удовлетворяющих неравенству $x > 5$. Отметим, что от *отрезка* $[a; b]$, состоящего из всех чисел, удовлетворяющих неравенствам $a \leq x \leq b$ (рис. 93), *интервал* $(a; b)$ отличается тем, что он не содержит концевые точки $x = a$ и $x = b$, т. е. определяется неравенствами $a < x < b$ (рис. 94).

Вообще, когда говорят, что рассматривается некоторая *числовая функция* $f(x)$, то имеют в виду задание некоторого отображения $f\colon \boldsymbol{A} \to \boldsymbol{R}$, где $\boldsymbol{A} \subset \boldsymbol{R}$. Числовые функции могут задаваться не только формулами, но и описательно. Такова, например, *функция Дирихле*, определяемая следующим образом:

$$f(x) = \begin{cases} 0, \text{ если число } x \text{ иррационально,} \\ 1, \text{ если число } x \text{ рационально.} \end{cases}$$

Например, $f(1) = 1$, $f\left(\frac{7}{2}\right) = 1$, $f(\sqrt{2}) = 0$, $f(\pi) = 0$. Несмотря на то, что эта функция не записана формулой, она принимает вполне определенное значение (0 или 1) для любого $x \in \boldsymbol{R}$.

В качестве еще одного примера рассмотрим так называемую характеристическую функцию множества. Пусть $\boldsymbol{A}$ — некоторое подмножество числовой прямой $\boldsymbol{R}$. Зададим числовую функцию $\varphi(x)$, определенную на всей числовой прямой $\boldsymbol{R}$, следующим образом:

$$\varphi(x) = \begin{cases} 1, \text{ если } x \in \boldsymbol{A}, \\ 0, \text{ если } x \notin \boldsymbol{A}. \end{cases}$$

Это и есть *характеристическая функция* множества $\boldsymbol{A}$. Если, например, $\boldsymbol{A}$ есть множество всех рациональных чисел, то характеристическая функция этого множества совпадает с рассмотренной выше *функцией Дирихле*.

Отображения рассматриваются и в геометрии. Все *геометрические преобразования*, играющие важную роль в этой области математики, являются отображениями. Например, *центральная симметрия* с центром O есть отображение, которое каждой точке A, отличной от O, ставит в соответствие такую точку B, что отрезок AB имеет O своей серединой. Точке же O ставится в соответствие сама точка O, т. е. как говорят, точка O является *неподвижной точкой* этого отображения. На рис. 95 образом точки A при центральной симметрии относительно точки O является точка B, образом точки B является A, образом точки C является D, образом точки D является C, образом точки O является O. Обозначая эту симметрию буквой s, мы можем написать: $s(A) = B$, $s(B) = A$, $s(C) = D$, $s(D) = C$, $s(O) = O$.

Поворот r плоскости $\boldsymbol{P}$ на угол $\frac{\pi}{2}$ вокруг центра O также представляет собой некоторое отображение r: $\boldsymbol{P} \to \boldsymbol{P}$ этой плоскости на ту же самую плоскость. Точку A, отличную от O, это отображение переводит в такую точку $B = r(A)$, что A и B лежат на одной окружности с центром O и величина угла AOB, отсчитываемая в положительном направлении (т. е. против хода часовой стрелки), равна $\frac{\pi}{2}$ (рис. 96). Точка O переводится рассматриваемым поворотом в ту же точку O, т. е. O является неподвижной точкой этого отображения. Вообще, любое геометрическое преобразование f плоскости $\boldsymbol{P}$ (осевая симметрия, параллельный перенос, гомотетия и многие другие) представляет собой некоторое отображение f: $\boldsymbol{P} \to \boldsymbol{P}$. Иначе говоря, каждое геометрическое преобразование f плоскости $\boldsymbol{P}$ есть как бы «геометрическая функция», которая, однако, в отличие от числовой

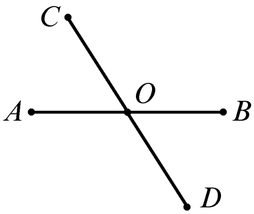


Рис. 95

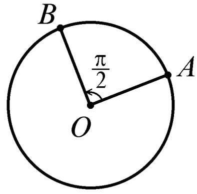


Рис. 96

функции, имеет своей областью определения не некоторое множество чисел, а множество $\boldsymbol{P}$, состоящее из точек, и образами также являются не числа, а точки, т. е. для любой точки $A \in \boldsymbol{P}$ ее образ $B = f(A)$ также есть некоторая точка плоскости $\boldsymbol{P}$.

Таким образом, рассмотрение не только множеств, но и отображений сближает алгебру и геометрию, т. е. составляет некоторый общий взгляд на эти разделы школьной математики.

Попытки модернизировать школьный курс математики, весьма сильно отстававший от современной науки, были сделаны во многих странах мира в середине XX столетия. Чтобы рассказать об этом, мы начнем издалека.

Три русских писателя, братья Жемчужниковы и Алексей Константинович Толстой, выпустили интересную и весьма известную книгу, назвав ее сочинениями Козьмы Пруткова. Этим псевдонимом они охарактеризовали напыщенного и самонадеянного вымышленного автора, который назидательно высказывал вроде бы банальные, но, если вдуматься, очень глубокие истины. А в середине XX столетия группа французских математиков (Картан, Дьедонне и ряд других ученых) создала «математического Козьму Пруткова», издав серию книг под псевдонимом Никола́ Бурбаки́. Художники создали и портрет этого вымышленного самодовольного автора, генерала в отставке, который решил остаток жизни посвятить написанию трактатов о математических науках. Книги Бурбаки по разным разделам математики являются ультрасовременными и содержат много интересных идей, объединяющих общим взглядом различные ветви математики, главным образом на базе теории множеств и топологии. Эти десятки трактатов Бурбаки внесли существенный вклад в формирование математических взглядов современных ученых, особенно молодых. Любопытен эпизод, связанный с международным математическим конгрессом, проходившим в 1966 году в Москве. При подготовке этого конгресса оргкомитет разослал во все страны мира извещение о предстоящем математическом форуме. Каждый математик, имевший научные публикации, получил приглашение принять участие в работе конгресса, для чего он должен был написать о своем согласии и выслать оргвзнос в размере 40 американских долларов. Получил такое приглашение и Никола Бурбаки, автор десятков широко известных книг. Эта шутка московского оргкомитета была подхвачена французскими математиками, которые не пожалели 40 долларов и сообщили, что известный французский математик выражает свое согласие участвовать в работе конгресса. И перед началом одного из пленарных заседаний конгресса в фойе было вывешено кричащее объявление, извещавшее о том, что в этот день, после заседаний, в пресс-центре конгресса состоится встреча с известным французским математиком Никола Бурбаки...

Модные теоретико-множественные концепции Бурбаки сопровождались и некоторой ревизией математических обозначений. Так, числовой интервал, издавна записываемый в виде $(a; b)$, Бурбаки предложил обозначать символом $]a; b[$, в котором «вывернутые» квадратные скобки должны были символизировать то, что, в отличие от отрезка $[a; b]$ интервал $]a; b[$ не содержит своих концевых точек, т. е. описывается неравенствами $a < x < b$.

Всеобщее увлечение идеями Бурбаки совпало по времени с процессом модернизации школьного образования (а может быть, не только «совпало» по времени, но и было органически связано с этим процессом). Новые учебники, изданные в ряде стран мира, уделяли существенное внимание рассмотрению множеств и отображений. В алгебре это выражалось, в частности, в том, что рассматривалось *множество* всех решений уравнения или неравенства, и оно записывалось с помощью знаков $\in$, $\subset$, $\bigcup$, $\bigcap$. Мы это видели на примере неравенства (6) и его решения (7) в пункте 12.

Разумеется, дело не в обозначениях, а в самих идеях множества и отображения, в удобстве *математического языка*, использующего эти понятия и объединяющего алгебру с геометрией. В геометрии идеи этого круга были связаны с прогрессивной тенденцией, восходящей к замечательному немецкому математику и педагогу Феликсу Клейну, который говорил о широком использовании геометрических преобразований взамен устаревших евклидовых приемов геометрических доказательств (об этом еще будет идти более подробный разговор в последующих беседах).

В России попытка модернизировать школьный курс математики и связанная с нею реформа школьного обучения были особенно драматичны. Инициатором и активным руководителем реформы школьного математического образования стал выдающийся математик XX столетия академик Андрей Николаевич Колмогоров. Правда, созданный по новой программе курс геометрии оказался трудным. Но в идейном плане это был совершенно новый курс, содержавший современные идеи и способствовавший более глубокому осмыслению математики, чем прежние учебники. И хотя новые учебники были трудными для школьного учителя, но в течение 2 – 3 лет работы по этим учебникам школьные преподаватели их освоили, «пропустили через свое сердце» и стали на три головы выше в своем идейном понимании математики.

И тут грянул гром среди ясного неба. Начался «крестовый поход» против колмогоровских идей и против употребления в школе множеств и других новых для школы понятий, возглавленный академиком Л. С. Понтрягиным. Личность Л. С. Понтрягина весьма противоречива в истории науки. Талантливый математик, получивший ряд первоклассных результатов в топологии и других областях математики, он в последние годы жизни придерживался весьма реакционных взглядов в области организации науки и школьного образования.

Понтрягин написал статью, в которой попытался философскими, социологическими фразами и малоубедительными соображениями о школьном преподавании ниспровергнуть колмогоровскую концепцию математического образования, остановить наметившиеся прогрессивные тенденции в школьном преподавании (и, в частности, введение понятий множества и отображения). Эта статья была поддержана идеологическим отделом большевистской партии. К началу понтрягинской статьи, ниспровергавшей идеи Колмогорова, были присоединены философские фразы, и она была опубликована в журнале «Коммунист».

Вскоре после этого в школах был введен учебник геометрии, написанный академиком А. В. Погореловым. Это был лишь несколько обновленный вариант учебника геометрии, написанного еще в первой половине века талантливым преподавателем А. П. Киселевым. Великолепный в плане обучения, учебник А. П. Киселева стал уже несовременным по содержанию. И тогда А. В. Погорелов несколько «подштопал» обветшавший, но хорошо написанный киселевский учебник, добавив свою систему аксиом (как оказалось, математически противоречивую). Никаких новых задач, никакой замены архаичных евклидовых методов геометрических доказательств в этом учебнике не было. Но под нажимом партийных органов учебник Погорелова был введен в школьное преподавание. Начатая А. Н. Колмогоровым тенденция обновления учебников в духе Клейна и Бурбаки была прервана. Однако его прогрессивная деятельность все же привела к некоторому улучшению последующих учебников; в частности, были введены векторы и геометрические преобразования. К этому мы еще вернемся в беседе 15.

Задачи и упражнения

41. Являются ли отображениями следующие соответствия для множества живущих людей?

а) Каждому человеку ставится в соответствие его дочь.

б) Каждому человеку ставится в соответствие его мать.

в) Каждому человеку ставится в соответствие его год рождения.

42. Пусть f и g — характеристические функции множеств $\boldsymbol{A}$ и $\boldsymbol{B}$, расположенных во множестве $\boldsymbol{U}$. Докажите, что характеристической функцией множества $\boldsymbol{A} \cap \boldsymbol{B}$ является функция $h = f \cdot g$. Запишите характеристическую функцию множества $\boldsymbol{A} \cup \boldsymbol{B}$ через функции f и g.

43. Через $\Theta(x)$ обозначают характеристическую функцию луча $[0; \infty)$, она называется функцией Хэвисайда. Докажите тождество: $|x| = x(2\Theta(x) - 1)$.

44. Участникам математической олимпиады было предложено пять задач. Является ли функцией соответствие, сопоставляющее каждому участнику:

а) номера решенных им задач;

б) сумму номеров решенных им задач?

45. Найдите область определения функции $f(x) = \sqrt{3 - \sqrt{x^2 - 16}}$.

16. Некоторые виды отображений

Пусть f: $\boldsymbol{A} \to \boldsymbol{B}$ — некоторое отображение. Возьмем произвольный элемент $b \in \boldsymbol{B}$ и рассмотрим в $\boldsymbol{A}$ подмножество, состоящее из всех тех

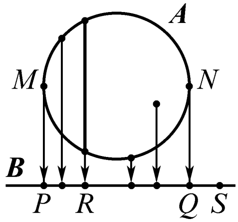


Рис. 97

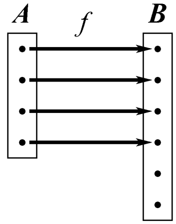


Рис. 98

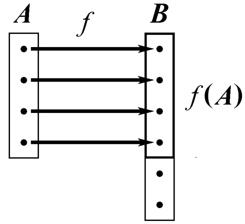


Рис. 99

элементов, образ которых совпадает с b. Это подмножество называется *прообразом* элемента b и обозначается через $f^{-1}(b)$.

Так, для отображения, показанного на рис. 90, прообраз элемента $p \in \boldsymbol{B}$ состоит из двух элементов: $f^{-1}(p) = \{b, c\}$; прообраз каждого из элементов r, s состоит лишь из одного элемента: $f^{-1}(r) = a$, $f^{-1}(s) = d$; что же касается $q \in \boldsymbol{B}$, то его прообраз пуст — совсем не содержит элементов: $f^{-1}(q) = \varnothing$.

В качестве второго примера рассмотрим в плоскости круг $\boldsymbol{A}$ и прямую $\boldsymbol{B}$, а через $g\colon \boldsymbol{A} \to \boldsymbol{B}$ обозначим *ортогональное проектирование* круга на эту прямую (рис. 97). В этом случае прообраз каждой из точек P, Q содержит лишь один элемент: $g^{-1}(P) = M$, $g^{-1}(Q) = N$; далее для любой другой точки R отрезка PQ ее прообраз $g^{-1}(R)$ есть *отрезок*; если же $S \in \boldsymbol{B} \setminus PQ$, то прообраз $g^{-1}(S)$ есть пустое множество (через PQ мы здесь обозначаем *отрезок* с концами P, Q).

Отображения различаются по своим свойствам, прежде всего, тем, каковы прообразы различных точек. Если прообраз каждого элемента содержит *не более одного* элемента (т. е. либо содержит один элемент, либо пуст), то $f\colon \boldsymbol{A} \to \boldsymbol{B}$ называется *вложением* (или, в соответствии с французской терминологией, *инъективным отображением*, рис. 98). Таким образом, если $f\colon \boldsymbol{A} \to \boldsymbol{B}$ — вложение, то, отождествляя каждый элемент $a \in \boldsymbol{A}$ с его образом $f(a) \in \boldsymbol{B}$, мы как бы «вкладываем» множество $\boldsymbol{A}$ в $\boldsymbol{B}$ в качестве его подмножества (рис. 99), откуда и происходит термин «вложение». Вообще говоря, если $f\colon \boldsymbol{A} \to \boldsymbol{B}$ — вложение, то во множестве $\boldsymbol{B}$ могут быть «лишние» элементы, не являющиеся образами каких-либо элементов из $\boldsymbol{A}$. Иными словами, вложение $f\colon \boldsymbol{A} \to \boldsymbol{B}$ устанавливает взаимно однозначное соответствие между множеством $\boldsymbol{A}$ и некоторым подмножеством $f(\boldsymbol{A})$ множества $\boldsymbol{B}$.

Далее, если для любого $b \in \boldsymbol{B}$ прообраз $f^{-1}(b)$ непуст, то $f\colon \boldsymbol{A} \to \boldsymbol{B}$ называется *наложением* (или *сюръективным отображением*, в соответствии с французской терминологией), или еще *отображением на* множество $\boldsymbol{B}$ (рис. 100). Иначе говоря, отображение $f\colon \boldsymbol{A} \to \boldsymbol{B}$ представляет собой наложение, если в $\boldsymbol{B}$ нет «лишних» элементов, т. е.

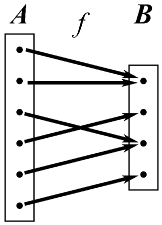


Рис. 100

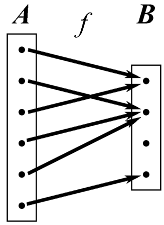


Рис. 101

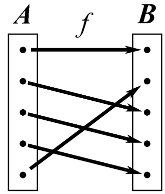


Рис. 102

каждый элемент множества $\boldsymbol{B}$ является образом хотя бы одного элемента из $\boldsymbol{A}$.

Вообще, пусть $f\colon \boldsymbol{A} \to \boldsymbol{B}$ — произвольное отображение; через $f(\boldsymbol{A})$ обозначим образ множества $\boldsymbol{A}$, т. е. множество всех элементов из $\boldsymbol{B}$, являющихся образами элементов из $\boldsymbol{A}$. Если $f(\boldsymbol{A})$ *совпадает* с множеством $\boldsymbol{B}$, то f является *наложением*; если же в $\boldsymbol{B}$ есть «лишние» элементы, т. е. $f(\boldsymbol{A})$ — *собственное* подмножество множества $\boldsymbol{B}$, то f наложением не является (рис. 101).

Наконец, отображение, одновременно являющееся вложением и наложением, называется *взаимно однозначным* (или *биективным*). Такое отображение $f\colon \boldsymbol{A} \to \boldsymbol{B}$ устанавливает взаимно однозначное соответствие между множествами $\boldsymbol{A}$ и $\boldsymbol{B}$ (рис. 102). Ясно, что взаимно однозначное отображение множества $\boldsymbol{A}$ на множество $\boldsymbol{B}$ существует только в том случае, если эти множества имеют одинаковую мощность. Можно сказать и иначе: отображение $f\colon \boldsymbol{A} \to \boldsymbol{B}$ в том и только в том случае взаимно однозначно, если уравнение $f(x) = b$ имеет для любого $b \in \boldsymbol{B}$ единственное решение $x \in \boldsymbol{A}$.

Рассмотрим теперь понятие графика отображения. Пусть $f\colon \boldsymbol{A} \to \boldsymbol{B}$ — некоторое отображение множества $\boldsymbol{A}$ в множество $\boldsymbol{B}$. Для каждого элемента $x \in \boldsymbol{A}$ рассмотрим пару $(x; f(x))$. Так как $f(x) \in \boldsymbol{B}$, то эта пара является элементом произведения $\boldsymbol{A}\times\boldsymbol{B}$ (рис. 103). Множество всех таких пар и есть *график* рассматриваемого отображения f. Таким образом, если $\boldsymbol{A}$ представлять себе *строкой* в нижней части чертежа, а $\boldsymbol{B}$ — *столбцом* слева (как на рис. 103, 104), то в прямо-

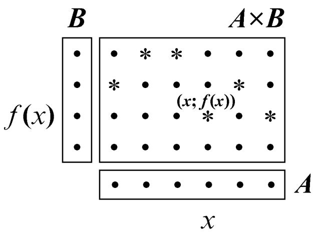


Рис. 103

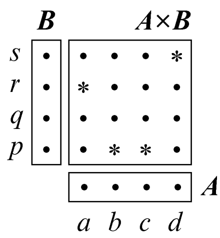


Рис. 104

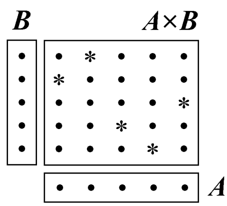


Рис. 105

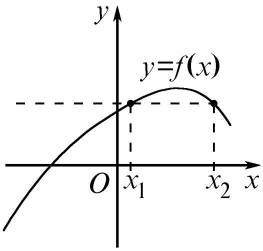


Рис. 106

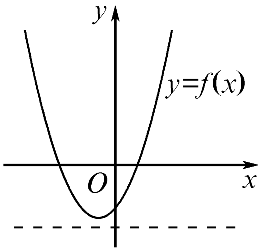


Рис. 107

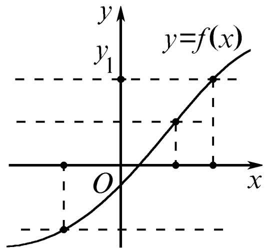


Рис. 108

угольной таблице, изображающей произведение $\boldsymbol{A}\times\boldsymbol{B}$, в столбце над элементом $x\in\boldsymbol{A}$ должен быть отмечен лишь один элемент $(x;y)$, а именно тот, в котором y есть образ элемента x. Все отмеченные таким образом элементы произведения $\boldsymbol{A}\times\boldsymbol{B}$ и составляют, вместе взятые, график отображения f. Рис. 104 изображает график отображения, указанного на рис. 90.

На рис. 104 имеются два элемента b, c, образы которых совпадают: $f(b)=f(c)=p$, т. е. в произведении $\boldsymbol{A}\times\boldsymbol{B}$ имеется строка, содержащая более чем один элемент графика. Это означает, что отображение f не является вложением. На рис. 104 имеется в произведении $\boldsymbol{A}\times\boldsymbol{B}$ строка, совсем не содержащая элементов графика. Это означает, что отображение $f\colon\boldsymbol{A}\to\boldsymbol{B}$ не является наложением. Лишь в том случае, когда каждая строка произведения $\boldsymbol{A}\times\boldsymbol{B}$ содержит *ровно один* элемент графика, отображение $f\colon\boldsymbol{A}\to\boldsymbol{B}$ взаимно однозначно (рис. 105).

В случае *числовой функции* $f\colon\boldsymbol{A}\to\boldsymbol{R}$ (где $\boldsymbol{A}\subset\boldsymbol{R}$) ее график, согласно сказанному выше, есть множество всех таких пар $(x;y)$, что x принадлежит области определения функции f (т. е. множеству $\boldsymbol{A}$), а $y=f(x)$ есть образ точки x, т. е. значение функции f в точке x. Это — обычное определение графика функции.

На рис. 106 имеются две точки x_1, x_2, для которых их образы (т. е. значения функции f) совпадают: $f(x_1)=f(x_2)$. Иными словами, соответствующая «строка», т. е. прямая, параллельная оси абсцисс и проходящая через точку $(x_1;f(x_1))$, содержит две (или более) точки графика. Это означает, что отображение f не является вложением, т. е. по данному y соответствующее значение x не восстанавливается однозначно. А на рис. 107 имеется «строка», совсем не содержащая элементов графика, т. е. функция f отображает множество $\boldsymbol{A}$ не на всю числовую прямую $\boldsymbol{R}$.

Впрочем, если числовую функцию $f(x)$ рассматривать как отображение $f\colon\boldsymbol{A}\to\boldsymbol{B}$, где $\boldsymbol{A}$ — область ее определения, а $\boldsymbol{B}=f(\boldsymbol{A})$ — не вся числовая прямая, а *образ* множества $\boldsymbol{A}$ (т. е. *множество всех значений* функции f), то функция f автоматически становится наложением. Таким образом, числовую функцию можно в этом смысле всегда считать наложением, т. е. сопоставляющей каждому $x \in \boldsymbol{A}$ некоторую точку множества значений $\boldsymbol{B}$. Взаимная однозначность будет в том и только в том случае, если для каждого $y_1 \in \boldsymbol{B}$ соответствующая прямая, состоящая из всех точек $(x; y_1)$, $x \in \boldsymbol{R}$, пересекает график в единственной точке (рис. 108).

Задачи и упражнения

46. Будут ли наложениями или вложениями отображения, определенные в упражнениях 41в и 44б?

47. Является ли наложением или вложением отображение, определенное в упражнении 45, если его рассматривать как отображение множества $[-5; -4] \cup [4; 5]$ в множество действительных чисел?

48. Какое из следующих отображений множества действительных чисел в себя является наложением или вложением: а) $y = x^2$; б) $y = x^3$; в) $y = 2^x$?

49. Для функции $f(x)$, определенной в упражнении 45, найдите прообраз отрезка [1; 2].

50. Постройте график функции $f(x) = [x]$, где $[x]$ — наибольшее целое число, не превосходящее числа x (эта функция называется *целой частью* числа x).

17. Обратное отображение

В случае, когда отображение f: $\boldsymbol{A} \to \boldsymbol{B}$ взаимно однозначно, можно определить *обратное отображение* f^{-1}: $\boldsymbol{B} \to \boldsymbol{A}$. Ввиду важности этого понятия остановимся на нем подробнее.

Пусть f: $\boldsymbol{A} \to \boldsymbol{B}$ — некоторое взаимно однозначное отображение. Тогда для каждого элемента $y \in \boldsymbol{B}$ его прообраз $f^{-1}(y)$ содержит ровно один элемент x множества $\boldsymbol{A}$, т. е. имеется только один элемент $x \in \boldsymbol{A}$, для которого $f(x) = y$. На рис. 109 от элемента x к элементу y ведет стрелка. Поставив в соответствие элементу y этот элемент $x = f^{-1}(y)$, мы получаем некоторое отображение множества $\boldsymbol{B}$ на множество $\boldsymbol{A}$. Схематическое его изображение мы получим, если на всех стрелках, ведущих от элементов множества $\boldsymbol{A}$ к элементам множества $\boldsymbol{B}$, мы изменим направление на противоположное (рис. 110). Получаемое отображение $\boldsymbol{B} \to \boldsymbol{A}$ обозначается через f^{-1} и называется отображением, *обратным* отображению f. Иначе говоря, обратное

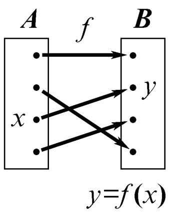


Рис. 109

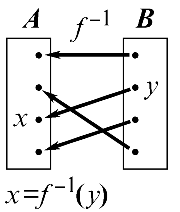


Рис. 110

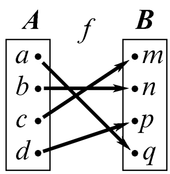


Рис. 111

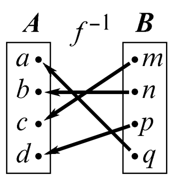


Рис. 112

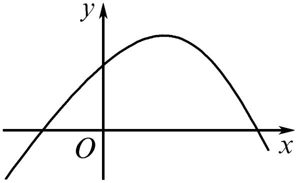


Рис. 113

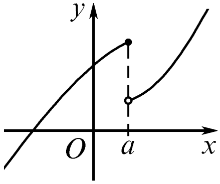


Рис. 114

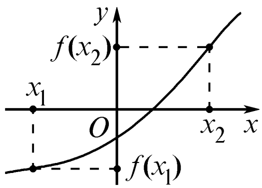


Рис. 115

отображение определяется следующим образом: если $f(x) = y$, то $f^{-1}(y) = x$.

На рис. 111 схематически показано взаимно однозначное отображение f: $\boldsymbol{A} \to \boldsymbol{B}$. Оно определяется соотношениями: $f(a) = q$, $f(b) = n$, $f(c) = m$, $f(d) = p$. Обратное отображение f^{-1} имеет вид: $f^{-1}(m) = c$, $f^{-1}(n) = b$, $f^{-1}(p) = d$, $f^{-1}(q) = a$ (рис. 112). Отображение f «переводит» элемент $a \in \boldsymbol{A}$ в элемент $q \in \boldsymbol{B}$, а обратное отображение f^{-1} как бы «возвращает» его на прежнее место: $f(a) = q$, $f^{-1}(q) = a$.

Для числовых функций нахождение их обратных функций представляет собой важную и интересную задачу. Чаще всего она рассматривается для *непрерывных монотонных* функций. О непрерывности мы поговорим в другой раз. Здесь же ограничимся наглядным пояснением, что числовая функция непрерывна, если ее график представляет собой плавную линию без разрывов (рис. 113), т. е. как бы может быть вычерчен без отрыва карандаша от бумаги. Функция на рис. 114 непрерывной не является — вблизи точки $x = a$ она имеет *разрыв*.

Монотонными же называют *возрастающие* функции (т. е. такие, что при $x_2 > x_1$ справедливо неравенство $f(x_2) > f(x_1)$, рис. 115) и *убывающие* функции (т. е. такие, что при $x_2 > x_1$ справедливо противоположное неравенство $f(x_1) > f(x_2)$, рис. 116). У возрастающих функций график при движении вправо поднимается вверх, а у убывающих он при движении вправо опускается вниз. На рис. 117 изображен график функции, которая не является монотонной.

Итак, рассмотрим непрерывную и монотонную функцию, скажем, возрастающую. Пусть, например, эта функция определена на отрезке

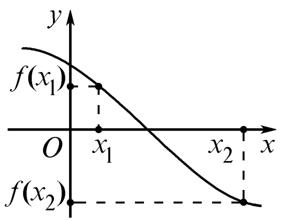


Рис. 116

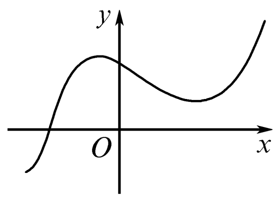


Рис. 117

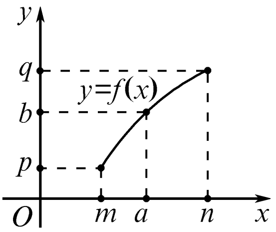


Рис. 118

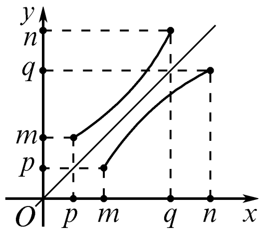


Рис. 119

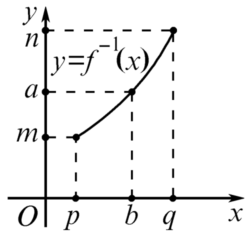


Рис. 120

$[m; n]$, а множеством ее значений является отрезок $[p; q]$ (рис. 118), т. е. рассматриваемая функция имеет вид $f\colon \boldsymbol{A} \to \boldsymbol{B}$, где $\boldsymbol{A} = [m; n]$ и $\boldsymbol{B} = [p; q]$. Наглядно понятно, что каждая прямая, параллельная оси абсцисс и проходящая через некоторую точку отрезка $[p; q]$ (расположенного на оси ординат), пересекает график функции только в одной точке, т. е. имеется *единственная* точка $a \in [m; n]$, для которой $f(a) = b$. Иначе говоря, $f^{-1}(b) = a$ есть вполне определенная точка отрезка $\boldsymbol{A} = [m; n]$. Этим в данном случае и определяется обратная функция $f^{-1}\colon \boldsymbol{B} \to \boldsymbol{A}$.

Заметим теперь, что при построении графика функции мы обычно точку, принадлежащую области определения функции, берем на оси абсцисс, а значение функции в этой точке откладываем по оси ординат (см., например, рис. 118). Но при рассмотрении обратной функции $f^{-1}\colon \boldsymbol{B} \to \boldsymbol{A}$ мы берем точку b, принадлежащую области определения обратной функции (т. е. отрезку $\boldsymbol{B} = [p; q]$), а она находится на оси ординат; соответствующее значение $a = f^{-1}(b)$ обратной функции оказывается отложенным по оси абсцисс. Чтобы исправить это неудобство, мы применим к графику, изображенному на рис. 118, *осевую симметрию*, причем в качестве оси симметрии возьмем биссектрису первого и третьего координатных углов (рис. 119). Эта симметрия удобна тем, что при ее применении ось абсцисс и ось ординат меняются местами. В частности, образ отрезка $\boldsymbol{B} = [p; q]$, расположенного на оси ординат, окажется расположенным на оси абсцисс.

Теперь рис. 118 примет вид, показанный на рис. 120, и станет более удобным: точка $b \in [p; q]$ окажется расположенной на оси абсцисс, а соответствующее значение $a = f^{-1}(b)$ обратной функции будет отложено по оси ординат. Иначе говоря, рис. 120 даст нам как раз график обратной функции f^{-1}. Вспомнив, что рис. 120 получен из рис. 118 с помощью осевой симметрии, мы получаем следующее удобное и важное правило.

Пусть линия L представляет собой график непрерывной и монотонной функции f; тогда линия L′, симметричная L относительно биссект-

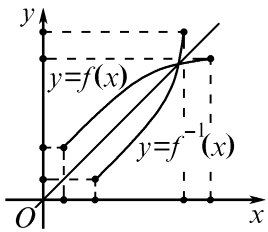


Рис. 121

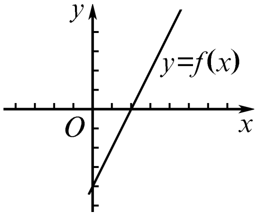


Рис. 122

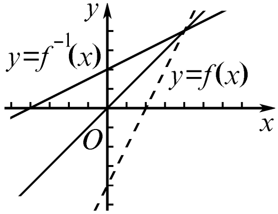


Рис. 123

рисы первого и третьего координатных углов, является графиком обратной функции f^{-1}.

Более коротко: график непрерывной монотонной функции и график ее обратной функции симметричны друг другу относительно указанной биссектрисы (рис. 121). Это правило применимо не только к функциям, определенным на отрезке, но также и к случаям, когда областью определения функции является луч или вся прямая.

В качестве примера рассмотрим функцию

$$f(x) = 2x - 4. \tag{3}$$

Она определена на всей числовой прямой, непрерывна и является монотонной (возрастающей). График ее приведен на рис. 122. Областью ее определения (и множеством значений) является вся числовая прямая, т. е. эта функция имеет вид $f\colon \boldsymbol{R} \to \boldsymbol{R}$. Найдем ее обратную функцию.

Если задано произвольное число $b \in \boldsymbol{R}$, то $a = f^{-1}(b)$ есть прообраз точки b, т. е. такое число $a \in \boldsymbol{R}$, что $f(a) = b$. Иными словами, $a = f^{-1}(b)$ есть такое число, для которого (согласно (3)) справедливо равенство $2a - 4 = b$. Из этого равенства, рассматриваемого как уравнение относительно a, мы находим: $a = \frac{1}{2}b + 2$. Таким образом, $f^{-1}(b) = \frac{1}{2}b + 2$. Это верно для любого $b \in \boldsymbol{R}$, т. е. обратная функция имеет вид $f^{-1}(x) = \frac{1}{2}x + 2$. Ее график симметричен графику функции (3) относительно биссектрисы первого и третьего координатных углов (рис. 123).

В качестве второго примера рассмотрим функцию

$$f(x) = \sqrt{x}. \tag{4}$$

Она определена лишь при $x \geq 0$, т. е. на луче $[0; \infty)$, непрерывна и является монотонной (возрастающей). График ее показан на рис. 124. Значения рассматриваемой функции также принадлежат лучу $[0; \infty)$, т. е. она имеет вид $f\colon [0; \infty) \to [0; \infty)$. Найдем ее обратную функцию.

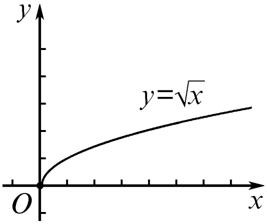


Рис. 124

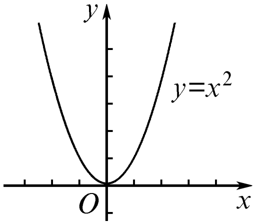


Рис. 125

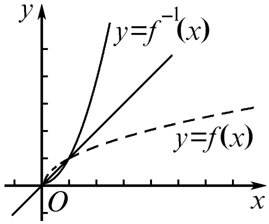


Рис. 126

Если задано произвольное число $b \in [0; \infty)$, то $a = f^{-1}(b)$, есть прообраз точки b, т, е. такое число $a \in [0; \infty)$, что $f(a) = b$. Иными словами, a есть такое число, что $b = \sqrt{a}$ (см. (4)), и потому $a = b^2$, т. е. $f^{-1}(b) = b^2$. Это верно для любого $b \in [0; \infty)$, т. е. обратная функция f^{-1} определена лишь на луче $[0; \infty)$, а ее значения даются формулой $f^{-1}(x) = x^2$. Заметим, что функция $y = x^2$ определена для *всех* действительных x (рис. 125), а не только на луче $[0; \infty)$. Рассматривая ее *только* на луче $[0; \infty)$, мы и получаем искомую обратную функцию $f^{-1}(x)$. Ее график (рис. 126) симметричен графику функции (4) относительно биссектрисы первого и третьего координатных углов.

Ограничимся пока этими примерами. Важные примеры обратных функций мы будем иметь в последующих беседах при рассмотрении тригонометрических и показательных функций. Существенную роль играют обратные отображения и в геометрии, о чем также будет идти речь впоследствии.

Задачи и упражнения

51. Рассмотрим отображение множества действительных чисел в себя, определенное следующим образом: $f(x) = \dfrac{1}{x}$, если $x \neq 0$, и $f(0) = 0$.

Существует ли у этого отображения обратное?

52. Укажите отрезки действительной прямой, на которых функция $f(x) = x^3 - x^2$ имеет обратную.

53. Рассмотрим функцию $f(x) = x + [x]$ (функция $[x]$ определена в упражнении 50), отображающую множество действительных чисел в себя. Имеет ли эта функция обратную? Имеет ли она обратную на интервале (100; 101)?

54. Существует ли отрезок, на котором функция Дирихле имеет обратную?

55. Пусть функции $f(x)$ и $g(x)$ отображают множество действительных чисел на себя и имеют обратные. Верно ли, что и их сумма также имеет обратную?

18. Композиция отображений

Рассмотрим три множества $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, и пусть даны некоторые отображения $f\colon \boldsymbol{A} \to \boldsymbol{B}$, $g\colon \boldsymbol{B} \to \boldsymbol{C}$. Их можно записать в виде «цепочки»:

$$A \xrightarrow{f} B \xrightarrow{g} C. \tag{5}$$

Для каждого элемента $a \in A$ определен его образ $b = f(a)$, содержащийся в множестве B. В свою очередь для этого элемента $b \in B$ определен его образ $c = g(b)$, содержащийся в множестве C (рис. 127). Поставим в соответствие элементу a сразу этот элемент c, минуя промежуточное множество B (пунктирная линия на рис. 127). Мы получаем таким образом некоторое отображение множества A в множество C (рис. 128). Оно называется *композицией* отображений f и g и обозначается через $g \circ f$.

Обратите внимание: первое применяемое отображение (т. е. f) пишется справа, а отображение g, применяемое вторым, — слева. Такая форма записи применяется по следующим соображениям. Элемент c, который в результате композиции ставится в соответствие элементу a, имеет вид $c = g(b)$, где $b = f(a)$ — образ элемента a при первом отображении. Таким образом, $c = g(b) = g(f(a))$. В этой записи $g(f(a))$ отображение g оказывается записанным слева, и именно в таком порядке записываются f и g при обозначении композиции: f справа, g слева. Итак, по определению

$$(g \circ f)(a) = g(f(a))$$

для любого $a \in A$.

Например, если f — *центральная симметрия* плоскости β относительно точки P, а g — *центральная симметрия* относительно точки Q (рис. 129), то композиция $g \circ f$ отображений $f\colon \beta \to \beta$ и $g\colon \beta \to \beta$ представляет собой параллельный перенос на вектор, вдвое больший, чем вектор $\overrightarrow{PQ}$. Иначе говоря, для любой точки $A \in \beta$ ее образ C при отображении $g \circ f$ определяется равенством $\overrightarrow{AC} = 2\,\overrightarrow{PQ}$ (рис. 130).

Еще один пример получается, если взять в качестве f *осевую симметрию* относительно прямой k, а в качестве g — *осевую симметрию* относительно прямой l, причем k и l пересекаются в некоторой точке O. Из рассмотрения рис. 131, на котором отмечены равные углы, нетрудно заключить, что точка A переводится композицией $g \circ f$ в такую точку C, что угол AOC вдвое больше угла α между прямыми k и l, причем точки A и C находятся на одинаковом рас-

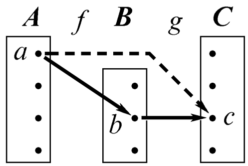


Рис. 127

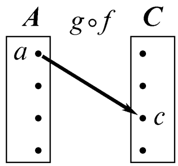


Рис. 128

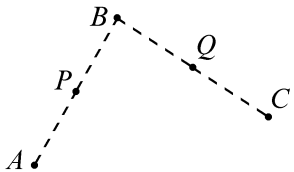


Рис. 129

стоянии от точки O. Иначе говоря, композиция $g \circ f$ представляет собой поворот плоскости β вокруг точки O на угол 2α (рис. 132).

Приведенные примеры показывают, что операция композиции отображений играет важную роль при рассмотрении геометрических преобразований, а мы еще будем в последующих беседах говорить, следуя Феликсу Клейну, о том, что геометрические преобразования — это подлинная основа всей элементарной (и не только элементарной) геометрии. Композиция отображений играет важную роль и в алгебре, где она связана с понятием сложной функции. Рассмотрим этот вопрос подробнее.

Пусть $f(x)$ и $g(x)$ — две числовые функции. Вначале для простоты рассмотрим случай, когда областью определения каждой из функций f, g является вся числовая прямая, т. е. они представляют собой отображения f: $\boldsymbol{R} \to \boldsymbol{R}$, g: $\boldsymbol{R} \to \boldsymbol{R}$. Тогда можно рассматривать *сложную функцию* $g(f(x))$, т. е. композицию $g \circ f$ этих отображений. Иначе говоря, сложная функция получается, если в функцию $g(y)$ подставить вместо y значение $y = f(x)$.

Рассмотрим, например, биквадратный многочлен

$$h(x) = x^4 - 7x^2 + 12. \tag{6}$$

Здесь удобно рассмотреть вспомогательную функцию $y = f(x) = x^2$. Тогда многочлен (6) принимает вид $y^2 - 7y + 12$, и это дает повод ввести в рассмотрение еще одну функцию $g(x) = x^2 - 7x + 12$. Композиция $g \circ f$ этих функций, т. е. сложная функция

$$g(f(x)) = g(y) = y^2 - 7y + 12 = x^4 - 7x^2 + 12$$

совпадает с многочленом (6). Таким образом, $h = g \circ f$.

Подобно этому, функция $h(x) = \sin 5x$ есть композиция двух функций $f(x) = 5x$ и $g(x) = \sin x$. В самом деле, полагая $y = f(x) = 5x$, мы находим: $g(f(x)) = g(y) = \sin y = \sin 5x$, т. е. $h = g \circ f$.

Представление заданной функции в виде композиции более простых функций часто используется при решении уравнений. Чтобы рассмотреть этот вопрос, вначале проведем некоторое общее рассуждение. Рассмотрим снова «цепочку» отображений (5). Пусть c —

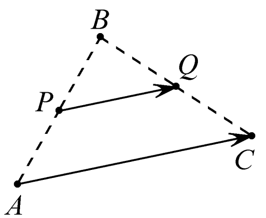


Рис. 130

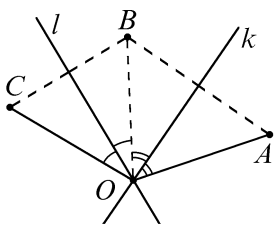


Рис. 131

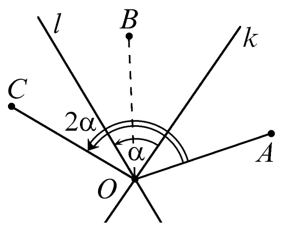


Рис. 132

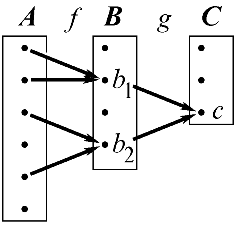


Рис. 133

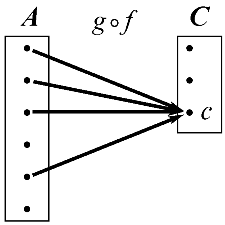


Рис. 134

некоторый элемент множества $\boldsymbol{C}$, и пусть $g^{-1}(c) = \{b_1, b_2, ...\} \subset \boldsymbol{B}$ — его прообраз, т. е. b_1, b_2, ... — все те элементы множества $\boldsymbol{B}$, которые переводятся отображением g в элемент c (рис. 133). В свою очередь мы можем рассматривать прообразы $f^{-1}(b_1)$, $f^{-1}(b_2)$, ... этих элементов при отображении f. Как показывает рис. 134, *объединение* всех этих прообразов как раз представляет собой прообраз элемента c при отображении $g \circ f$, т. е.

$$(g \circ f)^{-1}(c) = f^{-1}(b_1) \cup f^{-1}(b_2) \cup \dots\ . \tag{7}$$

Пусть теперь задано уравнение

$$g(f(x)) = 0, \tag{8}$$

где $g\colon \boldsymbol{R} \to \boldsymbol{R}$ и $f\colon \boldsymbol{R} \to \boldsymbol{R}$ — некоторые числовые функции. Это уравнение имеет вид $h(x) = 0$, где $h = g \circ f$ — сложная функция, стоящая в левой части уравнения (8).

Решить уравнение $h(x) = 0$ — значит найти множество *всех* его корней, т. е. действительных чисел, для которых значение функции h равно нулю. Иначе говоря, решение этого уравнения есть прообраз точки 0, т. е. $h^{-1}(0)$, или, что то же самое, $(g \circ f)^{-1}(0)$. Согласно сказанному выше, надо сначала найти прообраз $g^{-1}(0)$, т. е. найти все корни b_1, b_2, ... уравнения $g(x) = 0$. После этого, согласно формуле (7), надо найти все прообразы $f^{-1}(b_1)$, $f^{-1}(b_2)$, ...; их *объединение* как раз и будет множеством всех корней уравнения (8). Но прообраз $f^{-1}(b_1)$ есть множество всех чисел $x \in \boldsymbol{R}$, для которых $f(x) = b_1$, т. е. *множество всех корней* уравнения $f(x) = b_1$. То же справедливо для $f^{-1}(b_2)$ и т. д.

В результате мы получаем следующее правило: *для решения уравнения* (8) *надо сначала найти все корни* b_1, b_2, ... *уравнения* $g(x) = 0$, *затем рассмотреть отдельно все уравнения*

$$f(x) = b_1,\ f(x) = b_2,\ \ldots; \tag{9}$$

тогда объединение множеств корней всех получаемых уравнений (9) *и представляет собой множество всех корней уравнения* (8).

Например, для решения биквадратного уравнения

$$x^4 - 7x^2 + 12 = 0, \tag{10}$$

т. е. уравнения (8), в котором $f(x) = x^2$, $g(x) = x^2 - 7x + 12$, надо сначала решить квадратное уравнение $g(x) = 0$, т. е. уравнение $x^2 - 7x + 12 = 0$. Оно имеет два корня $b_1 = 3$, $b_2 = 4$. Затем, согласно сформулированному выше правилу, надо рассмотреть отдельно уравнения (9), т. е. в данном случае уравнения $x^2 = 3$, $x^2 = 4$. Множество корней первого из них есть $\{\sqrt{3}, -\sqrt{3}\}$, а множество корней второго есть $\{2, -2\}$. Следовательно, их объединение $\{-2, -\sqrt{3}, \sqrt{3}, 2\}$ и есть множество всех корней биквадратного уравнения (10).

Смысл рассмотренного приема состоит в том, что мы заменяем левую часть уравнения композицией *более простых* функций, в результате чего исходное сложное уравнение сводится к решению нескольких *более простых*. В школьных задачниках можно найти много уравнений, допускающих применение этого приема.

Существенно заметить, что сложная функция $g(f(x))$ может рассматриваться и для таких функций $f(x)$, $g(x)$, области определения которых не совпадают со всей числовой прямой. Область определения функции $g(f(x))$ состоит в этом случае из всех тех $x \in \boldsymbol{R}$, для которых, во-первых, x принадлежит области определения функции f, и, во-вторых, соответствующее значение $f(x)$ принадлежит области определения функции g. Сформулированное выше правило решения уравнения (8) остается справедливым и в этом случае.

Рассмотрим в качестве примера уравнение

$$\sqrt{x^2 - x - 5} - 1 = 0. \tag{11}$$

Его левую часть можно представить в виде $\sqrt{y} - 1$, где $y = x^2 - x - 5$. Иначе говоря, взяв функции $f(x) = x^2 - x - 5$, $g(x) = \sqrt{x} - 1$, мы как раз запишем уравнение (11) в форме (8). Здесь область определения функции $g(x)$ не совпадает со всей числовой прямой, однако, как мы отмечали выше, это не препятствует применению описанного выше приема. Согласно описанию этого приема мы прежде всего должны рассмотреть уравнение $g(x) = 0$, т. е. в данном случае $\sqrt{x} - 1 = 0$. Как легко видеть, это уравнение имеет единственный корень $b_1 = 1$. Следовательно, имеется только одно уравнение (9), а именно уравнение $f(x) = b_1$, т. е.

$$x^2 - x - 5 = 1. \tag{12}$$

Решая его, находим два корня $x_1 = -2$, $x_2 = 3$. Таким образом, множество всех корней исходного уравнения имеет вид $\{-2; 3\}$.

Заметим, что для решения уравнения (11) могут быть применены и другие приемы. Один из них, часто применяемый при решении иррациональных уравнений (т. е. уравнений, содержащих знаки радикалов), состоит в том, что квадратный радикал надо «уединить», а затем возвести обе части уравнения в квадрат. Иначе говоря, надо переписать уравнение (11) в виде $\sqrt{x^2 - x - 5} = 1$, а затем от него перейти «возведением в квадрат» к уравнению $x^2 - x - 5 = 1$.

Как мы видим, получается то же уравнение (12), что и прежде, и это дает те же корни $x_1 = -2$, $x_2 = 3$. Однако возведение обеих частей уравнения в квадрат может привести к появлению так называемых *посторонних корней*, в связи с чем такой способ решения требует *проверки* каждого найденного корня (что, правда, в данном случае несложно). Прием же, рассмотренный ранее (связанный с рассмотрением сложной функции), *никогда* не приводит ни к потере корней, ни к появлению «посторонних» корней, и потому является более предпочтительным (в тех случаях, когда его удается применить).

Еще один прием решения уравнений заключается в использовании так называемой «области допустимых значений переменной» (сокращенно ОДЗ). Под этим понимается множество всех тех $x \in \boldsymbol{R}$, для которых все обозначенные действия выполнимы. В случае уравнения (11) это означает рассмотрение множества всех $x \in \boldsymbol{R}$, для которых выражение под знаком квадратного радикала *неотрицательно*:

$$x^2 - x - 5 \geq 0.$$

Сразу же бросается в глаза, что это неравенство существенно сложнее уравнения (12) (например, квадратный трехчлен в левой части имеет корни $\frac{1}{2}(1 \pm \sqrt{21})$); к тому же это неравенство вовсе не помогает решению уравнения (11).

Термин ОДЗ, употребляемый только в школе, применяется во многих пособиях для учащихся и даже в школьных учебниках, но его надо признать неудачным. Ниже, в беседах, посвященных математическим рассуждениям, мы остановимся на этом подробнее, здесь же отметим лишь терминологические соображения.

Выполнимость обозначенных действий для данного $x \in \boldsymbol{R}$ означает, что x принадлежит *области определения* числовой функции (или пересечению областей определения — например, в случае уравнения типа $f(x) = g(x)$). Уже по этой причине использование (только в школе!) термина ОДЗ вместо общепринятого математического термина *область определения* означает ненужное противопоставление школьного математического преподавания традициям науки. Кроме того, при рассмотрении функции $y = f(x)$ слово *значение* употребляется по отношению к y, т. е. это *образ* точки x при отображении f, это значение откладывается по оси ординат. Применение же термина «допустимое значение» для точки x, принадлежащей области определения (и потому изображаемой на оси абсцисс), означает лишь создание в психологическом плане условий для путаницы понятий.

Ведь когда мы называем равенство $(x+y)^2 = x^2 + 2xy + y^2$ *тождеством*, мы тем самым выражаем наше *отношение* к этому равенству, имея в виду, что оно справедливо для любых x, y, т. е. речь идет о некоторой теореме. Когда мы называем равенство $f(x) = 0$ «уравнением», это также определяет *наше отношение* к этому равенству, означает постановку *вопроса* о том, каково множество всех $x \in \boldsymbol{R}$, удовлетворяющих этому равенству. И даже если бы термин ОДЗ выражал что-то математически новое, сам термин «область допустимых значений» этого *вопроса* не содержит и вряд ли уместен с языковой и психологической точек зрения. А иногда пишут еще «область допустимых значений уравнения» ...

В заключение скажем несколько слов о связи обратных отображений с понятием композиции отображений.

Пусть $f\colon \boldsymbol{A} \to \boldsymbol{B}$ — некоторое взаимно однозначное отображение и $f^{-1}\colon \boldsymbol{B} \to \boldsymbol{A}$ — соответствующее обратное отображение. Тогда для любого $a \in \boldsymbol{A}$ соотношения $f(a) = b$ и $f^{-1}(b) = a$ равносильны — при отображении f элемент a переводится в b, стало быть, b получается из a. Следовательно, композиция $f^{-1} \circ f$ переводит элемент a снова в тот же элемент a, т. е. $f^{-1}(f(a)) = f^{-1}(b) = a$. Иными словами, $f^{-1} \circ f$ есть *тождественное отображение* множества $\boldsymbol{A}$, т. е. оно переводит каждый элемент $a \in \boldsymbol{A}$ в себя. Это выражают записью $f^{-1} \circ f = 1_{\boldsymbol{A}}$. Аналогично отображение $f \circ f^{-1}\colon \boldsymbol{B} \to \boldsymbol{B}$ представляет собой тождественное отображение множества $\boldsymbol{B}$, т. е. $f \circ f^{-1} = 1_{\boldsymbol{B}}$.

Итак, для взаимно однозначного отображения $f\colon \boldsymbol{A} \to \boldsymbol{B}$ и его обратного отображения $f^{-1}\colon \boldsymbol{B} \to \boldsymbol{A}$ справедливы соотношения

$$f^{-1} \circ f = 1_{\boldsymbol{A}}, \quad f \circ f^{-1} = 1_{\boldsymbol{B}}.$$

Верно и обратное: если отображения $f\colon \boldsymbol{A} \to \boldsymbol{B}$ и $g\colon \boldsymbol{B} \to \boldsymbol{A}$ удовлетворяют соотношениям

$$g \circ f = 1_{\boldsymbol{A}}, \quad f \circ g = 1_{\boldsymbol{B}}, \tag{13}$$

то отображения f, g взаимно однозначны и являются взаимно обратными, т. е. $g = f^{-1}$.

Заметим, что выполнение только *одного* из соотношений (13) не достаточно для того, чтобы f и g были взаимно обратными биективными отображениями. Так, отображения, показанные на рис. 135, 136, удовлетворяют первому из соотношений (13), однако взаимно однозначными не являются (первое из них — вложение, второе — наложение).

Рассмотрим в качестве примера функцию $f\colon \boldsymbol{R} \to \boldsymbol{R}$, заданную формулой $f(x) = 2x - 4$ (см. (3)), и обратную к ней функцию

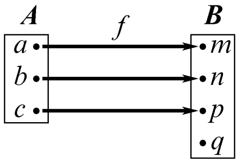


Рис. 135

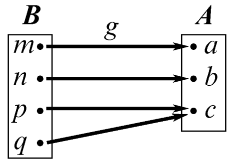


Рис. 136

$$g(x) = f^{-1}(x) = \frac{1}{2}\,x + 2.$$

Тогда $g(f(x)) = \frac{1}{2}f(x) + 2 = \frac{1}{2}\,(2x - 4) + 2 = x$ для любого $x \in \boldsymbol{R}$, и потому $g \circ f = 1_{\boldsymbol{R}}$. Аналогично проверяется, что $f \circ g = 1_{\boldsymbol{R}}$.

Для функции (4) и обратной к ней функции $g = f^{-1}$ (рис. 126) находим для любого $x \in [0; \infty)$:

$$g(f(x)) = (f(x))^2 = (\sqrt{x})^2 = x.$$

т. е. $g(f(x)) = x$ для любого $x \in [0; \infty)$, и потому $g \circ f = 1_{[0;\,\infty)}$.

Точно так же для любого $x \in [0; \infty)$ получаем

$$f(g(x)) = f(x^2) = \sqrt{x^2} = |x| = x$$

(поскольку $x \geq 0$), т. е. $f(g(x)) = x$, и потому $f \circ g = 1_{[0;\,\infty)}$.

Задачи и упражнения

56. Докажите, что композиция симметрий относительно двух заданных параллельных прямых есть параллельный перенос перпендикулярно этим прямым на расстояние, вдвое большее расстояния между этими прямыми.

57. Опишите преобразование плоскости, являющееся композицией поворота f на угол α вокруг точки O и поворота g на угол $-\alpha$ вокруг точки O'.

58. Заданы функции $f(x) = \Theta(x)$ (см. упр. 43) и $g(x) = x^2 - 1$. Постройте графики функций $y = f(g(x))$ и $y = g(f(x))$.

59. Найдите все корни уравнения $x^4 - 2x^3 - 4x^2 + 5x + 6 = 0$.

60. Какое количество различных корней может иметь уравнение $x^2 + p|x| + q = 0$ в зависимости от значений p и q?

19. Классификация

Идея классификации хорошо известна; по своему существу она связана с множествами и отображениями. Волк, лошадь, горилла, дельфин принадлежат к классу *млекопитающих*; орел, голубь, гусь — классу *птиц*; крокодил, черепаха, ящерица — классу *пресмыкающихся*; щука, карась, осетр — классу *рыб*; амеба, инфузория «туфелька», вибрион холеры — классу *простейших*. Обозначим через $\boldsymbol{A}$ множество всех биологических видов животных, а через $\boldsymbol{B}$ = {млекопитающие, птицы, рыбы, пресмыкающиеся, простейшие} — множество всех классов. Тогда каждому элементу множества $\boldsymbol{A}$, т. е. виду животных, естественно сопоставляется некоторый элемент множества $\boldsymbol{B}$ — тот класс, которому этот вид принадлежит. Получающееся отображение $e\colon \boldsymbol{A} \to \boldsymbol{B}$ является наложением (поскольку ни один класс не является пустым множеством); оно называется *естественным отображением*. Каждый элемент $a \in \boldsymbol{A}$ является *представителем* того класса, которому он принадлежит. Например, щука — представитель класса рыб. Понятие *класса* возникает в результате абстракции отождествления. Все виды животных, вскармливающих своих детенышей молоком, объединены в одно подмножество множества $\boldsymbol{A}$. Абстрагируясь от индивидуальных черт каждого вида (хищник это или травоядное, парнокопытное или нет, и т. п.), мы объединяем всех этих представителей животного мира по основному признаку, как бы отождествляем их в этом отношении, и это дает один из классов. Название «млекопитающие» отражает основной признак. (Правда, есть исключение: пингвин вскармливает своих детенышей молоком, однако он отнесен — по другим признакам — к классу птиц.) Два представителя, отнесенные к одному классу, называются (в смысле этой классификации) *эквивалентными*. Иначе говоря, эквивалентность означает принадлежность одному и тому же классу.

Вообще, идея классификации состоит в том, что, имея заданное множество $\boldsymbol{A}$, мы объединяем его элементы в классы по определенному *признаку эквивалентности*, причем так, что получающиеся классы попарно не пересекаются и охватывают *все* элементы множества $\boldsymbol{A}$. Иначе говоря, каждый элемент попадает в один и только один из классов, т. е. является *представителем* какого-то одного класса.

Какими же свойствами должен обладать признак эквивалентности (или, как говорят математики, *отношение эквивалентности*), чтобы с его помощью можно было осуществить классификацию, т. е. разбить множество $\boldsymbol{A}$ на классы, объединение которых дает все множество $\boldsymbol{A}$ и которые попарно не пересекаются? Иначе говоря, каким должно быть отношение эквивалентности, чтобы, объединяя вместе эквивалентные элементы, мы получили разбиение множества $\boldsymbol{A}$ на классы?

Условимся для краткости считать, что запись $a \sim b$ означает: «элемент a эквивалентен элементу b». Далее, если a — элемент множества $\boldsymbol{A}$, то через $[a]$ условимся обозначать *класс элемента* a, т. е. множество всех тех $x \in \boldsymbol{A}$, для которых $x \sim a$:

$$[a] = \{x \in \boldsymbol{A}\colon x \sim a\}.$$

Чтобы выяснить, какими свойствами должно обладать отношение эквивалентности, рассмотрим некоторые примеры.

На рис. 137 изображены ученики a, b, c, d, e, f, выстроившиеся по росту. Рассмотрим отношение «не ниже» и будем его обозначать символом $\sim$. Например, $a \sim b$, $b \sim c$, $d \sim f$, но отношение $f \sim a$ не соблюдается (т. е. неверно, что «f не ниже a»). Посмотрим, дает ли это отношение разбиение на классы. Класс элемента b содержит, кроме самого b, еще a и c, поскольку каждый из учеников a и с *не* *ниже* b, т. е. $a \sim b$, $c \sim b$. Таким образом, должно быть $[b] = \{a, b, c\}$. Аналогично $[d] = \{a, b, c, d, e\}$. Это означает, что классы $[b]$ и $[d]$ *различны*, но имеют общие элементы. Однако это противоречит идее классификации, поскольку различные классы должны быть непересекающимися (т. е. два класса должны либо совпадать, либо не иметь общих элементов). Это произошло потому, что рассмотренное отношение «не ниже» не обладает свойством *симметричности*, т. е. из $a \sim b$ не вытекает $b \sim a$ (или иначе, $a \sim b$ и $b \sim a$ означают не одно и то же). Этот пример показывает, что для осуществления классификации отношение эквивалентности $\sim$ должно быть симметричным:

$$\text{если } a \sim b, \text{ то } b \sim a. \tag{14}$$

Однако одного свойства симметричности не достаточно. Чтобы убедиться в этом, рассмотрим еще один пример.

Пусть $\boldsymbol{A}$ — некоторое множество людей, и для двух элементов a, b этого множества $a \sim b$ означает «a и b знакомы» (под этим мы будем понимать, что каждый из них знает другого). На рис. 138 изображено некоторое множество $\boldsymbol{A}$ людей. Наличие отрезка, соединяющего двух из них, означает, что они знакомы; например, $a \sim b$, $b \sim c$, $b \sim d$.

Отношение знакомства $\sim$ обладает свойством симметричности: «x знаком с y» и «y знаком с x» означает одно и то же (т. е. означает: «x и y знакомы»). Из рассмотрения рис. 138 видно, что $[a] = \{a, b, f\}$, $[b] = \{a, b, c, d\}$ (мы, естественно, считаем, что каждый человек «знаком с самим собой», т. е. $a \sim a$). Опять классификация не получается: классы $[a]$ и $[b]$ имеют общие элементы a, b, но не совпадают. Происходит это потому, что для данного отношения не выполнено свойство *транзитивности*:

$$\text{если } a \sim b \text{ и } b \sim c, \text{ то } a \sim c. \tag{15}$$

В самом деле, на рис. 138 a и b знакомы, b и c тоже знакомы, но a и c не знакомы.

Итак, чтобы послужить основой для классификации, отношение эквивалентности должно быть симметричным и транзитивным.

Заметим еще, что при рассмотрении последнего примера мы нашли «естественным» выполнение свойства *рефлексивности*:

$$a \sim a \tag{16}$$

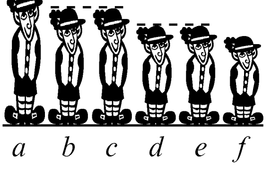


Рис. 137

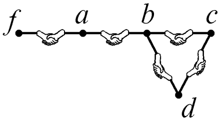


Рис. 138

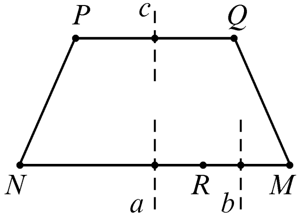


Рис. 139

(каждый человек знаком с самим собой). Однако, несмотря на его «естественность», свойство рефлексивности должно быть четко указано как одно из свойств отношения эквивалентности. Рассмотрим пример, показывающий важность этого свойства. Для этого возьмем множество $\boldsymbol{A}$ всех прямых, на плоскости и записью $a \sim b$ (для двух элементов из $\boldsymbol{A}$) будем обозначать отношение «прямая a параллельна b». Это свойство симметрично (если одна прямая параллельна другой, то и вторая параллельна первой) и вроде бы транзитивно (две прямые, параллельные третьей, параллельны друг другу). Рефлексивно ли это отношение? Это зависит от определения параллельности прямых — если прямая «сама себе параллельна», то рефлексивность есть.

В старом учебнике геометрии А. П. Киселева параллельными назывались прямые, лежащие в одной плоскости и не имеющие общих точек. Поэтому только две *различные* прямые могли быть параллельными, т. е. две *совпадающие* прямые параллельными не считались. В результате отношение параллельности не было рефлексивным. То же было повторено в учебнике А. В. Погорелова (и в ряде других учебников). К чему это приводит? На рис. 139 изображена равнобедренная трапеция $MNPQ$ и точка R на ее основании. Обозначим через a серединный перпендикуляр к отрезку MN, через b — серединный перпендикуляр к отрезку MR, а через c — серединный перпендикуляр к отрезку PQ. Тогда $a \parallel b$ и $b \parallel c$, т. е., обозначая временно это отношение знаком ~, мы находим $a \sim b$, $b \sim c$. Однако прямые a и c совпадают, т. е. не параллельны, согласно терминологии Киселева и Погорелова. Выходит, что утверждение «две прямые, параллельные третьей, параллельны» у них должно быть сформулировано в виде: «две прямые, параллельные третьей, либо параллельны, либо совпадают». Подобных сложностей, связанных с нерефлексивностью принятого определения отношения параллельности, в учебнике А. В. Погорелова немало. Без рефлексивности отношение параллельности не устанавливает классификации прямых: класс $[a]$ прямой a состоит из всех прямых, параллельных a, но не содержит самой прямой a (по Киселеву и Погорелову), т. е. $a \notin [a]$.

В учебнике по курсу А. Н. Колмогорова совпадающие прямые считаются параллельными, т. е. свойство рефлексивности соблюдается и получается полная классификация: каждый класс содержит все параллельные между собой прямые, без всяких исключений.

Итак, для того, чтобы быть основой для классификации, отношение эквивалентности должно быть рефлексивным, симметричным и транзитивным, т. е. для любых трех элементов a, b, c рассматриваемого множества (которое мы хотим разбить на классы с помощью этого отношения эквивалентности) *должны* выполняться свойства (16), (14), (15).

Замечательно, что и обратное утверждение справедливо, т. е. если отношение эквивалентности, введенное в множестве $\boldsymbol{A}$, рефлексивно, симметрично и транзитивно, то оно осуществляет в $\boldsymbol{A}$ классификацию. Иначе говоря, рассматривая для каждого $a \in \boldsymbol{A}$ его класс эквивалентности $[a]$, мы получаем разбиение множества $\boldsymbol{A}$ на классы, каждые два из которых либо совпадают, либо не пересекаются (как и требует классификация).

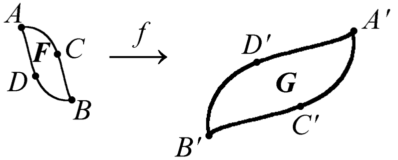


Рис. 140

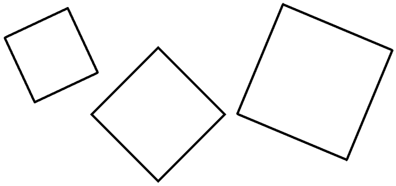


Рис. 141

Отношение параллельности (рефлексивное, как у Колмогорова) является примером такого отношения эквивалентности, т. е. оно разбивает все прямые на классы попарно параллельных прямых. Вот еще примеры.

Будем называть фигуры на плоскости эквивалентными, если они имеют одинаковую площадь. Легко видеть, что это отношение рефлексивно, симметрично и транзитивно (Евклид говорил: «порознь равные третьему, равны между собой» — в данном случае «равные» в смысле площади). Значит, это отношение эквивалентности разбивает все плоские фигуры на классы эквивалентности: в один класс попадают все фигуры одинаковой площади (или, как иногда говорят, *равновеликие* фигуры).

В качестве следующего примера отметим *подобие* фигур. Будем писать $\boldsymbol{F} \sim \boldsymbol{G}$, если фигура $\boldsymbol{F}$ подобна фигуре $\boldsymbol{G}$, т. е. существует такое взаимно однозначное соответствие $\boldsymbol{F} \to \boldsymbol{G}$ (рис. 140), при котором расстояния между парами соответствующих точек имеют одно и то же отношение (называемое коэффициентом подобия):

$$\frac{A'B'}{AB} = \frac{A'C'}{AC} = \frac{B'D'}{BD} = \ldots = k.$$

Отношение подобия $\boldsymbol{F} \sim \boldsymbol{G}$ рефлексивно, симметрично и транзитивно, т. е. является отношением эквивалентности, разбивающим все фигуры на классы подобных фигур. Например, все квадраты образуют один класс эквивалентности (т. е. все квадраты подобны между собой, и любая фигура, подобная квадрату, также является квадратом, рис. 141). Ромбы же входят в один класс в том и только в том случае, если они имеют равные углы (рис. 142). Все окружности также образуют один класс (они все подобны, рис. 143).

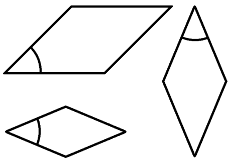


Рис. 142

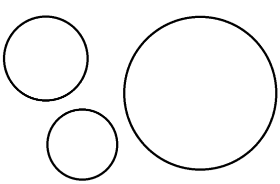


Рис. 143

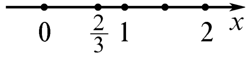


Рис. 144

Наконец, отметим еще два примера, имеющие существенное значение. Первый из них касается *рациональных чисел*, т. е. чисел, представимых в виде дроби $\frac{p}{q}$, где p и q — целые, причем $q \neq 0$. Две дроби $\frac{p}{q}$ и $\frac{r}{s}$ считаются эквивалентными, если $ps = qr$. Например, эквивалентными являются дроби $\frac{2}{3}$ и $\frac{4}{6}$, поскольку $2 \cdot 6 = 3 \cdot 4$. Это отношение эквивалентности является рефлексивным, симметричным и транзитивным, т. е. множество всех дробей разбивается на классы эквивалентности. Каждый такой класс эквивалентности и есть рациональное число.

Дроби $\frac{2}{3}$ и $\frac{4}{6}$ представляют *одно и то же* рациональное число, поскольку эти дроби эквивалентны. Именно рациональные числа, а не отдельные дроби изображаются точками на числовой оси. Ведь нет отдельных точек для $\frac{2}{3}$, $\frac{4}{6}$ или $\frac{10}{15}$. Эти дроби (и все другие, эквивалентные им) изображаются одной точкой (рис. 144). Надо было бы писать $\left[\frac{2}{3}\right] = \left[\frac{4}{6}\right]$, т. е. *класс*, содержащий дробь $\frac{2}{3}$, *совпадает с классом*, содержащим дробь $\frac{4}{6}$ (и именно этот класс, т. е. рациональное число, изображается точкой на числовой оси). Однако для простоты мы опускаем квадратные скобки, т. е. пишем $\frac{2}{3} = \frac{4}{6}$ (вместо более четкой записи $\frac{2}{3} \sim \frac{4}{6}$).

Заметим еще, что сложение (а также умножение) рациональных чисел выполняется *по представителям*. Пусть, например, $r_1 = \left[\frac{2}{3}\right]$, $r_2 = \left[\frac{1}{5}\right]$, т. е. r_1 — рациональное число, представителем которого является дробь $\frac{2}{3}$, а r_2 — число, представляемое дробью $\frac{1}{5}$. По правилу сложения дробей $\frac{2}{3} + \frac{1}{5} = \frac{2 \cdot 5 + 1 \cdot 3}{15} = \frac{13}{15}$ и мы принимаем за *сумму* $r_1 + r_2$ рассматриваемых рациональных чисел класс получающейся дроби $\frac{13}{15}$, т. е. $r_1 + r_2 = \left[\frac{13}{15}\right]$.

Если бы мы взяли других представителей чисел r_1 и r_2, например, $\frac{4}{6}$ и $\frac{3}{15}$, то, согласно правилу сложения дробей, мы нашли бы (взяв

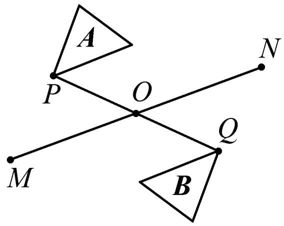


Рис. 145

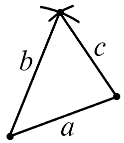


Рис. 146

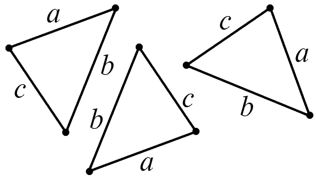


Рис. 147

30 в качестве общего знаменателя): $\frac{4}{6}+\frac{3}{15}=\frac{20}{30}+\frac{6}{30}=\frac{26}{30}$. Но $\frac{26}{30}$ и $\frac{13}{15}$ — эквивалентные дроби, т. е. они представляют одно и то же рациональное число $\left[\frac{13}{15}\right]=\left[\frac{26}{30}\right]$. Иначе говоря, беря различных представителей чисел r_1 и r_2 и выполняя сложение этих дробей (представителей), мы получаем дроби, представляющие *одно и то же* рациональное число, которое и принимается за сумму $r_1 + r_2$.

Таким образом, хотя сложение рациональных чисел определяется по представителям, но оно определено *математически корректно*, т. е. результат (сумма $r_1 + r_2$) не зависит от выбора представителей чисел r_1 и r_2. В связи с этим мы обычно опускаем квадратные скобки; надо было бы писать $\left[\frac{2}{3}\right]+\left[\frac{1}{5}\right]=\left[\frac{13}{15}\right]$, но для краткости мы пишем проще, т. е. без скобок: $\frac{2}{3}+\frac{1}{5}=\frac{13}{15}$.

Заключительный пример — понятие *конгруэнтности* геометрических фигур. Каждая фигура (например, на плоскости) представляет собой некоторое *множество точек*. Как всегда, два множества $\boldsymbol{A}$ и $\boldsymbol{B}$ считаются *равными*, если они совпадают, т. е. любая точка множества $\boldsymbol{A}$ принадлежит также множеству $\boldsymbol{B}$ и обратно. Точно так же две геометрические фигуры $\boldsymbol{A}$, $\boldsymbol{B}$ следует считать *равными*, если они совпадают, т. е. если $\boldsymbol{A}$ и $\boldsymbol{B}$ — это *одна и та же* фигура. Фигуры же, которые могут быть совмещены друг с другом при помощи некоторого движения, называются по принятой в математике традиции не равными, а *конгруэнтными*. Например, на рис. 145 фигуры $\boldsymbol{A}$ и $\boldsymbol{B}$ конгруэнтны: симметрия s относительно точки O переводит фигуру $\boldsymbol{A}$ в фигуру $\boldsymbol{B}$, т. е. $s(\boldsymbol{A})=\boldsymbol{B}$. Но $\boldsymbol{A}$ и $\boldsymbol{B}$ — это все же *различные* фигуры, т. е. на рис. 145 имеются *две* фигуры $\boldsymbol{A}$ и $\boldsymbol{B}$, а вовсе не одна и та же фигура; «равными» (т. е. совпадающими) считать эти две фигуры было бы неправильным. Конгруэнтность фигур $\boldsymbol{A}$ и $\boldsymbol{B}$ обозначается обычно записью $\boldsymbol{A}\cong\boldsymbol{B}$. Примем и мы это обозначение. Таким образом, на рис. 145

$$\boldsymbol{A} \cong \boldsymbol{B},\ s(\boldsymbol{A}) = \boldsymbol{B},\ s(\boldsymbol{B}) = \boldsymbol{A},$$

т. е. фигуры $\boldsymbol{A}$ и $\boldsymbol{B}$ конгруэнтны, а образ фигуры $\boldsymbol{A}$ при симметрии s равен фигуре $\boldsymbol{B}$, т. е. совпадает с ней (и, точно так же, образ фигуры $\boldsymbol{B}$ совпадает с $\boldsymbol{A}$).

В учебнике А. В. Погорелова (вслед за более ранним учебником А. П. Киселева) понятия равенства и конгруэнтности не различаются, во всех случаях говорится о «равных» фигурах. Иными словами, «равными» называются (по Погорелову) не только две совпадающие фигуры, но и две различные фигуры, если они могут быть совмещены движением («наложением» по Киселеву). Например, любые две точки «равны»; любые две прямые «равны». При погореловской терминологии использование общепринятых функциональных обозначений (не только удобных, но и сближающих алгебру с геометрией) становится математически некорректным и бессмысленным. Так, применяя к рисунку 145 функциональные обозначения, мы можем написать $s(M) = N$, $s(P) = Q$, т. е. точка $s(M)$ (образ точки M) *равна* N, совпадает с ней; точка $s(P)$ равна точке Q, т. е. $s(P)$ и Q — это *одна и та же точка*. Так обстоит дело в общепринятой математической терминологии. Но по Погорелову любые две точки «равны» (например, $M = N = P = Q$), и потому функциональная запись $s(M) = N$ *ничего не выражает*, поскольку $s(M)$ есть точка, N — тоже точка и потому они, конечно, «равны». С такой же обоснованностью можно было бы написать $s(M) = P$, $s(M) = Q$ и т. п. (ведь любые же две точки «равны»!). Таким образом, упрощенчество, связанное с отказом от термина *конгруэнтность*, перечеркивает идею функциональных обозначений. Мы далее сохраняем общепринятый математический термин конгруэнтность и обозначение $\boldsymbol{A} \cong \boldsymbol{B}$, что дает право в полной мере использовать функциональные обозначения.

Теперь укажем, как это связано с идеей классификации. Отношение конгруэнтности геометрических фигур рефлексивно, симметрично и транзитивно, т. е. оно разбивает все множество фигур на классы. Мы обычно говорим, что задача о построении треугольника по длинам трех сторон a, b, c имеет *единственное* решение (при условии, конечно, что каждое из чисел a, b, c меньше суммы двух других). Это построение легко осуществить циркулем и линейкой (рис. 146). Но ведь в действительности имеется *бесконечно много* треугольников с данными длинами сторон (рис. 147), а вовсе не «один» такой треугольник. Однако все треугольники, дающие решение этой задачи, конгруэнтны между собой, т. е. составляют *один класс* конгруэнтных фигур.

Правильнее было бы сказать, что *с точностью до конгруэнтности* задача имеет единственное решение, т. е. решением является один *класс* конгруэнтных треугольников. Задача же *построить треугольник по двум сторонам и высоте, сходящимся в одной вершине*, имеет два решения (рис. 148). Точнее, два решения *с точностью до конгруэнтности*, т. е. имеются *два класса* конгруэнтных треугольников, служащих решениями этой задачи.

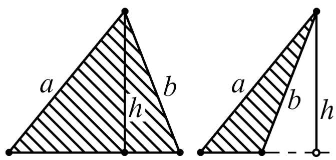


Рис. 148

Задачи и упражнения

61. Может ли служить отношением эквивалентности целых чисел наличие у них общего делителя, большего 1?

62. Покажите, что во множестве целых чисел, больших единицы, следующее отношение является отношением эквивалентности: наибольший из делителей числа, не равный самому числу, один и тот же. Опишите класс чисел, у которых этот делитель равен 7, и класс чисел, у которых он равен 12.

63. Является ли классификацией следующее разбиение четырехугольников: трапеции, параллелограммы, ромбы, остальные четырехугольники?

64. Является ли отношением эквивалентности во множестве треугольников равенство радиусов вписанных окружностей? А у четырехугольников?

65. Два города назовем эквивалентными, если количество жителей одного отличается от количества жителей другого не более, чем на 5000 человек. Является это отношение отношением эквивалентности?

Беседа 5. Упорядоченные множества

20. Понятие упорядоченного множества

Прежде чем ввести отношение порядка, которому посвящен этот пункт, мы снова рассмотрим два примера, хорошо нам известных. В качестве первого примера возьмем множество $\boldsymbol{R}$ всех действительных чисел. В этом множестве введены *неравенства*. Мы говорим, что число *a меньше b* и пишем $a < b$ (или $b > a$, *b больше a*), если разность $b - a$ есть положительное число. Графически неравенство $a < b$ означает, что точка b расположена на числовой оси правее точки a (рис. 149). Неравенства обладают свойством транзитивности (рис. 150):

$$\text{если } a < b \text{ и } b < c, \text{ то } a < c. \qquad (1)$$

Иногда рассматриваются также нестрогие неравенства $a \le b$ (или $b \ge a$). Запись $a \le b$ (словами: a меньше или равно b) означает, что выполнено какое-либо одно из условий $a < b$, $a = b$. Например, неравенство $5 \le 7$ верно (одно из условий $5 < 7$, $5 = 7$ верно, а именно первое из них); неравенство $4 \le 4$ также верно (одно из условий $4 < 4$, $4 = 4$ верно, а именно второе из них). Иными словами, запись $a \le b$ означает, что число a не больше b, т. е. a либо совпадает с b, либо расположено на числовой оси левее b. Так, числовой отрезок $[a; b]$ состоит из всех действительных чисел x, удовлетворяющих *двойному* нестрогому неравенству $a \le x \le b$. Таким образом, все числа x, расположенные *между* a и b (рис. 151), принадлежат этому отрезку; концевые точки a и b также ему принадлежат.

В качестве второго примера рассмотрим некоторое универсальное множество $\boldsymbol{U}$ и обозначим через $\boldsymbol{S}$ множество всех его подмножеств.

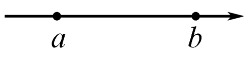


Рис. 149

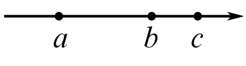


Рис. 150

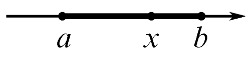


Рис. 151

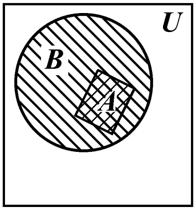


Рис. 152

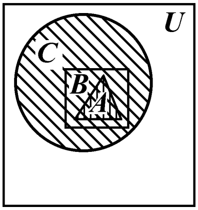


Рис. 153

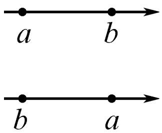


Рис. 154

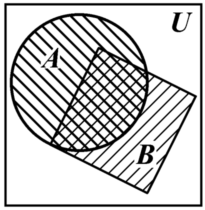


Рис. 155

Для двух элементов $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{S}$ (т. е. подмножеств множества $\boldsymbol{U}$) условимся писать $\boldsymbol{A} < \boldsymbol{B}$, если $\boldsymbol{A}$ — собственное подмножество множества $\boldsymbol{B}$ (рис. 152). Нестрогое неравенство $\boldsymbol{A} \leq \boldsymbol{B}$ означает, что либо $\boldsymbol{A} < \boldsymbol{B}$, либо $\boldsymbol{A} = \boldsymbol{B}$, т. е. $\boldsymbol{A}$ либо есть собственное подмножество множества $\boldsymbol{B}$, либо же $\boldsymbol{A}$ совпадает с $\boldsymbol{B}$. Иначе говоря, нестрогое неравенство $\leq$ равносильно включению $\subset$. Введенные таким образом в множестве $\boldsymbol{S}$ неравенства обладают свойством транзитивности (1) (рис. 153).

В обоих рассмотренных примерах (в множестве $\boldsymbol{R}$ и в множестве $\boldsymbol{S}$) введены неравенства между элементами, причем свойство транзитивности выполняется в обоих случаях. Однако между этими примерами имеется и существенное различие. В $\boldsymbol{R}$ для любых двух элементов a, b (т. е. двух действительных чисел) обязательно справедливо какое-либо одно из условий

$$a < b, \quad a = b, \quad a > b, \tag{2}$$

т. е. точка a (если она не совпадает с b) расположена либо слева, либо справа от b (рис. 154). В множестве же $\boldsymbol{S}$ имеются *несравнимые* элементы, т. е. такие подмножества $\boldsymbol{A}$, $\boldsymbol{B}$ универсального множества $\boldsymbol{U}$, что ни одно из них не является подмножеством другого (рис. 155), т. е. *ни одно* из соотношений (2) места не имеет. Таким образом, в обоих множествах $\boldsymbol{R}$, $\boldsymbol{S}$ выполнено *условие исключительности*: для любых двух элементов a, b имеет место не более чем одно из соотношений (2), т. е. эти соотношения исключают друг друга.

Теперь будут понятны следующие определения. Множество $\boldsymbol{M}$, в котором по некоторому правилу введены неравенства, т. е. указано, для каких элементов $a, b \in \boldsymbol{M}$ справедливо соотношение $a < b$ (и тогда $b > a$), называется *упорядоченным множеством*, если в нем выполнено условие транзитивности (1) и условие исключительности, т. е. для любых a, b выполнено не более чем одно из соотношений (2). Само отношение $a < b$ (удовлетворяющее условиям транзитивности и исключительности) называется *отношением порядка*. В общем случае в упорядоченном множестве могут существовать *несравнимые элементы* a, b, т. е. такие, для которых ни одно из соотношений (2) места не имеет. Таковым было упорядоченное множество $\boldsymbol{S}$. Если же для любых двух $a, b \in \boldsymbol{M}$ выполнено одно из соотношений $a < b$, $a = b$, $a > b$, т. е. несравнимых элементов нет, то $\boldsymbol{M}$ называется *цепью*. Таковым было упорядоченное множество $\boldsymbol{R}$.

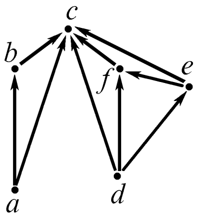


Рис. 156

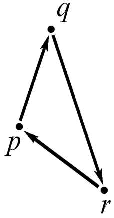


Рис. 157

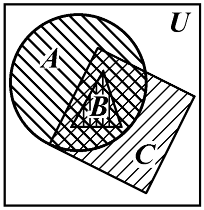


Рис. 158

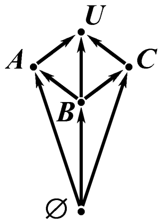


Рис. 159

На рис. 156 упорядоченное множество $\boldsymbol{T}$, содержащее элементы a, b, c, d, e, f, показано в виде *направленного графа*: каждая стрелка идет от меньшего элемента к большему (для наглядности стрелки направлены вверх). Таким образом, если $a < b$, $b < c$ и, следовательно, $a < c$ (в соответствии со свойством транзитивности (1)), то в треугольнике с вершинами a, b, c вектор $\overrightarrow{ac}$ равен сумме векторов, идущих по двум другим сторонам ($\overrightarrow{ac} = \overrightarrow{ab} + \overrightarrow{bc}$). То же справедливо для треугольника def и других аналогичных треугольников. Однако такого треугольника pqr, в котором стрелки идут, скажем, в направлении обхода по часовой стрелке (рис. 157), быть не может, поскольку это означало бы выполнение неравенства $p < q$, $q < r$, $r < p$ — вопреки свойству транзитивности.

Заметим, что в упорядоченном множестве $\boldsymbol{T}$ (рис. 156) имеется один *максимальный элемент* c (т. е. не существует в $\boldsymbol{T}$ элемента, большего c) и два *минимальных элемента* a, d (т. е. не существует элемента, меньшего a, и также не существует элемента, меньшего d). В упорядоченном множестве $\boldsymbol{S}$ имеется один максимальный элемент $\boldsymbol{U}$ и один минимальный элемент $\varnothing$, т. е. для любого подмножества $\boldsymbol{X}$ множества $\boldsymbol{U}$ справедливы неравенства $\boldsymbol{X} \le \boldsymbol{U}$ и $\boldsymbol{X} \ge \varnothing$. Для наглядности на диаграмме Эйлера–Венна (рис. 158) показаны три подмножества $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ множества $\boldsymbol{U}$, а на рис. 159 — соответствующий направленный граф. А в $\boldsymbol{R}$ нет ни максимальных, ни минимальных элементов. В самом деле, какое бы число $a \in \boldsymbol{R}$ мы ни взяли, всегда в $\boldsymbol{R}$ найдется *большее* число x, т. е. число a не является максимальным.

Задачи и упражнения

66. Для двух человек A и B запишем $A > B$, если B является потомком A. Будет ли в этом случае множество людей упорядоченным?

67. В волейбольном турнире каждая команда сыграла с каждой по одной игре (ничьих в волейболе не бывает). Запишем $A > B$, если команда A выиграла у команды B. Будет ли в этом случае множество команд упорядоченным?

68. На конкурсе красоты выступало 5 претенденток на звание «Мисс изящество». Судили конкурс трое судей, каждый из которых расположил претенденток в порядке предпочтения. Для претенденток A и B запишем $A > B$, если большее число судей поставило A выше, чем B. Всегда ли множество претенденток окажется упорядоченным? Обязательно ли в нем найдутся максимальный и минимальный элементы?

69. Для точек $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$ (на координатной плоскости) запишем $M_1 > M_2$ при $x_1 > x_2$, а если $x_1 = x_2$, то при $y_1 > y_2$. Станет ли это множество упорядоченным?

70. Двадцать восьмиклассников выстроились в шеренгу, за ними встала шеренга из двадцати девятиклассников. Оказалось, что каждый девятиклассник выше стоящего перед ним восьмиклассника. Через некоторое время школьники в каждой шеренге перестроились, встав по росту. Докажите, что вновь перед каждым девятиклассником окажется восьмиклассник, меньший его по росту.

21. Минимальные элементы и математическая индукция

В историческом процессе познания мира (а иногда, в естественных науках, и поныне) важную роль играла *экспериментальная индукция*, т. е. заключение о том, что многократно наблюдаемое явление должно (в аналогичных условиях) повторяться и впредь. Тысячелетиями люди замечали, что вечером солнце исчезает за горизонтом на западе, а утром снова появляется на востоке. И это (так же как смена времен года) было *экспериментально установленным фактом* — задолго до того, как была выяснена причина этих явлений, связанная с вращением Земли и ее движением вокруг Солнца. Экспериментальная индукция имеет значение и в математике, но роль ее здесь несколько иная.

Квадратный трехчлен $f(n) = n^2 - n + 41$ при подстановке в него (вместо n) последовательных натуральных чисел 1, 2, 3, ... принимает следующие значения: $f(1) = 41$, $f(2) = 43$, $f(3) = 47$, $f(4) = 53$, $f(5) = 61$, $f(6) = 71$, ... Все эти значения являются *простыми числами*. Если продолжить вычисления, то можно убедиться, что при подстановке вместо n следующих чисел 7, 8, 9, 10 снова получаются простые числа. Можно ли считать «экспериментально установленным», что для любого натурального n значение $f(n)$ является простым числом? Математика отвечает на этот вопрос отрицательно. Какое бы *конечное* число проб мы ни сделали, это не служит доказательством того, что для *любого* натурального n значение $f(n)$ является простым числом. Это — лишь гипотеза, предположение. Но достаточно найти один *контрпример* (т. е. натуральное n, для которого $f(n)$ не будет простым), и высказанное общее утверждение будет ниспровергнуто, т. е. будет показано, что сформулированная гипотеза ошибочна. И такой пример нетрудно найти: $f(41) = 41^2 - 41 + 41 = 41^2$, т. е. число $f(41)$ не является простым. Кстати, это — *минимальное* n, являющееся контрпримером, т. е. для всех $n = 1, 2, 3, \ldots.\ 40$ число $f(n)$ — простое.

Еще более впечатляющим является квадратный трехчлен $g(n) = n^2 - 79n + 1601$. И здесь последовательные значения $g(1) = 1523$, $g(2) = 1447$, $g(3) = 1373$ являются простыми. В этом нетрудно убедиться с помощью микрокалькулятора. Действительно, попробовав делить 1523 на простые числа 2, 3, 5, 7, 11, ..., 37, мы найдем, что на них 1523 не делится; но тогда 1523 не делится и на большие простые числа, поскольку если m не простое, то у него непременно имеется простой множитель, не превосходящий $\sqrt{m}$ (кстати, почему?). Таким образом, 1523 — простое число, и можно убедиться, что дальнейшие значения 1447, 1373 тоже являются простыми. Правда, эти значения убывают, но все же отрицательных значений трехчлен $g(n)$ не имеет, поскольку его дискриминант $79^2 - 4 \cdot 1601 = -163$ отрицателен; в этом также можно убедиться, рассматривая график функции $g(x) = x^2 - 79x + 1601$ (рис. 160), принимающей наименьшее значение при $x = \frac{79}{2}$. Например, значение $g(40) = 41$, так же как $g(39) = 41$, снова представляет собой простое число (это будет *наименьшее* значение, которое принимает $g(n)$ для натуральных n).

И все же гипотеза о том, что для любого натурального n значение $g(n)$ является простым числом, ошибочна. Контрпримером является число 80: $g(80) = 80^2 - 79 \cdot 80 + 1601 = 1681 = 41^2$ (и это — минимальное натуральное n, служащее контрпримером, т. е. для всех $n = 1, 2, 3, ..., 79$ число $g(n)$ — простое).

Наконец, рассмотрим третий, весьма эффектный пример.

Великий математик XVII столетия Пьер Ферма (о котором мы уже говорили выше) высказал предположение, что *для любого натурального n число* $2^{2^n} + 1$ — *простое*. Эта его гипотеза основывалась на том, что при $n = 1, 2, 3, 4$ получающиеся числа $2^{2^1} + 1 = 2^2 + 1 = 5$, $2^{2^2} + 1 = 2^4 + 1 = 17$, $2^{2^3} + 1 = 2^8 + 1 = 257$, $2^{2^4} + 1 = 2^{16} + 1 = 65\,537$ являются, в самом деле, простыми. При подстановке следующего значения $n = 5$ получается число $2^{2^5} + 1 = 2^{32} + 1 = 4\,296\,961\,297$. Проверить «вручную» отсутствие у него простых делителей затруднительно. Однако Леонард Эйлер обнаружил изящным рассуждением, что это число *не простое*: оно делится на 641. Тем самым получился контрпример к гипотезе Ферма.

Вот как это было установлено Эйлером. Прежде всего, он замечает, что $2^{32} + 1 = 2^{28} \cdot 16 + 1 = 2^{28}(641 - 625) + 1$, т. е.

$$2^{32} + 1 = 2^{28} \cdot 641 - (2^{28} \cdot 625 - 1). \tag{3}$$

Стоящее в скобках число Эйлер записывает в виде

$$2^{28} \cdot 625 - 1 = 2^{28} \cdot 5^4 - 1 = (2^7 \cdot 5)^4 - 1.$$

Наконец, он пользуется тем, что многочлен $x^{2n} - 1$ делится на $x^2 - 1$ и, значит, делится на $x + 1$. В частности, $(2^7 \cdot 5)^4 - 1$ делится на $2^7 \cdot 5 + 1$, т. е. на 641. Таким образом, число $2^{32} + 1$ представлено, согласно (3), в виде разности двух чисел, каждое из которых делится на 641, а потому и число $2^{32} + 1$ делится на 641.

Проследить за этими преобразованиями нетрудно, но остается загадкой, почему именно так действовал Эйлер, какая фантазия и сила интуиции подсказали ему, что $2^{32} + 1$ делится на таинственное число 641 и, следовательно, не является простым.

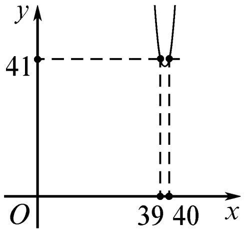


Рис. 160

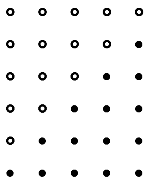


Рис. 161

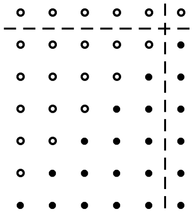


Рис. 162

Итак, если имеется некоторая общая гипотеза («для любого натурального n справедливо то-то и то-то ... »), то ошибочность этой гипотезы устанавливается указанием хотя бы одного контрпримера. Для того же, чтобы подтвердить истинность этой гипотезы, необходимо провести некоторое общее рассуждение (и никакое конечное число подтверждающих примеров не заменяет такого общего доказательства).

Каким же образом может быть проведено математически корректное общее рассуждение, доказывающее, что некоторый факт имеет место для *любого* натурального n? Наиболее употребительным для этого является *метод математической индукции*. Сущность его мы рассмотрим на следующем примере: *для любого натурального n сумма n первых натуральных чисел равна* $\frac{n(n+1)}{2}$, т. е.

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}. \tag{4}$$

Существует красивое геометрическое рассуждение, показывающее правильность этой формулы. Изобразим последовательные вертикальные столбики, содержащие 1, 2, 3, 4, 5 элементов (черные кружки на рис. 161). Затем поместим над ними добавочные вертикальные столбики, содержащие 5, 4, 3, 2, 1 элементов (белые кружки на рис. 161). Количество черных кружков равно $S_5 = 1 + 2 + 3 + 4 + 5$, и столько же имеется белых кружков, т. е. всего на рисунке изображено $2S_5$ кружков. Но ясно, что у нас получилось 5 столбиков, в каждом из которых имеется 5 + 1 кружков, т. е. всего в этой прямоугольной диаграмме содержится $5{\cdot}(5 + 1)$ кружков. Значит, $2S_5 = 5{\cdot}(5 + 1)$, т. е. $S_5 = \frac{5\cdot(5+1)}{2}$. Это означает, что для $n = 5$ формула (4) справедлива. Для $n = 6$ (рис. 162) к прямоугольной (5×6) диаграмме точек для $n = 5$ добавится столбик справа и строка сверху. Получится прямоугольная диаграмма 6×7, поэтому $S_6 = \frac{6\cdot(6+1)}{2}$. И так далее. Но вот это «и так далее» как раз представляет собой суть метода математической индукции.

Покажем теперь, как можно подробно изложить индуктивное рассуждение, доказывающее справедливость формулы (4). Обозначим через T_n утверждение о том, что формула (4) (для рассматриваемого натурального n) верна. Например, на рис. 161 дана иллюстрация справедливости утверждения T_5, а на рис. 162 — иллюстрация утверждения T_6.

Справедливость утверждения T_n для $n = 1$ (или для небольших $n = 2, 3, ...$) непосредственно проверяется. Но как убедиться в справедливости этого утверждения для *любого* натурального n?

Здесь применяется прием рассуждения, называемый *переходом* от n к $n + 1$. Допустим, что формула (4) верна для некоторого натурального n, т. е. что утвержение T_n истинно, и докажем ее справедливость для следующего натурального числа $n + 1$, т. е. установим утверждение T_{n+1}. В самом деле,

$$1 + 2 + ... + n + (n + 1) = (1 + 2 + ... + n) + (n + 1) =$$
$$= \frac{n(n+1)}{2} + (n + 1) = \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2},$$

а это и означает, что утверждение T_{n+1} также истинно. Но теперь «ясно», что утверждение T_n верно для *любого* натурального n. Ведь T_1 верно, а потому рассматриваемое утверждение верно для следующего натурального числа $n = 2$, т. е. T_2 истинно. Но тогда, поскольку T_2 истинно, это утверждение верно и для следующего натурального числа, т. е. T_3 истинно. А так как T_3 истинно, то верно и следующее утверждение T_4 и т. д. Поэтому утверждение T_n верно для любого натурального n.

Это и есть рассуждение, проведенное методом математической индукции. Таким образом, чтобы доказать, что некоторое утверждение T_n справедливо для любого натурального n, мы, во-первых, убеждаемся в справедливости его для $n = 1$, и, во-вторых, осуществляем переход от n к $n + 1$, т. е. доказываем, что если для некоторого n верно T_n, то тогда верно и T_{n+1}. Если это сделано, то утверждение T_n считается доказанным «по индукции».

Следует заметить, что возможность применения метода математической индукции, какой бы «понятной» она ни представлялась, является *аксиомой*, описывающей некоторое свойство натурального ряда $\boldsymbol{N} = \{1, 2, 3, ...\}$. Она означает, что, переходя от натурального числа 1 к следующему числу 2, затем к следующему, затем еще к следующему и т. д., мы *переберем все натуральные числа*.

Важно заметить, что существенны оба элемента математической индукции: и *начало индукции* (т. е. проверка истинности утверждения T_1), и *общий шаг* индукции (т. е. переход от n к $n + 1$: в предположении, что T_n верно, показать, что в таком случае T_{n+1} также верно). Если мы не проведем общего шага индукции, а лишь проверим, что

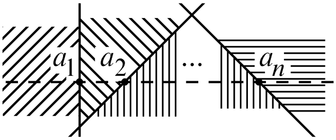


Рис. 163

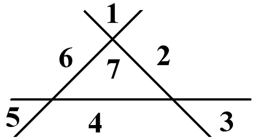


Рис. 164

утверждение T_n верно для нескольких первых значений n, то это означает, что мы ограничиваемся *экспериментальной индукцией*, которая, как мы видели, не дает корректного математического доказательства. Не менее важно и *начало* индукции; пренебрежение им может привести к ошибочным выводам. Мы рассмотрим это на следующем примере.

Задача, которую мы хотим решить, состоит в том, чтобы найти, *на сколько частей разбивают плоскость n прямых, если никакие две из них не параллельны и никакие три не пересекаются в одной точке*.

Обозначим это число частей через P_n, т. е. пусть n прямых (удовлетворяющих указанным условиям) разбивают плоскость на P_n частей. Проведем теперь $(n+1)$-ю прямую (рис. 163). Она пересечет каждую из проведенных ранее прямых (поскольку прямые попарно не параллельны), причем все n точек ее пересечения с проведенными ранее прямыми различны (поскольку никакие три прямые не пересекаются в одной точке). Обозначим эти точки пересечения через a_1, a_2, a_3, ..., a_n. Если мы будем двигаться по $(n+1)$-й прямой «из бесконечности», проходя последовательно точки a_1, a_2, a_3, ..., a_n и вновь «удаляясь в бесконечность», то легко сможем подсчитать, сколько новых частей возникло в результате проведения $(n+1)$-й прямой. В самом деле, подходя к точке a_1, мы заметим, что та из прежних частей, по которой мы двигались, разделилась на две части, т. е. добавилась одна новая часть. Подходя затем к точке a_2, мы обнаружим, что еще одна из прежних частей разделилась на две, т. е. опять появилась одна новая часть и т. д. Таким образом, подойдя к точке a_n, мы обнаружим, что появилось n новых частей. Наконец, продолжая движение по $(n+1)$-й прямой «в бесконечность», мы получим еще одну новую часть. Итак, проведение $(n+1)$-й прямой добавляет $n+1$ новую часть, т. е.

$$P_{n+1} = P_n + (n+1). \tag{5}$$

Соотношение такого типа, т. е. соотношение, выражающее P_{n+1} через предыдущее значение, называют *рекуррентным*.

Но ведь в предыдущем примере, т. е. при доказательстве соотношения (4), мы как раз, осуществляя переход от n к $n+1$, добавляли к сумме $S_n = 1 + 2 + 3 + ... + n$ число $n + 1$. Получается, что и в рассматриваемом примере мы получаем то же значение, т. е. $P_n = \frac{n(n+1)}{2}$. Но верно ли это заключение? На рис. 164 показан случай трех прямых. По вроде бы доказанной формуле мы должны получить $P_3 = \frac{3 \cdot 4}{2} = 6$, т. е. три прямые должны разбить плоскость на 6 частей. Но это неверно: рис. 164 показывает, что три прямые разбивают плоскость на *семь* частей! За счет чего возникла эта ошибка? Именно за счет того, что мы осуществили только переход от n к $n + 1$, пренебрегая *началом индукции*.

Корректное решение выглядит следующим образом. Одна прямая разбивает плоскость на две части, т. е. $P_1 = 2$ (в то время как S_1 было равно 1). Далее, добавив вторую прямую, мы (согласно рекуррентному соотношению (5)) добавим еще две части, т. е. $P_2 = 4$, (тогда как ранее было $S_2 = 1 + 2 = 3$. Вроде бы получается, что P_n всегда на единицу больше, чем S_n, т. е.

$$P_n = \frac{n(n+1)}{2} + 1. \tag{6}$$

Это — наша *гипотеза*, которую надо доказать методом математической индукции. При этом начало индукции уже проведено, т. е. P_1 как раз получается по формуле (6). Остается осуществить переход от n к $n + 1$. Но ясно, что это делается дословно так же, как и в предыдущем примере, поскольку прибавление единицы одновременно к S_n и S_{n+1} ничего не меняет (педантичный читатель может заново провести переход от n к $n + 1$). Теперь математическая индукция проведена полностью и мы получаем окончательный результат: число частей, на которые n прямых разбивают плоскость (при указанных условиях их расположения), дается формулой (6).

В заключение мы отметим *принцип наименьшего числа* и его связь с методом математической индукции. Указанный принцип состоит в том, что *в любом непустом подмножестве множества $\boldsymbol{N}$ натуральных чисел имеется наименьший элемент.*

Справедливость этого принципа нетрудно обосновать. Пусть $\boldsymbol{M}$ — непустое подмножество множества $\boldsymbol{N}$. Так как $\boldsymbol{M}$ непусто, то найдется натуральное число m, принадлежащее подмножеству $\boldsymbol{M}$. Рассмотрим теперь пересечение $\boldsymbol{M} \cap \{1, 2, ..., m\}$. Оно является конечным множеством, содержащимся в $\boldsymbol{N}$, и потому в нем имеется наименьший элемент m_0 (это достаточно «очевидно», но может быть строго доказано индукцией по числу элементов рассматриваемого конечного подмножества). Остается заметить, что m_0 является наименьшим элементом не только конечного подмножества

$\boldsymbol{M} \cap \{1, 2, ..., m\}$, но и всего множества $\boldsymbol{M}$, что доказывает справедливость принципа наименьшего числа.

Таким образом, если возможность проведения доказательств методом математической индукции принимается в качестве аксиомы, то принцип наименьшего числа *доказывается* как теорема. Верно и обратное: если принять принцип наименьшего числа за аксиому, то возможность применения метода математической индукции *доказывается* как теорема. В самом деле, допустим, что утверждение T_n (для которого проведено начало индукции и переход от n к $n+1$) верно не для всех натуральных n. Обозначим через $\boldsymbol{M}$ множество всех n, для которых T_n неверно. Оно, по предположению, непусто. Согласно принципу наименьшего числа в множестве $\boldsymbol{M}$ имеется *наименьшее* число m. Следовательно, обозначая через n число $m-1$, мы заключаем, что T_n верно, а T_{n+1} неверно. Но это невозможно, поскольку был проведен переход от n к $n+1$. Таким образом, предположение о непустоте множества $\boldsymbol{M}$ приводит к противоречию. Следовательно, $\boldsymbol{M}$ пусто, т. е. утверждение T_n верно для всех натуральных n.

Теперь мы можем ввести еще одно, очень важное определение. Пусть некоторое упорядоченное множество является цепью. Это множество называется *вполне упорядоченным*, если в нем имеет место принцип наименьшего элемента, т. е. любое непустое подмножество имеет наименьший элемент. Согласно сказанному выше, множество $\boldsymbol{N}$ всех натуральных чисел является вполне упорядоченным. Множество же всех действительных чисел (или всех рациональных чисел) является упорядоченным, но не вполне упорядоченным. В самом деле, интервал $(a; b)$, являющийся подмножеством множества $\boldsymbol{R}$, *не имеет* наименьшего элемента. В заключительном пункте этой беседы вполне упорядоченные множества будут играть особо важную роль. Однако, если он покажется сложным, читатель может его пропустить без ущерба для понимания дальнейшего изложения.

Задачи и упражнения

71. Пользуясь принципом математической индукции, докажите, что сумма всех нечетных чисел от 1 до $2n+1$ равна $(n+1)^2$.

72. Докажите, что $1^2 + 2^2 + ... + n^2 = \dfrac{n(n+1)(2n+1)}{6}$.

73. Докажите, что $1^3 + 2^3 + ... + n^3 = \dfrac{n^2(n+1)^2}{4}$.

74. Найдите ошибку в следующем рассуждении, проведенном по методу математической индукции.

«Пусть имеется неограниченное число разноцветных шариков. Докажем, что если мы возьмем несколько шариков, то все они будут одного цвета. Для $n=1$ это верно. Пусть это верно для $n=k$, докажем, что это верно для $n=k+1$. Возьмем $k+1$ шарик и отложим один из них. Останется k шариков. По предположению все они одного цвета. Отложим еще один шарик и вернем первоначально отложенный. Получим вновь k шариков, которые по предположению имеют один цвет. Значит, отложенный в первый раз шарик имеет тот же цвет, что и остальные. Итак, все $k+1$ шариков имеют одинаковый цвет. По принципу индукции отсюда следует, что в любом наборе шарики будут одинакового цвета.»

75. Докажите, что неравенство $n^3 - 4 > 1000n^2 + 3n$ выполняется при любом $n \geq 2000$.

22. Трансфинитные числа и аксиома выбора

В общеупотребительном языке мы различаем *количественные* и *порядковые* числительные. Например, в конечном множестве может быть десять элементов, а если оно упорядочено и притом является цепью, то в нем можно рассмотреть *четвертый* или *седьмой* элементы. Подобно этому помимо мощностей, являющихся для бесконечных множеств аналогом количественных числительных, Георг Кантор ввел еще *трансфинитные числа* («сверхконечные числа»), обобщающие порядковые числительные.

Натуральный ряд чисел 1, 2, 3, ... является упорядоченным (и даже вполне упорядоченным) множеством, содержащим бесконечно много элементов. Чтобы охарактеризовать его «порядковый тип», вводится первое бесконечное трансфинитное число ω. Его располагают вслед за всеми натуральными числами: 1, 2, 3, ..., ω.

Получающееся упорядоченное множество имеет порядковый тип $\omega + 1$ (т. е. вслед за натуральными числами еще один элемент). Присоединив это новое трансфинитное число $\omega + 1$, мы получаем упорядоченное множество 1, 2, 3, ..., ω, $\omega + 1$.

Его порядковым типом является $\omega + 2$ (т. е. еще два элемента вслед за натуральными числами). Продолжая, мы получаем последовательность трансфинитных чисел

$$1,\ 2,\ 3,\ \ldots,\ \omega,\ \omega+1,\ \omega+2,\ \omega+3,\ \ldots\ .$$

Порядковый тип этого множества можно обозначить как $\omega + \omega$ или $\omega \cdot 2$. Теперь можно рассмотреть дальнейшие трансфинитные числа

$$1,\ 2,\ 3,\ \ldots,\ \omega,\ \omega+1,\ \ldots,\ \omega\cdot 2,\ \omega\cdot 2+1,\ \omega\cdot 2+2,\ \ldots,$$
$$\omega\cdot 3,\ \omega\cdot 3+1,\ \ldots,\ \omega\cdot 4,\ \ldots,\ \omega\cdot n,\ \ldots\ .$$

Вслед за ними идет трансфинитное число $\omega \cdot \omega = \omega^2$, затем где-то встретятся ω^3, ... ω^4, ...

После ω, ω^2, ω^3, ... ω^n идет трансфинитное число ω^ω, затем $\omega^\omega + 1$, ... Продолжая, мы получим числа

$$1,\ 2,\ 3,\ \ldots,\ \omega,\ \ldots,\ \omega^2,\ \ldots,\ \omega^\omega,\ \ldots,\ \omega^{\omega\cdot 2},\ \ldots,\ \omega^{\omega\cdot 3},\ \ldots,\ \omega^{\omega\cdot n},\ \ldots,$$

после которых идет ω^{ω^2}. Затем будут числа

$$1,\ 2,\ 3,\ \ldots,\ \omega,\ \ldots,\ \omega^\omega,\ \ldots,\ \omega^{\omega^2},\ \ldots,\ \omega^{\omega^3},\ \ldots,\ \omega^{\omega^n},\ \ldots,$$

вслед за которыми идет $\omega^{\omega^{\omega}}$. А после всех чисел

$$1,\ 2,\ 3,\ \ldots,\ \omega,\ \ldots,\ \omega^{\omega},\ \ldots,\ \omega^{\omega^{\omega}},\ \ldots,\ \omega^{\omega^{\omega^{\omega}}},\ \ldots$$

будет идти трансфинитное число, для которого уже нет принятых алгебраических обозначений и для него надо будет ввести какое-то новое обозначение, скажем γ, и т. д.

Каждое из указанных трансфинитных чисел обладает тем свойством, что *перед ним* имеется лишь счетное множество трансфинитных чисел. Поэтому их называют *счетными* трансфинитными числами. Мы можем теперь взять множество всех счетных трансфинитных чисел. За ними всеми следует трансфинитное число, которое принято обозначать символом Ω. Это — первое *несчетное* трансфинитное число, т. е. ему предшествует несчетное множество трансфинитных чисел. Но и на этом процессе образования трансфинитных чисел не обрывается. Вслед за числом Ω идут числа $\Omega+1$, $\Omega+2$, ..., $\Omega+\omega$, ..., $\Omega+\omega+n$, ..., затем $\Omega+\Omega$ (т. е. $\Omega\cdot 2$) и т. д. И какое бы бесконечное множество $\boldsymbol{M}$ мы ни взяли, всегда мы в конце концов можем найти такое трансфинитное число, что все предшествующие ему трансфинитные числа образуют множество *большей мощности*, чем $|\boldsymbol{M}|$.

При этом, сколько бы мы ни построили трансфинитных чисел, их множество оказывается не только цепью, но и вполне упорядоченным множеством, т. е. в нем справедлив принцип минимального элемента. Иначе говоря, какое бы мы ни взяли подмножество множества трансфинитных чисел, в нем всегда найдется наименьший элемент.

И еще полезно заметить, что имеются трансфинитные числа двух видов. К первому виду относятся трансфинитные числа, которые имеют *непосредственно предшествующее* число. Таковы, например, $\omega+1$, $\omega+13$, $\omega\cdot 2+3$, $\omega^2+\omega\cdot 2+5$, $\omega^{\omega}+\omega\cdot 5+8$ и т. п. Трансфинитные числа второго вида (так называемые *предельные* трансфинитные числа) не имеют непосредственно предшествующих. Таковы, например, ω, $\omega\cdot 3$, ω^2, $\omega^2+\omega$, $\omega^{\omega}+\omega^2+\omega\cdot 3$ и т. п.

А теперь расскажем о трансфинитной индукции. Пусть ρ — некоторое трансфинитное число. Предположим, что для каждого трансфинитного числа α, меньшего ρ, сформулировано некоторое утверждение T_{α}, и мы хотим доказать, что это утверждение верно для *любого* $\alpha<\rho$. Как это можно было бы сделать? Для этого применяется *трансфинитная индукция*, которая состоит в следующем.

Во-первых, надо проверить утверждение T_{α} для *наименьшего* трансфинитного числа $\alpha=1$, т. е. доказать утверждение T_1. Это — начало трансфинитной индукции.

Во-вторых, надо осуществить *индуктивный переход* (вспомните «переход от n к $n+1$» при обычной индукции), т. е. доказать, что если

β — произвольное трансфинитное число, меньшее ρ, и если для всех $\alpha < \beta$ утверждение T_α верно, то верно и утверждение T_β.

Если трансфинитная индукция проведена (т. е. осуществлены начало индукции, а также индуктивный переход), то утверждение T_α считается доказанным для всех трансфинитных чисел, меньших ρ. Заметим, что если $\rho = \omega$, то множество всех трансфинитных чисел, меньших ρ, совпадает с натуральным рядом $\{1, 2, 3, ...\}$, и в этом случае трансфинитная индукция сводится к обычной математической индукции. Если же ρ таково, что множество всех трансфинитных чисел, меньших ρ, несчетно, то приходится прибегать к трансфинитнои индукции. Для удобства множество всех трансфинитных чисел, меньших ρ, обозначим через $(0; \rho)$; это — *интервал в множестве трансфинитных чисел*. Таким образом, трансфинитная индукция позволяет доказать утверждение T_α для всех $\alpha \in (0; \rho)$.

Правомерность применения трансфинитной индукции вытекает из того, что множество трансфинитных чисел вполне упорядочено, т. е. в нем справедлив принцип наименьшего элемента. В самом деле, допустим, что утверждение T_α (для которого проведены начало индукции и индуктивный переход) справедливо не для всех $\alpha \in (0; \rho)$, т. е. множество $\boldsymbol{M}$ всех $\alpha \in (0; \rho)$, для которых T_α неверно, является непустым. Пусть β — *наименьший* элемент множества $\boldsymbol{M}$, т. е. T_β неверно, но для всех $\alpha < \beta$ утверждение T_α верно. Однако существование такого β противоречит индуктивному переходу (возможность которого предполагается доказанной). Полученное противоречие означает, что множество $\boldsymbol{M}$ не может быть непустым. Следовательно, $\boldsymbol{M}$ пусто, т. е. T_α верно для любого $\alpha \in (0; \rho)$. Тем самым правомерность метода трансфинитной индукции установлена.

В качестве примера применения трансфинитной индукции докажем следующую теорему: для любого множества $\boldsymbol{A}$ *существует такое трансфинитное число* α, *что интервал* $(0; \alpha)$ *эквивалентен множеству* $\boldsymbol{A}$, *т. е. существует взаимно однозначное отображение* $f\colon (0; \alpha) \to \boldsymbol{A}$.

В самом деле, допустим, что эта теорема неверна. Выберем такое трансфинитное число ρ, что интервал $(0; \rho)$ имеет большую мощность, чем множество $\boldsymbol{A}$. Докажем теперь, что можно для любого $\alpha \in (0; \rho)$ выбрать элемент $x_\alpha \in \boldsymbol{A}$, отличный от ранее выбранных элементов. Это утверждение мы обозначим через T_α и будем его доказывать с помощью трансфинитной индукции. Начало индукции ясно: мы выбираем произвольный элемент множества $\boldsymbol{A}$ и обозначаем его через x_1. Осуществим теперь индуктивный переход. Пусть элементы x_ξ уже выбраны для всех $\xi < \alpha$; тогда, полагая $f(\xi) = x_\xi$, мы получаем *вложение* $f\colon (0; \alpha) \to \boldsymbol{A}$. Это вложение *не является* взаимно однозначным отображением интервала $(0; \alpha)$ на *все* множество A (поскольку мы предположили, что доказываемая теорема неверна). Значит, можно найти в A элемент, *не являющийся* образом какого-либо элемента $\xi \in (0; \alpha)$. Мы выберем такой элемент и обозначим его через x_α. Этот элемент отличен от всех ранее построенных элементов x_ξ, т. е. утверждение T_α верно. Этим осуществлен индуктивный переход. Проведенная трансфинитная индукция доказывает, что T_α верно для *всех* $\alpha \in (0; \rho)$. Иначе говоря, полагая $f(\alpha) = x_\alpha$, мы получаем *вложение* f интервала $(0; \rho)$ в множество A. Однако это невозможно, поскольку мощность интервала $(0; \rho)$ больше мощности множества A. Полученное противоречие и показывает, что предположение о ложности рассмотренной теоремы *ошибочно*, чем и завершается доказательство.

В этом доказательстве имеется некоторый пробел (который, возможно, не был замечен читателем). Однако, прежде чем указать этот пробел и поговорить о его исправлении, мы поясним смысл доказанной теоремы.

Интервал $(0; \rho)$ трансфинитных чисел представляет собой *вполне упорядоченное* множество. Значит, согласно доказанной теореме, *любое* множество A может быть поставлено во взаимно однозначное соответствие с некоторым вполне упорядоченным множеством. Иначе говоря, *любое множество может быть вполне упорядочено*. Можно сказать и иначе: элементы любого множества A могут быть обозначены в виде x_ξ, где ξ пробегает некоторый интервал $(0; \alpha)$ трансфинитных чисел, т. е. элементы любого множества A могут быть *пронумерованы* трансфинитными числами из этого интервала. Это уже кажется невероятным с интуитивной точки зрения. Например, все действительные числа (их несчетное множество) могут быть пронумерованы одно за одним — пусть не обычными целочисленными номерами, а трансфинитными, но все же пронумерованы. Интуиция протестует против такой возможности.

Может быть, где-то в доказательстве рассмотренной теоремы была допущена ошибка? Нет, доказательство было проведено корректно, но в нем неявно была использована *аксиома выбора*, сформулированная в 1904 году математиком Э. Цермело. В несколько упрощенном виде ее использование можно пояснить так. Пусть в процессе рассуждения рассматривается некоторое семейство множеств; тогда представляется возможным выбрать в каждом из этих множеств по одному элементу. В приведенном выше доказательстве (которое также принадлежит Цермело) такими множествами были *дополнения* к множествам уже пронумерованных элементов $\{x_\alpha\}$, и для каждого трансфинитного числа β из интервала $(0; \rho)$ мы выбрали в соответствующем дополнении один элемент x_β. Именно аксиома выбора привела к результату, с которым интуиция не хочет мириться, т. е. аксиома выбора находится в противоречии с нашей интуицией.

В этой связи уместно упомянуть о замечательном результате, который в 1963 году был получен английским математиком П. Коэном. Но прежде чем о нем рассказать, мы сформулируем *континуум-гипотезу*, высказанную в 1878 году Георгом Кантором. Как мы видели в беседе 2, мощность континуума $\mathfrak{C}$ *больше* $\aleph_0$, т. е. больше мощности любого счетного множества. В то же время выше (в рассказе о трансфинитных числах) мы говорили о числе Ω, которое является *первым* несчетным трансфинитным числом. Это означает, что интервал $(0; \Omega)$ имеет мощность большую, чем $\aleph_0$, и притом эта мощность — самая *маленькая* из тех, которые больше $\aleph_0$. Эту мощность обозначают через $\aleph_1$. Следовательно, $\mathfrak{C} \geq \aleph_1$, поскольку $\mathfrak{C} \geq \aleph_0$, а $\aleph_1$ — наименьшая из мощностей, которые больше $\aleph_0$. Гипотеза континуума состоит в том, что $\mathfrak{C}$ *совпадает* с $\aleph_1$, т. е. что $\mathfrak{C}$ и есть та наименьшая мощность, которая превосходит $\aleph_0$. Верно ли это? Ответ на этот вопрос (ждавший своего решения почти столетие), а также на вопрос, «верна ли» аксиома выбора, и был дан П. Коэном. Он установил, что на эти вопросы *невозможно* дать ответ. Несколько более точно: если к основным аксиомам теории множеств (так называемым аксиомам Цермело—Френкеля) мы присоединим в качестве аксиомы континуум-гипотезу, то мы получим столь же непротиворечивую теорию множеств, как и в том случае, если мы присоединим к основным аксиомам предложение, противоположное континуум-гипотезе (т. е. предположение, что $\mathfrak{C} > \aleph_1$). Более того, если к основным аксиомам мы присоединим аксиому выбора, то получим *столь же непротиворечивую* теорию множеств, как и в том случае, если мы присоединим к основным аксиомам *отрицание* аксиомы выбора. Получается, что мы обладаем по крайней мере *четырьмя* различными теориями множеств (с континуум-гипотезой или без нее, с аксиомой выбора или без нее), и все они в одинаковой степени непротиворечивы! О том, что такое непротиворечивость, мы поговорим в беседе 13.

В заключение расскажем о *лемме Цорна*, применение которой часто бывает удобнее, чем непосредственное использование трансфинитной индукции. Прежде всего введем одно определение. Пусть $\boldsymbol{M}$ — упорядоченное множество. Цепь $\boldsymbol{A}$, содержащаяся в $\boldsymbol{M}$, называется *ограниченной*, если в $\boldsymbol{M}$ имеется такой элемент q (принадлежащий или не принадлежащий цепи $\boldsymbol{A}$), что $a \leq q$ для любого $a \in \boldsymbol{A}$.

Лемма Цорна. *Если все цепи, содержащиеся в упорядоченном множестве $\boldsymbol{M}$, ограничены, то $\boldsymbol{M}$ содержит (хотя бы один) максимальный элемент.*

Мы докажем эту лемму с помощью трансфинитной индукции. Предположим, что все цепи, содержащиеся в упорядоченном множестве $\boldsymbol{M}$, ограничены, но $\boldsymbol{M}$ не содержит ни одного максимального элемента, и приведем это предположение к противоречию. Выберем такое трансфинитное число ρ, что интервал $(0; \rho)$ имеет большую мощность, чем множество $\boldsymbol{M}$. Мы докажем, что можно для любого $\alpha \in (0; \rho)$ выбрать элемент $x_\alpha \in \boldsymbol{M}$, который больше любого из ранее выбранных элементов, т. е. $x_\xi < x_\alpha$ при $\xi < \alpha$. Это утверждение обозначим через T_α и будем его доказывать с помощью трансфинитной индукции. Начало индукции ясно: мы выбираем произвольный элемент множества $\boldsymbol{M}$ и обозначаем его через x_1. Осуществим теперь индуктивный переход. Пусть элементы x_ξ уже выбраны для всех $\xi < \alpha$. Нетрудно видеть, что элементы x_ξ, взятые для всех $\xi < \alpha$, образуют цепь. В самом деле, пусть ξ и η — два различных трансфинитных числа, принадлежащих интервалу $(0; \alpha)$. Так как множество трансфинитных чисел вполне упорядочено, то в множестве $\{\xi, \eta\}$ имеется наименьший элемент; пусть, скажем $\xi < \eta$. Тогда x_ξ выбран ранее, чем x_η, и потому $x_\xi < x_\eta$. Таким образом, для любых $\xi, \eta \in (0; \alpha)$ элементы x_ξ, x_η связаны одним из соотношений $x_\xi < x_\eta$, $x_\eta < x_\xi$ (в зависимости от того, будет ли $\xi < \eta$ или $\eta < \xi$). Это означает, что множество всех x_ξ, взятых для $\xi \in (0; \alpha)$, представляет собой цепь. Так как, по предположению, все цепи, содержащиеся в $\boldsymbol{M}$, ограничены, то существует в $\boldsymbol{M}$ такой элемент q, что $x_\xi \le q$ для всех $\xi \in (0; \alpha)$. Но мы предположили, что в $\boldsymbol{M}$ нет ни одного максимального элемента. Значит, элемент q не является максимальным в $\boldsymbol{M}$ и потому существует такой элемент $x_\alpha \in \boldsymbol{M}$, что $q < x_\alpha$. Но тогда $x_\xi \le q < x_\alpha$ для любого $\xi \in (0; \alpha)$, т. е. x_α больше всех ранее выбранных элементов. Этим осуществлен индуктивный переход.

Проведенная трансфинитная индукция означает, что элементы x_α можно выбрать для всех $\alpha \in (0; \rho)$, причем они попарно различны (поскольку $x_\xi < x_\alpha$ при $\xi < \alpha$). Иначе говоря, полагая $f(\alpha) = x_\alpha$ для любого $\alpha \in (0; \rho)$, мы получаем вложение интервала $(0; \rho)$ в множество $\boldsymbol{M}$. Однако это невозможно, поскольку мощность интервала $(0; \rho)$ *больше* мощности множества $\boldsymbol{M}$. Полученное противоречие означает, что предположение об отсутствии в $\boldsymbol{M}$ максимальных элементов ложно, т. е. в $\boldsymbol{M}$ *найдется* хотя бы один максимальный элемент, чем и завершается доказательство.

Читатель заметил, конечно, что мы и здесь применили аксиому выбора. Вообще, большинство математиков предпочитает работать в такой теории множеств, в которой аксиома выбора выполняется.

Глава II. КОМБИНАТОРИКА

Беседа 6. Размещения, сочетания и родственные задачи

23. Размещения с повторениями

Комбинаторику можно назвать *количественной теорией конечных множеств*. Допустим, мы имеем некоторое конечное универсальное множество $\boldsymbol{U}$ и рассматриваем его подмножества некоторого специального вида (скажем, содержащие данное число элементов). *Сколько* таких подмножеств имеется в $\boldsymbol{U}$? Или рассматриваются отображения $\boldsymbol{U} \to \boldsymbol{A}$ некоторого специального вида (где $\boldsymbol{A}$ — заданное конечное множество). Сколько таких отображений существует? Решение этих и аналогичных задач и составляет содержание комбинаторики.

В этом пункте мы рассмотрим вторую из этих задач. Пусть множество $\boldsymbol{U}$ содержит k элементов, скажем $\boldsymbol{U} = \{1, 2, ..., k\}$, а множество $\boldsymbol{A}$ содержит n элементов. Чтобы задать отображение $f\colon \boldsymbol{U} \to \boldsymbol{A}$, мы должны указать образы всех элементов 1, 2, ..., k, т. е. указать некоторые элементы $a_1 = f(1)$, $a_2 = f(2)$, ..., $a_k = f(k)$ множества $\boldsymbol{A}$. Иначе говоря, мы должны задать некоторую *последовательность* a_1, a_2, ..., a_k, состоящую из k элементов множества $\boldsymbol{A}$, т. е. задать некоторое «слово», «буквами» которого являются элементы множества $\boldsymbol{A}$, причем рассматриваются только слова, состоящие из k букв.

Пусть, например, $k = 2$, причем «алфавит» $\boldsymbol{A}$ состоит всего из трех букв: $\boldsymbol{A} = \{a, b, c\}$. В этом случае имеется девять слов:

$$aa,\ ab,\ ac;\ ba,\ bb,\ bc;\ ca,\ cb,\ cc.$$

В самом деле, на первое место мы можем поставить любую из трех букв a, b, c, и, независимо от этого, мы можем на второе место поставить любую из трех букв, откуда и получается $3 \cdot 3 = 9$ двухбуквенных слов. А если алфавит $\boldsymbol{A}$ содержит n «букв», то имеется $n \cdot n = n^2$ двухбуквенных слов. Это можно пояснить и так: каждое двухбуквенное слово представляет собой пару $(a_1; a_2)$ из двух элементов $a_1 \in \boldsymbol{A}$, $a_2 \in \boldsymbol{A}$, т. е. такое слово есть элемент произведения $\boldsymbol{A} \times \boldsymbol{A}$, а это произведение содержит n^2 элементов.

Далее, из алфавита $\boldsymbol{A}$, содержащего n букв, можно составить n^3 трехбуквенных слов. В самом деле, мы имеем n^2 двухбуквенных слов, и к каждому из них можно приписать на третьем месте любую из n букв, т. е. получается $n^2 \cdot n = n^3$ трехбуквенных слов. Теперь уже понятно, что k-буквенных слов, при составлении которых используется n букв, всего имеется n^k (на самом деле это «понятно» означает, что можно провести *математическую индукцию* по k). Например, из русского алфавита, содержащего 32 буквы (включая ё, й, ь, ъ), можно составить $32^4 = 1048576$ четырехбуквенных слов (если, конечно, не обращать внимания на семантическую значимость, т. е. допускать любые слова, скажем, «иуэо» или «ьььь»).

Рассмотренную задачу можно охарактеризовать как задачу о *размещении n* букв на k последовательных местах, причем в одном слове могут *повторяться* одинаковые буквы. Поэтому говорят о *размещениях с повторениями*.

Возвращаясь к отображениям, мы можем сформулировать полученный результат следующим образом: *существует n^k отображений множества $\boldsymbol{U} = \{1, 2, ..., k\}$ в множество $\boldsymbol{A}$, содержащее n элементов.*

Можно рассмотреть более общий случай, когда используются «буквы» из двух (или нескольких) «алфавитов». Пусть, например, рассматриваются слова из $p+q$ букв, в которых первые p букв принадлежат алфавиту $\boldsymbol{A}$, содержащему m букв, а последующие q букв принадлежат алфавиту $\boldsymbol{B}$, содержащему n букв.

Тогда первые p букв дают слово из алфавита $\boldsymbol{A}$ (таких слов имеется m^p), а последующие q букв составляют слово из алфавита $\boldsymbol{B}$ (таких слов имеется n^q). Так как любому способу записать первые p букв можно сопоставить любой из способов записи последующих q букв, то всего имеется $m^p n^q$ требуемых слов.

Пусть, например, номер автомашины содержит четыре цифры и три буквы (скажем, 6257 МТА), причем используются десять цифр 0, 1, 2, ..., 9 и 27 букв русского алфавита (поскольку ё, й, ь, ъ, ы исключаются). Тогда на первых четырех местах мы можем написать 10^4 комбинаций цифр, из которых, однако, надо исключить неиспользуемую комбинацию 0000, т. е. имеется 9999 допустимых комбинаций цифр. Буквенных же комбинаций имеется 27^3. Таким образом, можно написать $9999 \cdot 27^3 = 196\,810\,317$ всех мыслимых автомобильных номеров рассматриваемого вида.

Отдельного рассмотрения заслуживает случай, когда «алфавит» $\boldsymbol{A}$ состоит только из двух букв $\boldsymbol{A} = \{0, 1\}$. Сколько в этом случае существует отображений $\boldsymbol{U} \to \boldsymbol{A}$? Ответ не кажется интересным: поскольку $n = 2$, то из рассмотренного ранее общего результата вытекает, что имеется 2^k отображений $f\colon \boldsymbol{U} \to \{0, 1\}$, если, как и прежде, $\boldsymbol{U}$ содержит k элементов. Однако этот подсчет допускает интересную трактовку.

Каждое отображение $f\colon \boldsymbol{U} \to \{0, 1\}$ однозначно определяется, если указано подмножество $\boldsymbol{M} = f^{-1}(1)$ множества $\boldsymbol{U}$. В самом деле, тогда $f^{-1}(0)$ есть дополнение $c\boldsymbol{M}$ множества $\boldsymbol{M}$ и, значит, для каждого $u \in \boldsymbol{U}$ известен его образ: $f(u) = 1$, если $u \in \boldsymbol{M}$, и $f(u) = 0$, если $u \in (c\boldsymbol{M})$. Иначе говоря, каждое отображение $f\colon \boldsymbol{U} \to \{0, 1\}$ представляет собой *характеристическое отображение* некоторого подмножества $\boldsymbol{M} \subset \boldsymbol{U}$. Сколько подмножеств, столько и отображений, а отображений, как мы знаем, имеется 2^k. Следовательно, *множество $\boldsymbol{U}$,* *содержащее k элементов, имеет 2^k подмножеств* (включая пустое множество и само множество $\boldsymbol{U}$).

В связи с этим множество всех подмножеств множества $\boldsymbol{U}$ принято обозначать через $2^{\boldsymbol{U}}$ (оно содержит $2^{|\boldsymbol{U}|}$ элементов). Это обозначение применяется и в том случае, когда множество $\boldsymbol{U}$ бесконечно.

В этой связи отметим (хотя это и не относится к комбинаторике *конечных* множеств) следующую теорему Кантора: *для всякого множества $\boldsymbol{U}$ множество всех его подмножеств имеет бо́льшую мощность, чем само множество $\boldsymbol{U}$*, т. е. $|2^{\boldsymbol{U}}| > |\boldsymbol{U}|$ (см. упр. 20). В частности, если мы рассмотрим всевозможные подмножества множества $\boldsymbol{R}$ действительных чисел, то они образуют множество, мощность которого *больше*, чем мощность множества действительных чисел, т. е. $2^{\mathfrak{C}} > \mathfrak{C}$ (иногда мощность $2^{\mathfrak{C}}$ называют *гиперконтинуумом*).

Заметим еще, что $2^{\aleph_0} = \mathfrak{C}$ — это также было установлено Кантором. Иначе говоря, *в множестве натуральных чисел имеется столько же всевозможных подмножеств, сколько имеется действительных чисел*. Рассмотрим идею доказательства. Для всякого подмножества $\boldsymbol{M} \subset \boldsymbol{N}$ можно рассмотреть соответствующее «размещение с повторениями», т. е. последовательность из нулей и единиц, в которой единицы стоят на тех местах, номера которых принадлежат множеству $\boldsymbol{M}$, а нули — на остальных местах. Например, множеству $\boldsymbol{M}$, состоящему из всех четных чисел, соответствует последовательность 01010101... . Теперь условимся каждую такую последовательность считать бесконечной *двоичной дробью*, у которой целая часть равна нулю. В частности, множеству всех четных чисел соответствует двоичная дробь 0,01010101... (можно доказать, что она изображает число $\frac{1}{3}$). Таким образом, каждой последовательности из нулей и единиц (т. е. каждому подмножеству множества $\boldsymbol{N}$) соответствует некоторое действительное число, принадлежащее отрезку [0; 1]. Этим определяется некоторое отображение φ множества всех подмножеств множества $\boldsymbol{N}$ в отрезок, т. е. $\varphi\colon 2^{\boldsymbol{N}} \to [0; 1]$. Если бы это отображение было взаимно однозначным, то мы и получили бы доказательство равенства $2^{\aleph_0} = \mathfrak{C}$. В действительности отображение φ является наложением, но взаимно однозначным оно не является; например, бесконечные двоичные дроби 0,10000... и 0,01111... изображают одно и то же число $\frac{1}{2}$. Эту некорректность можно исправить, т. е. превратить идею в строгое доказательство, но как это сделать, мы здесь говорить не будем.

Задачи и упражнения

76. Сколько существует семизначных чисел, не содержащих цифру 7?

77. Шифр устроен следующим образом: каждой цифре соответствует одна из трех букв (см. таблицу), а знаку * — одна из двух букв или пробел.

0	1	2	3	4	5	6	7	8	9	*
а	г	ж	й	м	п	т	х	ш	ы	ю
б	д	з	к	н	р	у	ц	щ	ь	я
в	е	и	л	о	с	ф	ч	ъ	э	_

Рассмотрим запись 5545643*8265. Сколькими способами можно ее восстановить? Какой из них содержателен?

78. Попробуйте найти хотя бы одну осмысленную расшифровку следующей записи: 5343934*150413*6*8156215044414*305041080.

79. Автомобильные номера содержат три буквы и пять цифр, например, Р 502 ВВ 77. Сколько может быть различных автомобильных номеров такого вида?

80. Известна легенда об изобретателе шахмат, который попросил у индийского раджи в награду следующее количество зерен пшеницы: одно на первую клетку, два на вторую, четыре на третью и т. д., удваивая каждый раз количество зерен. В результате получилось количество зерна, которое на земле можно получить лишь за несколько столетий.

Рассмотрим следующую задачу. На каждую клетку шахматной доски либо кладется одно зерно, либо не кладется зерна. Что больше: количество вариантов раскладки или количество зерен, потребованных изобретателем шахмат?

24. Системы счисления

Упомянутая выше двоичная система обозначения чисел (как и другие системы счисления) имеет непосредственное отношение к размещениям с повторениями.

Когда мы говорим «пять тысяч семьсот шестьдесят три», мы имеем в виду число $5{\cdot}1000 + 7{\cdot}100 + 6{\cdot}10 + 3 = 5{\cdot}10^3 + 7{\cdot}10^2 + 6{\cdot}10 + 3$, т. е. число, разложенное по *степеням десятки* с коэффициентами (цифрами), не превосходящими 9. Использование числа 10 в качестве основания системы счисления имеет историческое происхождение, несомненно связанное (полностью или частично) с наличием десяти пальцев на двух руках человека. Имеются и другие системы счисления. Сравнение денежных единиц *мина* и *шеккель*, бывших в ходу у древних шумеров и аккадян, привели к созданию *шестидесятеричной* системы счисления, остатки которой мы встречаем и сегодня (минута содержит 60 секунд, а час содержит 60 минут, т. е. 60^2 секунд; аналогично при измерении углов: градус, минута, секунда). Для представления цифровой информации на различных носителях (числовая информация на дисках компьютеров, звуковая информация на компакт-дисках и т. п.) широко используются *двоичная*, *восьмеричная* и *шестнадцатиричная* системы счисления.

В общем случае при использовании системы счисления, основанием которой служит натуральное число $n > 1$, числа записываются в форме

$$c = a_1 n^{k-1} + a_2 n^{k-2} + \ldots + a_{k-1} n + a_k, \tag{1}$$

где $a_1, a_2, \ldots, a_k$ — *цифры* числа c, каждая из которых может принимать значения $0, 1, \ldots, n-1$ (причем обычно первая слева цифра берется отличной от нуля, поскольку при $a_1 = 0$ можно в записи (1) просто отбросить первое слагаемое и считать, что число c имеет меньшее количество цифр). *Любое* натуральное число *однозначно* записывается в виде (1). В самом деле, рассмотрим все k-значные числа в системе с основанием n, т. е. числа вида (1), или, короче,

$$\overline{a_1a_2 \dots a_k}^{(n)}, \qquad (2)$$

где черта сверху и верхний индекс n означают, что рассматривается не произведение $a_1a_2 \dots a_k$, а запись числа в системе с основанием n. При такой записи мы будем допускать, что и первые цифры *могут быть нулями*: например, «*четырехзначными*» в десятичной системе будут при таком соглашении числа 0001, 0019, 0162, 5837 и т. п., т. е. все числа, являющиеся при обычном понимании (т. е. без нулей слева) *не более* чем четырехзначными.

Количество k-значных чисел (2), т. е. количество *слов* из k букв в алфавите $\boldsymbol{A} = \{0, 1, \dots, n-1\}$, как мы видели выше, равно n^k; но так как запись $\overline{00 \dots 0}^{(n)}$ изображает нуль, а не натуральное число, то получается $n^k - 1$ записей натуральных чисел в виде (2) (или, более подробно, в виде (1)). И как раз *следующее* натуральное число, т. е. n^k, представляет собой первое $(k + 1)$-значное число и записывается в виде $1 \cdot n^k + 0 \cdot n^{k-1} + \dots + 0 \cdot n + 0$. Это рассуждение можно превратить с помощью индукции в корректное доказательство того, что *любое* натуральное число может быть записано (и притом однозначно) в системе счисления с основанием n.

В такой системе счисления действия выполняются по той же схеме, что и в привычной нам десятичной системе; надо только иметь под руками таблицу сложения и таблицу умножения цифр в этой системе. Например, в *восьмеричной* системе имеются 8 цифр, а таблицы сложения и умножения выглядят следующим образом:

сложение								
	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	10
2	2	3	4	5	6	7	10	11
3	3	4	5	6	7	10	11	12
4	4	5	6	7	10	11	12	13
5	5	6	7	10	11	12	13	14
6	6	7	10	11	12	13	14	15
7	7	10	11	12	13	14	15	16

умножение								
	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	10	12	14	16
3	0	3	6	11	14	17	22	25
4	0	4	10	14	20	24	30	34
5	0	5	12	17	24	31	36	43
6	0	6	14	22	30	36	44	52
7	0	7	16	25	34	43	52	61

Например, $5 + 6 = \overline{13}^{(8)}$, поскольку $5 + 6 = \overline{11}^{(10)} = 1 \cdot 8 + 3 = \overline{13}^{(8)}$; далее, $5 \cdot 7 = \overline{43}^{(8)}$, поскольку $5 \cdot 7 = \overline{35}^{(10)} = 4 \cdot 8 + 3 = \overline{43}^{(8)}$.

Имея таблицы сложения и умножения, нетрудно производить обычные действия столбиком, повторяя про себя, скажем, «три пишем, пять в уме», и т. д. Например, в восьмеричной системе (мы не пишем для простоты записи верхний индекс (8)):

$$\begin{array}{r} {}_{+}\begin{array}{r}127\\ \underline{673}\end{array}\\ 1022 \end{array} \qquad \begin{array}{r} {}_{\times}\begin{array}{r}132\\ \underline{27}\end{array}\\ 1166\\ \underline{264}\\ 4026 \end{array}$$

Особенно просто выполняются действия в *двоичной* системе, поскольку в ней очень компактно выглядят таблицы сложения и умножения:

сложение		
	0	1
0	0	1
1	1	10

умножение		
	0	1
0	0	0
1	0	1

Правда, эта простота не дается даром: запись чисел в двоичной системе гораздо длиннее их десятичной записи: $\overline{14}^{(10)} = \overline{1110}^{(2)}$, $\overline{37}^{(10)} = \overline{100101}^{(2)}$ и т. п. Однако простота таблиц сложения и умножения оказывается чрезвычайно удобной для применения этой системы в электронных схемах компьютеров.

Интересно, что гениальный философ и математик Готфрид Вильгельм Лейбниц высоко ценил возможности двоичной системы. Лаплас писал: «В своей бинарной арифметике Лейбниц видел прообраз творения. Ему представлялось, что единица — божественное творение, а нуль — небытие, и что Высшее Существо создает все сущее из небытия подобно тому, как единица и нуль в его системе выражают все числа».

В заключение рассмотрим одну шуточную задачу, бытовавшую в довоенном мехматском фольклоре Московского университета. В то время на велосипеды выдавались шестизначные номера. Номер считался «несчастливым», если в нем есть хотя бы одна цифра 8 (поскольку «восьмерка» — один из дефектов велосипедного колеса). Вопрос заключается в том, чтобы решить, каких номеров больше, счастливых или несчастливых. На первый взгляд кажется, что поскольку 8 — лишь одна цифра из десяти возможных, то счастливых номеров в несколько раз больше. Но проведем точные расчеты.

Счастливый номер — это шестибуквенное слово в «алфавите», содержащем *девять* букв (все цифры, кроме восьмерки). Число таких слов равно $9^6 = 531\,441$, но слово 000000 непригодно в качестве велосипедного номера, т. е. число счастливых номеров равно 531 440. Остальные номера — несчастливые, их $999\,999 - 531\,440 = 468\,559$, т. е. ненамного меньше, чем счастливых. А если бы номера были семизначными, то счастливых номеров было бы меньше, чем несчастливых.

Задачи и упражнения

81. Запишите число 2002 в двоичной системе счисления.

82. Докажите, что число, записанное в троичной системе счисления, делится на 2, если сумма его цифр делится на 2.

83. Числа $1101^{(2)}$ и $11011^{(2)}$ записаны в двоичной системе счисления. Чему равна их сумма в этой системе? А в десятичной?

84. При каком основании системы счисления имеет решение следующий числовой ребус: КИТО + КИОТО = ТОКИО? (Одинаковым буквам должны соответствовать одинаковые цифры, разным—разные; Кито, Киото и Токио — крупные города Японии.)

85. В какой системе счисления $13^2 = 171$?

25. Размещения без повторений

В пункте 23 мы рассмотрели всевозможные отображения $f\colon \boldsymbol{U} \to \boldsymbol{A}$. Число таких отображений равно n^k, где k — число элементов множества $\boldsymbol{U} = \{1, 2, ..., k\}$, а n — число элементов множества $\boldsymbol{A}$. Теперь мы рассмотрим не произвольные отображения $f\colon \boldsymbol{U} \to \boldsymbol{A}$, а только *вложения*. Сколько вложений $f\colon \boldsymbol{U} \to \boldsymbol{A}$ существует?

Как мы знаем, произвольное отображение $f\colon \boldsymbol{U} \to \boldsymbol{A}$ представляет собой некоторое *слово* из k букв, причем буквами являются элементы «алфавита» $\boldsymbol{A}$. При этом буквы $f(1) = a_1$, $f(2) = a_2$, ..., $f(k) = a_k$ могут быть частично совпадающими, т. е. в слове могут быть *повторяющиеся* буквы. Отображение $f\colon \boldsymbol{U} \to \boldsymbol{A}$ в том случае является вложением, если в нем нет повторяющихся букв, т. е. все элементы множества $f(\boldsymbol{U})$ *попарно различны*. Иначе говоря, чтобы получить *вложение* $f\colon \boldsymbol{U} \to \boldsymbol{A}$, мы должны задать некоторую последовательность, состоящую из k элементов a_1, a_2, ..., a_k множества $\boldsymbol{A}$, но так, чтобы в этой последовательности не было повторяющихся элементов. Такие последовательности называются *размещениями без повторений* (состоящими из k элементов множества $\boldsymbol{A}$). Разумеется, когда мы говорим о *последовательности*, то имеем в виду *упорядоченное* множество элементов множества $\boldsymbol{A}$, так что перестановки элементов друг с другом не допускаются. Например, $a_1a_2a_3 \ldots a_k$ и $a_2a_1a_3 \ldots a_k$ — это *различные* размещения.

Отличие между размещениями с повторениями и размещениями без повторений можно пояснить еще следующим образом. Предположим, что у нас имеется набор «образцов» букв (т. е. элементов, входящих в алфавит $\boldsymbol{A}$). При употреблении какой-либо буквы мы изготавливаем ее копию, поэтому каждая буква может быть изготовлена во *многих экземплярах*, так что при наборе слова мы можем использовать любое число экземпляров каждой буквы. Тогда слова получаются с возможными повторениями букв, т. е. мы получаем размещения *с повторениями*.

Совершенно иная ситуация возникает, если каждый элемент множества $\boldsymbol{A}$ уникален, т. е. имеется лишь в *единственном* экземпляре. Тогда, если мы уже использовали элемент a в качестве буквы некоторого слова, то его уже *больше нет* среди оставшихся элементов множества $\boldsymbol{A}$, т. е. повторно использовать его в качестве буквы в этом слове мы уже не можем. Это и дает размещения *без повторений*. Например, если мы хотим построить в некоторой последовательности пять учеников данного класса, то мы не можем поставить некоторого ученика на двух разных местах в одной и той же последовательности (ведь это только в фантастических рассказах у человека могут быть «двойники», т. е. один и тот же человек может существовать в нескольких экземплярах!). Поэтому последовательность пяти построившихся учеников представляет собой *размещение без повторений* (содержащее пять элементов данного множества, т. е. класса).

Теперь, выяснив природу размещений без повторений, мы ответим на вопрос о *числе* таких размещений. Итак, нам нужно составить последовательность $a_1, a_2, ..., a_k$ элементов данного множества (без повторений). Так как множество $\boldsymbol{A}$ содержит n элементов, то у нас есть n возможностей выбрать *первый* элемент последовательности. После этого остался $n-1$ элемент множества $\boldsymbol{A}$, и каждый из них может быть взят в качестве *второго* элемента последовательности. Значит, у нас имеется $n(n-1)$ возможностей выбрать первые два элемента последовательности. Теперь осталось $n-2$ элемента множества $\boldsymbol{A}$, и каждый из них может быть выбран в качестве *третьего* элемента последовательности. Поэтому у нас имеется $n(n-1)(n-2)$ возможностей выбрать первые три элемента последовательности. Аналогично, имеется $n(n-1)(n-2)(n-3)$ возможностей выбрать первые четыре элемента последовательности и т. д.

Продолжая, мы получаем выражение

$$n(n-1)(n-2)\cdot\ldots\cdot(n-k+1), \tag{3}$$

дающее число размещений без повторений по k элементов, взятых из n элементов множества $\boldsymbol{A}$. Разумеется, при этом предполагается, что $k \le n$. В самом деле, ведь в множестве $\boldsymbol{A}$ имеется только n элементов, и потому выбрать из них больше чем n элементов (без повторений) невозможно.

Если, в частности, $k = n$, т. е. мы хотим последовательно расположить все n элементов множества $\boldsymbol{A}$, то, согласно (3), число таких последовательностей равно $n(n-1)(n-2)\cdot\ldots\cdot2\cdot1$, т. е. равно произведению всех натуральных чисел от 1 до n (включительно). Это произведение обозначается через $n!$, т. е. $n! = 1\cdot2\cdot\ldots\cdot(n-1)n$.

Итак, число всех последовательностей, в виде которых можно расположить все n элементов множества $\boldsymbol{A}$, равно $n!$. Эти последовательности называют также *перестановками* (из n элементов множества $\boldsymbol{A}$).

В качестве примера рассмотрим простой случай, когда $n = 3$, т. е. множество $\boldsymbol{A} = \{a, b, c\}$ содержит лишь три элемента, а $k = 2$, т. е. рассматриваются размещения по два элемента. Число размещений *с повторениями* равно $3^2 = 9$. Вот эти размещения: aa, ab, ac; ba, bb, bc; ca, cb, cc.

Далее, число размещений *без повторений*, согласно формуле (3), равно $3\cdot2 = 6$. Эти размещения ab, ac; ba, bc; ca, cb получаются, если

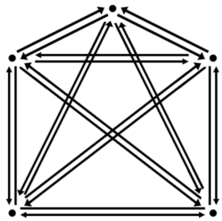


Рис. 165

Рис. 166

в предыдущем списке вычеркнуть размещения, в которых имеются повторения.

Наконец, число *перестановок*, т. е. размещений без повторений, содержащих все три элемента, равно $3! = 1 \cdot 2 \cdot 3 = 6$. Вот эти перестановки: *abc*, *acb*, *bac*, *bca*, *cab*, *cba*.

В виде упражнения читатель может подсчитать, сколько имеется размещений и перестановок из четырех и из пяти элементов (т. е. для $n = 4$ или $n = 5$).

В заключение рассмотрим несколько примеров.

На плоскости даны n точек $a_1, a_2, ..., a_n$ (причем для наглядности будем считать, что никакие три из них не лежат на одной прямой). Сколько существует векторов (направленных отрезков), идущих от какой-либо одной из этих точек к некоторой другой? Чтобы задать вектор, надо выбрать одну из точек $a_1, a_2, ..., a_n$ в качестве его начала, а другую — в качестве его конца, т. е. надо рассматривать *размещения* (без повторений, поскольку вектор идет от одной из точек к *некоторой другой*) по два элемента, составленные из элементов множества $\boldsymbol{A} = \{a_1, a_2, ..., a_n\}$. Согласно (3), число таких размещений (т. е. искомых векторов) равно $n(n - 1)$. Если, например $n = 5$, то получается $5 \cdot 4 = 20$ векторов (рис. 165).

Второй пример. На собрании членов кооператива присутствовали n человек. Надо выбрать председателя, его заместителя и секретаря. Сколькими способами это можно сделать? Речь идет о последовательности, содержащей три элемента из заданных n элементов, т. е. о размещениях без повторений (поскольку один человек не может быть председателем и в то же время его заместителем или секретарем) по 3 элемента из заданных n. Число таких размещений равно $n(n - 1)(n - 2)$, что и дает ответ в рассмотренной задаче. Если, например, на собрании присутствовали 20 человек, т. е. $n = 20$, то число возможностей равно $20 \cdot 19 \cdot 18 = 6840$.

Рассмотрим теперь пример на число перестановок, т. е. последовательностей, в которых участвуют *все* элементы множества $\boldsymbol{A}$ (без повторений). На гулянье собрались 7 девушек и баянист. Когда баянист заиграл хороводную, девушки решили стать в кружок (хоровод). Сколько есть у них возможностей организовать хоровод? Для решения отметим одну из девушек (скажем, старшую). Теперь надо указать, где будут располагаться остальные девушки: кто будет первой по кругу (против хода часовой стрелки) от старшей, кто будет второй, третьей, ..., шестой (рис. 166). Значит, задача заключается в том, чтобы пересчитать все способы расположить шесть девушек в последовательность (первая по кругу от старшей, вторая, ... шестая). Иначе говоря, речь идет о всех *перестановках* из n элементов, где $n = 6$. Число таких перестановок равно $n! = 6! = 1{\cdot}2{\cdot}3{\cdot}4{\cdot}5{\cdot}6 = 720$.

Мы рассмотрели два крайних случая: размещения с *повторениями*, когда каждый элемент может повторяться в последовательности любое число раз, и размещения *без повторений*, когда каждый элемент может встречаться в последовательности не более чем один раз. Могут представляться и различные промежуточные случаи, рассмотрение которых оказывается более сложным. Например, представим себе, что каждая «буква» (т. е. элемент множества $\boldsymbol{A}$) может копироваться не более двух раз. Тогда в качестве размещений могут быть взяты последовательности, в каждой из которых любая буква может встречаться *не более* двух раз.

Задачи и упражнения

86. Сколько имеется способов проставить оценки четырем студентам так, чтобы все они получили разные отметки?

87. Сколько существует пятизначных чисел, у которых цифры убывают, например, как у числа 76310?

88. 25 шахматистов проводят турнир. Его исходом считается указание того, кто занял первое место, кто — второе и кто — третье. Сколько исходов может иметь турнир?

89. В городе 10 музеев. Турист решил их все посетить, но никак не может решить в каком порядке это сделать. Сколько у него вариантов?

90. Сколькими способами можно сделать трехцветный флаг с горизонтальными полосами одинаковой ширины, если имеются материалы шести различных цветов?

26. Сочетания без повторений

Рассмотрим теперь первую из задач, упомянутых в начале этой беседы. Она формулируется более детально следующим образом. Имеется множество $\boldsymbol{A}$, содержащее n элементов; сколько в нем имеется подмножеств, каждое из которых содержит k элементов? Число таких подмножеств обозначается через $\binom{n}{k}$.

Решение этой задачи, т. е. нахождение явного выражения для $\binom{n}{k}$, можно свести к рассмотрению ранее изученных задач. В самом деле, пусть $\{b_1, b_2, ..., b_k\} \subset \boldsymbol{A}$ — некоторое подмножество, содержащее k элементов. Подмножество не является последовательностью, т. е. элементы в нем не считаются как-либо упорядоченными. Однако мы можем упорядочить эти элементы, т. е. расположить их в некоторой последовательности. Так как в рассматриваемом подмножестве имеется k элементов, то всевозможных таких последовательностей

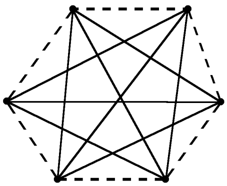


Рис. 167

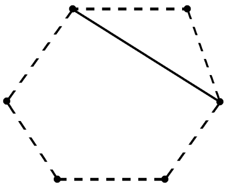


Рис. 168

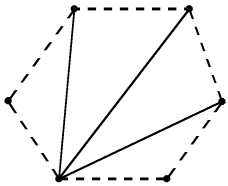


Рис. 169

(т. е. *перестановок* элементов b_1, ..., b_k) имеется $k!$. Если же мы возьмем в A всевозможные подмножества, содержащие k элементов, и для каждого из этих подмножеств возьмем все перестановки его элементов, то мы получим *все* размещения, состоящие из k элементов множества A. Отсюда получается соотношение

$$\binom{n}{k} k! = n(n-1)\ldots(n-k+1),$$

которое и дает решение интересующей нас задачи:

$$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k!}. \tag{4}$$

Это решение можно записать и иначе. Если умножить числитель и знаменатель полученной дроби на число $(n-k)! = 1\cdot 2\cdot \ldots \cdot(n-k)$, то в числителе получится число $n!$. Значит,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{5}$$

Формула (5) представляется «более красиво» записанной, чем (4). Впрочем, в некоторых случаях выражение (4) более удобно для вычислений, чем (5).

Иногда говорят не о *подмножествах*, содержащих k элементов, а о «сочетаниях» из n элементов по k (где n — число всех элементов, имеющихся в множестве A). Формула (4) (или (5)) дает выражение для числа сочетаний.

В качестве примера рассмотрим следующую задачу: сколько диагоналей имеется в выпуклом n-угольнике (рис. 167)? Так как в n-угольнике имеется n вершин, то в его множестве вершин существует $\binom{n}{2}$ подмножеств, содержащих по два элемента.

Каждое такое подмножество определяет либо сторону, либо диагональ (рис. 168). Значит, число всех сторон плюс число всех диагоналей равно $\binom{n}{2}$. А так как у n-угольника имеется n сторон, то число его диагоналей равно

$$\binom{n}{2} - n = \frac{n(n-1)}{2!} - n = \frac{n(n-3)}{2}. \tag{6}$$

Впрочем, этот же результат можно получить и с помощью другого комбинаторного рассуждения. Из каждой вершины n-угольника исходят $n-3$ диагонали (они идут из вершины A ко всем вершинам, кроме самой вершины A и двух смежных с ней вершин, рис. 169). Так как всего имеется n вершин, и из каждой исходят $n-3$ диагонали, то мы таким образом насчитаем $n(n-3)$ диагоналей. Однако при таком подсчете мы засчитаем каждую диагональ *дважды* (с одного и с другого конца). Значит, число диагоналей вдвое меньше, чем $n(n-3)$, что и дает полученный ранее результат (6).

Рассмотрим еще один геометрический пример. На плоскости даны n точек, никакие три из которых не расположены на одной прямой; сколько имеется треугольников с вершинами в этих точках? Ответ очевиден: чтобы получить треугольник, надо взять в качестве вершин какие-либо три из данных точек; поэтому число всевозможных треугольников с вершинами в этих точках равно $\binom{n}{3} = \frac{n(n-1)(n-2)}{3!}$.

Например, для семи точек (т. е. при $n = 7$) получается 35 треугольников.

К задаче о числе сочетаний без повторений можно подойти и иначе. Выбрать из множества $\boldsymbol{A} = \{a_1, a_2, ..., a_n\}$ некоторое подмножество, содержащее k элементов, — это все равно что *разбить* множество $\boldsymbol{A}$ на две части, первая из которых содержит k элементов, а вторая — остальные $n-k$ элементов. Число таких разбиений (т. е. число сочетаний из n элементов по k) равно $\binom{n}{k}$.

А сколько будет разбиений множества $\boldsymbol{A}$ на три части, первая из которых содержит k элементов, вторая содержит l элементов (отличных от выбранных k), а третья — остальные $n-k-l$ элементов? Ответ нетрудно получить с помощью формулы (5). В самом деле, выбрать из $\boldsymbol{A}$ какие-либо k элементов можно $\binom{n}{k}$ способами. Далее, выбрать из оставшихся $n-k$ элементов какие-либо l элементов можно $\binom{n-k}{l}$ способами. А третье множество выбирать уже не нужно: в него войдут $n-k-l$ оставшиеся элементов. Значит, разбить множество $\boldsymbol{A}$ на три группы указанного вида можно

$$\binom{n}{k} \cdot \binom{n-k}{l} = \frac{n!}{k!(n-k)!} \cdot \frac{(n-k)!}{l!(n-k-l)!} = \frac{n!}{k!l!(n-k-l)!}$$

способами.

Аналогичное рассуждение позволяет получить и более общий результат. Пусть дано некоторое множество $\boldsymbol{A}$, содержащее n элементов, и мы хотим разбить его на q групп элементов, первая из которых содержит k_1 элементов, вторая содержит k_2 элементов, ... q-я группа содержит k_q элементов (при этом, когда мы говорим о «разбиении», то предполагаем, что эти группы попарно не имеют общих элементов, а их объединение есть все множество $\boldsymbol{A}$, т. е. $k_1 + k_2 + \ldots + k_q = n$). Сколько таких разбиений существует? Рассуждая так же, как и выше, мы найдем, что число таких разбиений равно

$$\frac{n!}{k_1!k_2!\ldots k_q!}. \tag{7}$$

Например, во время военной игры девять мальчиков решили разбиться на три группы, первая из которых (разведка) должна содержать трех из них, вторая (группа захвата) — четырех, а третья (штаб) — двух мальчиков. Сколькими способами они могут разбиться на группы такого вида? Согласно (7), число способов равно $\frac{9!}{3!4!2!} = 1260$.

Задачи и упражнения

91. Докажите, что $\binom{n}{k} = \binom{n}{n-k}$.

92. Предлагается тест из 10 вопросов. Известно, что на половину из них следует ответить «да», а на вторую половину — «нет». Сколькими способами можно ответить на вопросы теста при данном условии?

93. Сколько различных слов (не обязательно осмысленных) можно получить, переставляя буквы в слове «МАТЕМАТИКА»?

94. Из 20 бусин различного цвета собирается ожерелье. Сколькими различными способами можно расположить на нем бусины?

95. В шахматном турнире Женя и Саша сыграли одинаковое количество партий, заболели и выбыли из турнира. Остальные участники турнира доиграли до конца. Всего было сыграно 28 партий. Играли ли Женя и Саша в турнире между собой?

27. Сочетания с повторениями

Теперь мы рассмотрим ситуацию, которая приведет нас к другому типу комбинаторных задач. Имеются рублевые марки четырех цветов: желтые, зеленые, красные и синие. На почтовый конверт надо наклеить 9 рублевых марок. Сколькими способами это можно сделать? Если мы условимся выклеивать марки в одну линию (скажем, по верхнему краю конверта), то получим уже известную нам задачу: составить слово из девяти букв, используя алфавит {Ж, З, К, С} (начальные буквы цветов). Это — задача на размещения с повторениями, рассмотренная в п. 23. Она имеет n^k решений, где в данном случае $n = 4$, $k = 9$, т. е. задача имеет $4^9 = 262\,144$ решения.

Однако рассмотрим другой вариант этой задачи — нам надо купить эти марки, а порядок расположения марок на конверте не играет роли (их можно клеить где угодно), важен лишь цветовой состав выбираемых девяти марок. Тогда задача принимает следующую форму. Имеется папка (скажем, на почте), содержащая достаточно много марок типа Ж, З, К и С. Сколькими способами можно выбрать из этой папки набор (неупорядоченный!), содержащий девять марок?

Для решения этой задачи будем характеризовать каждый интересующий нас набор следующим образом. Сначала выпишем последовательно столько единиц, сколько в наборе имеется марок типа Ж. Затем поставим нуль, после которого запишем столько единиц, сколько в наборе марок типа З. После этого снова поставим нуль и столько единиц, сколько взято марок типа К. Наконец, нуль и столько единиц, сколько взято марок типа С. Например, запись 111101101011 означает, что в набор вошли четыре марки Ж, две З, одна К и две С, а запись 111110001111 означает, что был взят набор, в котором пять марок Ж и четыре С, а З и К не было взято совсем.

Каждая получаемая запись содержит 9 единиц и 3 нуля, причем набор букв (типов марок) Ж, З, К, С, описываемый этой записью, определяется тем, на каких местах стоят нули. Сколько есть записей (т. е. слов, состоящих из девяти единиц и трех нулей), столько и существует интересующих нас наборов из букв Ж, З, К, С. Но число слов из девяти единиц и трех нулей равно числу сочетаний из 12 элементов по 3: надо из 12 мест выбрать три, на которых будут стоять нули. Значит, число решений равно $\binom{12}{3}=\frac{12\cdot 11\cdot 10}{1\cdot 2\cdot 3}=220$. Как видим, отказ от рассмотрения порядка, в котором выклеиваются марки, существенно уменьшил число решений (т. е. наборов марок).

В рассмотренной задаче используются *сочетания с повторениями*. В общем случае эта комбинаторная задача может быть сформулирована следующим образом. Рассматривается алфавит $\boldsymbol{A}$ из n букв и имеется папка, содержащая достаточно большое число экземпляров каждой буквы. Нас интересует, сколькими способами можно выбрать из этой папки неупорядоченный набор, содержащий k букв. Такие наборы называются *сочетаниями с повторениями* из n элементов по k. Ответ получается дословно так же (выписыванием последовательностей из нулей и единиц), как и в рассмотренной выше задаче о почтовых марках: число сочетаний с повторениями из n элементов по k равно

$$\binom{n+k-1}{n-1}=\frac{(n+k-1)(n+k-2)\,...\,(k+1)}{(n-1)!}.$$

Так, в задаче о почтовых марках мы имели $n=4$, $k=9$, т. е. $n+k-1=12$ и потому число решений равно $\binom{12}{3}=220$.

Как и в случае размещений, мы рассмотрели два крайних случая задач на сочетания. В первом случае (сочетания без повторений) каждый элемент множества $\boldsymbol{A}$ был уникален, т. е. имелся в единственном экземпляре, и нас интересовали всевозможные подмножества

множества $\boldsymbol{A}$, содержащие k элементов. Во втором случае (сочетания с повторениями) элементы множества $\boldsymbol{A}$, т. е. «буквы», были изготовлены в неограниченном количестве экземпляров («касса букв»), и нас интересовало, сколько различных наборов, содержащих k букв, можно составить с помощью этой кассы. Возможны и различные промежуточные случаи. Скажем, буква a_1 может иметься в кассе в количестве m_1 экземпляров, буква a_2 — в количестве m_2 экземпляров, ..., буква a_n — в количестве m_n экземпляров, и нас интересует, сколько сочетаний из k элементов может быть составлено с помощью этой кассы. Формулы для решения этой задачи являются более сложными.

Задачи и упражнения

96. В магазине продаются пирожные четырех сортов; эклеры, слоеные, песочные и бисквитные. Сколькими различными способами можно составить набор из десяти пирожных, если в наборе должно быть хотя бы одно пирожное каждого сорта?

97. В джаз-оркестре из 7 человек должны быть трубачи, ударники и саксофонисты (хотя бы по одному). Сколько различных составов таких оркестров может быть?

98. В фондовом магазине предлагаются акции семи компаний. Сколькими способами можно купить десять акций?

99. Король Эмирии учредил четыре военных ордена. Сколькими способами он может наградить семью орденами своего военного министра?

100. ООН формирует батальон миротворческих сил для поддержания мира в одной из стран. В создании этого батальона готовы принять участие 12 стран. Сколькими способами можно сформировать батальон, если в нем 6 взводов и каждая страна готова послать необходимое количество взводов?

28. Бином Ньютона

Биномом (или двучленом) называют сумму двух слагаемых, например, $x+a$. В школе хорошо известны формулы для квадрата и куба такого бинома:

$$(x+a)^2 = x^2 + 2ax + a^2,\ (x+a)^3 = x^3 + 3x^2a + 3xa^2 + a^3.$$

Речь идет о нахождении формулы для $(x+a)^n$ в случае произвольного натурального n.

По определению $(x+a)^n = (x+a)(x+a)...(x+a)$, где в правой части имеется n множителей. Нам надо написать выражение, получающееся при раскрытии скобок. Это выражение является суммой одночленов, каждый из которых получается, когда из каждой скобки берется одно из слагаемых (x или a) и они перемножаются, а затем приводятся подобные члены; поэтому слагаемые имеют вид $c_k x^{n-k} a^k$. Если в каждой из скобок $(x+a)$ мы возьмем первое слагаемое x, то получится произведение x^n. Далее, если в одной из скобок взять a, тогда как во всех остальных взять x, то получится произведение $x^{n-1}a$. Сколько раз оно войдет в сумму, получающуюся при раскрытии скобок? Ясно, что оно получится столько раз, сколькими способами можно отметить *одну* скобку (в которой мы берем a) из имеющихся n скобок, т. е. столько, сколько имеется *одноэлементных* подмножеств в множестве $\{1, 2, ..., n\}$. Поэтому при раскрытии скобок получится $\binom{n}{1} = n$ слагаемых $x^{n-1}a$:

$$(x + a)^n = x^n + nx^{n-1}a + ... \;.$$

Следующим возможным слагаемым будет $x^{n-2}a^2$. Таких слагаемых при раскрытии скобок получится столько, сколькими способами можно отметить *две* скобки (в которых мы берем a) из имеющихся n скобок, т. е. получится $\binom{n}{2} = \frac{n(n-1)}{2}$ слагаемых $x^{n-2}a^2$:

$$(x + a)^n = x^n + \binom{n}{1}x^{n-1}a + \binom{n}{2}x^{n-2}a^2 + ... \;.$$

Вообще, слагаемое $x^{n-k}a^k$ получится столько раз, сколькими способами можно выбрать k скобок (в которых мы берем a) из имеющихся n скобок, т. е. оно получится $\binom{n}{k}$ раз. Это и дает формулу Ньютона:

$$(x + a)^n =$$

$$= x^n + \binom{n}{1}x^{n-1}a + \binom{n}{2}x^{n-2}a^2 + ... + \binom{n}{k}x^{n-k}a^k + ... + \binom{n}{n-1}xa^{n-1} + a^n. \quad (8)$$

Здесь последнее слагаемое a^n получается, если из *всех* скобок взять a; это можно сделать $\binom{n}{n}$ способами (т. е. единственным способом), и потому это слагаемое входит с коэффициентом $\binom{n}{n} = 1$.

Аналогичными рассуждениями может быть получена более общая формула

$$(x_1 + x_2 ... + x_q)^n = \sum \frac{n!}{k_1!k_2! ... k_q!} x_1^{k_1} x_2^{k_2} ... x_q^{k_q}, \quad (9)$$

в которой знак $\sum$ (суммирование) распространен на всевозможные способы представить n в виде $n = k_1 + k_2 + ... + k_q$. В самом деле, произведение $x_1^{k_1} x_2^{k_2} ... x_q^{k_q}$ получится при раскрытии скобок столько раз, сколько есть разбиений множества $\{1, 2, ..., n\}$ (т. е. множества номеров скобок) на q групп, в первой из которых имеется k_1 элементов (это будут те скобки, из которых мы берем x_1), во второй k_2 элементов (это будут скобки, из которых мы берем x_2), ..., в q-й k_q элементов. Число таких разбиений дается формулой (7), откуда и вытекает соотношение (9).

Например, если $n = 3$ т. е. имеется только три слагаемых x_1, x_2, x_3, то получается следующее соотношение:

$$(x_1 + x_2 + x_3)^n =$$

$$= x_1^3 + x_2^3 + x_3^3 + 3x_1^2x_2 + 3x_1^2x_3 + 3x_1x_2^2 + 3x_1x_3^2 + 3x_2^2x_3 + 3x_2x_3^2 + 6x_1x_2x_3.$$

Коэффициент при последнем слагаемом получается как $\frac{3!}{1!\,1!\,1!} = 6$.

Вернемся к формуле Ньютона (8). Полагая в ней $x = 1$, $a = 1$, мы получаем интересное соотношение

$$1 + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{k} + \ldots + \binom{n}{n-1} + 1 = 2^n. \qquad (10)$$

Впрочем, это соотношение станет очевидным, если мы заметим, что 1 есть число *пустых* подмножеств в множестве $\boldsymbol{A} = \{1, 2, ..., n\}$, далее $\binom{n}{1}$ есть число *одноэлементных* подмножеств в $\boldsymbol{A}$, ..., $\binom{n}{k}$ — число k-*элементных* подмножеств в $\boldsymbol{A}$ и т. д., так что в целом сумма в левой части (10) означает число *всевозможных* подмножеств множества $\boldsymbol{A}$. Но мы знаем, что число всех подмножеств n-элементного множества $\boldsymbol{A}$ равно 2^n (это мы видели в п. 23), т. е. как раз получается соотношение (10) — без применения формулы Ньютона.

Однако при $x = 1$, $a = -1$ из (8) получается соотношение, которое не столь очевидно:

$$1 - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \binom{n}{4} - \ldots = 0,$$

т. е. в биноминальной формуле (8) сумма коэффициентов, стоящих на четных местах, равна сумме коэффициентов на нечетных местах. Это особенно интересно для *четных* n:

$$(x + a)^4 = x^4 + 4x^3a + 6x^2a^2 + 4xa^3 + a^4;$$
$$(x + a)^6 = x^6 + 6x^5a + 15x^4a^2 + 20x^3a^3 + 15x^2a^4 + 6xa^5 + a^6.$$

Если выписать строка за строкой коэффициенты, встречающиеся в формуле (8), для $n = 0, 1, 2, ...$, то получается таблица, называемая *треугольником Паскаля*:

$$\begin{array}{c}
1\\
1\quad 1\\
1\quad 2\quad 1\\
1\quad 3\quad 3\quad 1\\
1\quad 4\quad 6\quad 4\quad 1\\
1\quad 5\quad 10\quad \underline{\mathit{10}\quad \mathit{5}}\quad 1\\
1\quad 6\quad 15\quad 20\quad \mathit{15}\quad 6\quad 1
\end{array}$$

Эта таблица может быть легко построена и без формулы (8), если заметить, что в ней *каждое число равно сумме двух соседних с ним чисел верхней строки* (см., например, выделенные числа).

То, что это действительно так, может быть доказано следующим образом. В n-й строке имеется коэффициент $\binom{n}{k}$ (как видно из формулы (8)), а над ним, в $n-1$-й строке стоят числа $\binom{n-1}{k-1}$ и $\binom{n-1}{k}$, т. е. фрагмент таблицы имеет вид

$$\begin{array}{c}
\cdots\binom{n-1}{k-1}\binom{n-1}{k}\cdots\\
\cdots\binom{n}{k}\cdots
\end{array}$$

Сформулированное выше правило (способ построения треугольника Паскаля) означает, что справедливо соотношение

$$\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k}. \tag{11}$$

Действительно, $\binom{n}{k}$ есть число *всевозможных сочетаний* из n элементов множества $\{1, 2, ..., n\}$ по k элементов. Но все эти сочетания можно разбить на две группы: содержащие элемент n и не содержащие элемент n. Сочетаний, *содержащих* элемент n, будет $\binom{n-1}{k-1}$, поскольку к взятому элементу n надо добавить еще $k-1$ элементов из $\{1, 2, ..., n-1\}$. А сочетаний, *не содержащих* элемент n, будет $\binom{n-1}{k}$. Это и дает формулу (11), т. е. доказывает описанное выше правило построения треугольника Паскаля.

Заметим в заключение, что формула (8) была известна и до Ньютона. Ее знали в средние века Омар Хайям и другие среднеазиатские математики. Блез Паскаль, о котором мы еще будем говорить в связи с понятием *вероятности* (и который ввел рассматривавшийся выше *треугольник Паскаля*), также владел этой формулой задолго до Ньютона. Однако Ньютону принадлежит некоторое обобщение формулы (8): он нашел выражение для $(x+a)^n$ в случае *нецелого* показателя n.

Задачи и упражнения

101. Найдите разложение для $(2x + a)^4$.

102. При какой степени x коэффициент будет наибольшим в биноме $(2 + x)^{10}$?

103. Чему равен коэффициент при члене $x^2y^3z^2$ в выражении $(x + y + z)^7$?

104. Докажите равенство $\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \binom{n}{2}^2 + \ldots + \binom{n}{n}^2$.

105. Постройте треугольник Паскаля, используя «двоичную арифметику»: $0 + 0 = 0,\ 0 + 1 = 1 + 0 = 0,\ 1 + 1 = 10$. Исследуйте получающиеся треугольники Паскаля.

29. Производящие функции

В формуле (8) коэффициентами при степенях переменного x являются числа $\binom{n}{k}$, дающие решение комбинаторной задачи о сочетаниях (без повторений). Говорят поэтому, что $(x + a)^n$ является *производящим многочленом* для этой комбинаторной задачи. Производящие функции можно рассмотреть и для других комбинаторных задач.

Рассмотрим, например, следующую задачу: сколькими способами можно уплатить сумму k копеек, используя каждую монету достоинством в 1, 2, 3, 5, 10, 15, 20 и 50 копеек не более одного раза? Обозначим это число способов через b_k (например, b_{11} — число способов уплатить 11 копеек указанными монетами). Мы можем тогда составить многочлен

$$1 + b_1x + b_2x^2 + b_3x^3 + \ldots + b_{106}x^{106}$$

(число 106 — это сумма достоинств *всех* взятых монет, т. е. максимальная сумма, которую мы можем уплатить). Это и есть производящий многочлен для рассматриваемой комбинаторной задачи. Но как найти числа $b_1, b_2, \ldots, b_{106}$? Ведь это и есть интересующий нас вопрос. Оказывается, этот производящий многочлен можно записать и в другом виде:

$$1 + b_1x + b_2x^2 + b_3x^3 + \ldots + b_{106}x^{106} =$$
$$= (1 + x)(1 + x^2)(1 + x^3)(1 + x^5)(1 + x^{10})(1 + x^{15})(1 + x^{20})(1 + x^{50}). \quad (12)$$

Здесь показателями степени в правой части служат как раз достоинства взятых монет. Раскрывая скобки в правой части, мы получим явное выражение взятого производящего многочлена, и это даст возможность найти все коэффициенты b_k, т. е. полностью решить поставленную комбинаторную задачу.

Откуда же получается формула (12)? При раскрытии скобок в правой части получатся слагаемые вида $x^{m_1+m_2+\ldots+m_p}$, где $m_1, m_2, \ldots, m_p$ — некоторые из чисел 1, 2, 3, 5, 10, 15, 20, 50. Сколько раз встретится слагаемое, у которого показатель $m_1 + m_2 + \ldots + m_p$ равен заданному k, столько и будет способов уплатить сумму k копеек. Иными словами, x^k встретится при раскрытии скобок столько раз, сколько есть способов уплатить k копеек, т. е. b_k раз. Это и означает справедливость формулы (12).

Рассмотрим теперь другую комбинаторную задачу. Известно, что любую сумму k рублей, где $k \geq 8$, можно уплатить 3- и 5-рублевыми банкнотами (раньше такие банкноты были в ходу в России). В самом деле, 8, 9 и 10 рублей уплатить легко, а потом можно добавлять одни лишь трешки. Однако *сколькими способами* можно уплатить сумму k рублей, используя только трешки и пятерки? Обозначим это число способов через C_k и составим соответствующую производящую функцию:

$$1 + C_1x + C_2x^2 + \ldots + C_kx^k + \ldots ;$$

она будет представлять собой уже не многочлен, а *бесконечный* ряд (если читатель слышал кое-что о *сходимости* бесконечных рядов, пусть он пока забудет это — мы сейчас будем производить «формальное» действие, не заботясь о сходимости). Однако, как явно вычислить этот ряд, т. е. получить решение интересующей нас задачи для любого k? В этом случае вместо (12) мы получаем другую формулу:

$$1 + C_1x + C_2x^2 + \ldots + C_kx^k + \ldots =$$
$$= (1 + x^3 + x^6 + x^9 + x^{12} + \ldots)(1 + x^5 + x^{10} + x^{15} + \ldots). \qquad (13)$$

Перемножая ряды в правой части (т. е. производя формальное умножение последовательных членов этих рядов), мы и получаем явное выражение производящей функции:

$$1 + x^3 + x^5 + x^8 + x^9 + x^{10} + x^{11} + x^{12} +$$
$$+ x^{13} + x^{14} + 2x^{15} + x^{16} + x^{17} + 2x^{18} + \ldots .$$

Например, x^{18} получается дважды: $x^{18} = x^3 \cdot x^{15}$ и $x^{18} = x^{18} \cdot 1$; это и дает два способа уплатить 18 рублей: одна трешка и остальное пятерками или же только трешки без пятерок. Детали читатель может продумать самостоятельно.

А если задача заключается в уплате суммы k рублей трешками, пятерками и десятками, то производящая функция примет вид

$$(1 + x^3 + x^6 + x^9 + \ldots)(1 + x^5 + x^{10} + x^{15} + \ldots)(1 + x^{10} + x^{20} + x^{30} \ldots).$$

В приведенных задачах рассматривались разбиения числа k на слагаемые некоторого заданного вида, причем *порядок* этих слагаемых не учитывался. В самом деле, можно уплатить сначала одну трешку и затем три пятерки или, наоборот, сначала дать три пятерки и потом трешку — это будет один и тот же способ уплатить 18 рублей трешками и пятерками. Нам важно лишь, какие купюры и в каком количестве используются, без учета порядка их уплаты. Есть однако, как мы видели выше, комбинаторные задачи, в которых существенно учитывается порядок. Рассмотрим производящие функции для некоторых из таких задач.

Как мы знаем, число размещений с повторениями из n элементов по k равно n^k. Здесь n, т. е. число элементов множества $\boldsymbol{A} = \{0, 1, ..., n\}$, из которых мы составляем размещения, предполагается фиксированным, а k (т. е. длина рассматриваемых «слов») может принимать значение 0, 1, 2, ... Соответствующая производящая функция имеет вид бесконечного ряда

$$1 + nx + n^2x^2 + ... + n^kx^k + ...,$$

т. е. коэффициентом при x^k является число размещений с повторениями по k элементов (т. е. число слов длины k). Заметим теперь, что справедливо следующее формальное тождество (еще раз подчеркиваем, что вопроса сходимости рассматриваемых рядов мы не затрагиваем):

$$(1 - z)(1 + z + z^2 + z^3 + ... + z^k + ...) = 1. \tag{14}$$

В самом деле, раскрывая скобки, мы получаем в левой части

$$(1 - z) + (z - z^2) + (z^2 - z^3) + (z^3 - z^4) + ...,$$

откуда видно, что все слагаемые, кроме 1, взаимно уничтожаются. В частности, при $z = nx$ мы получаем из равенства (14)

$$\frac{1}{1 - nx} = 1 + nx + n^2x^2 + n^3x^3 + ... + n^kx^k + ...,$$

откуда видно, что рассмотренная выше производящая функция для размещений с повторениями из n элементов может быть записана в виде $\dfrac{1}{1 - nx}$.

В заключение рассмотрим еще одну комбинаторную задачу с учетом порядка выбора элементов. Имеются (в достаточном количестве) рублевые, четырехрублевые и десятирублевые почтовые марки. На конверт надо наклеить четыре марки на общую сумму 22 рубля. Сколькими способами это можно сделать, если учитывается порядок расположения марок (выклеиваемых по правому краю конверта)? Для решения мы сразу выпишем соответствующую производящую функцию $(x + x^4 + x^{10})^4$. Чтобы пояснить ее смысл, мы запишем ее в виде

$$(x + x^4 + x^{10})(x + x^4 + x^{10})(x + x^4 + x^{10})(x + x^4 + x^{10})$$

и условимся, что первая скобка характеризует первую марку, вторая — вторую и т. д. Если, например, мы взяли из первой скобки x, из

1
4
10
10

Рис. 170

1	1	1	10	10	10	4	4	4	10
1	10	10	10	1	1	4	4	10	4
10	1	10	1	10	1	4	10	4	4
10	10	1	1	1	10	10	4	4	4

Рис. 171

второй x^4, а из третьей и четвертой взяли x^{10}, то это соответствует выклеиванию марок, показанному на рис. 170; здесь получился набор марок на сумму 25 рублей, т. е. на большую сумму, чем следовало. Если же мы раскроем скобки и найдем, сколько раз встретится x^{22}, то мы и получим число способов требуемой выклейки марок. Иначе говоря,

$$(x + x^4 + x^{10})^4 = r_4x^4 + r_5x^5 + r_6x^6 + \ldots + r_{22}x^{22} + \ldots + r_{40}x^{40}, \qquad (15)$$

где r_k означает число способов, которыми можно (при поставленных условиях) выклеить марки на сумму k рублей. Раскрыв скобки, мы найдем, что $r_{22} = 10$, т. е. возможны 10 способов требуемой расклейки марок (рис. 171). Конечно, найти эти способы выклейки марок можно было бы и без использования производящей функции, но если мы хотим найти решение сразу для всех k, то проще всего раскрыть скобки в выражении, стоящем в левой части равенства (15).

Мы рассмотрели лишь несколько отдельных примеров. Вообще же метод производящих функций является важным приемом решения задач комбинаторики (и не только комбинаторики).

Задачи и упражнения

106. Найдите производящую функцию для геометрической прогрессии a_0, a_0q, a_0q^2, ...

107. С помощью производящей функции докажите тождество

$$\binom{n+m}{s} = \binom{n}{0}\binom{m}{s} + \binom{n}{1}\binom{m}{s-1} + \ldots + \binom{n}{k}\binom{m}{s-k} + \ldots + \binom{n}{s}\binom{m}{0}.$$

Считается, что при $s - k < 0$ число $\binom{m}{s-k}$ равно нулю; кроме того, считается, что $0! = 1$ и потому $\binom{n}{0} = \frac{n!}{0!\,n!} = 1$.

108. Найдите производящую функцию для натурального ряда.

109. С помощью производящей функции, задаваемой биномом Ньютона, докажите равенство $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$.

110. Докажите, что $1\cdot(2n + 1) - 2\cdot 2n + 3\cdot(2n - 1) - \ldots + (2n + 1)\cdot 1 = n + 1$.

30. Принцип Дирихле

Зададим следующий риторический вопрос, ответ на который заранее ясен. Множество $\boldsymbol{A}$ содержит $n+1$ элемент, а множество $\boldsymbol{B}$ содержит n элементов; существует ли *вложение* $\boldsymbol{A} \to \boldsymbol{B}$? Ответ, очевидно, отрицателен. Ведь множество $\boldsymbol{A}$ содержит больше элементов, чем $\boldsymbol{B}$, и потому вложить $\boldsymbol{A}$ в множество $\boldsymbol{B}$ в качестве его подмножества невозможно. Иначе говоря, никакое отображение $f\colon \boldsymbol{A} \to \boldsymbol{B}$ не является (при указанных условиях) вложением, т. е. найдутся два элемента a_1, $a_2 \in \boldsymbol{A}$, имеющие один и тот же образ: $f(a_1) = f(a_2)$ (рис. 172).

Несмотря на полную очевидность сказанного, дадим еще одно наглядное пояснение. Элементы множества $\boldsymbol{A}$ будем представлять себе *зайцами*, а элементы множества $\boldsymbol{B}$ — *клетками*. Отображение будет состоять в том, что мы сажаем зайцев в клетки, т. е. каждому зайцу поставлена в соответствие (в качестве его образа) та клетка, в которую он посажен. Поскольку зайцев больше чем клеток, это отображение не может быть вложением, т. е. не может оказаться так, что каждому зайцу предоставлена отдельная клетка, в которой сидит только он. Иными словами, если $n+1$ зайцев рассадить в n клеток, то хотя бы в одной клетке окажется не менее двух зайцев (рис. 173).

Указанный факт, несмотря на полную его очевидность, нередко используется в математических рассуждениях. Он называется *принципом Дирихле* (немецкий математик Петер Дирихле часто применял его в своих исследованиях по теории чисел). Приведем несколько примеров использования принципа Дирихле.

На плоскости даны три прямых и восемь точек, не лежащих на этих прямых. Можно ли среди этих точек найти такие две, что соединяющий их отрезок не пересекает ни одной из данных прямых? Для решения заметим, что если прямые не пересекаются в одной точке и попарно не параллельны, то они разбивают плоскость на *семь* частей, если же прямые пересекаются в одной точке или среди них есть параллельные, то число частей будет меньшим (рис. 174). Части плоскости будем считать «клетками», а точки — «зайцами». Так как

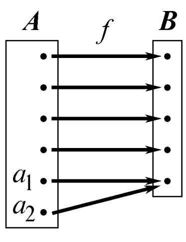


Рис. 172

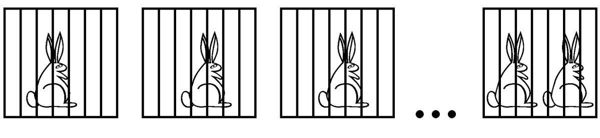


Рис. 173

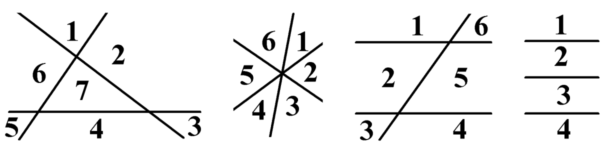


Рис. 174

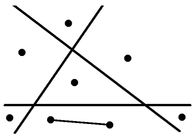


Рис. 175

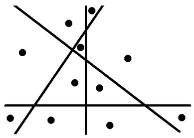


Рис. 176

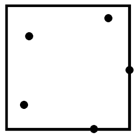


Рис. 177

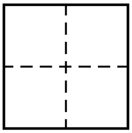


Рис. 178

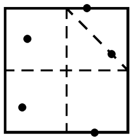


Рис. 179

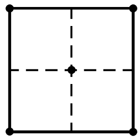


Рис. 180

точки не лежат на заданных прямых, то каждая из них попадает в некоторую клетку. Итак, у нас есть 8 зайцев, рассаженных в 7 или меньшее число клеток. Следовательно, в какой-либо из клеток окажется не менее двух зайцев, т. е. найдутся две точки, лежащие в одной и той же части плоскости (рис. 175). Отрезок, соединяющий эти две точки, не пересекается ни с одной из данных прямых. Таким образом, ответ на поставленный вопрос утвердителен.

Используя пример, рассмотренный в п. 21, читатель может обобщить этот результат следующим образом: если на плоскости даны n прямых и $\frac{n(n+1)}{2}+2$ точек, не лежащих на этих прямых, то среди этих точек найдутся такие две, что соединяющий их отрезок не пересекает ни одной из данных прямых. Заметим, что это — *наименьшее* число точек, для которых сформулированный результат верен: если взять на одну точку меньше, т. е. $\frac{n(n+1)}{2}+1$ точек, то может оказаться, что все зайцы сидят в отдельных клетках (случай $n=4$ изображен на рис. 176), и потому любой отрезок, соединяющий две из этих точек, непременно пересекает хотя бы одну из данных прямых.

Следующий пример мы также возьмем из области геометрии. Внутри (или на границе) квадрата со стороной 1 взяты пять точек (рис. 177). Надо доказать, что среди этих пяти точек найдутся две точки, расстояние между которыми не превосходит $\frac{1}{\sqrt{2}}$. Для доказательства разделим данный квадрат на четыре квадрата половинного размера (рис. 178), и эти меньшие квадраты будем считать «клетками».

Так как число точек («зайцев») равно 5, то, согласно принципу Дирихле, найдутся два зайца, сидящие в одной клетке, т. е. две точки, лежащие в одном квадрате-четвертушке (рис. 179). Но тогда расстояние между ними не превосходит диагонали этого маленького квадрата, т. е. это расстояние не больше $\frac{1}{\sqrt{2}}$.

Заметим, что в доказанном утверждении число $\frac{1}{\sqrt{2}}$ не может быть уменьшено: на рис. 180 изображены 5 точек в квадрате, среди которых *нет* двух точек, находящихся на расстоянии, меньшем $\frac{1}{\sqrt{2}}$ друг от друга.

Теперь расскажем о применении принципа Дирихле к десятичной записи рациональных чисел. Результат формулируется следующим образом: *всякое рациональное число* $\frac{p}{q}$ *изображается бесконечной периодической десятичной дробью, которая содержит не более* $q-1$ *цифр в периоде*. Например, обращая дробь $\frac{1}{7}$ в десятичную, т. е. производя деление числа 1 на 7, получаем следующие цифры:

$$\begin{array}{r|l} 1{,}0 & \underline{\;\;\;\;7} \\ \underline{\;\;7} & 0{,}14 \\ 30 & \\ \underline{28} & \\ \ldots & \end{array}$$

Продолжая деление, находим: $\frac{1}{7} = 0{,}142857142857142857\ldots$, т. е. получается периодическая дробь, имеющая период из шести цифр 142857 (здесь $q = 7$, т. е. период содержит как раз $q-1$ цифру). А для дроби $\frac{1}{3}$ находим десятичное представление $\frac{1}{3} = 0{,}33333\ldots$; здесь $q = 3$, т. е. $q-1=2$, но период содержит лишь одну цифру, т. е. меньше чем $q-1$.

Доказательство сформулированного общего результата можно получить с помощью принципа Дирихле следующим образом. При выполнении деления $p{:}q$ «углом» мы получаем некоторый *остаток* (как, например, выше, при делении 1:7), затем приписываем справа 0, снова делим, снова получаем остаток и т. д. Если где-то встретится остаток, равный нулю (т. е. в каком-то месте деление осуществляется без остатка), то $\frac{p}{q}$ записывается в виде *конечной* десятичной дроби. Например, $\frac{7}{20} = 0{,}350000\ldots$, т. е. в этом случае период содержит лишь одну цифру 0.

Рассмотрим теперь случай, когда все время получаются остатки, отличные от нуля. В этом случае остатками (при делении на q) могут быть числа 1, 2, ..., $q-1$, т. е. возможны $q-1$ различных остатков.

Их мы будем считать «клетками». Теперь рассмотрим q последовательно получающихся остатков: 1-й, 2-й, ..., q-й. Числа 1, 2, ..., q (номера получающихся остатков) будем считать «зайцами». Каждый из этих остатков равен одному из чисел $1, 2, ..., q-1$, т. е. он попадает в одну из клеток. А так как зайцев больше, чем клеток, то найдется клетка, в которую попадает *больше одного зайца*. Иначе говоря, найдутся среди чисел $1, 2, ..., q$ такие два числа, скажем, k и $k+r$, где $k+r \le q$, и потому $r < q$, что k-й остаток *совпадает* с $(k+r)$-м остатком. Но каждая последующая цифра в процессе деления определяется лишь тем, какой был предыдущий остаток. Поэтому, начиная с $(k+r)$-го места, цифры будут *те же*, что и с k-го места, потом с $(k+2r)$-го места снова пойдут те же r цифр и т. д. Таким образом, в десятичной записи числа $\frac{p}{q}$, начиная с k-го места, одни и те же r цифр будут все время повторяться. Так как $r < q$, то это и означает, что $\frac{p}{q}$ записывается бесконечной десятичной дробью, которая (начиная с k-й цифры) будет периодической, причем период содержит не более $q-1$ цифры.

В заключение расскажем о *малой теореме Ферма*. Пусть p — простое число и a — натуральное число, не делящееся на p. Рассмотрим числа $a, a^2, ..., a^p$. Это — зайцы. Каждое из этих чисел дает некоторый остаток при делении на p, который может быть равен $1, 2, ..., p-1$. Эти остатки — клетки. Так как число зайцев больше числа клеток, то найдутся два зайца, сидящие в одной и той же клетке. Иначе говоря, найдутся такие натуральные числа m и k, что a^m и a^{m+k} дают одинаковые остатки при делении на p, причем $m+k \le p$, т. е. $k \le p-1$. Это означает, что число $a^{m+k} - a^m = a^m(a^k - 1)$ делится на p, и потому $a^k - 1$ делится на p. Мы можем предполагать, что k — *наименьшее* натуральное число, обладающее этим свойством (или выбрать среди таких k наименьшее). Тогда k является *делителем* числа $p-1$, т. е. $p-1 = kl$ при некотором натуральном l (читатель может попробовать доказать это, используя идею классификации, изложенную в п. 19; мы к этому вернемся при рассмотрении *конечных полей*). Из тождества

$$z^l - 1 = (z-1)(1 + z + ... + z^{l-2} + z^{l-1}) \tag{16}$$

при $z = a^k$ получаем $a^{p-1} - 1 = (a^k - 1)(1 + ...)$, откуда видно, $a^{p-1} - 1$ делится на p. Это и есть малая теорема Ферма: *Если p — простое число и a — натуральное число, не делящееся на p, то число $a^{p-1} - 1$ делится на p.*

Чтобы показать, как работает эта теорема, рассмотрим следующую задачу: *доказать, что* $10^{2001} + 1$ *делится на* 7. Число $p = 7$ —

простое, и потому, согласно малой теореме Ферма, число $10^6 - 1$ делится на 7. Так как 1998 делится на 6, $1998 = 6 \cdot 333$, то из равенства (16) при $z = 10^6$, $l = 333$ получаем, что $10^{1998} - 1$ делится на 7. Иначе говоря, 10^{1998} при делении на 7 дает остаток 1. Следовательно, число $10^{2001} = 10^{1998} \cdot 10^3$ дает при делении на 7 такой же остаток, как и число 10^3. Поэтому задача принимает вид: доказать, что $10^3 + 1$ делится на 7, т. е. 1001 делится на 7. Это проверяется непосредственно.

Таким же образом решаются и другие аналогичные задачи. Например: доказать, что число $3^{100\,000} - 1$ делится на 11; доказать, что число $100^{999\,999} - 8$ делится на 17, и т. п.

Задачи и упражнения

111. В комнате собралось 5 человек. Докажите, что среди них найдутся двое, имеющие среди собравшихся одинаковое число знакомых.

112. В классе 31 ученик, всем вместе 434 года. Можно ли выбрать 20 учеников, которым вместе не меньше 280 лет?

113. Редакция журнала получила в январе 160 писем. Докажите, что в один из дней было получено больше 5 писем.

114. Докажите, что равносторонний треугольник нельзя покрыть двумя меньшими равносторонними треугольниками.

115. Можно ли написать 55 различных двузначных чисел так, чтобы сумма любых двух из них не равнялась 100?

Беседа 7. События и вероятности

31. События

Игра в кости, карточные и другие азартные игры издавна привлекали интерес некоторого круга людей. Естественно, основной вопрос состоял в том, как делать ставки в игре, какой сделать ход, чтобы выигрыш был наиболее вероятен. Именно поэтому развитие теории вероятности было связано, в первую очередь, с азартными играми, и в качестве иллюстраций использовались именно игровые ситуации. В середине XVII столетия Блез Паскаль, Пьер Ферма и другие математики заложили научные основы этой теории, не сомневаясь, что теория случайных явлений найдет важные приложения во многих сферах человеческой деятельности. Однако и сейчас при изложении основных понятий теории вероятности ситуации с бросанием костей, подбрасыванием монет или выбором нескольких карт из колоды служат удобным материалом для примеров и иллюстраций.

Игральная кость — кубик с округленными углами, на гранях которого нанесены точки, изображающие числа от 1 до 6 (рис. 181). В детских играх обычно используется одна кость, при игре в нарды — две, а при игре в кости — и большее их количество. «Выпавшей» считается верхняя грань брошенной кости (на рис. 181 выпала «пятерка»).

Бросание кости представляет собой опыт, испытание, а его результат (выпавшее число очков) — исход этого испытания. При бросании монеты могут быть два исхода: $\boldsymbol{O}$ — выпадение орла или

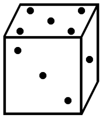


Рис. 181

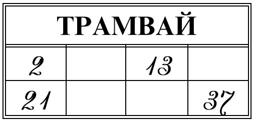


Рис. 182

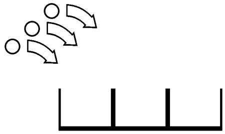


Рис. 183

Р — выпадение решетки («решка»). Если через данную остановку проходят четыре маршрута трамвая (рис. 182), то угадывание номера показавшегося вдали трамвая является (если расписание нам неизвестно) испытанием, которое может иметь 4 исхода. Если мы наудачу выбираем одного ученика из списка учащихся данного класса и интересуемся месяцем, в котором он родился, то мы производим испытание, которое может иметь 12 исходов.

Рассмотрим теперь более сложные примеры испытаний. Бросим две кости, которые будем считать различными, скажем, одна (первая) — синяя, а другая (вторая) — красная (или одну кость бросим дважды: первое бросание, второе бросание). В этом случае исходами являются пары (1; 3), (2; 5), (4; 4), (6; 3) и т. п. — всего 36 исходов. Если мы бросаем монету, то при трехкратном бросании имеются 8 возможных исходов: ***ООО***, ***ООР***, ***ОРО***, ***ОРР***, ***РОО***, ***РОР***, ***РРО***, ***РРР***. Вообще, при n-кратном бросании монеты имеются 2^n возможных исходов этой серии испытаний, а при n-кратном бросании кости возможны 6^n исходов (вспомните размещения с повторениями!). Наконец, рассмотрим еще серию примеров с шарами. Имеются три ящика — левый, средний, правый; мы берем три одинаковых (т. е. неразличимых) шара и наудачу бросаем их в эти ящики (рис. 183). У такого испытания имеются 10 исходов (рис. 184). Если же шары пронумерованы, то будет 3^3 исходов. Вообще, если k пронумерованных шаров наудачу брошены в n пронумерованных ящиков, то имеется n^k исходов (размещения с повторениями).

И еще пример с шарами. В мешке лежат n пронумерованных шаров (или n шаров разного цвета), и мы наудачу (не глядя) берем из мешка k шаров. Это испытание имеет $\binom{n}{k}$ исходов (сочетания без повторений).

Вернемся снова к бросанию одной кости. Обозначим через $\boldsymbol{E}_1$ событие «выпадает единица», через $\boldsymbol{E}_2$ — событие «выпадает двойка»

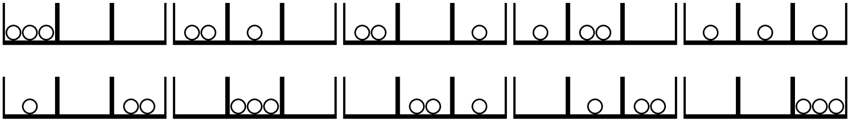


Рис. 184

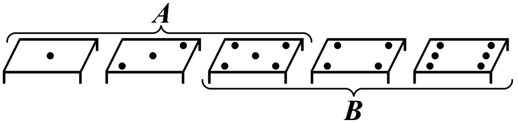


Рис. 185

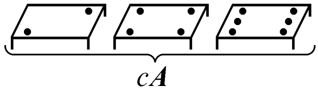


Рис. 186

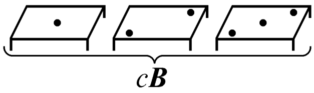


Рис. 187

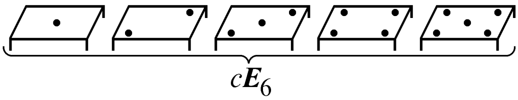


Рис. 188

и т. д. Эти события являются *случайными*: при бросании кости событие E_5 может в зависимости от случая либо произойти (выпадает пятерка), либо не произойти (выпадает другое число). Можно рассмотреть и другие события, скажем, событие A — выпадение нечетного числа, событие B — выпадение не менее четырех очков. События A и B *разложимы*: каждое из них может быть представлено в виде *объединения* некоторых событий E_1, E_2, ..., E_6 (рис. 185):

$$A = E_1 \cup E_3 \cup E_5;\; B = E_4 \cup E_5 \cup E_6.$$

События же E_1, E_2, ..., E_6 являются *элементарными* (неразложимыми), т. е. они не предствляются в виде объединения других событий. Вообще, объединение двух событий состоит в том, что произойдет *хотя бы одно* из них. Например, объединение рассмотренных выше событий A и B заключается в выпадении любого числа очков, кроме двойки (рис. 185).

Пересечение двух событий состоит в том, что наступят, произойдут *оба* эти события. Например, пересечение $A \cap B$ рассмотренных выше событий заключается в выпадении пятерки, т. е. $A \cap B = E_5$. Наконец, для каждого события определено его *дополнительное* (комплементарное) событие, которое, как и дополнение множества, мы будем обозначать буквой c. Например, cA означает, что *не наступит* событие A, т. е. выпадет четное число очков (рис. 186); cB означает выпадение *не больше трех* очков (рис. 187); cE_6 означает выпадение *менее шести* очков (рис. 188) и т. п.

Как видим, над событиями можно производить те же операции (объединение, пересечение, дополнение), что и над множествами. Это и неудивительно: если мы обозначим через U объединение всех элементарных событий:

$$U = E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6, \tag{1}$$

то любое событие будет *подмножеством* этого универсального множества $\boldsymbol{U}$, и потому для событий выполнимы операции $\cap$, $\cup$, c. Заметим, что само $\boldsymbol{U}$ является *достоверным событием*, т. е. это событие происходит в любом случае, при любом исходе испытания. Что же касается пустого множества $\varnothing$ (оно ведь тоже является подмножеством универсального множества $\boldsymbol{U}$), то оно является *невозможным событием*, т. е. оно не происходит, каков бы ни был исход испытания.

Мы подробно рассмотрели элементарные и разложимые события, а также операции над событиями для случая бросания одной кости. Возьмем теперь случай двух бросаний одной кости: первое бросание, второе бросание. В этом случае каждое элементарное событие представляет собой выпадение той или иной пары значений: (4; 6), (5; 5), (6; 2) и т. п. А вот суммарное выпадение 10 очков при двух бросаниях есть объединение трех элементарных событий: (4; 6), (5; 5) и (6; 4). Вообще, любое событие есть объединение нескольких элементарных, и для любых событий определены операции $\cap$, $\cup$, c.

Если мы выбираем (или, как говорят, осуществляем выборку) k шаров из имеющихся в мешке n пронумерованных шаров, то имеется $\binom{n}{k}$ возможных исходов испытания, т. е. $\binom{n}{k}$ элементарных событий. В качестве неэлементарных событий здесь могут быть, например, следующие: наличие шара № 1 в выборке; отсутствие шаров № 2 и № 3 в выборке; совпадение суммы номеров выбранных шаров с заданным числом q и т. д.

В приведенных нами примерах испытаний было возможно лишь конечное множество исходов, т. е. множество элементарных (неразложимых) событий было конечным. Все другие рассматриваемые события являлись объединением конечного числа элементарных событий. Задачи о случайных событиях при таких условиях проведения испытаний рассматриваются довольно часто, хотя возможны испытания и с бесконечным множеством исходов, т. е. бесконечным множеством элементарных (неразложимых) событий.

Задачи и упражнения

116. Из колоды в 36 игральных карт вытаскиваются две карты. Сколько есть различных способов сделать это? В скольких из них обе карты будут одной масти?

117. Пусть в условиях задачи 116 рассматриваются такие события: $\boldsymbol{A}$ — обе карты — трефы, $\boldsymbol{B}$ — обе карты — дамы, $\boldsymbol{C}$ — одна из карт — туз. Опишите события $\boldsymbol{A} \cup \boldsymbol{B}$, $\boldsymbol{A} \cap \boldsymbol{B}$, $\boldsymbol{A} \cap \boldsymbol{C}$.

118. Из комплекта домино вытаскиваются две кости. В скольких случаях сумма значений на костях будет равна 15?

119. В электрической цепи имеются четыре лампочки (рис. 189). Событие $\boldsymbol{A}_k$ — перегорела лампочка № k. Запишите событие, состоящее в том, что ток по цепи не может пройти.

120. При условиях задачи 119 запишите событие, состоящее в том, что перегорели ровно две лампочки, но ток по цепи может пройти.

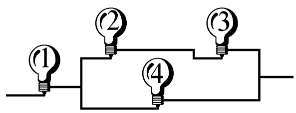


Рис. 189

32. Классическое понятие вероятности

Если монета ровная, неизогнутая, то можно ожидать, что при ее многократном бросании орел и решка будут выпадать одинаково часто, т. е. примерно в половине случаев будет выпадать орел, а в половине случаев — решка. Поэтому считают, что для такой монеты вероятность выпадения орла равна $\frac{1}{2}$ и вероятность выпадения решки тоже равна $\frac{1}{2}$.

Подобно этому естественно ожидать, что при многократном бросании игральной кости на долю каждого из чисел 1, 2, ..., 6 будет приходиться примерно шестая часть общего числа испытаний (т. е. бросаний кости). Поэтому считают, что вероятность выпадения каждой грани (т. е. каждого из чисел 1, 2, ..., 6) равна $\frac{1}{6}$. Разумеется, это следует ожидать только в случае идеально правильной кости: достаточно сместить центр тяжести кубика в сторону одной из граней, например, вставив внутрь кубика вблизи одной из граней свинцовый шарик, и кубик будет подобно игрушке «Ванька-встанька» наиболее часто падать на эту грань, т. е. одно из чисел 1, 2, ..., 6 будет выпадать чаще других.

Будем, однако, считать игральную кость идеальной. Какова в таком случае будет вероятность события $\boldsymbol{A}$ — выпадения нечетного числа очков? Легко понять, что эту вероятность следует считать равной $\frac{3}{6}$, т. е. $\frac{1}{2}$. Ведь всего имеется шесть возможных и одинаково вероятных исходов, т. е. наступления событий $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_6$, причем три из этих шести случаев (а именно $\boldsymbol{E}_1$, $\boldsymbol{E}_3$, $\boldsymbol{E}_5$) «благоприятствуют» событию $\boldsymbol{A}$, т. е. оно наступает в трех случаях из шести возможных. Событие же $\boldsymbol{C}$ («выпадение не менее пяти очков») имеет вероятность $\frac{2}{6}=\frac{1}{3}$, поскольку ему «благоприятствуют» два случая из шести (выпадение пятерки или шестерки). Вероятность события обычно обозначают буквой P (от слова probability — вероятность). Таким образом, в случае идеально правильной игральной кости справедливы равенства

$$P(\boldsymbol{E}_1) = P(\boldsymbol{E}_2) = \ldots = P(\boldsymbol{E}_6) = \frac{1}{6};\quad P(\boldsymbol{A}) = \frac{1}{2},\quad P(\boldsymbol{C}) = \frac{1}{3}.$$

Вообще, если при некоторых условиях проведения испытаний имеется n элементарных событий $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_n$ и если мы считаем эти элементарные события *равновероятными*, то каждое из этих элементарных событий имеет вероятность $\frac{1}{n}$:

$$P(\boldsymbol{E}_1) = P(\boldsymbol{E}_2) = \ldots = P(\boldsymbol{E}_n) = \frac{1}{n}.$$

Для события же, которому «благоприятствуют» k из этих элементарных событий (т. е. для события, которое является объединением k различных элементарных событий), вероятность его наступления (при каждом испытании) равна $\frac{k}{n}$. Иначе говоря, в условиях равновероятности исходов элементарных событий *вероятность любого события равна отношению числа благоприятных исходов к общему числу возможных исходов испытания*. Это и есть классическое определение вероятности. О роли предположения равновероятности мы еще поговорим в следующей беседе, а сейчас рассмотрим некоторые примеры.

Игральная кость бросается четыре раза подряд. Какова вероятность того, что ни разу не выпадет единица? Возможными исходами в этом испытании являются *четверки* чисел $(a; b; c; d)$, каждое из которых может принимать значения 1, 2, ..., 6. Все исходы (т. е. элементарные события) равновероятны, а число их равно $n = 6^4 = 1296$. Событию $\boldsymbol{M}$, состоящему в том, что ни разу не выпадет единица, соответствуют исходы, в которых каждое из чисел a, b, c, d может принимать *пять* значений 2, 3, 4, 5, 6. Число таких исходов равно $k = 5^4 = 625$. Таким образом, вероятность события $\boldsymbol{M}$ равна $P(\boldsymbol{M}) = \frac{5^4}{6^4} = \frac{625}{1296} = 0{,}482...$.

Дополнительному событию $c\boldsymbol{M}$, т. е. выпадению *хотя бы одной* единицы при четырех последовательных бросаниях кости, соответствуют остальные исходы, число которых равно $1296 - 625 = 671$. Следовательно, $P(c\boldsymbol{M}) = \frac{671}{1296} = 0{,}517...$

Как видим, выпадение единицы *хоть раз* при четырех бросаниях немного вероятнее, чем невыпадение единицы ни разу.

С рассмотренным примером связана интересная история (*парадокс де Мере*). Игрок Шевалье де Мере практиковал, в числе других игр, следующую. Он 24 раза бросал пару костей; если при этом хоть раз выпадут две единицы, де Мере получал определенный выигрыш (в зависимости от ставки). Своему партнеру де Мере предлагал бросить в ответ одновременно 4 кости; если при этом выпадала хотя бы одна единица, партнер получал такой же выигрыш. Де Мере полагал, что вероятность выигрыша для обоих *одинакова*, т. е. что игра является честной. Однако на практике оказалось, что де Мере чаще проигрывал, чем выигрывал, и он обвинял в этом математиков, считая их виновными в дезинформации.

Проведем соответствующие вычисления. Выпадение двух единиц при бросании двух костей есть событие, имеющее один шанс из 36. Дополнительному событию, т. е. невыпадению пары (1; 1), соответствуют 35 случаев из 36. Число исходов при 24 бросаниях пары костей равно 36^{24}, а число неблагоприятных исходов (пара единиц не выпадает ни разу) равно 35^{24}. Значит, вероятность того, что де Мере *не выигрывает*, равна

$$\frac{35^{24}}{36^{24}}=\left(\frac{35}{36}\right)^{24}=\left(\left(1-\frac{1}{36}\right)^{36}\right)^{2/3}. \tag{2}$$

Математики установили, что выражение $\left(1-\frac{1}{m}\right)^m$ при увеличении числа m все более приближается к некоторому пределу, который равен $\frac{1}{e}$, где $e = 2{,}7182818280...$ — число, играющее важную роль во многих разделах математики:

$$\lim_{m\to\infty}\left(1-\frac{1}{m}\right)^m=\frac{1}{e}. \tag{3}$$

Сравнивая (2) и (3), мы заключаем, что вероятность невыигрыша для де Мере примерно равна $\left(\frac{1}{e}\right)^{2/3}$.

Что же касается его партнера, то для него вероятность невыигрыша (ни на одной из четырех костей не выпала единица), равна, как мы выше видели,

$$\frac{5^4}{6^4}=\left(\frac{5}{6}\right)^4=\left(\left(1-\frac{1}{6}\right)^6\right)^{2/3}\approx\left(\frac{1}{e}\right)^{2/3}. \tag{4}$$

Выходит, что вероятность невыигрыша *примерно одинакова* для обоих (а потому и вероятность выигрыша примерно одинакова). Однако примерное равенство не есть *точное* равенство, и в этом кроется причина ошибки, которую допустил де Мере. В действительности выражение $\left(1-\frac{1}{m}\right)^m$ *возрастает* при увеличении числа m (в этом легко убедиться, взяв калькулятор и принимая $m = 2, 3, 4, 5, ...$). Поэтому, хотя каждое из чисел (2) и (4) примерно равно $\left(\frac{1}{e}\right)^{2/3}$, но все же число (4) *меньше* числа (2), т. е. вероятность невыигрыша у партнера чуть меньше, чем у самого де Мере.

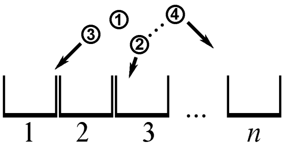


Рис. 190

В качестве второго примера рассмотрим следующую задачу о бросании шаров. Имеются n ящиков, в которые случайным образом бросаются n шаров (рис. 190); какова вероятность того, что в каждый ящик попадет один шар? Здесь каждое бросание всех n шаров представляет собой испытание, причем удобно считать, что шары пронумерованы и ящики пронумерованы. Число исходов испытания равно n^n (размещения с повторениями), а число благоприятных исходов (когда в каждый ящик попадает один шар, т. е. расположение шаров в ящиках представляет собой некоторую *перестановку* n шаров) равно $n!$ Следовательно, искомая вероятность равна $\frac{n!}{n^n}$. Например, при $n = 7$ вероятность оказывается равной

$$\frac{7!}{7^7} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{7 \cdot 7 \cdot 7 \cdot 7 \cdot 7 \cdot 7 \cdot 7} = 0{,}0061\ldots \ . \tag{5}$$

Это можно интерпретировать следующим образом. Предположим, что в некотором городе за каждую неделю (от понедельника до воскресенья) происходит 7 автомобильных ДТП (дорожно-транспортных происшествий). Естественно назвать неделю «нормальной», если в течение этой недели каждый день происходит ровно одно ДТП. Какова вероятность того, что наудачу взятая календарная неделя является нормальной? Это как раз задача о случайном бросании семи шаров (т. е. ДТП) в семь ящиков (дней недели), и потому ответ дается формулой (5). Как видим, эта вероятность весьма мала: всего 6 шансов из 1000. И 994 шанса из 1000 за то, что неделя не будет нормальной, т. е. найдется день, в который произойдет не менее двух ДТП (а значит, найдется и такой день, когда не будет ни одного ДТП).

В рассмотренном примере вероятность того, что в каждый ящик попадет один шар, оказалась равной $\frac{n!}{n^n}$, т. е. в ответе участвует факториал числа n. Во многих других случаях вычисление вероятности сводится к иным комбинаторным задачам, и в ответе участвуют сочетания или размещения, т. е. числа, также выражающиеся через факториалы (см. формулы (4) – (7) в предыдущей беседе). В связи с этим при вероятностных расчетах нередко приходится находить факториалы, и важным вычислительным средством является следующая *формула Стирлинга*, дающая приближенное значение факториала:

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n = (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n}. \tag{6}$$

Здесь приближенное равенство понимается в том смысле, что *отношение* выражений, стоящих слева и справа от знака $\approx$, стремится к единице при неограниченном возрастании n. Доказательство формулы Стирлинга мы рассмотрим впоследствии (оно проводится с помощью *интегралов*).

Заключительный пример относится к области биологии. В моравском городе Брно (столице Словакии) настоятель Августинского монастыря Грегор Мендель (1822 – 1884) занимался выращиванием садового гороха и других растений. Его аналитический, творческий ум заставлял относиться к работе вдумчиво, и Мендель сделал гениальное открытие, на несколько десятилетий опередившее свое время.

Строение цветка садового гороха приспособлено к самоопылению, т. е. пыльца от другого цветка не проникает к рыльцу, и семяпочки оплодотворяются только собственной пыльцой. Поэтому каждый сорт гороха размножается как *чистая линия*, с устойчивым сохранением признаков: например, имеются сорта с гладкими горошинами и сорта с морщинистыми горошинами. Однако возможно искусственное перекрестное оплодотворение: пыльник цветка удаляется, а потом на рыльце наносится пыльца *другого* растения.

Производя искусственное скрещивание гладкого и морщинистого сортов, Мендель заметил, что у получающихся *гибридов первого поколения* все семена были гладкими, морщинистость исчезала. Это означает, что свойство иметь гладкие семена является более сильным (как мы бы сейчас сказали, *доминантным*) признаком, а морщинистость — более слабым (*рецессивным*). На следующий год Мендель посадил эти гибридные семена и позволил выросшим растениям самоопыляться. К его удивлению примерно *четверть* получившихся горошин (*гибридов второго поколения*) оказались морщинистыми. Скрытый, рецессивный признак вновь проявился в отношении 1:3. Мендель провел опыты и с другими признаками (сорта с желтыми и сорта с зелеными горошинами и т. п.) — результат был таким же. И Мендель гениально предугадал механизм наследственности и его законы.

Признаки передаются от родителей «дискретными единицами наследственности» (*генами*, как мы теперь говорим, от слова *genus* — род). Для каждого простого признака растение имеет, как предположил Мендель, *две* единицы наследственности, скажем, WW (случай доминантной чистой линии) или ww (рецессивная чистая линия). Это показано на рис. 191. Каждое растение получает *одну* единицу наследственности от материнского растения (через семяпочку) и *одну* единицу от отцовского растения (через пыльцу). Поэтому все гибриды первого поколения имеют состав единиц (*генотип*) Ww. В силу доминантности преобладает W, т. е. горошины оказываются гладкими. Среди гибридов второго поколения, полученных на следующий год, имеются (примерно в равном количестве) растения с генотипами WW, Ww, wW, ww (рис. 192), поскольку родители передают каждому потомку либо ген W (с вероятностью 1/2), либо ген w (тоже с вероятностью 1/2). Но растения с генотипами WW, Ww, wW обладают *фенотипически* (т. е. внешне) доминантным признаком — они имеют

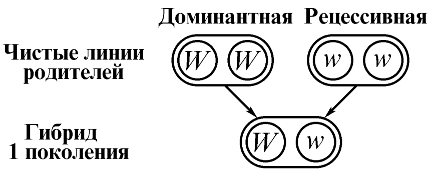


Рис. 191

		1 родитель	
		W	w
2 родитель	W	W W	w W
	w	W w	w w

Рис. 192

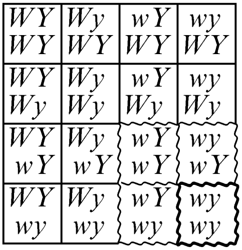


Рис. 193

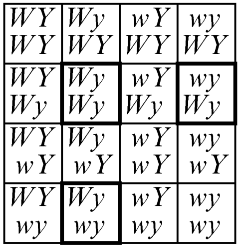


Рис. 194

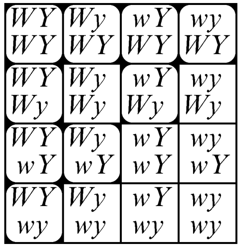


Рис. 195

гладкие горошины, и лишь растения *ww* обладают рецессивным признаком (морщинистые горошины). Это и дает менделевское отношение 3:1, т. е. примерно три четверти гибридов второго поколения имеют гладкие горошины и лишь четверть — морщинистые.

Мендель рассмотрел также одновременное наследование *двух* признаков (гладкие — морщинистые, а также желтые — зеленые) и получил среди гибридов второго поколения отношение 9:3:3:1. Это также хорошо подтверждается вероятностными соображениями. Желтизна горошин определяется доминантным геном Y, а зеленый цвет — рецессивным геном y. При этом признак «гладкость — морщинистость» не является *сцепленным* с признаком цветовой окраски — эти признаки наследуются *независимо* один от другого. Поэтому среди гибридов второго поколения, имеющих морщинистую форму горошин, будет то же распределение 3:1 по цвету, т. е. примерно три четверти из них имеют желтые горошины (доминантный признак Y), а четверть — зеленые (рис. 193). Иначе говоря, вероятность получения морщинистых зеленых горошин равна 1/16, а желтые морщинистые горошины получаются с вероятностью 3/16. Та же вероятность 3/16 соответствует гладким зеленым горошинам (рис. 194). И, наконец, в оставшихся 9 случаях из 16 горошины будут гладкие желтые (рис. 195). Это и дает менделевское распределение 9:3:3:1.

Работа Г. Менделя была доложена Брненскому обществу естествоиспытателей в 1865 году. Однако затем она была надолго забыта. Лишь в 1900 году его результаты были переоткрыты и подтверждены рядом других опытов в работах Августа Вейсмана, Хуго Де Фриза, Карла Корренса (которые затем «раскопали» работу Менделя и ссылались на его приоритет). К этому времени была уже детально исследована структура клетки и обнаружены содержащиеся в ядре *хромосомы* («окрашиваемые тела», выявленные при обработке препаратов клеток некоторыми красителями). При изучении поведения хромосом во время деления клеток была выдвинута гипотеза о том, что гены нанизаны на на хромосомные нити подобно бусинкам. А во второй половине нашего столетия была полностью расшифрована химическая структура молекул ДНК — двойных спиралей, составляющих основу хромосом. Были четко выделены те участки двойных спиралей, которые являются генами, и выяснено, что содержащуюся в них наследственную информацию можно записать с помощью *генетического кода* — алфавита, содержащего четыре буквы А, Г, Т, Ц (по первым буквам названий азотистых оснований, входящих в состав нуклеиновых кислот: аденин, гуанин, тимин, цитозин). При этом наследственная информация, содержащаяся в каждом гене, записывается в виде слова из сотен букв этого алфавита.

В каждой клетке (кроме *гамет* — половых клеток) хромосомы группируются *парами*, благодаря чему одному и тому же простому признаку соответствуют *два* гена (содержащихся по одному на каждой из «параллельных» хромосом). Такие гены называются *аллельными* между собой. Это и есть две единицы наследственности, предугаданные Менделем. При образовании гаметы происходит расщепление, и в нее случайным образом попадает одна хромосома из каждой пары, что также было предугадано Менделем. В результате оплодотворения (слияния мужской и женской гамет) возникает *зигота*, т. е. клетка, содержащая полный (двойной) набор *хромосом*, из которой и развивается индивидуум. Например, в клетках садового гороха (Pisum sativum), с которым производил свои опыты Мендель, есть семь пар хромосом. В одной из этих пар имеются, наряду с другими, два гена, контролирующие гладкость или морщинистость горошин и существующие в двух формах: W или w. Если хотя бы один из этих генов находится в форме W, то горошины гладкие. В другой паре хромосом имеются два гена, контролирующие цвет горошин; если хотя бы один из них присутствует в доминантной форме Y, то цвет горошин — желтый. Это и обусловливает правильность законов Менделя.

Задачи и упражнения

121. Два раза бросается игральная кость. Какова вероятность того, что сумма выпавших очков равна 7?

122. В коробке лежат 10 шаров: 3 красных, 3 синих и 4 зеленых. Из коробки случайным образом достали 2 шара. Какова вероятность, что они одинакового цвета?

123. Из комплекта домино убраны дубли. Случайно выбираются две кости. Какова вероятность, что их можно приложить друг к другу по правилам домино?

124. Что вероятнее: выиграть у равносильного противника 3 партии из 4 или 5 партий из 8? Ничьих не бывает.

125. Какова вероятность того, что четырехзначный номер автомашины содержит ровно две одинаковые цифры? (Мы считаем, что номер 0000 возможен.)

33. Свойства вероятности

Классическое определение вероятности исходит из предположения о *равновероятности* элементарных событий. Однако это предположение оправдывается не всегда. Рассмотрим, например, правильную четырехугольную пирамиду, на боковых гранях которой написаны числа 1, 2, 3, 4, а на основании — число 5 (рис. 196). Если эту пирамиду бросить на плоскость стола, то она упадет на какую-

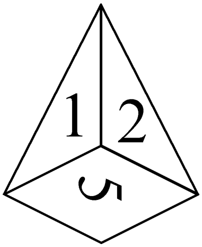


Рис. 196

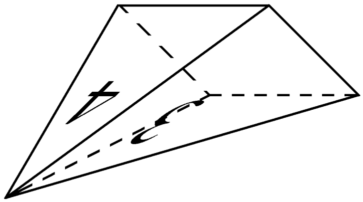


Рис. 197

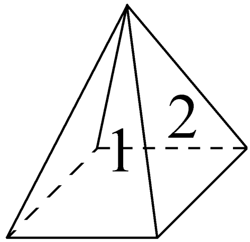


Рис. 198

либо грань, и написанное на этой грани число будем считать «выпавшим». Если пирамида выполнена из однородного материала и имеет идеальную геометрическую форму, то мы можем принять, что выпадение чисел 1, 2, 3, 4, написанных на боковых гранях (рис. 197), равновероятно. Однако выпадение числа 5, написанного на основании (рис. 198), будет более редким исходом, если высота пирамиды больше стороны основания. Поэтому следует принять, что элементарные события $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, $\boldsymbol{E}_3$, $\boldsymbol{E}_4$ (выпадение боковых граней с числами 1, 2, 3, 4) имеют одну и ту же вероятность, которая однако больше, чем вероятность элементарного события $\boldsymbol{E}_5$ (выпадение пятерки). Скажем, при определенной высоте пирамиды может оказаться, что каждое из событий $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, $\boldsymbol{E}_3$, $\boldsymbol{E}_4$ имеет вдвое большую вероятность, чем $\boldsymbol{E}_5$, т. е. вероятности этих событий соотносятся как 2:2:2:2:1. Иначе говоря,

$$P(\boldsymbol{E}_1) = P(\boldsymbol{E}_2) = P(\boldsymbol{E}_3) = P(\boldsymbol{E}_4) = \frac{2}{9},\ P(\boldsymbol{E}_5) = \frac{1}{9}.$$

Теперь мы можем сформулировать более общее понимание вероятности, чем то, которое содержится в классическом определении. Проводится некоторое испытание, имеющее n исходов, т. е. допускающее n элементарных событий $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_n$. По каким-либо соображениям (геометрическим, физическим или даже чисто произвольно) каждому элементарному событию $\boldsymbol{E}_i$ сопоставляется некоторое неотрицательное число $P(\boldsymbol{E}_i)$, называемое его вероятностью, причем числа эти выбраны так, что

$$P(\boldsymbol{E}_1) + P(\boldsymbol{E}_2) + \ldots + P(\boldsymbol{E}_n) = 1. \tag{7}$$

Далее, вероятностью произвольного разложимого события $\boldsymbol{A}$ называется *сумма* вероятностей тех элементарных событий, которые составляют событие $\boldsymbol{A}$:

$$P(\boldsymbol{A}) = \sum_i P(\boldsymbol{E}_i), \tag{8}$$

где суммирование распространяется на все элементарные события, содержащиеся в событии $\boldsymbol{A}$. В частности, для достоверного события $\boldsymbol{U} = \boldsymbol{E}_1 \cup \boldsymbol{E}_2 \cup \ldots \cup \boldsymbol{E}_n$ его вероятность, в силу определения (8), равна $P(\boldsymbol{E}_1) + P(\boldsymbol{E}_2) + \ldots + P(\boldsymbol{E}_n)$, т. е. $P(\boldsymbol{U}) = 1$ (см. (7)). Невозможное же событие $\varnothing$ имеет вероятность 0 (поскольку для такого события правая часть формулы (8) не содержит ни одного слагаемого). Для любого события $\boldsymbol{A}$ справедливо двойное неравенство $0 \le P(\boldsymbol{A}) \le 1$.

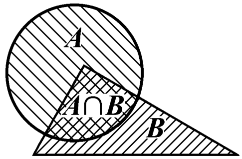


Рис. 199

Для любых двух событий $\boldsymbol{A}$, $\boldsymbol{B}$ справедливо соотношение

$$P(\boldsymbol{A} \cup \boldsymbol{B}) = P(\boldsymbol{A}) + P(\boldsymbol{B}) - P(\boldsymbol{A} \cap \boldsymbol{B}) \tag{9}$$

(рис. 199). Из этого вытекает, что для любых двух событий $\boldsymbol{A}$, $\boldsymbol{B}$ имеет место *неравенство Буля*

$$P(\boldsymbol{A} \cup \boldsymbol{B}) \le P(\boldsymbol{A}) + P(\boldsymbol{B})$$

(и аналогично для объединения не двух, а большего числа событий).

Если $\boldsymbol{A} \cap \boldsymbol{B} = \varnothing$, т. е. одновременное наступление обоих событий $\boldsymbol{A}$ и $\boldsymbol{B}$ невозможно, то $\boldsymbol{A}$ и $\boldsymbol{B}$ называются *несовместными событиями*. В этом случае формула (9) принимает более простой вид

$$P(\boldsymbol{A} \cup \boldsymbol{B}) = P(\boldsymbol{A}) + P(\boldsymbol{B}), \tag{10}$$

т. е. вероятность объединения двух несовместных событий равна *сумме* их вероятностей. То же верно для нескольких событий, если они *попарно несовместны*.

В частности, дополнительные события $\boldsymbol{A}$ и $c\boldsymbol{A}$ несовместны, т. е. к ним применима формула (10). Так как при этом их объединение $\boldsymbol{A} \cup (c\boldsymbol{A})$ является достоверным событием, т. е. имеет вероятность 1, то из (10) вытекает, что

$$P(c\boldsymbol{A}) = 1 - P(\boldsymbol{A}). \tag{11}$$

Рассмотрим пример, который относится к жизни *популяций* (сообществ). Кровь человека, помимо основных четырех групп, имеет еще одну характеристику — так называемый *резус-фактор*. Он может быть *положительным* или *отрицательным* и контролируется одним геном в доминантной форме R или рецессивной форме r. Отрицательный резус — рецессивный признак (генотип rr); генотипы RR и Rr соответствуют положительному резусу.

Известно, что отрицательным резусом стабильно обладает примерно 1/7 популяции. Можно ли по этим данным определить, какая часть населения обладает генотипом RR (*доминантная гемозигота*) и какая — генотипом Rr (*гетерозигота*)?

Для решения условимся соответствующие гены, имеющиеся у индивидуума, нумеровать: «первый» ген, «второй» ген. Вероятность того, что у случайно взятого индивидуума первый ген имеет рецессивную форму r, обозначим через p, т. е. p — частота, с которой встречается рецессивный ген r в популяции. Дополнительное событие (первый ген доминантен, т. е. имеет форму R) наступает с вероятностью $q = 1 - p$. Те же вероятности и для второго гена. Следовательно, вероятность того, что у случайно взятого индивидуума резус отрицателен, т. е. что он имеет генотип rr, равна p^2. Но по условию эта вероятность равна $\frac{1}{7} \approx 0{,}143$, т. е. $p^2 \approx 0{,}143$, и потому $p \approx \frac{1}{\sqrt{7}} \approx 0{,}378$. Таким образом, $q = 1 - p \approx 0{,}622$. Из этого вытекает, что вероятность генотипа RR равна $q^2 \approx 0{,}387$. Наконец, гетерозиготный генотип Rr (или, что то же самое, rR — ведь мы лишь условно обозначили гены, как «первый» и «второй») равна $2pq \approx 0{,}470$ (впрочем, это также вытекает из того, что сумма вероятностей всех трех генотипов равна единице).

Учет резус-фактора очень важен. Если роженица имеет *отрицательный* резус (генотип rr), а резус-фактор плода *положителен* (т. е. он имеет генотип Rr), то кровосмешение во время родов приводит к *гемолитической болезни* ребенка (и зачастую к его смерти). Поэтому при указанных генотипах применяется кесарево сечение, исключающее кровосмешение. Не менее важен учет резус-фактора и при переливании крови.

Наконец, еще один пример. В Бразилии встречаются случаи *ахейроподии* — ужасного генетического заболевания, при котором ребенок рождается без кистей рук и без стоп. Ахейроподия вызывается некоторым рецессивным геном, обозначим его буквой a — в отличие от соответствующего доминантного гена A. Статистика показывает, что ребенок, больной ахейроподией, рождается в одном случае из 10 000, т. е. с вероятностью 0,0001. Какая часть бразильской популяции обладает генотипом AA (доминантная гемозигота) и какая генотипом Aa (гетерозигота)?

Решение — дословно то же, что и в предыдущем примере, только число 1/7 надо заменить на 0,0001. Вероятность того, что у случайно взятого индивидуума «первый» ген рецессивен, т. е. имеет форму a, равна в этом случае $p = 0{,}01$; дополнительное событие (ген A) имеет вероятность $q = 0{,}99$. Следовательно, вероятность генотипа AA равна $q^2 \approx 0{,}98$; вероятность же гетерозиготного генотипа Aa равна $2pq \approx 0{,}02$. Таким образом, примерно два процента бразильцев (внешне нормальных) являются носителями опасного гена a.

Задачи и упражнения

126. Допустим, что имеется «неправильная» игральная кость, для которой вероятность выпадения каждой грани пропорциональна числу, указанному на этой грани. Чему равна вероятность выпадения грани с тройкой?

127. Из слова «наугад» выбирается наугад одна буква. Какова вероятность, что это будет гласная?

128. Десять человек садятся за круглый стол. Какова вероятность, что два данных человека окажутся рядом?

129. В «Спортлото» на карточке с числами от 1 до 49 игрок отмечает 6 чисел; при проведении тиража также определяются 6 «выигрышных» чисел; выигрывают карточки, в которых с результатами тиража совпали 3, 4, 5 или 6 чисел. Какова вероятность угадать 6 номеров из 49 при игре в «Спортлото»?

130. Из колоды в 36 карт вытащили 6 карт. Какова вероятность, что среди них: а) есть все четыре масти; б) нет ровно одной масти?

34. Условная вероятность

Начнем с рассмотрения примера. В классе 30 учеников, из них 10 являются участниками фотокружка. В этом классе 12 девочек, причем из них только 3 участвуют в работе фотокружка. При этих условиях вероятность того, что наудачу выбранный из списка ученик является участником фотокружка, равна $\frac{10}{30}=\frac{1}{3}$.

Допустим однако, мы заметили, наудачу выбрав фамилию ученика из списка, что это — девочка. Какова при этом условии вероятность того, что выбранный учащийся работает в фотокружке? Поскольку девочек в классе только 12, причем из них в фотокружке работает 3, то эта вероятность равна $\frac{3}{12}=\frac{1}{4}$. Таким образом, вероятность события $\boldsymbol{A}$ (выбранный ученик — участник кружка) равна $\frac{1}{3}$, но при условии, что произошло событие $\boldsymbol{C}$ (выбранный ученик — девочка), эта вероятность изменилась и стала равной $\frac{1}{4}$. Вероятность события $\boldsymbol{A}$ при условии, что наступило событие $\boldsymbol{C}$, обозначается через $P(\boldsymbol{A}/\boldsymbol{C})$ и называется *условной вероятностью*. В рассматриваемом случае эта вероятность равна $\frac{3}{12}=\frac{3/30}{12/30}$. Но ведь 3/30 — это вероятность события $\boldsymbol{A} \cap \boldsymbol{C}$ (т. е. что выбранный ученик и участвует в фотокружке, и является девочкой), а 12/30 — вероятность события $\boldsymbol{C}$ (рис. 200). Таким образом,

$$P(\boldsymbol{A}/\boldsymbol{C})=\frac{P(\boldsymbol{A} \cap \boldsymbol{C})}{P(\boldsymbol{C})}. \tag{12}$$

Мы рассмотрели испытание (выбор фамилии ученика из списка), имеющее 30 равновероятных исходов. По аналогии с этим и в общем случае (независимо от того, являются ли исходы испытания равновероятными или нет) формулу (12) считают определением условной вероятности, т. е. дробь, стоящую в правой части, принимают за вероятность события $\boldsymbol{A}$ при условии наступления события $\boldsymbol{C}$.

Рассмотрение условных вероятностей позволяет написать формулу полной вероятности, которая полезна при решении многих задач. Допустим, что при некотором испытании выделены события $\boldsymbol{C}_1$, $\boldsymbol{C}_2$, ..., $\boldsymbol{C}_n$, которые попарно несовместны (т. е. никакие два не могут наступить одновременно), а их объединение является достовер-

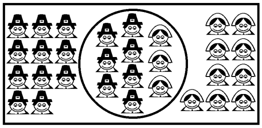


Рис. 200

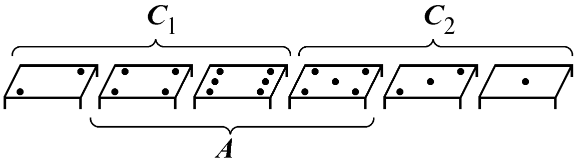


Рис. 201

ным событием (т. е. при проведении испытания какое-либо одно из событий C_1, C_2, ..., C_n непременно наступит). Тогда для любого события A мы имеем

$$A = A \cap (C_1 \cup C_2 \cup \dots \cup C_n) =$$
$$= (A \cap C_1) \cup (A \cap C_2) \cup \dots \cup (A \cap C_n).$$

Так как события $A \cap C_1$, $A \cap C_2$, ..., $A \cap C_n$ попарно несовместны, то их вероятности складываются, т. е.

$$P(A) = P(A \cap C_1) + P(A \cap C_2) + \dots + P(A \cap C_n).$$

Но, согласно (12), для каждого $i = 1, 2, \dots, n$ справедливо равенство $P(A \cap C_i) = P(A/C_i) \cdot P(C_i)$. Следовательно, предыдущее соотношение принимает вид

$$P(A) = P(A/C_1) \cdot P(C_1) + P(A/C_2) \cdot P(C_2) + \dots + P(A/C_n) \cdot P(C_n). \quad (13)$$

Это и есть формула *полной вероятности*.

Пусть, например, C_1 — выпадение четного числа очков при бросании игральной кости, а C_2 — выпадение нечетного числа очков. События C_1 и C_2 несовместны, а их объединение является достоверным событием. Через A обозначим событие, заключающееся в выпадении не менее четырех очков (т. е. выпадение четверки, пятерки или шестерки). Тогда $P(A) = \frac{1}{2}$, $P(C_1) = \frac{1}{2}$, $P(C_2) = \frac{1}{2}$. Далее, если известно, что выпало четное число очков (т. е. наступило событие C_1), то условная вероятность события A равна $\frac{2}{3}$, поскольку из трех случаев, благоприятствующих событию C_1 (выпадение двойки, четверки или шестерки), событие A имеет место в двух случаях (выпадение четверки или шестерки, рис. 201). Таким образом, $P(A/C_1) = \frac{2}{3}$. Аналогично $P(A/C_2) = \frac{1}{3}$. Подставляя эти значения в формулу (13), получаем соотношение $\frac{1}{2}=\frac{2}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{2}$ (как легко видеть, правильное). Это иллюстрирует формулу полной вероятности (13).

В качестве еще одного примера рассмотрим двухдетные семьи. Пол ребенка будем обозначать одной буквой М, Д (мальчик — девочка). Для простоты условимся считать, что рождение как мальчика, так и девочки имеет одну и ту же вероятность (равную $\frac{1}{2}$). Какова вероятность того, что оба ребенка — мальчики, если известно, что в семье есть мальчик? Может показаться интуитивно, что эта вероятность равна $\frac{1}{2}$. Ведь один ребенок — мальчик; значит другой будет с вероятностью $\frac{1}{2}$ тоже мальчиком.

Однако такое рассуждение ошибочно; точнее, оно дает решение совсем другой задачи. Если бы мы, например, знали, что старший ребенок — мальчик, то, действительно, вероятность наличия двух мальчиков была бы равна $\frac{1}{2}$. Но нам лишь сказано, что в семье есть мальчик, т. е. *хотя бы один* ребенок — мальчик (неизвестно, старший или младший), а это уже другая ситуация. Произведем точный расчет. Условимся первой буквой обозначать пол старшего ребенка, а второй буквой — пол младшего. Событие $\boldsymbol{A}$ — наличие *двух* мальчиков, событие $\boldsymbol{C}$ — наличие *хотя бы одного* мальчика. Наша задача заключается в нахождении условной вероятности $P(\boldsymbol{A}/\boldsymbol{C})$. Наудачу выбранной двухдетной семье соответствуют четыре возможных исхода ММ, МД, ДМ, ДД. Но нам известно, что наступило событие $\boldsymbol{C}$: в семье есть мальчик, т. е. остаются три исхода ММ, МД, ДМ. Эти исходы *равновероятны* (поскольку никакой дополнительной информации у нас нет). Таким образом, событию $\boldsymbol{A}$ (два мальчика) благоприятствует *один* случай из *трех*, т. е. условная вероятность $P(\boldsymbol{A}/\boldsymbol{C})$ равна $\frac{1}{3}$, (а не $\frac{1}{2}$). Иными словами, из двухдетных семей, имеющих хотя бы одного мальчика, примерно *одна треть* семей имеет двух мальчиков.

В заключение рассмотрим еще один пример из области генетики. В библейские времена браки между близкими родственниками (отца с дочерью, брата с сестрой) считались вполне респектабельными. Позднее браки между родственниками (даже брак двоюродных брата и сестры) стали редкими и представлялись нежелательными. Формула полной вероятности проливает свет на причину этого явления. Рассмотрим это на примере ахейроподии в бразильской популяции.

Допустим, родные брат и сестра (внешне здоровые, как и их родители) вступают в брак; какова вероятность события $\boldsymbol{B}$ — рождения у них больного ребенка (с генотипом *aa*)? Рассмотрение этого вопроса надо начать с генотипов родителей этих брата и сестры —

дедушки и бабушки рождающегося ребенка. В следующей таблице приведены возможные варианты их генотипов (события $\boldsymbol{C}_1$ — $\boldsymbol{C}_4$) и соответствующие вероятности (из п. 23):

Событие	Дедушка	Бабушка	Вероятность
$\boldsymbol{C}_1$	AA	AA	$0{,}98 \cdot 0{,}98 = 0{,}9604$
$\boldsymbol{C}_2$	Aa	AA	$0{,}02 \cdot 0{,}98 = 0{,}0196$
$\boldsymbol{C}_3$	AA	Aa	$0{,}98 \cdot 0{,}02 = 0{,}0196$
$\boldsymbol{C}_4$	Aa	Aa	$0{,}02 \cdot 0{,}02 = 0{,}0004$

При условии наступления события $\boldsymbol{C}_1$ брат и сестра имеют оба генотип AA и вероятность рождения у них больного ребенка равна нулю, т. е. $P(\boldsymbol{B}/\boldsymbol{C}_1) = 0$.

При условии наступления события $\boldsymbol{C}_2$ от бабушки достанется здоровый ген A, тогда как от дедушки с вероятностью $\frac{1}{2}$ (каждому) ген a. Если и брат, и сестра будут иметь генотип Aa (это произойдет с вероятностью $\frac{1}{4}$), то у них может родиться больной ребенок (с генотипом aa), причем с вероятностью $\frac{1}{4}$ (по Менделю, как на рис. 192). Таким образом, при условии наступления события $\boldsymbol{C}_2$ вероятность рождения больного ребенка у брата и сестры равна $\frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$, т. е. $P(\boldsymbol{B}/\boldsymbol{C}_2) = \frac{1}{16}$.

То же будет при условии наступления события $\boldsymbol{C}_3$: $P(\boldsymbol{B}/\boldsymbol{C}_3) = \frac{1}{16}$.

Наконец, при условии наступления события $\boldsymbol{C}_4$ вероятность того, что брат (внешне здоровый, т. е. имеющий хотя бы один ген A) имеет генотип AA, равна $\frac{1}{3}$ (как и в предыдущем примере — только вместо М и Д надо рассматривать A и a). Значит, дополнительное событие — генотип Aa у брата — имеет вероятность $\frac{2}{3}$. То же и для его сестры. Следовательно, вероятность того, что оба они (брат и сестра) имеют генотип Aa, равна $\frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9}$. При таких их генотипах у них опять с вероятностью $\frac{1}{4}$ родится больной ребенок. Таким образом, при условии наступления события $\boldsymbol{C}_4$ вероятность рождения больного ребенка у брата и сестры равна $\frac{4}{9} \cdot \frac{1}{4} = \frac{1}{9}$, т. е. $P(\boldsymbol{B}/\boldsymbol{C}_4) = \frac{1}{9}$.

Теперь по формуле полной вероятности находим:

$$P(\boldsymbol{B}) = P(\boldsymbol{B}/\boldsymbol{C}_1) + P(\boldsymbol{B}/\boldsymbol{C}_2) + P(\boldsymbol{B} / \boldsymbol{C}_3) + P(\boldsymbol{B}/\boldsymbol{C}_4) =$$

$$= 0{,}9604 \cdot 0 + 0{,}0196 \cdot \frac{1}{16} + 0{,}0196 \cdot \frac{1}{16} + 0{,}0004 \cdot \frac{1}{9} = 0{,}00289\ldots .$$

Таким образом, вместо обычной в популяции вероятности рождения больного ребенка $P = 0{,}0001$ (при случайных браках) у родных брата и сестры эта вероятность оказывается равной 0,00289 — почти в 29 раз больше! Аналогично проводятся расчеты для браков других близких родственников.

Сказанное относится не только к ахейроподии, но также к синдрому Дауна и другим генетическим заболеваниям. Именно этим объясняется вырождение у малочисленных изолированно живущих народностей (или в случаях, когда допускаются лишь узко кастовые браки). И имеющийся у некоторых северных народов обычай класть гостя в постель с женой хозяина является для них не аморальным, а оправданным, поскольку это способствует внесению здоровых генов в популяцию.

Выведем теперь *формулу Байеса*, имеющую важные приложения в теории вероятностей. Пусть $\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_n$ — полная группа попарно несовместных событий, т. е. событие $\boldsymbol{C}_1 \cup \boldsymbol{C}_2 \cup \ldots \cup \boldsymbol{C}_n$ достоверно, а каждые два различные события $\boldsymbol{C}_i$, $\boldsymbol{C}_j$ несовместны. Пусть, далее, $\boldsymbol{A}$ — какое-либо событие, и нас интересует условная вероятность $P(\boldsymbol{C}_i / \boldsymbol{A})$. Согласно формуле (12)

$$P(\boldsymbol{C}_i / \boldsymbol{A}) = \frac{P(\boldsymbol{C}_i \cap \boldsymbol{A})}{P(\boldsymbol{A})} = \frac{P(\boldsymbol{A} / \boldsymbol{C}_i) \cdot P(\boldsymbol{C}_i)}{P(\boldsymbol{A})}.$$

Если теперь записать $P(\boldsymbol{A})$ по формуле полной вероятности (13), то мы получим

$$P(\boldsymbol{C}_i / \boldsymbol{A}) = \frac{P(\boldsymbol{A} / \boldsymbol{C}_i) \cdot P(\boldsymbol{C}_i)}{P(\boldsymbol{A} / \boldsymbol{C}_1) \cdot P(\boldsymbol{C}_1) + P(\boldsymbol{A} / \boldsymbol{C}_2) \cdot P(\boldsymbol{C}_2) + \ldots + P(\boldsymbol{A} / \boldsymbol{C}_n) \cdot P(\boldsymbol{C}_n)}.$$

Это и есть формула Байеса. Ее применение мы рассмотрим в последующих задачах.

Задачи и упражнения

131. В письменном столе 6 ящиков. С вероятностью $\frac{1}{2}$ в столе лежит письмо. В первых четырех ящиках письма не оказалось. Какова вероятность, что оно лежит в пятом?

132. В школе три первых класса. В классе «А» 10 мальчиков и 10 девочек, в классе «Б» — 13 мальчиков и 8 девочек, в классе «В» — 7 мальчиков и 12 девочек. Какова вероятность, что встреченная вами первоклассница учится в классе «А»?

133. Три стрелка стреляли в одну мишень. В мишени оказалось две пробоины. Первый попадает с вероятностью 0,9, второй — с вероятностью 0,8, третий — с вероятностью 0,7. Какова вероятность, что промахнулся именно третий?

134. В вазе лежали 10 конфет «Мишка» и 5 конфет «Каракум». Три конфеты взял младший брат, затем старший берет наугад одну конфету, не зная, какие именно конфеты взял младший. Какова вероятность, что он вытащит «Каракум»?

135. Игральная кость бросается два раза. Какова вероятность того, что сумма очков будет больше десяти, если при бросаниях ровно один раз выпадала шестерка? А при условии, что при первом бросании выпала шестерка?

35. Независимые события и серии испытаний

Мы уже знаем, что при определенных условиях происходит *сложение* вероятностей: если события $\boldsymbol{A}$ и $\boldsymbol{B}$ несовместны, то вероятность события $\boldsymbol{A} \cup \boldsymbol{B}$ равна сумме их вероятностей (формула (10)). А теперь мы рассмотрим случай умножения вероятностей. Будем говорить, что событие $\boldsymbol{A}$ не зависит от события $\boldsymbol{C}$, если вероятность события $\boldsymbol{A}$ остается одной и той же вне зависимости от того, произошло или не произошло событие $\boldsymbol{C}$, т. е. если $P(\boldsymbol{A}) = P(\boldsymbol{A}/\boldsymbol{C})$. Иначе говоря, событие $\boldsymbol{A}$ не зависит от $\boldsymbol{C}$, если $P(\boldsymbol{A}) = \dfrac{P(\boldsymbol{A} \cap \boldsymbol{C})}{P(\boldsymbol{C})}$ (см. формулу (12)), т. е. если выполнено соотношение

$$P(\boldsymbol{A} \cap \boldsymbol{C}) = P(\boldsymbol{A}) \cdot P(\boldsymbol{C}). \tag{14}$$

Это равенство *симметрично* относительно обоих событий $\boldsymbol{A}$ и $\boldsymbol{C}$, поэтому если $\boldsymbol{A}$ не зависит от события $\boldsymbol{C}$, то и $\boldsymbol{C}$ не зависит от $\boldsymbol{A}$. В связи с этим говорят короче: события $\boldsymbol{A}$ и $\boldsymbol{C}$ *независимы*.

Равенство (14) можно считать определением независимости, т. е. два события называются *независимыми*, если вероятность их одновременного наступления равна *произведению* вероятностей этих событий. Но как можно убедиться в независимости событий $\boldsymbol{A}$ и $\boldsymbol{C}$? Можно, конечно, отдельно вычислить вероятности всех трех событий $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{A} \cap \boldsymbol{C}$ и проверить, выполняется ли равенство (14). Однако чаще всего о независимости событий судят, исходя из некоторых общих соображений о том, что наступление события $\boldsymbol{C}$ «не влияет» на вероятность события $\boldsymbol{A}$.

Пусть, например, два раза подряд бросается монета. Обозначим через $\boldsymbol{C}$ выпадение орла при первом бросании, а через $\boldsymbol{A}$ — выпадение орла при втором бросании. Мы считаем «из общих соображений», что наступление или ненаступление события $\boldsymbol{C}$, связанного с первым бросанием монеты, никак не сказывается на результате второго бросания (ведь после первого бросания с монетой ничего не произошло, и мы отдельно, независимо бросаем ее второй раз). Иначе говоря, исход первого бросания и исход второго бросания не связаны друг с другом, *независимы*, т. е. первое и второе бросания происходят «при неизменных условиях». И так как при каждом отдельном бросании выпадение орла имеет вероятность $\frac{1}{2}$, то событие $\boldsymbol{A} \cap \boldsymbol{C}$, т. е. выпадение орла при обоих бросаниях, имеет, согласно (14), вероятность $P(\boldsymbol{A} \cap \boldsymbol{C}) = P(\boldsymbol{A}) \cdot P(\boldsymbol{C}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Как видим, мы здесь использовали формулу (14) не для проверки того, что события $\boldsymbol{A}$ и $\boldsymbol{C}$ независимы, а для того, чтобы *вычислить* вероятность события $\boldsymbol{A} \cap \boldsymbol{C}$, будучи заранее уверенными в том, что события $\boldsymbol{A}$ и $\boldsymbol{C}$ независимы.

Фактически мы уже несколько раз использовали эти соображения, не указывая явно, что речь идет о независимых событиях. Вспомним, например, вероятностное обоснование законов Менделя. Все гибриды первого поколения имеют генотип Ww. Каждое растение следующего (второго) поколения гибридов с вероятностью $\frac{1}{2}$ получает материнский ген w (через семяпочку) и *независимо от этого* с вероятностью $\frac{1}{2}$ получает отцовский ген w (пыльца). Значит, согласно формуле (14) вероятность появления растения с генотипом ww (морщинистые горошины) равна $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. А дополнительное событие (гладкие горошины) имеет согласно (11) вероятность $\frac{3}{4}$. Это и дает менделевское отношение 3:1.

Важным примером применения формулы умножения вероятностей являются *повторные независимые* испытания, рассмотренные выдающимся математиком Яковом Бернулли (1654 – 1705). Проводится серия: первое испытание, второе испытание, ..., n-е испытание, причем они проводятся при совершенно одинаковых условиях и их результаты никак не сказываются на последующих испытаниях. Например, можно рассматривать последовательность бросаний монеты (или последовательность бросаний игральной кости). Рассматриваются два дополнительных события $\boldsymbol{S}$ и $\boldsymbol{F}$, причем событие $\boldsymbol{S}$ («успех») имеет — при каждом испытании — вероятность p, а дополнительное событие $\boldsymbol{F} = c\boldsymbol{S}$ («неудача») имеет вероятность $q = 1 - p$. В случае бросания монеты можно принять за $\boldsymbol{S}$ выпадение орла, а за $\boldsymbol{F}$ — выпадение решки: тогда $p = \frac{1}{2}$ и $q = 1 - p = \frac{1}{2}$. При бросании игральной кости можно за $\boldsymbol{S}$ принять, например, выпадение единицы (успех), а за $\boldsymbol{F}$ — выпадение любого другого числа очков; в этом случае $p = \frac{1}{6}$, $q = 1 - p = \frac{5}{6}$. Испытания считаются независимыми, поэтому можно применять формулу умножения вероятностей.

Рассмотрим некоторые типичные задачи, связанные с серией независимых повторных испытаний. Вероятность того, что при всех n

испытаниях нам будет сопутствовать успех, равна p^n. В самом деле, вероятность успеха при каждом (первом, втором, ..., n-м) испытании равна p, а вероятность *пересечения* этих событий (т. е. наступления успеха при *всех* испытаниях) равна, в силу независимости испытаний, *произведению* их вероятностей, т. е. равна $p \cdot p \cdot \ldots \cdot p$ (где число множителей равно n).

Аналогично, вероятность того, что при всех n испытаниях нас будет преследовать неудача, равна q^n. Следовательно, дополнительное событие (т. е. наступление хотя бы одного успеха при n испытаниях) имеет вероятность $1 - q^n$.

Например, при бросании кости (успех — выпадение единицы, т. е. $p = \frac{1}{6}$, $q = \frac{5}{6}$) вероятность наступления успеха при всех n испытаниях (каждый раз выпадала единица) равна $\left(\frac{1}{6}\right)^n$. Вероятность неудачи при всех n испытаниях *(ни разу* не выпала единица) равна $\left(\frac{5}{6}\right)^n$. А дополнительное событие (при n бросаниях *хотя бы раз* выпала единица) имеет вероятность $1 - \left(\frac{5}{6}\right)^n$.

Если же бросается пара костей и успехом считается выпадение двух единиц, то вероятность успеха $p = \frac{1}{36}$, а вероятность неудачи $q = 1 - p = \frac{35}{36}$. В случае, когда пара костей последовательно бросается n раз, вероятность успеха при всех n испытаниях равна $\left(\frac{1}{36}\right)^n$, а вероятность неудачи при всех испытаниях (*ни разу* не выпала пара единиц) равна $\left(\frac{35}{36}\right)^n$. А дополнительное событие (при n бросаниях *хотя бы раз* выпала пара единиц) равна $1 - \left(\frac{35}{36}\right)^n$. В частности, при $n = 24$ вероятность неудачи при всех испытаниях равна (как мы видели при рассмотрении игры де Мере)

$$\left(\frac{35}{36}\right)^n = \left(1 - \frac{1}{36}\right)^{24} = \left(\left(1 - \frac{1}{36}\right)^{36}\right)^{2/3} \approx \left(\frac{1}{e}\right)^{2/3}.$$

Особенно интересна при рассмотрении испытаний Бернулли задача о том, с какой вероятностью при n испытаниях успех осуществится ровно k раз. Случаи $k = 0$ или $k = n$ были рассмотрены выше, а сейчас мы обратимся к рассмотрению произвольного k (где $0 \le k \le n$).

Сначала определим вероятность того, что *первым* k испытаниям будет сопутствовать успех, а всем *последующим* $n-k$ испытаниям — неудача. Так как успех или неудача при каждом испытании представляют собой события, *независимые* от результатов других испытаний, то, согласно формуле умножения (см. (14)), искомая вероятность равна $p \cdot p \cdot \ldots \cdot p \cdot q \cdot q \cdot \ldots \cdot q$, где первые k множителей равны p, а последующие $n-k$ множителей равны q, т. е. эта вероятность равна $p^k q^{n-k}$. Та же вероятность соответствует событию, которое состоит в том, что при испытаниях с заданными k номерами (скажем, первыми k нечетными) осуществится успех, а при всех остальных испытаниях — неудача. Общее же событие «n испытаний привели k раз к успеху и $n-k$ раз к неудаче» есть *объединение* стольких элементарных событий (рассмотренного типа), сколькими способами можно k букв S расставить на n местах (а на остальных местах поставить букву F). Это число способов равно $\binom{n}{k}$ (сочетания без повторений). Каждое из этих элементарных событий имеет вероятность $p^k q^{n-k}$, а друг с другом они несовместны. Поэтому по формуле сложения вероятностей (см. (10)) вероятность получить ровно k успехов при n испытаниях равна

$$P(k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k} \approx \sqrt{\frac{n}{2\pi k(n-k)}} \cdot \frac{n^n p^k q^{n-k}}{k^k (n-k)^{n-k}}. \quad (15)$$

Последнее выражение получается с помощью формулы Стирлинга (6).

В качестве иллюстрации этой формулы заметим, что сумма *всех* чисел $P(k)$ (по $k = 0, 1, \ldots, n$) должна быть равна единице, поскольку осуществление хоть какого-то числа успехов (от 0 до n) есть достоверное событие. И в самом деле, сумма всех чисел $P(k)$ по $k = 0, 1, \ldots, n$ согласно формуле бинома Ньютона (см. п. 28), равна $(p+q)^n = 1^n = 1$.

Рассмотрим пример. Монета бросается 10 раз. Какова вероятность того, что ровно 5 раз выпадает орел и 5 раз решка? Здесь «успех» (орел) имеет вероятность $p = \frac{1}{2}$, «неудача» (решка) тоже имеет вероятность $q = \frac{1}{2}$. В данном случае $n = 10$, $k = 5$. По формуле (15) находим искомую вероятность:

$$\binom{10}{5} p^5 q^5 = 252 \cdot \left(\frac{1}{2}\right)^{10} = \frac{63}{256} = 0{,}246\ldots\ .$$

Результат может показаться странным: ведь вероятность выпадения орла равна $\frac{1}{2}$ и естественно ожидать, что при 10 бросаниях выпадут как раз 5 орлов и 5 решек. А расчет дал для вероятности

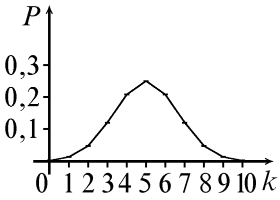


Рис. 202

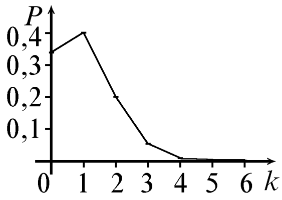


Рис. 203

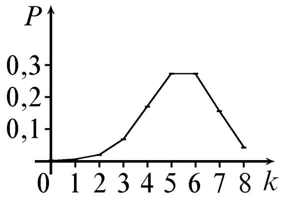


Рис. 204

этого события значение 0,246, т. е. примерно $\frac{1}{4}$. Иначе говоря, всего один шанс из четырех за то, что при 10 бросаниях монеты орел выпадет ровно 5 раз.

Аналогично вероятность выпадения орла ровно 4 раза (при десяти бросаниях) равна $\binom{10}{4}\cdot\left(\frac{1}{2}\right)^{10}=\frac{105}{512}=0{,}205...$. Такова же вероятность выпадения орла ровно 6 раз. Значит, вероятность того, что орел выпадет 4, 5 или 6 раз, примерно равна $0{,}205+0{,}246+0{,}205=0{,}656$, т. е. примерно равна $\frac{2}{3}$. Значит, имеется два шанса из трех за то, что при 10 бросаниях монеты орел выпадет 4, 5 или 6 раз.

На рис. 202 графически показаны вероятности выпадения орла k раз при десяти бросаниях монеты. «Колоколообразная» линия, соединяющая показанные точки, характерна для поведения вероятностей при испытаниях Бернулли.

В качестве второго примера рассмотрим шестикратное бросание игральной кости. «Успех» (выпадение единицы) имеет вероятность $p=\frac{1}{6}$, тогда как «неудача» (выпадение не менее двух очков) — вероятность $q=\frac{5}{6}$. Вероятность одного успеха при шести бросаниях ($n=6$, $k=1$) равна $\binom{6}{1}p^1q^5=6\cdot\left(\frac{1}{6}\right)\cdot\left(\frac{5}{6}\right)^5=\left(\frac{5}{6}\right)^5=0{,}401...$, т. е. эта вероятность примерно равна $\frac{2}{5}$. На рис. 203 графически показаны вероятности k успехов в этой задаче. И здесь наблюдается похожая колоколообразная кривая, но смещенная влево (это связано с тем, что вероятность p меньше половины). А если «успехом» при бросании кости мы будем считать выпадение не менее трех очков (здесь $p=\frac{2}{3}$), то при восьми бросаниях ($n=8$) мы получим картину, показанную на рис. 204. В этом случае колоколообразная кривая смещена *вправо*.

В заключение укажем формулу, которая была найдена известным французским математиком Симоном Пуассоном (1781 – 1840) и которая дает при определенных условиях приближенное значение вероятности (15). Условия, при которых удобно применять формулу Пуассона, следующие: k невелико по сравнению с n (это обозначают так: $k \ll n$, читается: «k много меньше n») и, кроме того, вероятность успеха p имеет вид $p = \frac{\lambda}{n}$, где λ также невелико по сравнению с n (т. е. $\lambda \ll n$). Используя соотношения $\binom{n}{k} = \frac{n(n-1)\,...\,(n-k+1)}{k!}$, $p = \frac{\lambda}{n}$, $q = 1 - \frac{\lambda}{n}$, мы можем переписать выражение (15) в виде

$$P(k) = \frac{n(n-1)\,...\,(n-k+1)}{n^k} \cdot \frac{1}{k!} \cdot \lambda^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

Здесь справа в первой дроби и в числителе, и в знаменателе стоят k множителей, примерно равных n, и потому эта дробь приближенно равна 1. Далее,

$$\left(1 - \frac{\lambda}{n}\right)^{n-k} = \left(\left(1 - \frac{\lambda}{n}\right)^{\frac{n}{\lambda}}\right)^{\frac{\lambda(n-k)}{n}} \approx \left(\frac{1}{e}\right)^{\frac{\lambda(n-k)}{n}} = \left(\left(\frac{1}{e}\right)^{\lambda}\right)^{\frac{n-k}{n}} \approx \left(\frac{1}{e}\right)^{\lambda} = e^{-\lambda}$$

(поскольку число $m = \frac{n}{\lambda}$ велико, а дробь $\frac{n-k}{n}$ близка к 1). Окончательно получаем: при $p = \frac{\lambda}{n}$, $k \ll n$, $\lambda \ll n$

$$\binom{n}{k} p^k q^{n-k} \approx \frac{1}{k!}\lambda^k e^{-\lambda}. \tag{16}$$

Это и есть формула Пуассона, дающая приближенное значение вероятности того, что при n последовательных испытаниях Бернулли «успех» реализуется k раз.

Рассмотрим примеры применения формулы Пуассона. Вернемся снова к шестикратному бросанию игральной кости, причем успехом будем по-прежнему считать выпадение единицы, т. е. $n = 6$, $p = \frac{1}{6}$. В формуле Пуассона принимается, что $p = \frac{\lambda}{n}$, т. е. в данном случае $\lambda = 1$. Теперь вероятности k успехов в этой задаче, вычисленные по формуле Пуассона, показаны на рис. 205 (точки), а для сравнения черточками показаны точные значения этих вероятностей (как на рис. 203). Мы видим, что для всех k значения вероятностей близки друг к другу, хотя вряд ли можно считать, что $n = 6$ «велико» по сравнению с $k = 0, 1, 2, 3, \ldots$. А при больших значениях n совпадение обеих частей приближенного равенства (16) еще более хорошее.

В качестве второго примера возьмем некоторую (случайно подобранную) группу из 100 бразильцев, которые все фенотипически нор-

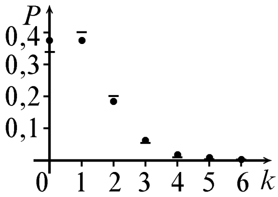


Рис. 205

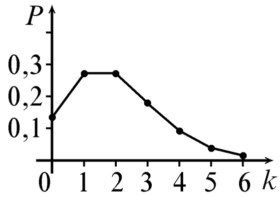


Рис. 206

мальны. Какова вероятность P_k того, что среди них имеются k лиц, имеющих генотип Aa, т. е. имеющих один рецессивный ген a, связанный с ахейроподией? Конечно, нас интересуют лишь небольшие значения k (скажем, $k = 0,\ 1,\ 2,\ 3,\ 4$).

Вероятность иметь ген a в данном случае, как мы видели, равна $p = 0{,}02$. Так как $n = 100$, то по формуле $p = \dfrac{\lambda}{n}$ мы должны принять $\lambda = 2$. Здесь, в самом деле, k и λ невелики по сравнению с n. Теперь по формуле Пуассона находим следующие приближенные значения вероятностей (рис. 206): $P_0 \approx 0{,}135$, $P_1 \approx 0{,}270$, $P_2 \approx 0{,}270$, $P_3 \approx 0{,}180$.

Рассмотрим еще один пример. В клубе 500 человек. Какова вероятность P_k того, что для k членов клуба их день рождения совпадает с праздником рождества (25 декабря)? Здесь вероятность «успеха» (т. е. совпадение дня рождения с праздником рождества) $p = \dfrac{1}{365}$, а число испытаний $n = 500$. Следовательно, $\lambda = pn = \dfrac{1}{365}\cdot 500 = 1{,}3699$, т. е. λ невелико по сравнению с n. По формуле (16) находим следующие приближенные значения вероятностей: $P_0 \approx \dfrac{1}{0!}\lambda^0 e^{-\lambda} = e^{-\lambda} \approx 0{,}254$, $P_1 \approx \dfrac{1}{1!}\lambda^1 e^{-\lambda} \approx 0{,}348$, $P_2 \approx 0{,}239$, $P_3 \approx 0{,}109$, $P_4 \approx 0{,}137$, ...

В частности, вероятность того, что *хотя бы один* из членов клуба родился в день рождества, равна $1 - P_0 \approx 0{,}746$, т. е. примерно три шанса из четырех за то, что кто-то из членов клуба отмечает день рождения в праздник рождества.

Наконец, приведем еще одну иллюстрацию. Допустим, что по статистике в час пик (12^{00} — 13^{00}) в городе происходит дорожно-транспортное происшествие (ДТП) с вероятностью 0,732. Тогда вероятность ДТП в данную минуту равна $p = \dfrac{1}{60}\cdot 0{,}732 = 0{,}0122$. Какова вероятность того, что в какую-то минуту произойдут одновременно два ДТП? А три? А ни одного ДТП?

Здесь $n = 60$ (исследуются 60 минут часа пик, т. е. последовательные испытания состоят в том, что берется первая минута, затем вторая, третья, ..., 60-я). В этой ситуации $\lambda = np = 60 \cdot 0{,}0122 = 0{,}732$, т. е. λ невелико по сравнению с n. По формуле (16) находим:

$$P_0 \approx e^{-\lambda} \approx 0{,}481, \quad P_1 \approx \frac{1}{1!}\lambda e^{-\lambda} \approx 0{,}352,$$

$$P_2 \approx \frac{1}{2!}\lambda^2 e^{-\lambda} \approx 0{,}129, \quad P_3 \approx \frac{1}{3!}\lambda^3 e^{-\lambda} \approx 0{,}03.$$

Аналогичным образом по формуле Пуассона рассчитываются частоты телефонных вызовов (если λ невелико, т. е. линия не очень загружена): вероятность P_0 того, что в данную минуту часа нет *ни одного* вызова, вероятность P_1 поступления ровно одного вызова в данную минуту, вероятность P_2 двух одновременных вызовов (и тогда если линия занята первым вызывающим, то второй слышит частые гудки, а в современных телефонах можно поговорить втроем) и т. д. В частности, $1 - P_0$ есть вероятность того, что в течение этой минуты будет хотя бы один вызов.

Задачи и упражнения

136. Вероятность попадания торпеды в цель равна 1/3. Сколько следует выпустить торпед, чтобы вероятность попадания в цель хотя бы одной была больше, чем 0,9?

137. На экзамене студенту задается 10 вопросов, на которые он должен ответить «да» или «нет». Какова вероятность, что он ответит правильно не меньше, чем на 7 вопросов, если он отвечает наугад?

138. Рейс из A в B совершают два самолета: двухмоторный и четырехмоторный. Вероятность выхода из строя каждого из моторов равна p. Двухмоторный самолет может продолжать полет и на одном моторе, а четырехмоторный — на трех и на двух моторах, а на одном уже не может. При каких значениях p безопаснее лететь на двухмоторном самолете?

139. В булочке с изюмом в среднем бывает 3 изюминки. Какова вероятность, что в купленной булочке не окажется изюминок?

140. Некоторую редкую группу крови имеет в среднем один человек на тысячу. У скольких людей нужно взять кровь на исследование, чтобы с вероятностью, большей 0,5, обнаружить человека с кровью этой группы?

Беседа 8. Случайные величины

36. Математическое ожидание и дисперсия

Сумма очков, выпавших на паре брошенных костей (скажем, синей и красной), может принимать значения 2, 3, ..., 12 с вероятностями, указанными в следующей таблице:

сумма очков	2	3	4	5	6	7	8	9	10	11	12
вероятность	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Например, сумме очков 5 соответствуют возможности (1; 4), (2; 3), (3; 2), (4; 1), каждая из которых реализуется с вероятностью 1/36, так что сумма очков 5 выпадает с вероятностью 4/36 = 1/9. Аналогично вычисляются и другие вероятности. Сумма очков, выпавших на паре брошенных костей, представляет собой *случайную величину*, которая может принимать значения 2, 3, ..., 12 с указанными в таблице вероятностями.

Если случайная величина X может в результате испытания принимать числовые значения x_1, x_2, ..., x_n с вероятностями p_1, p_2, ..., p_n, то число

$$E(x) = x_1 p_1 + x_2 p_2 + ... + x_n p_n \tag{1}$$

называется *математическим ожиданием* этой случайной величины (от слова Expectation —ожидание).

Поясним смысл этого понятия на следующем примере. Колесо рулетки имеет по окружности длину 1. Эта окружность разделена на n дуг, имеющих длины p_1, p_2, ..., p_n (так что $p_1 + p_2 + ... + p_n = 1$), и на этих дугах написаны некоторые положительные числа x_1, x_2, ..., x_n (рис. 207). Если игрок запускает рулетку (производит испытание), то вращающаяся стрелка укажет после остановки некоторую дугу, и написанное на этой дуге число будет означать *выигрыш* игрока. Здесь случайная величина X (выигрыш) может принимать значения x_1, x_2, ..., x_n с вероятностями p_1, p_2, ..., p_n. Число (1) означает *ожидаемый выигрыш* (или «средний» выигрыш).

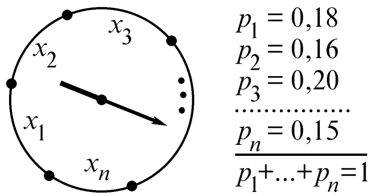


Рис. 207

Пусть теперь испытание $\boldsymbol{S}$ состоит в пятикратном бросании монеты, и число выпадений орла считается «выигрышем». Таким образом, рассматривается случайная величина S_5 (выигрыш), которая может принимать значения $k = 0, 1, 2, 3, 4, 5$ с вероятностями, которые несложно вычисляются по формуле (15), приведенной в предыдущей беседе:

число выпадений орла	0	1	2	3	4	5
вероятность	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Математическое ожидание числа выпадений орла, согласно (1), равно $E(S_5) = 0 \cdot \frac{1}{32} + 1 \cdot \frac{5}{32} + 2 \cdot \frac{10}{32} + 3 \cdot \frac{10}{32} + 4 \cdot \frac{5}{32} + 5 \cdot \frac{1}{32} = \frac{80}{32} = \frac{5}{2}$.

Результат этот очевиден: так как вероятность выпадения орла при каждом бросании равна $\frac{1}{2}$, то при пяти бросаниях естественно «ожидать» $5 \cdot \frac{1}{2}$ выпадений орла. А при n бросаниях монеты математическое ожидание числа выпадений орла равно $n \cdot \frac{1}{2}$.

В общем случае для серии испытаний Бернулли картина аналогична. При каждом из n испытаний событие $\boldsymbol{S}$ («успех») наступает с вероятностью p, а дополнительное событие («неудача») — с вероятностью $q = 1 - p$. Число успехов при проведении n испытаний есть случайная величина — обозначим ее через S_n, — которая может принимать значения с вероятностями, вычисляемыми по формуле (15) из предыдущей беседы. Следовательно, ее математическое ожидание равно

$$E(S_n) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k q^{n-k}$$

(сумма по всем $k = 0, 1, ..., n$). Читатель может проверить с помощью формулы бинома Ньютона, что $E(S_n) = np$. Впрочем, интуитивно этот результат понятен: вероятность успеха $\boldsymbol{S}$ при каждом испытании равна p, и потому при n испытаниях можно «ожидать» np успехов.

Рассмотрим, наконец, *пуассоновскую* случайную величину Y, которая может принимать значения $k = 0, 1, 2, ...$ с вероятностями

$$P_k = P(Y = k) = \frac{1}{k!} \lambda^k e^{-\lambda} \tag{2}$$

($k = 0, 1, 2, ...$; см. формулу (16) в предыдущей беседе). Здесь случайная величина Y может принимать *бесконечное* множество значений 0, 1, 2, ... Однако сумма их вероятностей по-прежнему равна единице:

$$P_0 + P_1 + P_2 + ... = e^{-\lambda} + \frac{\lambda}{1!} e^{-\lambda} + \frac{\lambda^2}{2!} e^{-\lambda} + \frac{\lambda^3}{3!} e^{-\lambda} + ... = 1.$$

Это вытекает из разложения функции e^x в ряд (доказываемого в математическом анализе): $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + ...$.

С помощью того же ряда читатель может вычислить математическое ожидание пуассоновской случайной величины Y. Оно оказывается равным $E(Y) = \lambda$. Впрочем, интуитивно это понятно. Мы производим n испытаний Бернулли, вероятность успеха при каждом испытании равна $p = \frac{\lambda}{n}$, где $\lambda << n$. Поэтому $E(S_n) = np = \lambda$. В пределе (при $n \to \infty$) случайная величина S_n превращается в пуассоновскую величину Y, и мы получаем $E(Y) = \lambda$.

В качестве иллюстрации вернемся к игре Шевалье Де Мере. Он 24 раза бросал пару костей и считал успехом выпадение двух единиц. Здесь вероятность успеха $p = \frac{1}{36}$. При $n = 24$ мы имеем $pn = \frac{1}{36} \cdot 24 = \frac{2}{3}$, т. е. число $\lambda = pn$, в самом деле, невелико по сравнению с n. Поэтому можно применять приближенную формулу (16) из предыдущей беседы. В частности, при $k = 0$ получается вероятность

$$P_0 \approx \frac{1}{0!} \lambda^0 e^{-\lambda} = e^{-\lambda} = e^{-\frac{2}{3}} = \left(\frac{1}{e}\right)^{\frac{2}{3}},$$

т. е. вероятность *невыигрыша* для Де Мере примерно равна $\left(\frac{1}{e}\right)^{\frac{2}{3}}$, как мы и видели раньше (см. соотношение (2) в предыдущей беседе). Далее, математическое ожидание количества выигрышей

$$E(S_n) \approx E(Y) = \lambda = \frac{2}{3}.$$

Проигрыши де Мере объяснялись не «ошибочностью» математики, а тем, что он не учитывал *приближенность* замены бернуллиевой случайной величины S_n предельной (пуассоновской) случайной величиной Y.

Второй важной характеристикой случайной величины является ее *дисперсия*, показывающая *разброс* возможных значений вокруг математического ожидания (т. е. «среднего» значения). Пусть случайная величина X имеет математическое ожидание $E(X) = \mu$ (см. (1)). Тогда случайная величина $X - \mu$ представляет собой *отклонение* величины X от ее среднего значения . Это отклонение может быть положительным, отрицательным или нулем. Квадрат же этого отклонения есть неотрицательная случайная величина; ее математическое ожидание обозначается через $D(X)$ и называется *дисперсией* случайной величины X:

$$D(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \ldots + p_n(x_n - \mu)^2. \quad (3)$$

Раскрывая скобки и учитывая равенство $p_1 + p_2 + \ldots + p_n = 1$ и формулу (1), находим формулу для вычисления дисперсии:

$$D(X) = p_1(x_1)^2 + p_2(x_2)^2 + \ldots + p_n(x_n)^2 - \mu^2 = E(X^2) - \mu^2. \quad (4)$$

В качестве примера снова рассмотрим пятикратное бросание монеты. Здесь, как мы видели, $E(S_5) = \mu = \frac{5}{2}$. Далее, квадрат $(S_5)^2$ случайной величины S_5 принимает значения 0^2, 1^2, 2^2, 3^2, 4^2, 5^2 с теми же вероятностями, что и раньше:

квадрат числа выпадений орла	0	1	4	9	16	25
вероятность	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

Поэтому $E((S_5)^2) = 0\cdot\frac{1}{32} + 1\cdot\frac{5}{32} + 4\cdot\frac{10}{32} + 9\cdot\frac{10}{32} + 16\cdot\frac{5}{32} + 25\cdot\frac{1}{32} = \frac{15}{2}$, и, следовательно, по формуле (4) дисперсия $D(S_5)$ равна $\frac{15}{2} - \left(\frac{5}{2}\right)^2 = \frac{5}{4}$. Это можно понять так: $D(S_5) = 5\cdot\frac{1}{2}\cdot\frac{1}{2} = npq$, где $n = 5$ есть число бросаний, а $p = q = \frac{1}{2}$ — вероятность успеха (и неудачи) при каждом бросании. Это справедливо и для любой серии испытаний Бернулли: случайная величина S_n, означающая число успехов при n испытаниях, в каждом из которых вероятность успеха равна p, имеет, как мы видели, математическое ожидание np; дисперсия же ее равна npq:

$$E(S_n) = np, \quad D(S_n) = npq. \tag{5}$$

Читатель может убедиться в этом, проводя тождественные преобразования с помощью формулы бинома Ньютона. Однако удобнее применить другой прием. Продифференцировав формулу бинома Ньютона $\sum\limits_{k=0}^{n} \binom{n}{k} x^k = (1 + x)^n$ (сумма по $k = 0, 1, \ldots, n$), а затем умножив результат на x, находим $\sum\limits_{k=0}^{n} k\binom{n}{k} x^k = nx(1 + x)^{n-1}$. Снова применяя ту же операцию (дифференцирование с последующим умножением на x), после упрощений получаем $\sum\limits_{k=0}^{n} k^2\binom{n}{k} x^k = nx(1 + x)^{n-2} + n^2x^2(1 + x)^{n-2}$.

Если теперь в двух полученных соотношениях положить $x = \frac{p}{q}$, а затем умножить результаты на q^n (где $q = 1 - p$), то мы найдем:

$$\sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} = np, \quad \sum_{k=0}^{n} k^2\binom{n}{k} p^k q^{n-k} = npq + n^2p^2.$$

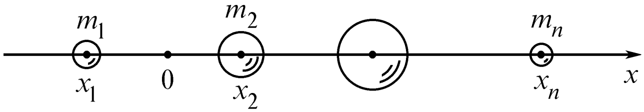


Рис. 208

Первое из этих соотношении означает, что $\mu = E(S_n) = np$ (см. (1)), а второе означает, что $E((S_n)^2) = npq + \mu^2$, т. е. $D(S_n) = npq$ (согласно (4)). Читатель, возможно, обратил внимание на то, что мы здесь использовали *производящую функцию* и ее дифференцирование.

Укажем, наконец, что для пуассоновской случайной величины Y (см. (2)) дисперсия, как и математическое ожидание, оказывается равной λ:

$$E(Y) = \lambda, \quad D(Y) = \lambda. \tag{6}$$

В этом можно убедиться с помощью ряда для e^x, причем и в этом случае удобно дважды применить ранее рассматривавшуюся операцию (дифференцирование с последующим умножением на x).

Смысл понятия дисперсии можно еще пояснить с помощью механической аналогии. Предположим, что единичная масса разделена на частичные массы $m_1, m_2, ..., m_n$, которые расположены на числовой оси в точках с координатами $x_1, x_2, ..., x_n$ (рис. 208). Тогда координата *центра тяжести* этой системы масс равна

$$\frac{m_1x_1 + m_2x_2 + ... + m_nx_n}{m_1 + m_2 + ... + m_n} = m_1x_1 + m_2x_2 + ... + m_nx_n = \mu.$$

Иначе говоря, если рассматривать $x_1, x_2, ..., x_n$ как *значения* случайной величины X, а $m_1, m_2, ..., m_n$ — как *вероятности* этих значений, то координата μ центра тяжести есть *математическое ожидание* рассматриваемой случайной величины: $E(X) = \mu$. Далее, согласно (3),

$$D(X) = m_1(x_1 - \mu)^2 + m_2(x_2 - \mu)^2 + ... + m_n(x_n - \mu)^2,$$

т. е. дисперсия $D(X)$ представляет собой момент инерции рассматриваемой системы масс относительно ее центра тяжести (который с механической точки зрения как раз характеризует разброс масс относительно центра тяжести системы).

Задачи и упражнения

141. Чему равняется математическое ожидание и дисперсия числа очков, выпадающих при бросании игрального кубика?

142. Ответьте на те же вопросы, если кубик имеет смещенный центр тяжести и вероятность выпадения каждой грани пропорциональна числу, указанному на этой грани.

143. Известна следующая игра. Игрок называет число от 1 до 6. Второй бросает три игральных кости (или три раза подряд одну кость). Если выпадает один раз названное число очков — первый получает от второго один доллар, при двух костях с этим числом очков — два доллара и при трех — три. Если названное число очков не выпадает, то первый платит один доллар второму. Чему равно математическое ожидание величины выигрыша первого игрока?

144. Найдите математическое ожидание величины выигрыша в «Спортлото».

145. Подбрасывается монета до тех пор, пока не выпадет герб. Найдите математическое ожидание количества подбрасываний монеты.

37. Нормальное распределение

На рассматривавшихся ранее рисунках 202 – 206 показаны «колоколообразные» кривые для испытаний Бернулли с различными n, p, q. Схожесть этих кривых естественно приводит к вопросу: нельзя ли указать одну «универсальную» колоколообразную кривую, из которой любая бернуллиева кривая получалась бы сдвигом и растяжениями? Утвердительный ответ на этот вопрос был получен в начале XVIII столетия английским математиком Муавром де Абрахамом и, в более усовершенствованной форме, французским математиком Пьером Лапласом в начале XIX столетия. В этом пункте мы поясним их результат.

Для удобства сначала укажем ту функцию, график которой служит искомой «универсальной кривой»:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \qquad (7)$$

ее график показан на рис. 209. Укажем некоторые свойства этой функции и ее графика.

График функции $\varphi(x)$ *симметричен относительно оси ординат*. Это непосредственно вытекает из того, что функция $\varphi(x)$ *четна* (поскольку в ее выражение входит лишь x^2, т. е. при подстановке x и $-x$ получается одно и то же значение функции).

Функция $\varphi(x)$ *достигает максимума в точке* $x = 0$ *и монотонно убывает до нуля при увеличении аргумента*. График функции $\varphi(x)$ представляет собой «колоколообразную» кривую. Это также очевидно, поскольку показатель степени $-\frac{x^2}{2}$ в формуле (7) отрицателен и возрастает по модулю до бесконечности при $x \to \infty$.

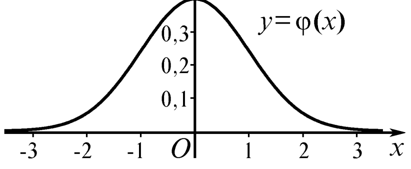


Рис. 209

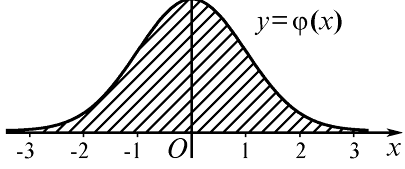


Рис. 210

Площадь, заключенная между графиком функции и осью абсцисс, равна 1 (рис. 210). Иначе говоря,

$$\int_{-\infty}^{\infty} \varphi(x)dx = 1. \qquad (8)$$

Теперь обратимся к теореме Муавра—Лапласа. Рассуждения будут носить интуитивный, поисковый характер. Пусть S_n — случайная величина, означающая число успехов при проведении n последовательных испытаний Бернулли, причем в каждом испытании вероятность успеха равна p. Как мы знаем, математическое ожидание этой случайной величины равно $E(S_n) = np$. Будем для простоты считать, что число np является целым, и найдем вероятность того, что число успехов будет равно np. Используя формулу (15) предыдущей беседы (при $k = np$, $n - k = nq$), мы находим:

$$P(np) \approx \sqrt{\frac{n}{2\pi \cdot np \cdot nq}} \cdot \frac{n^n \cdot p^{np} q^{nq}}{(np)^{np}(nq)^{nq}} = \frac{1}{\sqrt{2\pi npq}}.$$

Можно убедиться (при помощи той же формулы (15)), что если k отличается от np, то вероятность $P(k)$ наступления k успехов при n испытаниях будет *меньше*, чем $\frac{1}{\sqrt{2\pi npq}}$. Таким образом, функция $P(k)$ ($k = 0, 1, ..., n$) достигает наибольшего значения при $k = np$.

Теперь заметим, что правая часть формулы (15) определена *при всех* $k > 0$, а не только при целых (левую часть тоже можно доопределить при нецелых k, но это требует сложного математического аппарата). Рассмотрим функцию $P(x)$ и попытаемся сдвигами и растяжениями (сжатиями) приблизить ее к $\varphi(x)$. Функция $\varphi(x)$ достигает наибольшего значения при $x = 0$ (рис. 209), поэтому кривую, соответствующую испытаниям Бернулли (рис. 202–206), целесообразно для сравнения с $\varphi(x)$ сдвинуть влево на расстояние np, т. е. рассмотреть функцию $P^*(x) = P(x + np)$ (рис. 211, который построен для случая $n = 100$, $p = 0{,}25$). Тогда функция $P^*(x)$, как и функция $\varphi(x)$, будет иметь наибольшее значение при $x = 0$. Но наибольшее значение функции $P^*(x)$, т. е. $P^*(0) = P(np)$, как мы видели, равно $\frac{1}{\sqrt{2\pi npq}}$, тогда как наибольшее значение функции $\varphi(x)$ равно $\varphi(0) = \frac{1}{\sqrt{2\pi}}$. Следовательно, чтобы сделать наибольшие значения рассматриваемых функций одинаковыми, надо вместо $P^*(x)$ взять функцию $\sqrt{npq}\, P^*(x)$.

Далее, как мы знаем, $P(0) + P(1) + ... + P(n) = 1$, поскольку вероятность того, что *хоть какое-нибудь* число успехов (от 0 до n) наступит при n испытаниях, равна единице. Это означает, что площадь под

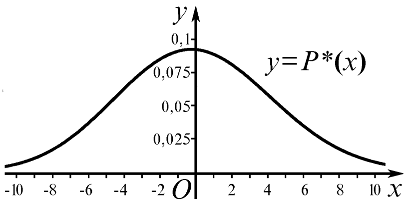


Рис. 211

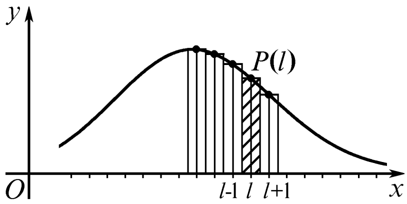


Рис. 212

графиком кривой $P(x)$ равна единице (на рис. 212 площадь заштрихованного прямоугольника равна $P(l)$). То же верно и для сдвинутой функции $P^*(x)$, а для функции $\sqrt{npq}\,P^*(x)$ это неверно — все ее ординаты в $\sqrt{npq}$ раз больше ординат функции $P^*(x)$, поэтому площадь под графиком функции $\sqrt{npq}\,P^*(x)$ равна $\sqrt{npq}$. Это можно исправить следующим образом: *сжать* график функции $\sqrt{npq}\,P^*(x)$ к оси ординат в $\sqrt{npq}$ раз, т. е. рассмотреть функцию

$$\sqrt{npq}\,P^*(x\sqrt{npq}) = \sqrt{npq}\,P(x\sqrt{npq} + np), \tag{9}$$

площадь под ее графиком равна единице, как и для функции $\varphi(x)$ (график функции $\sqrt{npq}\,P^*(x\sqrt{npq})$ приведен на рис. 213).

Итак, функция (9) имеет максимальное значение при $x = 0$, это максимальное значение равно $\frac{1}{\sqrt{2\pi}}$, как и для функции $\varphi(x)$, и площадь под графиком функции (9) равна 1, как и для функции $\varphi(x)$. Иначе говоря, обе функции «похожи». Это и дает геометрическую форму теоремы Муавра—Лапласа: *при $n \to \infty$ график функции* (9) *все более тесно прижимается к графику функции* $\varphi(x)$. Доказательство может быть получено предельным переходом (при $n \to \infty$) с помощью формулы (15).

А теперь сформулируем теорему Муавра—Лапласа в *интегральной форме*. Так как кривые (9) и $\varphi(x)$ близки, то площадь под кривой (9) на участке $a \le x \le b$ примерно равна аналогичной площади для кривой (7) (рис. 214), т. е. примерно равна $\int\limits_a^b \varphi(x)dx$. Но площадь под кривой (9) на указанном участке равна сумме вероятностей $P(l)$ по всем l, для которых $a\sqrt{npq} + np \le l \le b\sqrt{npq} + np$. Это ясно из рис. 212, на котором площадь заштрихованного прямоугольника равна $P(l)$. Иначе говоря, вероятность того, что значения случайной величины S_n (т. е. число успехов при n испытаниях) заключено между $a\sqrt{npq} + np$ и $b\sqrt{npq} + np$, примерно равна указанному выше интегралу (и стремится к этому интегралу при $n \to \infty$)

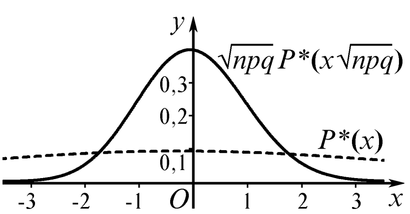


Рис. 213

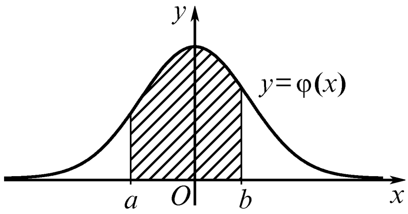


Рис. 214

$$P(a\sqrt{npq} + np \le S_n \le b\sqrt{npq}) + np) \approx \int_a^b \varphi(x)dx. \tag{10}$$

Это и есть *предельная теорема Муавра—Лапласа.*

Часто в приложениях используют функцию

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt, \tag{11}$$

т. е. через $\Phi(x)$ обозначают площадь под кривой $\varphi(t)$ на участке от $-\infty$ до x (рис. 215). Ясно тогда, что правая часть равенства (10) равна $\Phi(b) - \Phi(a)$ (рис. 214).

В заключение остановимся на понятии *распределения* случайной величины и, в частности, определим смысл термина *нормальное распределение*. Пусть X — случайная величина, которая может принимать конечное число значений $x_1, x_2, ..., x_n$. Говорят, что задано *распределение* случайной величины X, если заданы те вероятности $p_1, p_2, ...,$ p_n, с которыми величина X принимает эти значения: $P(X = x_1) = p_1$, $P(X = x_2) = p_2$, ..., $P(X = x_n) = p_n$. Например, случайная величина S_n, т. е. число успехов при n испытаниях Бернулли, имеет распределение $P(X = k) = \binom{n}{k}p^k q^{n-k}$ (см. формулу (15) в предыдущей беседе). Его называют *биномиальным* распределением, поскольку в его выражение входят биномиальные коэффициенты $\binom{n}{k}$.

Термин *распределение* употребляется и в том случае, когда случайная величина принимает бесконечное множество значений. Например, правая часть формулы (16), рассмотренной в предыдущей беседе, задает пуассоновское распределение $P(X = k) = \frac{1}{k!}\lambda^k e^{-\lambda}$, приближенно описывающее число успехов при большом числе испытаний Бернулли, когда вероятность успеха при каждом испытании имеет вид

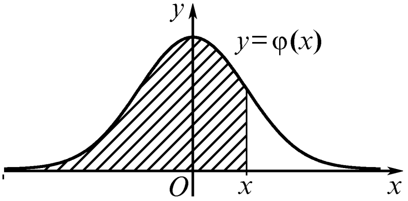


Рис. 215

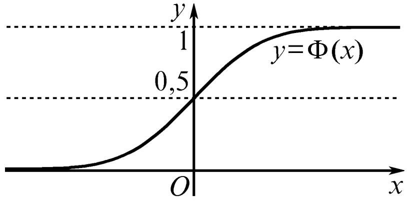


Рис. 216

$P=\frac{\lambda}{n}$, где λ мало в сравнении с n. Здесь случайная величина может принимать, вообще говоря, бесконечное (счетное) множество значений.

Но случайная величина может принимать и несчетное множество значений. Например, если мы бросаем шарик в длинную узкую (по ширине равную диаметру шарика) коробку, то точка касания шарика с дном коробки может оказаться в любой точке некоторого *отрезка* на основании коробки. Обычно такое *непрерывное* распределение можно описать с помощью функции $\varphi(x)$, которая представляет собой не вероятность того, что F точно примет значение x, а *плотность вероятности* в точке a. Говорят, что случайная величина F имеет *плотность вероятности* $\varphi(x)$, если вероятность того, что F примет значение, принадлежащее отрезку $[a; b]$, равна $\int\limits_a^b \varphi(x)dx$.

Таким образом, чем меньше отрезок $[a; b]$, тем меньше вероятность того, что F примет значение, принадлежащее этому отрезку. Вероятность того, что F примет *какое-нибудь* действительное значение, равна (согласно (8)) единице, как и следует ожидать, а вероятность того, что F примет *в точности заданное* значение a, равна $\int\limits_a^a \varphi(x)dx=0$.

Непрерывное распределение с плотностью $\varphi(x)$, определенное равенством (7), принято называть *нормальным распределением*.

Связь биноминального и нормального распределений (теорема Муавра—Лапласа) может быть в иной форме пояснена с помощью понятия *функции распределения*. Пусть X — некоторая случайная величина. Обозначим через $f(a)$ вероятность того, что случайная величина X примет значение, *не превосходящее* a: $f(a)=P(X\le a)$. Тогда $f(x)$, очевидно, представляет собой *неубывающую функцию*. Она называется *функцией распределения* случайной величины X. Для любой случайной величины X ее функция распределения удовлетворяет условиям $\lim\limits_{x\to-\infty} f(x)=0$, $\lim\limits_{x\to+\infty} f(x)=1$ (т. е. вероятность того, что x примет значение, «не превосходящее $+\infty$», равна единице).

Для нормального распределения соответствующая функция распределения равна $\Phi(x)$ (см. (11)), ее график показан на рис. 216. Теперь интегральная теорема Муавра—Лапласа может быть сформулирована следующим образом: *с точностью до сдвига и растяжения функция распределения для бернуллиевой случайной величины примерно совпадает с нормальной функцией распределения* $\Phi(x)$ (чем больше n, тем точнее совпадение).

Биноминальное, пуассоновское и нормальное распределения представляют собой наиболее важные распределения — как в самой теории вероятностей, так и в ее приложениях.

Задачи и упражнения

146. Постройте график функции распределения для случайной величины, равной числу очков, выпадающих на грани игрального кубика.

147. Постройте аналогичный график для кубика со смещенным центром тяжести (см. задачу 142).

148. Если бы проводилась игра «Спортлото» 3 из 6 (см. задачу 129), то как выглядела бы функция распределения для числа угаданных номеров в этой игре?

149. Объясните, почему многие инженеры считают, что при нормальном распределении значения, большие 3, практически не возникают.

150. При каких значениях a и b функция $f(x)=a\operatorname{arctg}x+b$ может быть функцией распределения некоторой случайной величины?

38. Закон больших чисел

Допустим, что производится n бросаний монеты. Рассчитывать, что успех (орел) осуществится ровно в половине случаев, мы, конечно, не вправе. Однако какова вероятность того, что частота выпадений орла будет заключена между 0,49 и 0,51, т. е. $\left|\frac{S_n}{n}-\frac{1}{2}\right|<0{,}01$? Ответ легко может быть получен с помощью интегральной теоремы Муавра—Лапласа.

Здесь вероятность успеха $p=\frac{1}{2}$ и легко найти те значения a, b, для которых (см. (10)) $\frac{a\sqrt{npq}+np}{n}=0{,}49$ и $\frac{b\sqrt{npq}+np}{n}=0{,}51$ (где $p=q=\frac{1}{2}$). Это будет при $a=-0{,}02\sqrt{n}$, $b=0{,}02\sqrt{n}$, т. е.

$$P\left(\left|\frac{S_n}{n}-\frac{1}{2}\right|<0{,}01\right)\approx\int\limits_{-0{,}02\sqrt{n}}^{0{,}02\sqrt{n}}\varphi(x)dx. \tag{12}$$

Из этого следует, что $P\left(\left|\frac{S_n}{n}-\frac{1}{2}\right|<0{,}01\right)\to 1$ при $n\to\infty$, поскольку при $n\to\infty$ пределы интегрирования в (12) захватывают все большую часть числовой прямой. Например, при $n = 10\,000$ вероятность (12) оказывается равной 0,9545 (это легко установить с помощью таблицы значений функции $\Phi(x)$, имеющейся во многих учебниках теории вероятностей), а при $n = 40\,000$ она становится равной 0,99994.

Аналогичный расчет показывает, что для испытаний Бернулли, в которых вероятность успеха при каждом испытании равна p, для любого $\varepsilon > 0$ справедливо соотношение

$$P\left(\left|\frac{S_n}{n}-p\right|<\varepsilon\right)\approx\int\limits_{-\varepsilon\sqrt{n}/\sqrt{pq}}^{\varepsilon\sqrt{n}/\sqrt{pq}}\varphi(x)dx\to 1 \text{ при } n\to\infty. \qquad (13)$$

Иными словами, *вероятность того, что средняя доля успехов при испытаниях Бернулли отклоняется от p более, чем на заданное положительное число* ε, *стремится к нулю при* $n\to\infty$.

Это утверждение известно под названием *закона больших чисел*, который оправдывает наше интуитивное представление о вероятности как о частоте наступления успеха.

Закон больших чисел можно осмыслить еще следующим образом. Обозначим через X_i случайную величину, означающую исход i-го испытания, $i = 1, \ldots, n$. Она принимает значение 1 (успех) с вероятностью p и значение 0 (неудача) с вероятностью $q = 1 - p$. Случайную величину X_i можно назвать *числом успехов при проведении i-го испытания*. Теперь ясно, что $S_n = X_1 + X_2 + \ldots + X_n$, т. е. число успехов в n испытаниях равно числу индексов i, для которых величина X_i принимает значение 1. Теперь закон больших чисел может быть сформулирован следующим образом.

Пусть $X_1, X_2, \ldots$ — *последовательность взаимно независимых случайных величин с одинаковыми распределениями и математическим ожиданием* $\mu = E(X_i)$. *Тогда для* $\varepsilon\to 0$

$$P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\mu\right|<\varepsilon\right)\to 1 \textit{ при } n\to\infty. \qquad (14)$$

Выше мы обосновали справедливость этого утверждения лишь в случае, когда X_i принимают лишь два возможные значения 1 и 0 с вероятностями p и $q = 1 - p$ соответственно (так что $\mu = p$). Но теперь это обстоятельство не постулируется, т. е. указанная формулировка относится к случаю, когда распределение случайной величины X *произвольно*. Тем не менее, и в такой более общей формулировке закон больших чисел оказывается справедливым. Это было установлено русским математиком А. Я. Хинчиным в 1929 году.

Пусть, например, рассматривается последовательность бросаний не монеты, а игральной кости, и случайная величина — число очков, выпавшее при i-м бросании. Здесь математическое ожидание величины X_i равно $E(X_i) = \frac{7}{2}$, а дробь $\frac{X_1 + X_2 + ... + X_n}{n}$ есть среднее число очков, выпавших после n бросаний. Закон больших чисел в формулировке Хинчина утверждает, что при большом n это среднее число очков с вероятностью, близкой к 1, будет близким к $\frac{7}{2}$. Если при этом дисперсия $D(X_i)$ конечна для *всех* i, то можно указать оценки, аналогичные (10); это было установлено Линденбергом в 1922 году. Однако в формулировке Хинчина (14) закон больших чисел справедлив без каких бы то ни было предположений о дисперсии.

Более того, был установлен закон больших чисел в *усиленной форме*. Для случая испытаний Бернулли он утверждает, что для любого $\varepsilon > 0$ с вероятностью 1 неравенство

$$\left| \frac{S_n}{n} - p \right| > \varepsilon \tag{15}$$

осуществляется при неограниченном возрастании n лишь для *конечного* числа испытаний, т. е. левая часть неравенства (15) принимает большие значения крайне редко. Этот закон был сформулирован и доказан Кантелли в 1917 году. Еще более сильный результат (так называемый *закон повторного логарифма*) был найден в 1924 году А. Я. Хинчиным, а затем еще усилен (в 1929 году) выдающимся русским математиком А. Н. Колмогоровым. Мы здесь эти результаты не формулируем; отметим лишь, что в формулировке Колмогорова допускается рассмотрение последовательности случайных величин X_1, X_2, ... с *различными* распределениями.

В заключение мы приведем одно приложение закона больших чисел, найденное в 1909 году известным французским математиком Борелем и впоследствии усиленное А. Я. Хинчиным.

Пусть x — действительное число, удовлетворяющее неравенствам $0 \le x < 1$, и пусть

$$x = 0{,}a_1 a_2 a_3 \ldots \tag{16}$$

— его десятичное разложение. Будем проводить последовательность независимых испытаний Бернулли, где n-е испытание состоит в том, что мы рассматриваем *n-ю цифру* числа x. Испытание будем считать успешным, если эта цифра равна 5. Таким образом, с каждым числом (16) связана бесконечная последовательность указанных испытаний Бернулли. Здесь при каждом испытании вероятность успеха равна

$p = \frac{1}{10}$ (поскольку 5 — одна из десяти возможных цифр 0, 1, ..., 9). Обозначим через $S_n(x)$ количество успехов в n испытаниях, т. е. количество пятерок среди первых n цифр числа x. Таким образом, отношение $\frac{S_n(x)}{n}$ есть *частота* появления пятерки среди первых n цифр числа x. Усиленный закон больших чисел утверждает, что *почти для всех чисел* x отношение $\frac{S_n(x)}{n}$ стремится к $\frac{1}{10}$ при $n \to \infty$. Здесь вместо пятерки можно взять любую другую цифру, т. е. почти для всех x частота появления каждой из десяти цифр стремится к $\frac{1}{10}$ при $n \to \infty$. Иначе говоря, почти все числа обладают тем свойством, что (при $n \to \infty$) в них *с одинаковой частотой*, равной $\frac{1}{10}$, встречаются все десять цифр. То же верно не только в десятичной системе, но и в системе счисления с любым другим основанием. Таким образом, почти все числа *нормальны*, т. е. в любой системе счисления в них все цифры встречаются одинаково часто (когда мы берем все большее и большее число цифр после запятой). Здесь термин «почти все» понимается в смысле *меры Лебега*, о которой мы поговорим в другой раз.

Так что можно сказать, что действительных чисел, не являющихся нормальными, «ничтожно мало». Однако, как это ни странно, указать *хотя бы одно* нормальное число не удается!

Беседа 9. Информация

39. Чет — нечет

Мы начнем с рассмотрения следующей известной задачи. Конь стоит в углу шахматной доски (рис. 217); может ли он попасть в противоположный угол, побывав ровно один раз на каждой клетке шахматной доски? Вообще, в любой ли клетке может конь завершить путь, побывав ровно один раз на каждой клетке?

Для решения этой задачи заметим, что если конь стоит на черной клетке, то следующим ходом он попадает на белую (рис. 218) и наоборот. Иначе говоря, после каждого хода коня цвет клетки, на которой он находится, меняется. Следовательно, после нечетного числа ходов цвет клетки меняется (конь попадает на белую клетку), а после четного конь снова попадает на черную: нечет (белая) — чет (черная) — нечет (белая) — чет (черная) — ... Но ведь если конь находится на первой клетке (черной), то, чтобы побывать по разу на всех 64 клетках, ему нужно сделать 63 хода, т. е. *нечетное* число ходов. Значит, побывав по разу на каждой клетке, конь завершит путь на белой клетке, т. е. попасть в противоположный угол (на черную клетку) он не сможет. Это дает *отрицательный* ответ на первый из поставленных вопросов.

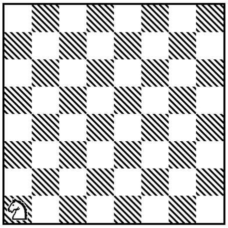


Рис. 217

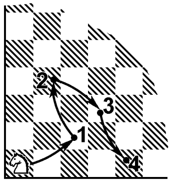


Рис. 218

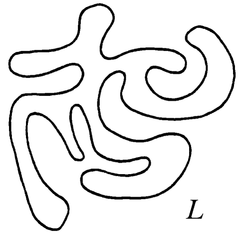


Рис. 219

А на второй вопрос проведенное рассуждение дает такой ответ: если и возможно, побывав на каждой клетке один раз, завершить путь в заданной клетке, то только в том случае, когда заданная клетка — белая. Иначе говоря, для возможности предписанного обхода доски *необходимо*, чтобы заданная конечная клетка была белой. В рассматриваемой задаче это необходимое условие является также и *достаточным* (что бывает далеко не всегда), т. е. для любой заданной белой клетки *существует* путь коня, обходящий по разу все клетки и оканчивающийся в заданной белой клетке. Однако это не вытекает из проведенного выше рассуждения, а требует отдельного обоснования.

Таким образом, чтобы знать, существует ли решение рассмотренной задачи, надо иметь ответ типа «да» — «нет» только на один вопрос: является ли белой заданная конечная клетка.

Рассмотрим еще один пример. На плоскости задана замкнутая линия L, не пересекающая себя. Какая бы она ни была сложная (рис. 219), эта линия разбивает плоскость на две области: внутреннюю (она заштрихована на рис. 220) и внешнюю. Две точки, принадлежащие одной и той же области, могут быть соединены ломаной, которая не пересекается с L (точки A и B на рис. 220), тогда как две точки, лежащие в разных областях, не могут быть соединены такой ломаной. Это утверждение (теорема Жордана) на самом деле справедливо для *любой* замкнутой линии, которая не пересекает себя, но об этом мы поговорим впоследствии, а сейчас ограничимся случаем, когда L — ломаная. Возьмем некоторую точку P, не лежащую на линии L (рис. 221). Как узнать, находится она во внутренней или во внешней области? Для этого нужно провести луч k, исходящий из точки P и не

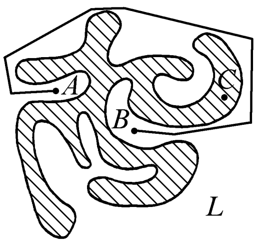


Рис. 220

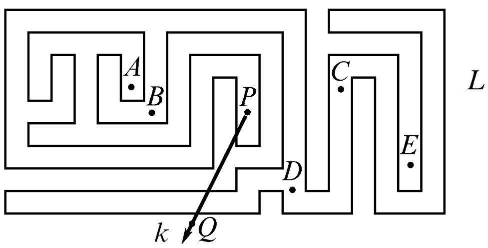


Рис. 221

проходящий ни через одну из вершин ломаной L. Точка Q, лежащая далеко на луче k (в бесконечности), принадлежит *внешней* области. Если же мы будем двигаться по лучу от точки P к Q, то при каждом пересечении линии L будем переходить из одной области в другую: нечет — чет — нечет — чет ... (внутренняя — внешняя — внутренняя — внешняя ...). Таким образом, если луч k пересекает линию L нечетное число раз, то точка P находится во внутренней области, если четное — во внешней. Иными словами, чтобы знать, в какой области находится точка P, надо иметь ответ типа «да» — «нет» только на один вопрос: является ли четным число точек пересечения луча k с ломаной L.

В качестве еще одного примера читатель может доказать, что любые две замкнутые ломаные линии на плоскости, у которых никакая вершина ломаной не принадлежит другой ломаной, имеют четное число точек пересечения.

Подобные задачи, в которых ответ меняется при каждом шаге (при каждом ходе коня, при каждом пересечении линии L и т. п.), т. е. определяется четностью числа шагов, встречаются нередко; их решение полностью определяется ответом типа «да» — «нет» только на один вопрос.

Задачи и упражнения

151. Можно ли покрыть шахматную доску доминошками (каждая закрывает ровно 2 клетки) так, чтобы остались свободными две угловые клетки на одной диагонали?

152. Можно ли разменять 25 рублей при помощи десяти купюр в 1, 3 и 5 рублей?

153. В ряд выписаны числа от 1 до 10. Можно ли расставить между ними знаки «+» и «–» так, чтобы значение полученного выражения было равно нулю?

154. Можно ли нарисовать 9-звенную ломаную так, чтобы каждое звено пересекалось ровно с одним из остальных звеньев?

155. По кругу стоят 99 корзин с арбузами. Может ли случиться, что в любой паре соседних корзин число арбузов отличается на единицу?

40. Количество двоичных цифр

Хорошо известна детская игра, в которой один играющий задумывает некоторый предмет или понятие, а другой должен отгадать его, задавая вопросы и получая на них только ответы «да» — «нет». Например: «живое?» — нет; «его едят?» — да; «его едят так, не поджаривая и не отваривая?» — да; «сладкое?» — нет; «белое?» — да; «это белый хлеб?» — да.

Однако в такой общей форме отгадывание может оказаться очень трудным делом. Как-то один из авторов предложил второму сыграть с ним в эту игру. «Живое?» — нет; «твердое?» — нет; «жидкое?» — нет; «газообразное?» — нет; «это абстрактное понятие?» — да; «черта характера?» — нет; «это какой-то процесс?» — да; «связанный с искусством?» — нет; «это связано с движением?» — да; «звук сопровождает этот процесс?» — нет; «процесс бесшумный?» — нет. Отгадывающий был поставлен в тупик, но продолжил вопросы. «Мы можем наблюдать этот процесс?» — нет. «Но все же это какой-то процесс?» — нет. Отгадывающий возмутился, поскольку ранее было указано, что задуман какой-то процесс. После еще нескольких противоречивых ответов он признал себя побежденным, и его коллега с улыбкой сообщил, что он задумал следующий процесс ответа на вопросы: если вопрос начинается с гласной буквы, он отвечает «да», а если с согласной — «нет».

Ясно, с чем здесь были связаны затруднения отгадывающего: не был четко очерчен круг задумываемых предметов, понятий, процессов. В более математизированном виде эта игра отгадывания может быть описана следующим образом. Имеется некоторое универсальное множество $\boldsymbol{U}$, состоящее, скажем, из 16 предметов. Сколько вопросов (с ответом «да» — «нет») следует задать, чтобы наверняка отгадать задуманный элемент этого множества? Легко понять, что следует применить следующую стратегию отгадывания. Мы разбиваем множество $\boldsymbol{U}$ на два подмножества $\boldsymbol{A}$ и $\boldsymbol{B}$, содержащие по 8 элементов, и спрашиваем: «Содержится ли задуманный элемент в множестве $\boldsymbol{A}$?». Любой ответ («да» или «нет») определяет *восемь* элементов, среди которых содержится задуманный. Теперь разделим эти восемь элементов на два подмножества $\boldsymbol{C}$ и $\boldsymbol{D}$, содержащие по 4 элемента, и зададим второй вопрос: «Содержится ли задуманный элемент в множестве $\boldsymbol{C}$?». Ответ определит четыре элемента, среди которых находится задуманный. Эти четыре элемента мы разбиваем на два подмножества по 2 элемента в каждом и, задав третий вопрос, определяем два элемента, среди которых находится задуманный. Наконец, мы выбираем один из этих двух элементов и задаем четвертый вопрос: «Этот ли элемент был задуман?» Получив ответ («да» или «нет»), мы сразу же находим задуманный элемент. Итак, задав 4 вопроса, мы отгадываем задуманное.

Эту стратегию отгадывания можно описать следующим образом. Будем для краткости обозначать ответ «да» цифрой 1, а ответ «нет» — цифрой 0, и запишем полученные ответы (на 4 вопроса) один за одним. Например, получив последовательно ответы «нет», «да», «да», «нет», мы можем записать это в виде 0110, т. е. в виде некоторого одного числа в двоичной системе счисления. Разным последовательностям ответов соответствуют разные двоичные числа от 0000 до 1111 (т. е. от 0 до 15) — всего 16 чисел. Это делает понятным, почему, задав 4 вопроса, мы сможем отгадать, какой из 16 предметов был задуман.

Можно и еще более математизировать игру отгадывания: мы просим занумеровать 16 элементов множества $\boldsymbol{U}$ двоичными числами от 0000 до 1111 и задаем 4 вопроса в следующей форме: «Равна ли единице первая цифра?», «Вторая цифра?», «Третья?», «Четвертая?». На каждый из таких вопросов мы получаем 1 бит информации (слово «бит» возникло из английского словосочетания binary digit — двоичная цифра). Итак, для отгадывания задуманного предмета (из 16 элементов множества $\boldsymbol{U}$) надо получить 4 бита информации.

А можно ли из 65 предметов отгадать задуманный, если мы задаем шесть вопросов (с ответами «да» — «нет»), т. е. получаем 6 битов информации? Шесть битов информации задают шестибуквенное слово, записываемое в алфавите $\{0, 1\}$. Например, 001101, 100111 и т. п. Всего таких слов имеется $2^6 = 64$, т. е. у нас в наличии 64 «клетки». А «кроликов», т. е. задумываемых предметов, имеется 65. Значит, как бы мы ни задавали вопросы, какие-нибудь 2 кролика окажутся в одной клетке, и потому заданных 6 вопросов может оказаться недостаточно, чтобы однозначно указать задуманный предмет. Если же в множестве $\boldsymbol{U}$ содержится не более чем 64 предмета, то шести битов информации достаточно для однозначного определения задуманного предмета.

Вообще, пусть множество $\boldsymbol{U}$ содержит q предметов. Сколько битов информации нужно, чтобы отгадать задуманный элемент множества $\boldsymbol{U}$? Рассмотренные выше примеры дают подход к решению этого вопроса.

Пусть k — такое натуральное число, что $2^{k-1} < q \le 2^k$. Тогда $k-1$ бит информации задает 2^{k-1} «клеток», т. е. $(k-1)$-буквенных слов из алфавита $\{0, 1\}$, а этих «клеток» недостаточно, чтобы рассадить в них q «кроликов», не более одного в каждой «клетке». Иначе говоря, $k-1$ битов информации может оказаться недостаточно для однозначного определения задуманного предмета. В то же время k битов информации достаточно для отгадывания. Например, чтобы определить задуманный предмет из 100 возможных, достаточно 7 битов информации, поскольку $2^6 < 100 \le 2^7$.

Неравенства $2^{k-1} < q \le 2^k$ означают, что $k-1 < \log_2 q \le k$, т. е. k — наименьшее натуральное число, для которого $k \ge \log_2 q$. Таким образом, чтобы отгадать задуманный элемент множества $\boldsymbol{U}$, содержащего q элементов, надо иметь не менее $\log_2 q$ битов информации. Например, $\log_2 100 \approx 6{,}664$, и так как нецелое числе вопросов с ответами типа «да» — «нет» мы задать не можем, то следует задать семь вопросов. Иначе говоря, нужно получить 7 битов информации для угадывания одного задуманного из 100 элементов.

В заключение рассмотрим еще два примера. На листе бумаги начерчена шахматная доска (рис. 222). Сколькими взмахами ножниц можно разрезать ее на отдельные клетки, если после каждого разреза (рис. 223) разрешается класть отдельные полученные куски друг на друга и затем производить новый разрез (рис. 224). Для решения этого вопроса заметим, что каждая клетка после очередного разреза остается либо слева, либо справа от ножниц. Иначе говоря, на вопрос, где была интересующая нас клетка после очередного разреза, мы даем ответ 0 или 1 («слева» — «справа»). Сколько разрезов, столько и битов информации. Так как число клеток равно 64 и $\log_2 64 = 6$, то для отделения всех клеток надо не менее 6 битов информации, т. е.

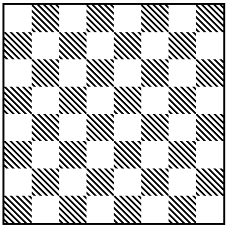


Рис. 222

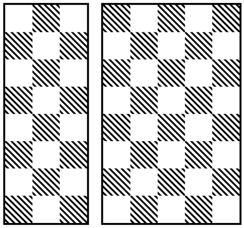


Рис. 223

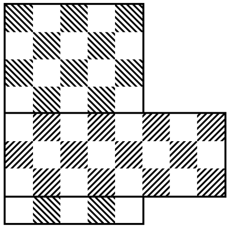


Рис. 224

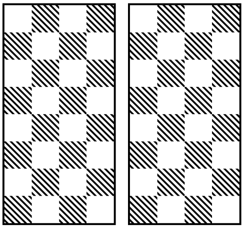


Рис. 225

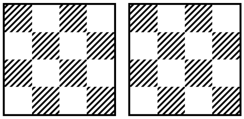


Рис. 226

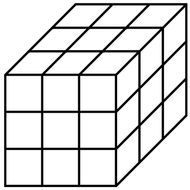


Рис. 227

не менее шести разрезов. Нетрудно убедиться, что шести разрезов и достаточно (первые два разреза показаны на рис. 225, 226).

В качестве второго примера возьмем куб 3×3×3 (рис. 227). Он состоит из 27 единичных кубов. Поставим задачу рассечь его на единичные кубы несколькими взмахами сабли (рис. 228), т. е. несколькими плоскими сечениями. При этом, как и в предыдущей задаче, разрешается после каждого разреза перекладывать части (рис. 229). Учитывая, что $\log_2 27 \approx 4{,}752$, мы заключаем, что потребуется не менее 5 битов информации, т. е. не менее пяти взмахов сабли. Однако в данном случае пяти взмахов недостаточно. В самом деле, рассмотрим *внутренний* единичный куб (расположенный в центре куба 3×3×3). Он имеет 6 граней, и любой взмах сабли позволяет открыть *лишь одну* из его граней. Поэтому, чтобы отделить внутренний единичный куб, нужно шесть взмахов.

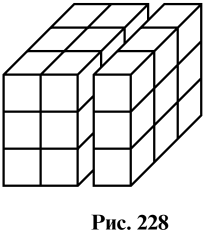


Рис. 228

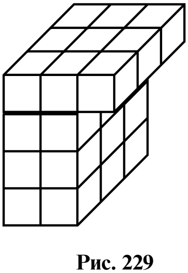


Рис. 229

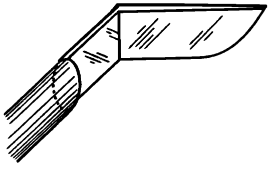


Рис. 230

Заметим, что если куб $3\times3\times3$ сделан из пластилина, а вместо сабли мы используем нож, лезвие которого загнуто под прямым углом (рис. 230), то пятью взмахами такого ножа *можно* разрезать куб на 27 единичных кубов. Плоских же разрезов (взмахов сабли) нужно, как мы видим, не 5, а 6.

Таким образом, подсчет информации дает лишь нижнюю оценку числа операций (ответов типа «да» — «нет», взмахов сабли и т. д.). Вопрос о том, является ли найденное число операций достаточным, требует отдельного рассуждения.

Задачи и упражнения

156. Среди 30 шариков имеется один радиоактивный. Имеется прибор, позволяющий установить для любой группы шариков наличие в ней радиоактивного. За какое наименьшее количество проверок всегда можно выделить радиоактивный шарик? А если радиоактивных шариков 2?

157. В начале книги, содержащей 1000 страниц, приведена теория и задачи, напечатанные крупным шрифтом, а дальше идут ответы, напечатанные мелким шрифтом. Мы пытаемся поскорее дойти до ответов, открывая книгу на некоторой странице. За сколько попыток можно найти страницу, с которой начинаются ответы?

158. Имеется предмет, масса которого равна целому числу граммов не более 200, и набор гирь в 1 г, 2 г, 4 г, 8 г, 16 г, 32 г, 64 г и 128 г. За какое наименьшее количество взвешиваний гарантированно можно найти массу этого предмета?

159. При каком наименьшем количестве гирь можно взвесить любой предмет массой от 1 г до 1 кг с точностью до 1 грамма? Гири можно класть лишь на одну из чашек весов.

160. Проводится турнир теннисистов по олимпийской системе (проигравший выбывает). Сколько туров будет в этом соревновании и сколько игр, если участвует 50 теннисистов?

41. Задачи на взвешивание

В задачном фольклорном творчестве студентов Московского университета имеется ряд задач о нахождении фальшивых монет при помощи взвешиваний на двухчашечных весах без гирь. Вот одна из этих задач.

Среди 27 монет одинакового достоинства одна — фальшивая, более тяжелая. С помощью трех взвешиваний на чашечных весах без гирь найти фальшивую монету.

Прежде чем переходить к решению, напомним, что вопросы с ответами «да» — «нет» позволяют разбить универсальное множество $\boldsymbol{U}$ на два подмножества $\boldsymbol{A}$, $\boldsymbol{B}$ и с помощью одного вопроса выяснить, в каком из этих подмножеств находится задуманный элемент. Именно с помощью деления на два подмножества решались задачи, рассмотренные в предыдущем пункте, и именно этим объясняется использование логарифмов *при основании* 2 для оценки необходимого числа битов информации.

Весы с двумя чашками (рис. 231) позволяют разбить рассматриваемое множество монет $\boldsymbol{U}$ на *три* подмножества:

$\boldsymbol{A}$ — монеты, помещаемые на левую чашку весов;

B — монеты, помещаемые на правую чашку;

C — монеты, оставленные в стороне (не взвешиваемые).

Рис. 231

Теперь понятно, что нам выгоднее всего распределить монеты на три *равные* кучки: 9 монет на одну чашку; 9 на другую и 9 в стороне. Если весы уравновесятся, то все монеты на весах настоящие, т. е. фальшивая монета находится среди 9 монет, оставленных в стороне. А если весы не уравновесятся, то фальшивая монета находится на более тяжелой чашке. Итак, после первого взвешивания мы выделим 9 монет, среди которых находится фальшивая. Теперь снова делим монеты на три равные кучки: 3 монеты на одну чашку, 3 на другую и 3 в стороне. В результате второго взвешивания мы выделим 3 монеты, среди которых находится фальшивая. А при третьем взвешивании мы кладем на каждую чашку по одной монете и одну оставляем в стороне. Это и позволит однозначно указать фальшивую монету, т. е. получить решение задачи.

Вообще, если даны 3^k монет, причем известно, что среди них имеется одна фальшивая (более тяжелая), то n взвешиваний на чашечных весах без гирь позволят однозначно указать фальшивую монету. То же справедливо для q монет, где $3^{k-1} < q \le 3^k$.

Сравним рассмотренную задачу с предыдущими. Как мы видели в предыдущем пункте, для угадывания одного задуманного из q элементов нужно получить $\log_2 q$ битов информации. В частности, при $q = 3^k$ мы находим, что для нахождения («угадывания») одной фальшивой (более тяжелой) монеты среди 3^k монет нужно иметь $\log_2 3^k = k \log_2 3$ битов информации. Так как нахождение фальшивой монеты может быть осуществлено k взвешиваниями, то это означает, что k взвешиваний дают не менее $k \cdot \log_2 3$ битов информации. Следовательно, каждое взвешивание дает $\log_2 3 \approx 1{,}584$ битов информации. Эта численная оценка позволяет рассматривать и другие задачи.

Пусть, например, даны 6 монет, из которых две — фальшивые (более тяжелые). Так как $\binom{6}{2} = 15$, то имеется 15 способов выбрать две монеты из шести, и надо найти *один* из этих 15 способов (т. е. выбрать две фальшивые монеты). Значит, надо иметь $\log_2 15 \approx 3{,}907$ битов информации. Поскольку 3 взвешивания дают $3 \cdot \log_2 3 \approx 4{,}752$ битов, то можно ожидать, что тремя взвешиваниями удастся найти обе фальшивые монеты. Читателю предоставляется в качестве упражнения найти способ это сделать.

Пусть теперь даны 6 монет, среди которых три — фальшивые (более тяжелые). Так как $\binom{6}{3} = 20$, то имеется 20 способов выбрать три монеты из шести и надо найти один из этих 20 способов (т. е. выбрать три фальшивые монеты).

Значит, надо иметь $\log_2 20 \approx 4{,}322$ битов информации. Поскольку 2 взвешивания дают примерно $2 \cdot \log_2 3 \approx 3{,}168$ битов, а 3 взвешивания дают $3 \cdot \log_2 3 \approx 4{,}752$ битов, то двух взвешиваний явно недостаточно, но можно ожидать, что тремя взвешиваниями удастся найти все три фальшивые монеты. Однако это сделать не удается.

В самом деле, если положить по две монеты на каждую чашку, и при этом чашки уравновесятся, то на каждой чашке и среди двух оставшихся в стороне монет имеется ровно по одной фальшивой. Но оставшимися двумя взвешиваниями не удается выявить по одной фальшивой монете из этих трех кучек. (И если взвешивать не по две, а по одной или по три монеты, то тоже ничего не получится.) Это еще раз подтверждает, что подсчет информации дает лишь нижнюю оценку числа операций (в данном случае — взвешиваний). Иными словами, подсчет информации указывает заведомо необходимое число операций, но не гарантирует достаточность этого числа операций для решения задачи.

В заключение рассмотрим еще одну задачу, также пришедшую из фольклора мехмата Московского университета.

Из 12 монет одинакового достоинства одна — фальшивая (отличающаяся по весу); найти фальшивую монету тремя взвешиваниями.

Отличие этой задачи от предыдущих состоит в том, что мы не знаем, тяжелее или легче фальшивая монета, чем каждая из остальных монет. Здесь всего имеется 24 возможности (любая из 12 монет может оказаться фальшивой, причем она может быть либо легче, либо тяжелее каждой из остальных). Поэтому для решения нужно получить не менее $\log_2 24 \approx 4{,}584$ битов информации. Так как 3 взвешивания дают 4,752 битов, то похоже, что тремя взвешиваниями удастся решить задачу. Это, действительно, удается сделать (хотя перебор различных возможностей, появляющихся при взвешивании, является не очень простым).

А что если число монет равно 13, а не 12? В этом случае число возможностей равно 26 (а не 24), и потому для решения нужно получить $\log_2 26 \approx 4{,}703$ битов информации. Так как 3 взвешивания дают 4,752 битов, то кажется правдоподобным, что и в этом случае задача может быть решена тремя взвешиваниями. Однако это не так!

Если, скажем, при первом взвешивании положить на обе чашки по 4 монеты, то может оказаться, что весы уравновесятся, а тогда фальшивая монета находится среди оставшихся в стороне пяти монет, что дает *десять* возможностей, и потому оставшимися двумя взвешиваниями выделить фальшивую монету не удастся. Это — еще один пример, показывающий, что подсчет информации дает лишь нижнюю оценку числа операций.

Задачи и упражнения

161. Имеется 5 грузиков разных масс и рычажные весы без гирь. Требуется расположить грузики по возрастанию массы. Сколько взвешиваний потребуется сделать?

162. Среди 10 одинаковых по виду монет могут быть одна или две фальшивые, отличающиеся от остальных по весу, но может и не быть ни одной фальшивой монеты. Все фальшивые монеты одинаковые, но легче они или тяжелее настоящих — неизвестно. Можно ли за три взвешивания узнать, есть ли среди этих 10 монет фальшивые и какие монеты тяжелее — фальшивые или настоящие?

163. Имеется 10 мешочков, в каждом из которых по 10 монет. В одном мешке все монеты фальшивые, а в остальных — настоящие. Настоящая монета весит 10 г, а фальшивая 9 г. Как за одно взвешивание определить мешочек с фальшивыми монетами, если пользоваться пружинными весами, указывающими вес с точностью до одного грамма?

164. Условие предыдущей задачи чуть изменили — в каждом из 10 мешочков по 9 монет. Можно ли теперь за одно взвешивание определить мешочек с фальшивыми монетами?

165. Условие предыдущей задачи тоже чуть изменили — в каждом из 10 мешочков по n монет. При каких n можно за одно взвешивание определить мешочек с фальшивыми монетами?

42. Понятие об энтропии

В предыдущих пунктах мы производили подсчет числа битов информации для некоторых комбинаторных задач. Здесь мы приведем вероятностную трактовку этих подсчетов.

Вспомним задачу об отгадывании одного задуманного предмета из 16 заданных предметов. Обозначим эту задачу через α. Как мы видели, для ее решения нужно получить 4 бита информации (4 ответа типа «да» — «нет»). Условимся в связи с этим говорить, что в задаче α *неопределённость* (или, как говорят, *энтропия*) равна четырем битам, и писать $H(\alpha) = 4$. Аналогично, для отгадывания одного из 2^k предметов нужно получить k бит информации, т. е. для этой задачи (обозначим ее через β) энтропия равна k бит: $H(\beta) = k$.

Вообще, если задано некоторое множество $\boldsymbol{U}$, содержащее q элементов, то задача γ об отгадывании одного задуманного элемента этого множества имеет энтропию $H(\gamma) = \log_2 q$. Иными словами, неопределенность (энтропия), вносимая *всеми* q элементами множества $\boldsymbol{U}$, равна $\log_2 q$, и потому можно считать, что каждый отдельный элемент множества вносит неопределенность, равную $\frac{1}{q}\log_2 q$, т. е. равную $-\frac{1}{q}\log_2\frac{1}{q}$. Заметим теперь, что каждый элемент множества

$\boldsymbol{U}$ с вероятностью $p = \frac{1}{q}$ может оказаться искомым («задуманным»). Таким образом, вместо $-\frac{1}{q}\log_2\frac{1}{q}$ можно написать $-p\log_2 p$, т. е. каждый отдельный элемент множества $\boldsymbol{U}$ вносит неопределенность, равную $-p\log_2 p$. Заметим, что число $-p\log_2 p$ *положительно*, поскольку $0 < p < 1$ и потому $\log_2 p < 0$.

Мы можем описать это следующим образом. Обозначим через $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_q$ элементарные события, состоящие в том, что был задуман первый, второй, ..., q-й элемент множества $\boldsymbol{U}$. События $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_q$ равновероятны, т. е. любое из них имеет вероятность $p = \frac{1}{q}$. Как мы отмечали выше, каждое отдельное событие вносит неопределенность $-p\log_2 p$, а суммарная неопределенность в q раз больше, т. е. она равна $q(-p\log_2 p) = pq\cdot(-\log_2 p) = -\log_2 p = \log_2 q = H(\gamma)$, как мы и видели выше.

Наконец, перейдем к общему случаю, в котором элементарные события не обязательно предполагаются равновероятными. Итак, пусть производится некоторое испытание γ и $\boldsymbol{E}_1$, $\boldsymbol{E}_2$, ..., $\boldsymbol{E}_q$ — элементарные события (исходы) этого испытания. Вероятности этих элементарных событий обозначим через p_1, p_2, ..., p_q. Тогда событие $\boldsymbol{E}_1$ вносит неопределенность $-p_1\log_2 p_1$, событие $\boldsymbol{E}_2$ вносит неопределенность $-p_2\log_2 p_2$ и т. д., т. е. общая неопределенность (энтропия) испытания равна

$$H(\gamma) = -p_1\log_2 p_1 - p_2\log_2 p_2 - \ldots - p_q\log_2 p_q. \qquad (1)$$

Приведенные рассуждения вовсе не служат «доказательством» формулы (1). Они приведены лишь с целью показать целесообразность такого понимания энтропии. Формулу (1) следует считать *определением* энтропии.

Например, в случае бросания монеты могут представиться *два* элементарных события $\boldsymbol{E}_1$, $\boldsymbol{E}_2$ (выпадение орла или решки), каждое из которых имеет вероятность $\frac{1}{2}$, т. е. $p_1 = p_2 = \frac{1}{2}$. Следовательно, по формуле (1) для этого испытания δ энтропия оказывается равной

$$H(\delta) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1,$$

т. е. неопределенность составляет один бит.

В качестве еще одного примера рассмотрим урну, в которой лежат 3 белых шара и 1 черный. Испытание ε заключается в том, что наудачу выбирается один из этих шаров. Здесь имеются два элементарных события E_1, E_2 — выбор белого или черного шара; вероятности их соответственно равны $\frac{3}{4}$ и $\frac{1}{4}$. Следовательно, по формуле (1)

$$H(\varepsilon) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} = -\frac{3}{4}(\log_2 3 - 2) + \frac{2}{4} = 2 - \frac{3}{4}\log_2 3 \approx 0{,}812.$$

Как видим, энтропия оказалась *меньше* одного бита, что и понятно, так как в урне больше белых шаров, чем черных, т. е. чаще при таком испытании должен появляться белый шар, и потому неопределенность исхода этого испытания меньше, чем 1 бит. Заметим, что если бы в урне все шары были белыми, то энтропия испытания (заключающегося в выборе из урны какого-либо шара) была бы равна *нулю*: всегда будет вынут *белый* шар, т. е. при таком испытании нет никакой неопределенности (или, иначе, исход опыта заранее известен).

Полезно отметить, что, согласно определению (формула (1)), энтропия представляет собой математическое ожидание неопределенностей отдельных исходов испытания, т. е. математическое ожидание случайной величины, которая принимает значения $-\log_2 p_1$, $-\log_2 p_2$, ..., $-\log_2 p_q$ с вероятностями p_1, p_2, ..., p_q.

Нетрудно проверить, что если испытания α и β *независимы*, то энтропия испытания $\alpha\beta$, заключающегося в последовательном (или одновременном) выполнении обоих испытаний α и β, равна *сумме* их энтропий:

$$H(\alpha\beta) = H(\alpha) + H(\beta). \tag{2}$$

Например, если α имеет q равновероятных исходов, а β имеет r равновероятных исходов, то испытание $\alpha\beta$ имеет qr равновероятных исходов, и потому $H(\alpha\beta) = \log_2 qr = \log_2 q + \log_2 r = H(\alpha) + H(\beta)$. Доказательство формулы (2) в общем случае можно предоставить читателю в качестве упражнения.

Если испытания α и β не являются независимыми, то можно определить *условную энтропию* $H(\beta/\alpha)$, т. е. неопределенность испытания β *при условии*, что произведено испытание α. Эта условная энтропия *не превосходит* энтропии испытания β, т. е. $0 \le H(\beta/\alpha) \le H(\beta)$. Формула (2) заменяется тогда соотношением

$$H(\alpha\beta) = H(\alpha) + H(\beta/\alpha).$$

которое справедливо в любом случае (т. е. без предположения о том, что испытания α и β независимы).

Соотношение $H(\beta/\alpha) \le H(\beta)$ не противоречит тому, что для *некоторых* исходов испытания α условная энтропия может быть и *больше*, чем $H(\beta)$. Рассмотрим следующий пример, который, впрочем, представляет собой лишь иную формулировку примера, приведенного в п. 34. Пусть испытание β означает рождение ребенка (здорового или больного ахейроподией) в семье бразильских молодоженов. Это испытание имеет два исхода: больной ребенок рождается с вероятностью $p_1 = 0{,}0001$, а здоровый — с вероятностью $p_2 = 0{,}9999$. Следовательно, энтропия испытания β равна

$$H(\beta) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 \approx 0{,}0005.$$

Эта энтропия очень мала; ведь исход «испытания» почти предопределен: 9999 шансов против 1 за то, что ребенок родится здоровым (в отношении ахейроподии). Теперь рассмотрим испытание α, заключающееся в том, чтобы выяснить, не являются ли молодожены родными братом и сестрой. Это испытание может иметь два исхода: $\boldsymbol{E}_1$— да, являются и $\boldsymbol{E}_2$ — нет, не являются, причем вероятность $p(\boldsymbol{E}_1)$ очень мала. Как мы видели, если для испытания α осуществляется исход $\boldsymbol{E}_1$, т. е. молодожены являются братом и сестрой, то в этом случае вероятность рождения больного ребенка равна $p_1' = 0{,}00289$, тогда как вероятность рождения здорового составляет $p_2' = 0{,}99711$, и потому

$$H(\beta/\boldsymbol{E}_1) = -p_1' \log_2 p_1' - p_2' \log_2 p_2' \approx 0{,}0292 .$$

Таким образом, энтропия $H(\beta/\boldsymbol{E}_1)$ больше энтропии $H(\beta)$, и это свидетельствует о возрастании неопределенности, т. е. о существенно большей вероятности рождения больного ребенка.

В заключение — несколько слов об истории возникновения понятия энтропии и о ее роли в различных прикладных вопросах. В своем первоначальном, простейшем виде понятие энтропии возникло в статистической механике и было введено Паулем Эренфестом. В этой модели рассматривается газ, находящийся в сосуде, который разделен на две части проницаемой пленкой. Если общее число молекул газа равно n, а в первой части сосуда имеется k молекул, то во второй части $n-k$ молекул. Распределить молекулы на такие две части можно $\binom{n}{k}$ способами, и каждое такое распределение можно рассматривать как исход некоторого испытания. Энтропия этого испытания равна $\log\binom{n}{k}$. В таком виде понятие энтропии (как меры неопределенности, беспорядка, хаоса) и было введено Эренфестом. Основание логарифма здесь не играет роли, поскольку при изменении основания число $\log\binom{n}{k}$ лишь умножается на некоторый множитель (модуль перехода от логарифмов при одном основании к логарифмам при другом основании), т. е. лишь меняется *единица*, с помощью которой мы измеряем энтропию. Всеобщее выравнивание температуры («тепловая смерть Вселенной») соответствует *максимуму* энтропии, тогда как жизнь, согласно У. Р. Эшби, имеет стремление приближать биохимические процессы к *минимуму* энтропии. Гейзенберговское соотношение неопределенности также является одним из факторов, способствующих применению понятия энтропии в физике.

Однако эренфестовское понимание энтропии, несмотря на его плодотворность в статистической физике, было лишь первоначальной догадкой на уровне той величины $\log_2 q$ (где q — число равновероятных исходов испытания), которая рассматривалась в п. 39 в связи с первоначальным подсчетом информации. Современное определение энтропии в виде формулы (1) и развитие на этой основе теории информации и теории кодирования связано с именем замечательного американского математика и инженера Клода Шеннона, который в 1947 – 1948 годах опубликовал ряд основополагающих работ в этой области. Конечно, шенноновская энтропия (1) зависит лишь от вероятностей $p_1, p_2, ..., p_q$ различных исходов испытаний, но не учитывает, являются ли сами эти исходы близкими друг к другу или же они имеют большой разброс. Однако именно эти свойства энтропии, т. е. ее чисто статистический характер особенно важны для теории передачи сообщений по линиям связи и, вообще, для рассмотрения процессов, связанных с передачей и хранением информации любого вида.

В качестве примера рассмотрим исследование *психических реакций* человека на появление некоторых сигналов. Время Δt реакции на простой сигнал (скажем, нажатие кнопки при вспыхивании лампочки) составляет даже при определенной натренированности не менее 0,1 секунды. Если же может вспыхнуть любая из q лампочек (с равной вероятностью) и требуется нажать соответствующую кнопку, то время реакции возрастает и становится равным примерно $\Delta t \cdot \log_2 q$, т. е. это время реакции приблизительно пропорционально эренфестовской энтропии $\Delta t \cdot \log_2 q$. Для иллюстрации на рис. 232 приведен график, полученный американским психологом Р. Хайнманом для случая не более восьми равновероятно вспыхивающих лампочек ($q \leq 8$), где по оси абсцисс откладывается H — энтропия в битах, а по оси ординат T — время реакции в микросекундах. При неравновероятном вспыхивании лампочек действует шенноновская энтропия (1).

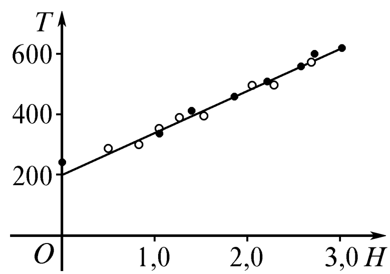


Рис. 232

Шеннон использовал введенную им энтропию для исследования *кодов*, применяемых при передаче информации. Известная «азбука Морзе» представляет собой фактически троичный код, поскольку она использует три сигнала: точку, тире и длинную паузу между буквами (двукратная длинная пауза означает пробел, т. е. отделяет слова друг от друга). Этот код требует большего времени для передачи сообщения, чем двоичный код Бодо, применяемый в телеграфных аппаратах (1— включение тока, 0 — пауза, причем оба эти элементарные сигнала имеют одинаковую длительность). Каждый элемент текста (буква, десятичная цифра, знак препинания, пробел) кодируется по Бодо *шестью* последовательными двоичными цифрами 0 или 1 (отсутствие или наличие тока). Так как шестью двоичными цифрами можно закодировать 64 комбинации, то этого достаточно для обозначения всех элементов текста. А декодирование вполне аналогично «отгадыванию» одного задуманного из 64 элементов, т. е. оно требует шести битов информации (Равна ли единице первая цифра шестерки? Равна ли единице вторая цифра? ... Шестая?). При таком «стандартном» способе кодирования передача текста, содержащего N элементов, требует N шестерок, т. е. $6N$ битов информации.

Однако 32 буквы русского алфавита, 10 цифр и 8 знаков препинания составляют всего 50 (а не 64) элементов текста, т. е. энтропия (если все знаки считаются равновероятными) равна $\log_2 50 \approx 5{,}644$ битов, что существенно меньше, чем 6 битов. А так как знаки не являются равновероятными (например, вероятность использования буквы Ы существенно меньше вероятности для П и т. д.), то шенноновская энтропия будет еще меньше. Если, скажем, она равна 5 битам, то передача текста, содержащего N элементов, потребует $5N$ битов информации, тогда как код Бодо использует $6N$ битов.

Это показывает возможность создания *более* экономичного кода передачи информации, чем код Бодо. Кроме экономичности, разумеется, важной чертой кода является его надежность, т. е. возможность однозначного декодирования, расшифровки. Таков шенноновский подход к созданию кодов передачи информация. На этой основе К. Шеннон и независимо от него Р. Фано предложили в 1948 – 1949 годах новые коды передачи информации, существенно более экономичные, чем код Бодо. Несколько позже Д. Харман разработал (для английских текстов) *наиболее экономичный* способ кодирования, также основанный на энтропийных соображениях.

В теории информации изучается, кроме того, проблема передачи сообщений при наличии помех, в решении которой основные результаты принадлежат К. Шеннону. Созданы также способы кодирования, которые позволяют обнаруживать и исправлять ошибки в поступающих сообщениях. Таким образом, появление понятия энтропии привело к созданию теории кодирования, которая является новой ветвью математики, важной как в теоретическом, так и в прикладном отношении.

Беседа 10. Комбинаторные задачи о графах

43. Графы и их элементы

На рис. 233 и других этого пункта изображены некоторые *графы*. Каждый граф имеет *вершины* и *ребра*, причем каждое ребро соединяет какие-либо две вершины. Чаще всего рассматривают графы, у которых множество вершин конечно и множество ребер также конечно. На рисунках вершины изображаются маленькими кружочками, а ребра — соединяющими их дугами (в частности, отрезками).

На рис. 234 граф содержит 6 вершин и 9 ребер, причем некоторые ребра изображены пересекающимися. Однако следует представлять себе, что этих «лишних» точек пересечения как бы не существует. Например, можно считать, что ребра являются некоторыми дугами в пространстве, и лишь за счет изображения графа на плоскости возникли эти «лишние» пересечения ребер. Иначе говоря, если p и q — два ребра некоторого графа и A — общая точка этих ребер, то A является концом каждого из этих ребер (рис. 235). Можно дать и иное пояснение: если в некотором графе имеется ребро p с концевыми вершинами A и B (рис. 236), то никакая внутренняя точка X дуги p не является ни вершиной этого графа, ни точкой, принадлежащей какому-либо другому ребру.

В графе, изображенном на рис. 237, вершины A и B соединены четырьмя ребрами, а вершины B и C — двумя. В таком случае говорят, что имеется *четырехкратное* ребро, соединяющее вершины A и B, и *двукратное* ребро, соединяющее B и C. Мы условимся в дальнейшем под термином «граф» понимать граф без кратных ребер

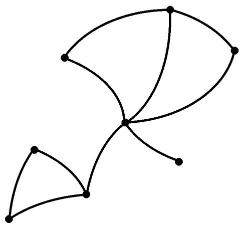


Рис. 233

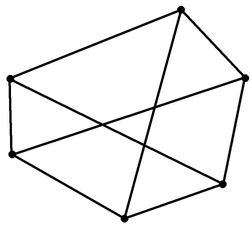


Рис. 234

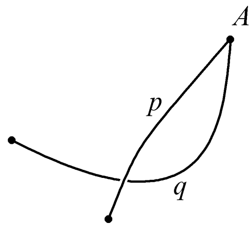


Рис. 235

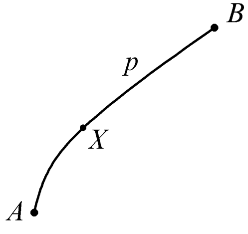


Рис. 236

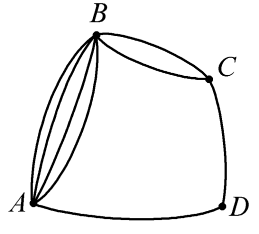


Рис. 237

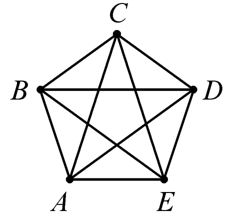


Рис. 238

(если не оговорено специально, что допускается рассмотрение кратных ребер). Таким образом, если A и B — вершины графа, то либо существует *одно* ребро, соединяющее эти вершины, либо же вершины A и B ребром не соединены.

Граф с n вершинами, в котором *каждые* две вершины соединены ребром, называется *полным графом* с n вершинами. На рис. 238 изображен полный граф с пятью вершинами (напомним, что «лишние» точки пересечения ребер, связанные с изображением этого графа на плоскости, считаются как бы не существующими; например, ребра AC и BD *не имеют* общих точек в этом графе). Легко видеть, что полный граф с n вершинами содержит $\binom{n}{2}$ ребер (т. е. ровно столько, сколько имеется *пар* вершин). Если же граф с n вершинами не является полным, то его можно получить следующим образом: взять *полный* граф с этими n вершинами (т. е. каждые две вершины соединить ребром), а затем ненужные ребра удалить. Из этого вытекает, что *любой граф с n вершинами имеет не более $\binom{n}{2}$ ребер.*

Пусть, например, проводятся игровые спортивные соревнования — чемпионат, в котором участвует n команд, причем по условиям чемпионата каждые две команды должны провести между собой одну игру. Будем считать команды *вершинами* графа, а игру между двумя командами — *ребром*, содержащим соответствующие вершины. После завершения чемпионата мы получим *полный граф* с n вершинами, поскольку для любых двух вершин (команд) имеется соединяющее их ребро (игра между этими командами). Всего в течение чемпионата должно быть проведено $\binom{n}{2}$ игр — по числу ребер полного графа с n вершинами. Если же мы рассмотрим некоторый момент, когда часть игр уже проведена, но чемпионат еще не закончился, то получим граф с n вершинами, в котором проведены не все ребер, а лишь некоторые из них, соответствующие уже проведенным играм.

В качестве второго примера рассмотрим *граф знакомства*. Имеются n персон — будем считать их *вершинами графа*. Две вершины условимся соединять *ребром*, если соответствующие персоны *знакомы* друг с другом (для четкости будем считать, что два человека знакомы, если каждый из них знает другого по имени и хоть раз обменялся с ним рукопожатием). В зависимости от выбора рассматриваемой группы персон в этом графе может быть от 0 до $\binom{n}{2}$ ребер.

Пусть G — некоторый граф и a — его вершина. *Индексом* вершины a называется число ребер графа G, имеющих a своим концом, т. е. число ребер *примыкающих* к вершине a (или, как еще говорят, инцидентных вершине a). Этот индекс мы будем обозначать через $\operatorname{ind} a$ (или, более полно, $\operatorname{ind}_G a$). Например, на рис. 239 имеем $\operatorname{ind} A = 3$, $\operatorname{ind} B = 5$. В графах, изображенных на рис. 234 и 240, каждая вершина имеет индекс 3, а в графе на рис. 238 каждая вершина имеет индекс 4.

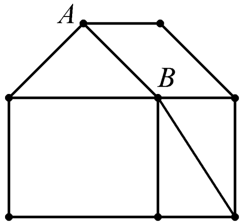


Рис. 239

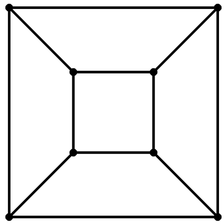


Рис. 240

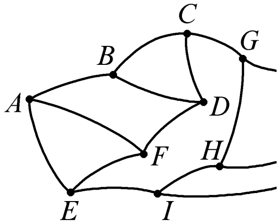


Рис. 241

Вообще, в полном графе с n вершинами каждая вершина имеет индекс $n-1$.

Следующая несложно доказываемая теорема устанавливает связь между индексами вершин конечного графа и числом его ребер: *во всяком конечном графе сумма индексов всех его вершин равна удвоенному числу ребер графа*. Иначе говоря,

$$\sum_a \operatorname{ind} a = 2P, \tag{1}$$

где сумма в левой части распространена на все вершины графа, а P — число его ребер. В самом деле, в сумме, стоящей в левой части равенства (1), каждое ребро засчитывается *дважды*: с того и с другого конца. Потому эта сумма и равна *удвоенному* числу ребер.

Из доказанной формулы (1) вытекает, в частности, что если в конечном графе G все вершины имеют один и тот же индекс k, то справедливо соотношение

$$Bk = 2P, \tag{2}$$

где B — число вершин графа G.

В качестве примера рассмотрим следующую комбинаторную задачу из довоенного мехматского фольклора Московского университета. Можно ли 77 абонентов соединить телефонными линиями так, чтобы каждый абонент был соединен линиями ровно с тремя другими (на рис. 241 изображен фрагмент такой телефонной сети)? Ответ, как показывает формула (2), отрицательный. Действительно, если бы искомый граф (вершинами которого являются абоненты, а ребрами — соединяющие их линии связи) существовал, то в нем каждая вершина имела бы индекс 3, и потому левая часть в (2) была бы равна $77 \cdot 3$, т. е. была бы *нечетным* числом. Однако это невозможно, поскольку правая часть в (2) *четна*.

Более общим образом, из формулы (1) вытекает, что в любом конечном графе число вершин, имеющих нечетные индексы, является четным.

Задачи и упражнения

166. Семеро коллег послали друг другу поздравительные новогодние открытки, причем каждый послал открытки трем коллегам. Могло ли случиться, что каждый из них получил открытки именно от тех своих коллег, которым он послал поздравления?

167. 30 команд участвуют в первенстве по футболу в один круг (каждые две команды встречаются один раз). Докажите, что в любой момент состязаний найдутся две команды, сыгравшие одинаковое число матчей.

168. Существует ли граф с пятью вершинами, у которого индексы вершин все различны и равны 0, 1, 2, 3 и 4?

169. Докажите, что в любом графе число вершин с нечетными индексами всегда четно.

170. Пусть в графе с n вершинами только одна пара вершин имеет одинаковые индексы. Докажите, что в этом случае у графа имеется либо вершина индекса 0 (отдельная вершина), либо вершина индекса $n-1$.

44. Цепи и циклы в графах

Последовательность ребер графа называется *цепью*, если эти ребра можно вычертить одно за другим, не отрывая карандаша и не проходя какое-либо ребро дважды. Если ребра p_1, p_2, ..., p_k образуют цепь, то их концы можно обозначить таким образом: $p_1 = A_0A_1$, $p_2 = A_1A_2$, ..., $p_k = A_{k-1}A_k$, т. е. начало каждого следующего ребра совпадает с концом предыдущего. Вершины A_0 и A_k называются *концами* цепи $A_0A_1 \ldots A_k$. Если цепь замкнута, т. е. конечная вершина A_k совпадает с начальной вершиной A_0, то цепь называется *циклом*.

Подчеркнем, что никакое ребро не может входить в цепь два или большее число раз. На рис. 242 ребра p_1, p_2, p_3, p_4, p_2, p_5 цепь не образуют. Ребра p_1, p_2, p_2, p_3 на рис. 243 также не образуют цепь. В то же время цепь может пересекать себя, т. е. дважды (или большее число раз) проходить через некоторые вершины. На рис. 244 цепь дважды проходит через вершину $A_2 = A_5$, т. е. содержит «петлю» $A_2A_3A_4A_5$. Если же цепь $A_0A_1 \ldots A_k$ не пересекает себя, т. е. все вершины A_0, A_1, ..., A_k различны (рис. 245), то цепь называется *простой*. В частности, цикл без самопересечений называется *простым циклом*.

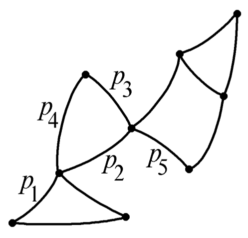


Рис. 242

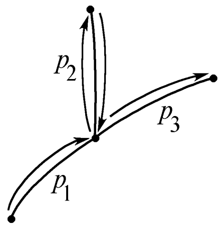


Рис. 243

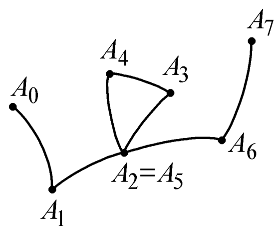


Рис. 244

Ясно, что если цепь с концами не является простой, т. е. имеет самопересечения, то можно, выбросив петли, получить простую цепь с теми же концами a_0, a_k. Так, на рис. 246 вместо $A_0A_1A_2A_3A_4A_5A_6A_7$ можно взять простую цепь $A_0A_1A_2A_6A_7$ с теми же концами.

Граф называется *связным*, если любые две его вершины могут быть соединены цепью. Несвязный граф может быть разбит на несколько отдельных связных графов, называемых *компонентами* исходного графа. На рис. 246 изображен граф, состоящий их трех компонент.

Пусть A и B — две вершины связного графа G. Наименьшее из таких чисел k, что в G существует цепь из k ребер, соединяющая вершины *A и B, называется комбинаторным расстоянием* между вершинами *A и B в графе G*. *Наибольшее* из комбинаторных расстояний между вершинами графа G называется *комбинаторным диаметром* графа G. Например, полный граф с n вершинами имеет комбинаторный диаметр 1. Граф, изображенный на рис. 247, имеет комбинаторный диаметр 2. Граф, представляющий собой *простой контур* с n вершинами (рис. 248), имеет своим комбинаторным диаметром число $\frac{n}{2}$ или $\frac{n-1}{2}$ (в зависимости от четности числа n). Доказать это предоставляется читателю.

Самая первая работа по теории графов появилась в 1736 году. Она принадлежит перу величайшего математика Леонарда Эйлера, о некоторых результатах которого мы уже говорили выше. Эйлер исследовал *головоломку о Кенигсбергских мостах*. Город Кенигсберг (ныне Калининград), расположенный на двух берегах реки Прегель и на двух островах, имел в то время семь мостов (рис. 249). Головоломка состояла в том, чтобы совершить прогулку по городу, маршрут которой проходил бы ровно один раз по каждому мосту. Сопоставим с планом города некоторый граф: на рис. 250 вершины L и R соответствуют левому и правому берегам, вершины A и B — островам, а каждое ребро — мосту. Заметим, что граф содержит кратные ребра: вершина B соединена с каждой из вершин L, R двумя ребрами. Задача заключается в том, чтобы решить, можно ли нарисовать этот граф

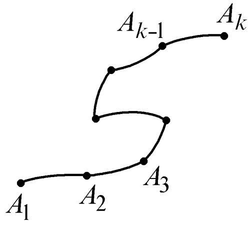


Рис. 245

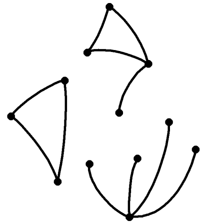


Рис. 246

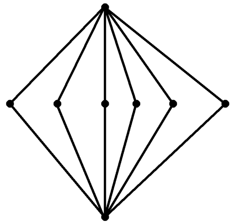


Рис. 247

«одним росчерком», т. е. найти в нем самопересекающуюся цепь, проходящую (по одному разу) по всем ребрам.

Можно поставить следующую, более общую проблему, которая и была решена Эйлером. Связный граф G, возможно, содержащий кратные ребра, называется *уникурсальным* («однoросчерковым»), если его весь можно пройти непрерывным движением, не проходя одно и то же ребро дважды, т. е. если в нем существует самопересекающаяся цепь, содержащая в определенной последовательности все ребра. Головоломка о кенигсбергских мостах как раз и сводится к вопросу, является ли граф на рис. 250 уникурсальным.

Эйлер предложил следующее решение проблемы уникурсальности. Предположим, что связный граф G уникурсален, т. е. в нем существует путь, начинающийся в некоторой вершине A и оканчивающийся в некоторой вершине Z (возможно, совпадающей с A), который по одному разу проходит каждое ребро. Пусть C — произвольная вершина графа G, отличная от A и Z. Рассмотрим все ребра, имеющие вершину C своим концом. Рассматриваемый путь, обходящий граф G, подходя к вершине C по некоторому ребру, сразу же уходит из C по другому ребру (рис. 251). Если впоследствии он снова приходит в C по третьему ребру, то тут же уходит из C по четвертому (рис. 252) и т. д. Из этого видно, что в вершине C сходится четное число ребер, т. е. индекс вершины C четен. Итак, индексы всех вершин графа G, кроме может быть, A и Z, четны. Из аналогичных соображений вытекает, что вершины A и Z, если они не совпадают, имеют нечетные индексы. Если же A и Z совпадают (т. е. мы заканчиваем путь в той же вершине, из которой исходили), то и эта вершина имеет четный индекс.

Итак, если связный граф уникурсален, то все его вершины, кроме, может быть, двух, имеют четные индексы. Это и есть эйлерово условие уникурсальности. Оно, как показывают проведенные выше рассуждения, является *необходимым*, т. е. связный граф, который ему не удовлетворяет, уникурсальпым быть не может. На рис. 253 приведены примеры графов, не являющихся уникурсальными.

Однако эйлерово условие не только необходимо, но и достаточно, т. е. *связный конечный граф G в том и только в том случае уникурсален, если он содержит не более двух вершин с нечетными индексами*. Заметим, что только одну вершину с нечетным индексом граф иметь не

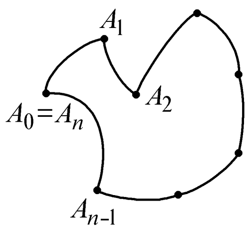


Рис. 248

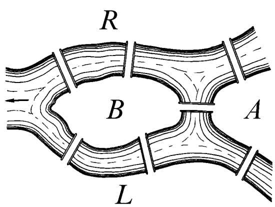


Рис. 249

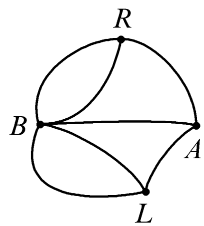


Рис. 250

может (в силу заключительного утверждения, содержащегося в предыдущем пункте).

Доказательство того, что эйлерово условие является не только необходимым, но и достаточным условием уникурсальности, чуть сложнее, но проводится аналогичным образом. В самом деле, пусть G — граф, все вершины которого имеют четные индексы, и A — произвольная его вершина. Будем, начиная от A, проводить цепь «как попало», пока это возможно. Сколько бы раз мы ни проходили вершину C, отличную от A, всегда останется хотя бы одно ребро, по которому можно уйти из вершины C (поскольку C имеет четный индекс). Поэтому вычерчиваемая цепь не может окончиться нигде, кроме A (а окончиться она должна, поскольку ребер конечное число). Если полученная замкнутая цепь K проходит по *всем* ребрам, наша цель достигнута.

Если же это не так, то найдется такая вершина Q цепи K, к которой примыкает ребро p', *не вошедшее в цепь* K (рис. 254). В этом случае мы, идя вдоль цепи K от вершины A и достигнув вершины Q, не пойдем дальше вдоль цепи K, а сначала совершим вспомогательную прогулку, т. е. пойдем по ребру p' и будем идти «как попало» по ребрам, не вошедшим в цепь K. При этом мы обязательно вернемся к вершине Q (больше нигде эта вспомогательная прогулка окончиться не может в силу соображений четности). Вернувшись в вершину Q, мы затем пойдем дальше вдоль цепи K. В результате получится вместо K другая цепь, проходящая по *большему* числу ребер. Если эта новая цепь содержит все ребра, то цель достигнута. Если же нет, — мы снова применим тот же прием к цепи и т. д. В конце концов (в силу конечности числа ребер) мы получим цепь, обходящую все ребра графа G. Этим достаточность доказана для графа, все вершины которого имеют четные индексы.

Наконец, если граф G содержит две вершины A и Z с нечетными индексами, мы можем применить то же рассуждение, начав путь из вершины A и закончив его в Z.

Доказанная теорема позволяет, в частности, получить *правило обхода лабиринта*. Рассмотрим некоторый лабиринт (рис. 255). Пусть нам поставлена задача дойти до сокровища, запрятанного в глубине этого лабиринта, и вернуться обратно. Поскольку мы не знаем, где именно запрятано сокровище, мы должны обойти весь лабиринт,

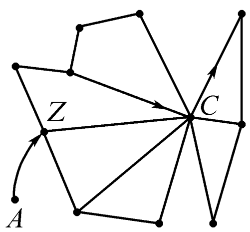


Рис. 251

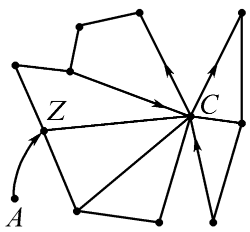


Рис. 252

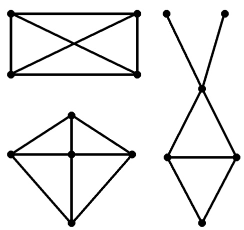


Рис. 253

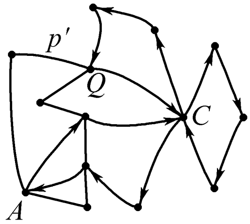


Рис. 254

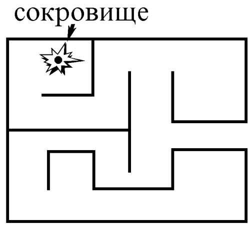


Рис. 255

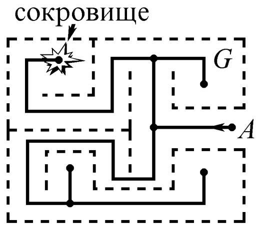


Рис. 256

включая возможные тупики. Проведем средние линии всех коридоров лабиринта — это будут ребра графа G, а вершинами станут перекрестки и тупики лабиринта (рис. 256). Теперь заменим каждое ребро графа G *двукратным* ребром (рис. 257). Получившийся граф обозначим через G^*. В графе G^* каждая вершина имеет *четный* индекс и потому он уникурсален (по теореме Эйлера), т. е. можно найти цепь, обходящую все его ребра. Если такую цепь нам удастся найти, то это означает, что мы пройдем *дважды* каждое ребро графа G, т. е. обойдем весь лабиринт, проходя каждый коридор два раза (например, туда и обратно).

Итак, начнем путешествие по лабиринту, отправляясь от вершины A графа G^* (от входа в лабиринт; рис. 258). Чтобы не проходить никакой коридор более двух раз, условимся после прохождения каждого перекрестка ставить стрелку (по направлению движения) на стене того коридора, по которому мы ушли от этого перекрестка. Тогда второй раз мы в том же направлении по этому коридору не пойдем (а если когда-нибудь снова попадем на этот перекресток, то уйдем по другому коридору, еще не помеченному стрелкой). Это обеспечит нам возможность заведомо *выйти* из лабиринта (ибо цепь может окончиться только в начальной вершине A).

Однако это еще не все: хотя мы непременно выберемся из лабиринта, а не будем блуждать в нем вечно, но, возможно, мы обойдем *не весь* лабиринт и не доберемся до сокровища (рис. 258). Чтобы избежать этого, условимся, что когда мы в *первый раз* будем попадать на какой-то перекресток, мы будем ставить кружок на стене того

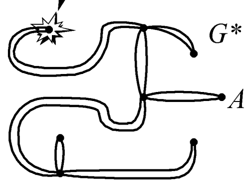


Рис. 257

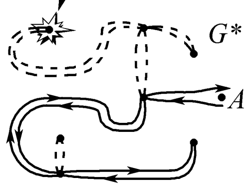


Рис. 258

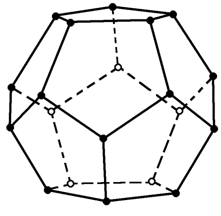


Рис. 259

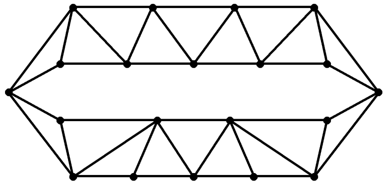


Рис. 260

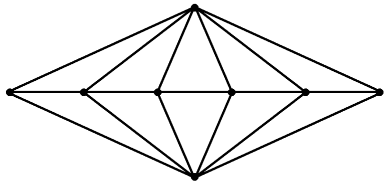


Рис. 261

коридора, по которому мы на этот перекресток пришли. Теперь правило обхода лабиринта гласит: от каждого перекрестка разрешается уходить по коридору, помеченному кружком, только после того, как все остальные коридоры, отходящие от этого перекрестка, уже были пройдены. Читатель докажет самостоятельно, что при соблюдении этого правила мы обойдем весь лабиринт и непременно завладеем сокровищем.

В заключение сформулируем проблему, в некотором смысле противоположную проблеме уникурсальности. Вопрос заключается в том, чтобы решить, существует ли в данном связном конечном графе G такой *простой цикл*, который проходит через все его вершины. Иными словами, существует ли в графе G «прогулка», позволяющая посетить все его вершины, причем каждую вершину только один раз. В отличие от задачи уникурсальности здесь надо пройтись не по всем ребрам, а через все вершины.

Проблема эта была на частном примере предложена известным ирландским математиком сэром Уильямом Гамильтоном в 1859 году. Гамильтон предложил обойти все вершины правильного додекаэдра (рис. 259), двигаясь по его ребрам и проходя каждую вершину лишь один раз. Головоломка эта имеет положительное решение (т. е. требуемый цикл на додекаэдре существует). Вскоре головоломку забыли, но проблему, сформулированную выше, назвали именем ее изобретателя: граф, в котором существует простой цикл, обходящий все вершины, называется *гамильтоновым* графом. В отличие от проблемы уникурсальности, эта проблема в общем случае не решена, т. е. неизвестно условие (необходимое и достаточное), при выполнении которого граф является гамильтоновым. Примеры гамильтоновых графов приведены на рис. 260, 261.

Задача эта имеет интересное обобщение: на ребрах гамильтонова графа указаны некоторые числа, означающие «стоимости проезда» по этим ребрам, и требуется найти «наиболее дешевый» гамильтонов цикл (т. е. обойти все вершины по разу с наименьшими затратами). Проблема эта, известная под названием «задачи о странствующем торговце», не решена в общем случае, т. е. метод нахождения требуемого пути в общем случае неизвестен.

Задачи и упражнения

171. Докажите, что граф с n вершинами, представляющий собой простой контур, имеет своим комбинаторным диаметром число $\frac{n}{2}$ или $\frac{n-1}{2}$ (в зависимости от четности числа n).

172. Является ли уникурсальным граф, изображенный на рисунке 260?

173. Для каждого из 5 правильных многогранников рассмотрим граф, состоящий из всех его вершин и ребер. Какие из этих графов уникурсальны и какие гамильтоновы?

174. Верно ли, что можно обойти граф, пройдя каждое ребро ровно 3 раза, в том и только в том случае, если граф уникурсальный?

175. Является ли гамильтоновым граф, изображенный на рисунке 260?

45. Плоские графы

Любой граф можно изобразить на плоскости, взяв нужное число вершин и проведя линии, соединяющие некоторые пары вершин. Однако при этом проведенные линии (ребра графа) могут иметь «лишние» пересечения (как на рис. 234 и 238). Если граф можно так начертить на плоскости, что лишних пересечений не будет, то граф называется *плоским*.

Прежде чем изучать плоские графы, уточним, что значит «можно начертить граф на плоскости». На рис. 262 и 263 два графа G_1 и G_2 изображены на плоскости по-разному, но «устроены» они одинаково: в графе G_1 вершины A_1 и A_3 соединены ребром, и в графе G_2 соответствующие вершины B_1 и B_3 тоже соединены; в графе G_1 вершины A_2 и A_5 не соединены ребром и в графе G_2 соответствующие вершины B_2 и B_5 тоже не соединены и т. д. «Одинаково устроенные» графы математики называют *изоморфными* (от греческих слов ςσοξ — одинаковый и μορφη — форма).

Более точно, два графа G_1 и G_2 называются *изоморфными*, если они имеют одинаковое число вершин и при этом можно взаимно однозначно отобразить множество вершин первого графа на множество вершин второго таким образом, что две вершины графа G_1 тогда и только тогда соединены ребром, когда образы этих вершин соединены ребром в графе G_2. Например, на рис. 264 и 265 графы кажутся «непохожими», но в действительности они изоморфны (чтобы убедиться в этом, достаточно каждой вершине первого графа поставить в соответствие во втором графе вершину с тем же номером).

Теперь понятие плоского графа можно точно определить следующим образом: граф называется *плоским*, если существует на плоскости изоморфный ему граф, изображенный без «лишних» пересечений. Например, граф, показанный на рис. 262, является плоским, поскольку он изоморфен графу (рис. 263), расположенному на плоскости без «лишних» пересечений. Графы на рис. 266, 267 также являются плоскими, в чем читатель может убедиться самостоятельно.

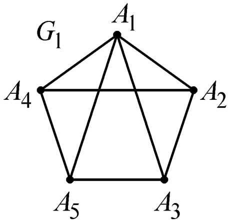


Рис. 262

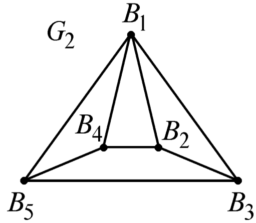


Рис. 263

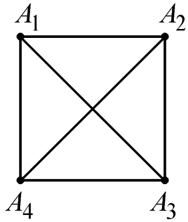


Рис. 264

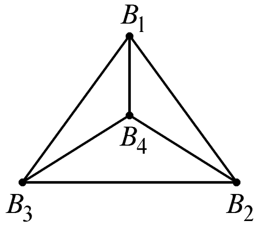


Рис. 265

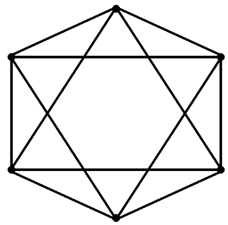


Рис. 266

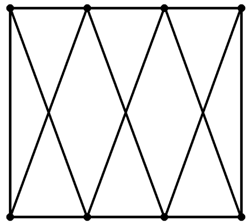


Рис. 267

А теперь мы рассмотрим два графа, которые не являются плоскими. Первый из них («домики и колодцы») получается следующим образом. На плоскости берутся шесть точек $Д_1$, $Д_2$, $Д_3$ («домики») и $К_1$, $К_2$, $К_3$ («колодцы»). Эти точки являются *вершинами* графа, а его *ребрами* являются девять дуг («тропинок»), которые соединяют каждый домик с каждым колодцем (рис. 268). Как мы видим, на рис. 268 две «тропинки» имеют точку пересечения. Оказывается, расположить «тропинки» на плоскости совсем без пересечений не удается. Иными словами, рассматриваемый граф не является плоским.

Вторым примером графа, не являющегося плоским, может служить полный граф с пятью вершинами (рис. 238). Как бы мы ни провели девять непересекающихся ребер этого графа на плоскости (рис. 269), десятое ребро обязательно пересечет хотя бы одно из ранее проведенных ребер. Разумеется, сколько бы попыток, заканчивающихся неудачей, мы ни сделали, это не дает доказательства того, что рассматриваемый граф не является плоским.

Приведем идею общего рассуждения, доказывающего, что этот граф не является плоским. Начертим в виде ломаных линий все 10 ребер полного графа с пятью вершинами A_1, A_2, A_3, A_4, A_5 (с «лишними» пересечениями, рис. 269). Два ребра будем называть *несмежными*, если они не имеют общих концов. Теперь обозначим через I число точек пересечения по всем парам несмежных ребер. На рис. 269 мы имеем $I = 1$, а для рис. 238 справедливо равенство $I = 5$. В обоих случаях число I *нечетно*. Мы докажем, что и при *любом* способе проведения ребер число I будет нечетным (мы при этом предполага-

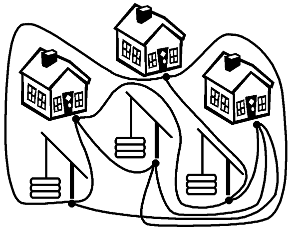


Рис. 268

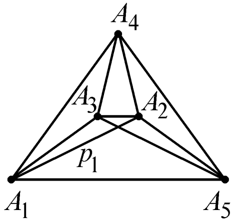


Рис. 269

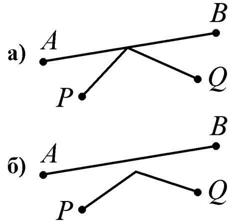


Рис. 270

ем, что каждые два несмежные ребра пересекаются «по-настоящему», т. е. таких общих точек, как на рис. 270а, нет — от них легко избавиться малым шевелением ребер, рис. 270б).

Предположим, что мы меняем положение, скажем, ребра p_1, т. е. заменяем p_1 другой ломаной p_1', с теми же концами A_1, A_2 (рис. 271). Вместе взятые, ребра p_1 и p_1' образуют замкнутый контур. Ребра p_2, p_3, p_4, несмежные с ребром p_1', также образуют замкнутый контур. Но любые два замкнутых контура на плоскости имеют *четное* число точек пересечения (об этом мы говорили в п. 39). Следовательно, число точек пересечения ребра p_1 с этим контуром $p_2p_3p_4$ имеет ту же четность, что и число точек пересечения ребра p_1' с этим контуром. Это означает, что при замене положения ребра p_1 любым новым положением p_1' четность числа *I не меняется*. То же справедливо не только для p_1, но и для любого другого ребра.

Из этого следует, что если мы последовательно заменим все ребра начерченного на плоскости полного графа с пятью вершинами любыми новыми положениями ребер, то четность числа I не изменится. А так как на рис. 269 число I нечетно, то оно *останется нечетным* при любом способе вычерчивания ребер. Значит, в любом случае $I \neq 0$, т. е. без «лишних» пересечений ребер этот граф вычертить на плоскости *невозможно*.

Аналогично доказывается, что и граф «домики и колодцы» не является плоским, т. е. не может быть изображен на плоскости без «лишних» пересечений ребер.

Интересно отметить, что полный граф с пятью вершинами и граф «домики и колодцы» являются «эталонами» графов, не вложимых в плоскость: *если граф не является плоским, то он непременно содержит один из двух указанных графов*. Это было доказано польским математиком Казимиром Куратовским. Например, граф на рис. 234 изоморфен графу «домики и колодцы» и потому он не является плоским. Граф на рис. 272 содержит граф «домики и колодцы», следовательно, он не плоский. Полный граф с шестью вершинами (содержащий полный граф с пятью вершинами) также не является плоским.

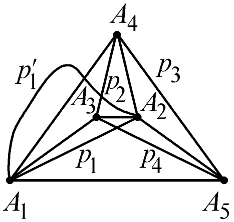


Рис. 271

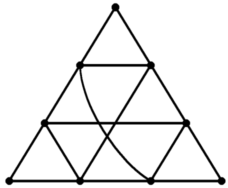


Рис. 272

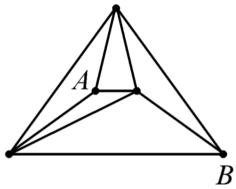


Рис. 273

Если же граф с пятью вершинами не является полным, то он — плоский. Иначе говоря, если мы потребуем, чтобы каждые две из пяти вершин, кроме одной пары вершин A, B, были соединены ребром, то такой граф можно расположить на плоскости без пересечения ребер (рис. 273). Как мы видим, на рис. 273 все ребра изображены прямолинейными отрезками. Это справедливо и в общем случае: каждый плоский граф можно таким образом начертить на плоскости, что все его ребра будут прямолинейными отрезками (если, конечно, граф не содержит кратных ребер).

Задачи и упражнения

176. В четырехугольнике отмечены середины всех сторон, проведены его диагонали и средние линии. Рассмотрим граф, вершинами которого являются вершины четырехугольника и отмеченные точки, а ребрами — отрезки его сторон и проведенные отрезки. Является ли этот граф плоским?

177. Правильный 7-угольник дополнили наименьшими диагоналями. Докажите, что полученный граф не является плоским.

178. Правильный 8-угольник дополнили наименьшими диагоналями. Докажите, что полученный граф может быть вложен в плоскость.

179. Обобщите результаты задач 177 и 178.

180. Правильный $2n$-угольник дополнили наибольшими диагоналями. Докажите, что при $n > 2$ полученный граф — не плоский.

46. Формула Декарта—Эйлера

Пусть G — некоторый связный граф, расположенный в плоскости, B — число его вершин, P — число ребер. Через Γ обозначим число областей («граней»), на которые этот граф разбивает плоскость. К числу этих областей относится и внешняя, неограниченная область, простирающаяся «до бесконечности». На рис. 274 изображен граф, для которого $B = 20$, $P = 30$, $\Gamma = 12$. Легко заметить, что сумма чисел B и Γ на две единицы превосходит число P, т. е.

$$B - P + \Gamma = 2. \tag{3}$$

Для графа на рис. 275 имеем $B = 10$, $P = 13$, $\Gamma = 5$, т. е. соотношение (3) также выполнено. То же можно усмотреть и для графа на рис. 276. Леонард Эйлер доказал, что формула (3) справедлива для *любого* связного графа на плоскости. Точнее, его теорема говорила не о «графах», а была сформулирована в иных терминах.

Именно, рассмотрим произвольный выпуклый многогранник $\boldsymbol{M}$ и над его верхней гранью F, очень близко к ней, расположим лампочку (рис. 277). Если теперь мы представим себе, что ни внутри, ни на гранях многогранника $\boldsymbol{M}$ никакого материала нет (или он совершенно прозрачен), а вот ребра многогранника изготовлены в виде проволочного скелета, то тень от этого скелета на горизонтальной плоскости (*центральная проекция* многогранника $\boldsymbol{M}$, рис. 278) будет представлять собой некоторый граф G. Вершин и ребер у этого графа столько же, сколько и у многогранника $\boldsymbol{M}$, а число областей, на которые граф G делит плоскость, равно числу граней многогранника $\boldsymbol{M}$. В самом деле, каждая грань многогранника $\boldsymbol{M}$, кроме верхней грани F, дает на тени некоторую область, а грани F сопоставим внешнюю, неограниченную область графа G. Значит, для *любого выпуклого многогранника* $\boldsymbol{M}$ *справедливо соотношение* (3), *где* B, P, Γ — *количество вершин*, *ребер*, *граней этого многогранника*. Это и есть обычная формулировка теоремы Эйлера о выпуклых многогранниках. Читатель легко проверит справедливость этой формулы для куба, тетраэдра, октаэдра, призмы, пирамиды и других выпуклых многогранников. Теорема эта была, однако, известна еще Рене Декарту, за несколько десятилетий до Эйлера.

Заметим, что формулировка теоремы для *графов* является более общей, чем для *многогранников*, поскольку не каждый связный плоский граф может быть получен как «тень скелета» некоторого выпуклого многогранника. Например, граф на рис. 278 не является такой тенью, ибо он содержит вершины индекса 1, которых у тени многогранника быть не может.

Ниже мы приведем доказательство формулы Декарта—Эйлера (3), но прежде нам нужно будет рассмотреть одно вспомогательное предложение. Конечный связный граф, не содержащий циклов, называется *деревом*. Так как дерево является связным графом, то любые две его вершины могут быть соединены цепью. А так как дерево не содержит циклов, то любые две его вершины можно соединить *единственной* цепью. Итак, дерево — это граф, любые две вершины которого могут быть соединены *единственной* цепью. Это свойство дерева может быть принято за его определение (эквивалентное первоначальному).

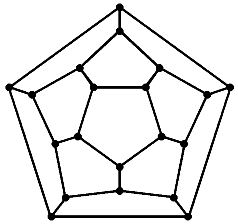


Рис. 274

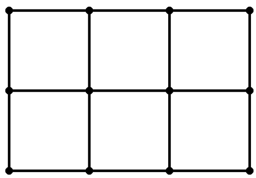


Рис. 275

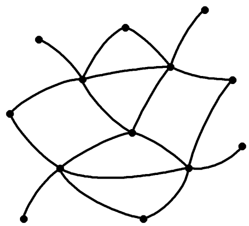


Рис. 276

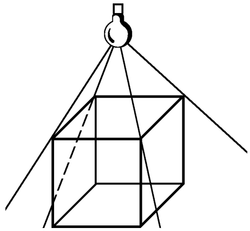


Рис. 277

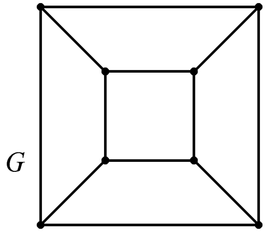


Рис. 278

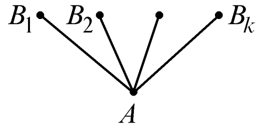


Рис. 279

Всякое дерево является плоским графом. Это можно пояснить следующим описанием дерева. Возьмем в дереве G произвольную вершину A и рассмотрим все ребра, имеющие A одним своим концом (рис. 279), а другим — точки B_1, B_2, ..., B_k, которые мы расположим «этажом выше». Каждая из точек B_i либо имеет индекс 1 (если ребро AB_i является конечным, «тупиковым»), либо же из нее «на следующий этаж» идут новые ребра B_iC_j (рис. 280). Каждая из точек C_1, C_2, ..., C_q, в которых оканчиваются эти новые ребра, также либо имеет индекс 1 (является «тупиком»), либо из нее выходят новые ребра. Продолжая таким образом, мы и начертим на плоскости дерево G.

Докажем теперь, что для каждого дерева G справедливо соотношение

$$B - P = 1. \tag{4}$$

где B — число вершин, а P — число ребер дерева. Для доказательства проведем индукцию по числу вершин дерева. Если дерево содержит только *одну* вершину A, то $B = 1$, $P = 0$, т. е. соотношение (4) справедливо. Это — начало индукции. Совершим теперь переход от n к $n + 1$. Предположим, что для любого дерева, содержащего n вершин, формула (4) справедлива, и рассмотрим дерево G, содержащее $n + 1$ вершину. Пусть CD — какое-либо «тупиковое» ребро дерева G, т. е. вершина D имеет индекс 1 (рис. 281). Отбросив ребро CD, мы получаем дерево G', содержащее n вершин. Для него, по предположению индукции, формула (4) верна. Но при переходе от дерева G' к G мы добавляем *одно* новое ребро CD и *одну* новую вершину D, т. е. разность между числом вершин и ребер не изменяется. Следовательно, и для дерева G формула (4) остается справедливой. Таким образом, по индукции, формула (4) верна для любого дерева.

Теперь мы можем изложить идею доказательства формулы Декарта—Эйлера (3). Для доказательства проведем индукцию по числу Γ, т. е. по числу областей, на которые связный граф G разбивает плоскость. Пусть сначала $\Gamma = 1$, т. е. на плоскости имеется только *одна* область, определяемая связным графом G. Тогда граф G является *деревом* (в противном случае в нем имелся бы цикл, вследствие чего

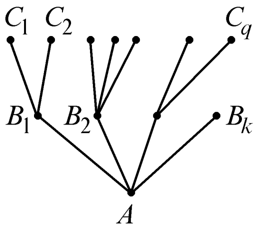


Рис. 280

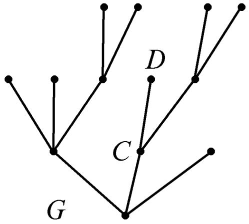


Рис. 281

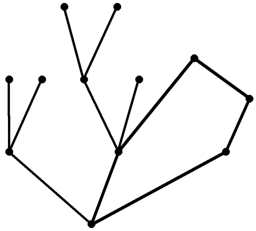


Рис. 282

плоскость оказалась бы разбитой на две или большее число областей, рис. 283). Так как G — дерево, то справедливо соотношение (4); кроме того, $Г = 1$, и потому $В - Р + Г = (В - Р) + Г = 1 + 1 = 2$, т. е. в этом случае соотношение (3) справедливо. Это — начало индукции.

Совершим теперь переход от n к $n + 1$. Предположим, что для любого связного графа, разбивающего плоскость на n областей, формула (3) справедлива, и рассмотрим связный граф G, разбивающий плоскость на $n + 1$ область. Пусть AB — какое-либо ребро, по которому примыкают друг к другу две из этих областей, скажем, области F_1 и F_2 (рис. 283). Если мы уберем из графа G ребро AB (оставив вершины A и B), т. е. уберем «перемычку», разделяющую области F_1 и F_2, то эти области соединятся в одну область F (рис. 284), а остальные области останутся без изменения. Иначе говоря, убрав ребро AB, мы получим вместо G связный граф G', который разбивает плоскость не на $n + 1$, а на n областей. Следовательно, по предположению индукции, для графа G' формула (3) справедлива. Но у графа G *столько же* вершин, сколько у графа G', тогда как ребер у G на одно больше; чем у графа G', и областей («граней») тоже на одну больше. Значит, при переходе от G' к графу G число $В - Р + Г$ *не изменится*, т. е. формула (3) останется справедливой и для графа G. Проведенная индукция показывает, что формула (3) справедлива для любого связного графа на плоскости.

Следует заметить, что проведенное рассуждение содержит *основную идею*, которая для получения корректного доказательства должна быть дополнена некоторыми уточненными деталями. Так, рассматривая граф G, разбивающий плоскость на $n + 1$ часть, мы взяли в нем ребро AB, к которому примыкают две *различные* области F_1 и F_2 (рис. 283). Однако, как показывает рис. 285, могут существовать ребра, к которым с обеих сторон примыкает *одна и та же* область. Значит, в качестве AB можно взять не любое ребро графа G. Однако требуемое ребро (к которому примыкают две *различные* области) все же можно найти. В самом деле, в графе G существует некоторый цикл (этот граф не может быть деревом, поскольку он разбивает плоскость на $n + 1$ часть, т. е. не менее чем на две области). Если взятый цикл

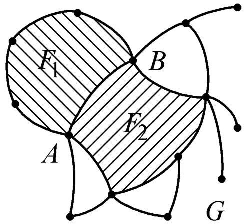


Рис. 283

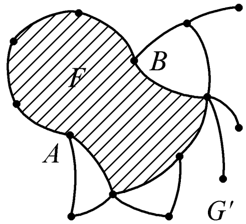


Рис. 284

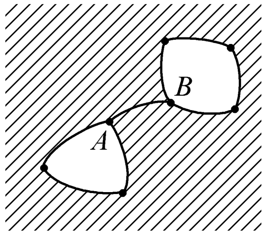


Рис. 285

является самопересекающимся, то, взяв лишь часть ребер, можно из этого цикла выделить *простой* цикл. Так как простой цикл разбивает плоскость на две области (внутреннюю и внешнюю), то любое ребро AB, входящее в этот цикл, обладает требуемым свойством, т. е. примыкающие к нему с двух сторон области являются различными (рис. 283). Этим заполняется указанный пробел в доказательстве. Имеются и другие, более мелкие уточнения, которые следует сделать, чтобы проведенное рассуждение стало полным, корректным доказательством. Мы оставляем это внимательному читателю.

Задачи и упражнения

181. Верна ли формула Декарта—Эйлера для несвязных графов? Как она должна формулироваться, если граф состоит из n связных частей (компонент)?

182. Внутри n-угольника выбрано m точек. Некоторые из этих $m+n$ точек как-то соединены друг с другом прямолинейными отрезками. В результате n-угольник разбился на треугольники. Чему равно число треугольников?

183. В плоском графе вершины имеют индекс 3, а грани имеют не более 5 ребер. Каково наибольшее возможное количество граней, являющихся пятиугольниками?

184. Как будет формулироваться теорема Декарта—Эйлера для связного графа на торе (см. рис. 471 на с. 340)?

185. Сформулируйте теорему Декарта—Эйлера для несвязного графа на торе.

47. Правильные многогранники и паркеты

На рис. 286 изображены пять выпуклых *правильных* многогранников. Их называют также *платоновыми* телами (хотя они были известны еще задолго до Платона). У каждого правильного многогранника грани являются правильными n-угольниками и из каждой вершины исходят k ребер. Эти числа полностью характеризуют правильный многогранник. Например, у куба $n=4$ и $k=3$, т. е. все грани — квадраты и в каждой вершине сходятся три ребра. А у икосаэдра $n=3$, $k=5$.

Почему существует (с точностью до подобия) только пять платоновых тел? Теория графов позволяет дать простой ответ на этот вопрос. Вершины и ребра правильного многогранника образуют граф (скелет многогранника), который можно центрально спроектировать на плоскость (как на рис. 277; на рис. 274 изображена центральная проекция скелета додекаэдра). Эта проекция представляет собой *правильный граф* в том смысле, что каждая его грань имеет n ребер (включая и внешнюю, неограниченную грань), а из каждой вершины исходят k ребер. Число вершин, ребер и граней этого графа будем обозначать через B, P, Γ. Если *каждую* грань (включая и внешнюю, неограниченную) мы обойдем по контуру, считая ребра, то всего насчитаем $n\Gamma$ ребер. При этом каждое ребро мы засчитаем *дважды* (так как к нему примыкают *две* грани). Это дает соотношение

$$n\Gamma = 2P. \tag{5}$$

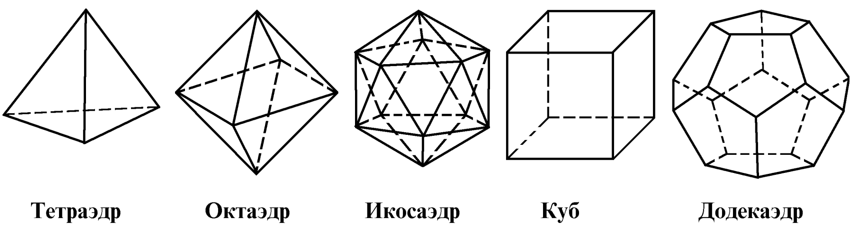
Из соотношений $Bk = 2P$ (см. (2)) и (5), учитывая равенство $B + \Gamma = 2 + P$, вытекающее из формулы Декарта—Эйлера, получаем:

$$\frac{1}{k} + \frac{1}{n} = \frac{B}{2P} + \frac{\Gamma}{2P} = \frac{2+P}{2P} = \frac{1}{P} + \frac{1}{2}.$$

Эта формула позволит найти все правильные графы, т. е. найти все те $n \geq 2$ и $k \geq 2$, которые удовлетворяют найденному соотношению

$$\frac{1}{k} + \frac{1}{n} = \frac{1}{P} + \frac{1}{2}. \tag{6}$$

Пусть сначала $k = 2$. Тогда из (6), (2) и (5) следует, что $P = n$, $\Gamma = 2$, $B = n$. Этим определяется граф в виде замкнутой ломаной с n звеньями (рис. 287). Можно считать, что ему соответствует «многогранник» на сфере, у которого две «грани» (северное и южное полушария), а n вершин и n ребер расположены на экваторе (рис. 288). Аналогично при $n = 2$ мы получаем из (6), (2) и (5) следующие значения: $P = k$, $\Gamma = k$, $B = 2$.. Соответствующий граф (рис. 289) содержит одно k-кратное ребро. Можно считать, что ему соответствует «многогранник» на сфере, у которого имеются две вершины (северный и южный полюсы) и k ребер (меридианов), тогда как каждая грань представляет собой «двуугольник», ограниченный двумя меридианами (рис. 290).



Тетраэдр **Октаэдр** **Икосаэдр** **Куб** **Додекаэдр**

Рис. 286

Пусть теперь $k = 3$. Тогда уравнение (6) принимает вид $\frac{1}{n} = \frac{1}{P} + \frac{1}{6}$. Так как здесь правая часть *больше* $\frac{1}{6}$, то $n < 6$. Таким образом, предоставляются следующие три возможности (для которых B и Γ вычисляются по формулам (2) и (5):

$$k = 3, \quad n = 3, \quad P = 6, \quad B = 4, \quad \Gamma = 4; \tag{7}$$

$$k = 3, \quad n = 4, \quad P = 12, \quad B = 8, \quad \Gamma = 6; \tag{8}$$

$$k = 3, \quad n = 5, \quad P = 30, \quad B = 20, \quad \Gamma = 12. \tag{9}$$

Графы (7) – (9) соответствуют правильному тетраэдру, кубу и правильному додекаэдру (рис. 286).

Далее, пусть $n = 3$. Тогда аналогично мы получим еще раз граф (7) и два новых графа:

$$n = 3, \quad k = 4, \quad P = 12, \quad B = 6, \quad \Gamma = 8; \tag{10}$$

$$n = 3, \quad k = 5, \quad P = 30, \quad B = 12, \quad \Gamma = 20. \tag{11}$$

Эти графы соответствуют правильному октаэдру и правильному икосаэдру (рис. 286).

Наконец, случаи, когда одновременно $n \geq 4$ и $k \geq 4$, невозможны, поскольку тогда левая часть в (6) *не больше* $\frac{1}{2}$, тогда как правая часть *больше* $\frac{1}{2}$. Таким образом, все правильные графы (а потому и правильные многогранники) найдены.

В проведенном выше рассуждении были изучены *конечные* правильные графы. Существуют, кроме того, *бесконечные* правильные графы на плоскости. На рис. 291 – 294 показаны *паркеты*, т. е. замощения плоскости конгруэнтными многоугольниками (треугольниками, четырехугольниками, шестиугольниками). На рис. 291, 294 стороны имеют неодинаковую длину (а грани не являются правильными многоугольниками). Тем не менее, граф, составленный из сторон многоугольников, и в этих случаях является *комбинаторно правиль-*

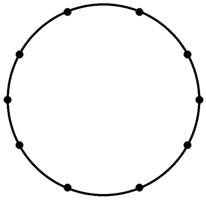


Рис. 287

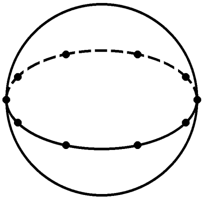


Рис. 288

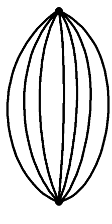


Рис. 289

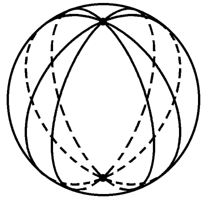


Рис. 290

ным: каждая вершина имеет один и тот же индекс k, и каждая грань имеет одно и то же число сторон n.

А существуют ли комбинаторно правильные паркеты, составленные из многоугольников, у которых число сторон n отлично от 3, 4, 6? Ответ на этот вопрос может быть как положительным, так и отрицательным — в зависимости от точки зрения.

Прежде всего покажем, что *конгруэнтными* многоугольниками можно настлать правильный паркет на плоскости (*евклидовой*!) только в том случае, если число сторон n многоугольника равно 3, 4 или 6. В самом деле, пусть $\boldsymbol{M}$ — такой многоугольник, что конгруэнтными ему многоугольниками можно настлать паркет на плоскости, n — число сторон многоугольника $\boldsymbol{M}$. Через h обозначим диаметр многоугольника $\boldsymbol{M}$, т. е. наибольшее из расстояний между его вершинами, а через S — его площадь. Мы предположим, что граф, составленный из ребер этого паркета, является правильным, и обозначим через k индекс каждой вершины этого графа.

Обозначим теперь через $\boldsymbol{K}$ круг радиуса r и рассмотрим все многоугольники паркета, имеющие с $\boldsymbol{K}$ общие точки. Они образуют конечный граф G (рис. 295), число вершин, ребер и граней которого обозначим через B, P, $Г$ (причем к числу граней причисляется и внешняя, неограниченная область F_∞). Числа B, P, $Г$ удовлетворяют формуле (3).

Так как все рассмотренные многоугольники (т. е. все грани, определяемые графом G, кроме F_∞) покрывают круг $\boldsymbol{K}$, то сумма их площадей больше площади этого круга:

$$(Г-1)S > \pi r^2. \tag{12}$$

Далее, все многоугольники, имеющие общие точки с окружностью круга $\boldsymbol{K}$ (назовем эти многоугольники «крайними»: они на рис. 295 заштрихованы), содержатся в кольце, образованном окружностями радиусов $r+h$ и $r-h$. Площадь этого кольца равна $\pi(r+h)^2 - \pi(r-h)^2 = 4\pi rh$ и потому число γ крайних многоугольников удовлетворяет условию $\gamma S < 4\pi rh$. Умножив это неравенство на неравенство, обратное (12), получаем

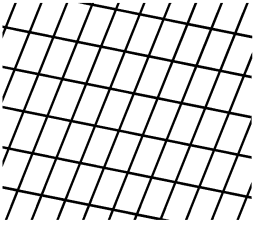


Рис. 291

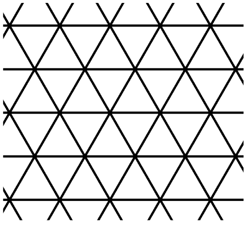


Рис. 292

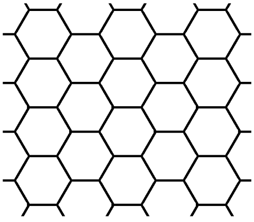


Рис. 293

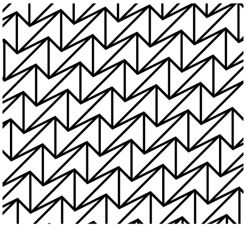


Рис. 294

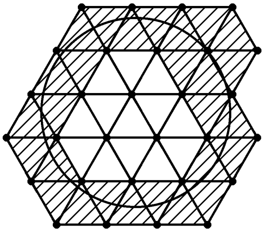


Рис. 295

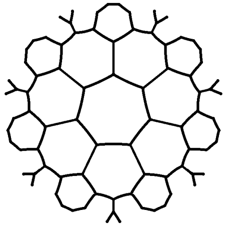


Рис. 296

$$\frac{\gamma}{\Gamma - 1} < \frac{4h}{r}. \tag{13}$$

Посчитав ребра по всем граням, определяемым графом G, мы насчитаем $n(\Gamma - 1) + \rho$ ребер, где ρ — число ребер наружного контура графа G, причем каждое ребро мы засчитаем дважды. Следовательно, $2P = n(\Gamma - 1) + \rho$, т. .е. $2P - n(\Gamma - 1) = \rho$. Но число ρ (т. е. число ребер внешнего контура графа G) меньше, чем число *всех* сторон, посчитанных по *всем* крайним многоугольникам, т. е. меньше чем $n\gamma$. Таким образом,

$$0 < 2P - n(\Gamma - 1) < n\gamma. \tag{14}$$

Теперь посчитаем все вершины по всем граням графа G, кроме внешней грани F_∞. Мы насчитаем $n(\Gamma - 1)$ вершин, причем каждую вершину, не лежащую на внешнем контуре, мы засчитаем k раз, а каждую вершину внешнего контура — менее k раз. Значит, число $kB - n(\Gamma - 1)$ положительно, но меньше чем взятое k раз число вершин внешнего контура, т. е.

$$0 < kB - n(\Gamma - 1) < kn\gamma. \tag{15}$$

Из (14) и (15) получаем (учитывая (13)):

$$0 < \frac{P}{\Gamma - 1} - \frac{n}{2} < \frac{n}{2} - \frac{\gamma}{\Gamma - 1} < \frac{n}{2} \cdot \frac{4h}{r}, \quad 0 < \frac{B}{\Gamma - 1} - \frac{n}{k} < \frac{n\gamma}{\Gamma - 1} < n \cdot \frac{4h}{r},$$

и потому

$$\lim_{r \to \infty} \frac{P}{\Gamma - 1} = \frac{n}{2}, \tag{16}$$

$$\lim_{r \to \infty} \frac{B}{\Gamma - 1} = \frac{n}{k}. \tag{17}$$

Наконец, записав формулу Декарта—Эйлера в виде

$$\frac{P}{\Gamma-1}=\frac{B}{\Gamma-1}+1-\frac{1}{\Gamma-1}$$

и переходя к пределу при $r\to\infty$, мы получаем (в силу (16), (17)) $\frac{n}{2}=\frac{n}{k}+1$, т. е.

$$\frac{1}{n}+\frac{1}{k}=\frac{1}{2}. \tag{18}$$

Но уравнение (18) имеет в натуральных числах n, k лишь следующие три решения: $n=3$, $k=6$; $n=4$, $k=4$; $n=6$, $k=3$. Это и означает, что правильный граф, осуществляющий на евклидовой плоскости паркет из конгруэнтных n-угольников, существует лишь при $n=3$, $n=4$ или $n=6$.

В приведенном выше доказательстве мы использовали *конгруэнтность* многоугольников, составляющих паркет, а также формулу площади круга, взятую из *евклидовой* геометрии. И то, и другое существенно. На рис. 296 показан фрагмент бесконечного правильного графа, дающего паркет из *семиугольных* областей. Правда, эти области не конгруэнтны между собой, да и паркет будет настлан не на всей плоскости, а лишь внутри круга. Но на *плоскости Лобачевского* такой паркет может быть осуществлен из *конгруэнтных* между собой правильных семиугольников и притом на всей плоскости. Вообще, для любых $n\ge 3$, $k\ge 3$ удовлетворяющих условию $\frac{1}{n}+\frac{1}{k}<\frac{1}{2}$, можно построить на плоскости Лобачевского бесконечный правильный граф, имеющий в каждой вершине индекс k и разбивающий всю плоскость на конгруэнтные правильные n-угольники.

В заключение отметим, что существуют также *полуправильные многогранники* (так называемые *архимедовы тела*), т. е. выпуклые многогранники, гранями которых являются правильные многоугольники *двух* типов. Так, на рис. 297 изображен *кубоктаэдр*, гранями которого служат квадраты и равносторонние треугольники. Формула Декарта—Эйлера позволяет найти все архимедовы тела. Их существует 16 типов, причем 14 из них определены, с точностью до подобия, однозначно, а два типа (правильные призмы и антипризмы,

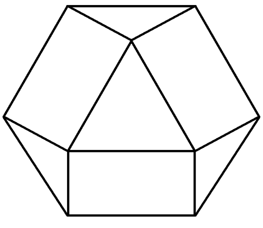


Рис. 297

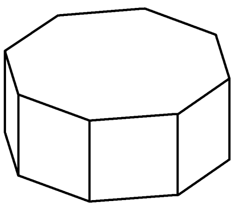


Рис. 298

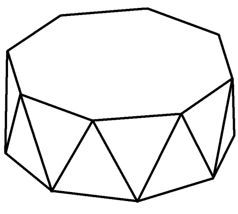


Рис. 299

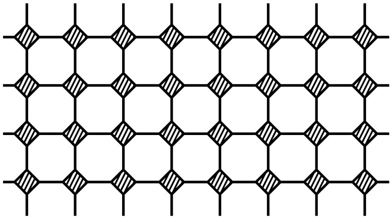


Рис. 300

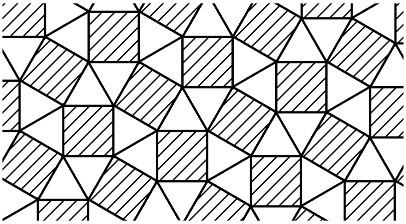


Рис. 301

рис. 298, 299) содержат бесконечно много представителей: их основаниями могут быть правильные многоугольники с любым числом сторон.

На плоскости можно также рассматривать паркеты, составленные из правильных многоугольников не одного, а двух типов. Они также могут быть все найдены с помощью формулы Декарта—Эйлера, причем на евклидовой плоскости таких паркетов имеется лишь конечное число (два примера приведены на рис. 300, 301), а на плоскости Лобачевского их бесконечно много.

Замечательной особенностью всех этих паркетов является их *периодичность*. Например, существует паркет, составленный из правильных двенадцатиугольников и треугольников, причем любой двенадцатиугольник может быть с помощью движения совмещен с любым другим двенадцатиугольником и при этом движении весь паркет снова целиком совместится с самим собой. Но в 1973 году Роджер Пенроуз обнаружил *квазипериодические* паркеты, которые сразу же привлекли внимание математиков и кристаллографов. На рис. 302 изображено разбиение ромба с углами 72° и 108° на два четырехугольника (один из которых— невыпуклый).

С помощью этих четырехугольников и с условием, что никакие два из них не складываются в виде ромба, как на рис. 302, можно построить бесконечное (даже континуальное) множество паркетов, причем ни один из этих паркетов не будет периодическим, хотя некоторые из них имеют симметрию пятого порядка (т. е. совмещаются сами с собой при повороте вокруг некоторой точки на угол

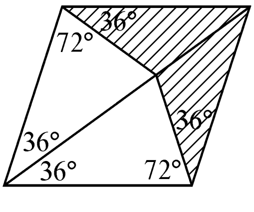


Рис. 302

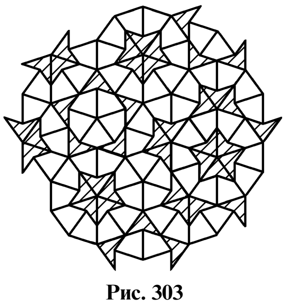


Рис. 303

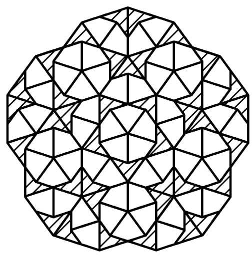


Рис. 304

$\frac{360^\circ}{5}$). Два примера таких паркетов приведены на рис. 303, 304. Однако любой паркет Пенроуза *квазипериодичен*, т. е. любой *конечный* (и притом как угодно большой) кусок паркета повторяется в этом паркете *бесконечно много раз*. Возможно, мы являемся свидетелями постепенного оформления новой ветви математики и кристаллографии — *теории квазикристаллов*.

Задачи и упражнения

186. Взят правильный многогранник. Какой многогранник получится, если его вершинами являются центры граней взятого многогранника, а ребрами — отрезки, соединяющие центры соседних граней взятого многогранника?

187. Докажите, что плоскость можно замостить конгруэнтными четырехугольниками произвольной формы.

188. Можно ли замостить плоскость конгруэнтными пятиугольниками?

189. Можно ли замостить плоскость конгруэнтными выпуклыми семиугольниками?

190. Можно ли замостить плоскость выпуклыми семиугольниками, не обязательно конгруэнтными?

48. Проблема четырех красок

Пусть G — конечный граф на плоскости. Он разбивает плоскость на конечное число областей («стран»). Поставим задачу раскрасить страны в разные цвета, причем таким образом, чтобы получилась «политическая карта». Разумеется, чтобы страны были хорошо видны, необходимо граничащие страны окрашивать в *разные* цвета. Однако в целях экономии количества красок разрешается не граничащие страны окрашивать одним цветом; при этом страны, имеющие общие точки в вершинах графа, но не имеющие общих ребер («границ»), будем считать не граничащими (страны A и $Б$ на рис. 305). Какое минимальное количество красок надо иметь, чтобы можно было раскрасить *любую* карту на плоскости?

Эта задача была сформулирована в 1852 году лондонским студентом Гутри, который обнаружил, что для различения графств на карте Англии достаточно *четырех* красок, и выдвинул гипотезу о том, что четырех цветов достаточно для раскраски *любой* карты. Спустя почти сорок лет английский математик Хивуд доказал, что любую карту на плоскости можно раскрасить в *пять* цветов. Постепенно проблема четырех красок привлекала интерес все большего количества математиков. В 1968 году Оре и Стемпл доказали, что любую карту, *имеющую не более* 40 *стран*, можно раскрасить в четыре цвета.

В 1976 году было распространено мнение, что проблема четырех красок получила *компьютерное* решение (положительное) в работах американских математиков К. Аппеля и В. Хакена. С помощью компьютера (который «помогал» им постепенно совершенствовать первоначальную программу) они разбили все возможные карты почти на 2000 четко определенных типов. Для каждого из этих типов (кроме трех, которые были исследованы «вручную», поскольку компьютер с ними не справлялся) компьютер решал задачу: может ли в рассматриваемом типе карт найтись такая, которая не раскрашивается в четыре цвета? Когда, выполнив десятки миллиардов арифметических и логических операций, компьютер давал ответ «нет», переходили к следующему типу карт и т. д. Получив ответ «нет» для всех типов карт, Аппель и Хакен объявили, что ими получено компьютерное решение проблемы четырех красок.

Новое компьютерное решение было предложено в 1978 году Д. Коэном. Число типов карт было у него существенно меньше, причем результат компьютерных вычислений по каждому типу и подтипу он получал не только в виде готового «нет», а в форме, допускающей «ручную» проверку. Исследование подтипа считалось завершенным, когда компьютер находил достаточно обозримый путь проверки окончательного «нет».

По мнению Коэна *проверка* найденного им решения проблемы четырех красок могла бы быть выполнена одним человеком в течение двух-трех лет (!) ежедневной восьмичасовой работы.

Однако гарантии правильности этих компьютерных решений нет. Ведь в каком-то, скажем, семнадцатом типе карт компьютер мог ответить «нет» не в результате безупречного анализа, а из-за сбоя в работе (что при длительной работе компьютера бывает нередко). Не зная об этом, вычислители переходят к восемнадцатому, девятнадцатому типу карт, фактически *пропустив* исследование семнадцатого типа. Не будет гарантии правильности решения даже в том случае, если мы, затратив много месяцев, повторим уникальный компьютерный эксперимент: может быть, где-то в многомиллиардной цепочке вычислений, связанных с тем же семнадцатым типом, и в нашем компьютере произойдет сбой?

Решение Коэна кажется более предпочтительным, поскольку оно допускает «ручную» проверку. Однако человек, занимавшийся с утра до вечера в течение двух лет нудной проверкой, вряд ли может гарантировать, что он нигде не допустил ни одной ошибки.

В довершение всего в последнее время возникли сомнения в самом *принципе* программирования: похоже, что нет уверенности в переборе всех мыслимых карт с помощью тех типов, которые были указаны. Одним словом, сегодня нет достаточной уверенности в полноте компьютерного решения. Все большее число скептиков склоняются сегодня к мнению о том, что проблема пока остается открытой.

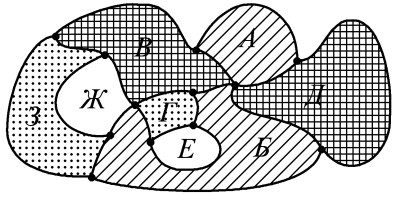


Рис. 305

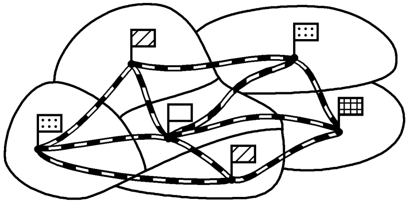


Рис. 306

Существует и иная формулировка проблемы четырех красок, также связанная с графами. Пусть на плоскости начерчена карта, требующая для своей раскраски C цветов. Выберем внутри каждой страны точку («столицу»). Для каждой из двух граничащих стран проведем по территории этих стран «железную дорогу», соединяющую их столицы (рис. 306), причем так, чтобы различные железные дороги не пересекались. Вместо окраски страны в некоторый цвет мы можем водрузить в ее столице флаг, имеющий этот цвет; если при этом две столицы соединены железной дорогой (т. е. страны граничат), то их флаги должны иметь разный цвет. Таким образом, надо покрасить вершины *дуального* графа G^* (ребрами которого являются железные дороги) таким образом, чтобы любые две *смежные* вершины (соединенные ребром) имели разный цвет. Ясно, что *хроматическое число* графа G^*, т. е. наименьшее число цветов, требуемых для его раскраски, равно C. Таким образом, проблема четырех красок может быть переформулирована следующим образом: доказать, что для любого плоского графа G^*, т. е. графа, располагаемого на плоскости без «лишних» пересечений ребер, его хроматическое число не более четырех.

Проблема четырех красок связана с рассмотрением именно *плоских* графов. Для графов, которые не располагаются на плоскости, хроматическое число может быть и бóльшим четырех. Например, полный граф с n вершинами имеет хроматическое число n (поскольку любые две вершины соединены ребром, все вершины должны быть окрашены в *разные* цвета). Однако, возможны и такие случаи, когда неплоские графы имеют маленькое хроматическое число. Например, граф «домики и колодцы» (рис. 268) имеет хроматическое число 2: домики надо окрасить одним цветом, а колодцы — другим.

Задачи и упражнения

191. Докажите, что карту на плоскости, полученную в результате проведения нескольких прямых, можно правильно окрасить двумя красками.

192. Докажите, что карту на плоскости, полученную в результате проведения нескольких окружностей, можно правильно окрасить двумя красками.

193. Карта на плоскости получена путем проведения нескольких окружностей и их диаметров (по одному в каждой окружности). Докажите, что полученную карту можно правильно раскрасить тремя красками.

194. На карте A страны являются многоугольниками. Проведена прямая, не проходящая ни через одну из вершин. Эта прямая разбивает некоторые страны на части, что дает карту B. Докажите, что если карта A раскрашиваема четырьмя красками, то и карта B тоже раскрашиваема четырьмя красками.

195. Покажите, что на торе можно расположить семь стран так, что каждые две граничат между собой.

49. Ориентированные графы

Граф называется *ориентированным* (или направленным) если на каждом его ребре указано (стрелкой) некоторое направление. Мы уже упоминали ориентированные графы в п. 20 в рассказе об упорядоченных множествах: элементы упорядоченного множества рассматриваются как *вершины* графа, и ребро, идущее в направлении от A к B, имеется в том и только в том случае, если $A < B$. Однако при таком способе изображения упорядоченного множества мы обычно рисуем слишком много стрелок (рис. 307). Именно, если $A < B$ и $B < C$, то в графе должны быть указаны не только ребра $\overrightarrow{AB}$ и $\overrightarrow{BC}$, но также и ребро $\overrightarrow{AC}$, поскольку в силу условия транзитивности из $A < B$, $B < C$ следует $A < C$. Будем говорить, что наличие ребра $\overrightarrow{AC}$ является *следствием* наличия ребер $\overrightarrow{AB}$ и $\overrightarrow{BC}$ *в силу транзитивности*.

Условимся теперь *не изображать* ребра, являющиеся следствием в силу транзитивности. Тогда граф, изображающий отношение порядка на рис. 307, будет выглядеть существенно более обозримым образом (рис. 308). А недостающие ребра (являющиеся следствием по транзитивности) можно легко восстановить: соотношение $X < Y$ имеет место на рис. 308 в том и только в том случае, если найдется в графе ориентированная цепь, идущая от X к Y (с учетом направления стрелок). Заметим, что в графе на рис. 308 нет ориентированных циклов. Это и понятно: если бы существовал ориентированный цикл $\overrightarrow{XY}$, $\overrightarrow{YZ}$, ..., $\overrightarrow{VW}$, $\overrightarrow{WX}$, то это означало бы, что $X < Y$, $Y < Z$, ..., $V < W$, $W < X$, т. е. было бы $X < W$, $W < X$, что в упорядоченном множестве невозможно.

Таким образом, ориентированный граф, изображающий отношение порядка, не содержит ориентированных циклов. Справедливо и обратное: если дан ориентированный граф, не содержащий ориентированных циклов, то, дополнив его ребрами по транзитивности, мы получаем некоторое отношение порядка в множестве его вершин. Это отношение порядка как раз и описывается указанным выше правилом: $X < Y$ если существует в графе ориентированная цепь (хотя бы одна), ведущая от X к Y. Так, для отношения порядка, определяемого графом на рис. 309, мы имеем $A < C$ (поскольку имеется ориентированная цепь от A к C). А элементы A и M *несравнимы*, так как от одного к другому не ведет никакая ориентированная цепь.

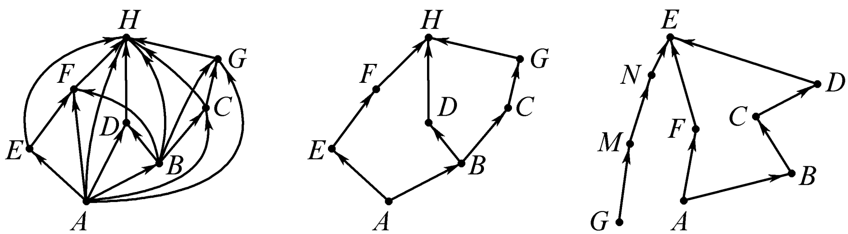


Рис. 307 **Рис. 308** **Рис. 309**

В графе, изображающем отношение порядка, отсутствие стрелок, являющихся следствиями по транзитивности, означает, что мы соединяем вершины U и V стрелкой $\overrightarrow{UV}$ только в том случае, если V является элементом, *непосредственно следующим* за U, т. е. $U < V$, но нет элемента X, расположенного между ними ($U < X < V$).

Граф, изображающий отношение порядка, называется *диаграммой Хассе*, если он не содержит стрелок, являющихся следствиями по транзитивности, и при этим все стрелки идут в направлении *снизу вверх*, т. е. вершина, изображающая *больший* элемент, расположена *выше*, чем меньший. Так на рис. 310 изображена некоторая диаграмма Хассе. Она получается следующим образом: мы рассматриваем куб, диагональ AH которого перпендикулярна горизонтальной плоскости (рис. 311), и считаем, что из двух вершин куба та является большей, которая расположена выше (а две вершины, расположенные на одной высоте, несравнимы).

На рис. 310 граф изображен с «лишними» пересечениями ребер. Можно этот граф, точнее, изоморфный ему, расположить на плоскости и без пересечений (рис. 312), однако это уже не будет диаграммой Хассе: хотя вершина H является *наибольшей*, но на рис. 312 она *не расположена* выше других. Возникает естественный вопрос: в каком случае отношение порядка в некотором конечном упорядоченном множестве $\boldsymbol{M}$ может быть изображено диаграммой Хассе *без самопересечений*? Для случая, когда $\boldsymbol{M}$ содержит единственный максимальный и единственный минимальный элемент, ответ дается следующей теоремой, которую доказал канадский математик Прим.

Пусть $\boldsymbol{M}$ — упорядоченное множество, содержащее единственный максимальный и единственный минимальный элемент, и пусть G — ориентированный граф (без следствий по транзитивности), изображающий отношение порядка в $\boldsymbol{M}$. Обозначим через G_1 неориентированный граф, получающийся добавлением к G одного ребра, соединяющего минимальный и максимальный элементы. Диаграмма Хассе для упорядоченного множества $\boldsymbol{M}$ в том и только в том случае может быть изображена *без пересечений*, если граф G_1 является плоским.

Например, для отношения порядка в множестве вершин куба, показанного на рис. 310, граф G_1 получается добавлением к графу рис. 310 ребра AH (рис. 313). Этот граф не является плоским (он содержит граф «домики и колодцы», рис. 314). Поэтому, согласно теореме Прима, диаграмма Хассе для рассматриваемого отношения порядка *не может быть* изображена без самопересечений.

Возможность расположения некоторого графа (ориентированного или нет) *на плоскости* без пересечений может показаться математической игрушкой, малоинтересной для практики. Однако это не так. Например, при изготовлении печатных плат для радиоприемников, магнитофонов и других приборов на пластине диэлектрика монтируется некоторое количество контактов (для выводов от емкостей, трансформаторов и т. д.), которые, как вершины графа, нужно соеди-

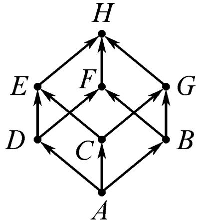


Рис. 310

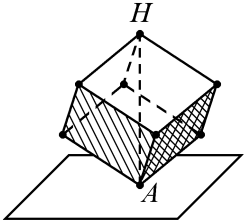


Рис. 311

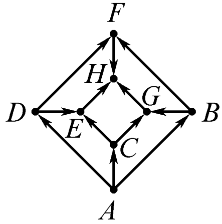


Рис. 312

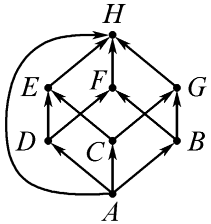


Рис. 313

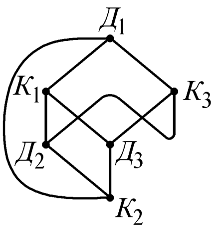


Рис. 314

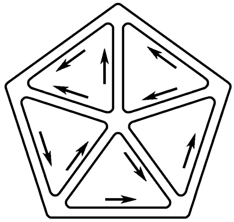


Рис. 315

нить (соответственно схеме) сетью токопроводящих дорожек. Если получающийся граф является плоским, то его можно изготовить травлением или напылением металла (предварительно защитив не подлежащие соединению места лаком или каким-либо экраном) и это существенно упрощает и удешевляет технологию. Если же граф не является плоским, приходится перекидывать «мостики» или делать схему многослойной с изолирующими прослойками диэлектрика.

Диаграмма Хассе особенно удобна при расположении вершин «снизу вверх» в порядке возрастания узловых напряжений (в этом случае близкие друг к другу узлы имеют не очень большую разность потенциалов, и это уменьшает вероятность пробоя).

Задачи и упражнения

196. Закончился круговой турнир баскетболистов, в котором не было ничьих. Всегда ли можно занумеровать команды так, чтобы каждая команда была победителем во встрече с командой со *следующим* номером?

197. Дорожки бульвара расположены так, как показано на рисунке 315. По ним бегают спортсмены в направлениях, указанных стрелками, причем по первой дорожке каждую минуту пробегает один спортсмен, по второй — два и т. д. по десятой — десять спортсменов. Время бега по каждой дорожке равно одной минуте. Расставьте на рисунке номера дорожек.

198. Несколько команд разыграли первенство по волейболу. Докажите, что если какие-нибудь две команды одержали одинаковое количество побед, то найдутся такие команды A, B и C, что A выиграла у B, B — у C, а C у A.

199. Между 5 планетами Солнечной системы установлено космическое сообщение. Ракеты летают по маршрутам: Земля $\to$ Меркурий, Земля $\to$

Плутон, Плутон → Венера, Венера → Марс, Марс → Меркурий, Марс → Земля и Меркурий → Плутон. Можно ли добраться от Меркурия до Земли?

200. Докажите, что всякий полный ориентированный граф имеет простой ориентированный путь, проходящий через все его вершины.

50. Конечные позиционные игры

Интересно применение ориентированных графов к теории игр. Чаще всего игра двух (или нескольких) участников заключается в том, что играющие поочередно делают ходы, которые по определенным правилам позволяют переходить от одной *игровой позиции* к некоторой другой. Игра с *полной информацией* не зависит от случайных элементов (бросание кости, расклад карт, который партнеры друг другу не показывают, и т. д.), причем множество возможных игровых позиций конечно (хотя, может быть, необозримо велико). Правила игры задают некоторый ориентированный граф, по ребрам которого можно переходить от одной позиции к другой. Течение игры (партия) представляет собой ориентированную цепь, проходящую по некоторым вершинам графа (т. е. по некоторым игровым позициям) и обрывающуюся в некоторой завершающей позиции, которая однозначно расценивается как либо «выигрыш первого», либо «выигрыш второго», либо «ничья». Будем предполагать, что трехкратное повторение одной и той же позиции расценивается правилами как «ничья» (такое соглашение имеется в шахматах). Тогда каждая партия оканчивается после конечного числа ходов.

В этих условиях справедлива следующая теорема: каждая игра двух участников (при фиксированной начальной позиции и четко оговоренных правилах) при правильном выборе стратегий игроков либо имеет ничейный исход, либо же несправедлива по отношению к одному из играющих. Иными словами, либо первый игрок (делающий начальный ход) при правильной стратегии фатально выигрывает (что бы ни делал второй), либо первый фатально проигрывает (при правильной стратегии второго), либо же — при правильной игре — ни один не может одолеть другого и игра фатально оканчивается ничейным исходом. Однако в случае необозримо большого (как в шахматах) числа игровых позиций при отсутствии хорошей теории нахождение «правильной стратегии» (т. е. однозначного правильного ответа на каждый ход противника) выше возможностей человеческого интеллекта или возможностей компьютерного перебора позиций. В случае сравнительно несложных игр нахождение «правильной стратегии» возможно.

Рассмотрим в качестве примера следующую игру. Имеются две кучки предметов (скажем, пуговиц), в первой из которых 7, а во второй 8 предметов. При каждом ходе играющий берет либо один предмет из первой кучки, либо один предмет из второй кучки, либо же по одному предмету из каждой кучки. Выигрывает тот, кто забирает последний предмет (или оба последних предмета, если в каждой кучке осталось по одному).

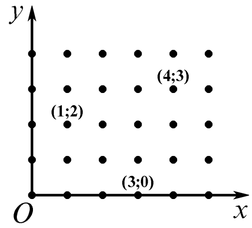


Рис. 316

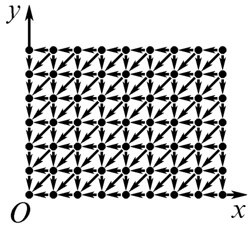


Рис. 317

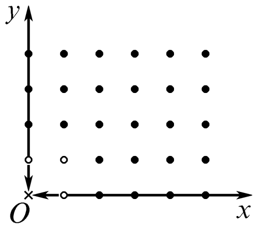


Рис. 318

В этой игре каждая позиция задается парой $(x; y)$, где x — число предметов в первой кучке, а y — во второй. Иначе говоря, игровые позиции характеризуются целочисленными *точками* первого квадранта (рис. 316). Из каждой позиции *внутри* квадранта возможны три хода: на одну единицу влево (берем предмет из первой кучки), на одну единицу вниз (берем из второй кучки) или же по диагонали вниз налево. Если же игровая позиция находится на одной из осей (но не в начале координат), то возможен лишь один ход. Это и дает весь ориентированный граф возможных игровых позиций (вершин) и возможных ходов (ребер, рис. 317). Заметим, что в этой игре ничейных исходов нет (кто-то непременно возьмет последний предмет).

Теперь найдем выигрышную стратегию, производя исследование игры «с конца». На рис. 318 крестиком обозначен конец игры (все предметы взяты), а три кружка обозначают выигрышные позиции, т. е. если игрок попал в какую-либо из этих позиций, то он забирает все и выигрывает. Далее, на рис. 319 поставлены еще два крестика, изображающие *проигрышные* позиции: если игрок попал в такую позицию, то он может из нее перейти только к одному из кружков, т. е. он отдает другому игроку выигрышную позицию. Окаймляя эти крестики кружками, мы получаем еще шесть выигрышных позиций: если игрок попал в одну из этих позиций, то у него есть ход, после которого другой игрок окажется в позиции, отмеченной крестиком (рис. 320). Продолжая, мы обнаружим, что все точки с обеими *четными* координатами помечены крестиком, а остальные — кружками (рис. 321). Так как начальное положение (7; 8) отмечено кружком (*не обе* координаты четны), то *первый* игрок (которому предложено сде-

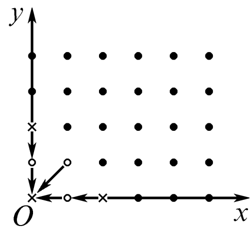


Рис. 319

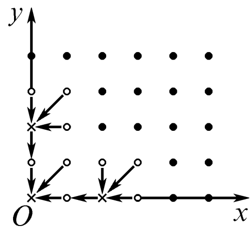


Рис. 320

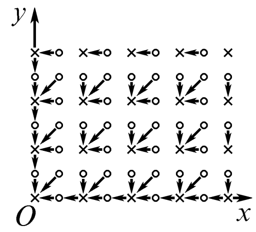


Рис. 321

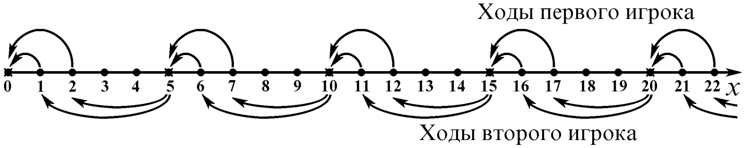


Рис. 322

лать начальный ход) выигрывает: он берет один предмет из первой кучки, после чего второй игрок оказывается в позиции (6; 8), помеченной крестиком. Затем *после любого* хода второго игрока первый снова может сделать обе координаты четными и т. д. — вплоть до выигрыша.

Итак, предложенная игра несправедлива в отношении второго игрока: он при указанной начальной позиции всегда проигрывает. Если же предложить другой вариант игры: те же правила, но с начальным положением (8; 10), то игра будет несправедлива в отношении первого игрока.

В качестве второго примера рассмотрим игру с одной кучкой, содержащей 21 предмет. Правила игры: игрок A, делающий ход первым, может взять один или два предмета (по своему усмотрению). Второй игрок B может взять 3 или 4 предмета. Затем ходит A (и берет 1 или 2 предмета), затем снова B (3 или 4) и т. д. Игрок, забирающий последний предмет, считается выигравшим; если же должен ходить B, но в кучке осталось *менее* трех предметов, т. е. B не в состоянии сделать ход («патовая позиция»), то игра считается закончившейся вничью.

Граф игры показан на рис. 322. Заметим, что при любом ходе одного из игроков другой может своим ходом дополнить число взятых предметов до 5 (т. е. сделать так, что оба вместе возьмут во время этих двух ходов 5 предметов). Теперь ясно, что при начальной позиции 21 игрок A заведомо может выиграть: он берет один предмет, предлагая игроку B позицию 20. Что бы B ни сделал, игрок A следующим ходом предложит ему позицию 15, затем 10, 5 и, наконец, последним ходом перейдет к позиции 0, т. е. заберет все.

Читатель может проследить аналогично, что если заданная начальная позиция имеет вид $5n+1$ или $5n+2$, то выигрывает A. Далее, если начальная позиция имеет вид $5n$, то выигрывает B. Наконец, если начальная позиция имеет вид $5n+3$ или $5n+4$, то при правильном поведении игроков ни один из них не может одолеть другого, игра закончится патовым положением (ничья). Итак, при каждой фиксированной начальной позиции игра либо несправедлива в отношении одного на игроков, либо оканчивается вничью (если игроки не делают ошибок).

Теперь поговорим немного об играх типа «математические развлечения». Вот одно из таких развлечений.

При поездке в трамвае (или автобусе) пассажир компостирует проездной талон, содержащий некоторый шестизначный номер. Задача заключается в том, чтобы, расставив между этими цифрами знаки арифметических действий, получить в результате число 100.

Например, если номер талона имеет вид 482973, то это можно сделать так: $4 \cdot 8 + 2 + 9 \cdot 7 + 3 = 100$.

Есть и другие возможности (с применением скобок):

$$4 + 8 - 2 + 9 \cdot (7 + 3) = 100,$$
$$(4 + 8 - 2) \cdot 9 + 7 + 3 = 100,$$
$$4 \cdot (8 - 2 + 9 + 7 + 3) = 100.$$

Можно также две рядом стоящие цифры рассматривать как двузначное число: $(4 + 8) : 2 + 97 - 3 = 100$ и т. д.

Множество таких развлечений можно найти в книгах замечательного популяризатора Я. И. Перельмана. Решать подобные задачи гораздо интереснее, чем выполнять рутинные упражнения на сложение чисел, и не случайно задачи такого типа стали появляться уже в школьных учебниках.

Увлекательны и другие, ставшие уже классическими, математические развлечения — лабиринты, задачи на разрезание и складывание фигур, нахождение нескольких стертых цифр при выполнении умножения столбиком и т. п. Все эти игры укладываются в разобранный нами тип конечных позиционных игр с полной информацией, а процесс нахождения решения (достижения завершающей позиции) способствует совершенствованию вычислительных навыков, развитию элементов числовой комбинаторики, геометрической интуиции.

Особенно интересна в этом плане *компьютерная* реализация таких игр. За последние десятилетия рынок наводнился тысячами разнообразных компьютерных игр и развлечений. В некоторых из них надо сбить самолет «противника», поймать «преступника», убежать от хищных животных, выиграть гонку и т. п. Такими играми достигается тренировка быстроты реагирования (в связи с чем они нередко применяются в качестве тренажеров). Но есть и другие игры — действительно позиционные, в которых цель достигается за некоторое число шагов, причем на каждом шаге надо сравнительно небольшое число возможностей выстроить в виде «правильной» цепочки, приводящей к очередному успеху (например: найти ключ, выпустить на свободу «недоваренного» кота и т. п.). Компьютер оказался очень удобным инструментом для реализации таких игр — каждая позиция игры (вершина графа) изображается на мониторе, в каждой позиции простым нажатием клавиши мыши выбирается допустимый ход (ребро графа), переводящий в другую позицию, причем часто компьютер сам сообщает, правильный ли сделан ход.

Подобную реализацию допускают многие математические игры и развлечения и таких компьютерных реализаций для интеллектуальных, развивающих игр появляется все больше и больше. В компьютерном оформлении игра (скажем, выбор предметов из одной или двух кучек, задача на переливание жидкостей и т. п.) становится гораздо более привлекательной, а реализация игры — более удобной и наглядной. Осваивающим программирование можно предложить такое упражнение: «поставить на компьютер» игры, описанные в этой беседе.

Задачи и упражнения

201. Разберите игру в «крестики-нолики» на доске 3×3, при которой каждый играющий может своим ходом поставить как «крестик», так и «нолик». Выигрывает, как обычно, тот, после хода которого на доске оказывается три одинаковых значка, стоящих на одной прямой.

202. Двое играют в такую игру: из кучки, в которой находится 25 спичек, берут по очереди — каждый себе — одну, две или три спички. Выигрывает тот, кто берет последнюю спичку. Кто выигрывает в этой игре — начинающий или его противник?

203. Рассмотрим игру «двойные шахматы», в которой все правила остаются прежними, но каждый игрок делает сразу два хода. Докажите, что начинающий не проигрывает при правильной игре.

204. У ромашки n лепестков. Двое по очереди отрывают лепестки. Разрешается отрывать один лепесток или два соседних лепестка. Выигрывает тот, кто оторвет последний лепесток. Кто из играющих может обеспечить себе победу?

205. На окружности расставлено 20 точек, являющихся вершинами правильного двадцатиугольника. За один ход разрешается соединить любые две из них отрезком, не пересекающим (внутри окужности) отрезков, проведенных ранее. Проигрывает тот, кто не сможет сделать ход. Кто?

51. Понятие о сетевом планировании

Подготавливается план осуществления некоторого мероприятия, скажем, постройка дачного домика. Вся работа разбивается на отдельные операции: составление плана, прокладка дороги по участку, подвоз материалов, сооружение фундамента, изготовление бревенчатого сруба, сборка стропил для крыши, настилка пола, установка оконных блоков, внутренняя отделка, установка дверных блоков, кровельные работы, оборудование чердака, установка водоподогревателя и отопления, проводка электричества, дымоход, лестница на чердак, запоры на окна и двери. Составленный проект изображается в виде ориентированного графа, вершинами которого служат результаты реализации отдельных *операций*, а ребра изображают *выполнение* операций и их зависимость. Для каждой операции определяются *предшествующие* ей операции. Например, установке оконных блоков предшествуют изготовление бревенчатого сруба и настилка полов, тогда как установление оконных блоков и установление дверных блоков не зависят друг от друга (т. е. эти вершины ребром не соединены). В результате появляется ориентированный граф («сетевой план», или ПЕРТ), на ребрах которого надписываются планируемые сроки выполнения каждой операции (рис. 323). Этот граф не должен содержать ориентированных циклов (поскольку никакая операция не может предшествовать самой себе), т. е. он устанавливает *упорядочение* в множестве его вершин. Стрелки, представляющие собой следствия по транзитивности, не изображаются.

На рис. 324 предыдущий граф изображен в виде диаграммы Хассе, в которой вершины располагаются снизу вверх в виде нескольких ярусов. Такое изображение графа всегда можно выполнить (возможно, с пересечениями ребер). Это позволяет более наглядно представить зависимость операций друг от друга и их последовательность: например, вершина **4** не может быть достигнута, пока не пройдены вершины **2** и **3**.

Для исполнителя проекта особо важную роль играет *критический путь*, который представляет собой *самую неблагоприятную* (самую напряженную в смысле времени) ориентированную цепь, ведущую от начальной операции к завершающей, т. е. цепь, которая имеет *максимальную* сумму длительностей операций. На рис. 323, 324 критический путь вычерчен жирно. Сумма длительностей операций, взятая вдоль критического пути, представляет собой время *выполнения* проекта в целом.

Вершины, по которым проходит критический путь, называются *критическими операциями*. На их выполнение руководитель проекта должен обратить особое внимание, так как любая задержка в выполнении какой-либо из этих операций означает задержку завершения всей работы («сдачи объекта»). Напротив, некритические операции допускают некоторое запаздывание в их выполнении (что дает руководителю некоторые возможности переброски рабочей силы для поддержки критических операций).

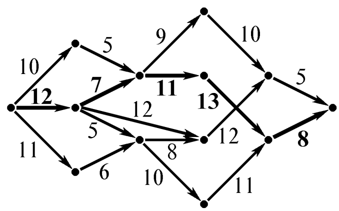


Рис. 323

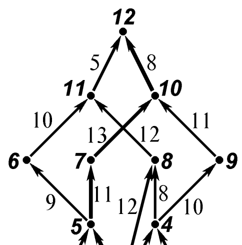


Рис. 324

Рассмотренный метод (*сетевое планирование*) представляет собой важное практическое применение теории графов. Для достаточно сложных проектов, содержащих много вершин, нахождение критического пути может представлять собой трудоемкую задачу, без решения которой однако нельзя обоснованно судить о времени возможной реализации проекта. В связи с этим созданы специальные математические *алгоритмы* нахождения критического пути. Некоторые из них были разработаны при участии Ричарда Беллмана, руководителя математического отдела американской корпорации RAND.

Сказанное выше относится к случаю, когда длительности выполнения операций точно определены (детерминированы). На практике же могут иметь место случаи, когда некоторые длительности выполнения операций зависят от случайных обстоятельств

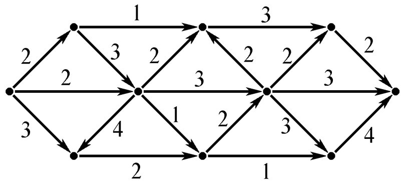


Рис. 325

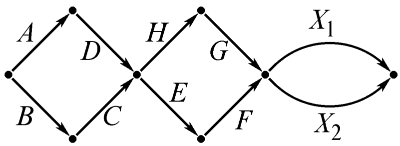


Рис. 326

(задержка выполнения поставок смежниками из других ведомств, влияние погодных условий и т. п.). В таком случае задача принимает стохастический (вероятностный) характер. Если известно распределение вероятностей для длительности выполнения отдельных операций, то можно вычислить *математическое ожидание* времени завершения проекта. Обычно распределение вероятностей оценивается (а иногда и корректируется в процессе работы) экспертами (метод *экспертных оценок*), а вычисления проводятся с помощью компьютеров, т. е. возникают «человеко-машинные» системы сетевого планирования.

Задачи и упражнения

206. Укажите на сетевом графике (рис. 325) критический путь. Чему равно время осуществления проекта?

207. Составьте сетевой график подготовки и проведения шахматного турнира в школе.

208. При проведении ремонта квартиры выполняются следующие операции: а) покупка обоев; б) покупка краски и других малярных материалов; в) шпаклевка потолков и стен; г) принятие решения о ремонте; д) договор с малярами; е) оклеивание стен обоями; ж) окраска дверей и оконных рам; з) побелка потолков; и) шпаклевка дверей и оконных рам.

Составьте сетевой график проведения ремонта квартиры.

209. При компьютерном решении некоторой задачи последовательность вычислений удалось разбить на несколько этапов. При этом каждый этап требует результатов некоторых других этапов и дает данные для проведения некоторых других этапов. На рис. 326 приведен сетевой график процесса решения этой задачи. Для решения задачи задействована сеть из нескольких работающих параллельно компьютеров, результаты вычислений каждого поступают в общую базу данных, но каждый этап вычислений должен выполняться целиком на одном компьютере. Если каждый этап вычислений выполняется в течение 10 секунд машинного времени и в сети достаточно компьютеров, то сколько машинного времени займут вычисления вдоль критического пути? Какое минимальное количество компьютеров будет при этом задействовано?

210. При перестановке мебели в малогабаритной квартире для установки нового шкафа оказалось необходимым произвести следующие операции: а) из прихожей вешалка убирается на кухню; б) шкаф заносится в прихожую; в) из гостиной сервант переносится в спальню; г) на место серванта ставится шкаф; д) вешалка возвращается в прихожую; е) диван переносится на кухню; ж) на место дивана ставится сервант; з) диван переезжает в спальню.

Составьте сетевой график и определите те операции, которые можно выполнить до приезда шкафа.

ГЛАВА III. РАССУЖДЕНИЯ

Беседа 11. Теоремы

52. Существование и общность

Предложение, о котором можно однозначно установить, истинно оно или ложно, называют *высказыванием*. Это, конечно, не точное определение, а лишь пояснение. Например, предложения:

A: «$3 + 6 = 9$»;

B: «Москва — столица России»;

C: «через любые две точки проходит единственная прямая»;

D: «для любых трех точек A, B, C сумма длин отрезков AB и BC не меньше длины отрезка AC» (рис. 327)

являются истинными высказываниями, а предложения:

E: «Париж — столица Италии»;

F: «$2\times2 = 5$»

— ложными высказываниями.

Истинность или ложность некоторого высказывания вовсе не должны быть немедленно очевидными. Например, высказывание ***C*** представляет собой в евклидовой геометрии некоторую *аксиому*, т. е. истинность этого высказывания принимается в этой геометрии без доказательства. А высказывание ***D*** представляет собой *теорему* евклидовой геометрии, и истинность этого высказывания устанавливается при помощи некоторого рассуждения (доказательства).

В качестве еще одного примера рассмотрим высказывание

G: «на 1 000 000 000-м месте после запятой в десятичном разложении числа π стоит цифра 5».

Мы не можем сказать, истинно это высказывание или ложно, поскольку такого количества цифр числа π сегодня не знает никто. Но принципиально решить этот вопрос можно, т. е. можно узнать, истинно или ложно утверждение ***G***. Например, известно равенство $\frac{\pi}{4} = 1 - \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots$ и существует точная оценка того, какова величина ошибки, которую мы сделаем, оборвав этот ряд на n-м члене, т. е. принципиально возможно вычислить любую цифру числа π (хотя требуемое для нахождения 1 000 000 000-й цифры этого числа количество вычислений столь велико, что нынешние компьютеры не могут их выполнить в обозримый срок).

А утверждение «Сегодня хорошая погода» высказыванием не будет, даже если уточнить дату, под которой понимается «сегодня», и уточнить в каком именно месте земной поверхности обнаружива-

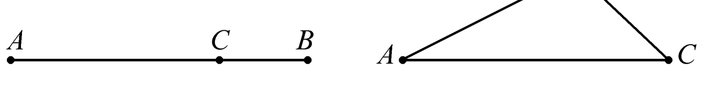


Рис. 327

ется «хорошая погода». Ведь смысл термина «хорошая погода» однозначно не определен, и потому вопрос об истинности или ложности высказанного утверждения может не иметь однозначного ответа. Никакое вопросительное предложение также не является высказыванием. Например, бессмысленно задавать вопрос о том, истинно или ложно вопросительное предложение «Сколько будет $57+168$?».

Первое систематическое исследование о высказываниях и операциях над ними (о чем мы будем вести разговор ниже) было предпринято гениальным древнегреческим ученым Аристотелем, который заложил начала логики. Математическое исследование этих вопросов ведет свое начало от основополагающего труда Джорджа Буля, изданного в Лондоне в 1854 году. Этот труд Буля положил начало математической логики, систематическое развитие которой было достигнуто работами многих математиков XX века. Некоторое первоначальное представление о специфике этой области математики и о ее приложениях (в частности, к школьному курсу математики) мы и предложим читателю в этой главе. А в дальнейших беседах мы несколько расширим эти первоначальные представления.

Кроме высказываний в логике рассматриваются *предикаты*. Предикат — это как бы высказывание, содержащее неизвестную (или несколько неизвестных).

Рассмотрим следующее предложение, содержащее переменную n: «n делится на 7». Обозначим это предложение через $A(n)$. Предложение $A(n)$ высказыванием не является: о его истинности мы ничего не можем сказать, пока не указано, какое конкретное значение принимает n. Например, при $n=1$ получается высказывание $A(1)$, которое словами выражается так: «1 делится на 7». Это высказывание ложно. Высказывание же $A(14)$, т. е. «14 делится на 7» — истинно.

Итак, $A(n)$ представляет собой предложение, которое при каждом конкретном n превращается в некоторое высказывание. Такие предложения в математике называются *предикатами*. В данном случае предикат $A(n)$ содержит переменную n, которая может принимать значения из множества натуральных чисел. Поэтому говорят, что предикат $A(n)$ задан на множестве натуральных чисел.

Замечание. Слово «предикат» в переводе с латинского означает «сказуемое» (оно употребляется также в немецком, английском и других языках). Например, рассмотренное выше предложение $A(n)$ содержит в качестве подлежащего букву n, а вся остальная часть предложения представляет собой сказуемое «делится» с подчиненными ему словами. Подставляя вместо n то или иное подлежащее (в данном случае числительные 1, 14 и т. д.), мы получаем некоторые высказывания. В дальнейшем в этом пункте высказывания и предикаты мы иногда будем заключать в фигурные скобки.

Аристотель ограничился в своей логике рассмотрением предикатов только от одной переменной. Но после работ Буля в рассмотрение вошли и предикаты от нескольких переменных. Вот пример предиката от двух переменных: $\{$Писатель W написал книгу $b\}$. Это пока еще не высказывание, поскольку нам не указано, о каком писателе и о какой книге идет речь. Но если вместо W и b мы подставим какие-либо конкретные значения, получится некоторое высказывание. Например, подставив W = Л. Н. Толстой, b = «Война и мир», мы получим истинное высказывание

{Писатель Л. Н. Толстой написал книгу «Война и мир»}.

А при подстановке W = И. С. Тургенев, b = «Анна Каренина» получится ложное высказывание.

Дадим теперь интерпретацию предикатов на языке теории множеств. Рассмотренный выше предикат $\boldsymbol{A}(n)$ задан на множестве $\boldsymbol{N}$ всех натуральных чисел. Множество $\boldsymbol{A} \subset \boldsymbol{N}$, состоящее из чисел, делящихся на 7: $\boldsymbol{A} = \{7, 14, 21, 28, 35, 42, 49, \ldots\}$, представляет собой множество всех тех значений n, для которых $\boldsymbol{A}(n)$ превращается в *истинное* высказывание. Его дополнение $c\boldsymbol{A} = \boldsymbol{N} \setminus \boldsymbol{A}$ представляет собой множество всех тех значений n, при подстановке которых рассматриваемый предикат превращается в *ложное* высказывание. Например, $84 \in \boldsymbol{A}$, потому что $\boldsymbol{A}(84)$ — истинное высказывание (число 84 делится на 7), а $37 \in c\boldsymbol{A}$, потому что $\boldsymbol{A}(37)$ — ложное высказывание.

Вообще, если задано некоторое универсальное множество $\boldsymbol{U}$, на котором определен предикат $\boldsymbol{B}(x)$, то с точки зрения теории множеств это означает, что выделено некоторое подмножество $\boldsymbol{B} \subset \boldsymbol{U}$, состоящее из всех $x \in \boldsymbol{U}$, при подстановке которых $\boldsymbol{B}(x)$ превращается в *истинное* высказывание. Его дополнение $c\boldsymbol{B}$ состоит из всех $x \in \boldsymbol{U}$, при подстановке которых $\boldsymbol{B}(x)$ превращается в *ложное* высказывание. Это множество $\boldsymbol{B}$ называется *множеством истинности* предиката $\boldsymbol{B}(x)$.

Рассмотрим теперь предикаты от двух переменных. Указанный выше предикат о писателях и книгах можно записать так:

$$\boldsymbol{L}(W, b) = \{\text{Писатель } W \text{ написал книгу } b\}.$$

Здесь переменная W пробегает множество $\boldsymbol{A}$ всех писателей, а переменная b пробегает множество $\boldsymbol{B}$ всех книг. Каждая пара $(W; b)$ представляет собой элемент множества $\boldsymbol{A}\times\boldsymbol{B}$, т. е. произведения множеств $\boldsymbol{A}$ и $\boldsymbol{B}$:

B (книги)	*A* (писатели)			
	Л. Н. Толстой	И. С. Тургенев	А. С. Пушкин	И. А. Гончаров
«Война и мир»	✔ ●	●	●	●
«Анна Каренина»	✔ ●	●	●	●
«Обрыв»	●	●	●	✔ ●
«Евгений Онегин»	●	●	✔ ●	●

Здесь точками обозначены элементы произведения $\boldsymbol{A}\times\boldsymbol{B}$, т. е. всевозможные пары $(W; b)$, а «галочками» отмечено множество истинности $\boldsymbol{L} \subset \boldsymbol{A} \times \boldsymbol{B}$, т. е. множество тех пар $(W; b)$, при подстановке которых рассматриваемый предикат превращается в *истинное* высказывание. Дополнение же $\boldsymbol{A} \times \boldsymbol{B} \setminus \boldsymbol{L}$ состоит из пар, при подстановке которых получается *ложное* высказывание.

Если рассматривается предикат $\boldsymbol{K}(a, b, c)$ от трех переменных $a \in \boldsymbol{A}$, $b \in \boldsymbol{B}$, $c \in \boldsymbol{C}$, то его множество истинности представляет собой некоторое подмножество $\boldsymbol{K} \subset \boldsymbol{A} \times \boldsymbol{B} \times \boldsymbol{C}$.

Множества истинности предикатов нередко обозначают с помощью фигурных скобок. Например, запись $\boldsymbol{A} = \{x \in \boldsymbol{N}\colon\ \boldsymbol{A}(x)\}$ означает, что $\boldsymbol{A}$ есть множество всех тех $x \in \boldsymbol{N}$, для которых имеет место $\boldsymbol{A}(x)$, т. е. множество истинности этого предиката. Более подробно:

$$\boldsymbol{A} = \{x \in \boldsymbol{N}\colon\ x \text{ делится на } 7\}.$$

В качестве второго примера рассмотрим запись

$$\boldsymbol{M} = \{M \in \boldsymbol{R}^2\colon\ AM = BM\}.$$

Она означает множество всех точек M (взятых на плоскости $\boldsymbol{R}^2$), для которых предикат $AM = BM$ превращается в истинное высказывание, т. е. множество всех точек плоскости, равноудаленных от двух заданных точек A и B. Как известно из школьного курса геометрии, это множество представляет собой *серединный перпендикуляр* к отрезку AB.

Вернемся к высказываниям $\boldsymbol{C}$ и $\boldsymbol{D}$, рассмотренным выше. В них используются выражения «любые две точки», «для любых трех точек». Часто приходится формулировать и другие высказывания, в которых используются слова «все», «всякий», «каждый», «любой» и т. п. В математике принято вместо этих слов использовать знак $\forall$ (перевернутая первая буква A английского слова *All* — все), который называется *знаком общности*.

Рассмотрим следующие два утверждения:

(1) *сумма величин углов треугольника ABC равна* 180°;

(2) *в треугольнике ABC угол A прямой*.

Чтобы подчеркнуть общность утверждения (1), т. е. его справедливость для любого треугольника ABC, используется знак $\forall$: $(\forall\ \triangle ABC)\ (\angle A + \angle B + \angle C = 180°)$.

Читается: «для любого треугольника ABC сумма величин углов A, B и C равна 180°». Это — истинное высказывание.

Утверждение (2) такой общностью не обладает. И если перед этим утверждением записать $(\forall\ \triangle ABC)$, то получается ложное высказывание. Например, в треугольнике, изображенном на рисунке 328, угол A не прямой.

В качестве еще одного примера возьмем сформулированное выше высказывание $\boldsymbol{D}$. В нем используется предикат $AB + BC \geq AC$, в котором имеются *три* переменные: точки A, B, C. Само же высказывание $\boldsymbol{D}$ можно записать следующим образом:

$$(\forall \text{ т. } A, B, C)\ (AB + BC \geq AC).$$

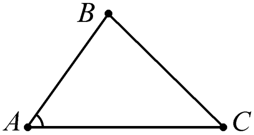


Рис. 328

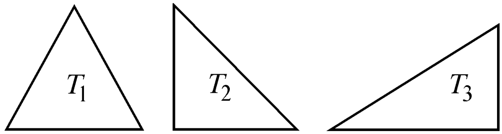


Рис. 329

Знак общности в этой записи показывает, что неравенство треугольника справедливо для *любых* точек A, B и C, а обозначение «т.» — сокращение слова «точка» (в дальнейшем для сокращения слова «прямая» будет применяться обозначение «пр.»).

Вообще, если к некоторому предикату, содержащему, скажем, переменную x, добавить слева запись $(\forall\ x)$, т. е. «навесить» на переменную x знак общности (или, как говорят математики, *квантор общности*), то получится некоторое высказывание, истинное или ложное.

Возьмем, например, следующий предикат, определенный на множестве $\boldsymbol{H}$ всех людей: $\boldsymbol{A}(x) = \{x$ имеет две ноги$\}$.

Навесив на переменную x квантор общности, мы получим высказывание $(\forall\ x \in \boldsymbol{H})\ \boldsymbol{A}(x)$, т. е. словами: «*любой* человек x имеет две ноги». Это высказывание, к сожалению, ложно.

Рассмотрим теперь еще один квантор, применяемый в математической логике — так называемый *квантор существования*.

Нам часто приходится формулировать высказывания, в которых используются слова: *хотя бы один*, *найдется*, *существует* и т. д. В математике принято вместо этих слов использовать знак $\exists$ (перевернутая первая буква E английского слова *Exists* — «существует»).

Рассмотрим, например, три предиката, заданные на множестве всех треугольников:

$\boldsymbol{A}(T) = \{$в треугольнике T сумма величин углов равна $180°\}$;

$\boldsymbol{B}(T) = \{$один из углов треугольника T — прямой$\}$;

$\boldsymbol{C}(T) = \{$два угла треугольника T — прямые$\}$.

Предикат $\boldsymbol{A}(T)$ для всякого конкретного треугольника T превращается в истинное высказывание. Иными словами, высказывание $(\forall\ T)\ \boldsymbol{A}(T)$ истинно (в любом треугольнике сумма величин углов равна $180°$).

Предикат $\boldsymbol{B}(T)$ не для всякого конкретного треугольника T превращается в истинное высказывание; например, для треугольника T_1 на рис. 329 высказывание $\boldsymbol{B}(T_1)$ ложно. Поэтому мы не можем утверждать, что *для любого* T имеет место $\boldsymbol{B}(T)$ — не в любом треугольнике один из углов прямой. Иными словами, высказывание $(\forall\ T)\ \boldsymbol{B}(T)$ ложно. Вместе с тем *существуют* такие треугольники, для которых $\boldsymbol{B}(T)$ истинно (например, треугольники T_2, T_3 на рис. 329). Этот факт записывают с помощью квантора существования $\exists$:

$(\exists\ T)\ \boldsymbol{B}(T)$. Словами эта запись читается так: существует T, для которого справедливо $\boldsymbol{B}(T)$ (т. е. существует треугольник T, один из углов которого — прямой).

И, наконец, предикат $\boldsymbol{C}(T)$ ни для какого треугольника T не превращается в истинное высказывание (в Евклидовой геометрии).

Задачи и упражнения

211. Является ли высказыванием утверждение «У марсиан по три глаза»?

212. Найдите множество истинности следующего высказывания: сумму в n рублей можно уплатить купюрами в 3 и 5 рублей.

213. Используя кванторы, запишите высказывание: «Уравнение $ax=b$ имеет решение». Найдите множество истинности этого высказывания.

214. Используя кванторы, запишите высказывание: «Уравнение $ax=b$ имеет положительный корень». Найдите множество истинности этого высказывания.

215. Используя кванторы, запишите высказывание: «Существует целое число, которое делится на любое другое целое число, отличное от нуля». Истинно ли это высказывание?

53. Структура теоремы

В школьном курсе математики устанавливаются различные свойства геометрических фигур. В качестве примера рассмотрим следующее хорошо известное утверждение: *диаметр*, *перпендикулярный хорде*, *делит эту хорду пополам* (рис. 330). Иными словами, если через центр окружности провести прямую, перпендикулярную хорде AB, то проведенная прямая пройдет через середину этой хорды.

Проверить это можно было бы опытным путем: начертить окружность, провести в ней хорду и перпендикулярный ей диаметр, а затем, перегнув чертеж по диаметру, убедиться, что половинки хорды совмещаются. Такую проверку можно было бы повторить несколько раз для различных окружностей и хорд.

Однако сколько бы таких проверок мы ни произвели, это не может нас убедить в справедливости высказанного утверждения. Ведь по смыслу оно имеет *общий* характер, т. е. утверждается, что *для любой* хорды перпендикулярный к ней диаметр делит ее пополам. А общее утверждение, относящееся к любой хорде, не может быть доказано несколькими пробами. Для того, чтобы установить истинность некоторого общего утверждения, нужно провести *рассуждение*, убеждающее нас в его справедливости. Мы уже говорили об этом в п. 21 при рассказе о методе математической индукции. И справедливость сформулированного выше утверждения (о хорде и перпендикулярном ей диаметре) была установлена в курсе геометрии именно с помощью некоторого рассуждения, которое убедило нас в том, что это высказывание справедливо для любой окружности и любой ее хорды.

В математике часто встречаются общие высказывания, справедливость которых устанавливается при помощи рассуждений. Такие высказывания называют *теоремами*, а рассуждение, которое убеждает в справедливости теоремы, называют *доказательством*.

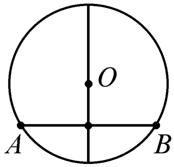


Рис. 330

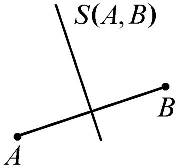


Рис. 331

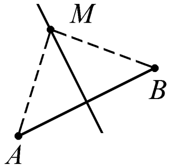


Рис. 332

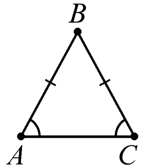


Рис. 333

Таким образом, высказывание «диаметр, перпендикулярный хорде, делит эту хорду пополам» является теоремой.

Нередко в формулировках теорем встречаются выражения «если ... то», «из ... следует ...», «из ... вытекает ...» и т. д. Например, рассмотренную выше теорему можно сформулировать так: *если* диаметр перпендикулярен хорде, *то* он делит эту хорду пополам (или: из перпендикулярности диаметра к хорде *вытекает*, что он делит ее пополам).

Условимся серединный перпендикуляр к отрезку AB (т. е. ось симметрии точек A и B, рис. 331), обозначать через $S(A, B)$. Рассмотрим в качестве еще одного примера следующую теорему: если расстояние от точки M до точек A и B одинаково, то точка M принадлежит серединному перпендикуляру к отрезку AB (рис. 332). Тогда рассматриваемая теорема кратко формулируется так:

$$\text{если } AM = BM, \text{ то } M \in S(A, B).$$

При этом теорема имеет *общий* характер, т. е. она справедлива для *любых* точек A, B, M. Иными словами, какие бы три точки A, B, M мы ни взяли, если только $AM = BM$, то непременно точка M принадлежит прямой $S(A, B)$. Поэтому теорема может быть записана так:

$$(\forall \text{ т. } A \neq B, M)\ (AM = BM) \Rightarrow (M \in S(A, B)). \qquad (3)$$

Здесь знак $\Rightarrow$ используется вместо словесных выражений «если ... то ...», «из ... следует ...» и т. д. Он называется *знаком импликации*.

Аналогичным образом могут быть записаны и другие теоремы: сначала ставится *разъяснительная часть* теоремы, т. е. запись вида $(\forall\ \triangle ABC)$, $(\forall \text{ т. } A, B, M)$ и т. п., а затем пишутся два утверждения, соединенные знаком $\Rightarrow$. Первое из этих утверждений (стоящее после разъяснительной части и перед знаком $\Rightarrow$) называется *условием* теоремы, второе (стоящее после знака $\Rightarrow$) называется *заключением* теоремы. Например, одно из свойств равнобедренного треугольника может быть записано в виде следующей теоремы (рис. 333): $(\forall\ \triangle ABC)\ (AB = CB) \Rightarrow (\angle A = \angle C)$.

Здесь разъяснительная часть состоит в том, что рассматривается любой треугольник, вершины которого обозначены буквами A, B и C. Запись $AB = CB$ выражает условие теоремы (в рассматриваемом треугольнике ABC стороны AB и CB имеют одинаковые длины, т. е.

конгруэнтны). Наконец, запись $\angle A = \angle C$ выражает заключение теоремы (величины углов A и C равны, т. е. эти углы конгруэнтны).

Нередко эту теорему кратко формулируют так: в равнобедренном треугольнике величины углов при основании равны. Здесь подразумевается (но явно это не высказано), что теорема имеет *общий* смысл, т. е. в любом треугольнике ABC, если только $AB = CB$, то обязательно и $\angle A = \angle C$. Общий характер этой теоремы явно указан в записи ($\forall\ \triangle ABC$).

Выше мы говорили, что импликация выражается словами: «если ... то», «вытекает», «следует» ... Но в каком случае мы можем применять эти слова? В каком случае импликация имеет место? Принято считать, что если $\boldsymbol{A}$ и $\boldsymbol{B}$ — два высказывания, то импликация только в том случае является ложным высказыванием, если $\boldsymbol{A}$ истинно, а $\boldsymbol{B}$ — ложно. Иными словами, из истины не вытекает ложь. Во всех остальных случаях импликация считается истинным высказыванием. Это отмечено в следующей *таблице истинности*, которую следует рассматривать как *определение* операции импликации:

$\boldsymbol{A}$	$\boldsymbol{B}$	$\boldsymbol{A \Rightarrow B}$
И	**И**	**И**
И	**Л**	**Л**
Л	**И**	**И**
Л	**Л**	**И**

Последние две строки этой таблицы означают, что из ложного высказывания вытекает любое высказывание (как истинное, так и ложное). По меткому выражению Ф. Хаусдорфа, автора одной из первых книг по теории множеств и топологии (к тому же великолепно написанной), «если дважды два — пять, то существуют ведьмы».

Возьмем в качестве иллюстрации теорему (3). Пусть A, B, M — три произвольные точки, $A \neq B$. Если окажется, что расстояния AM и BM не одинаковы, то высказывание $AM = BM$ ложно, а из ложного высказывания вытекает *что угодно* и, в частности, $M \in S(A, B)$. Если же расстояния AM и BM одинаковы (рис. 332), то, как показывает несложное рассуждение, включение $M \in S(A, B)$ справедливо. Таким образом, *в любом случае* импликация, указанная в теореме (3) имеет место. Это и выражено в записи (3) с помощью указания ($\forall$ т. $A \neq B, M$), т. е. могут быть взяты *любые* конкретные точки A, B, M (где A и B различны), и тогда импликация $(AM = BM) \Rightarrow (M \in S(A, B))$ представляет собой *истинное* высказывание.

Часто разъяснительная часть теоремы начинается словом «пусть», а переход к условию теоремы отмечается словом «тогда». Вот как в этом случае формулируются теоремы, рассмотренные выше.

а) Пусть A, B, M — три произвольные точки; тогда если расстояния от точки M до точек A и B одинаковы, то M принадлежит серединному перпендикуляру отрезка AB.

б) Пусть ABC — произвольный треугольник; тогда если стороны AB и CB конгруэнтны, то величины углов A и C равны.

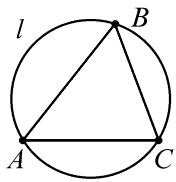


Рис. 334

Важно заметить, что как условие теоремы, так и ее заключение являются *предикатами*, а разъяснительная часть постулирует *общий характер* теоремы. Опять рассмотрим в качестве примера теорему (3). Здесь $AM = BM$ (условие теоремы) есть предикат от трех переменных A, B, M. При подстановке вместо A, B, M *конкретных* точек он превратится в некоторое высказывание — истинное или ложное (в зависимости от того, какие точки взяты). Заключение теоремы $M \in S(A, B)$ также представляет собой предикат от трех переменных A, B, M. Знак импликации, соединяющий эти предикаты, и квантор общности ($\forall$ т. $A \neq B, M$), стоящий слева, дают *высказывание* (а не предикат), т. е. дают полную формулировку этой теоремы.

Кроме теорем, записываемых в форме импликации (как, например, (3)) в математике рассматриваются также теоремы существования, формулируемые символически с помощью квантора $\exists$. Приведем несколько типично «школьных» примеров теорем существования.

Запись

$$(\forall \ \triangle ABC)\ (\exists!\ \text{окр.}\ l)\ (A, B, C \in l) \tag{4}$$

выражает теорему о существовании *описанной окружности* для любого треугольника (рис. 334). Знак $\exists!$ (квантор существования с восклицательным знаком) означает: «существует и притом только один». Таким образом, словесно теорема формулируется следующим образом: для любого треугольника существует, и притом только одна, окружность, проходящая через все три его вершины.

Запись

$$(\forall \ a \neq 0, b)\ (\exists!\ x)\ (ax = b) \tag{5}$$

утверждает, что для любых чисел a и b, первое из которых отлично от нуля, *существует и притом только одно* число x, являющееся решением уравнения $ax = b$. Для полноты можно было бы указать, что рассматриваются *действительные* числа, т. е. записать теорему более подробно:

$$(\forall \ a, b \in \boldsymbol{R}, a \neq 0)\ (\exists!\ x \in \boldsymbol{R})\ (ax = b).$$

В качестве еще одного примера укажем *основную теорему алгебры*. В ней идет речь о многочленах, коэффициентами которых могут быть произвольные *комплексные числа* (множество комплексных чисел, обозначаемое обычно буквой $\boldsymbol{C}$, является расширением множества действительных чисел). Теорема (установленная Гауссом) утверждает, что каждый многочлен произвольной степени $n \geq 1$ (с действительными или комплексными коэффициентами) имеет *хотя бы один корень*

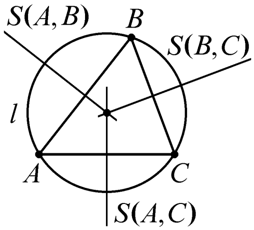


Рис. 335

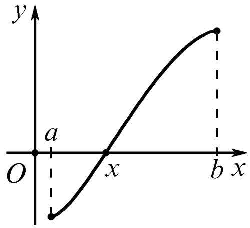


Рис. 336

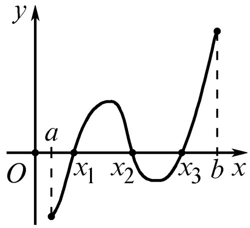


Рис. 337

— действительный или комплексный. Вот символическая запись этой теоремы:

$$(\forall\ n \geq 1;\ a_1, \ldots, a_n \in \boldsymbol{C})\ (\exists\ x \in \boldsymbol{C})\ (x^n + a_1 x^{n-1} + \ldots + a_n = 0). \qquad (6)$$

Основная теорема алгебры представляет собой типичную теорему существования, т. е. в ней лишь утверждается наличие хотя бы одного корня, но ничего не говорится о том, как его можно найти. В этом она существенно отличается от теорем (4) и (5). Например, для теоремы (4) имеется четко описанный прием, позволяющий *найти* эту описанную окружность (надо провести к сторонам перпендикуляры через их середины, и тогда точка пересечения этих перпендикуляров как раз будет центром описанной окружности (рис. 335)). Для теоремы (6) такого общего приема нет: основная теорема алгебры только *утверждает* существование корня.

Несмотря на то, что теоремы существования, как правило, лишь утверждают *факт существования* какого-либо математического объекта без указания способа нахождения этого объекта (как это было, например, в случае теоремы (6)), они играют очень важную роль в современной математике. Приведем простой пример.

Известная теорема Больцано (теорема о промежуточном значении) записывается следующим образом:

$$(\forall\ \text{непр. ф.}\ f(x)\ \text{на}\ [a;\, b])\ (f(a) < 0, f(b) > 0) \Rightarrow (\exists\ x \in [a;\, b])\ (f(x) = 0).$$

Эта теорема имеет форму импликации, но в заключении теоремы использован квантор существования. Разъяснительная часть теоремы говорит о том, что рассматривается произвольная функция, *непрерывная* на отрезке $[a;\, b]$. Теорема утверждает, что если в точках a и b функция принимает значения разных знаков (в точке a отрицательное, а в точке b положительное), то *существует* на отрезке $[a;\, b]$ хотя бы один корень уравнения $f(x) = 0$, т. е. график этой функции хотя бы в одной точке пересекает ось абсцисс (рис. 336, 337). Например, всякий многочлен третьей степени $f(x) = x^3 + ax^2 + bx + c$ (или, вообще, нечетной степени) имеет хотя бы один действительный корень, поскольку для больших по величине отрицательных x значения много-

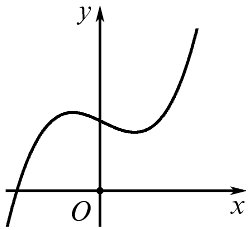


Рис. 338

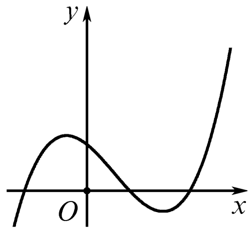


Рис. 339

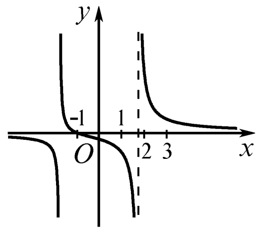


Рис. 340

члена $f(x)$ отрицательны, а для больших положительных x они положительны (рис. 338, 339).

Чтобы пояснить, как работает эта теорема, рассмотрим уравнение

$$x^3 + x^2 - 2x - 3 = 0. \qquad (7)$$

Многочлен $f(x)$, стоящий в левой части этого уравнения, принимает при $x = 1$ отрицательное значение: $f(1) = -3$. А при $x = 2$ он принимает положительное значение: $f(2) = 5$. Значит, на отрезке [1; 2] имеется корень уравнения. Разобьем отрезок [1; 2] на десять равных частей точками 1,1; 1,2; 1,3; ... . Вычисляя значения $f(x)$ в этих точках (это легко сделать с помощью калькулятора), мы найдем, что $f(1{,}5) = -0{,}375 < 0$, а $f(1{,}6) = 0{,}456 > 0$. Значит, искомый корень расположен на отрезке [1,5; 1,6]. Разобьем и этот отрезок на 10 равных частей точками 1,51; 1,52; ... . Теперь находим $f(1{,}54) < 0$, $f(1{,}55) > 0$. Таким образом, искомый корень примерно равен $x = 1{,}54$.

Можем ли мы быть уверенными, что таким путем мы вычисляем именно корень взятого уравнения, а не что-то иное? Такую уверенность дает нам теорема Больцано. Всякий многочлен есть *непрерывная* функция, а потому имеется корень на отрезке [1; 2], на отрезке [1,5; 1,6], на отрезке [1,54; 1,55], ... Таким образом, сам факт *существования* корня делает наши действия осмысленными.

А теперь рассмотрим уравнение $\dfrac{x+1}{x^2-3} = 0$.

Обозначая функцию, стоящую в левой части, через $g(x)$, находим $g(1) = -1 < 0$, $g(2) = 3 > 0$, т. е. в концах отрезка [1; 2] функция принимает значения разных знаков. Исследуя значения функции в точках 1,1; 1,2; ..., находим: $g(1{,}7) = -24{,}5... < 0$, $g(1{,}8) = 11{,}6... > 0$. Далее получаем $g(1{,}73) = -384{,}507... < 0$, $g(1{,}74) = 99{,}275... > 0$. Вроде, как в предыдущем случае, мы бы должны все более точно вычислять корень рассматриваемого уравнения (равный примерно 1,73...), однако подозрительно, что значения функции все более удаляются от нуля!

Дело в том, что рассматриваемая функция $g(x)$ *не является непрерывной* на отрезке [1; 2]; это видно из поведения ее графика (рис. 340). Поэтому у нас нет оснований делать вывод о *существовании* корня, в отличие от предыдущего случая, когда мы рассматривали уравнение (7).

В данном случае все ясно: *знаменатель* функции $g(x)$ принимает нулевое значение при $x=\sqrt{3}\approx 1{,}73$, а потому функция, хотя и меняет знак при переходе через $x=\sqrt{3}$, но «обращается в бесконечность» при этом значении (рис. 340). Однако и в более сложных случаях факт *существования* корня (или какого-либо другого математического объекта) является важным фактором при решении различных задач во многих вопросах математики и ее приложений.

Задачи и упражнения

216. Запишите с помощью кванторов следующую теорему: «Сумма любых трех последовательных целых чисел делится на 3».

217. Запишите с помощью кванторов следующую теорему: «Для любого треугольника существует единственная вписанная окружность».

218. Запишите утверждение: «В течение любого часа существует момент, во время которого минутная и часовая стрелка совпадают». Истинно ли это утверждение?

219. Запишите утверждение: «Среди любых трех целых чисел найдутся два, сумма которых четна». Истинно ли это утверждение?

220. Используя теорему Больцано, докажите, что при любых значениях a, b и c многочлен x^3+ax^2+bx+c имеет действительный корень.

54. Отрицание

Из всякого высказывания можно получить новое высказывание, отрицая его, т. е. утверждая, что исходное высказывание не имеет места, не выполняется. Для записи отрицания используется знак $\neg$ (его называют знаком отрицания или, короче, знаком «не»).

Обозначим через $\boldsymbol{A}$ высказывание $5\cdot 7=35$. Его отрицание можно записать так: $5\cdot 7\neq 35$. В данном случае высказывание $\boldsymbol{A}$ истинно, а $\neg\boldsymbol{A}$ ложно.

Возьмем теперь следующее высказывание $\boldsymbol{B}$: «в нашем классе более 100 учеников». Его отрицание $\neg\boldsymbol{B}$ можно прочитать словами так: «в нашем классе не более 100 учеников». Обычно высказывание $\boldsymbol{B}$ ложно, а его отрицание $\neg\boldsymbol{B}$ истинно.

В качестве геометрического примера рассмотрим следующее высказывание: «в прямоугольном треугольнике с катетами, имеющими длины 4 и 5, длина гипотенузы равна 9». Его отрицание утверждает, что в прямоугольном треугольнике с катетами, имеющими длины 4 и 5, длина гипотенузы *не равна* 9. Здесь первоначальное высказывание ложно, а его отрицание истинно.

Из двух высказываний, одно из которых является отрицанием другого, одно истинно, а другое ложно.

Это положение математической логики называют *законом исключенного третьего*. Иными словами, либо $\boldsymbol{A}$, либо не $\boldsymbol{A}$; третьего быть не может.

Операцию отрицания можно применять не только к высказываниям, но и к предикатам. Обозначим, например, через $\boldsymbol{A}(T)$ следующий предикат, заданный на множестве всех треугольников: «треугольник T — равносторонний». Отрицание $\neg A(T)$ тоже представляет собой предикат: «треугольник T не является равносторонним».

Дважды отрицая какое-либо высказывание A (или, как говорят, взяв его двойное отрицание), мы получаем высказывание $\neg\neg A$, которое означает то же самое, что и первоначальное высказывание A.

Например, отрицанием высказывания $5\cdot 7 = 35$ служит высказывание $5\cdot 7 \neq 35$. Взяв еще раз отрицание, получаем высказывание $\neg\,(5\cdot 7 \neq 35)$. Оно означает, «неверно, что $5\cdot 7 \neq 35$», т. е., иными словами, оно утверждает, что $5\cdot 7 = 35$. Таким образом, двойное отрицание $\neg\,\neg\,(5\cdot 7 = 35)$ означает то же самое, что и первоначальное высказывание $5\cdot 7 = 35$.

Заметим, что существует в математической логике школа *интуиционистов*, которые не признают абсолютной истинности закона исключенного третьего, т. е. считают неправомочной (во всяком случае, не очевидной) возможность отождествлять двойное отрицание $\neg\neg A$ с исходным высказыванием A. Из представителей отечественной науки к школе интуиционизма примыкал известный математик А. А. Марков.

Таким образом, интуиционистская логика *отлична* от излагаемой здесь (и принимаемой подавляющим большинством математиков) формальной логики, которая принимает закон исключенного третьего одной из логических аксиом. В дальнейшем мы будем оставаться в рамках формальной логики и, без всякого сомнения, будем отождествлять высказывания $\neg\neg A$ и A.

Сказанное означает, что операция отрицания определяется следующей таблицей истинности:

A	$\neg A$
И	**Л**
Л	**И**

Из этой таблицы нетрудно заключить, что $\neg\neg A$ совпадает с A, т. е. эта таблица относится всецело к формальной логике.

В том случае, если высказывание выражено простым предложением с одним сказуемым, для образования его отрицания достаточно добавить «не» к сказуемому (а если частица «не» уже стояла перед сказуемым, то для образования отрицания достаточно опустить эту частицу). Например:

A = {Париж находится в Италии};

$\neg A$ = {Париж не находится в Италии}.

Другой пример получим, рассматривая рис. 341, на котором точка C — середина стороны BD треугольника ABD:

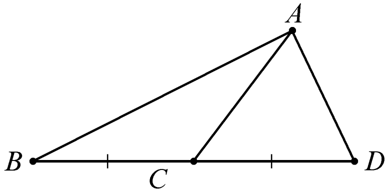


Рис. 341

$\boldsymbol{B}$ = {площади треугольников ABC и ADC на рис. 341 равны};

$\neg\boldsymbol{B}$ = {площади треугольников ABC и ADC на рис. 341 не равны}.

Однако такой простой способ образования отрицаний применим только в том случае, если в высказывании нет знаков $\forall$, $\exists$ (или равнозначных им слов). Чтобы пояснить сложность образования отрицаний для высказываний, содержащих кванторы, обратимся к следующему древнему софизму. В нем говорится о лжецах (всегда говорящих неправду), остальные же люди (не лжецы) всегда говорят правду. Вот этот софизм.

Один критянин (т. е. житель острова Крит) утверждал, что *все критяне лжецы*. Но он сам критянин и потому лжец. Значит, он сказал неправду, т. е. все критяне вовсе не лжецы, а *правдивые люди*. Но тогда он (как критянин) тоже правдивый человек. Значит, он все же сказал правду, и потому *все критяне лжецы*. Но это значит, что и он лжец, т. е. он сказал неправду, а потому *критяне правдивы*...

Как же выйти из этого порочного круга, т. е. как раскрыть этот софизм? На этот вопрос мы ответим чуть позже, когда детально рассмотрим взаимодействие отрицания с кванторами.

Рассмотрим, например, высказывание

$\boldsymbol{P}$ = {в любом треугольнике ABC величины углов A и B равны}.

Если просто добавить «не» к сказуемому, то получится высказывание

$\boldsymbol{Q}$ = {в любом треугольнике ABC величины углов A и B не равны}.

Легко видеть, что высказывание $\boldsymbol{Q}$ не является отрицанием высказывания $\boldsymbol{P}$ (ведь оба высказывания $\boldsymbol{P}$ и $\boldsymbol{Q}$ — ложные, а из двух высказываний $\boldsymbol{P}$ и $\neg\boldsymbol{P}$ одно должно быть истинным).

Как же составить отрицание высказывания $\boldsymbol{P}$? Запишем это высказывание знаками:

$$\boldsymbol{P} = (\forall\ \triangle ABC)\ (\angle A = \angle B).$$

Отрицание его получается, если поставить знак $\neg$ перед всем высказыванием:

$$\neg\ \boldsymbol{P} = \neg\ (\forall\ \triangle ABC)\ (\angle A = \angle B). \tag{8}$$

Словами высказывание $\neg\ \boldsymbol{P}$ можно прочитать так: «не для любого треугольника ABC величины углов A и B равны». Это означает, что найдется (существует) $\triangle ABC$, в котором не имеет места соотношение $\angle A = \angle B$:

$$\neg\, \boldsymbol{P} = (\exists\ \triangle ABC)\ \neg\ (\angle A = \angle B). \tag{9}$$

Итак, чтобы образовать отрицание высказывания, начинающегося квантором $\forall$, можно либо поставить знак $\neg$ перед высказыванием (см (8)), либо же заменить квантор $\forall$ на $\exists$ и при этом поставить знак $\neg$ перед остальной частью высказывания (см. (9)).

Аналогично можно получить отрицание высказывания, начинающегося знаком $\exists$. Рассмотрим в качестве примера следующее высказывание:

$$\boldsymbol{M} = (\exists\ x \in \boldsymbol{R})\ (x^2 < 0), \tag{10}$$

т. е. существует действительное число x, квадрат которого отрицателен. Его отрицание можно получить, поставив знак $\neg$ перед всем высказыванием:

$$\neg\, \boldsymbol{M} = \neg\ (\exists\ x \in \boldsymbol{R})\ (x^2 < 0). \tag{11}$$

Это читается так: «не существует действительного числа x, для которого x^2 отрицательно». Это означает, что, какое бы число x мы ни взяли, его квадрат не будет отрицательным числом:

$$\neg\, \boldsymbol{M} = (\forall\ x \in \boldsymbol{R})\ \neg\ (x^2 < 0). \tag{12}$$

Итак, чтобы образовать отрицание высказывания, начинающегося знаком $\exists$, можно либо поставить знак $\neg$ перед всем высказыванием (см. (11)), либо же заменить знак $\exists$ на $\forall$ и при этом поставить знак $\neg$ перед остальной частью высказывания (см. (12)).

Если мы условимся знаки $\forall$ и $\exists$ называть *противоположными*, то правило составления отрицания можно сформулировать так:

Чтобы образовать отрицание высказывания, начинающегося одним из знаков $\forall$ *или* $\exists$, *можно либо поставить знак* $\neg$ *перед всем высказыванием, либо* 1) *заменить начальный знак* ($\forall$ *или* $\exists$) *на противоположный и* 2) *поставить знак* $\neg$ *перед остальной частью высказывания.*

Теперь мы можем вернуться к рассмотренному выше софизму. Обозначим через $\boldsymbol{K}$ множество всех критян, а через $\boldsymbol{A}(x)$ обозначим предикат «x — лжец». Тогда высказанное критянином утверждение можно записать так:

$$(\forall\ x \in \boldsymbol{K})\ \boldsymbol{A}(x). \tag{13}$$

Первый вывод, указанный в софизме, совершенно правилен: если все критяне лжецы, то он сам лжец, т. е. он сказал неправду. Значит, написанное высказывание ложно, т. е. истинным является его отрицание: $\neg\ (\forall\ x \in \boldsymbol{K})\ \boldsymbol{A}(x)$. Иначе говоря, *не все* критяне лжецы, т. е. существует (хотя бы один) критянин, не являющийся лжецом:

$$(\exists\ x \in \boldsymbol{K})\ \neg\ \boldsymbol{A}(x).$$

Итак, *существует* критянин, являющийся правдивым человеком. Но это вовсе не означает, что *все* критяне правдивы. Поэтому ниоткуда *не следует*, что тот критянин, который высказал утверждение (13), правдив. Иначе говоря, мы не вправе заключить, что он сказал правду. На этом и обрывается приведенное в софизме рассуждение, т. е. никакого замкнутого круга не получается.

В заключение рассмотрим роль контрпримеров при решении вопроса о том, верно ли некоторое утверждение, содержащее кванторы $\forall$, $\exists$.

Пусть $\boldsymbol{B}(x)$ — некоторый предикат, заданный на множестве $\boldsymbol{M}$. Поставим вопрос: истинно ли высказывание $(\forall\ x \in \boldsymbol{M})\ \boldsymbol{B}(x)$? Если множество $\boldsymbol{M}$ конечно, можно попытаться перебрать все его элементы и убедиться, что для любого элемента $m \in \boldsymbol{M}$ высказывание $\boldsymbol{B}(m)$ истинно. Если же множество $\boldsymbol{M}$ бесконечно (или конечно, но содержит очень много элементов), то такой перебор неосуществим, и потому в справедливости высказывания $(\forall\ x \in \boldsymbol{M})\ \boldsymbol{B}(x)$ можно убедиться только общим рассуждением.

Если разные попытки доказать справедливость высказывания $(\forall\ x \in \boldsymbol{M})\ \boldsymbol{B}(x)$ не приводят к успеху, то возникает предположение, что, быть может, это высказывание неверно, т. е. истинным является его отрицание $\neg\ (\forall\ x \in \boldsymbol{M})\ \boldsymbol{B}(x)$. А как можно убедиться в правильности этого предположения? Для этого заметим, что высказывание $\neg\ (\forall\ x \in \boldsymbol{M})\ \boldsymbol{B}(x)$ означает то же самое, что и $(\exists\ x \in \boldsymbol{M})\ \neg\ \boldsymbol{B}(x)$. А вот для подтверждения этого утверждения достаточно указать *один* элемент $m \in \boldsymbol{M}$, для которого высказывание $\boldsymbol{B}(m)$ ложно, т. е., чтобы убедиться в ошибочности общего высказывания $(\forall\ x \in \boldsymbol{M})\ \boldsymbol{B}(x)$, достаточно привести один контрпример.

Рассмотрим еще один пример. Хорошо известна школьная теорема о том, что диагонали ромба перпендикулярны. Справедлива ли обратная теорема: «если диагонали четырехугольника F перпендикулярны, то он является ромбом»? Это высказывание можно записать в виде

$$(\forall\ F)\ (\boldsymbol{A}(F) \Rightarrow \boldsymbol{B}(F)), \tag{14}$$

где $\boldsymbol{A}(F)$ и $\boldsymbol{B}(F)$ — следующие предикаты:

$\boldsymbol{A}(F)$ = {диагонали четырехугольника F перпендикулярны};

$\boldsymbol{B}(F)$ = {четырехугольник F — ромб}.

Чтобы убедиться в ложности высказывания (14), нужно доказать, что истинно его отрицание $\neg\ (\forall\ F)\ (\boldsymbol{A}(F) \Rightarrow \boldsymbol{B}(F))$, т. е. что не для каждого четырехугольника F из справедливости $\boldsymbol{A}(F)$ вытекает справедливость $\boldsymbol{B}(F)$. Иначе говоря, нужно убедиться, что существует четырехугольник, у которою диагонали перпендикулярны, но который вместе с тем ромбом не является, т. е. для которого $\boldsymbol{A}(F)$ истинно и одновременно $\boldsymbol{B}(F)$ ложно. Таким образом, высказывания $\neg\ (\forall\ F)\ (\boldsymbol{A}(F) \Rightarrow \boldsymbol{B}(F))$ и $(\exists\ F)\ (\boldsymbol{A}(F)$ и $\neg\ \boldsymbol{B}(F))$ означают одно и то же (или, как еще говорят, *равносильны*).

Чтобы убедиться в ложности высказывания

$$(\forall\ F)\ (\boldsymbol{A}(F) \Rightarrow \boldsymbol{B}(F)), \tag{15}$$

достаточно проверить истинность высказывания

$$(\exists\ F)\ (\boldsymbol{A}(F)\ \text{ и }\ \neg\ \boldsymbol{B}(F)).$$

Иными словами, чтобы убедиться в ложности высказывания вида (15), достаточно, как и всегда, привести один контрпример, причем в этом случае контрпримером будет элемент F, для которого одновременно истинны высказывания $\boldsymbol{A}(F)$ и $\neg\ \boldsymbol{B}(F)$. В случае рассмотренного выше предложения (14) контрпримером должен быть четырехугольник F, для которого $\boldsymbol{A}(F)$ истинно (диагонали перпендикулярны), а $\boldsymbol{B}(F)$ ложно (F не является ромбом). Такой контрпример есть, он приведен на рис. 342. Этот контрпример показывает, что теорема «если диагонали четырехугольника перпендикулярны, то он является ромбом» неверна.

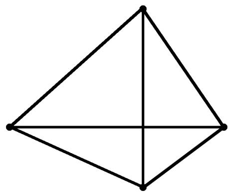


Рис. 342

Задачи и упражнения

221. Постройте отрицание высказывания: «Нет ничего лучше плохой погоды» (название одного из болгарских кинофильмов).

222. Постройте отрицание высказывания: «Каждому овощу свое время».

223. Постройте отрицание высказывания: «Среди ста двузначных натуральных чисел найдутся два равных». Какое из этих двух высказываний истинно?

224. Запишите теорему, обратную теореме Пифагора.

225. Запишите теорему, обратную теореме Больцано.

55. Необходимое и достаточное условие

В предыдущем пункте мы рассмотрели пример *обратной теоремы*. Вообще для каждой теоремы, записанной в форме импликации

$$(\forall\ x \in \boldsymbol{M})\ (\boldsymbol{P}(x) \Rightarrow \boldsymbol{Q}(x)), \tag{16}$$

можно составить обратную ей теорему, меняя местами условие и заключение (при сохранении той же разъяснительной части):

$$(\forall\ x \in \boldsymbol{M})\ (\boldsymbol{Q}(x) \Rightarrow \boldsymbol{P}(x)).$$

В качестве еще одного примера возьмем следующую теорему о натуральных числах:

$$(\forall\ n \in \boldsymbol{N})\ (n \text{ делится на } 6) \Rightarrow (\text{последняя цифра числа } n \text{ — четная}).$$

Эта теорема справедлива, но обратная ей теорема

$$(\forall\ n \in \boldsymbol{N})\ (\text{последняя цифра числа } n \text{ четная}) \Rightarrow (n \text{ делится на } 6)$$

неверна.

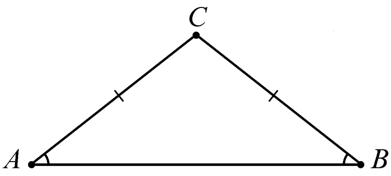


Рис. 343

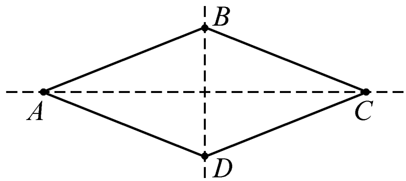


Рис. 344

Таким образом, доказав некоторую теорему, мы еще не можем утверждать, что справедлива обратная теорема. Справедливость обратной теоремы требует отдельного доказательства.

Рассмотрим теперь теоремы, для которых обратные теоремы также справедливы. Вот пример:

$(\forall\ n \in \boldsymbol{N})$ (n делится на 9) $\Rightarrow$ (сумма цифр числа n делится на 9).

Обратной является теорема

$(\forall\ n \in \boldsymbol{N})$ (сумма цифр числа n делится на 9) $\Rightarrow$ (n делится на 9).

В этом случае и исходная теорема, и обратная ей — обе справедливы. Приведем геометрический пример. Следующая теорема выражает одно из свойств равнобедренного треугольника (рис. 343):

$$(\forall\ \triangle ABC)\ (BC = AC) \Rightarrow (\angle A = \angle B). \tag{17}$$

Обратная ей теорема также справедлива:

$$(\forall\ \triangle ABC)\ (\angle A = \angle B) \Rightarrow (BC = AC), \tag{18}$$

т. е. если в треугольнике величины углов при основании равны, то этот треугольник — равнобедренный.

В математике принято формулировать теоремы, имеющие форму импликации, с помощью слов *необходимо*, *достаточно*. Рассмотрим, например, снова теорему о перпендикулярности диагоналей ромба (рис. 344):

$$(\forall\ ABCD)\ (ABCD \text{ — ромб}) \Rightarrow (AC \perp BD). \tag{19}$$

Эту теорему можно разъяснить так: достаточно знать, что четырехугольник $ABCD$ является ромбом, и тогда мы можем утверждать, что его диагонали перпендикулярны. Или иначе: для перпендикулярности диагоналей четырехугольника достаточно, чтобы он был ромбом.

Вообще, если сказано, что выполнение некоторого утверждения $\boldsymbol{P}$ является *достаточным* для $\boldsymbol{Q}$, то это означает, что утверждается справедливость теоремы, в которой $\boldsymbol{P}$ — *условие*, а $\boldsymbol{Q}$ — *заключение* (см. (16)).

Теорему (19) можно разъяснить еще и так: если $ABCD$ — ромб, то непременно его диагонали перпендикулярны. В математике принято эту формулировку выражать по-другому, используя слово «необходимо»: для того чтобы четырехугольник был ромбом, необходимо, чтобы его диагонали были перпендикулярны.

Вообще, если сказать, что выполнение некоторого утверждения $\boldsymbol{Q}$ является *необходимым* для $\boldsymbol{P}$, то это означает, что утверждается справедливость теоремы, в которой $\boldsymbol{P}$ — *условие*, а $\boldsymbol{Q}$ — *заключение* (см. (16)).

Если для некоторой теоремы справедлива также и обратная ей теорема, то ее формулировку можно выразить по-другому, используя слова *необходимо и достаточно*. Например, теорема (17) и обратная ей теорема (18) обе справедливы. Иными словами, для того чтобы треугольник был равнобедренным, необходимо, чтобы у него величины углов при основании были равны (см. (17)); в то же время для того, чтобы треугольник был равнобедренным, достаточно, чтобы у него величины углов при основании были равны (см. (18)). Это кратко записывается в виде

$$(\forall\ \triangle ABC)\ (BC = AC) \Leftrightarrow (\angle A = \angle B)$$

и читается словами так: для того, чтобы треугольник был равнобедренным, необходимо и достаточно, чтобы у него величины углов при основании были равны.

А рассмотренный выше признак делимости на 9 можно выразить так: для того, чтобы натуральное число делилось на 9, необходимо и достаточно, чтобы сумма его цифр делилась на 9:

$$(\forall\ n \in \boldsymbol{N})\ (n \text{ делится на } 9) \Leftrightarrow (\text{сумма цифр числа } n \text{ делится на } 9). \quad (20)$$

Заметим еще, что слова «необходимое условие» часто заменяются словами: «только в том случае», «только тогда». Например, теорему (19) можно сформулировать еще и так:

четырехугольник только в том случае может быть ромбом, если его диагонали перпендикулярны.

Слова «необходимо и достаточно» нередко заменяются словами «тогда и только тогда, когда» или «в том и только в том случае, если». Например, четырехугольник в том и только в том случае является параллелограммом, если его диагонали, пересекаясь, делятся пополам.

Вместо того, чтобы сказать «достаточное условие», «необходимое условие», иногда говорят «достаточный признак», «необходимый признак». Иногда даже говорят просто *признак*, считая ясным, о каком из признаков (достаточном или необходимом) идет речь. Например, теорема (20) называется *признаком делимости* на 9; этот признак является необходимым и достаточным для делимости числа на 9.

В школьном обиходе термин «признак» также используется. Например, мы говорим о *признаке параллельности двух прямых*. Под этим понимается следующая теорема: две прямые в том и только в том случае параллельны, если при пересечении с третьей прямой

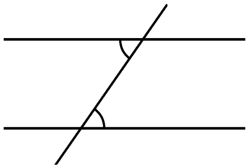


Рис. 345

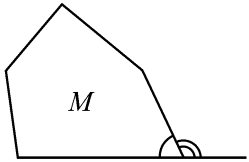


Рис. 346

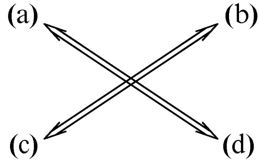


Рис. 347

(«секущей») они образуют конгруэнтные накрест лежащие углы (рис. 345). Таким образом, признак параллельности прямых является *необходимым и достаточным* условием.

В заключение рассмотрим *метод доказательства от противного* и связанное с ним понятие *противоположной теоремы*.

Рассмотрим в качестве примера следующую теорему: выпуклый многоугольник не может иметь более трех острых углов.

Для доказательства предположим, что это утверждение неверно, т. е., что верно его отрицание: существует выпуклый многоугольник M, имеющий не менее четырех острых углов.

Рассмотрим внешние углы выпуклого многоугольника M. Так как угол, смежный с острым, является тупым (рис. 346), то у многоугольника M не менее четырех тупых внешних углов. Значит, сумма величин всех внешних углов выпуклого многоугольника M больше 360°. Но это противоречит теореме о том, что сумма внешних углов выпуклого многоугольника всегда равна 360°. Полученное противоречие доказывает теорему.

Приведенное рассуждение было проведено методом доказательства от противного: мы предположили, что доказываемое утверждение неверно (т. е. верно его отрицание), и показали, что это приводит к противоречию. Тем самым была установлена ложность отрицания и, следовательно, истинность доказываемого утверждения.

Обратимся теперь к понятию противоположной теоремы. Из каждой теоремы, имеющей форму импликации

$$(\forall\ x \in \boldsymbol{M})\ (\boldsymbol{P}(x) \Rightarrow \boldsymbol{Q}(x)), \tag{21a}$$

(см. (16)), можно получить еще три теоремы, показанные на следующей схеме:

$$(\forall\ x \in \boldsymbol{M})\ (\boldsymbol{Q}(x) \Rightarrow \boldsymbol{P}(x)), \tag{21b}$$
$$(\forall\ x \in \boldsymbol{M})\ (\neg\ \boldsymbol{P}(x) \Rightarrow \neg\ \boldsymbol{Q}(x)), \tag{21c}$$
$$(\forall\ x \in \boldsymbol{M})\ (\neg\ \boldsymbol{Q}(x) \Rightarrow \neg\ \boldsymbol{P}(x)). \tag{21d}$$

По отношению к исходной теореме (21a) теорема (21b) является обратной. Теорема (21c) получается из исходной теоремы (21a) заменой условия и заключения их отрицаниями, она называется *противоположной* исходной теореме (21a). Теорема (21d) — противоположная обратной.

Проиллюстрируем схему (21) на примере теоремы о перпендикулярности диагоналей ромба. Для этого обозначим через $\boldsymbol{M}$ множество всех четырехугольников и рассмотрим следующие предикаты:

$\boldsymbol{P}(x) = \{x$ — ромб$\}$;

$\boldsymbol{Q}(x) = \{$диагонали четырехугольника x перпендикулярны$\}$.

Тогда четыре теоремы, показанные на схеме (21) означают следующие утверждения:

(21a): если четырехугольник является ромбом, то его диагонали перпендикулярны;

(21b): если диагонали четырехугольника перпендикулярны, то он является ромбом;

(21c): если четырехугольник не является ромбом, то его диагонали не перпендикулярны;

(21d): если диагонали четырехугольника не перпендикулярны, то он не является ромбом.

Обратная (21b) и противоположная (21c) теоремы здесь неверны, это показывает контрпример, приведенный на рис. 342.

Теорема (21d), противоположная обратной, так же как и исходная теорема (21a), обе справедливы.

Вообще, теорема, противоположная обратной, *равносильна* исходной теореме, т. е. если истинна одна теорема, то истинна и другая. Поэтому, если доказательство исходной теоремы вызывает трудности, можно вместо нее доказать теорему, противоположную обратной. В этом и заключается метод доказательства от противного.

Итак, подведем итоги. По отношению к теореме (21a) теорема (21b) является обратной, теорема (21c) — противоположной, теорема (21d) — противоположной обратной. Как мы уже знаем, теоремы (21a) и (21d) равносильны.

Если же за исходную мы примем теорему (21b), то теорема (21a) будет обратной. Поэтому теоремы (21b) и (21c) также равносильны (рис. 347).

Для того, чтобы установить, что $\boldsymbol{A}(x)$ является необходимым и достаточным условием для $\boldsymbol{B}(x)$, нужно доказать, во-первых, одну из теорем (21a) или (21d) и, во-вторых, одну из теорем (21b) или (21c).

Заметим еще, что при формулировке обратной теоремы (а также теорем, противоположных к исходной и обратной) весьма существенную роль играет разъяснительная часть теоремы. Рассмотрим в качестве примера все ту же теорему о перпендикулярности диагоналей ромба. Обозначим через $\boldsymbol{M}$ множество всех четырехугольников, а через $\boldsymbol{\Pi}$ — множество всех параллелограммов, и рассмотрим следующие две теоремы:

$$(\forall\ x \in \boldsymbol{M})\ (\boldsymbol{P}(x) \Rightarrow \boldsymbol{Q}(x)), \tag{22}$$

$$(\forall\ x \in \boldsymbol{\Pi})\ (\boldsymbol{P}(x) \Rightarrow \boldsymbol{Q}(x)). \tag{23}$$

Казалось бы, они различаются лишь по форме, но по существу выражают одно и то же: если четырехугольник (или параллелограмм) является ромбом, то его диагонали перпендикулярны. Кратко эту теорему выражают так (пренебрегая разъяснительной частью): *диагонали ромба перпендикулярны*.

Однако при формулировании обратной и противоположной теорем именно *форма* записи оказывается весьма существенной. По отношению к (22) обратной является теорема $(\forall\ x \in \boldsymbol{M})\ (\boldsymbol{Q}(x) \Rightarrow \boldsymbol{P}(x))$, т. е. если диагонали *четырехугольника* перпендикулярны, то он является ромбом. Эта теорема неверна (рис. 342). По отношению к (23) обратной является теорема $(\forall\ x \in \boldsymbol{\Pi})\ (\boldsymbol{Q}(x) \Rightarrow \boldsymbol{P}(x))$, т. е. если диагонали *параллелограмма* перпендикулярны, то он является ромбом. Эта теорема верна, в отличие от предыдущей!

Иначе говоря, теорема (23) содержит фактически необходимое и достаточное условие: $(\forall\ x \in \boldsymbol{\Pi})\ (\boldsymbol{Q}(x) \Leftrightarrow \boldsymbol{P}(x))$, т. е. *параллелограмм* в том и только в том случае является ромбом, если его диагонали перпендикулярны.

Задачи и упражнения

226. Является ли теорема, обратная теореме Пифагора, необходимым и достаточным условием прямоугольности треугольника?

227. Для теоремы: «Диагонали ромба перпендикулярны» составьте обратную, противоположную и обратную к противоположной теоремы. Какие из них истинны?

228. Для теоремы: «Диагонали прямоугольника равны» составьте обратную, противоположную и обратную к противоположной теоремы. Какие из них истинны?

229. Утверждение: «Если целое число оканчивается на 0, то оно делится на 5». Является оно необходимым, достаточным или необходимым и достаточным для делимости числа на 5? Составьте к этому утверждению обратное, противоположное и обратное к противоположному. Какие из них истинны?

230. Верно ли утверждение: «Два треугольника конгруэнтны тогда и только тогда, когда две стороны и заключенный между ними угол одного конгруэнтны двум сторонам и соответствующему углу другого»?

56. Конъюнкция и дизъюнкция

Нередко приходится рассматривать случаи, когда одновременно выполняются два каких-либо свойства. Возьмем, например, высказывание: существуют равнобедренные прямоугольные треугольники (рис. 348). Это высказывание означает, что существует $\triangle ABC$, который одновременно обладает двумя свойствами: угол C прямой и $AC = BC$.

В математике одновременное выполнение двух свойств принято называть *конъюнкцией* этих свойств и обозначать знаком $\wedge$ (читается «и»). Таким образом, сформулированное выше высказывание можно записать так: $(\exists\ \triangle ABC)\ ((\angle C = 90°) \wedge (AC = BC))$.

Еще одним примером использования конъюнкции служит определение пересечения множеств. В самом деле, элемент принадлежит пересечению множеств $\boldsymbol{M}$ и $\boldsymbol{P}$, если он, во-первых, принадлежит множеству $\boldsymbol{M}$ и, во-вторых, принадлежит множеству $\boldsymbol{P}$:

$$\boldsymbol{M} \cap \boldsymbol{P} = \{x: (x \in \boldsymbol{M}) \wedge (x \in \boldsymbol{P})\}. \tag{24}$$

Здесь фигурные скобки означают слово «множество», а двоеточие можно заменить словосочетанием «для которых» (т. е. двоеточие *поясняет*, какие x следует взять). Словами запись (24) можно прочитать так: пересечение $\boldsymbol{M} \cap \boldsymbol{P}$ есть множество тех элементов x, которые принадлежат множеству $\boldsymbol{M}$ и множеству $\boldsymbol{P}$ (рис. 349). Союз «и» в записи (24) выражен знаком конъюнкции $\wedge$.

Из сказанного ясно, что операция конъюнкции определяется таблицей истинности

A	B	$A \wedge B$
И	И	И
И	Л	Л
Л	И	Л
Л	Л	Л

т. е. высказывание $\boldsymbol{A} \wedge \boldsymbol{B}$ истинно *только* в том случае, если истинны *оба* высказывания $\boldsymbol{A}$, $\boldsymbol{B}$.

Кроме знака конъюнкции $\wedge$ в математике часто используется знак *дизъюнкции* $\vee$ (читается «или»). Знак $\vee$ не совсем точно передается союзом «или», которым мы пользуемся в обычной речи. Возьмем, например, фразу «Я уезжаю, но за книгой зайдет моя сестра или брат». Здесь, скорее всего, союз «или» понимается в смысле «либо-либо», т. е. зайдет *либо* сестра, *либо* брат; возможность, что они зайдут вместе, не предусматривается. Иначе говоря, в обычной речи союз «или» чаще всего имеет *разделительный* оттенок. В математике знак дизъюнкции $\vee$ не имеет разделительного смысла, т. е. если два свойства соединены знаком $\vee$, то это значит, что выполняется *хотя бы одно* из этих свойств, т. е. либо имеет место первое свойство (но не второе), либо имеет место второе свойство (но не первое), либо же (и в этом отличие) имеют место *оба* свойства одновременно.

Рассмотрим в качестве примера объединение множеств $\boldsymbol{M}$ и $\boldsymbol{P}$. Имеет место один из трех случаев (см. рис. 350):

1) либо $x \in \boldsymbol{M}$, но $x \notin \boldsymbol{P}$ (точка x_1);

2) либо $x \in \boldsymbol{P}$, но $x \notin \boldsymbol{M}$ (точка x_2);

3) либо $x \in \boldsymbol{M}$ и $x \in \boldsymbol{P}$ (точка x_3).

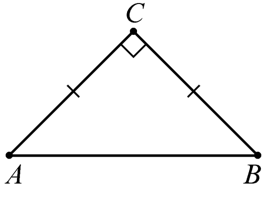


Рис. 348

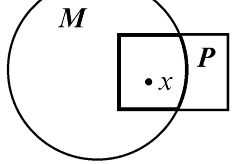


Рис. 349

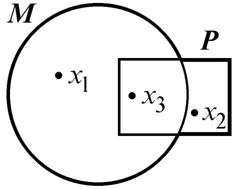


Рис. 350

Запись $(x \in \boldsymbol{M}) \vee (x \in \boldsymbol{P})$ как раз и означает, что имеет место какой-либо из этих трех случаев.

Из сказанного ясно, что операция дизъюнкции определяется таблицей истинности

A	B	$A \vee B$
И	И	И
И	Л	И
Л	И	И
Л	Л	Л

т. е. высказывание $\boldsymbol{A} \vee \boldsymbol{B}$ ложно *только* в случае, если ложны оба высказывания $\boldsymbol{A}$, $\boldsymbol{B}$. Поэтому $\boldsymbol{M} \cup \boldsymbol{P} = \{x\colon (x \in \boldsymbol{M}) \vee (x \in \boldsymbol{P})\}$.

Нередко конъюнкция и дизъюнкция применяются при формулировке теорем. Вспомним, например, теорему о том, что катет, лежащий против угла 30°, равен половине гипотенузы (рис. 351). Более детально, если треугольник является прямоугольным и один из его углов содержит 30°, то длина катета, лежащего против этого угла, равна половине длины гипотенузы:

$$(\forall\ \triangle ABC)\ (\angle C = 90°) \wedge (\angle A = 30°) \Rightarrow (BC = \frac{1}{2}AB).$$

Здесь в формулировке теоремы использована конъюнкция.

Вот еще пример: если произведение двух чисел равно нулю, то хотя бы один из сомножителей равен нулю, т. е.

$$(\forall\ a, b)\ (ab = 0) \Rightarrow ((a = 0) \vee (b = 0)).$$

Здесь в заключении теоремы используется дизъюнкция. В тех случаях, когда нужно образовать *отрицание* утверждения, содержащего конъюнкцию или дизъюнкцию, можно воспользоваться следующими эквивалентностями:

$$\neg\ (\boldsymbol{P} \wedge \boldsymbol{Q}) \Leftrightarrow (\neg\ \boldsymbol{P}) \vee (\neg\ \boldsymbol{Q}), \tag{25}$$

$$\neg\ (\boldsymbol{P} \vee \boldsymbol{Q}) \Leftrightarrow (\neg\ \boldsymbol{P}) \wedge (\neg\ \boldsymbol{Q}), \tag{26}$$

которые справедливы для любых высказываний (или предикатов) $\boldsymbol{P}$, $\boldsymbol{Q}$. Соотношение (25) можно пояснить следующим образом: если неверно, что справедливы оба утверждения $\boldsymbol{P}$ и $\boldsymbol{Q}$, то это означает, что неверно $\boldsymbol{P}$ или неверно $\boldsymbol{Q}$.

Например, четырехугольник $ABCD$ называется параллелограммом, если каждые две его противоположные стороны параллельны (рис. 352), т. е. $(AB \parallel CD) \wedge (AD \parallel BC)$. Если же $ABCD$ не является параллелограммом, т. е. $\neg\ ((AB \parallel CD) \wedge (AD \parallel BC))$, то это означает, согласно (25), что $(\neg\ (AB \parallel CD)) \vee (\neg\ (AD \parallel BC))$, т. е. *хотя бы одно* из условий $AB \parallel CD$, $AD \parallel BC$ не выполняется. Иными словами, если $ABCD$ — не параллелограмм, то либо не выполняется условие $AB \parallel CD$ (рис. 353), либо не выполняется условие $AD \parallel BC$ (рис. 354), либо же оба эти условия не выполнены (рис. 355).

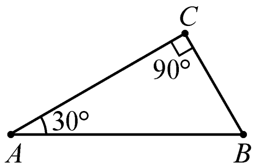


Рис. 351

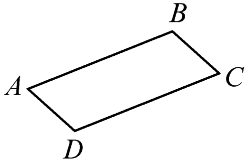


Рис. 352

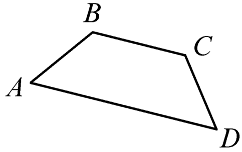


Рис. 353

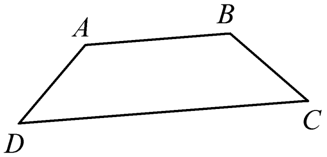


Рис. 354

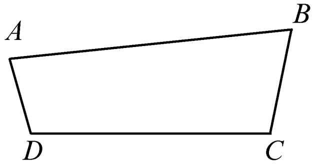


Рис. 355

Рассмотрим еще один иллюстрирующий пример. Будем считать, что вылет самолета состоится в том и только в том случае, если, во-первых, есть горючее и, во-вторых, погода летная. Иначе говоря, необходимое и достаточное условие вылета имеет вид

$$(\text{есть горючее}) \wedge (\text{погода летная}).$$

Вылет не состоится, если это условие не выполнено, т. е.

$$\neg\ ((\text{есть горючее}) \wedge (\text{погода летная})).$$

Но, согласно (25), это эквивалентно тому, что

$$(\neg\ (\text{есть горючее})) \vee (\neg\ (\text{погода летная})).$$

Иными словами, вылет не состоится, если нет горючего *или* погода нелетная.

Разумеется, рассмотрение примеров, подтверждающих разумность соотношений (25), (26), не является их доказательством. Установить правильность этих соотношений нетрудно с помощью таблиц истинности. Следующая таблица показывает истинность высказываний $\neg\,(\boldsymbol{P} \wedge \boldsymbol{Q})$ и $(\neg\,\boldsymbol{P}) \vee (\neg\,\boldsymbol{Q})$ в зависимости от истинности высказываний $\boldsymbol{P}$, $\boldsymbol{Q}$:

$\boldsymbol{P}$	$\boldsymbol{Q}$	$\boldsymbol{P} \wedge \boldsymbol{Q}$	$\neg\,(\boldsymbol{P} \wedge \boldsymbol{Q})$	$\neg\,\boldsymbol{P}$	$\neg\,\boldsymbol{Q}$	$(\neg\,\boldsymbol{P}) \vee (\neg\,\boldsymbol{Q})$
И	**И**	**И**	**Л**	**Л**	**Л**	**Л**
И	**Л**	**Л**	**И**	**Л**	**И**	**И**
Л	**И**	**Л**	**И**	**И**	**Л**	**И**
Л	**Л**	**Л**	**И**	**И**	**И**	**И**

Мы видим, что соответствующие столбцы в этих таблицах истинности *одинаковы*, и это показывает, что оба высказывания $\neg\,(\boldsymbol{P} \wedge \boldsymbol{Q})$ и $(\neg\,\boldsymbol{P}) \vee (\neg\,\boldsymbol{Q})$ *эквивалентны*, т. е. означают одно и то же. Тем самым справедливость соотношения (25) подтверждена. Аналогично проверяется справедливость соотношения (26).

Выше мы рассматривали в качестве примера определение параллелограмма. Знаками его можно записать следующим образом:

$$(\forall\ ABCD)\,(ABCD \text{ — параллелограмм}) \underset{def}{\equiv} ((AB \parallel CD) \wedge (AD \parallel BC)).$$

Здесь запись $(\forall\ ABCD)$ (похожая на разъяснительную часть теоремы) показывает, что рассматривается произвольный четырехугольник $ABCD$. В связи с этим говорят, что «четырехугольник» является *родовым понятием* (при определении параллелограмма), т. е. рассматривается множество всех четырехугольников и среди них будут выделяться те, которые образуют новое понятие (параллелограмм). Далее, запись ($ABCD$ — параллелограмм) представляет собой введение нового *термина*, т. е. мы указываем, как будут называться те четырехугольники, которые мы выделяем. Запись $\underset{def}{\equiv}$ означает, что мы имеем дело с *определением* нового понятия (от английского слова *definition*). Наконец, запись $((AB \parallel CD) \wedge (AD \parallel BC))$ указывает *видовые отличия*, т. е. те отличительные свойства, которыми должен обладать четырехугольник, чтобы мы считали его параллелограммом.

Словесное определение содержит фактически все эти же части, но не выделенные явно:

четырех-угольник	называется	параллело-граммом,	если его противоположные стороны попарно параллельны
↑	↑	↑	↑
(*родовое понятие***)**	**(***def***)**	**(***термин***)**	**(***видовые отличия***)**

Для сравнения приведем определение трапеции: оно отличается от определения параллелограмма тем, что при формулировке видовых отличий конъюнкция заменяется дизъюнкцией:

$$(\forall\ ABCD)\,(ABCD \text{ — трапеция}) \underset{def}{\equiv} ((AB \parallel CD) \vee (AD \parallel BC)). \qquad (27)$$

Иными словами, четырехугольник является трапецией, если *хотя бы одна* пара его противоположных сторон является двумя параллельными отрезками. Заметим, что если *обе* пары противоположных сторон параллельны, то четырехугольник также является трапецией (в силу определения дизъюнкции), т. е. *параллелограмм — частный случай трапеции*. Отметим еще, что многие авторы заменяют это «дизъюнктивное» определение другим: четырехугольник называется трапецией, если *одна* пара его противоположных сторон — параллельные отрезки, а другая — нет. Но мы придерживаемся определения (27).

Как мы видим, структура определения очень похожа на структуру теоремы, но части, имеющиеся в записи определения, соединены другим символом (и называются по-другому). При этом теорему надлежит *доказывать*, тогда как определение в этом не нуждается.

Заметим еще, что при записи определений рекомендуется использовать «ближайшее» родовое понятие. Например, определение прямоугольника можно записать в следующих двух формах:

$$(\forall\ ABCD)\ (ABCD \text{ — прямоугольник}) \underset{def}{\equiv} (\angle A = \angle B = \angle C = 90°);$$

$$(\forall \text{ пар. } ABCD)\ (ABCD \text{ — прямоугольник}) \underset{def}{\equiv} (\angle A = 90°).$$

Первое из этих определений говорит, что *произвольный четырехугольник* называется прямоугольником, если у него есть три прямых угла (четвертый угол тогда тоже прямой, так как сумма углов четырехугольника равна 360°). Второе определение выделяет прямоугольники из *параллелограммов*, а не из всех четырехугольников: параллелограмм называется прямоугольником, если у него есть один прямой угол (тогда и все углы прямые — в силу свойств параллелограммов). Второе определение использует более близкое к прямоугольнику родовое понятие «параллелограмм», чем первое определение.

Рассмотрим следующий пример. Учащемуся предъявляются несколько различных фигур: круг, шестиугольник, треугольник, многоугольники, указанные на рис. 352 – 355. Затем предлагается вопрос: какие из этих фигур являются трапециями? Тот, кто уже овладел понятием трапеции, считает, что он «сразу» видит, какие из представленных ему фигур являются трапециями. Однако в действительности в его сознании идет работа, нередко скрытая от него самого (психологи называют такую работу «свернутой», или «автоматизированной»). Если осуществить рефлексию, т. е. «развернуть» эту деятельность с помощью психологического самоанализа, то эта деятельность в ее полном виде будет выглядеть следующим образом. Вначале проверяется принадлежность к родовому понятию, т. е. отбрасывается круг (это — не многоугольник), а также шестиугольник и треугольник (это — не четырехугольники). Остаются фигуры, изображенные на рис. 352 – 355. Теперь надо проверить наличие видовых отличий. У фигуры на рис. 355 *ни одна* пара противоположных сторон не обладает свойством параллельности, значит, это — не трапеция. Остаются фигуры на рис. 352 – 354. Они обладают видовыми отличиями, т. е. являются трапециями.

Некоторое затруднение может вызвать рис. 352, на котором не одна, а *обе* пары противоположных сторон являются параллельными. Осознание того, что параллелограмм — частный случай трапеции, весьма важно и связано с пониманием смысла дизъюнкции.

Описанная деятельность вполне соответствует установлению принадлежности к понятию «трапеция» с помощью определения (27), *адекватна* этому определению. Аналогичная деятельность адекватна не только определению трапеции, но и многим другим математическим определениям, а именно тем, в которых видовое отличие содержит два условия, соединенные дизъюнкцией.

Организация адекватной деятельности предусматривает исследование многих, разнообразных объектов, для которых выполняется или не выполняется принадлежность к родовому понятию, выполняются или не выполняются отдельные видовые отличия (дизъюнктивно или конъюнктивно связанные) и т. д. Выполняя такую деятельность, учащийся охватит все части определения, т. е. *усвоит* изучаемое понятие. Именно после полного усвоения у него будет создано впечатление, что он «сразу» видит, принадлежит или не принадлежит предъявляемый объект к изучаемому понятию, хотя в действительности он проверяет *все* факторы, имеющиеся в определении, но делает это свернуто, незаметно для него самого.

Задачи и упражнения

231. Что можно сказать об истинности высказывания $\neg(\boldsymbol{A} \wedge \boldsymbol{B} \wedge \boldsymbol{C}) \vee \boldsymbol{A}$, если $\boldsymbol{A}$ — истинно, а $\boldsymbol{B}$ и $\boldsymbol{C}$ — ложны?

232. Пусть $\boldsymbol{A}$ — ложное высказывание, а $\boldsymbol{B}$ и $\boldsymbol{C}$ — истинные. Будет ли истинным высказывание $(\boldsymbol{A} \wedge \boldsymbol{B}) \vee (\boldsymbol{B} \wedge \boldsymbol{C}) \vee (\boldsymbol{C} \wedge \boldsymbol{A})$? $(\boldsymbol{B} \wedge \boldsymbol{A}) \vee (\boldsymbol{C} \wedge \boldsymbol{A})$?

233. В каких случаях высказывания $\boldsymbol{A} \wedge (\boldsymbol{B} \vee \boldsymbol{C})$ и $(\boldsymbol{A} \wedge \boldsymbol{C}) \vee (\boldsymbol{A} \wedge \boldsymbol{B})$ одновременно истинны или ложны?

234. Пусть $\boldsymbol{A}$ — высказывание: «число делится на 3», $\boldsymbol{B}$ — «число делится на 2», $\boldsymbol{C}$ — «число делится на 6». Какие из следующих высказываний истинны: а) $\boldsymbol{C} \Rightarrow \boldsymbol{A} \vee \boldsymbol{B}$; б) $\boldsymbol{A} \wedge \boldsymbol{B} \Rightarrow \boldsymbol{C}$; в) $\neg\,\boldsymbol{C} \wedge \boldsymbol{B} \Rightarrow \neg\,\boldsymbol{A}$?

235. Покажите, что высказывания $\boldsymbol{A}$ и $(\boldsymbol{A} \wedge \boldsymbol{B}) \vee (\boldsymbol{A} \wedge \neg\,\boldsymbol{B})$ всегда одновременно истинны или ложны.

Беседа 12. Понятие об аксиоматическом методе

57. Возникновение аксиоматического метода в математике

Начальные геометрические знания были получены опытным путем. Первые сведения о геометрических фактах дошли до нас из глубокой древности. Например, формулы вычисления площадей земельных участков, имеющих форму прямоугольников, треугольников, трапеций, приведены в древнеегипетских математических папирусах и в клинописных таблицах древнего Вавилона.

Опытное происхождение геометрических фактов подтверждается тем, что в древних источниках упоминаются как верные, так и неточные сведения. Например, древние египтяне вычисляли площадь произвольного четырехугольника с последовательными сторонами a, b, c, d (рис. 356) по формуле $S = \frac{a+c}{2} \cdot \frac{b+d}{2}$, что дает неверный результат даже для параллелограммов, отличных от прямоугольников. По-видимому, на практике встречались лишь четырехугольники, имеющие близкую к прямоугольнику форму, а для них вносимая этой формулой погрешность невелика.

Получение новых геометрических фактов при помощи *рассуждений* (доказательств) относится к VI веку до нашей эры и связано с именем древнегреческого математика Фалеса (его называют также

Фалесом Милетским, поскольку он был родом из Милета). Собственно говоря, Фалес был не профессиональным математиком, а купцом. Плавая по Великому (Средиземному) морю на своих кораблях и занимаясь торговлей, он посвящал свободное время математике. И вот, этот купец сделал гениальное открытие: он обнаружил, что геометрические истины можно добывать не только опытным путем, с помощью измерений, но также чисто умозрительно. Ему приписывают установление свойств равнобедренного треугольника, доказательство равенства величин вертикальных углов (рис. 357), доказательство того, что вписанный угол, опирающийся на диаметр, — прямой (рис. 358), и другие начальные факты геометрии. Фалес, по-видимому, применял поворот части фигуры и перегибание чертежа, т. е. то, что в наши дни называют *движениями*.

Постепенно доказательства приобретают в геометрии все большее значение. К III веку до нашей эры геометрия становится *дедуктивной наукой*, т. е. наукой, в которой подавляющее большинство фактов устанавливается путем вывода, доказательства. К этому времени относится книга «Начала», написанная древнегреческим ученым *Евклидом*. В своей книге Евклид систематизировал известные к этому времени геометрические сведения. В ней доказываются многие теоремы, которые мы изучаем и сейчас, например: свойства параллелограммов и трапеций, теорема Пифагора, подобие многоугольников. Изложение материала в книге Евклида было настолько продуманным, что в течение двух тысячелетий она служила основным учебным пособием при изучении геометрии.

В IV веке до н. э. крупнейший мыслитель древности Аристотель создал основы логики. Аристотель учил, что изложение всякой теории должно начинаться с первоначальных предложений (аксиом), из которых путем рассуждений выводятся дальнейшие факты. Греческое слово «аксиома» означает «удостоенное, принятое положение». Весь набор (система) аксиом называется *аксиоматикой*. Таким образом, аксиомы — это первоначальные факты науки (геометрии), которые принимаются без доказательства и позволяют вывести из них все дальнейшие факты этой науки. Утверждения, выводимые из аксиом, называют теоремами.

При изложении геометрии Евклид полностью следует идеям Аристотеля. Точка зрения Евклида была примерно следующей. Взяв какую-либо теорему, можно проследить по ее доказательству, какие

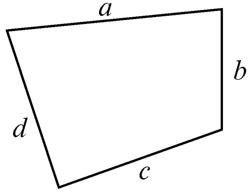


Рис. 356

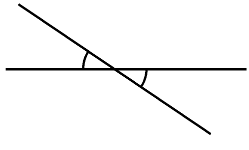


Рис. 357

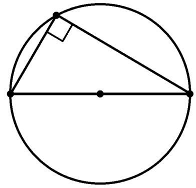


Рис. 358

ранее доказанные теоремы были использованы при ее выводе. Для этих ранее доказанных теорем в свою очередь можно указать те более простые факты, из которых они выводятся, и т. д. В конце концов получается некоторый список простых фактов (аксиом), которые, во-первых, позволяют, идя обратным путем, доказать, исходя из них, все изучаемые теоремы геометрии, а, во-вторых, они настолько просты, что не возникает вопроса о необходимости их «вывода».

Евклид выделил полтора десятка аксиом геометрии, на которых основывалось все его изложение. Среди сформулированных им аксиом имеются, например, следующие.

«Через две точки можно провести прямую».

«Совмещающиеся между собой равны».

«Порознь равные третьему, равны между собой».

Формулировка первой из этих аксиом понятна. Вторая из указанных аксиом говорит о том, что две фигуры, совмещающиеся между собой (подразумевается, при помощи движения), конгруэнтны. (Евклид не пользовался термином «конгруэнтность», а говорил о «равных» фигурах.) Третья из указанных аксиом имеет общематематический характер, т. е. она относится не только к геометрическим фигурам, но также и числам, площадям и иным величинам. Например, в применении к числам она означает, что если числа a и b порознь равны третьему числу c, т. е. $a=c$ и $b=c$, то эти числа равны между собой, т. е. $a=b$.

Эта аксиома вместе с «очевидными» условиями рефлексивности и симметричности равносильна *транзитивности*. Таким образом, этой аксиомой Евклид фактически вводил *отношение эквивалентности*.

С понятиями геометрии дело обстоит примерно так же, как с теоремами. Каждое новое понятие вводится (с помощью определения) на основе ранее введенных понятий. Эти ранее введенные понятия, в свою очередь, определяются через предыдущие и т. д. В конце концов мы приходим к небольшому числу *первоначальных понятий*, через которые, идя обратным путем, можно определить все встречающиеся в курсе геометрии понятия. Евклид попытался дать определение всех без исключения понятий геометрии, в том числе и первоначальных. Например, «точка», относящаяся к числу первоначальных понятий, определяется у Евклида следующим образом: «Точка есть то, что не имеет частей».

Нетрудно понять, однако, что эта фраза не может служить определением, так как сразу же возникают дальнейшие вопросы: какой смысл имеет слово «то»? Что такое «часть»? Что означает «иметь часть»? Неудивительно, что это туманное «определение» нигде дальше в изложении Евклида не используется. По существу, фраза Евклида представляет собой обращение к нашему жизненному опыту: Евклид убежден, что читатель имеет представление о чем-то предельно маленьком, неделимом далее, и предлагает это понятие именовать термином «точка». Таким образом, следует признать, что у Евклида понятие «точка» является первоначальным, вводимым без определения, а фраза Евклида является лишь пояснением, показывающим опытное происхождение этого понятия.

Современная точка зрения на аксиоматическое построение курса геометрии заключается в следующем.

Во-первых, перечисляются первоначальные (не определяемые) понятия.

Во-вторых, указывается список аксиом. При формулировке аксиом используются первоначальные понятия. Смысл аксиом в том, что они устанавливают некоторые связи и взаимоотношения между первоначальными понятиями.

После этого дальнейшие геометрические факты (теоремы) выводятся дедуктивным путем (доказываются) с помощью *правил вывода* (в частности, правил использования знаков $\forall$, $\exists$, $\Rightarrow$, $\neg$, $\wedge$, $\vee$, $\Leftrightarrow$).

Первоначальные понятия и аксиомы заимствованы из опыта и возникли в процессе долгого и сложного исторического развития науки. Первоначальные факты накапливаются в результате практической деятельности человека. Их проверяют, уточняют, систематизируют. Исключают из них те, которые могут быть выведены из других первоначальных фактов. Иногда обнаруживается, что оставшийся список первоначальных фактов (аксиом) — неполный, т. е. этих фактов недостаточно для вывода всех теорем, и тогда к этому списку добавляют недостающие аксиомы. В результате и получается полный набор аксиом (аксиоматика).

Это делает понятным, почему все последующие геометрические факты, хотя они и получаются на основе системы аксиом чисто умозрительным, дедуктивным путем, имеют тесную связь с жизнью, достаточно хорошо описывают свойства реального пространства, применяются в практической деятельности человека.

Евклид не сумел последовательно и до конца провести аксиоматическую точку зрения. Дело не только в том, что он не отделил первоначальные, неопределяемые понятия от последующих понятий. Основной недостаток изложения Евклида заключается в том, что приведенный им список аксиом был *неполным*.

После Евклида многие поколения математиков стремились улучшить данную им систему аксиом (аксиоматику) геометрии. Большую роль сыграли работы современника Евклида — древнегреческого ученого Архимеда, который сформулировал аксиомы, относящиеся к измерению геометрических величин. Из математиков более позднего времени большой вклад в усовершенствование аксиоматики геометрии внесли замечательный русский математик Н. И. Лобачевский, а также А. Лежандр, М. Паш и другие ученые.

Итог этих исследований был подведен выдающимся немецким математиком Д. Гильбертом, который на рубеже XIX и XX столетий завершил работу по аксиоматизации геометрии.

Полный список аксиом геометрии содержит около двух десятков аксиом и является достаточно сложным. Поэтому знакомство с аксиоматическим методом проведем на более простых примерах: *метрического пространства* и *группы*.

58. Метрические пространства

В этом пункте рассматривается аксиоматическое построение геометрии *метрических пространств*, широко используемых в современной математике (причем не только в геометрии).

Прежде всего мы должны перечислить первоначальные понятия. В теории метрических пространств имеется два первоначальных понятия: *точка* и *расстояние*. Расстояние между точками A и B обозначается через $d(A, B)$.

Далее мы должны перечислить аксиомы. В теории метрических пространств имеется всего три аксиомы.

Аксиома 1. *Расстояние между двумя любыми точками A и B есть число неотрицательное, причем $d(A, B) = 0$ в том и только в том случае, если $A = B$.*

Аксиома 2. *Для любых двух точек A и B справедливо равенство*

$$d(A, B) = d(B, A).$$

Аксиома 3. *Для любых трех точек A, B и C справедливо неравенство*

$$d(A, C) \le d(A, B) + d(B, C).$$

Аксиому 2 иногда называют *аксиомой симметричности расстояния*, а аксиому 3 — *аксиомой треугольника*.

Таким образом, метрическим пространством называется множество $\boldsymbol{M}$, элементы которого считаются точками и в котором задано расстояние, удовлетворяющее перечисленным трем аксиомам.

В качестве первого примера рассмотрим множество $\boldsymbol{P}$ всех точек плоскости. Расстояние между точками будем понимать в обычном смысле (единица длины предполагается заданной).

Множество $\boldsymbol{P}$ является метрическим пространством, т. е. расстояния между точками на плоскости удовлетворяют аксиомам 1, 2 и 3.

Рассмотрим другой пример метрического пространства. Будем теперь считать «точкой» произвольное действительное число, а за «расстояние» между a и b примем модуль разности этих чисел:

$$d(a, b) = |a - b|.$$

Докажем, что множество всех действительных чисел $\boldsymbol{R}$ есть метрическое пространство. Для этого надо проверить, что рассматриваемые «точки» и «расстояния» удовлетворяют аксиомам 1, 2 и 3.

Справедливость аксиом 1 и 2 легко проверяется. Например, справедливость аксиомы 2 проверяется так:

$$d(a, b) = |a - b| = |b - a| = d(b, a).$$

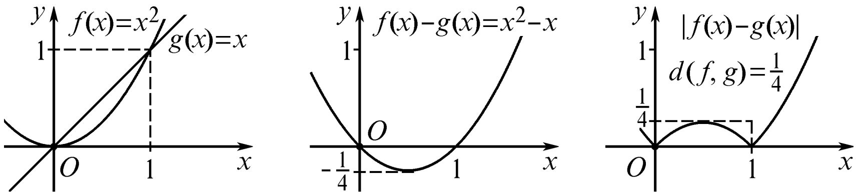


Рис. 359

Справедливость аксиомы 3 следует из неравенства $|x + y| \leq |x| + |y|$, доказываемого в курсе алгебры. Действительно, если в этом неравенстве положить $x = a - b$, $y = b - c$, то оно запишется в виде

$$|a - c| \leq |a - b| + |b - c|.$$

А это и означает, что $d(a, c) \leq d(a, b) + d(b, c)$.

В качестве третьего примера метрического пространства рассмотрим множество $\boldsymbol{S}$ всех многочленов от буквы x, т. е. «точкой» метрического пространства $\boldsymbol{S}$ будет произвольный многочлен $f(x) = a_0x^n + a_1x^{n-1} + ... + a_n$. Под «расстоянием» $d(f, g)$ между «точками» будем понимать наибольшее значение функции $|f(x) - g(x)|$ на отрезке $0 \leq x \leq 1$ (рис. 359). Проверка показывает, что и в этом случае введенное расстояние удовлетворяет аксиомам 1, 2 и 3, т. е. множество $\boldsymbol{S}$ представляет собой метрическое пространство.

Итак, мы рассмотрели три примера, или, как говорят математики, три *модели* метрического пространства, причем модель $\boldsymbol{P}$ имеет геометрический характер, а модели $\boldsymbol{R}$ и $\boldsymbol{S}$ — алгебраический. В разных моделях «точки» выглядят по-разному. Например, в модели $\boldsymbol{P}$ точки имеют привычный смысл, в модели $\boldsymbol{R}$ за «точки» принимаются числа, а в модели $\boldsymbol{S}$ — многочлены. В соответствии с этим в различных моделях по-разному определяются и расстояния.

В приложениях современной математики рассматриваются метрические пространства, «точками» которых являются линии, фигуры, траектории полета космических кораблей, плановые задания заводов и т. д. В чем же заключается выгода от введения общего термина «метрическое пространство» для всех этих столь различных между собой случаев? Дело в том, что, доказав (на основе аксиом 1, 2 и 3) некоторую теорему в геометрии метрических пространств, мы сможем утверждать, что эта теорема будет справедлива и в модели $\boldsymbol{P}$, и в модели $\boldsymbol{S}$, и в модели $\boldsymbol{R}$, и вообще в любой модели метрического пространства. Ведь в каждой модели аксиомы 1, 2 и 3 выполняются, а при доказательстве теорем в аксиоматической теории метрических пространств мы только этими аксиомами и пользуемся.

Таким образом, построение теории метрических пространств (или иной аксиоматической теории) дает выводы, применимые к различным областям: геометрии, алгебре, экономике, астронавтике и вообще в любом случае, когда выполняются рассматриваемые аксиомы. Положение вещей здесь в какой-то мере аналогично тому, которое мы имели при выводе геометрических фактов с помощью рассуждений. Например, измерив транспортиром меры углов конкретного треугольника и убедившись, что сумма этих мер равна, насколько позволяет судить точность используемых инструментов, 180° (и даже проведя такие измерения для многих конкретных треугольников), мы еще не можем утверждать, что это справедливо для любого треугольника. А геометрическое рассуждение (доказательство) позволяет, не проводя повторных измерений углов треугольника, утверждать, что этот вывод имеет *общий* характер, т. е. что в *каждом* треугольнике сумма мер углов равна 180°.

Подобно этому, развив аксиоматическую теорию, мы можем, не проводя повторных рассуждений, утверждать, что выводы этой аксиоматической теории имеют место в *каждой* модели метрического пространства. Это позволит, один раз доказав теоремы о метрическом пространстве, применять их в геометрии, алгебре, экономике, астронавтике, т. е. во всех тех областях, где появляются метрические пространства.

Разница, однако, будет в том, что отдельное доказательство (например, в геометрии) позволяет получить *теорему*, приложимую хотя и к различным фигурам, но в рамках одной области — геометрии. Аксиоматический же метод позволяет целые аксиоматически развитые *теории* применять в различных областях знаний. В этом состоит сила аксиоматического метода, широко используемого сейчас в различных областях как мощное орудие познания.

Здесь мы рассмотрим лишь две самые простые теоремы и одно определение из теории метрических пространств для того, чтобы показать, как проводится доказательство в аксиоматической теории.

Теорема 1. *Пусть* $a_1, a_2, \ldots, a_n$ — *произвольные точки метрического пространства* $\boldsymbol{M}$ (*где* $n \geq 3$). *Тогда*

$$d(a_1, a_n) \leq d(a_1, a_2) + d(a_2, a_3) + \ldots + d(a_{n-1}, a_n).$$

Доказательство. При $n = 3$ доказываемое неравенство имеет вид $d(a_1, a_3) \leq d(a_1, a_2) + d(a_2, a_3)$; оно вытекает непосредственно из аксиомы треугольника.

Теперь, для проведения доказательства по индукции, достаточно осуществить переход от n к $n + 1$. Пусть для некоторого $n \geq 3$ утверждение теоремы справедливо. Возьмем произвольные $n + 1$ точку $a_1, \ldots, a_n, a_{n+1}$. В силу предположения индукции справедливо неравенство $d(a_1, a_n) \leq d(a_1, a_2) + d(a_2, a_3) + \ldots + d(a_{n-1}, a_n)$. Кроме того, согласно аксиоме треугольника, $d(a_1, a_{n+1}) \leq d(a_1, a_n) + d(a_n, a_{n+1})$. Складывая эти неравенства и производя упрощения, мы получаем соотношение $d(a_1, a_{n+1}) \leq d(a_1, a_2) + d(a_2, a_3) + \ldots + d(a_{n-1}, a_n) + d(a_n, a_{n+1})$, которое и показывает, что утверждение теоремы справедливо для $n+1$ точки. Этим и завершается индукция.

Определение. *r-окрестностью* точки a в метрическом пространстве $\boldsymbol{M}$ называется множество всех точек $x \in \boldsymbol{M}$, находящихся от a на расстоянии, меньшем r.

Обозначать эту окрестность будем через $\boldsymbol{U}_r(a)$:

$$\boldsymbol{U}_r(a) = \{x \in \boldsymbol{M}\colon d(a, x) < r\}.$$

Теорема 2. *Обозначим через p расстояние между точками a и b. Тогда если $p < r$, то $(r-p)$-окрестность точки b содержится в r-окрестности точки a.*

Знаками эту теорему можно записать так:

$$(\forall\ a, b \in \boldsymbol{M};\ r > 0)\ (p = d(a, b) < r) \Rightarrow (\boldsymbol{U}_{r-p}(b) \subset \boldsymbol{U}_r(a)).$$

Доказательство. Пусть x принадлежит $(r-p)$-окрестности точки b, т. е. $d(b, x) < r-p$. Это неравенство можно переписать в виде $p + d(b, x) < r$, т. е. $d(a, b) + d(b, x) < r$. По аксиоме треугольника $d(a, x) \le d(a, b) + d(b, x)$. Поэтому имеем $d(a, x) < r$, т. е. точка x принадлежит r-окрестности точки a.

Итак, для любой точки $x \in \boldsymbol{U}_{r-p}(b)$ имеем $x \in \boldsymbol{U}_r(a)$, т. е. $\boldsymbol{U}_{r-p}(b) \subset \boldsymbol{U}_r(a)$.

Доказанные теоремы 1 и 2 справедливы в *любом* метрическом пространстве.

Рассмотрим, какой смысл имеют эти теоремы в метрическом пространстве $\boldsymbol{P}$ (на плоскости) и в метрическом пространстве $\boldsymbol{R}$ (в множестве всех действительных чисел).

На плоскости $\boldsymbol{P}$ расстояние $d(A, B)$ между точками A и B равно длине соединяющего их отрезка AB. Поэтому сумма

$$d(A_1, A_2) + d(A_2, A_3) + \ldots + d(A_{n-1}, A_n)$$

представляет собой сумму длин отрезков A_1A_2, A_2A_3, ..., $A_{n-1}A_n$, т. е. *длину ломаной* $A_1A_2 \ldots A_n$. Далее, $d(A_1, A_n)$ есть длина отрезка A_1A_n, соединяющего концы этой ломаной. Таким образом, в применении к плоскости $\boldsymbol{P}$ теорема 1 означает, что длина отрезка, соединяющего концы ломаной, не превосходит длины этой ломаной.

В метрическом пространстве $\boldsymbol{R}$ теорема 1 имеет следующий смысл: для любых действительных чисел a_1, a_2, ..., a_n справедливо неравенство $|a_1 - a_n| \le |a_1 - a_2| + |a_2 - a_3| + \ldots + |a_{n-1} - a_n|$.

Обратимся теперь к теореме 2.

На плоскости $\boldsymbol{P}$ множество $\boldsymbol{U}_r(A) = \{X\colon\ d(A, X) < r\}$ представляет собой *открытый круг* (круг без границы) с центром A и радиусом r. Теорема 2 утверждает, что если расстояние p между точками A и B

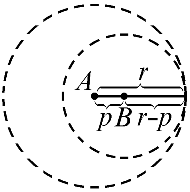


Рис. 360

Рис. 361

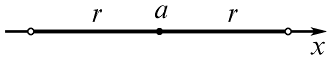
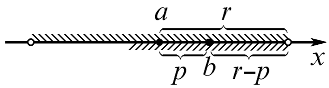


Рис. 362



меньше r, то круг радиуса $r-p$ с центром B содержится в круге радиуса r с центром A (окружности этих кругов касаются внутренним образом, рис. 360).

В метрическом пространстве $\boldsymbol{R}$ множество $U_r(A) = \{x\colon\ |x-a| < r\}$ представляет собой *интервал длины* $2r$, серединой которого служит точка a (рис. 361). Теорема 2 в этом случае означает, что если $|a-b| = p < r$, то интервал длины $2(r-p)$ с центром b содержится в интервале длины $2r$ с центром a (рис. 362).

Задачи и упражнения

236. Пусть точками пространства $\boldsymbol{\Gamma}$ являются города России, а в качестве расстояния между любыми двумя городами возьмем минимальную стоимость проезда из одного города в другой. Будет ли $\boldsymbol{\Gamma}$ метрическим пространством?

237. Будет ли метрическим пространством множество точек сферы, если в качестве расстояния между точками взять:

а) расстояние по прямой;

б) расстояние по дуге большого круга, проходящего через эти точки?

238. Рассмотрим множество кругов на плоскости и введем расстояние между любыми двумя из них как число, равное сумме площадей этих кругов без удвоенной площади их пересечения. Будет ли полученное пространство метрическим?

239. Рассмотрим на плоскости n различных точек. Расстояние между любыми двумя различными точками будем считать равным единице. Образуют ли при любом n эти точки метрическое пространство?

240. Будут ли образовывать метрическое пространство точки координатной прямой, если в качестве расстояния между точками взять $d(a, b) = |a-b|$?

59. Коммутативные группы

В этом пункте приводится еще один пример аксиоматической теории.

Рассмотрим некоторое множество $\boldsymbol{G}$, в котором каждым двум элементам a и b поставлен в соответствие третий элемент c, называемый суммой элементов a и b и обозначаемый через $a+b$. Таким образом, первоначальными понятиями являются *элемент* и *сумма элементов*. Операция нахождения суммы называется *сложением*. Перечислим теперь аксиомы. В рассматриваемой теории их четыре.

Аксиома 1. *Сложение коммутативно, т. е. для любых двух элементов a и b справедливо соотношение $a+b=b+a$.*

Аксиома 2. *Сложение ассоциативно, т. е. для любых трех элементов a, b и c справедливо соотношение $(a+b)+c=a+(b+c)$.*

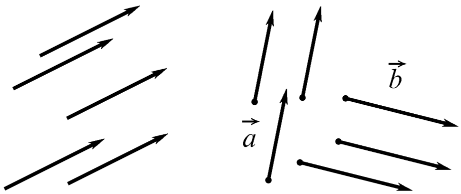


Рис. 363

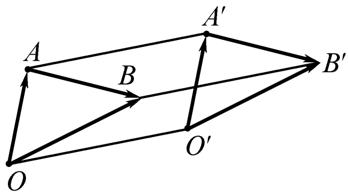


Рис. 364

Аксиома 3. *Существует элемент* 0, *обладающий тем свойством, что для любого a справедливо равенство* $a + 0 = a$.

Аксиома 4. *Для любого элемента a существует элемент, который в сумме с элементом a дает* 0. *Этот элемент называется противоположным элементу a и обозначается через* $-a$.

Множество $\boldsymbol{G}$, для элементов которого указано, что понимается под их суммой, и в котором выполняются сформулированные выше аксиомы 1 – 4, называется *коммутативной группой*. Аксиомы группы символами записываются следующим образом:

Аксиома 1. $(\forall\ a, b \in \boldsymbol{G})\ (a + b = b + a)$;

Аксиома 2. $(\forall\ a, b, c \in \boldsymbol{G})\ ((a + b) + c = a + (b + c))$;

Аксиома 3. $(\exists\ 0 \in \boldsymbol{G})\ (\forall\ a \in \boldsymbol{G})\ (a + 0 = a)$;

Аксиома 4. $(\forall\ a \in \boldsymbol{G})\ (\exists\ (-a) \in \boldsymbol{G})\ (a + (-a) = 0)$.

Рассмотрим множество $\boldsymbol{T}$ всех *параллельных переносов* плоскости. Под суммой двух параллельных переносов f и g будем в данном случае понимать их композицию $g \circ f$ (сначала f, потом g), т. е. будем писать $f + g = g \circ f$.

Из свойств параллельных переносов, изученных в курсе геометрии, непосредственно вытекает, что аксиомы 1 – 4 в данном случае выполняются. Таким образом, множество $\boldsymbol{T}$ всех параллельных переносов плоскости с указанной выше операцией сложения является коммутативной группой.

Рассмотрим теперь множество $\boldsymbol{V}$ всех векторов на плоскости. Напомним, что вектором называется класс всех направленных отрезков, имеющих одну и ту же длину и одно и то же направление (рис. 363). Сложение векторов определяется по представителям, т. е. выбирается направленный отрезок $\overrightarrow{OA}$, являющийся представителем вектора $\vec{a}$, и направленный отрезок $\overrightarrow{AB}$, являющийся представителем вектора $\vec{b}$ (тем самым конец первого направленного отрезка совпадает с началом второго, рис. 364), и тогда вектор, представителем которого служит направленный отрезок $\overrightarrow{OB}$, обозначается через $\vec{a} + \vec{b}$ и называется *суммой* векторов $\vec{a}$ и $\vec{b}$. Из школьного курса известно, что множество $\boldsymbol{V}$ всех векторов на плоскости удовлетворяет аксиомам 1 – 4 и потому представляет собой коммутативную группу. Кстати, именно тем, что не удается непосредственно определить сумму двух направленных отрезков (например, OA и $A'B'$ на рис. 364) и объясняется определение вектора не как одного направленного отрезка, а как класса направленных отрезков. Определение «вектором называется направленный отрезок» является математически некорректным.

Дальнейшими примерами коммутативных групп могут служить: множество всех поворотов плоскости вокруг заданной точки O (в качестве операции сложения рассматривается композиция поворотов); множество всех целых чисел $\boldsymbol{Z}$ (включая отрицательные числа) с обычной операцией сложения; множество всех многочленов степени не выше второй (при обычном сложении); множество всех многочленов вообще.

Приведем в качестве примера несколько теорем теории коммутативных групп.

Теорема 1. *Элемент* 0 *коммутативной группы* $\boldsymbol{G}$ *определен однозначно*: $(\exists!\, 0 \in \boldsymbol{G})\ (\forall\ a \in \boldsymbol{G})\ (a + 0 = a)$.

Доказательство. Допустим, напротив, что существует, помимо 0, еще один элемент $0'$, обладающий тем же свойством $(\forall\ a \in \boldsymbol{G})\ (a + 0' = a)$. Тогда $0 + 0' = 0$, поскольку $0'$ обладает указанным свойством. С другой стороны, $0 + 0' = 0' + 0$ (аксиома 1), и потому $0 + 0' = 0' + 0 = 0'$ (аксиома 3). Таким образом, 0 и $0'$ совпадают с одним и тем же элементом $0 + 0'$, т. е. $0 = 0'$.

Аналогично устанавливается единственность противоположного элемента.

Теорема 2. *Пусть* a_1, a_2, ..., a_n — *элементы коммутативной группы* $\boldsymbol{G}$. *Тогда в каком порядке мы бы их последовательно ни складывали друг с другом, результат будет одним и тем же.*

Поясним формулировку этой теоремы на примере четырех элементов a_1, a_2, a_3 и a_4. Теорема утверждает, что при вычислении суммы этих элементов мы можем располагать слагаемые и выполнять сложение в произвольном порядке. Например,

$$((a_1 + a_2) + a_3) + a_4 = (a_3 + a_4) + (a_2 + a_1). \tag{1}$$

Здесь в левой части сначала находится сумма $a_1 + a_2$, затем к ней прибавляется элемент a_3 и, наконец, к результату прибавляется a_4; в правой же части порядок расположения слагаемых и порядок выполнения действий другой.

В общем виде доказательство этой теоремы можно провести с помощью метода математической индукции.

Теорема 3. *Для любых двух элементов a и b* группы $\boldsymbol{G}$ *уравнение* $a + x = b$ *всегда имеет в группе* $\boldsymbol{G}$ *решение и притом единственное.*

Знаками формулировку этой теоремы можно записать так:

$$(\forall\ a, b \in \boldsymbol{G})\ (\exists!\, x \in \boldsymbol{G})\ (a + x = b).$$

Доказательство. Для установления существования достаточно указать элемент, являющийся решением (корнем) этого уравнения. Проверим, что таким корнем является элемент $x = b + (-a)$. В самом деле, применяя последовательно аксиомы 1, 2, 4, 1 и 3, получаем (для этого значения x):

$$a + x = a + (b + (-a)) = a + ((-a) + b) = (a + (-a)) + b = 0 + b = b + 0 = b.$$

Этим существование корня доказано.

Докажем единственность. Пусть x_1, x_2 — два корня рассматриваемого уравнения, т. е.

$$a + x_1 = b, \quad a + x_2 = b. \tag{2}$$

Докажем, что $x_1 = x_2$. Действительно, из равенств (2) следует, что $a + x_1 = a + x_2$. Прибавим к элементу $(-a)$ равные между собой элементы $a + x_1$ и $a + x_2$: $(-a) + (a + x_1) = (-a) + (a + x_2)$. В силу аксиомы 2 это равенство можно переписать в виде $((-a) + a) + x_1 = ((-a) + a) + x_2$, откуда, применяя аксиомы 1, 4 и 3, находим, что $x_1 = x_2$.

Замечание. Для краткости элемент $b + (-a)$ принято записывать без скобок в виде $b - a$ и называть *разностью* элементов b и a. Таким образом, по определению

$$b - a = b + (-a).$$

Доказанная теорема утверждает, что уравнение

$$a + x = b \tag{3}$$

имеет в коммутативной группе $\boldsymbol{G}$ единственное решение

$$x = b - a. \tag{4}$$

Иными словами, равенства (3) и (4) означают одно и то же, т. е. определяют один и тот же элемент x. Сравнивая форму записи равенств (3) и (4), мы получаем обоснование следующего хорошо известного правила.

Слагаемое из одной части равенства можно переносить в другую, меняя знак, стоящий перед этим слагаемым, на противоположный.

Доказанные теоремы 1, 2, 3 справедливы в любой коммутативной группе. Иначе говоря, какой бы пример (модель) коммутативной группы мы ни взяли, мы можем, не проводя заново доказательств, утверждать, что эти теоремы будут справедливы для этой модели коммутативной группы. В частности, они справедливы не только для рациональных чисел, но и для параллельных переносов, для векторов, для многочленов и т. д. Вообще, если нам в дальнейшем встретится пример коммутативной группы, мы можем применять теоремы 1, 2, 3, т. е. выполнять сложение в любом порядке, решать линейные уравнения, переносить слагаемые из одной части равенства в другую с изменением знака.

Коммутативные группы используются в разных областях приложений современной математики: в экономике, кристаллографии, ядерной физике, биологии и т. д.

В математике и в ее приложениях нередко встречаются также «некоммутативные группы», т. е. множества, в которых определена операция (сложение), удовлетворяющая только аксиомам 2, 3 и 4, но в которых соотношение коммутативности может выполняться не для всех элементов. При этом формулировки аксиом 3 и 4 видоизменяются: требуется существование не только правого нуля (как в аксиоме 3), но и левого нуля; не только правого, но и левого противоположного элемента. Единственность каждого из этих нулей (т. е. совпадение левого нуля и правого нуля) и единственность каждого из противоположных элементов имеют место и в этом случае, но доказываются чуть сложнее.

Примером группы может служить множество $\boldsymbol{D}$ всех движений плоскости, в котором в качестве операции сложения берется композиция: $f + g = g \circ f$. Элементом 0 служит тождественное преобразование $e \in \boldsymbol{D}$: для любого элемента $f \in \boldsymbol{D}$ имеем $e \circ f = f$, т. е. $f + 0 = f$. Это означает, что выполнена аксиома 3. Наконец, противоположным элементу f служит обратное преобразование f^{-1}: $f^{-1} \circ f = e$, т. е. $f + (-f) = 0$.

Группа $\boldsymbol{D}$ не является коммутативной. Если, например, f — поворот на 90° вокруг точки O, а g — параллельный перенос на вектор $\vec{a} \neq 0$, то точки $f(g(M))$ и $g(f(M))$ различны, т. е. $g + f \neq f + g$.

В качестве еще одного примера можно указать *группу подстановок* из n элементов, т. е. множество всех взаимно однозначных отображений множества $\boldsymbol{M} = \{1, 2, \dots n\}$ на себя. Элементы этой группы обычно задаются в виде *двухстрочной матрицы* вида

$$\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix},$$

где i_1 есть образ элемента 1, далее, i_2 — образ элемента 2 и т. д. При этом все элементы i_k различны (поскольку отображение взаимно однозначно), т. е. вторая строка представляет собой некоторую *перестановку* элементов 1, 2, ..., n. Суммой двух подстановок f, g, т. е. взаимно однозначных отображений $\boldsymbol{M} \to \boldsymbol{M}$, называется *композиция* этих отображений: $f + g = g \circ f$. Роль элемента 0 здесь играет *тождественная подстановка* e, а противоположным для элемента f является *обратное отображение* f^{-1}. Эта группа подстановок обозначается через S_n и называется *симметрической группой* порядка n. Она содержит $n!$ элементов. Таким образом, симметрическая группа (любого порядка) представляет собой пример конечной группы.

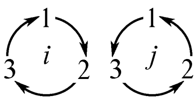


Рис. 365

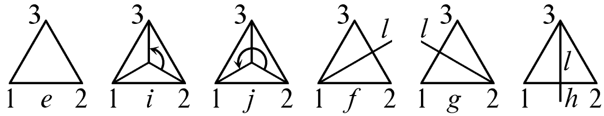


Рис. 366

Вообще, пусть $\boldsymbol{M}$ — произвольное множество. *Преобразованием* множества $\boldsymbol{M}$ называется всякое взаимно однозначное отображение множества $\boldsymbol{M}$ на себя. Для любых двух преобразований f, g множества $\boldsymbol{M}$ можно определить их композицию $g \circ f$, которую условимся обозначать через $f + g$ и называть *суммой* этих преобразований. Пусть теперь $\boldsymbol{G}$ — некоторое множество преобразований множества $\boldsymbol{M}$ (не обязательно совпадающее с множеством *всех* его преобразований). Говорят, что $\boldsymbol{G}$ представляет собой группу преобразований множества $\boldsymbol{M}$, если выполняются следующие три условия:

1) $e \in \boldsymbol{G}$; 2) $(\forall f \in \boldsymbol{G})\,(f^{-1} \in \boldsymbol{G})$; 3) $(\forall f, g \in \boldsymbol{G})\,(g \circ f \in \boldsymbol{G})$.

Нетрудно проверить, что если $\boldsymbol{G}$ — группа преобразований множества $\boldsymbol{M}$, то $\boldsymbol{G}$ является группой.

По существу, мы имеем здесь дело с новой аксиоматической теорией — теорией групп преобразований, в которой указанные выше условия 1 – 3 служат аксиомами. В качестве примера рассмотрим группу всех преобразований множества $\boldsymbol{M}$, состоящего из трех элементов: $\boldsymbol{M} = \{1, 2, 3\}$ (т. е. симметрическую группу порядка 3). В данном случае преобразований, т. е. взаимно однозначных отображений множества $\boldsymbol{M}$ на себя, существует шесть:

e	i	j	f	g	h
$1 \to 1$	$1 \to 2$	$1 \to 3$	$1 \to 1$	$1 \to 3$	$1 \to 2$
$2 \to 2$	$2 \to 3$	$2 \to 1$	$2 \to 3$	$2 \to 2$	$2 \to 1$
$3 \to 3$	$3 \to 1$	$3 \to 2$	$3 \to 2$	$3 \to 1$	$3 \to 3$

Здесь e — тождественное отображение, i, j — так называемые циклические перестановки (рис. 365), а каждое из преобразований f, g, h оставляет один элемент множества $\boldsymbol{M}$ неподвижным (т. е. переводит его в себя), а два другие переставляет между собой.

Заметим, что эта группа преобразований по существу совпадает с группой самосовмещений равностороннего треугольника. В самом деле, обозначим вершины равностороннего треугольника цифрами 1, 2, 3 (рис. 366). Группа самосовмещений этого треугольника состоит из поворотов на 0°, 120°, 240° вокруг центра треугольника (это дает преобразования e, i, j), а также трех осевых симметрий (например, g — симметрия относительно оси, проходящей через вершину 2).

Задачи и упражнения

241. Является ли коммутативной группой множество нечетных чисел с обычной операцией сложения $a + b$?

242. Является ли коммутативной группой множество отображений (поворотов и симметрий) правильного шестиугольника на себя?

243. Образуют ли группу нечетные числа относительно операции умножения $a \circ b = ab$?

244. Образует ли группу множество целых чисел, если в качестве операции $a \circ b$ берется остаток при делении числа $a + b$ на 3?

245. Будет ли группой множество целых чисел, если для любых чисел a и b их сумма считается равной нулю?

Беседа 13. Непротиворечивость, независимость, полнота

60. Непротиворечивость и понятие модели

Имеются важные принципиальные категории, относящиеся к любым системам аксиом. К ним относится в первую очередь понятие *непротиворечивости*. Мы рассмотрим его на простом примере.

Несколько восьмиклассников решили организовать шахматный турнир. Однако при обычной схеме (когда каждый участник играет с каждым из остальных и притом по две партии: белыми и черными) требуется очень много времени для проведения турнира. Поэтому мальчики решили провести турнир по упрощенной схеме: каждый должен сыграть ровно три партии с кем-либо из остальных участников (а белыми или черными — по жребию). Так как составить расписание турнира никак не удавалось, двое из мальчиков, Петя и Сережа, решили обратиться к учителю.

Учитель прежде всего поинтересовался, четно или нечетно число участников предполагаемого турнира. Оказалось, что число участников нечетно. Тогда учитель предложил сформулировать требования, которые ученики предъявили к турниру, в виде аксиом. Для этого потребовалось ввести три первоначальных (неопределяемых) понятия: «игрок», «партия», «участие игрока в партии». Аксиом получилось четыре:

Аксиома 1. *Число игроков нечетно.*

Аксиома 2. *Каждый игрок участвует в трех партиях.*

Аксиома 3. *В каждой партии участвуют два игрока.*

Аксиома 4. *Для каждых двух игроков имеется не более одной партии, в которой они оба участвуют.*

Из этих аксиом можно вывести ряд теорем. Первую из них предложил для примера сам учитель.

Теорема 1. *Число игроков не меньше пяти.*

Доказательство. Так как нуль — четное число, то по аксиоме 1 число игроков не равно нулю, т. е. существует хотя бы один игрок A. Этот игрок в силу аксиомы 2 участвует в трех партиях, причем в каждой из этих партий, помимо A, участвует еще один игрок (аксиома 3). Пусть B, C, D — игроки, отличные от A, которые участвуют в этих партиях. По аксиоме 4 все игроки B, C, D *различны* (если бы, например, было $B = C$, то оказалось бы, что имеются *две* партии, в которых участвуют игрок A и игрок $B = C$). Итак, мы уже нашли четырех игроков A, B, C, D. Но тогда, по аксиоме 1, число игроков не меньше пяти.

Следующую теорему доказал Петя. Для этого он определил новое понятие: если q — некоторая партия и A — один из участвующих в ней игроков, то пару $(q; A)$ назовем *выступлением* игрока.

Теорема 2. *Число всех выступлений игроков четно.*

Доказательство. Если в партии q участвуют игроки A и B, то мы получаем два выступления игроков: $(q; A)$ и $(q; B)$, т. е. каждая партия дает ровно два выступления игроков (аксиома 3). Значит, число всех выступлений игроков четно, так как оно вдвое больше числа всех партий.

Однако Сережа доказал теорему, противоречащую предыдущей:

Теорема 3. *Число всех выступлений игроков нечетно.*

Доказательство. По аксиоме 2 игрок A участвует ровно в трех партиях, скажем, q_1, q_2, q_3. Это дает три выступления игрока A: $(q_1; A)$, $(q_2; A)$, $(q_3; A)$. Отсюда следует, что число *всех* выступлений игроков равно $3n$, где n — число участников. Так как n нечетно (аксиома 1), то и $3n$ нечетно.

Учитель подвел итоги проведенного исследования. Взятая аксиоматика позволяет доказать ряд теорем (три из них были рассмотрены). Однако среди этих теорем имеются две, противоречащие друг другу: теорема 2 утверждает, что некоторое число *четно*, а теорема 3 гласит, что то же самое число *нечетно*. Аксиоматика, из которой можно вывести две противоречащие друг другу теоремы, называется *противоречивой*. Таким образом, взятая аксиоматика противоречива, т. е. требования, выдвинутые организаторами турнира, несовместимы. Неудивительно, что мальчики не сумели составить расписание турнира: такого расписания просто не существует.

Противоречивая аксиоматика не может служить основой построения содержательной теории. «Теория» основанная на противоречивой аксиоматике, считается несостоятельной, бессодержательной. Математика такие «теории» отвергает.

Зачем учитель предложил Пете и Сереже другую систему организации турнира, при которой каждый из участников должен сыграть, не три, а *четыре* партии с кем-либо из остальных участников. Иначе говоря, он предложил рассмотреть теорию, в которой те же первоначальные понятия (игрок, партия и участие игрока в партии); далее, аксиомы 1, 3, 4 остаются без изменения, а аксиома 2 формулируется по-новому:

Аксиома 2. *Каждый игрок участвует в четырех партиях.*

Однако Петя и Сережа не спешили выводить теоремы из этих аксиом: вдруг опять, после доказательства нескольких теорем, обнаружится противоречие. Учитель же заверил мальчиков, что, сколько бы теорем они ни выводили из этих аксиом, никогда противоречий не будет. Вот как он убедил их в этом.

Рассмотрим девятиугольник (рис. 367), в котором, кроме сторон, проведены девять диагоналей, соединяющих вершины через одну.

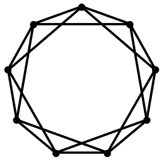


Рис. 367

Вершины девятиугольника будем считать «игроками», проведенные отрезки (стороны и диагонали) будем считать «партиями», а «игроками», участвующими в некоторой «партии», будем считать концы соответствующего отрезка. Мы получаем *модель* (или схему) интересующего нас турнира. Легко проверяется, что все четыре аксиомы (в том числе аксиома 2 в новой редакции) здесь выполняются. Итак, удалось построить модель, в которой выполняются все рассматриваемые аксиомы, причем эта модель построена из «материала» геометрии, т. е. науки, в непротиворечивости которой мы не сомневаемся.

Предположим теперь, что из рассматриваемых четырех аксиом можно вывести две теоремы, противоречащие друг другу. Тогда доказательства этих теорем можно было бы повторить и в построенной модели (ведь в этой модели все четыре аксиомы имеют место). В результате получилось бы, что, рассуждая о правильном девятиугольнике, мы можем получить две противоречащие друг другу теоремы. Но это означало бы, что *геометрия* — наука противоречивая, чего мы не допускаем. Таким образом, мы должны признать, что две противоречащие друг другу теоремы вывести из рассматриваемых четырех аксиом невозможно.

Вообще, пусть рассматриваются две теории P и Q, причем теория Q задается аксиоматически (и в ее непротиворечивости мы заранее не уверены), а P — это хорошо известная нам теория, в непротиворечивости которой мы не сомневаемся. Если из «материала» теории P удастся построить *модель*, в которой выполняются все аксиомы теории Q, то этим непротиворечивость теории Q считается установленной.

61. Математические примеры моделей

Приведем некоторые математические примеры построения моделей. Примем теорию целых чисел в качестве основной теории, в справедливости которой мы не сомневаемся (причем мы рассматриваем в этой теории не только натуральные числа, но также нуль и отрицательные целые числа). Множество всех целых чисел обозначается через $\boldsymbol{Z}$. Теорию же *поля рациональных чисел* $\boldsymbol{Q}$ мы хотим построить аксиоматически. Именно, в множестве $\boldsymbol{Q}$ определены сложение, умножение и неравенства, удовлетворяющие следующим аксиомам.

Аксиома 1. Множество $\boldsymbol{Q}$ содержит $\boldsymbol{Z}$ в качестве подмножества, причем выполняется *принцип перманентности Ганкеля* — имеющиеся в $\boldsymbol{Q}$ сложение, умножение и порядок совпадают на множестве $\boldsymbol{Z} \subset \boldsymbol{Q}$ с имеющимися в $\boldsymbol{Z}$ сложением, умножением и порядком.

Аксиома 2. $(\forall\, r_1, r_2 \in \boldsymbol{Q})\ (r_1 + r_2 = r_2 + r_1)$.

Аксиома 3. $(\forall\, r_1, r_2, r_3 \in \boldsymbol{Q})\ ((r_1 + (r_2 + r_3)) = ((r_1 + r_2) + r_3))$.

Аксиома 4. $(\forall\, r \in \boldsymbol{Q})\ (r + 0 = r)$.

Аксиома 5. $(\forall\, r \in \boldsymbol{Q})\ (\exists\, (-r) \in \boldsymbol{Q})\ (r + (-r) = 0)$.

Аксиомы 1 – 5 означают, что $\boldsymbol{Q}$ представляет собой *коммутативную группу* по сложению.

Аксиома 6. $(\forall\ r_1, r_2 \in \boldsymbol{Q})\ (r_1 r_2 = r_2 r_1)$.

Аксиома 7. $(\forall\ r_1, r_2, r_3 \in \boldsymbol{Q})\ ((r_1(r_2 r_3)) = ((r_1 r_2) r_3))$.

Аксиома 8. $(\forall\ r \in \boldsymbol{Q})\ (r \cdot 1 = r)$.

Аксиома 9. $(\forall\ r \neq 0, r \in \boldsymbol{Q})\ (\exists\ (r^{-1}) \in \boldsymbol{Q})\ (r \cdot (r^{-1}) = 1)$.

Аксиомы 6 – 9 означают, что множество всех положительных рациональных чисел (см. далее аксиому 13) является группой относительно операции умножения. Она называется *мультипликативной группой*, ее «нулем» является число 1, а «противоположным» для r является обратный элемент r^{-1}.

Аксиома 10. Сложение и умножение связаны свойством дистрибутивности: $(\forall\ r_1, r_2, r_3 \in \boldsymbol{Q})\ ((r_1 + r_2) r_3 = r_1 r_3 + r_2 r_3)$.

Вместе взятые, аксиомы 2 – 10 означают, что множество всех рациональных чисел представляет собой *поле*.

Аксиома 11. Множество $\boldsymbol{Q}$ линейно упорядочено, т. е. имеется отношение порядка (неравенства), которое обладает свойствами транзитивности и исключительности.

Аксиома 12. Сумма двух положительных чисел (т. е. чисел, которые больше нуля) является положительным числом: $(\forall\ r_1, r_2 \in \boldsymbol{Q})\ ((r_1 > 0) \wedge (r_2 > 0)) \Rightarrow (r_1 + r_2 > 0)$.

Аксиома 13. Произведение двух положительных чисел является положительным числом: $(\forall\ r_1, r_2 \in \boldsymbol{Q})\ ((r_1 > 0) \wedge (r_2 > 0)) \Rightarrow (r_1 r_2 > 0)$.

Построим теперь (из материала теории целых чисел) модель для этой аксиоматики. С этой целью рассмотрим множество $\boldsymbol{M}$ всех упорядоченных пар $(m; n)$, где $m, n \in \boldsymbol{Z}$, причем $n > 0$. Иными словами, вместо дроби $\dfrac{m}{n}$ мы будем писать пару $(m; n)$. Использование пар (вместо дробей) удобно тем, что мы будем лишены соблазна применять привычные (и кажущиеся «очевидными») свойства дробей. Доказательства формулируемых ниже теорем мы не приводим.

Определение 1. Пары $(m; n)$ и $(p; q)$ будем считать *эквивалентными*, если $mq = np$.

Теорема 1. *Введенное отношение эквивалентности рефлексивно, симметрично и транзитивно, т. е. оно определяет разбиение множества* $\boldsymbol{M}$ *на классы эквивалентности.*

Класс эквивалентности, содержащий элемент $(m; n)$, будем обозначать $[(m; n)]$, а множество всех классов эквивалентности мы и обозначим через $\boldsymbol{Q}$. Элементы множества $\boldsymbol{Q}$ будем называть рациональными числами.

Определение 2. Сложение и умножение в $\boldsymbol{Q}$ определим (по представителям) формулами

$$[(m; n)] + [(p; q)] = [(mq + np; nq)]; \qquad [(m; n)] \cdot [(p; q)] = [(mp; nq)].$$

Легко проверяется, что эти операции определены корректно, т. е. не зависят от выбора представителей.

Определение 3. Рациональное число $[(p; 1)]$ условимся отождествлять с целым числом p.

Теорема 2. *В силу этих определений в модели* $\boldsymbol{Q}$ *выполнен принцип перманентности Ганкеля* (аксиома 1).

Теорема 3. *В силу определений* 1, 2, 3 *операция сложения в модели* $\boldsymbol{Q}$ *удовлетворяет аксиомам* 2 – 5, *т. е.* $\boldsymbol{Q}$ *является коммутативной группой.*

Теорема 4. *В силу определений* 1, 2, 3 *операция умножения в модели* $\boldsymbol{Q}$ *удовлетворяет аксиомам* 6 – 10, *т. е.* $\boldsymbol{Q}$ *является полем.*

Теорема 5. *В силу определений* 1, 2, 3 *любое рациональное число имеет вид* $r = m \cdot n^{-1}$, *где* $m, n \in \boldsymbol{Z}$, *причем* $n > 0$. Иначе говоря, модель $\boldsymbol{Q}$ является *полем частных* над множеством $\boldsymbol{Z}$.

Определение 4. Рациональное число $r = [m; n]$ условимся считать *положительным*, если $m > 0$. (Напомним, что $n > 0$ по определению множества пар.) Далее, условимся для $r_1, r_2 \in \boldsymbol{Q}$ писать $r_1 > r_2$ (или, иначе, $r_2 < r_1$), если $r_1 = r_2 + d$, где число d положительно.

Теорема 6. *В силу определений* 1, 2, 3 *в модели* $\boldsymbol{Q}$ *выполнены аксиомы* 11 – 13, т. е. $\boldsymbol{Q}$ является упорядоченным полем.

Этим и завершается построение модели. Тем самым непротиворечивость рассматриваемой аксиоматики поля рациональных чисел установлена (при условии, что мы считаем непротиворечивой теорию целых чисел).

Заметим, что таким же путем устанавливается непротиворечивость теории целых чисел — при условии, что мы считаем непротиворечивой теорию натуральных чисел. С этой целью также рассматривается множество всех упорядоченных пар $(m; n)$, где m, n — произвольные натуральные числа (имеется в виду, что эта пара изображает разность $m - n$). Пары $(m; n)$ и $(p; q)$ считаются эквивалентными, если $m + q = n + p$. Затем в множестве получающихся классов эквивалентности вводится сложение и неравенства по представителям: $[(m; n)] + [(p; q)] = [(m + p; n + q)]$; $[(m; n)] > [(p; q)]$ если $m + q > n + p$.

Наконец, проверяется, что в этой модели выполняются все аксиомы теории целых чисел (которые получаются из рассмотренных выше аксиом 1 – 5 заменой $\boldsymbol{Q}$ на $\boldsymbol{Z}$ и вытекающими отсюда изменениями).

Итак, непротиворечивость теории *рациональных* чисел сводится (построением модели) к вопросу о непротиворечивости теории *целых* чисел. В свою очередь непротиворечивость этой последней сводится к непротиворечивости теории *натуральные* чисел. Что же дальше?

Поскольку натуральные числа возникают при пересчете предметов в конечных множествах, может быть, следует построить модель натуральных чисел, предполагая непротиворечивость теории множеств? А как тогда быть с самой теорией множеств?

Да и к тому же при рассмотрении моделей мы доказывали теоремы. Но что такое доказательство? Значит, надо обосновать и правила вывода, лежащие в основе математической логики? Мы здесь отметим лишь, что, согласно теореме, доказанной в начале XX столетия австрийским логиком Куртом Гёделем (1906 – 1978), в *любой* аксиоматической теории можно сформулировать теорему, которая не может быть ни доказана, ни опровергнута в пределах этой теории, без выхода за границы этой теории. Так что с чего-то все-таки надо начинать!

В связи с этим иногда произносят сакраментальную фразу: «В непротиворечивости теории натуральных чисел нас убеждает многовековая история науки». Пусть эта фраза служит читателю утешением!

62. Построение аксиоматики геометрии

В отличие от ранее рассматривавшихся аксиоматических теорий геометрия принадлежит к числу наиболее сложно аксиоматизируемых разделов математики. Система аксиом, впервые предложенная Евклидом, не давала окончательного решения вопроса. Мы отложим до следующего пункта проблему существования *неевклидовых геометрий*, связанную с V постулатом Евклида (или аксиомой о параллельных). Здесь же мы рассмотрим вопрос о *пополнении* евклидова списка евклидовых аксиом, оказавшегося явно недостаточным.

Дело в том, что Евклид, доказывая ту или иную теорему, нередко обращался к чертежу и аргументировал свои доводы словом «Смотри!». Например, вычисляя длины отрезков, на которые в прямоугольном треугольнике делится гипотенуза, к которой из вершины прямого угла проведен перпендикуляр (рис. 368), Евклид, нимало не сумнящеся, говорил, что основание перпендикуляра (точка P) лежит «между» концами гипотенузы.

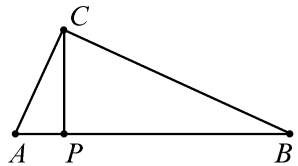


Рис. 368

Но что значит «между»? Проще всего сказать, что это одно из первоначальных, неопределяемых понятий геометрии. Но тогда надо указать аксиомы, описывающие свойства этого понятия — иначе мы не сможем о нем рассуждать! Ведь «Смотри!» или «очевидно, что» для строгой, аксиоматической теории неуместны.

Такими аксиомами могут быть, например, следующие.

Из трех точек на прямой одна и только одна находится между двумя другими.

Если точка C находится между A и B, а точка D находится между A и C, то D находится между A и B.

Легко дать иллюстрации этих аксиом: именно лишь иллюстрации, т. е. что-то вроде утешения для наших наглядных представлений (чтобы не было очень сухо и абстрактно), а вовсе не «подтверждения», «доказательства», «обоснования» и т. п. Ведь в аксиоматической теории все утверждения должны либо быть объявлены аксиомами, либо должны быть получены строгими логическими рассуждениями (доказательствами) на основе аксиом и ранее доказанных теорем. Никакой апелляции к наглядности, очевидности, интуиции быть не должно.

Особенно активно работа по пополнению евклидова списка аксиом проводилась в XIX столетии.

Исследованию вопроса о взаимном расположении точек на прямой и прямых на плоскости посвятил ряд работ немецкий математик Мориц Паш (1843 – 1930). В частности, он заметил, что при аксиоматическом построении геометрии по схеме Евклида необходимо пользоваться следующей аксиомой:

если прямая, лежащая в плоскости треугольника и не проходящая ни через одну из его вершин, пересекает одну из сторон треугольника, то она обязательно пересечет одну из двух других его сторон.

Аксиома Паша означает, что если прямая l «входит» в треугольник T (рис. 369а), то она не может остаться внутри этого треугольника (рис. 369б), а должна выйти из этого треугольника, пересекая какую-либо одну из двух других его сторон (рис. 369в, г). Евклид в подобных случаях ссылался на наглядность чертежа, а это при строго аксиоматическом построении геометрии недопустимо.

Работа по аксиоматизации евклидовой геометрии была завершена в самом конце XIX столетия известным немецким математиком Давидом Гильбертом (1862 – 1943). В своей книге «Основания геометрии» (1899) Д. Гильберт дает полный список аксиом евклидовой геометрии (21 аксиома) и доказывает непротиворечивость этой аксиоматики.

У Гильберта все аксиомы разбиты на пять групп. Первая группа аксиом (аксиомы соединения) содержит, в частности, аксиому о том, что через каждые две точки проходит прямая и что такая прямая — единственная; аксиому о том, что на каждой прямой имеются по крайней мере две точки; существуют по крайней мере три точки, не лежащие на одной прямой, и другие аксиомы.

Вторая группа аксиом (аксиомы порядка) говорит о взаимном расположении точек и прямых. Здесь содержится аксиома о том, что из трех точек на прямой одна и только одна лежит между двумя другими. К этой группе аксиом относится также аксиома М. Паша и ряд других аксиом.

Третья, четвертая и пятая группы аксиом носят соответственно названия: аксиомы конгруэнтности, аксиома параллельности и аксиомы непрерывности.

Но Гильберт не только подвел итог двухтысячелетнему исследованию аксиоматики в духе Евклида и дал полный список аксиом

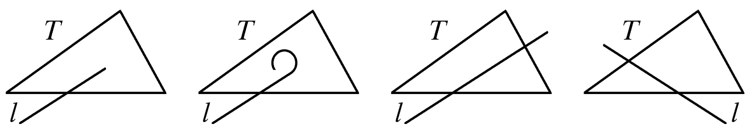


Рис. 369

геометрии. Гильберт также доказал *непротиворечивость* этой аксиоматики. Точнее, он доказал, что из материала теории действительных чисел можно построить *модель* геометрии Евклида. Значит, если теория действительных чисел признается непротиворечивой, то, как показывает построенная Гильбертом модель, геометрия Евклида также лишена противоречий. Что же касается теории действительных чисел, то ее непротиворечивость (как показывают модели, построенные Кантором и Дедекиндом) сводится к непротиворечивости *рациональных чисел*, а это, как мы видели, сводится к непротиворечивости теории натуральных чисел. Этим проблема непротиворечивости геометрии Евклида была решена.

Однако с появлением книги Гильберта работа по аксиоматизации геометрии вовсе не завершилась. Дело в том, что одна и та же теория может быть построена на основе *различных* систем аксиом. Это легко понять на примере аксиоматики коммутативной группы.

Выше были указаны четыре аксиомы коммутативной группы. Однако можно было бы определить коммутативную группу при помощи другой системы аксиом, содержащей следующие три аксиомы.

Аксиома 1*. $(\forall\ a, b \in \boldsymbol{G})\ (a + b = b + a)$.

Аксиома 2*. $(\forall\ a, b, c \in \boldsymbol{G})\ ((a + b) + c = a + (b + c))$.

Аксиома 3*. $(\forall\ a, b \in \boldsymbol{G})\ (\exists!\ x \in \boldsymbol{G})\ (a + x = b)$.

Нетрудно понять, что эта аксиоматика эквивалентна той, которая была указана ранее. В самом деле, аксиома 1* и аксиома 2* просто совпадают с аксиомами 1 и 2. Аксиома 3* выводится как теорема из аксиом 1 – 4 (мы это видели в п. 59).

Таким образом, если выполняются аксиомы 1 – 4, то выполняются и аксиомы 1* – 3*. В свою очередь, из аксиом 1* – 3* можно вывести в качестве теорем утверждения аксиом 1 – 4.

Например, покажем, как выводится аксиома 4. Применяя аксиому 3* к случаю $b = 0$, получаем, что для любого элемента a существует и притом только один элемент $x \in \boldsymbol{G}$, для которого $a + x = 0$, а это и есть аксиома 4.

Из сказанного видно, что все выводы, которые можно получить из аксиом 1 – 4, можно также получить из аксиом 1* – 3*, и наоборот. Таким образом, понятие коммутативной группы может быть определено по-разному: либо с помощью аксиоматики 1 – 4, либо с помощью аксиоматики 1* – 3*.

Аналогично дело обстоит и с аксиоматикой геометрии: могут быть построены разные системы аксиом, задающие одну и ту же теорию — геометрию Евклида. Это выяснилось уже в конце XIX в. В это время, одновременно с гильбертовской аксиоматикой, были предложены и другие системы аксиом евклидовой геометрии. Одна из них была предложена в 1899 г. М. Пиери, другая, чуть позже, — нашим соотечественником В. Ф. Каганом.

Работа по усовершенствованию аксиоматики геометрии продолжалась и в XX столетии. В 1911 г. Ф. Шур предложил аксиоматику, отличавшуюся от гильбертовской тем, что вместо аксиом конгруэнтности (III группа аксиом Д. Гильберта) он ввел аксиомы движений. Таким образом, аксиоматика Ф. Шура оказывается ближе к идеям Ф. Клейна, чем гильбертовская аксиоматика, и в этом смысле более прогрессивной.

Наиболее интересная с современной точки зрения аксиоматика была предложена выдающимся немецким математиком Германом Вейлем (1885 – 1955) в его книге «Пространство, время, материя», вышедшей в 1918 г. Вместо точек и прямых в качестве первоначальных неопределяемых понятий Г. Вейль принимает точки и векторы, а среди его аксиом имеются, в частности, аксиомы, указывающие свойства суммы векторов, произведения вектора на число, скалярного произведения векторов. Таким образом, многие из свойств операций над векторами, которые при обычном изложении геометрии *доказываются*, в аксиоматике Г. Вейля принимаются в качестве *аксиом*. Это позволяет значительно легче и короче (чем при использовании других аксиоматик) осуществить построение курса геометрии.

Когда-то царь Птолемей потребовал, чтобы Евклид указал ему «царский путь» в геометрии, ибо царю негоже идти тем же путем, что и все, а ему надлежит быстро достичь вершин знания: он ведь царь! Евклид ответствовал, что такого пути нет. Чтобы постичь геометрию, царь должен, как и прочие смертные, решать задачи и доказывать одну теорему за другой. Но вейлевская аксиоматика поистине открывает царский путь в геометрию, короче и яснее вводит понятия и факты.

К тому же, в отличие от аксиоматики Гильберта, направленной всецело в прошлое (как завершить дело, начатое свыше двух тысячелетий назад Евклидом?), аксиоматика Г. Вейля направлена в будущее — она связывает геометрию с *теорией векторных пространств* и другими разделами современной математики, применяемыми в физике, химии, математической экономике, биологии и других науках.

63. Геометрия Лобачевского

Как уже отмечалось, система аксиом геометрии, впервые предложенная Евклидом, не давала окончательного решения вопроса. Уточнение аксиоматики Евклида шло не только по линии добавления недостающих аксиом. После того как было обнаружено, что аксиома «Все прямые углы равны между собой» является ненужной (т. е. может быть доказана как теорема с помощью остальных аксиом), были сделаны попытки определить, нет ли у Евклида и других «лишних» аксиом. Особое внимание математиков привлекал *пятый постулат Евклида*. Этот постулат, который у самого Евклида формулировался довольно сложно, в XIX в. был заменен следующей аксиомой параллельности.

Пусть в плоскости даны прямая и лежащая вне этой прямой точка, тогда через эту точку можно провести к данной прямой только одну параллельную прямую.

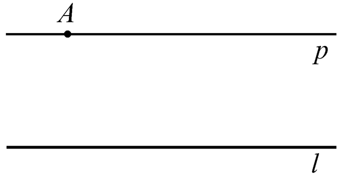


Рис. 370

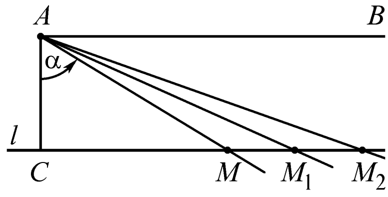


Рис. 371

Иначе говоря, если точка A не принадлежит прямой l, то (в плоскости, содержащей l и A) существует *не более одной* прямой p, для которой истинно $(A \in p) \wedge (l \cap p = \varnothing)$ (рис. 370).

По сравнению с другими аксиомами Евклида, аксиома параллельности формулируется наиболее сложно. Многие теоремы Евклида выражают гораздо более простые факты, например: «в равнобедренном треугольнике углы при основании конгруэнтны». Неудивительно, что многие математики, жившие после Евклида, пытались доказать, что эта аксиома является лишней, т. е. может быть доказана как теорема на основании остальных аксиом. Среди математиков, работавших над разрешением загадки пятого постулата, были Д. Саккери (XVIII в.), И. Ламберт (XVIII в.), Жан Бертран (XIX в.), А. Лежандр (XVIII – XIX вв.) и др. Особенно интересны работы французского математика Лежандра. Его перу принадлежит учебник геометрии, представляющий собой переработку евклидовых «Начал». Этот учебник неоднократно переиздавался. Почти в каждом издании своего учебника Лежандр помещал доказательство пятого постулата, однако после выхода книги выяснялось, что опубликованное доказательство ошибочно. В следующем издании Лежандр помещал новое доказательство. И хотя среди опубликованных Лежандром доказательств не оказалось ни одного верного, тем не менее его работы прояснили кое-что в проблеме пятого постулата.

Заслуга решения проблемы пятого постулата принадлежит нашему замечательному соотечественнику, профессору Казанского университета Николаю Ивановичу Лобачевскому. Первоначально он пытался, как и его предшественники, доказать пятый постулат Евклида, исходя из остальных аксиом. С этой целью он попробовал провести рассуждения методом доказательства «от противного».

Исходную идею Лобачевского можно изложить следующим образом. Рассмотрим в плоскости прямую l и не принадлежащую ей точку A. Проведем из точки A перпендикуляр AC к прямой l и луч AB, перпендикулярный AC (рис. 371). Рассмотрим некоторый луч AM, пересекающий прямую l в точке M. Меру угла CAM обозначим через α. Будем теперь неограниченно удалять точку M от точки C по прямой l (так что точка M будет пробегать положения M_1, M_2, ...). Угол α будет при этом возрастать, оставаясь все время меньше $\pi/2$. При неограниченном удалении точки M по прямой l луч будет все

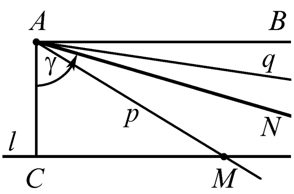


Рис. 372

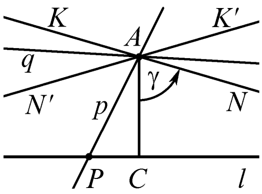


Рис. 373

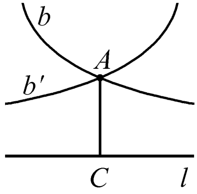


Рис. 374

более приближаться к некоторому предельному положению AN, причем логически могут представиться две возможности:

1) луч AN совпадает с лучом AB;

2) луч AN составляет с лучом AB некоторый острый угол.

Первый случай соответствует аксиоме параллельности: AB является *единственной* прямой, проходящей через точку A и не пересекающей l.

Посмотрим теперь, какое положение вещей мы имеем во втором случае. Обозначим меру острого угла CAN через γ (рис. 372). Луч AN *отделяет* лучи, пересекающие прямую l, от лучей, не пересекающих ее. Иначе говоря, всякий луч p, исходящий из точки A и проходящий внутри угла CAN, пересекает прямую l, а всякий луч q, проходящий внутри угла BAN, не пересекает ее. Сам же луч AN также с прямой l не пересекается. Если теперь продолжить луч AN за точку A и построить прямую AN', симметричную AN относительно AC, то получим при точке A две пары вертикальных углов (рис. 373). Всякая прямая, проходящая через точку A и расположенная внутри пары вертикальных углов NAN' и KAK', пересекается с l. Прямые же, проходящие через точку A и расположенные внутри пары вертикальных углов NAK', $N'AK$, не пересекаются с l.

Для того, чтобы доказать справедливость пятого постулата Евклида, нужно было бы установить, что случай, изображенный на рис. 373, невозможен.

Проводя доказательство «от противного», т. е. допуская, что имеет место случай, показанный на рис. 373, Лобачевский начал выводить различные следствия из этого допущения, надеясь, что рано или поздно он придет к противоречию, чем и завершится доказательство пятого постулата. Однако он доказал много десятков теорем, вытекающих из сделанного им допущения, не обнаружив при этом никаких логических противоречий. Доказанные Лобачевским теоремы были непривычны с точки зрения евклидовых представлений и вроде бы даже противоречили здравому смыслу. Тем не менее, все выводы были логически безупречны. И тогда Лобачевскому пришла в голову гениальная догадка: заменив аксиому параллельности ее отрицанием, т. е. предположив, что через точку $A \notin l$ можно провести более одной прямой, не пересекающей l, мы получаем *новую геометрию*, в корне отличающуюся от геометрии Евклида. В этой геометрии

(Лобачевский назвал ее «воображаемой») выполняются все аксиомы Евклида, кроме аксиомы параллельности, которая заменена ее отрицанием.

Приведем примеры теорем, которые имеют место в геометрии Лобачевского. Прежде всего заметим, что все теоремы, доказываемые в евклидовой геометрии *без использования* аксиомы параллельности, сохраняются и в геометрии Лобачевского. Например: вертикальные углы конгруэнтны; углы при основании равнобедренного треугольника конгруэнтны; из данной точки можно опустить на данную прямую только один перпендикуляр. Теоремы же евклидовой геометрии, при доказательстве которых применяется аксиома параллельности, в геометрии Лобачевского видоизменяются. Первой встретившейся теоремой, при доказательстве которой использовались свойства параллельных прямых, была теорема о сумме углов треугольника. Здесь ожидает нас первый «сюрприз»: в геометрии Лобачевского *сумма величин углов любого треугольника меньше* π.

Если два угла одного треугольника соответственно конгруэнтны двум углам другого, то в евклидовой геометрии конгруэнтны и третьи углы; как известно, такие треугольники подобны. В геометрии Лобачевского это уже не так. Более того, здесь вообще не существует подобных треугольников (не конгруэнтных между собой): если углы одного треугольника соответственно конгруэнтны углам другого, то в геометрии Лобачевского эти треугольники конгруэнтны.

Разность $\pi - (\angle A + \angle B + \angle C)$ называется дефектом треугольника ABC. Лобачевский доказал, что в его геометрии площадь треугольника пропорциональна дефекту, т. е. выражается формулой

$$S = k \cdot (\pi - (\angle A + \angle B + \angle C)), \tag{1}$$

где коэффициент k зависит от выбора единицы измерения площадей.

Термин «параллельные прямые» Лобачевский применяет не к любым непересекающимся прямым, а лишь к таким прямым, как l и b (или l и b') на рис. 374 (мы здесь изображаем прямые в виде искривленных линий, поскольку иначе с помощью наших евклидовых представлений не удается изобразить положение вещей, характерное для геометрии Лобачевского). Другие же пары непересекающихся прямых Лобачевский называет *расходящимися*.

Теорема о постоянстве расстояний между параллельными прямыми не сохраняется в геометрии Лобачевского: параллельные прямые неограниченно сближаются между собой в направлении параллельности и неограниченно удаляются друг от друга в противоположном направлении, т. е. расположение двух прямых на плоскости Лобачевского напоминает расположение двух искривленных линий (рис. 374).

Далее, две прямые в том и только в том случае являются расходящимися, если они имеют общий перпендикуляр (рис. 375). Заметим, что двух общих перпендикуляров две прямые в геометрии Лобачевского иметь не могут: ведь в геометрии Лобачевского сумма углов

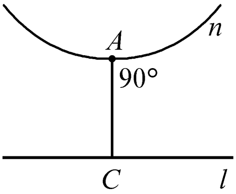


Рис. 375

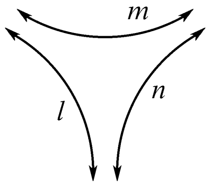


Рис. 376

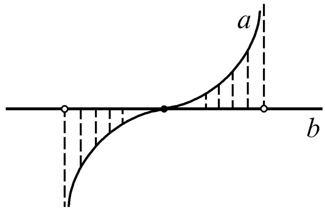


Рис. 377

четырехугольника меньше 2π, и потому не существует прямоугольников.

Интересно, что в геометрии Лобачевского существуют три прямые, попарно параллельные друг другу в различных направлениях (рис. 376). У треугольника, образованного этими прямыми, вершины как бы находятся в бесконечности, причем мера каждого угла равна 0. Отсюда следует, что дефект этого «треугольника» равен π и, следовательно, этот бесконечный «треугольник» имеет конечную площадь (см. формулу (1)).

И еще один факт из геометрии Лобачевского. Если прямые a и b пересекаются (но не перпендикулярны), то проекция прямой a на прямую b, т. е. множество оснований оснований перпендикуляров, проведенных к прямой b, представляет собой не всю прямую b (как в евклидовой геометрии), а лишь некоторый *интервал* на прямой b (рис. 377). Если же прямые a и b параллельны, то проекция прямой a на прямую b представляет собой *луч*.

Н. И. Лобачевский сделал сообщение об открытой им геометрии 11 февраля (по старому стилю) 1826 г. в своем докладе на заседании физико-математического отделения Казанского университета. Этот день и считается теперь датой рождений новой, «воображаемой» геометрии — геометрии Лобачевского.

64. Модель геометрии Лобачевского

Математический мир идей Лобачевского не воспринял. Ученые не были подготовлены даже к мысли о том, что может существовать какая-либо геометрия, отличная от евклидовой. Мужественно отстаивая правоту своих идей, Лобачевский опубликовал несколько книг и ряд статей, посвященных открытой им геометрии.

Лобачевский умер в 1856 г., так и не добившись признания своих идей.

Были, однако, два человека, которые не только поняли Лобачевского при его жизни, но и делят с ним заслугу открытия неевклидовой геометрии. Это были венгерский математик Янош Бойяи (1802 – 1860) и «король математики» Карл Фридрих Гаусс (1777 – 1855). Каждый из них, независимо от Лобачевского, пришел к тем же идеям. Бойяи написал к учебнику геометрии «добавление» («Appendix»), в котором он, несколько иначе и не столь полно, как это сделал Лобачевский, излагал идеи той же геометрии. Работа Бойяи вышла на несколько лет позже первой книги Лобачевского. Впоследствии Бойяи узнал о трудах Лобачевского и специально изучал русский язык, чтобы их прочитать, поскольку он опасался, что его обманывают, ссылаясь на какого-то русского, чтобы не признавать открытие, сделанное Бойяи.

Гаусс не опубликовал ни одной работы, относящейся к неевклидовой геометрии. То, что он владел идеями Лобачевского, было обнаружено лишь после смерти ученого при изучении его архива. Неподготовленность математического мира к восприятию этих идей была настолько велика, что даже гениальный Гаусс, к мнению которого прислушивались все, не рискнул опубликовать свои работы или выступить открыто в поддержку Лобачевского. Свое отношение к научному подвигу русского ученого он выразил тем, что добился избрания Лобачевского членом-корреспондентом Геттингенского королевского научного общества. Это была единственная научная почесть, которая выпала на долю Лобачевского при жизни.

Признание открытия Лобачевского пришло во второй половине XIX в., после появления работ итальянского математика Эудженио Бельтрами, английского математика Артура Кэли, немецкого математика Феликса Клейна, французского математика Анри Пуанкаре.

Лобачевский был абсолютно уверен, что сколько бы новых теорем ни доказывалось, никаких противоречий в его геометрии обнаружено не будет, но *доказательства* внутренней непротиворечивости своей геометрии он так и не представил. Такое доказательство было дано в работах названных выше ученых. Каждый из них построил *модель* геометрии Лобачевского. В частности модель, построенная в работах А. Кэли и Ф. Клейна, может быть описана следующим образом.

Рассмотрим в евклидовой плоскости круг K. В модели Кэли—Клейна рассматриваются только точки, лежащие внутри этого круга (а точки граничной окружности и точки, лежащие вне круга, не рассматриваются). Под «прямыми линиями» в этой модели понимаются всевозможные хорды круга K. В этой модели можно говорить о пересекающихся и непересекающихся прямых (на рис. 378 «прямые» a и b пересекаются, а a и c и также b и c не пересекаются). Непосредственно проверяется, что в рассматриваемой модели выполняется аксиома: через две точки можно провести прямую, и притом только одну (рис. 379). Расстояния между точками и величины углов опре-

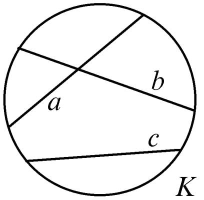


Рис. 378

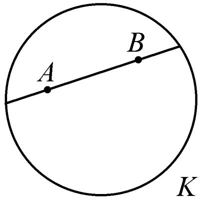


Рис. 379

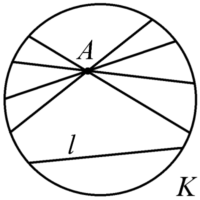


Рис. 380

деляются в модели Кэли—Клейна весьма сложными формулами. Но когда эти формулы были найдены, удалось проверить, что и все остальные аксиомы евклидовой геометрии, кроме аксиомы параллельности, в модели Кэли—Клейна также выполняются. Остается рассмотреть вопрос об аксиоме параллельности. Если бы оказалось, что в этой модели справедлива евклидова аксиома параллельности, то мы имели бы лишь еще одну модель все той же евклидовой геометрии. Если же в этой модели евклидова аксиома параллельности не выполняется, то перед нами — модель геометрии Лобачевского.

Легко видеть, что в рассматриваемой модели реализуется именно второй случай: через точку A, взятую вне l, проходит более одной «прямой», не пересекающейся с l (рис. 380). Таким образом, в модели Кэли—Клейна реализуется геометрия Лобачевского.

Как мы уже отмечали, длины отрезков и величины углов выражаются в модели Кэли—Клейна довольно сложными формулами. Поэтому хотя «прямая» изображается в модели Кэли—Клейна ограниченным отрезком, на ней можно откладывать отрезки любой «длины». Точно так же, хотя треугольники в модели Кэли—Клейна похожи на обычные евклидовы треугольники, сумма углов треугольника в этой модели, вычисленная по относящимся к этой модели формулам, оказывается меньше π.

Построение модели Кэли—Клейна полностью решило вопрос о непротиворечивости геометрии Лобачевского. В самом деле, каждая теорема геометрии Лобачевского интерпретируется в этой модели в виде некоторого утверждения, касающегося точек и хорд круга K. И если бы в геометрии Лобачевского существовало какое-либо внутреннее противоречие, это означало бы, что мы можем прийти к противоречию, рассуждая о точках и хордах круга K в обычной евклидовой плоскости. Поэтому если мы не сомневаемся в непротиворечивости геометрии Евклида, то мы обязаны признать, что и геометрия Лобачевского является столь же непротиворечивой.

Этим была полностью решена проблема пятого постулата: *евклидова аксиома параллельности не может быть доказана на основе остальных аксиом геометрии*. Действительно, все остальные аксиомы, кроме пятого постулата, выполняются как в евклидовой геометрии, так и в геометрии Лобачевского, но в одной из них аксиома параллельности имеет место, а в другой нет. Иными словами, аксиома параллельности (или пятый постулат Евклида) является *независимой* от остальных аксиом геометрии.

65. Изоморфизм моделей

Каждая модель представляет собой множество, на котором заданы один или несколько предикатов. Рассмотрим, например, понятие коммутативной группы. На множестве всех элементов коммутативной группы задан один предикат $p(a, b, c) \underset{def}{\equiv} (a + b = c)$. В группе $\boldsymbol{Z}$ целых чисел высказывание $p(2, 3, 5)$ истинно (т. е. $2 + 3 = 5$), а высказывание $p(3, 5, 7)$ ложно. И заданием этого предиката p (от трех переменных) операция сложения в группе полностью определяется.

В геометрии на плоскости рассматривается много предикатов. Например, обозначим через $\alpha(X, y)$ утверждение: «точка X принадлежит прямой y», т. е. $\alpha(X, y) \underset{def}{\equiv} (X \in y)$. Это — предикат от двух переменных; для прямых и точек, изображенных на рис. 381, мы можем с помощью этого предиката (подставляя те или иные точки и прямые) получить ряд высказываний: $\alpha(A, l)$ — истинно; $\alpha(A, m)$ — истинно; $\alpha(A, p)$ — ложно; $\alpha(B, l)$ — истинно; $\alpha(B, m)$ — ложно и т. д.

Еще один геометрический предикат β задается утверждением «прямые y_1 и y_2 имеют общую точку» (т. е. их пересечение непусто): $\beta(y_1, y_2) \underset{def}{\equiv} (y_1 \cap y_2 \neq \varnothing)$. Для прямых на рисунке 381 высказывание $\beta(l, m)$ истинно, а высказывание $\beta(l, p)$ ложно.

Можно также рассмотреть геометрический предикат «точка C расположена между A и B» (т. е. точки A, B, C лежат на одной прямой и при этом точка C принадлежит отрезку AB). Обозначим этот предикат через $\gamma(A, B, C)$. Например, на рис. 382 высказывание $\gamma(A, M, P)$ истинно (т. е. P лежит между A и M), а высказывания $\gamma(A, P, M)$ и $\gamma(B, M, P)$ ложны.

Между прочим, шесть прямых и семь точек, показанные на рис. 382, представляют собой модель, в которой выполняется аксиома «из трех точек, лежащих на одной прямой, одна и только одна расположена между двумя другими». Однако в этой модели аксиома «через каждые две точки проходит прямая» не выполняется (например, через точки M и N в этой модели не проходит прямая). Иначе говоря, если мы обозначим через $\sigma(X, Y)$ предикат «через точки X, Y проходит прямая», то высказывание $(\forall\ X, Y)\ \sigma(X, Y)$ в рассматриваемой модели ложно. В обычной же (евклидовой) геометрии высказывание $(\forall\ X, Y)\ \sigma(X, Y)$ истинно (т. е. через любые две точки X, Y проходит прямая).

Добавим, однако, к модели, изображенной на рис. 382, еще одну линию — окружность, проходящую через точки M, N, Q (рис. 383) и условимся эту окружность также считать «прямой». В получающейся модели высказывание $(\forall\ X, Y)\ \sigma(X, Y)$ уже будет истинным.

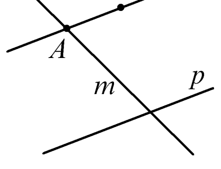


Рис. 381

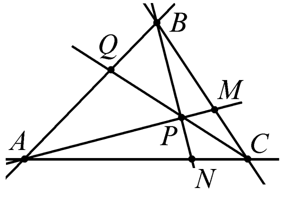


Рис. 382

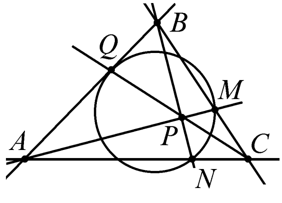


Рис. 383

Рассмотренные примеры позволяют дать общее описание весьма важного понятия *изоморфизма* моделей. Условимся записью $<A; \alpha, \beta, \gamma>$ выражать тот факт, что рассматривается модель, представляющая собой множество A, на котором заданы предикаты α, β, γ. Две модели $<A; \alpha, \beta, \gamma, ...>$ и $<A'; \alpha', \beta', \gamma', ...>$ условимся называть *подобными*, если, во-первых, число предикатов в первой модели *равно* числу предикатов во второй модели, и, во-вторых, предикат α зависит от стольких же переменных, что и предикат α', предикат β зависит от стольких же переменных, что и предикат β' и т. д.

Например, модели, изображенные на рис. 382 и 383, подобны, если в каждой из них рассматриваются три предиката:

$$\rho(X) \underset{def}{\equiv} \{X \text{ есть точка}\};$$

$$\tau(x) \underset{def}{\equiv} \{x \text{ есть прямая}\};$$

$$\sigma(X, Y) \underset{def}{\equiv} \{\text{через точки } X \text{ и } Y \text{ проходит прямая}\}.$$

При этом модель на рис. 382 задана на множестве, состоящем из 13 элементов (7 точек и 6 прямых), а модель на рис. 383 задана на множестве, содержащем 14 элементов (7 точек и 7 прямых).

Две подобные модели $<A; \rho, \tau, \sigma, ...>$ и $<A'; \rho', \tau', \sigma', ...>$ называются *изоморфными*, если можно установить такое взаимно однозначное соответствие $f: A \to A'$ между множествами A и A', что при отображении f, так же как и при обратном отображении f^{-1}, сохраняется истинность высказываний. Иначе говоря, высказывание $\rho(a)$ истинно тогда и только тогда, когда истинно $\rho'(f(a))$, высказывание $\sigma(a, b)$ истинно тогда и только тогда, когда истинно $\sigma'(f(a), f(b))$ и т. д.

Например, модели на рис. 382 и 383 не изоморфны, поскольку первая из них задана на множестве A, содержащем 13 элементов, а вторая — на множестве A', содержащем 14 элементов, так что установить между A и A' взаимно однозначное соответствие не удастся. Но если бы даже мы не обратили внимание на число элементов в множествах A и A', мы могли бы указать и другую причину, по которой эти модели не изоморфны: в первой из этих моделей высказывание $(\forall \text{ т. } X, Y)\, \sigma(X, Y)$ ложно, тогда как во второй модели оно истинно.

А теперь приведем примеры изоморфных моделей. Обозначим через $\boldsymbol{R}$ модель, состоящую из всех действительных чисел, для которых мы рассмотрим только один предикат $p(a, b, c) \underset{def}{\equiv} (a + b = c)$.

Далее, через $\boldsymbol{R'}$ обозначим множество всех положительных действительных чисел, для которых мы рассмотрим только один предикат $p'(a, b, c) \underset{def}{\equiv} (ab = c)$.

Например, высказывания $p(1, -5, -4)$, $p'(\sqrt{2}, \sqrt{3}, \sqrt{6})$ истинны, а высказывания $p(1, 2, 4)$, $p'(2, 2, 5)$ ложны.

Эти две модели $<\boldsymbol{R}; p>$ и $<\boldsymbol{R}'; p'>$ подобны: в каждой из них рассматривается множество, на котором задан только один предикат, зависящий от трех переменных. Более того, эти две модели изоморфны: если мы обозначим через $f\colon \boldsymbol{R} \to \boldsymbol{R}'$ отображение, выражаемое формулой $f(x) = 2^x$, то получаем взаимно однозначное соответствие, при котором сохраняется истинность.

В самом деле, если $p(a, b, c)$ истинно, т. е. $a + b = c$, то $2^a \cdot 2^b = 2^c$ (поскольку при перемножении степеней показатели складываются). Это означает, что $f(a) \cdot f(b) = f(c)$, т. е. высказывание $p'(f(a), f(b), f(c))$ истинно. Аналогично, если $p'(f(a), f(b), f(c))$ истинно, то и $p(a, b, c)$ истинно. Итак, $p(a, b, c) \Leftrightarrow p'(f(a), f(b), f(c))$, а это и означает, что рассматриваемое отображение f и обратное отображение f^{-1} сохраняют истинность. Иначе говоря, *аддитивная группа действительных чисел* (т. е. множество всех действительных чисел с единственной рассматриваемой операцией — сложением) *изоморфна мультипликативной группе положительных действительных чисел* (с единственной рассматриваемой операцией — умножением).

Еще один пример изоморфизма мы фактически рассмотрели в п. 59: симметрическая группа S_3 (т. е. группа всех подстановок из трех элементов) изоморфна группе всех самосовмещений равностороннего треугольника. Там ты говорили, что эти группы «по существу совпадают», но математически более корректно говорить, что эти две группы *изоморфны*.

66. Полнота аксиоматики

Применим теперь понятие изоморфизма моделей к рассмотрению аксиоматических теорий. Система аксиом называется *полной*, если любые две модели, в которых все эти аксиомы выполнены, являются изоморфными.

Аксиоматика коммутативной группы полной не является. В самом деле, целые числа (с единственной операцией — сложением) дают модель коммутативной группы. Действительные числа (тоже рассматриваемые только с одной операцией — сложением) — еще одна модель коммутативной группы. Однако в этих моделях *разное* число элементов (первая модель счетная, а вторая несчетная), и потому эти модели *неизоморфны*. Имеются также *конечные* коммутативные группы: например, множество из двух элементов (чет—нечет) с таблицей сложения

+	**Ч**	**Н**
Ч	**Ч**	**Н**
Н	**Н**	**Ч**

(т. е. Ч + Ч = Ч, Н + Н = Ч, Ч + Н = Н, Н + Ч = Н) представляет собой, как нетрудно проверить, коммутативную группу. Эта группа содержит только два элемента, и потому она не изоморфна ни аддитивной группе целых чисел, ни аддитивной группе действительных чисел.

Итак, аксиоматика коммутативной группы является неполной. Легко видеть, что аксиоматика метрического пространства также не является полной. Существуют метрические пространства, состоящие из конечного числа точек. Например, определив в множестве $\boldsymbol{M} = \{a, b, c\}$ расстояния соотношениями $d(a, b) = d(a, c) = d(b, c) = 1$ (и, разумеется, считая, что расстояние между двумя совпадающими точками равно нулю), мы получаем метрическое пространство, содержащее лишь три точки (проверка аксиом метрического пространства не представляет труда). Как мы видели в п. 58, вся плоскость представляет собой метрическое пространство. Прямая — тоже метрическое пространство. Все эти модели метрического пространства не изоморфны между собой, и это показывает, что система аксиом метрического пространства не является полной.

В п. 57 мы говорили, что список аксиом геометрии, предложенный Евклидом, был неполным. Тогда это утверждение имело интуитивно понятный смысл. Но теперь мы можем придать этому утверждению точный математический смысл. Именно, *система аксиом Гильберта является полной*, т. е. любые две модели геометрии, построенные с учетом всех аксиом Гильберта, *изоморфны*. Если же мы возьмем лишь часть из этих аксиом, то можем получить неполную аксиоматику. В частности, отбросив *аксиому измерения*, предложенную Архимедом, мы получаем систему аксиом, допускающую в качестве моделей не только евклидову геометрию, но и различные *неархимедовы геометрии*. Отсутствие аксиом расположения (включающих описание термина «точка лежит между двумя другими», а также аксиомы Паша, о которой речь шла в п. 62) также приводит к существованию моделей, не изоморфных евклидовой. К их числу относятся, например, *проективные* геометрии.

Наконец, отбрасывание пятого постулата (аксиомы параллельности) дает весьма интересную систему аксиом, которая описывает так называемую *абсолютную геометрию*. К ней относятся все теоремы геометрии, доказываемые без использования аксиомы параллельности: свойства равнобедренного треугольника, свойства перпендикуляров и наклонных, упоминавшуюся в п. 53 теорему о том, что диаметр, перпендикулярный хорде, делит ее пополам, и т. д. Аксиоматика абсолютной геометрии *неполна*. В самом деле, геометрия Евклида и геометрия Лобачевского представляют собой две неизоморфные модели абсолютной геометрии, т. е. в обеих этих геометриях все положения абсолютной геометрии имеют место. Можно сказать, что абсолютная геометрия — это и есть все то общее, что имеется в геометриях Евклида и Лобачевского.

Заметим в заключение, что полные и неполные аксиоматики имеют принципиально различную значимость. Гильбертова аксиоматика геометрии полна. Она описывает *только одну*, с точностью до изоморфности, модель, т. е. дает построение одной ветви знания. Смысл ее в том, что эта ветвь знания (в данном случае евклидова геометрия) получает свое до конца аксиоматизированное изложение, до конца выясненную логическую структуру. И как только аксиоматизация (в этом смысле, т. е. вплоть до получения *полной* системы аксиом) закончена, эта ветвь знания близка к своему полному завершению. Иными словами, если в этой науке (евклидовой геометрии) и возникают новые идеи, методы, понятия, то они выделяются в отдельную новую ветвь науки. Примерами могут служить топология, теория графов, дифференциальная геометрия, теория выпуклости, дискретная (или, иначе, комбинаторная) геометрия и другие ветви геометрии. Но в своем типичном понимании (в духе элементарной геометрии) рамки евклидовой геометрии с завершением ее полной аксиоматизации закрываются. И если появляются (трудами учителей и любителей математики) новые теоремы, относящиеся именно к евклидовой геометрии, они не делают погоды. Наука эта завершена. Обычно так и бывает: завершенная, в основном, в своем развитии наука получает в качестве венца свою полную аксиоматику. Кроме евклидовой геометрии, другим примером является теория действительного числа.

Совершенно иную роль играют неполные аксиоматики, и об этом мы уже частично говорили в п. 58 в связи с метрическими пространствами. Неполная аксиоматика имеет неизоморфные модели, имеет различные реализации в других областях знания. Например, теория групп применяется в ядерной физике, кристаллографии, математической экономике и других науках, не говоря уже о многих разделах самой математики. Понятия и факты теории групп, один раз разработанные, многократно живя новой жизнью в других разделах знания, являются важными кирпичами и блоками, используемыми при возведении зданий других наук.

И еще одно замечание о математической логике. Мы впоследствии будем иметь случай говорить о системе аксиом Цермело—Френкеля, служащей ее базой. Эта система аксиом *неполна*. И, подобно тому как, добавляя к аксиомам абсолютной геометрии либо евклидову аксиому параллельности, либо ее отрицание (т. е. аксиому Лобачевского), мы получаем *разные* геометрии, так и в математической логике добавление к аксиоматике Цермело—Френкеля новых положений дает «разные логики». В XX веке появилась замечательная работа Поля Коэна, который установил, что можно добавить к аксиомам Цермело—Френкеля либо аксиому выбора, либо ее отрицание, и в обоих случаях мы получим две «логики», которые в *одинаковой степени* непротиворечивы. Точнее, хотя их непротиворечивость до конца не установлена, но если принять непротиворечивость одной, то непротиворечива и другая, и обратно. Это в какой-то мере напоминает положение дел в геометрии. Аксиоматика абсолютной геометрии неполна. Добавляя к этой геометрии либо аксиому параллельности, либо ее отрицание, мы получаем две различные геометрии (Евклида или Лобачевского), которые в одинаковой степени непротиворечивы, т. е. если принять непротиворечивость одной, то непротиворечива и другая.

Глава IV. ПОИСК РЕШЕНИЙ

Беседа 14. Инсайт

67. Цикл озарения

Математическая логика вовсе не отвечает на вопрос, как мы мыслим при решении задач. По существу, математическая логика позволяет оформить *уже найденное* решение в таком виде, что мы можем убедить *другого человека* в правильности решения: шаг за шагом рассуждение проверяется на его строгость, логическую безупречность, и в конце доказательства получается подтверждение установленного факта. Но когда мы *ищем* еще не известное нам решение, ищем путь решения, мы, как правило, мыслим *иначе*. Как именно? На это должна ответить еще не созданная «математическая психология».

Условимся считать, что задания делятся на *упражнения* и *задачи*. Непременным условием того, чтобы задание можно было считать упражнением, является указание (явное или неявное) о способе его выполнения. Решающий должен быть осведомлен (перед выполнением упражнения или группы упражнений), хотя бы в общих чертах, о характере той деятельности, которую он должен выполнить, — иначе это не будет упражнением.

Иногда указание о способе выполнения включается в текст упражнения. Например: «Применяя формулу квадрата суммы или разности двух чисел, вычислите 21^2, 18^2, 32^2, 98^2». Здесь указание о применении формулы $(a \pm b)^2 = a^2 \pm 2ab + b^2$ сформулировано явно, чтобы избежать типичной ошибки: для вычисления, например 32^2, производится умножение $32 \cdot 32$ «в столбик», что представляет собой (при изучении алгебры) пустую трату времени и сил. В большинстве же случаев указание о способе выполнения в текст упражнения не включается. Однако тот факт, что это упражнение дается после того, как на предыдущих уроках был изучен определенный теоретический материал, является, как правило, неявным указанием на необходимость применить при выполнении упражнения именно этот материал.

Но помимо знания правил, необходимых для выполнения упражнения, обучаемый должен обладать определенной культурой, характеризующей уровень уже имеющихся у него знаний. Итак, определенный уровень знаний плюс четко очерченный круг применяемых правил (законов) — вот что характеризует упражнение. Как правило, при выполнении упражнения применение этих законов является стандартным, типичным.

Заметим, что обычно выполнение упражнения вовсе не сводится к одному мыслительному акту, а представляет собой определенную деятельность. Так, при нахождении числа 32^2 требуется, во-первых, представить себе, что число 32 надо записать в виде суммы или разности. Во-вторых, надо фактически осуществить такую запись: $32 = 30 + 2$. В-третьих, надо уточнить, что здесь *сумма* (а не разность), так что надо применить именно формулу *квадрата суммы*. В-четвертых, надо подставить в эту формулу конкретные значения: $32^2 = 30^2 + 2 \cdot 30 \cdot 2 + 2^2$. Наконец, в-пятых, надо произвести подсчет в правой части последнего равенства, чтобы получить окончательный ответ: $32^2 = 1024$. Вначале учащийся выполняет все эти операции обособленно, сознательно контролируя каждую из них. Однако после нескольких однотипных упражнений как бы «сразу» выполняются лишь двух последние из описанных выше операций. В действительности же это лишь означает, что навык в выполнении упражнений такого типа приобретен, в результате чего первые три операции выполняются *свернуто*, *автоматизированно*.

В отличие от упражнения *задача* носит более творческий характер. Разумеется, различие между упражнениями и задачами условно и в значительной степени субъективно. Вопрос о том, «стандартный» или «творческий» характер имеет определенное задание, зависит от склонностей, знаний, уровня общей культуры человека, выполняющего его. Более того, один и тот же человек после решения достаточного числа сходных между собой задач будет склонен относить их к разряду упражнений, так как они утратят для него творческий характер. И все же, несмотря на условность и относительность деления заданий на упражнения и задачи, между первыми и вторыми есть существенная разница. Решение задачи расчленяется на несколько отдельных этапов, проведение каждого из которых сводится к выполнению определенного упражнения, существенно более простого, чем вся задача в целом. Иными словами, для решения задачи нужно сформулировать и выполнить несколько упражнений в определенной (заранее не указанной) последовательности. При этом трудность решения задачи заключается не только (и не столько) в том, что надо выполнить несколько упражнений; основная трудность состоит в отыскании необходимой *последовательности* тех упражнений, выполнение которых ведет к решению задачи. Даже если выполнение упражнений, которые могут понадобиться, доведено до стадии навыка, вопрос о том, какие упражнения и в какой последовательности (подчас совершенно неожиданной) нужно выполнить для решения задачи, остается основным и нетривиальным.

Каждый из нас многократно наблюдал такую картину. Решающий долго и напряженно думает над задачей (скажем, по физике). Вдруг вспышка, озарение, *инсайт* (*insight* в переводе с английского как раз и означает «озарение»). Откуда-то в сознании появляется путь решения задачи, «видна» та последовательность упражнений, которую надо выполнить, чтобы получить решение задачи. Что же произошло, что означает это озарение? Попытаемся ответить на этот вопрос. Точнее, попытаемся построить математическую модель озарения. Разумеется, озарение — сложный мыслительный процесс. Предлагаемая модель (как, впрочем, и любая другая) огрубляет его, приближенно имитируя основные черты. Однако применение этой модели будет полезно при решении задач. Для пояснения предлагаемой модели озарения рассмотрим следующую задачу.

Задача 1. Доказать тождество

$$\cos 20° \cos 40° \cos 80° = \frac{1}{8}. \tag{1}$$

Замечаем, что 40° — это *двойной угол* по отношению к 20°. В связи с этим вспоминаем, что имеются специальные формулы синуса и косинуса двойного угла. Применить формулы для синуса двойного угла $\sin 2\alpha = 2\sin\alpha\cos\alpha$ можно было бы, если *домножить* выражение на $\sin 20°$ (так как $\cos 20°$ там уже есть). А чтобы выражение не изменилось, надо не только умножить, но и разделить на $\sin 20°$. Таким образом, у нас в сознании осуществился переход от первоначально поставленной задачи (1) к новой ее формулировке:

$$\frac{\sin 20° \cos 20° \cos 40° \cos 80°}{\sin 20°} = \frac{1}{8}. \tag{2}$$

Теперь, в соответствии с замыслом, можно применить формулу синуса двойного угла, т. е. заменить $\sin 20° \cos 20°$ на $\frac{1}{2}\sin 40°$. Итак, надо доказать, что

$$\frac{\sin 40° \cos 40° \cos 80°}{2\sin 20°} = \frac{1}{8}. \tag{3}$$

Но теперь мы видим, что можно еще раз применить формулу синуса двойного угла, т. е. переписать доказываемое равенство в виде

$$\frac{\sin 80° \cos 80°}{4\sin 20°} = \frac{1}{8}. \tag{4}$$

И опять можно уже в третий раз применить формулу синуса двойного угла, хотя мы с опаской замечаем, что величины углов увеличиваются. Все же двигаемся дальше, т. е. переписываем доказываемое равенство в виде

$$\frac{\sin 160°}{8\sin 20°} = \frac{1}{8}. \tag{5}$$

До сих пор хотя мы и рассуждали в нужном для решения задачи направлении, но не могли еще сказать, что он уже нашли решение. И вдруг — эврика! — в числителе и знаменателе сумма углов равна 180°, и потому $\sin 160° = \sin(180° - 160°) = \sin 20°$, что и завершает доказательство требуемого равенства.

Структурная схема рассуждений показана на рис. 384. Здесь заштрихованный прямоугольник обозначает исходный уровень знаний, а кружками обозначены формулировки, указанные в проведенном выше рассуждении. Пунктирная стрелка ведет к постановке задачи, тонкие сплошные стрелки показывают логические связи, постепенно

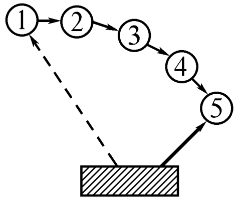


Рис. 384

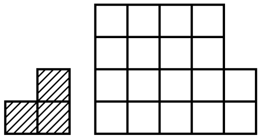


Рис. 385

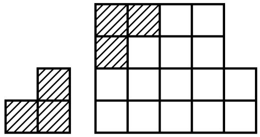


Рис. 386

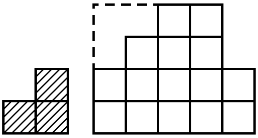


Рис. 387

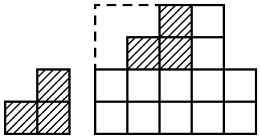


Рис. 388

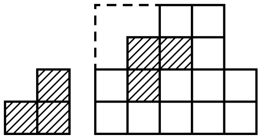


Рис. 389

накапливаемые в процессе рассуждения, жирная стрелка символизирует озарение, инсайт. Она замыкает цикл: именно замыкание цикла на структурной схеме, изображающей ход рассуждения, характеризует озарение.

Рассмотрим еще несколько примеров.

Задача 2. Замостить фигуру на рис. 385, содержащую 18 квадратиков, с помощью «уголков».

Решение осуществляется в несколько приемов. Например, уложив один кирпичик («прикинув», как он может располагаться в целой фигуре, рис. 386), мы уже осуществляем один шаг, как бы заменив исходную задачу другой, более простой, в которой требуется выложить кирпичиками меньшую фигуру (рис. 387). Тем самым наша задача преобразована в другую, тесно связанную с первоначальной, но все же другую. Теперь у нас меньший простор, меньшая по площади фигура и мы делаем выводы о путях дальнейшего решения.

Осуществляя такие шаги, постепенно можно прийти к отысканию решения, последовательно продвигаясь по более и более простым фигуркам к игровой ситуации, уже допускающей очевидное разрешение, либо возвращаясь назад (полностью или частично), если некоторая подпоследовательность шагов заводит попытку решения в тупик (рис. 388), и продвигаясь вперед другими шагами (рис. 389).

Схематически процесс решения можно изобразить в виде графа (дерева, рис. 390), в котором кружки изображают последовательно перебираемые игровые ситуации, каждая стрелка символизирует некоторый шаг рассуждения, нижний прямоугольник задает условие задачи, а штриховая линия ведет к искомой ситуации, т. е. изображает желаемое решение задачи. На рис. 390 показана не только последовательность стрелок, прямо ведущая к решению, но также побочные пути, которые заводят игрока в тупиковые ситуации. А жирная стрелка — инсайт, озарение. Она схематизирует заключительный шаг

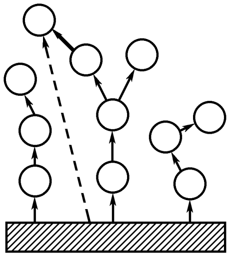


Рис. 390

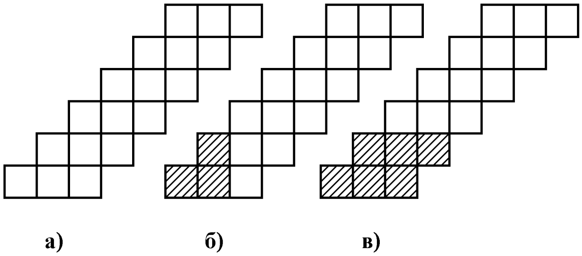


Рис. 391

системы пошаговых действий («Ага, эврика! Теперь я вижу, что на этом пути решение получится»).

Задача 3. Заполнить уголками фигуру на рис. 391а.

В результате различных попыток укладывания уголка в этой фигуре мы замечаем, что способ укладывания, показанный на рис. 391б дает *единственную* возможность заполнить левый нижний квадратик заданной фигуры. После этого имеется *единственный* способ заполнить нижний квадратик оставшейся части фигуры (рис. 391в). Оставшаяся фигура очень похожа на исходную, но стала меньше. А теперь происходит инсайт, позволяющий постепенно прийти к обобщению: если дана ступенчатая фигура такого вида, как на рис. 391а, причем у нее слева четное число «зубцов» (на рис. 391а слева шесть зубцов), то ее удается заполнить уголками. Если же у этой фигуры слева нечетное число зубцов, то ее не удастся заполнить уголками.

68. Сфера достижимости

Предложенная выше математическая модель озарения требует для своего более полного осуществления введения понятия «сферы достижимости». Мы опишем это понятие на примере следующей арифметической задачи.

Задача 4. По рации сообщили, что в далеком селе лежит тяжелобольной геолог, нуждающийся в срочном лечении. Автомашина со специализированной аппаратурой может двигаться лишь со скоростью 60 км/ч. Поэтому одновременно с ней выслали машину скорой помощи (СП) (135 км/ч), которая сможет доехать до села за 2 ч 40 мин, заберет больного и сразу же поедет обратно на встречу со спецмашиной. Через сколько времени больному будет оказана помощь?

С помощью имеющихся в задаче данных можно:

А. Определить расстояние до села (для этого надо скорость СП 135 км/ч умножить на время $2\frac{2}{3}$ ч).

Б. Найти разность скоростей машин, т. е. узнать, на сколько отстает за час спецмашина от машины СП.

В. Узнать, во сколько раз быстрее едет машина СП по сравнению со спецмашиной.

Г. Найти сумму скоростей машин, т. е. узнать, на сколько километров за час будут сближаться машины, когда они поедут навстречу друг другу.

Д. Узнать расстояние, пройденное спецмашиной к моменту, когда машина СП прибудет в село.

Е. Узнать, какую часть пути (от больницы до села) проезжает машина СП за час.

Каждое из этих данных можно получить в одно арифметическое действие, исходя из имеющихся данных (первый ярус, рис. 392).

С помощью полученных данных можно:

Ж. Определить (учитывая **А**), сколько времени нужно спецмашине, чтобы доехать до села.

З. Узнать (**Б**), на сколько километров отстанет спецмашина к тому времени, когда машина СП приедет в село (то же можно узнать из **А** и **Д**).

И. Узнать (**В**), какую часть всего пути проезжает спецмашина за час.

К. Узнать (**А** и **Д**), какую часть всего пути проехала спецмашина к моменту, когда машина СП приехала в село.

Л. Узнать (**А** и **Г**), на какую часть всего пути сближаются машины за час при движении навстречу друг другу.

Все это составляет второй ярус. Теперь (третий ярус, см. рис. 392) можно:

М. Узнать (**З** и **Г**), через сколько времени после выезда машины СП из села обе машины встретятся.

Н. Узнать (**К** или **А** и **З**), какую часть всего пути составляло расстояние между машинами в момент, когда машина СП прибыла в село.

О. Узнать (**З**), сколько еще времени (после приезда машины СП в село) потребуется спецмашине, чтобы доехать до села.

Аналогично строятся четвертый и следующие ярусы:

П. Учитывая **Н** и **Л**, можно вновь найти **М**.

Р. Можно установить (**М** или **П**), через сколько времени после выезда машин из больницы они встретятся.

С. Можно узнать (**Р**), на каком расстоянии от больницы встретились машины.

Т. Можно определить (**А** и **С** или **М**), на каком расстоянии от села встретились машины, и т. д.

Получается постепенно растущая область («сфера достижимости»), которая сначала захватывает первый ярус, затем второй, третий и т. д. Путь решения задачи можно считать найденным, когда постепенно растущая сфера достижимости «поглотит» точку, изображающую то, что спрашивается в задаче — в данном случае **Р**. К искомой точке можно прийти различными путями:

Б → З → Г → М → Р;

А → Д → З → Г → М → Р;

А → Д → К → Н → Г → Л → П → Р;

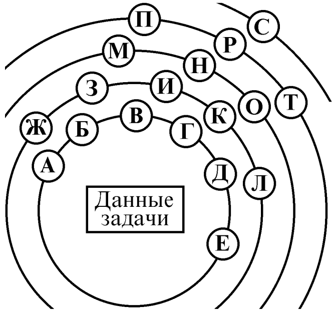


Рис. 392

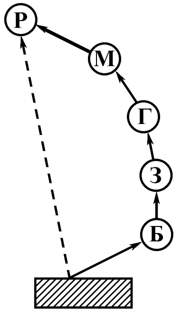


Рис. 393

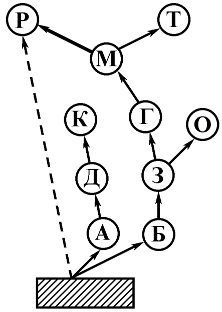


Рис. 394

А → Д → З → Н → Г → Л → П → Р;
Б → З → А → Н → Г → Л → П → Р.

Каждый из этих путей схематически изображает путь решения задачи. Например, полное решение (с вычислениями), следующее первому пути, можно изложить так.

За час спецмашина отстает на $135 - 60 = 75$ (км). Значит, за 2 ч 40 мин она отстанет на $75 \cdot 2\frac{2}{3} = 200$ (км), т. е. в момент выезда машины СП из села расстояние между машинами равно 200 км. Но за каждый час машины будут сближаться на $135 + 60 = 195$ (км). Следовательно, для встречи им понадобится 200:195 (ч), т. е. примерно 1 ч 2 мин. Таким образом, помощь больному будет оказана через 2 ч 40 мин + 1 ч 2 мин = 3 ч 42 мин.

Вернемся к вопросу об озарении. На первый взгляд кажется, что первому пути решения задачи соответствует схема на рис. 393 (жирная стрелка — озарение: «Теперь я и смогу узнать то, что требуется!»). Однако эта схема неполна. Возможно, схема рассуждения будет разветвленной (рис. 394), содержащей боковые ветви (ложные попытки, уводящие в сторону и постепенно отбрасываемые).

На рис. 394 видно, что жирная стрелка, символизирующая озарение, замыкает цикл, т. е. завершает замкнутый путь схематически изображенных мыслительных актов. До этого решающий делает шаги, не зная точно, по верному или ложному пути он следует (т. е. не может еще сказать, что он видит путь решения).

Еще одну иллюстрацию понятия сферы достижимости дает старинная головоломка о переливании жидкостей.

З а д а ч а 5. Бак емкостью 8 литров доверху наполнен маслом (рис. 395). Имеются, кроме того, два пустых бака, имеющих емкости 5 литров и 3 литра. Как можно, пользуясь этими баками, отмерить точно 4 литра масла (рис. 396)?

Решение получается следующим образом. Сначала мы можем, например, заполнить пятилитровый бак. В самом большом остается тогда $8 - 5 = 3$ литра масла. А можно поступить и иначе: полностью

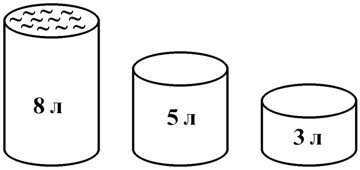


Рис. 395

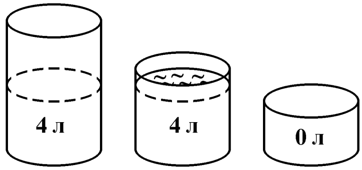


Рис. 396

залить маленький бак (тогда в большем останется 5 л). Таким образом, начальная позиция головоломки может быть обозначена записью: **А** = (8; 0; 0), а после переливания получится одна из записей

(1) **Б** = (3; 5; 0), **В** = (5; 0; 3).

Это — первый ярус сферы достижимости.

Теперь, имея игровую позицию **Б**, мы можем из второго бака перелить масло в третий (самый меньший), наполняя его полностью. Во втором баке останется $5 - 3 = 2$ литра. Получается запись **Г** = (3; 2; 3). Аналогично можно найти и другие позиции, получаемые из **Б**, **В**. Они составляют второй ярус:

(2) **Г** = (3; 2; 3), **Д** = (0; 5; 3), **Е** = (5; 3; 0).

Далее мы можем достичь двух новых позиций, исходя из позиции **Г** (сливая масло из меньшего бака в больший) или **Е**:

(3) **Ж** = (6; 2; 0), **З** = (2; 3; 3).

Четвертый ярус содержит две игровые позиции:

(4) **И** = (6; 0; 2), **К** = (2; 5; 1),

а пятый — позиции

(5) **Л** = (1; 5; 2), **М** = (7; 0; 1).

Дальнейшие переливания дают шестой ярус:

(6) **Н** = (1; 4; 3), **О** = (7; 1; 0),

а затем седьмой:

(7) **П** = (4; 4; 0), **Р** = (4; 1; 3).

Растущая сфера достижимости дошла до требуемой позиции **П**, это и дает решение задачи (рис. 396). Решение можно, таким образом, получить двумя путями:

А → **Б** → **Г** → **Ж** → **И** → **Л** → **Н** → **П**;

или

А → **В** → **Е** → **З** → **К** → **М** → **О** → **Р** → **П**.

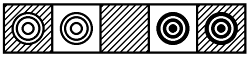


Рис. 397

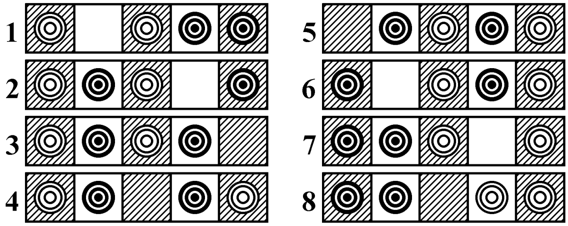


Рис. 398

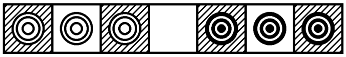


Рис. 399

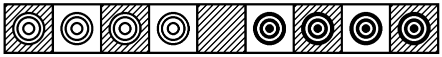


Рис. 400

Аналогичным образом решается задача в случае, если имеется бак вместимостью 12 литров, наполненный маслом, и два пустых бака 7 и 5 литров (или, вообще, бак вместимостью $4n$ литров, наполненный маслом, и два пустых бака вместимостью $2n+1$ и $2n-1$ литров). Возможны и другие задания.

Интересна головоломка, известная под названием «Игра Люка». Она также может быть проиллюстрирована понятием сферы достижимости. В самом простом варианте она формулируется так.

З а д а ч а 6. Имеется пятиклеточная доска, на которой стоят 2 белых шашки и 2 черных (рис. 397). За один ход разрешается либо передвинуть шашку на свободное соседнее поле, либо перепрыгнуть через одну шашку на свободное поле (перепрыгивать можно через шашку любого цвета, причем друг друга шашки не «едят»). Сколько ходов нужно, чтобы поменять местами белые и черные шашки?

Решение можно осуществить за 8 ходов (рис. 398), т. е. на восьмом ярусе сферы достижимости.

Можно предложить более сложный вариант игры с 6 шашками (рис. 399, 15 ходов) или еще более сложный с 8 шашками (рис. 400). Маленькое указание для нахождения наиболее коротких решений: не следует прыгать через шашку того же цвета.

С головоломкой «Ханойская башня» связана старинная буддийская легенда. В старинном храме укреплены три стержня, инкрустированные бриллиантами. На первом стержне нанизаны 64 золотых диска — внизу самый большой, а следующие — все меньше и меньше. Жрецы храма должны без устали перекладывать диски с одного стержня на другой по определенным правилам. И когда, следуя этим правилам, они перенесут все диски с первого стержня на третий, наступит конец света.

А правила, которым должны следовать жрецы, заключаются в следующем. За один прием разрешается перекладывать только один диск (т. е. сразу два или большее число дисков перекладывать не

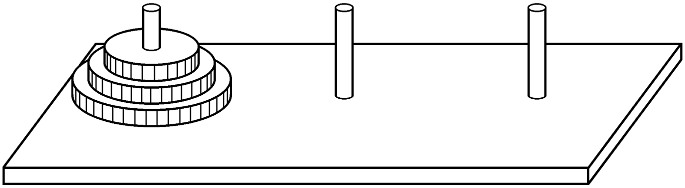


Рис. 401

разрешается); кроме того, не разрешается класть больший диск на меньший.

Задача 7. За какое наименьшее число ходов можно в «Ханойской башне» перенести все диски с первого стержня на третий?

Мы наметим решение «Ханойской башни» из трех дисков (рис. 401). Так как за один прием разрешается перекладывать только один диск, то можно с первого стержня снять только один (верхний) диск, и его можно поместить на второй или на третий стержень. Если верхний диск снят с первого стержня и надет на третий, то следующим ходом можно снять еще один диск с первого стержня, но надеть его на третий стержень не разрешается (нельзя класть больший диск на меньший); поэтому его можно надеть только на второй стержень. Затем можно маленький диск с третьего стержня перенести на другой стержень и т. д. Постепенно наращивается сфера достижимости, причем решение получается на ее седьмом ярусе, т. е. для переноса всех трех дисков с первого стержня на третий потребуется 7 ходов.

В «Ханойской башне» с четырьмя дисками для переноса всех четырех дисков с первого стержня на третий требуется уже 15 ходов. Далее, для башни с пятью дисками требуется 31 ход, а для башни с шестью дисками — 63 хода, и вообще для башни с n дисками требуется $2^n - 1$ ход. Таким образом, количество ходов очень быстро возрастает с увеличением числа дисков. Так, если каждую секунду перекладывать один диск и работать круглые сутки, то, чтобы решить головоломку «Ханойская башня» с двадцатью дисками, требуется более двенадцати дней, при тридцати дисках требуется более тридцати четырех лет, а тридцать пять дисков требуют уже более тысячи лет. Что же касается «Ханойской башни» с 64 дисками, то скорее погаснет Солнце, чем жрецы успеют решить эту головоломку!

69. Анализ и синтез

Современное общество ставит человека перед необходимостью находить и принимать решения в различных жизненных и производственных ситуациях. Производство и сбыт почти ежедневно требует от начальника цеха, инженера, рабочего, менеджера безотлагательно найти правильное решение. Нахождение решений важно и в семейной жизни, и в быту. От того, будут ли юноши и девушки знать методы поиска решений и обладать навыками их нахождения, зависит их психологическая подготовленность к жизни в обществе.

За последние десятилетия возникла новая наука — системный анализ, которая, в частности, рассматривает вопросы поиска и выбора решений в различных ситуациях не только при индивидуальном, но и при коллективном рассмотрении проблем, при использовании человеко-машинного диалога и т. д.

Методы нахождения решений и психическая деятельность, связанная с поиском решения, во многом сходны как в жизненных или производственных задачах, так и в школьных (по математике, физике, химии). Поэтому ознакомление еще в средней школе с методами поиска решений готовит к решению проблем в любой будущей деятельности, в жизни.

Важнейшими элементами любого метода поиска решения являются *анализ* и *синтез*. При решении математических задач синтез может использоваться в двух формах рассуждения: когда двигаются от данных к искомым фактам или когда элементы объединяют в одно целое. Точно так же и анализ может выступать в двух формах: когда в рассуждениях двигаются от искомых к данным задачи или когда целое (фигуру, выражение и т. п.) расчленяют на части.

Синтез, проведенный в форме постепенного «восхождения» от данных к искомому, позволяет изложить уже найденное решение четко и логично: из данных делается один вывод, затем другой, из них логически следует третий и т. д., а в конце цепочки выводов получается то, что требовалось вычислить, узнать или доказать. Это позволяет *убедить* слушающего в правильности, логической безупречности решения. Однако тому, кто слушает решение, не всегда понятно, как можно до всего этого догадаться. При синтетическом изложении «за кадром» остается вопрос о том, почему был выбран именно этот путь рассуждения и как (при самостоятельном поиске) избежать ложных, побочных шажков мысли. Можно сказать, что изложение решения следует законам математической логики, а не «математической психологии», позволяющей научить делать пусть маленькие, но открытия.

Анализ же в первую очередь направлен на *поиск* пути решения: после завершения анализа нередко требуется заново провести синтетическое рассуждение, чтобы оформить и изложить найденное решение. Но зато анализ позволяет показать, как можно самому решить задачу. Он в большой мере способствует развитию мышления и творческих способностей.

Анализ задачи состоит в том, что мы предполагаем ее уже решенной и находим различные следствия (или предпосылки) этого предположения, а затем, в зависимости от вида этих следствий, пытаемся найти путь отыскания решения поставленной задачи. Грубо говоря, «рецепт» проведения анализа состоит в последовательном проведении трех этапов рассуждений:

1) предположим, что задача решена;

2) посмотрим, какие из этого можно извлечь выводы;

3) теперь, сопоставляя полученные выводы, попытаемся найти путь для действительного решения задачи.

Применение этого «рецепта» анализа к решению задач на построение хорошо известно. Но и в других случаях — при решении уравнений, доказательстве неравенств и т. д. — мы постоянно применяем метод анализа (не всегда осознавая это).

Рассмотрим применение анализа при решении уравнений.

Задача 8. Решить уравнение

$$\sqrt{x+2} + \sqrt{x-1} = 3. \tag{1}$$

Типичное решение выглядит так. Последовательно упрощаем уравнение следующим образом:

$$\sqrt{x+2} = 3 - \sqrt{x-1}, \tag{2}$$

$$x + 2 = 9 - 6\sqrt{x-1} + x - 1, \tag{3}$$

$$6\sqrt{x-1} = 6, \tag{4}$$

$$\sqrt{x-1} = 1, \tag{5}$$

$$x - 1 = 1. \tag{6}$$

После этого находим, что $x = 2$, проводим проверку (подстановкой в уравнение (1)) и получаем ответ: единственным корнем уравнения (1) является $x = 2$.

Легко видеть, что последовательный переход от уравнения (1) к уравнению (6) представляет собой *анализ* предложенной задачи, хотя в явном виде никто на уроках алгебры об этом не говорит. Что же происходит в описанном процессе решения уравнения?

В полном виде этот процесс может быть осмыслен следующим образом. Мы вначале не знаем корней заданного уравнения (1), т. е. решение задачи нам неизвестно. Но предположим, что задача решена, и пусть число x_1 является искомым корнем этого уравнения. Тогда справедливо числовое равенство $\sqrt{x_1+2} + \sqrt{x_1-1} = 3$. Вычитая из обеих частей этого равенства число $\sqrt{x_1-1}$, получаем новое равенство $\sqrt{x_1+2} = 3 - \sqrt{x_1-1}$, которое показывает, что x_1 является также корнем уравнения (2). Далее мы последовательно убеждаемся, что x_1 является корнем каждого из уравнений (3), (4), (5), (6). Наконец, мы замечаем, что уравнение (6) принадлежит к известному нам типу (уравнение первой степени с одним неизвестным). На этом процесс анализа и заканчивается. Дальнейшее решение состоит в нахождении корня уравнения (6) и проверке.

Проведенный процесс анализа схематически показан на рис. 402. Штриховая стрелка изображает сделанное в начале анализа допущение («предположим, что задача решена, т. е. нам известен корень x_1 уравнения (1)»). Дальнейшие стрелки (тонкие) показывают следствия, которые мы выводим из этого предположения («тогда x_1 является корнем каждого из уравнений (2), ..., (6)»). Наконец, замечаем, что полученное уравнение (6) принадлежит известному типу («ага, такое

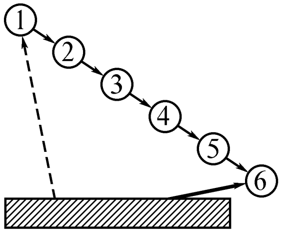


Рис. 402

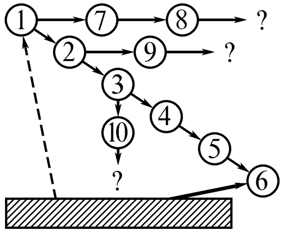


Рис. 403

уравнение мы уже умеем решать!»); это изображено на рис. 402 жирной стрелкой. Цикл замкнулся — анализ решения задачи закончен (дальнейшая часть решения на этой схеме не показана).

Подобным же образом проводится решение уравнений, принадлежащих к другим типам. Заданное уравнение постепенно преобразовывается, например, «расщепляется» на одно или несколько уравнений все более простых типов, пока не обнаружится, что уравнения, к которым мы в конце концов пришли, решаются известными нам методами.

Конечно, не всегда процесс анализа проходит так гладко. Нахождение описанного выше цикла, символизирующего завершение стадии анализа, может быть осложнено тем, что мы начнем производить преобразования, которые ведут по ложному пути, уводят от решения. Например, может показаться, что для решения уравнения (1) нужно обе его части умножить на «сопряженное выражение» $\sqrt{x+2}-\sqrt{x-1}$, подобно тому, как это делается при освобождении от иррациональности в знаменателе. В результате получается уравнение

$$(x+2)-(x-1)=3(\sqrt{x+2}-\sqrt{x-1}), \tag{7}$$

которое после упрощения приводится к виду

$$\sqrt{x+2}-\sqrt{x-1}=1. \tag{8}$$

Однако само по себе это уравнение *не проще* первоначального. Поэтому придется вернуться к уравнению (1) и искать другие пути. Возможны и иные попытки, не приводящие к решению (рис. 403).

Жирная стрелка на рис. 403, замыкающая цикл анализа, символизирует *озарение*, происходящее в сознании решающего. До этого момента он делает шаги, если и не вслепую, то, во всяком случае, не зная точно, по верному или ложному пути следует. Иными словами, до замыкания цикла нельзя сказать, что мы видим уже, как решить задачу. Замыкание цикла — озарение: «нашел!», «понял!», «эврика!».

Анализ представляет собой наиболее трудную, творческую стадию процесса решения задачи. Поэтому весьма важно *учить* процессу анализа. Один из возможных путей для этого состоит в том, чтобы *устно* проводить обсуждение процесса анализа. Можно предложить, *не решая уравнения*, объяснить, какие преобразования надо бы сделать для решения и почему в результате должно получиться уравнение уже известного типа. Например, в отношении уравнения (1) «работают» такие соображения: коэффициенты при x в обоих подкоренных выражениях одинаковы, так что, уединяя один радикал и возводя в квадрат, мы получим уравнение, в котором члены, содержащие x, в сумме дают нуль и останется только один радикал. Такое обсуждение требует меньшего времени, чем полное решение уравнения, а дает не меньше (если не больше) для осознания процесса решения.

70. Обратимый анализ

Рассмотрим теперь вопрос о применении анализа при доказательстве неравенств. Это дает нам представление о другой форме проведения анализа.

Задача 9. Доказать, что для любого натурального n справедливо неравенство

$$\sqrt[4]{n^2-1}\,(\sqrt{n+1}-\sqrt{n-1}) \le 1. \tag{1}$$

Умножая обе части на «сопряженный множитель» $\sqrt{n+1}+\sqrt{n-1}$ (он положителен) и производя дальнейшие преобразования, последовательно перепишем это неравенство следующим образом:

$$\sqrt[4]{n^2-1}\,((n+1)-(n-1)) \le \sqrt{n+1}+\sqrt{n-1}, \tag{2}$$

$$2\sqrt[4]{n^2-1} \le \sqrt{n+1}+\sqrt{n-1}, \tag{3}$$

$$2\sqrt{\sqrt{n^2-1}} \le \sqrt{n+1}+\sqrt{n-1}, \tag{4}$$

$$2\sqrt{\sqrt{n+1}\cdot\sqrt{n-1}} \le \sqrt{n+1}+\sqrt{n-1}, \tag{5}$$

$$\sqrt{\sqrt{n+1}\cdot\sqrt{n-1}} \le \frac{\sqrt{n+1}+\sqrt{n-1}}{2}. \tag{6}$$

Наконец, мы замечаем, что неравенство (6) нам *известно*: это — частный случай неравенства $\sqrt{ab} \le \frac{a+b}{2}$, справедливого для любых неотрицательных чисел a, b. На этом процесс анализа и заканчивается: установлена некоторая связь между неравенством (1), подлежащим доказательству, и известными фактами.

Но чего мы добились выполнением этих преобразований? В полном виде проведенные рассуждения можно осмыслить следующим образом. Предположим, что задача решена, т. е. справедливость неравенства (1) уже установлена. Тогда, домножая на «сопряженный множитель», находим, что неравенство (2) также должно быть справедливо. Далее последовательно убеждаемся, что должны быть справедливы неравенства (3), (4), (5), (6). Наконец, замечаем, что неравенство (6) нам *известно*. Цикл замкнулся, причем в данном случае цикл имеет тот же вид, что и раньше (рис. 402).

Получили ли мы в результате проведения анализа *доказательство* неравенства (1)? Разумеется, нет! Ведь, формально говоря, мы лишь установили, что из неравенства (1) (если бы мы его уже умели доказывать!) можно было бы вывести (6), т. е. частный случай хорошо известного неравенства. Вроде бы невелико приобретение! Однако мы нашли при этом цепочку неравенств (1) – (6), каждое из которых примыкает к предыдущему, т. е. как бы построили «лесенку» из шести ступенек, ведущую от известных фактов к неравенству (1). Возникает надежда, что если в этой цепочке (1) – (6) не только каждое неравенство *следует* из предыдущего, но и наоборот, из каждого неравенства можно было бы *вывести* предыдущее, то мы смогли бы, идя от (6) к (1), получить искомое *доказательство* неравенства (1). Иными словами, в процессе анализа был осуществлен поиск доказательства, в результате которого найден *возможный путь* доказательства:

$$(6) \Rightarrow (5) \Rightarrow (4) \Rightarrow (3) \Rightarrow (2) \Rightarrow (1).$$

Легко проследить, что в данном случае этот путь доказательства реализуется, т. е. действительно из (6) непосредственно вытекает (5), из (5) вытекает (4) и т. д.

Таким образом, решение поставленной задачи состоит из двух стадий:

1) анализа (в процессе которого осуществляется поиск пути доказательства);

2) доказательства.

Разумеется, при некотором навыке можно проводить обратимый анализ, т. е. *совместить* обе стадии: сразу устанавливать, что все «шаги» обратимы (т. е. устанавливать справедливость не только $(1) \Rightarrow (2)$, $(2) \Rightarrow (3)$ и т. д., но и $(1) \Leftrightarrow (2)$, $(2) \Leftrightarrow (3)$ и т. д.).

Но, может быть, стадия анализа является лишней (ведь все равно доказательство приходится проводить заново)? Конечно, нет! Вряд ли ученик (даже очень хороший) сможет по виду неравенства (1), не проводя анализа, определить, что надо исходить именно из неравенства (6), вывести из него (5) и т. д.

Может быть, в таком случае лишним является доказательство (ведь оно повторяет в обратном порядке рассуждения, проведенные на стадии анализа)? Опять нет! Ограничиться одним анализом нельзя, так как если хотя бы один шаг в цепочке $(6) \Rightarrow (5) \Rightarrow (4)$ ... окажется необратимым, то доказательства мы на этом пути не получим. Проверка обратимости всех шагов и представляет собой доказательство. Таким образом, обе стадии (анализ и доказательство) являются неотъемлемыми частями решения.

Заметим, что в рассматриваемом случае жирная стрелка на рис. 402, замыкающая цикл, означает *озарение* («ага, это неравенство мы уже умеем доказывать!»). В момент замыкания цикла происходит скачок от незнания к знанию: становится виден путь доказательства.

71. Анализ — поиск решения

За недостатком времени на уроке обычно не рассказывается о поиске доказательства теоремы. Между тем путь доказательства (включая проведение вспомогательных построений) часто бывает довольно неожиданным и вызывает у вдумчивого ученика естественный вопрос: «А как же можно было «открыть» это доказательство?». Для воспитания математической культуры учащихся полезно рассказать о поиске еще не известного доказательства, т. е. провести анализ стоящей перед нами проблемы.

Задача 10. Доказать следующую теорему: множество всех точек (на плоскости), равноудаленных от двух заданных точек A и B, есть прямая l, перпендикулярная отрезку AB и проходящая через его середину.

Для доказательства этой теоремы нужно установить следующие два факта:

1) если $M \in l$, то $MA = MB$;

2) если $M \notin l$, то $MA \neq MB$.

Справедливость первого факта непосредственно следует из соображений симметрии. Второй факт доказывается несколько сложнее; в этом случае с целью поиска доказательства целесообразно провести анализ.

Пусть $M \notin l$; для определенности будем считать, что точки M и A лежат по одну сторону от прямой l (рис. 404). Чертеж подсказывает, что, видимо, нужно доказывать неравенство

$$MB > MA. \tag{1}$$

Прежде всего замечаем, что отрезок MB пересекает прямую l в некоторой точке C (рис. 405). Так как $MB = MC + CB$, то неравенство (1) можно переписать в виде:

$$MC + CB > MA. \tag{2}$$

Теперь замечаем, что, поскольку $C \in l$, точка C (в силу уже доказанного) равноудалена от точек A и B. Это естественно приводит к мысли провести отрезок AC и написать равенство $CB = CA$. Неравенство (2) переписывается после этого в виде:

$$MC + CA > MA. \tag{3}$$

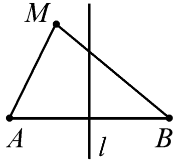


Рис. 404

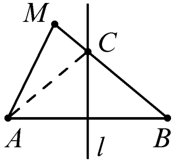


Рис. 405

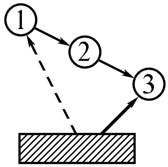


Рис. 406

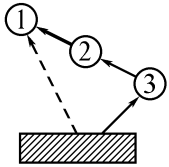


Рис. 407

Наконец, мы замечаем, что соотношение (3) нам известно — сумма длин двух сторон треугольника больше длины третьей стороны. Цикл замкнулся (рис. 406) — стадия анализа закончена. Идя в обратном порядке (3) $\Rightarrow$ (2) $\Rightarrow$ (1), мы получаем доказательство требуемого неравенства (1).

В данном случае процесс анализа несложен: он содержит всего три «ступеньки». Однако для понимания доказательства он очень важен. В самом деле, представим себе, что кто-то, доказывая нам неравенство (1), демонстрирует рис. 404 и затем вдруг проводит отрезок AC (рис. 405). Почему? Зачем? Далее говорится: «применим теорему о том, что сумма длин двух сторон треугольника больше длины третьей стороны». Почему именно эту теорему? К тому же на основании этой теоремы можно написать три различных неравенства, связывающих длины сторон треугольника. Почему же используется именно неравенство (3)? На эти вопросы может быть только один ответ: «Потому что таким путем удается получить нужное доказательство». Не видя стадию анализа, мы становимся пассивным наблюдателем, а не активным, сознательным участником процесса доказательства теоремы.

Заметим, что если нам объясняют лишь доказательство теоремы (не проводя анализа), то озарение (возникающее при *понимании* излагаемого доказательство), имеет психологически совершенно иной характер, показанный на рис. 407 («да, действительно, из уже установленного соотношения (2) следует то, что мы хотели доказать!»).

Таким образом, характер озарения, возникающего при проведении анализа (т. е. *поиска* доказательства) и при прослушивании *готового* доказательства, различен.

Задача 11. Двое играют в игру «Ицзяньши», которую в простейшем варианте можно описать следующими правилами. Первый играющий называет число 1 или 2. Второй прибавляет к названному числу еще 1 или 2 и называет сумму. Затем первый прибавляет к результату 1 или 2, затем второй и т. д. Выигрывает тот, кто назовет 10.

Чтобы найти стратегию, ведущую к выигрышу, проведем анализ. Сколько бы ни прибавил второй игрок, первый может добавить до трех: если второй сказал 1, то первый прибавляет 2, а если второй сказал 2, то первый прибавляет 1. Иначе говоря, первый игрок всегда может сделать так, что в итоге двух ходов прибавляется 3. Значит, если первый сумеет назвать число 7, то он обязательно выиграет: сколько бы ни сказал после этого второй игрок, первый добавит до 3 и назовет 10. Таким образом, 7 — это для первого игрока выигрышное число. Иными словами, мы осуществили преобразование задачи: правила те же, но выигрывает тот, кто назовет число 7.

Теперь ясно, что если первый назовет число 4, то он также выиграет (назвав следующим ходом 7). Это — второй шаг преобразования задачи. Но чтобы добраться до 4, первому надо начинать с числа 1: тогда, что бы ни прибавил второй, он сможет назвать 4. И это означает инсайт: «Эврика!». Назвав первым ходом число 1, а затем прибавляя числа так, чтобы получилось 4, 7, 10, первый игрок обязательно выиграет. Схема цикла озарения та же (рис. 407).

Далее можно применить обобщение: при тех же правилах считать выигравшим того, кто получит 20 (или 30); другой вариант — условиться, что каждый играющий прибавляет при очередном ходе какое-либо из чисел 1, 2, 3, 4, 5, а игра ведется до 50.

Классический вариант игры ицзяньши состоит в том, что каждый играющий при очередном ходе прибавляет какое-либо из чисел 1, 2, ..., 10, а игра ведется до 100. Здесь выигрышными являются числа 1, 12, 23, 34, 45, 56, 67, 78, 89, 100, получающиеся, если из 100 последовательно отнимать число 11. В самом деле, сколько бы ни прибавил при своем очередном ходе второй игрок, первый может добавить до 11 (если второй сказал 1, то первый прибавляет 10, если второй сказал 2, то первый прибавляет 9 и т. д.). Таким образом, при правильной игре первый всегда выигрывает.

72. Поиск решения нестандартных задач

В нестандартных задачах (которых, к сожалению, в школьном курсе почти нет!) процесс анализа становится менее тривиальным и потому особенно заметным. В качестве первого примера рассмотрим следующую задачу.

Задача 12. Доказать, что если в треугольнике две биссектрисы конгруэнтны, то этот треугольник — равнобедренный.

Эта задача, которая носит имя «теоремы Штейнера—Лемуса», в 1940 г. была предложена Д. О. Шклярским на занятии школьного математического кружка при МГУ для «чисто геометрического» решения (т. е. без использования формул, выражающих длины биссектрис). Один из авторов книги посещал этот кружок и ему удалось тогда найти такое решение, но для этого пришлось провести следующий анализ.

В треугольнике ABC (рис. 408) проведены биссектрисы. Известно, что $AD = CE$. Требуется доказать, что треугольник ABC — равнобедренный, т. е.

$$AB = BC. \tag{1}$$

Зная, что биссектрисы треугольника пересекаются в одной точке, проведем третью биссектрису — возможно, это пригодится. Чтобы доказать равенство длин каких-либо отрезков, часто используют следующий прием: устанавливают конгруэнтность треугольников, имеющих эти отрезки своими сторонами. В данном случае, видимо, лучше всего попытаться доказать, что

$$\triangle ABD \cong \triangle CBE, \tag{2}$$

так как в этих треугольниках сторонами являются отрезки AD, CE, длины которых по условию равны, и отрезки AB, BC, для которых это нужно доказать. Посмотрим, какие одинаковые элементы имеются у интересующих нас треугольников ABD и CBE (для удобства эти треугольники изображены отдельно на рис. 409а, б). В этих треугольниках длины оснований AD и CE по условию равны. Кроме того, угол B у них общий (т. е. на рис. 409 $\angle ABD = \angle CBE$). Наконец, оба треугольника имеют общую биссектрису BO. Итак, интересующие нас треугольники имеют по три соответственно конгруэнтных элемента. Но эти элементы таковы, что не подходит ни один из известных признаков конгруэнтности треугольников. Похоже на то, что придется доказать новый признак, который должен формулироваться следующим образом:

если в двух треугольниках конгруэнтны основания, конгруэнтны углы при вершине и конгруэнтны биссектрисы, проведенные из вершины, то эти треугольники конгруэнтны. (3)

Как же доказать утверждение (3)? Для удобства на рис. 410б один из треугольников симметрично отражен относительно биссектрисы, так что треугольники находятся теперь «в одинаковом положении». Используя метод доказательства «от противного», можно попытаться вместо (3) доказать следующее утверждение:

Если в двух треугольниках конгруэнтны основания и конгруэнтны углы при вершине, но сами треугольники не конгруэнтны, то у них не могут быть конгруэнтны биссектрисы, проведенные из вершины. (4)

У треугольников, о которых идет речь в утверждении (4), конгруэнтны основания и конгруэнтны углы при вершине. Это наводит на мысль использовать теорему о сегменте, вмещающем данный угол. Иными словами, если основания AD и $C'E'$ совместить, то обе вершины B и B' окажутся на одной окружности, проходящей через концы основания (рис. 411). При этом обе точки B, B' лежат по одну сторону перпендикуляра l, проведенного через середину основания, и не совпадают (т. е. $B \neq B'$), поскольку мы предполагаем, что треугольники не конгруэнтны. Заметим, наконец, что продолжения биссектрис проходят через точку N пересечения прямой l с окружностью. В результате мы приходим к следующей эквивалентной формулировке.

В окружности проведены диаметр MN и перпендикулярная ему хорда AD. На дуге DM (меньшей полуокружности) взяты две различ-

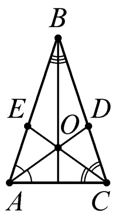


Рис. 408

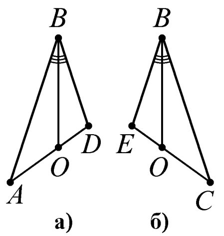


Рис. 409

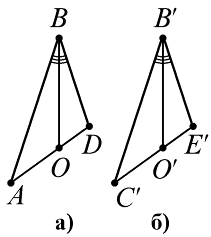


Рис. 410

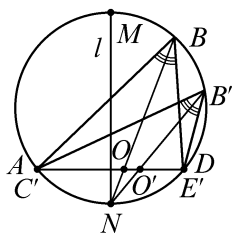


Рис. 411

ные точки B, B'. Если O и O' — точки пересечения прямых BN, $B'N$ с хордой AD, то

$$BO \neq B'O'. \tag{5}$$

Судя по чертежу, можно высказать более определенное суждение: если $\angle MNB < \angle MNB'$, то

$$BO > B'O'. \tag{6}$$

Чтобы доказать неравенство (6), удобно воспользоваться соотношениями $BO = BN - ON$, $B'O' = B'N - O'N$. Ясно, что если $\angle MNB < \angle MNB'$, то $BN > B'N$. Значит, неравенство (6) будет установлено, если мы обнаружим, что

$$ON < O'N. \tag{7}$$

Но (инсайт!) это неравенство представляет собой известный факт.

Цикл замкнулся — анализ завершен. Читатель без труда проверит, что, проводя рассуждения в обратном порядке, мы получим искомое доказательство равенства (1). Разумеется, при индуктивном изложении доказательства порядок рассуждений будет обратным. Но разве мыслимо было бы, не проводя анализа, догадаться, что для решения поставленной задачи надо начать с рассмотрения чертежа, показанного на рис. 411?

Заметим, что в приведенном примере логическая связь между ступеньками следующая: $(1) \Leftarrow (2) \Leftarrow (3) \Leftrightarrow (4) \Leftrightarrow (5) \Leftarrow (6) \Leftarrow (7)$. Однако психологически, с точки зрения последовательности появления идей по время рассуждения, связь между ступеньками имеет вид $(1) \to (2) \to (3) \to (4) \to (5) \to (6) \to (7)$, т. е. цикл имеет такой же вид, как на рис. 402 (только содержит семь ступенек).

Рассмотрим еще один пример нестандартной задачи.

Задача 13. Найти четырехзначное число, являющееся точным квадратом, у которого две первые цифры одинаковы и две вторые тоже одинаковы.

Допустим, что задача решена, и пусть a — цифра в разряде тысяч искомого четырехзначного числа, b — цифра в разряде единиц. Тогда

$$1000a + 100a + 10b + b = N^2, \tag{*}$$

где N — некоторое натуральное число. Следовательно,

$$11 \cdot (100a + b) = N^2. \tag{1}$$

Из этого вытекает, что

$$100a + b = 11k^2, \tag{2}$$

где $k = \dfrac{N}{11}$. Так как далее $100a + b = 99a + (a + b)$, то

$$a + b \text{ делится на } 11 \tag{3}$$

а потому

$$a + b = 11 \tag{4}$$

(ибо каждая из цифр a, b не превосходит 9, т. е. $a + b \leq 18$). Из (2) и (4) получаем

$$99a + 11 = 11k, \tag{5}$$

$$9a + 1 = k^2. \tag{6}$$

Остается выбрать из цифр $a = 1, 2, ..., 9$ такую, для которой $9a + 1$ является точным квадратом. Легкий перебор показывает, что этому условию удовлетворяет только

$$a = 7, \tag{7}$$

откуда в силу (4) получаем $b = 4$. Итак, решением является число 7744. Схема рассуждения показана на рис. 412.

73. Соединение анализа с синтезом

В рассмотренных примерах для решения использовался преимущественно либо анализ, либо синтез. Однако в большинстве случаев применение анализа осуществляется не в чистом виде, а в соединении с синтезом.

Мы это видели на примере доказательства неравенства (п. 70). Двигаясь от доказываемого неравенства «понятным рассуждением» к более простым неравенствам, мы на каждом этапе проверяем обратимость сделанного шага (рис. 413). Завершив рассуждение, т. е. добравшись, наконец, до истинного высказывания, мы не только осуществляем анализ (сплошные стрелки), но и имеем полную схему синтетического изложения решения (пунктирные стрелки, ведущие от данных к искомым).

Обратимый анализ позволяет не только получать решения задач на доказательство, но и выяснять справедливость (или ложность) некоторых гипотез. Пусть предложение (*) (рис. 413) является гипотезой и, проведя обратимый анализ, мы получили последовательно, скажем, предложения (1), (2), (3). Иначе говоря, мы установили, что

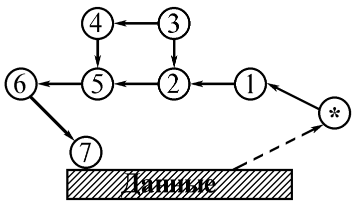


Рис. 412

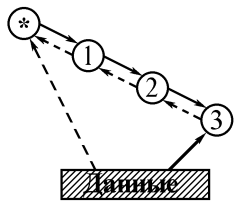


Рис. 413

(1) — необходимое (сплошная стрелка) и достаточное (пунктирная стрелка) условие справедливости гипотезы (*), далее, (2) — необходимое и достаточное условие справедливости предложения (1), а значит, и гипотезы (*), а заключительное предложение (3) этой цепочки является необходимым и достаточным условием справедливости (2), а потому и гипотезы (*).

Тогда если предложение (3) истинно, то гипотеза доказана. Если же цепочка привела к ложному предложению, т. е. (3) ложно, то исходная гипотеза опровергнута. Рассмотрим пример.

Задача 14. Существует ли приведенное квадратное уравнение $x^2 + px + q = 0$, у которого сумма коэффициентов p и q равна -13, а разность корней 6?

Примем гипотезу: такое уравнение существует, т. е.

$$p + q = -13, \quad x_2 - x_1 = 6. \tag{*}$$

Для проверки этой гипотезы проведем обратимый анализ. Так как разность корней равна $\sqrt{D} = \sqrt{p^2 - 4q}$, то

$$p + q = -13, \quad \sqrt{D} = 6. \tag{1}$$

Легко видеть (с учетом $x_1 + x_2 = -p$), что переход $(*) \Rightarrow (1)$ обратим, т. е. если существует уравнение $x^2 + px + q = 0$, для которого справедливо (1), то для этого уравнения справедливо и (*). Теперь выразим q через p и подставим это значение в выражение для $\sqrt{D}$; получим

$$\sqrt{p^2 + 4(13 + p)} = 6. \tag{2}$$

И этот переход $(1) \Rightarrow (2)$ обратим. Соотношение (2) преобразуем:

$$\sqrt{(p + 2)^2 + 48} = 6. \tag{3}$$

И здесь переход $(2) \Rightarrow (3)$ обратим. Однако из (3) получаем

$$(p + 2)^2 = 36 - 48 = -12, \tag{4}$$

и потому не существует p, удовлетворяющего условию (3). Значит, и первоначальная гипотеза была ложной, т. е. уравнения, о котором идет речь в задаче, не существует.

Заметим, что если бы мы видоизменили условие, указав, что разность корней равна 7, то в результате аналогичного обратимого анализа пришли бы к соотношению $(p + 2)^2 + 48 = 49$, из которого получили бы два корня $p_1 = -1$, $p_2 = -3$ (и соответственно $q_1 = -12$, $q_2 = -10$). Значит, при такой постановке задачи существуют два искомых уравнения: $x^2 - x - 12 = 0$, $x^2 - 3x - 10 = 0$.

Заметим, что если задача ставится определенно: *опровергнуть* некоторую гипотезу, — то проверять обратимость каждого шага не нужно. Получив из гипотезы (*) ряд *следствий* и обнаружив, что последнее следствие ложно, уже можно сделать вывод о ложности исходной гипотезы (*). По существу, такой прием рассуждения сводится к методу доказательства от противного. Рассмотрим пример.

Задача 15. Доказать, что радиус описанной окружности треугольника не может быть меньше шестой части его периметра.

Допустим, что существует треугольник, в котором

$$R < \frac{2p}{6} \tag{*}$$

($2p$ — периметр), и установим ложность этой гипотезы. Обозначим через a наибольшую сторону треугольника (или одну из наибольших). Тогда $2p \le 3a$, и потому из (*) получаем

$$R < \frac{a}{2} \tag{1}$$

(отметим, что переход (*) $\Rightarrow$ (1) необратим). Но тогда

$$2R < a \tag{2}$$

и потому в треугольнике OBC (рис. 414) мы имеем

$$OB + OC < BC. \tag{3}$$

Однако это утверждение ложно (оно противоречит неравенству треугольника). Этим установлена и ложность исконной гипотезы (рис. 415), т. е. треугольника, удовлетворяющего условию (*), не существует.

Имеется еще одна форма соединения анализа с синтезом, которая применяется для подтверждения, а не опровержения гипотез. Схема этой формы соединения анализа с синтезом показана на рис. 416. Здесь идущая от гипотезы (*) пунктирная стрелка (не указывающая на логическое следствие) означает поиск такого предложения (1), из которого вытекает (*). Найдя затем предложение (2), из которого вытекает (1), мы получаем следующий шаг рассуждения и т. д. Если мы в конце концов получим такое предложение, скажем, (3), которое является истинным, то (инсайт!) мы получаем подтверждение исходной гипотезы (*). Здесь пунктирные стрелки обозначают анализ (поиск нужных предложений), а сплошные стрелки — синтез, дающий доказательство предложения (*).

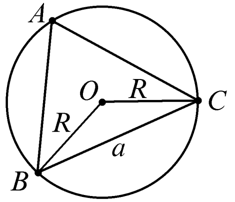


Рис. 414

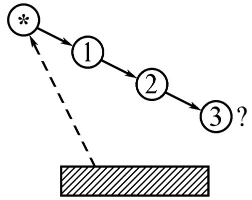


Рис. 415

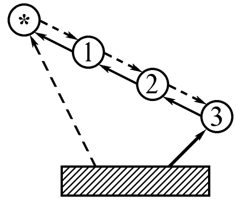


Рис. 416

Задача 16. Доказать неравенство

$$100^{200} > 200!. \qquad (*)$$

Для решения представим правую часть в несколько ином виде, выделив вначале четыре множителя, а остальные сгруппировав парами:

$$100^{200} > 200 \cdot 100 \cdot 1 \cdot 199 \cdot (2 \cdot 198) \cdot (3 \cdot 197) \cdot \ldots \cdot (99 \cdot 101). \qquad (1)$$

Здесь произведение первых четырех множителей справа меньше 100^4. Поэтому, если мы докажем, что каждое произведение в скобках меньше 100^2, то требуемое неравенство будет доказано. Иначе говоря, для подтверждения неравенства (1) достаточно установить, что

$$100^2 > k\,(200 - k). \qquad (2)$$

при каждом $k = 2, 3, \ldots, 99$.

Заметим, что переход (2) $\Rightarrow$ (1) необратим (например, из неравенства $17 \cdot 17 > (2 \cdot 9) \cdot (3 \cdot 5)$ не вытекает, что справедливо каждое из неравенств $17 > 2 \cdot 9$, $17 > 3 \cdot 5$), т. е. (2) есть именно достаточное условие для (1). Далее, неравенство (2) равносильно утверждению

$$(100 - k)^2 > 0, \qquad (3)$$

которое при $k = 2, 3, \ldots, 99$, очевидно, истинно. Тем самым неравенство (*) установлено (схема рассуждений показана на рис. 416).

Наиболее общим и плодотворным является такое соединение анализа с синтезом, при котором мы совершаем в рассуждении *попеременное* движение с двух сторон: то от данных в направлении искомого (синтез), то от искомого к данным (анализ), пока не сблизим получаемые предложения настолько, что будет осуществлен инсайт.

Задача 17. Доказать, что в прямоугольном треугольнике биссектриса прямого угла делит пополам угол между медианой и высотой, проведенными к гипотенузе.

Изучаем условие и выполняем чертеж (рис. 417). Замечаем, что полезно провести описанную окружность и продлить биссектрису треугольника до пересечения с этой окружностью в середине S дуги AC (рис. 418), откуда следует, что

биссектриса лежит между медианой и высотой. (1)

На рис. 419 этот первый шаг поиска решения обозначен (1), хотя на самом деле это не один «шажок» мысли, а целая цепочка умозаключений. Теперь сформулируем искомое (цель решения):

$$\angle 2 = \angle 3. \qquad (*)$$

Пока дальнейшее продвижение от данных в направлении искомого неясно. Пробуем идти «попятным» движением от искомого. Так

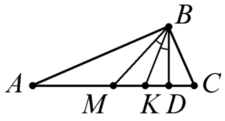


Рис. 417

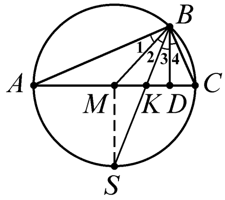


Рис. 418

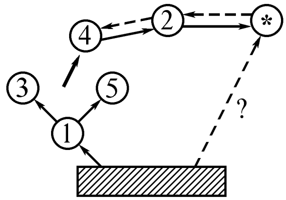


Рис. 419

как $\angle ABK = \angle KBC$, то для доказательства равенства (*) достаточно установить, что

$$\angle 1 = \angle 4. \tag{2}$$

Возвращаемся к данным. Способа доказать равенство (2) пока не видно, но мы замечаем, что $\angle 1$ равен *другому* углу:

$$\angle 1 = \angle A. \tag{3}$$

Возвращаемся к искомому. Из (3) видно теперь, что для подтверждения соотношения (2) достаточно установить равенство

$$\angle A = \angle 4. \tag{4}$$

Однако пока не видно, из чего можно было бы получить это равенство. Снова возвращаемся к данным. Замечаем, что $\angle A$ и $\angle 4$ — острые углы прямоугольных треугольников ABC и DBC с общим острым углом C, т. е. $\angle A$ и $\angle 4$ дополняют $\angle C$ до 90°:

$$\angle A = 90° - \angle C, \quad \angle 4 = 90° - \angle C. \tag{5}$$

Значит, (4) верно (эврика!), чем и завершается решение.

Поиск решения проходил по схеме соединения анализа с синтезом: движение шло попеременно в «прямом» и «попятном» направлениях, а инсайт осуществился в середине (рис. 419).

Заметим, что рассказ этого решения учащийся будет излагать не в том порядке шажков мысли, который реализован при решении, а в виде синтетической цепочки (1) $\Rightarrow$ (5) $\Rightarrow$ (4) $\Rightarrow$ (3) $\Rightarrow$ (2) $\Rightarrow$ (*).

Беседа 15. Наглядность. Аналогия. Интуиция

74. Формула наглядности — изоморфизм плюс простота

Остановимся теперь на категории *наглядности*, которая имеет важнейшее значение для поиска решения задачи.

Понимание наглядности и ее роли в учебном процессе за последнее время значительно изменилось и расширилось. Мы сейчас считаем, что иной раз запись, сделанная мелом на доске, или даже устный рассказ учителя могут быть более наглядными, чем демонстрация явления в его натуральном виде. Например, формула серной кислоты, написанная на доске, является в некотором смысле существенно более наглядной, чем созерцание прозрачной жидкости (серной кислоты), налитой в сосуд, стоящий на столе учителя. Формула соединения говорит подготовленному учащемуся неизмеримо больше о химических свойствах соединения, чем натуральный объект: она вскрывает химический состав молекулы, валентные связи между отдельными атомами, позволяют легко найти молекулярную массу и т. п.

Разберем на этом примере, в чем заключается преимущество формальной записи соединения. Нас интересует лишь одна сторона явления: количество и характер атомов в молекуле вещества. От всех остальных сторон мы отвлекаемся, абстрагируемся. В этой модели отсутствуют все несущественные для нас стороны реального объекта. Модель отражает только одно — количество и характер атомов в химическом соединении. Это единственная сторона, связывающая нашу модель с реальным объектом. Но зато она представлена в рассматриваемой модели адекватно, изоморфно изучаемому явлению. В данном случае это означает, что реальная молекула и наша модель являются изоморфными по своему составу, а только состав нас в данном случае и интересует.

Итак, только одна сторона изучаемого явления отражена в модели, причем отражена изоморфно, так что модель чрезвычайно упрощает рассматриваемый процесс, огрубляет его. Взамен утраченных сторон явления модель приобретает *простоту* восприятия. Эти две характерные черты модели (изоморфное отражение существенных черт явления и простота восприятия модели) и выражают *наглядность* модели.

В этом смысле применяемые в школе модели геометрических тел являются наглядными. Например, понятие прямоугольного параллелепипеда, изучаемого в начальных классах, является довольно сложным. Оно предполагает умение абстрагироваться от цвета, поверхности тела, температуры, шероховатости и т. п., что дается учащимся далеко не сразу. Реальный предмет (например, деревянный брусок) не является, конечно, прямоугольным параллелепипедом в математическом понимании, он служит лишь моделью. Эта модель изоморфно отражает основные свойства прямоугольного параллелепипеда как геометрического тела: у бруска столько же граней, ребер, вершин, сколько и у абстрактного параллелепипеда, такая же схема соединения (например, в каждой вершине сходятся по три ребра). Таким образом, изоморфизм здесь налицо. В то же время деревянный брусок, несомненно, отличается простотой восприятия — предметы такого вида учащиеся неоднократно наблюдали, они для них просты и привычны. Мы видим, что модель прямоугольного параллелепипеда (так же как и модели других геометрических тел) обладает свойством наглядности.

Еще один пример. При изучении прямоугольного параллелепипеда известную трудность для учащихся составляет понятие каркаса, т. е. фигуры, составленной из всех ребер и вершин параллелепипеда. «Увидеть» устройство этой фигуры, изучая деревянный брусок, не так

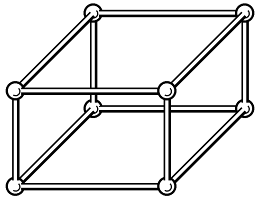


Рис. 420

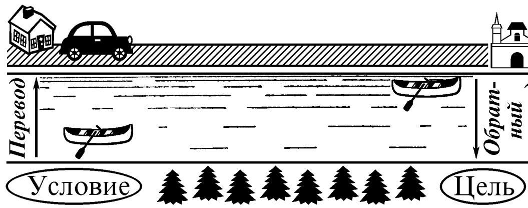


Рис. 421

легко. Поэтому на уроках применяется каркасная модель прямоугольного параллелепипеда (рис. 420). Изоморфизм здесь налицо (т. е. число вершин и ребер, а также схема соединения и параллельность на модели точно соответствует аналогичным элементам абстрактного каркаса). Неоспорима также и простота восприятия: невидимые ребра и вершины, которые ученику нелегко себе представить, рассматривая сплошной деревянный брусок, на этой модели хорошо видны; ясно видны четверки параллельных ребер, равенство противоположных ребер и т. п. Таким образом, каркасная модель оказывается (при изучении свойств параллелепипеда) *проще* для восприятия, чем сплошной деревянный брусок, а следовательно, и *нагляднее*.

Значение понятия изоморфизма как одной из двух существенных черт наглядности трудно переоценить. Требование правильного, адекватного отображения изучаемого явления с помощью прибора, таблицы и т. п. заключается в том, что это средство обучения должно представлять собой изоморфную модель. Благодаря этому выводы, полученные при рассмотрении модели, правильно, адекватно, изоморфно отображают свойства самого изучаемого явления (в отношении тех свойств реального явления, которые отображены в модели).

Простота модели, простота оперирования с ней означают, что с помощью модели *проще* придти к выводам (и притом к правильным выводам — в силу изоморфизма), чем при рассмотрении реального явления.

Переход к наглядной модели является важным средством решения математических задач, важным средством нахождения пути решения.

Можно пояснить применение наглядной модели следующим образом. Перед нами широкая река, один берег которой называется «первоначальной задачей», а второй — «наглядной моделью». Условие задачи и поставленная цель нам понятны. Однако решение, т. е. переход от условия к поставленной цели по прямой представляет трудности. Можно представить себе, что дремучий, труднопроходимый лес встал на пути от условия к поставленной цели (рис. 421). Однако на другом берегу проложен существенно *более простой* путь, идущий как раз в нужном направлении. Это обеспечивается тем, что берега параллельны (т. е. наглядная модель изоморфна изучаемому явлению). Теперь для того, чтобы решить задачу, мы прежде всего переправляемся на другой берег, т. е. осуществляем действие моделирования, как бы производим *перевод* тех данных, которые имеются в

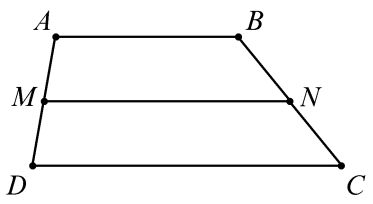


Рис. 422

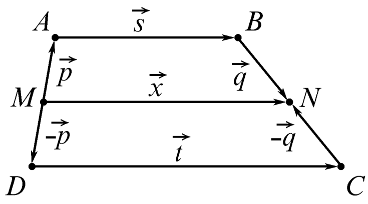


Рис. 423

условии задачи, на язык терминов, имеющихся в наглядной модели. Благодаря простоте оперирования с наглядной моделью, мы значительно легче, чем на берегу первоначальной задачи, продвигаемся шаг за шагом к достижению цели (т. е. выполняем учебные действия *преобразования*). Наконец, наступает инсайт: мы видим, что на Берегу наглядной модели мы подошли к решению, и требуется лишь переплыть реку, возвратиться на исходный берег, т. е. осуществить *обратный перевод* (с языка наглядной модели к исходным терминам). Таким образом, при применении наглядности для решения задачи мы имеем три стадии решения:

1) *моделирование*, т. е. подбор подходящей наглядной модели и перевод первоначальной задачи на язык терминов этой модели;

2) *оперирование* с условием задачи на языке наглядной модели;

3) *инсайт* (Эврика! В этих терминах мы получаем требуемое решение).

Рассмотрим некоторые примеры применения наглядных моделей при решении математических задач.

Пример 1. В выпуклом четырехугольнике точки M и N — середины противоположных сторон AD и BC (рис. 422). Доказать, что если $MN = \frac{1}{2}(AB + CD)$, то $ABCD$ — трапеция.

Заметим, что согласно хорошо известной школьной теореме, если $ABCD$ — трапеция, то длина средней линии MN равна полусумме длин оснований. Задача состоит в том, чтобы доказать *обратную* теорему.

Для решения вместо *точек* будем рассматривать *векторы*. Это и есть та наглядная модель, которую мы здесь хотим применить. Вряд ли векторы *проще* для восприятия, чем точки (хотя векторная модель геометрии *изоморфна* обычной). Однако векторная модель все же обладает определенной простотой: *оперировать* с векторами легче, чем с точками, углами треугольника, поскольку с векторами можно производить действия (сложение, умножение на число), алгебраические свойства которых просты и привычны.

Итак, надо осуществить перевод условия задачи на язык наглядной модели, язык векторов. Для этого обозначим вектор $\overrightarrow{MN}$ через $\vec{x}$ и выберем такие направления отрезков чертежа, чтобы $\vec{x}$ был равен сумме получающихся векторов (рис. 423):

$$\overrightarrow{p} + \overrightarrow{s} + \overrightarrow{q} = \overrightarrow{x}, \tag{1}$$

$$\overrightarrow{(-p)} + \overrightarrow{t} + \overrightarrow{(-q)} = \overrightarrow{x}. \tag{2}$$

Кроме того, по условию задачи

$$\frac{1}{2}(|\overrightarrow{s}| + |\overrightarrow{t}|) = |\overrightarrow{x}|. \tag{3}$$

Этим осуществлен перевод задачи на векторный язык.

Перейдем теперь к преобразованиям в рамках нашей наглядной модели. Вид равенств (1) и (2) подсказывает, что их целесообразно сложить:

$$\overrightarrow{s} + \overrightarrow{t} = 2\overrightarrow{x}. \tag{4}$$

Из равенств (3) и (4) получаем

$$|\overrightarrow{s} + \overrightarrow{t}| = 2|\overrightarrow{x}| = |\overrightarrow{s}| + |\overrightarrow{t}|. \tag{5}$$

Отложим теперь от произвольной точки O вектор $\overrightarrow{OE} = \overrightarrow{s}$ и от точки E вектор $\overrightarrow{EF} = \overrightarrow{t}$. Тогда $\overrightarrow{OF} = \overrightarrow{s} + \overrightarrow{t}$ (рис. 424). Равенство (5) показывает, что

$$OF = OE + EF. \tag{6}$$

Теперь (эврика!) полученное равенство (6) означает, что $E \in OF$, и потому $\overrightarrow{s} \parallel \overrightarrow{t}$. Но это означает (если перевести это на геометрический язык), что $AB \parallel CD$, т. е. $ABCD$ — трапеция.

Схема проведенного рассуждения показана на рис. 425, где двойная горизонтальная пунктирная прямая изображает «реку», разделяющую «Берег первоначальной задачи» и «Берег наглядной модели».

Использованная здесь наглядная модель (векторная интерпретация геометрии) весьма полезна и очень часто применяется при решении геометрических задач. Основная трудность при использовании этой наглядной модели состоит в том, чтобы удачно ввести обозначения нужных векторов, обеспечив тем самым перевод условия геометрической задачи на векторный язык.

П р и м е р 2. Доказать *теорему Архимеда*: медианы произвольного треугольника пересекаются в одной точке (называемой центром тя-

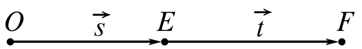


Рис. 424

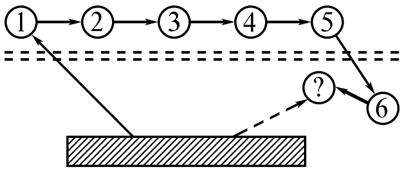


Рис. 425

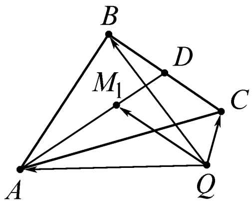


Рис. 426

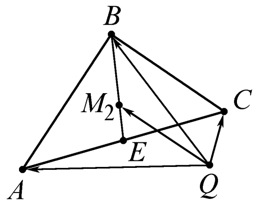


Рис. 427

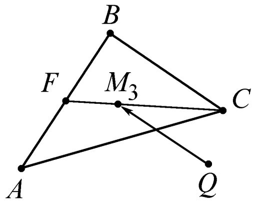


Рис. 428

жести треугольника), расстояние от которой до каждой вершины треугольника равно $\frac{2}{3}$ длины соответствующей медианы.

Для решения обозначим через D середину стороны BC треугольника ABC и введем векторы, как на рис. 426, где Q — произвольная точка. Через M_1 здесь обозначена такая точка, что $AM_1 = \frac{2}{3}AD$. Это — перевод задачи на векторный язык. Тогда

$$\overrightarrow{QM_1} = \overrightarrow{QA} + \overrightarrow{AM_1} = \overrightarrow{QA} + \frac{2}{3}\overrightarrow{AD} = \overrightarrow{QA} + \frac{2}{3}(\overrightarrow{AQ} + \overrightarrow{QD}) =$$
$$= \overrightarrow{QA} + \frac{2}{3}(-\overrightarrow{QA} + \overrightarrow{QD}) = \frac{1}{3}\overrightarrow{QA} + \frac{2}{3}\overrightarrow{QD} =$$
$$= \frac{1}{3}\overrightarrow{QA} + \frac{2}{3}\left(\frac{1}{2}\overrightarrow{QB} + \frac{1}{2}\overrightarrow{QC}\right) = \frac{1}{3}\overrightarrow{QA} + \frac{1}{3}\overrightarrow{QB} + \frac{1}{3}\overrightarrow{QC}.$$

Аналогично точка $M_2 \in BE$, для которой $BM_2 = \frac{2}{3}BE$ (рис. 427) удовлетворяет условию $\overrightarrow{QM_2} = \frac{1}{3}\overrightarrow{QA} + \frac{1}{3}\overrightarrow{QB} + \frac{1}{3}\overrightarrow{QC}$. Точно так же точка $M_3 \in CF$, для которой $CM_3 = \frac{2}{3}CF$ (рис. 428), удовлетворяет условию $\overrightarrow{QM_3} = \frac{1}{3}\overrightarrow{QA} + \frac{1}{3}\overrightarrow{QB} + \frac{1}{3}\overrightarrow{QC}$.

Но (эврика!) это означает, что $M_1 = M_2 = M_3$ — это одна и та же точка. Этим и доказано, что все три медианы пересекаются в одной точке M. Из доказательства видно также, что центр тяжести определяется равенством $\overrightarrow{QM} = \frac{1}{3}(\overrightarrow{QA} + \overrightarrow{QB} + \overrightarrow{QC})$, какова бы ни была точка Q.

Пример 3. Вычислить радиус сферы, описанной вокруг правильного тетраэдра с ребром p.

Для решения введем векторы $\overrightarrow{e_1} = \overrightarrow{OA}$, $\overrightarrow{e_2} = \overrightarrow{OB}$, $\overrightarrow{e_3} = \overrightarrow{OC}$, $\overrightarrow{e_4} = \overrightarrow{OD}$, где A, B, C, D — вершины правильного тетраэдра, а O — центр его описанной сферы. Тогда

$$OA = OB = OC = OD = r, \quad AB = AC = AD = BC = BD = CD = p.$$

Заметим теперь, что $\vec{e}_1 + \vec{e}_2 + \vec{e}_3 + \vec{e}_4 = \vec{0}$. Это вытекает из того, что вектор $\vec{e}_1 + \vec{e}_2 + \vec{e}_3 + \vec{e}_4$ не меняется ни при какой перестановке векторов $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, $\vec{e}_4$ (что возможно только в случае, если этот вектор равен нулю). Возведя равенство $\vec{e}_1 + \vec{e}_2 + \vec{e}_3 + \vec{e}_4 = \vec{0}$ в квадрат, получим (учитывая, что $\vec{e}_i^{\,2} = r^2$, $\vec{e}_i\,\vec{e}_j = r^2 \cos\alpha$ при $i \neq j$, где α — величина угла между векторами $\vec{e}_i$ и $\vec{e}_j$, рис. 429) равенство $4r^2 + 12r^2\cos\alpha = 0$. Отсюда получаем $\cos\alpha = -\frac{1}{3}$. Следовательно,

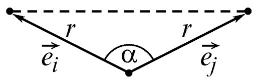


Рис. 429

$$p^2 = (\vec{e}_i - \vec{e}_j)^2 = 2r^2 - 2\vec{e}_i\,\vec{e}_j = 2r^2 - 2r^2\cos\alpha = \left(2 + \frac{2}{3}\right)r^2 = \frac{8}{3}r^2,$$

и потому $r = \frac{\sqrt{3}}{2\sqrt{2}}p$.

Векторная модель является важным и широко применяемым средством решения геометрических задач. Примеры использования этой модели были рассмотрены выше. Укажем еще одну наглядную модель, которая также широко применяется при решении задач — на этот раз алгебраических. Этой моделью является *графическое изображение функций*. Напомним построение этой модели подробнее.

Каждой числовой функции f соответствует ее график $\Gamma(f)$, т. е. множество таких точек $M(x; y)$ координатной плоскости, что x принадлежит области определения функции f и выполнено равенство $y = f(x)$. Сопоставляя каждой числовой функции f ее график $\Gamma(f)$, мы получаем взаимно однозначное отображение Γ множества всех числовых функций на совокупность всех подмножеств координатной плоскости, однозначно проектирующихся (параллельно оси ординат) на ось абсцисс. Это отображение Γ является изоморфизмом в отношении ряда свойств.

Перечислим несколько наиболее типичных свойств, которые легко увидеть при помощи рассмотрения графиков.

1) На некотором промежутке выполняется неравенство $f(x) > g(x)$; графически это выражается в том, что на рассматриваемом промежутке график $\Gamma(f)$ расположен *выше* графика $\Gamma(g)$ (рис. 430). Частным случаем является *положительность* функции f на некотором промежутке (ее график расположен на этом промежутке выше оси абсцисс, рис. 431). В качестве еще одного частного случая отметим, что если функция f ограничена снизу (т. е. существует константа c,

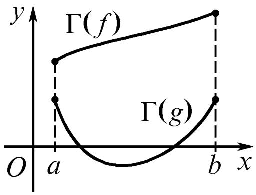


Рис. 430

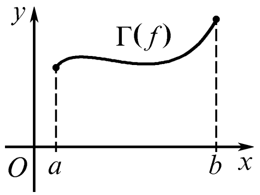


Рис. 431

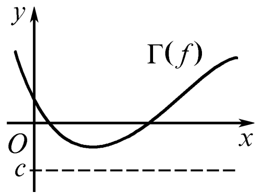


Рис. 432

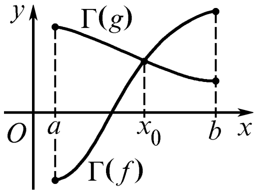


Рис. 433

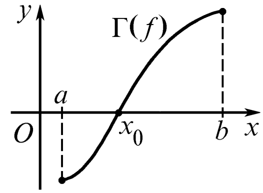


Рис. 434

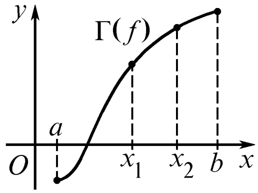


Рис. 435

удовлетворяющая условию $f(x) > c$ для всех x из области определения функции f, то график $\Gamma(f)$ расположен выше некоторой горизонтальной прямой $y = c$, где $c = \text{const}$ (рис. 432).

2) Число x_0 является корнем уравнения $f(x) = g(x)$, графически это выражается в том, что графики $\Gamma(f)$ и $\Gamma(g)$ имеют общую точку, абсцисса которой равна x_0 (рис. 433). Частным случаем этого равенства является утверждение о том, что x_0 — нуль функции f, т. е. корень уравнения $f(x) = 0$; это означает, что точка $A(x_0; 0)$ принадлежит графику функции $\Gamma(f)$ (рис. 434).

3) Функция f является на некотором промежутке возрастающей, т. е. $f(x_2) > f(x_1)$ при $x_2 > x_1$, где x_1, x_2 принадлежат рассматриваемому промежутку; графически это выражается в том, что на данном промежутке при движении вправо график $\Gamma(f)$ поднимается *вверх* (рис. 435). Аналогично рассматриваются убывающие функции. Частным случаем этих свойств является достижение функцией f локального максимума (или минимума) в точке x_0: если на некотором отрезке, примыкающем к точке x_0 слева, функция f является возрастающей, а на некотором отрезке, примыкающем к точке x_0 справа, она является убывающей (рис. 436), то функция f достигает в точке x_0 локального максимума.

Имеются и другие свойства функций, наглядно выражаемые геометрическими свойствами их графиков, но указанные выше свойства являются наиболее часто применяемыми.

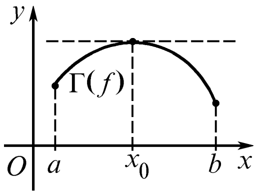


Рис. 436

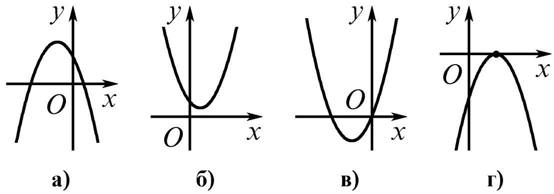


Рис. 437

Хорошо сформированное «графическое мышление» заключается в умении использовать графики в качестве наглядной модели при решении задач (в частности, алгебраических).

Весьма важным является искусство обратного перевода. Например, парабола, являющаяся графиком функции $y = ax^2 + bx + c$ (где $a \neq 0$), неограниченно уходит ветвями вверх при $a > 0$ и вниз при $a < 0$. Надо уметь по эскизу графика (рис. 437) делать выводы о знаках коэффициентов a, b, c.

Приведем примеры решения задач при помощи описываемой наглядной модели, т. е. посредством графиков функций.

Пример 4. Доказать, что если найдется число $x = x_0$, при котором трехчлен $x^2 + px + q$ принимает отрицательное значение, то корни этого трехчлена действительны и различны.

Для решения заметим, что график трехчлена (парабола) имеет ветви, направленные вверх (так как коэффициент при x^2 положителен). График проходит через некоторую точку нижней полуплоскости (поскольку функция при некотором x принимает отрицательное значение). Этим мы перевели задачу на язык наглядной модели. Теперь ясно, что этот график пересекает ось абсцисс в двух различных точках (рис. 438). Но это и означает (здесь мы осуществляем обратный перевод), что трехчлен имеет два различных действительных корня.

Можно предложить различные варианты рассмотренной задачи. Например, доказать, что при действительных a, b и отрицательном c корни уравнения $(x - a)(x - b) + c = 0$ действительны и различны (это следует из того, что при $x = a$ левая часть отрицательна). Или другой вариант: доказать, что если корни трехчлена $x^2 + px + q$ действительны и различны, то это же имеет место при $q' < q$ и для трехчлена $x^2 + px + q'$.

Пример 5. Доказать следующее утверждение: если квадратный трехчлен $x^2 + px + q$ принимает при $x = a$ и при $x = b$ отрицательные значения, то корни этого трехчлена действительны и удовлетворяют условию $|x_2 - x_1| \geq |b - a|$.

Для доказательства заметим, что график трехчлена имеет ветви, уходящие вверх, и заходит в нижнюю полуплоскость (поскольку трехчлен принимает в некоторых точках отрицательные значения). Сле-

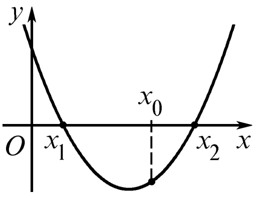


Рис. 438

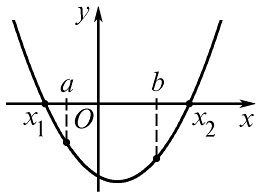


Рис. 439

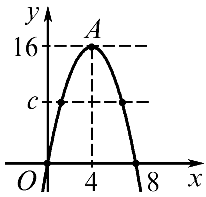


Рис. 440

довательно, этот график пересекает ось абсцисс в двух точках x_1, x_2. Отрицательные значения трехчлен принимает только между этими точками. Так как в точках a, b значения трехчлена отрицательны, то обе точки a, b расположены *внутри* отрезка $[x_1; x_2]$. Этим осуществлен перевод задачи на язык наглядной модели (рис. 439). Из свойств длины отрезка вытекает, что расстояние между точками a и b *меньше*, чем между точками x_1 и x_2. Но это означает (здесь мы осуществляем обратный перевод), что $|a - b| < |x_1 - x_2|$, чем и завершается решение.

Пример 6. Установить, при каких значениях c уравнение $x^4 - 8x^2 + c = 0$ имеет четыре различных корня.

Для решения нужно рассмотреть вспомогательное уравнение $z^2 - 8z + c = 0$, получающееся в результате замены $x^2 = z$. Если z_1, z_2 — корни вспомогательного уравнения, то корни исходного получаются в результате решения совокупности уравнений

$$\left[\begin{array}{l} x^2 = z_1, \\ x^2 = z_2. \end{array}\right.$$

Таким образом, исходное уравнение имеет четыре корня в том и только в том случае, если корни вспомогательного уравнения положительны и различны. Для решения вспомогательного уравнения достаточно найти точки пересечения графиков функций $y = 8x - x^2$ и $y = c$ (рис. 440). Так как вершина параболы $y = 8x - x^2$ находится в точке A (4; 16), а с осью ординат эта парабола пересекается в точке O (0; 0), то, как видно из чертежа, прямая $y = c$ пересекает параболу в двух различных точках с положительными абсциссами при $0 < c < 16$.

Аналогичным рассуждением легко определить, при каких c исходное уравнение имеет три корня, два корня или не имеет ни одного.

75. Наглядность и математическая эстетика

В чем состоит красота математического рассуждения? Почему одно решение задачи оставляет нас лишь спокойно удовлетворенными, тогда как другое вызывает эмоциональный подъем, поражает смелостью замысла и изяществом?

Красивое решение должно нас чем-то удивить, должно быть в чем-то неожиданным. Однако одной необычностью, неожиданностью красоту математического рассуждения не объяснить. Если ученик не заметил стандартного подхода к задаче и — совершенно неожиданно для всех — стал вымучивать решение длинным, окольным путем, то это вызовет только раздражение.

Для того чтобы понять, что еще, помимо необычности, должно характеризовать математическое рассуждение, чтобы оно произвело впечатление красивого, элегантного, мы обратимся к *наглядности*. При решении трудной задачи учащийся должен составить для себя наглядную модель описываемой в задаче ситуации. Трудность нестандартной задачи состоит в том, что для ее решения требуется *самостоятельно* подобрать изоморфную модель (равносильную задачу, сформулированную иначе), которая была бы проще первоначальной, т. е. представляла бы собой ее наглядную модель. Удачный подбор наглядной модели нередко предопределяет успех, а необычность этой модели, ее неожиданность воспринимается как красота и изящество решения. На уроке (или на занятии математического кружка) учащиеся пытаются решить трудную задачу. Вдруг кто-то находит выход из положения и идет к доске, чтобы рассказать решение. Но вместо того, чтобы сразу приступить к решению предложенной задачи, он неожиданно упоминает теорему, которая, казалось бы, никакого отношения к задаче не имеет и очень далека от нее. Настолько далека, что никому и в голову не пришло вспомнить о ней. И вдруг все, слушающие решение, с удивлением замечают, что применение этой теоремы позволяет получить иную версию предложенной задачи, как бы новую ее модель, причем модель наглядную. Простое заключительное рассуждение — и под возгласы «Как красиво!» решение завершается. И чем дальше от тематики задачи отстоит использованная теорема, чем более скрыта она от непосредственного мысленного взора, тем более удивительной кажется вначале мысль о ее применении, тем больше ощущение красоты найденного решения.

Таким образом, красота математического рассуждения складывается из наглядности и неожиданности, что и дает «формулу математической эстетики»:

$$\text{Красота} = \text{наглядность} + \text{неожиданность} =$$
$$= \text{изоморфизм} + \text{простота} + \text{неожиданность}.$$

Рассмотрим несколько иллюстрирующих примеров.

Задача 1. Дан угол K и точка P (рис. 441). Требуется через точку P провести прямую, отсекающую от угла K треугольник заданного периметра.

Решение основано на применении теоремы, которая, казалось бы, очень далека от ситуации, описываемой в условии задачи. Речь идет о теореме, которая утверждает, что касательные MA и MB, проведенные из внешней точки M к окружности, имеют одинаковую длину (где A и B — точки касания, рис. 442).

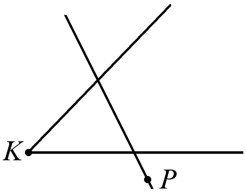


Рис. 441

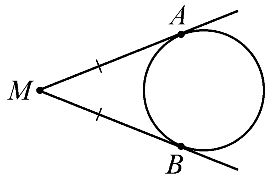


Рис. 442

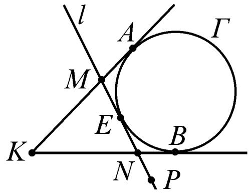


Рис. 443

Теперь перейдем к рассматриваемой задаче. Пусть l — какая-либо проходящая через P прямая, M и N — точки ее пересечения со сторонами угла. Проведем вневписанную окружность $Г$ треугольника KMN (рис. 443). В силу указанной выше теоремы $AM = ME$ и $EN = NB$. Поэтому периметр отсекаемого треугольника KMN равен

$$KM + KN + MN = KM + ME + EN + KN =$$
$$= KM + MA + NB + KN = KA + KB,$$

т. е. этот периметр равен длине двухзвенной ломаной AKB.

Сказанное может быть истолковано следующим образом. Каждому треугольнику, отсекаемому от угла K прямой, проходящей через точку P, соответствует его вневписанная окружность $Г$. Этим устанавливается взаимно однозначное соответствие между отсекаемыми треугольниками и окружностями, вписанными в угол K. При этом периметр отсекаемого треугольника равен длине соответствующей двухзвенной ломаной AKB, т. е. указанное взаимно однозначное соответствие является изоморфизмом относительно этих числовых характеристик. Мы приходим к задаче: *построить окружность $Г$, вписанную в угол K, для которой двухзвенная ломаная AKB (где A и B — точки касания) имеет заданную длину*. Полученная задача не только позволяет найти решение исходной, но и несравненно проще ее, т. е. она представляет собой наглядную модель исходной задачи. Теперь задача допускает простое решение.

Красота решения состоит в том, что эта наглядная модель, связанная с построением вневписанной окружности, является неожиданной.

Задача 2. Доказать, что для любых действительных чисел x_1, x_2, ..., x_n, y_1, y_2, ..., y_n, справедливо неравенство

$$(x_1y_1 + x_2y_2 + ... + x_ny_n)^2 \le (x_1^2 + x_2^2 + ... + x_n^2)(y_1^2 + y_2^2 + ... + y_n^2). \qquad (1)$$

Для доказательства неравенства (1) заметим, что его можно переписать в виде $B^2 - AC \le 0$, где

$$A = x_1^2 + ... + x_n^2, \quad B = x_1y_1 + ... + x_ny_n, \quad C = y_1^2 + ... + y_n^2. \qquad (2)$$

Запись $B^2 - AC \leq 0$ наводит на мысль рассмотреть квадратный трехчлен $Az^2 + 2Bz + C$, для которого выражение $B^2 - AC$ является дискриминантом. Неположительность дискриминанта означает, что этот трехчлен принимает значения одного знака и притом *неотрицательные* (поскольку $A \geq 0$). Итак, мы получаем эквивалентную задачу: *доказать, что квадратный трехчлен* $Az^2 + 2Bz + C$ *с коэффициентами* (2) *принимает лишь неотрицательные значения.* Остается заметить, что этот трехчлен можно записать в виде

$$(x_1^2 + \ldots + x_n^2)z^2 + 2(x_1y_1 + \ldots + x_ny_n)z + (y_1^2 + \ldots + y_n^2),$$

а затем в виде

$$(x_1^2z^2 + 2x_1y_1z + y_1^2) + \ldots + (x_n^2z^2 + 2x_ny_nz + y_n^2) =$$
$$= (x_1z + y_1)^2 + \ldots + (x_nz + y_n)^2,$$

после чего неотрицательность всех его значений становится очевидной.

Успех этого решения был определен отдаленным сходством доказываемого неравенства с видом дискриминанта квадратного трехчлена. Это навело на мысль использовать изоморфизм между множеством троек $(A; B; C)$ действительных чисел с условием $B^2 \leq AC$ и множеством квадратных трехчленов с условием неотрицательности всех значений трехчлена. Этот изоморфизм привел к эквивалентной задаче, которая оказалась очень простой, а неожиданность подмеченной связи со свойствами квадратного трехчлена приводит к ощущению красоты решения. Изящество этого решения особенно заметно, если перед тем как познакомиться с ним, читатель испробовал разные попытки доказательства неравенства (1). Например, можно было раскрыть скобки в (1), попытаться тем или иным способом произвести группировку слагаемых... И хотя на этом пути тоже можно при некоторой настойчивости получить доказательство, но насколько мучительнее напролом пробираться через «дремучий лес», чем двигаться к цели по комфортабельной дороге на «Берегу наглядной модели»!

З а д а ч а 3. К трем окружностям различных радиусов проведены общие внешние касательные (рис. 444). Доказать, что три точки пересечения этих касательных, взятых попарно, лежат на одной прямой.

Для решения рассмотрим три шара, для которых взятые окружности являются большими окружностями. Тогда плоскость α, в которой лежат окружности, проходит через центры шаров. Для каждых двух из этих шаров рассмотрим содержащий их конус, касающийся обоих шаров по окружности (рис. 445). Плоскость β, касающаяся «снизу» всех трех шаров, проходит через одну образующую каждого конуса и потому содержит вершину конуса. Таким образом, все три вершины конусов (точки A, B, C) лежат в плоскости β. А так как все они лежат и в плоскости α, то точки A, B, C лежат на одной прямой $\alpha \cap \beta$.

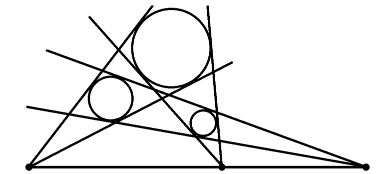


Рис. 444

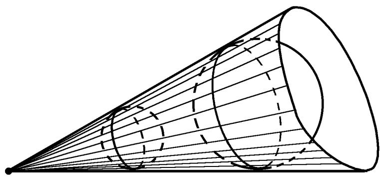


Рис. 445

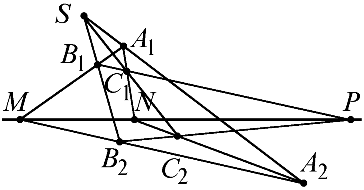


Рис. 446

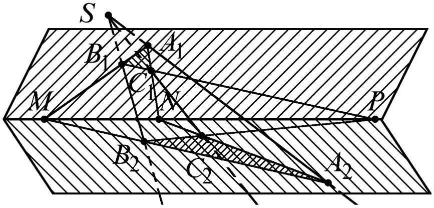


Рис. 447

Успех этого решения, его неожиданность и изящество связаны с выходом в пространство. Еще один красивый пример выхода в пространство дает теорема Дезарга. Мы ее рассмотрим в виде задачи.

З а д а ч а 4. Два треугольника $A_1B_1C_1$ и $A_2B_2C_2$, соответствующие стороны которых не параллельны, расположены так, что прямые A_1A_2, B_1B_2, C_1C_2 пересекаются в одной точке S. Через M, N, P обозначены точки пересечения прямых A_1B_1 и A_2B_2, A_1C_1 и A_2C_2, B_1C_1 и B_2C_2 (рис. 446). Доказать, что точки M, N, P расположены на одной прямой.

Для решения предположим сначала, что треугольники $A_1B_1C_1$ и $A_2B_2C_2$ расположены в пространстве, причем они лежат в двух непараллельных плоскостях α и β (рис. 447). В этом случае решение не представляет труда.

В самом деле, прямые A_1B_1 и A_2B_2, лежащие в одной плоскости A_1B_1S, пересекаются в некоторой точке M. Эта точка лежит в плоскости α (поскольку $A_1B_1 \subset \alpha$) и в плоскости β (поскольку $A_2B_2 \subset \beta$). Следовательно, $M \in \alpha \cap \beta$. Аналогично $N \in \alpha \cap \beta$ и $P \in \alpha \cap \beta$, т. е. все три точки M, N, P расположены на одной прямой $\alpha \cap \beta$. Чтобы завершить решение, достаточно заметить, что плоский чертеж, изображенный на рис. 446, представляет собой *проекцию* пространственного чертежа (рис. 447).

Читатель сам сможет найти те задачи, решения которых показалось ему «красивым», и убедиться, что красота решения укладывается в рамки приведенной выше «формулы математической эстетики».

76. Аналогия — общность аксиоматики

На уроках нередко можно услышать: «эта задача решается аналогично предыдущей». Кое-где выше мы также говорили, что решение задачи «аналогично» решению предыдущей. Старшеклассники понимают, что при несомненном наличии индивидуальных черт обе задачи обладают, тем не менее, похожими свойствами. Если бы речь шла лишь о раскрытии смысла термина «аналогия», то, вероятно, этим следовало бы ограничиться. Но аналогия нас прежде всего интересует как метод мышления. Мы нередко делаем умозаключения «по аналогии»; удачно подмеченная аналогия иной раз помогает найти решение трудной задачи.

Учитель при изучении темы о векторах говорит, что сложение векторов обладает свойствами, *аналогичными* свойствам сложения чисел. В чем же состоит эта аналогия?

Как сложение чисел, так и сложение векторов обладает свойствами коммутативности (переместительности) $x + y = y + x$ и ассоциативности (сочетательности) $x + (y + z) = (x + y) + z$. Иными словами, эти соотношения будут справедливыми и в том случае, если x, y, z — *числа* (например, целые), и в том случае, если x, y, z — *векторы*. Кроме того, в обоих случаях (для чисел и для векторов) существует нулевой элемент, обладающий свойством $x + 0 = x$ для любого x. Наконец, для любого x (числа или вектора) существует противоположный элемент $-x$, обладающий свойством $x + (-x) = 0$. Такими же свойствами обладает операция сложения в множестве всех рациональных чисел, в множестве всех действительных (или комплексных) чисел, в множестве всех многочленов и в ряде других случаев.

Таким образом, если принять эти свойства сложения за аксиомы, то мы обнаружим, что во всех рассмотренных моделях (в множестве действительных чисел, векторов, многочленов и т. д.) эти аксиомы выполняются. Иными словами, указанные модели *аналогичны* друг другу, и эта аналогия заключается в том, что они удовлетворяют одной и той же системе аксиом, а именно, системе аксиом *коммутативной группы*.

Заметим, что рассмотренные модели вовсе не изоморфны друг другу, а только аналогичны. Например, в множестве действительных чисел имеются неравенства, а в множестве комплексных чисел или многочленов ввести неравенства не удается. Имеются и другие свойства, отличающие эти модели друг от друга. Однако все те свойства сложения, которые вытекают из перечисленных выше аксиом, наблюдаются во всех моделях; в отношении этих свойств рассматриваемые модели аналогичны.

Например, в каждой из этих моделей можно записывать сумму нескольких слагаемых без скобок, т. е. писать

$$(a + b) + (c + d) = a + b + c + d,$$
$$a + (b + (c + d)) = a + b + c + d,$$
$$a + ((b + c) + d) = a + b + c + d,$$

где справа имеется в виду *последовательное* выполнение сложения $((a+b)+c)+d$. Это *вытекает* из аксиом коммутативной группы, и в этом — аналогичность всех рассмотренных моделей друг другу.

В качестве еще одного свойства отметим, что можно переносить слагаемые из одной части равенства в другую с изменением знака, т. е. если $a+b=c$, то $a=c+(-b)$ и т. п. Это также одно из следствий аксиоматики коммутативной группы, и потому указанное свойство имеет место *во всех* моделях коммутативной группы. И в этом отношении все модели (числа, векторы, многочлены) также аналогичны друг другу.

Заметим, что указанные свойства (скажем, возможность переносить слагаемые из одной части равенства в другую с изменением знака) в какой-то степени обосновываются в начальных классах, когда учащиеся знакомы только с *рациональными* числами. Но в дальнейшем (при изучении многочленов, векторов, комплексных чисел и т. п.) обоснование уже не проводится. Практика показывает, что учащиеся не умеют обосновывать справедливость этих свойств сложения. Необходимо предельно ясно объяснить, что эти свойства *вытекают из аксиом* коммутативной группы, так что их достаточно один раз доказать, а затем лишь проверять выполнение аксиом коммутативной группы. Так аналогия помогает более глубокому осмыслению курса алгебры.

Итак, две модели аналогичны, если они удовлетворяют одной и той же системе аксиом. Именно, они аналогичны в отношении всех тех свойств, которые описываются этой системой аксиом. Короче говоря, аналогия — это общность аксиоматики. Может, однако, показаться, что это основное положение не дает общего описания всех ситуаций, в которых можно говорить об аналогичности явлений. Действительно, аналогию мы нередко усматриваем и тогда, когда трудно говорить о точном описании явлений и, тем более, об их аксиоматизации. И тем не менее указанное определение аналогии является общим.

Правда, нередко общность аксиоматического описания не сразу усматривается. В этом случае можно говорить о «неявной» аналогии: мы имеем в виду какой-то набор свойств, в отношении которых рассматриваемые явления (которые еще следует превратить в модели для их точного описания) «одинаковы» или «аналогичны». Иными словами, если эти свойства точно описать аксиомами (а это всегда возможно), то мы и получим аксиоматику, которой удовлетворяют обе рассматриваемые модели явления, т. е. получим точное осмысление того, почему мы находим эти явления аналогичными.

Рассмотрим следующий пример. В курсе алгебры доказывается, что если A и B — две произвольные точки числовой оси, причем точка A имеет координату x, а точка B — координату x', то расстояние между точками A и B равно $|x'-x|$. Доказательство этого утверждения не излагается в учебниках полностью, а выглядит примерно так.

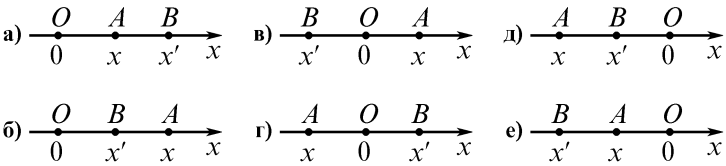


Рис. 448

Рассмотрим случай, когда обе точки A, B расположены на положительной полуоси, причем точка B расположена дальше от начала O, чем точка A, т. е. $x' > x > 0$ (случай а) на рис. 448). В этом случае $OA = x$, $OB = x'$, $AB = OB - OA = x' - x$, кроме того, $|x' - x| = x' - x$ (так как число $x' - x$ положительно); следовательно, $AB = |x' - x|$. Аналогичное доказательство проводится и при ином расположении точек A, B (случаи б) – е) на рис. 448).

Здесь говорится о том, что доказательство формулы проводится при различных случаях расположения точек A и B с помощью *аналогичных* рассуждений. Как и в рассмотренных ранее примерах, аналогия заключается в том, что эти рассуждения можно описать одинаковыми аксиомами. Вероятно, эту общность аксиоматики разные люди были бы склонны выражать по-разному, т. е. мы имеем здесь дело с неявной аналогией, точно выразить (аксиоматизировать) которую можно разными путями.

Например, ее можно выразить следующими аксиомами:

1) если точка A имеет координату x, то $OA = \pm x$ (знак «+» или «–» выбирается так, чтобы правая часть была неотрицательной);

2) для любых точек A, B на числовой прямой $AB = \pm OA \pm OB$ (знаки выбираются в зависимости от расположения точек A и B относительно O);

3) для любых чисел x, y имеем $|x - y| = \pm x \pm y$ (при соответствующем выборе знаков).

Во всех случаях а) – е) на рис. 449 эти аксиомы выполняются (разница только в выборе знаков); при учете же этих аксиом рассуждение будет во всех случаях одинаковым. Математики могут возразить, что, например, указанное выше положение 3 аксиомой обычно не признается, а выводится из других фактов, изучаемых в курсе алгебры. Это верно. Однако нас в данном случае интересует лишь то, какие положения в этих доказательствах используются в качестве заранее известных фактов. Они и служат в рассматриваемом случае исходными «аксиомами». Вообще, «локально» (например, при решении некоторой задачи) мы можем те факты, которые нужны для решения (в частности, изученные ранее теоремы), считать при решении этой задачи как бы «аксиомами» — в пределах рассуждения, связанного с решением задачи. Такие «расширенные» системы аксиом вовсе не противоречат духу математики. В рамках такой расширенной аксиоматики аналогичность двух задач становится понятной — как общность системы аксиом.

Но вернемся к рассмотренной задаче (рис. 448). В наличии общей аксиоматики и заключается «аналогичность» рассуждений. Ученик, которому изложили полное доказательство в случае а) и указали эту общую аксиоматику, легко сможет провести рассуждение и в другом случае. Более того, даже если сообщить только рассуждение для одного случая и сказать, что в других случаях «рассуждение аналогично», но не сообщать рассмотренную выше общую аксиоматику, учащийся сможет самостоятельно провести по аналогии рассуждения и для других случаев (причем при помощи почти тех же слов в рассуждении). Иными словами, даже неявная аналогия является важным психологическим фактором. И если предложить человеку, проведшему доказательство по аналогии, осмыслить, проанализировать, какие черты сходства помогли ему провести рассуждение в ином варианте расположения точек A и B, он перечислит эти черты сходства, т. е. сделает аналогию явной.

Заметим, что аналогия существенно используется при формировании абстрактных понятий. Рассмотрим в качестве примера наиболее простой вид абстракции — отождествление. Мы рассматриваем некоторый класс предметов, которые с определенной точки зрения *аналогичны* в отношении существенных, главных черт и отличаются друг от друга лишь несущественными, второстепенными свойствами. Отвлекаясь от второстепенных черт и отождествляя рассмотренные предметы между собой в отношении их главных свойств, т. е. отождествляя в своем сознании эти аналогичные предметы, мы получаем представление об общем абстрактном прототипе предметов рассматриваемого класса. Остается закрепить вновь возникшее абстрактное понятие единым обобщающим термином.

Однако, хотя при образовании нового понятия с помощью абстракции отождествления мы объединяем в своем сознании «аналогичные» между собой предметы, но аналогия эта почти всегда неявная.

В самом деле, если, например, речь идет о вилке, то мы не утруждаем себя требованием указать аксиоматику, на основе которой можно было бы о данном предмете судить, вилка это или не вилка. И тем не менее если на столе лежит целый ряд предметов (в том числе несколько вилок, с двумя и с бо́льшим числом зубьев, большие и маленькие, стальные или пластмассовые), то мы безошибочно отделим предметы, которые следует считать вилками. Все эти предметы аналогичны друг другу в отношении их существенных черт. Аналогия здесь неявная, но при желании можно сделать ее явной, т. е. выписать общую аксиоматику для этих предметов. Описание термина в толковых словарях в большинстве случаев можно рассматривать как аксиоматическое определение понятия (поскольку там дается описание основных, характерных черт), но эта аксиоматика обычно не расчленяется на отдельные положения (аксиомы).

В заключение отметим весьма важную для математики аналогию в отношении *размерности*. Как мы говорили во вводной беседе, многомерные пространства являются важным инструментом исследования не только в самой математике, но в ряде смежных наук — физике, химии, математической экономике и других.

Наблюдая свойства двумерных фигур (на плоскости) и трехмерных тел (в пространстве), математик «по аналогии» пытается судить о свойствах многомерных тел, постепенно вырабатывая «многомерную интуицию». И хотя переход от двух измерений к трем измерениям очень важен, но он далеко не достаточен для осмысления многомерных фактов «по аналогии». Подлинной основой такой аналогии служит именно *общность аксиоматики*. Имеется четко сформулированная система аксиом многомерного евклидова пространства, для которой двумерное пространство (плоскость), трехмерное пространство, четырехмерное и, вообще, n-мерное пространство являются *моделями*, причем неизоморфными между собой. Многомерная интуиция, подкрепленная этой прочной аксиоматической базой, и позволяет доказывать важные факты многомерной геометрии, находящие многочисленные приложения в математике и смежных науках. И аналогия, основанная на общности аксиоматики, играет здесь первостепенную роль.

Более подробный разговор об аксиоматике многомерных пространств мы поведем в другой раз, здесь же ограничимся одним примером, показывающим пользу многомерной геометрической аналогии при решении чисто алгебраической задачи.

Пример. Сумма n чисел равна единице. Каковы должны быть эти числа, чтобы сумма их квадратов была наименьшей?

Для решения обратимся сначала к случаю $n = 2$ и ответим на интересующий нас вопрос, используя геометрическую наглядную модель. При $n = 2$ рассматриваются числа x, y, удовлетворяющие условию $x + y = 1$, и требуется найти, в каком случае сумма квадратов $x^2 + y^2$ будет наименьшей.

Уравнение $x + y = 1$ определяет на координатной плоскости прямую l (рис. 449). Рассмотрим окружность S с центром в начале координат, которая касается этой прямой. Точку касания обозначим через A (рис. 450). Если точка M прямой l отлична от A, то она лежит вне окружности S, и потому OM больше радиуса r этой окружности, т. е. координаты x, y точки M удовлетворяют неравенству $x^2 + y^2 > r^2$. Если же M совпадает с A, то сумма $x^2 + y^2$ равна r^2, т. е. именно для точки A эта сумма принимает наименьшее значение. Точка A имеет координаты $x = y = \frac{1}{2}$. Это и есть решение поставленной алгебраической задачи при $n = 2$.

Пусть теперь $n = 3$. Уравнение $x + y + z = 1$ определяет в пространстве *плоскость* L (рис. 451). Рассмотрим сферу S с центром в начале координат O, касающуюся этой плоскости в некоторой точке A. Для любой точки $M \in L$, отличной от A, ее расстояние от точки O больше радиуса r сферы S, т. е. $OM^2 > r^2$, и потому $x^2 + y^2 + z^2 > r^2$. Если же

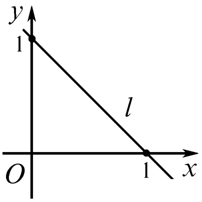


Рис. 449

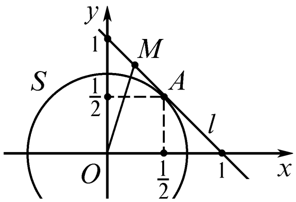


Рис. 450

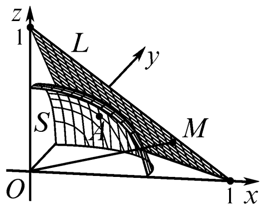


Рис. 451

M совпадает с A, то $x^2 + y^2 + z^2 = r^2$. Таким образом, именно для точки A сумма $x^2 + y^2 + z^2$ принимает наименьшее значение. Заметим теперь, что точка A имеет *равные* координаты: $x = y = z$. Это следует из того, что при повороте пространства вокруг точки O, переставляющем оси координат: $x \to y$, $y \to z$, $z \to x$, как плоскость L, так и сфера S переходят в себя, а потому их общая точка A остается неподвижной. Итак, для точки A имеем $x + y + z = 1$, причем $x = y = z$, и потому точка A имеет координаты $x = y = z = \frac{1}{3}$. Это и есть решение поставленной алгебраической задачи для $n = 3$.

Рассмотрим, наконец, произвольное n. Рассуждения будем вести в n-мерном пространстве, каждая точка которого имеет координаты $(x_1; x_2; ...; x_n)$. Уравнение $x_1 + x_2 + ... + x_n = 1$ определяет в этом пространстве «плоскость» L, имеющую размерность $n - 1$ (например, при $n = 3$, т. е. в трехмерном пространстве, такое уравнение определяет плоскость, имеющую размерность 2, т. е. на единицу меньшей размерности, чем все пространство). Математики называют «плоскости», имеющие размерность $n - 1$, *гиперплоскостями* в n-мерном пространстве. Рассмотрим *гиперсферу* S с центром в начале координат O, касающуюся гиперплоскости L в некоторой точке A. Гиперсфера состоит из всех точек, находящихся от O на одном и том же расстоянии r; число r называется радиусом этой гиперсферы.

Всякая точка M гиперплоскости L, кроме A, лежит вне гиперсферы S, т. е. находится от начала координат O на расстоянии, большем r. Если же M совпадает с A, то $OM = r$. Следовательно, сумма $x_1^2 + x_2^2 + ... + x_n^2$, т.е. *квадрат расстояния* точки M с координатами $(x_1; x_2; ...; x_n)$ от начала координат, для всех точек $M \in L$ принимает наименьшее значение при $M = A$. Заметим теперь, что все координаты точки A равны между собой: $x_1 = x_2 = ... = x_n$ (поскольку поворот пространства, переставляющий оси: $x_1 \to x_2$, ..., $x_{n-1} \to x_n$, $x_n \to x_1$, переводит гиперплоскость L в себя и гиперсферу S тоже в себя, а потому оставляет точку A неподвижной). Следовательно, точка A имеет координаты $x_1 = x_2 = ... = x_n = \frac{1}{n}$. Это и есть решение поставленной алгебраической задачи: при $x_1 + x_2 + ... + x_n = 1$ сумма квадратов $x_1^2 + x_2^2 + ... + x_n^2$ принимает наименьшее значение, когда $x_1 = x_2 = ... = x_n = \frac{1}{n}$.

Разумеется, это проведенное по аналогии геометрическое решение можно признать корректным лишь в том случае, если читатель владеет некоторыми понятиями и фактами n-мерной геометрии. И это дает хорошую иллюстрацию пользы «аналогии по размерности».

77. Прогнозирование

В заключение остановимся еще на одном моменте, который играет важную роль в процессе поиска решения. Допустим, во время раздумий над возможными путями решения задачи нам пришел в голову некоторый «шажок мысли». Правильным ли он является? Приближает это преобразование к нахождению решения или же это ложный путь, уводящий в сторону от правильного направления? Критерием в этом вопросе является прогнозирование, т. е. предвидение результата, получаемого в процессе анализа, синтеза, обобщения — мы прикидываем, стали ли мы ближе к нахождению решения, куда продвинулись с помощью сделанного шага. Происходит маленький «инсайтик»: «Ага, эврика! Кажется, так более вероятно получить решение, чем в предыдущей ситуации».

Маленький инсайтик, о котором шла речь, представляет собой *действие*, позволяющее ориентировочно судить о приближении к решению или, напротив, об удалении от возможного пути решения и вследствие этого о целесообразности отбрасывания этого хода. Благодаря этому часто удается не перебирать все узлы полного дерева ходов, а двигаться лишь по части узлов и сравнительно быстро придти к решению. Без проведения такого контрольно-оценочного действия происходит тупой перебор, бесконечные преобразования задачи, напоминающие блуждание в потемках, а не целенаправленный поиск решения.

Итак, формирование умения прогнозировать, предвидеть результаты, к которым приведет каждый отдельный шажок мысли, является важным компонентом развития мышления. С этой целью на уроках математики при обсуждении идеи решения, когда кто-либо из учащихся предлагает воспользоваться той или иной формулой, теоремой, тождественным преобразованием, целесообразно добиваться того, чтобы он обосновал разумность своего предложения и хотя бы в общих чертах указал, к чему оно приведет.

Пример 1. Доказать тождество

$$(1 + 2\cos 7x)\sin 3{,}5\,x = \sin 10{,}5\,x. \qquad (*)$$

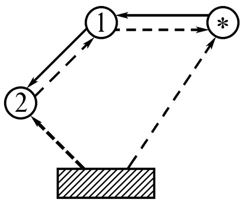


Рис. 452

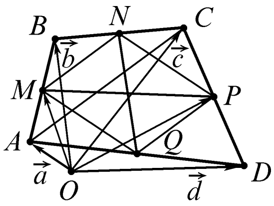


Рис. 453

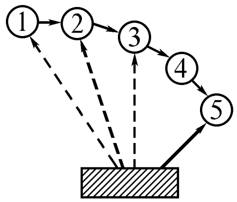


Рис. 454

Обсудим учебную ситуацию, возникшую в классе при решении этой задачи. Один из учащихся предложил обозначить $3{,}5x$ через a; тогда $7x = 2a$, $10{,}5x = 3a$ и слева будет стоять некоторое выражение от a и $2a$, а справа $\sin 3a$. После этого второй ученик посчитал целесообразным разложить правую часть по формуле $\sin(2a + a)$ и объяснил, что при этом уменьшится число аргументов: не будет $3a$, а слева и справа останутся a и $2a$. Учащиеся считали, что оба шага, выполненные один за другим, могут приблизить к решению, хотя еще не все детали предстоящих преобразований им были достаточно ясны.

Вызванный к доске учащийся реализовал путь, указанный в процессе прогнозирования:

$$(1 + 2\cos 2a)\sin a = \sin 3a, \tag{1}$$

$$(1 + 2\cos 2a)\sin a = \sin 2a\cos a + \cos 2a\sin a. \tag{2}$$

Дальше решение легко было доведено до конца.

Прогнозирование имело форму обратимого анализа, схема которого представлена на рис. 452; жирная штриховая стрелка изображает инсайт, т. е. озарение, которое испытали все учащиеся класса (эврика! — в результате этих преобразований мы приблизимся к решению).

Прогнозирование нередко сопровождается использованием наглядной модели и аналогии. В следующем примере мы рассматриваем ситуацию, когда используется векторная наглядная модель, о применении которой уже говорилось выше.

Пример 2. Доказать, что в произвольном четырехугольнике сумма квадратов длин диагоналей равна удвоенной сумме квадратов длин средних линий (т. е. отрезков, соединяющих середины противоположных сторон четырехугольника).

Во время обсуждения один учащийся сказал, что поскольку речь идет о квадратах длин отрезков, то надо применить скалярное произведение векторов. Однако учитель посчитал такое прогнозирование недостаточно определенным. Другой учащийся добавил, что надо обозначить через $\vec a$, $\vec b$, $\vec c$, $\vec d$ векторы, идущие от некоторой точки O к вершинам четырехугольника $ABCD$ (рис. 453). Тогда векторы $\vec a - \vec c$ и $\vec b - \vec d$ направлены по диагоналям, и можно легко получить векторы, идущие к серединам сторон, а через них выразить векторы, направленные по средним линиям. Учитель указал, что оба предложения, взятые вместе, составляют хороший прогноз решения. Оно выглядит следующим образом:

$$\overrightarrow{OM} = \frac{\vec{a} + \vec{b}}{2}, \ \overrightarrow{OP} = \frac{\vec{c} + \vec{d}}{2}, \ \overrightarrow{PM} = \frac{\vec{a} + \vec{b}}{2} - \frac{\vec{c} + \vec{d}}{2}.$$

Аналогично выражается $\overrightarrow{ON}$. Доказываемое утверждение записывается в виде

$$(\vec{a} - \vec{c})^2 + (\vec{b} - \vec{d})^2 = 2\left(\frac{\vec{a} + \vec{b}}{2} - \frac{\vec{c} + \vec{d}}{2}\right)^2 + 2\left(\frac{\vec{b} + \vec{c}}{2} - \frac{\vec{a} + \vec{d}}{2}\right)^2$$

и его справедливость легко проверяется.

Третий учащийся предложил другую идею решения. В четырехугольнике $MNPQ$ сторона MN — средняя линия треугольника ABC (и аналогично для других сторон). Это замечание связало стороны четырехугольника $MNPQ$ и диагонали исходного четырехугольника $ABCD$. Четвертый учащийся заметил, что отрезки MN и PQ равны и параллельны, т. е. $MNPQ$ — параллелограмм. Наконец, пятый добавил: средние линии PM и QN исходного четырехугольника служат диагоналями параллелограмма $MNPQ$, и потому уместно применить теорему о сумме квадратов диагоналей параллелограмма. Эти соображения, вместе взятые, содержат не только прогноз, но почти полное второе решение задачи.

Заметим, что прогнозирование, как правило, сопровождается «маленькими промежуточными озарениями». Так, задача 1, рассмотренная в п. 67, фактически имеет схему рассуждения, показанную на рис. 454, на котором пунктирные стрелки означают акты прогнозирования, являющиеся промежуточными инсайтами. В самом деле, при переходе от (1) к (2) учащийся понял (Эврика!), что умножение и деление на $\sin 20°$ позволяет применить формулу синуса двойного угла. А получив (3), учащийся совершает еще одно открытие — опять можно применить ту же формулу!

Следует отметить, что прогнозирование связано с приобретением опыта решения задач, умением творчески мыслить. Могут, однако, встретиться случаи, когда осуществить прогнозирование с целью выбрать из нескольких возможных логических шажков какой-либо один (более предпочтительный) не удается. В таких случаях не остается ничего лучшего, как произвести перебор этих возможных логических шажков, т. е. искать решение, испробовав один шажок, а если это не приводит к цели, испробовать второй, третий ... Так, в завершающей части решения задачи 13 в п. 72 было сказано: «остается выбрать из цифр $a = 1, ..., 9$ такую, для которой ...». Это означает перебор, и его в этом случае было осуществить проще, чем искать обоснованное прогнозирование для выбора из цифр 1, ..., 9 более предпочтительной. Впрочем, перебор значений $a = 1, ..., 9$ можно было осуществить и раньше — на основании соотношений (2), (4).

И все же перебор — лишь крайний случай. Основой выбора очередного логического шажка и всего пути решения (с использованием анализа, синтеза, обобщения, аналогии) является именно прогнозирование, а не перебор. При индивидуальном решении задач учащийся молча, «про себя» осуществляет прогноз и воплощает его в решении, а иногда даже как бы не замечает скрытую работу мысли на этапе прогноза. Обсуждение процесса поиска решения и, в частности, прогнозирование решения при активном участии всего класса — мощное средство развития навыков логического мышления учащихся.

Пример 3. Рассмотрим следующее игровое задание, о котором уже шла речь выше.

Вы едете на трамвае или автобусе и приобретаете билет (или компостирует талон), на котором имеется шестизначный номер. Скажем, достался билет с номером 327147. Ставится задача расставить между цифрами этого билета такие знаки действий, чтобы при их выполнении получилось 100. При этом между некоторыми цифрами можно не ставить никаких знаков действий, соединяя две рядом стоящие цифры в одно двузначное число.

Как подходить к решению этой задачи? Если поставить в первом промежутке знак плюс, то этим осуществляется некоторое преобразование задачи: уже получено число $3+2=5$ и остается составить 95 с помощью оставшихся четырех цифр. Это не удается, приходится пробовать вставлять другие знаки. Такие преобразования могут продолжаться долго, нудно и безуспешно. Но вдруг появляется новая идея: $(3+2)\cdot(\ldots)=100$. И сразу же осуществляется контрольно-оценочное действие: «Ага, так мы приближаемся к решению — ведь нужно лишь из оставшихся четырех цифр составить 20, а это может оказаться намного легче!» Например, это можно осуществить так: $(3+2)(-7-1+4\cdot7)$.

В результате возникает представление о структурировании задачи. Так как имеют место равенства $100=5\cdot20$, $100=4\cdot25$, $100=10\cdot10$, то можно попытаться разбить шесть цифр на две группы так, чтобы, скажем, в одной группе получилось 5, а в другой 20. Эта структурированность дает новый подход к обзору дерева всех возможных ходов, дает хорошую ориентировку в процессе решения. И при появлении нового билетика мы уже будем искать — нет ли возможности получить (справа или слева) 4, 5 или 10.

С точки зрения поиска пути решения эта структуризация облегчает осуществление прогнозирования и получение вывода о разумности сделанного хода. Однако возможности структуризации этим не ограничиваются. Еще одна идея структуризации состоит в том, что число 71 достаточно близко к 100 и нужно лишь распорядиться оставшимися четырьмя цифрами так, чтобы получилось 29. О таком способе решения уже шла речь выше; возможно и другое решение, реализующее ту же идею: $3-2+71+4\cdot7$.

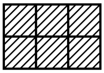


Рис. 455

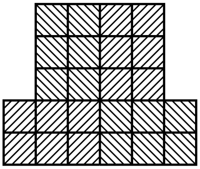


Рис. 456

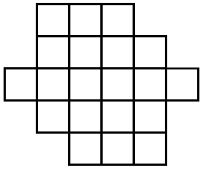


Рис. 457

Результат состоит не только в приобретении навыков выполнения ориентировочных действий, но также (в математическом плане) в развитии навыков числовой комбинаторики.

Пример 4. В задаче складывания целой фигурки из кирпичиков, имеющих форму уголков (см. задачу 2 в п. 67), можно заметить, что из двух уголков удается сложить прямоугольничек 2×3 (рис. 455). Теперь этот прямоугольничек становится *зна́ком*, *символом*, с помощью которого можно попытаться выложить всю заданную фигуру. В некоторых случаях это удается (рис. 456). Это означает, что такой прямоугольничек выступает агрегированным блоком, символом, с помощью которого задача решается в свернутом виде. Иначе говоря, пытаясь заполнить данную фигуру этими блоками (как на рис. 456), мы не разбивает *каждый* блок на уголки, а мыслим символами, т. е. *свертками-блоками*. Уже не нужно разбивать блоки на уголки, поскольку возможность заполнения фигуры такими блоками доказывает возможность заполнения и уголками, т. е. задача принципиально решена. Вместе с тем эта знаковая модель не является изоморфной первоначальной задаче. Если некоторая фигура не может быть заполнена блоками 2×3, то это еще не означает, что и первоначальная задача о заполнении фигуры уголками неразрешима (например, фигура на рис. 385 не может быть заполнена блоками 2×3, но заполнение ее уголками возможно). Иными словами, возможность заполнения фигуры агрегированными блоками является *достаточным*, но не необходимым условием возможности заполнения фигуры уголками.

Пример 5. Рассмотрим теперь задачу заполнения уголками другой фигуры (рис. 457). Эта фигура составлена из 20 единичных квадратиков (что сразу не видно, т. е. нужно мысленно расчленить фигуру на квадратики и произвести подсчет). В этом случае для решения задачи также можно использовать знаковую символику. Задача агрегируется, т. е. мы уже не рассматриваем фигуру с точки зрения различных ее геометрических свойств. Из всех свойств выступает только одно — *площадь* (т. е. число единичных квадратиков, из которых можно составить эту фигуру). В то же время уголок составлен из трех единичных квадратиков, т. е. он имеет площадь 3.

Теперь на уровне этой новой модели мы получаем новую постановку задачи: можно ли составить фигуру, имеющую площадь 20, из нескольких уголков, каждый из которых имеет площадь 3? И если мы в чистом виде вычленили эту модель, то легко получаем ответ:

составить эту фигуру из уголков нельзя, поскольку 20 не делится на 3. Иными словами, делимость площади на 3 является *необходимым* критерием возможности заполнения фигуры уголками. Однако этот необходимый критерий не является достаточным.

Такой способ получения доказательства невозможности нередко используется в математике. Именно, пусть рассматривается некоторый класс $\boldsymbol{M}$ математических объектов и среди них выделена часть объектов, обладающих интересующим нас (и достаточно сложным) свойством P. Рассматривается задача о том, как определить, обладает ли конкретно заданный объект класса $\boldsymbol{M}$ этим свойством P.

Если свойство P проверяется сложно, то вместо такой проверки нередко применяется некоторый необходимый критерий. Для этого подбирается другое свойство Q, более просто проверяемое. Если удается доказать, что все объекты, обладающие свойством P, неизменно обладают и свойством Q, то для «доказательства невозможности», т. е. для установления того факта, что взятый конкретный объект не обладает свойством P, достаточно проверить, что он не обладает свойством Q.

Именно так мы поступили в примере 5. Свойство P состоит в возможности заполнения фигуры уголками, а свойство Q заключается в делимости площади фигуры на 3. Для фигуры на рис. 457 свойство Q проверяется просто, и этот критерий удобен для доказательства невозможности: если свойство Q не выполняется, то и свойство P не имеет места. И идея «проверить делимость на 3» воспринимается в виде свернутой рекомендации, включающей в себя несколько действий, обозначенных одним символом «делимость на 3». Таким образом, применение необходимого критерия часто может служить мощным средством прогнозирования.

Вот еще пара задач, связанных с остатками при делении.

Пример 6. Капризная девочка нарисовала что-то на листе бумаги, но рисунок ей не понравился, и она разорвала его на четыре части. Потом одну из частей она разорвала еще на четыре части, потом один из кусочков еще на четыре части, потом еще и еще ... Через некоторое время около нее лежала куча кусочков бумаги. Могло ли оказаться, что число кусочков в этой куче равно 2001?

Решение опять связано с делимостью на 3. Сначала был один лист бумаги. После первого разрыва листа стало четыре куска, т. е. число частей увеличилось на 3. После следующего разрыва вместо одного из кусков появилось четыре меньших, т. е. опять число кусков увеличилось на 3. Продолжая таким образом, мы заключаем, что каждый раз число кусочков увеличивалось на 3. Поэтому число кусочков в куче получалось равным 1, 4, 7, ..., т. е. все время получались числа, дающие при делении на 3 остаток 1. А так как число 2001 при делении на 3 не дает единицу в остатке, то никогда число кусочков в куче не будет равным 2001. Мы опять получаем доказательство невозможности!

Пример 7. Змей-горыныч имел три головы. Илья Муромец в битве со Змеем срубал ему головы. Но каждый раз на месте срубленной головы тут же вырастало пять новых. Еще голову срубит — опять на ее месте вырастает пять новых. Могло ли случиться, что в некоторый момент битвы у Змея оказалась 2001 голова?

И здесь есть необходимый критерий: всегда число голов дает при делении на 4 остаток 3. А так как число 2001 не обладает этим свойством, то ни в какой момент битвы число голов у Змея не равнялось 2001.

78. Несколько слов о математической интуиции

Формирование навыков прогнозирования означает развитие *математической интуиции* и общей математической культуры. Человек с развитой математической интуицией быстрее и легче решает задачи, даже трудные, причем предлагаемые им решения чаще, чем у других, бывают необычными, красивыми. Математическая культура и интуиция предполагают наличие большого кругозора, умение по малейшим, незаметным признакам находить аналогии с другими (иногда очень далекими) областями математики, умение легко усматривать новые формулировки задач на другом языке, находить различные модели задачи, видеть среди этих моделей более простые, более наглядные. Интуиция — высшая ступень прогнозирования. Несколько примеров мы имели выше при разговоре о наглядности и аналогии. Ряд интересных дальнейших примеров можно найти в книгах «Как решить задачу», «Математика и правдоподобные рассуждения», «Математическое открытие», написанных американским математиком, выходцем из Венгрии, Д. Пойа.

Отметим еще один аспект, также тесно связанный с математической культурой и интуицией. Речь идет о том, чтобы выяснить, что значит свободно владеть некоторым математическим понятием. По мысли известного российского математика Д. К. Фаддеева, для хорошего владения математическим понятием недостаточно хорошо знать его определение и конгломерат основных теорем, в которых это понятие применяется. Существенным является также владение некоторым достаточно богатым набором примеров, с разных сторон иллюстрирующих смысл этого понятия. По существу, *конечное* число таких разнообразных примеров чаще всего заменяет нам формальное определение, влияет на глубину интуиции. Условимся в этом смысле говорить о *фаддеевском наборе* примеров, иллюстрирующих понятие.

Возьмем в качестве примера понятие коммутативной группы. Целые числа, группа из двух элементов (чет—нечет), вообще, группа вычетов по модулю p, затем рациональные числа, действительные числа, векторы ... Конечно, этот набор примеров является достаточно полным; он по-разному оттеняет понятие коммутативной группы (с точки зрения ее конечности или бесконечности, дискретности или непрерывности и т. п.). Но вот студент-математик познакомился с так называемыми p-адическими числами, совершенно не похожими по своим свойствам на перечисленные ранее примеры. Добавляя этот новый пример к своему фаддеевскому набору примеров коммутативных групп, студент обогащает свое владение понятием коммутативная группа. Вообще, с каждой новой теоремой фаддеевский набор может увеличиться (оставаясь конечным). И если человеку с богатым набором примеров предложат вопрос «Верно или неверно такое-то утверждение о коммутативных группах?», он прежде всего сопоставит это утверждение с имеющимся в его распоряжении фаддеевским набором. И это также — важный элемент математической культуры и интуиции.

Интересным примером может служить построенный К. Вейерштрассом пример непрерывной функции, которая ни в одной точке не имеет производной. Этот появившийся в XIX столетии пример существенно обогатил фаддеевский набор, относящийся к понятию «непрерывная функция». И другие теоремы Вейерштрасса (ограниченность непрерывной функции, заданной на отрезке; достижение наибольшего и наименьшего значений, свойство равномерной непрерывности, возможность приближения непрерывных функций многочленами и др.) существенно обогащают относящийся к этому понятию фаддеевский набор. И несмотря на конечность таких наборов, связанных в сознании математика с каждым понятием, именно эти наборы составляют богатство нашего математического миросозерцания, именно они позволяют охватить мысленным взором дискретное и непрерывное в математике, определяют глубину интуиции как в самой математике, так и в ее приложениях.

К «фаддеевским наборам» относятся не только различные *примеры* объектов, принадлежащих к понятию, но также необходимые и достаточные признаки понятия вместе с относящимися сюда *контрпримерами*. Поясним это, взяв понятие «параллелограмм». Известен целый ряд свойств параллелограммов (т. е. необходимых признаков этого понятия): противоположные стороны параллелограмма конгруэнтны, противоположные углы конгруэнтны, диагонали делят друг друга пополам, биссектрисы противоположных углов параллелограмма параллельны и т. п. Можно отдельные необходимые признаки комбинировать и ставить вопрос о справедливости обратной теоремы, т. е. вопрос о том, является ли взятая комбинация достаточным признаком. Вот несколько примеров таких вопросов.

а) Является ли четырехугольник параллелограммом, если две его противоположные стороны параллельны, а две другие конгруэнтны?

б) Является ли четырехугольник параллелограммом, если у него стороны (не обязательно противоположные) попарно конгруэнтны?

в) Является ли четырехугольник параллелограммом, если у него одна из диагоналей делится другой диагональю пополам и, кроме того, две противоположные стороны конгруэнтны?

г) Является ли четырехугольник параллелограммом, если у него две биссектрисы углов параллельны и, кроме того, два противоположных угла конгруэнтны?

Все эти примеры нарочито подобраны так, что они имеют отрицательный ответ, т. е. взятая комбинация необходимых признаков не является достаточным признаком. Контрпримеры приведены на

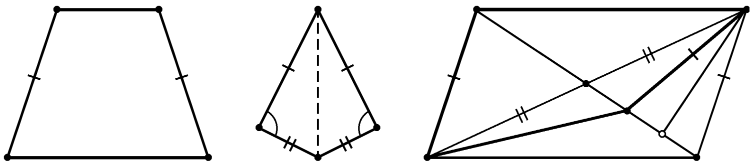


Рис. 458 а) б), г) в)

рис. 458 (в контрпримере к вопросу г) биссектрисы даже совпадают). Эти контрпримеры обогащают смысл понятия «параллелограмм».

Вообще, не только типичные (или экзотические) примеры объектов, принадлежащих понятию, но и различные контрпримеры, оттеняющие смысл этого понятия (или, во всяком случае, навык построения таких контрпримеров), составляют «фаддеевский набор», и чем он богаче, тем глубже наше владение рассматриваемым понятием. Большое количество контрпримеров собрано в книге Б. Гелбаума и Дж. Олмстеда «Контрпримеры в анализе».

Кроме указанных контрпримеров, показывающих, что некоторая комбинация необходимых признаков не составляет достаточного признака, важны и интересны также контрпримеры другого рода, оттеняющие условия теорем (в частности, необходимых признаков). Возьмем следующую хорошо известную теорему: если две противоположные стороны четырехугольника параллельны и конгруэнтны, то этот четырехугольник — параллелограмм. Условие этой теоремы представляет собой *конъюнкцию* двух отдельных требований:

1) две противоположные стороны четырехугольника параллельны;

2) *эти же* стороны конгруэнтны.

Попробуем выяснить, существенно ли для справедливости теоремы, что рассматриваются *эти же* противоположные стороны. Иными словами, не будет ли справедлива следующая теорема: если две противоположные стороны четырехугольника параллельны, а две другие его противоположные стороны конгруэнтны, то этот четырехугольник — параллелограмм? Равнобедренная трапеция является контрпримером, показывающим ложность этой гипотезы.

Вообще, если в условии некоторой теоремы имеется конъюнкция нескольких отдельных требований, то целесообразно выяснить, насколько существенным является каждое из этих требований, т. е. можно ли построить контрпример, отбросив одно требование. Более детально: если теорема имеет вид $(\forall\ x)\ (A(x) \wedge B(x) \wedge C(x)) \Rightarrow D(x)$, то возникает вопрос, существует ли контрпример, показывающий существенность требования $A(x)$, т. е. контрпример, показывающий ложность теоремы $(\forall\ x)\ (B(x) \wedge C(x)) \Rightarrow D(x)$? Иными словами, существует ли x_0, для которого $B(x_0)$ и $C(x_0)$ истинны, а $D(x_0)$ ложно?

Тот же вопрос о существенности ставится и в отношении требований $B(x)$ и $C(x)$. Такая работа (с каждой теоремой!) значительно повышает математическую культуру, кругозор и интуицию.

РЕШЕНИЯ ЗАДАЧ И УПРАЖНЕНИЙ

Глава I

1. У мадемуазель Рембо одна кошка, одна собака, один попугай и нет тараканов.

2. Способ определения множества перегоревших лампочек таков. Следует испробовать все шесть положений и отметить лампочки, которые не загорались ни при каких положениях. Они, и только они являются перегоревшими.

3. Нетрудно понять, что если кто-то из присутствующих сказал правду, то и следующие за ним также сказали правду, а если кто-то из присутствующих солгал, то и все, говорившие перед ним, также солгали. Заметим также, что первый заведомо солгал, а последний сказал правду. Таким образом, последние k человек сказали правду, а первые $12 - k$ солгали. Первой прозвучавшей правдой было «Здесь не более k правдивых людей», это сказал человек, говоривший $(12 - k + 1)$-м по порядку, а судя по тому, *что* он сказал, он говорил $(k + 1)$-м. Поэтому $(12 - k) + 1 = k + 1$, откуда $k = 6$. Итак, первые шесть человек — лжецы, последние шесть составляют в этой комнате множество правдивых людей.

4. Это множество ограничено дугами окружности радиуса 3 км с центром в точке A и касательными к этой окружности из точек B и C, находящихся на дороге на расстоянии 6 км от точки A (см. рис. 459).

5. Это множество состоит из окружности с центром в точке A радиуса AB, кроме точек B и C и им диаметрально противоположных, а также из еще шести точек M_1, M_2, ..., M_6, изображенных на рис. 460.

6. Подсчитаем число танцевавших пар. С одной стороны, их число равно утроенному числу кавалеров, а с другой — утроенному числу дам, следовательно, кавалеров было столько же, сколько и дам.

7. Если в билете первой катушки заменить цифры их дополнениями до 9, т. е. вместо цифры x ставить цифру $9 - x$, то получим взаимно однозначное соответствие между номерами билетов в первой и второй катушке. При таком преобразовании счастливые билеты переходят в счастливые, следовательно, их поровну в обеих катушках.

8. Каждому многоугольнику, не содержащему красную точку, сопоставим многоугольник с теми же вершинами и еще одной вершиной — в красной точке. При таком соответствии треугольникам, одна из вершин которых находится в красной точке, не соответствуют никакие многоугольники с вершинами лишь в белых точках. Следовательно, больше многоугольников, содержащих красную точку.

9. Из цифр 1 и 2 можно составить следующие числа, соответствующие буквам в алфавите: 1, 2, 11, 12, 21, 22, т. е. буквам А, Б, Й, К, У, Ф. Единственное осмысленное слово в данном случае: ФУФАЙКА.

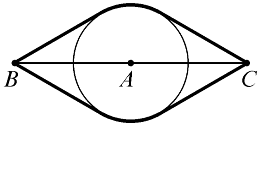


Рис. 459

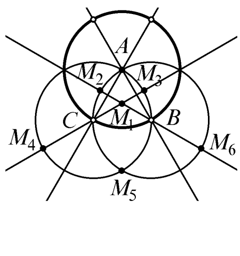


Рис. 460

10. Расположим красные палочки в одну линию, а рядом с ними синие, как на рис. 461. Теперь распилим красные палочки в стыках синих, а синие — в стыках красных и получим искомое расположение.

11. Поскольку множество рациональных чисел счетно, то перенумеруем их: p_1, p_2, ..., p_n, ... Точка с рациональными координатами на плоскости определяется парой рациональных чисел $(p_i; p_j)$, таким образом, для доказательства счетности точек плоскости с рациональными координатами достаточно доказать счетность множества пар чисел $(p_i; p_j)$. Такое доказательство можно провести точно так же, как проводилось доказательство счетности рациональных чисел. А именно, сначала возьмем пару $(p_1; p_1)$, затем последовательно пары $(p_1; p_2)$ и $(p_2; p_1)$, далее $(p_1; p_3)$, $(p_2; p_2)$, $(p_3; p_1)$, т. е. сначала пару с суммой номеров 2 (она одна), затем пары с суммой номеров 3 (их две), далее — с суммой номеров 4 и т. д. Для определенности пары с одной и той же суммой номеров n располагаем в порядке возрастания первого номера от 1 до $n - 1$. Нетрудно видеть, что таким образом окажутся пронумерованными все пары рациональных чисел, а следовательно, и все точки на плоскости с рациональными координатами.

12. Каждый элемент этого множества определяется парой $(M; r)$, где M — точка на плоскости с рациональными координатами, а r — положительное рациональное число. В предыдущей задаче было доказано, что множество точек M счетно. Занумеруем их: m_1, m_2, ..., m_n, ... Множество чисел r является частью множества рациональных чисел. Эти числа также можно занумеровать: r_1, r_2, ..., r_n, ... Осталось доказать счетность множества пар $(m_n; r_k)$, что проводится в точности так же, как и в задаче 11: первой берется пара $(m_1; r_1)$, затем пара $(m_1; r_2)$, далее $(m_2; r_1)$ и т. д.

13. В п. 9 было доказано, что множество рациональных чисел счетно, в задаче 11 — что множество упорядоченных пар рациональных чисел счетно. Теперь нетрудно доказать счетность множества наборов из трех рациональных чисел $(p; q; r)$, затем из четырех, далее из пяти и так дойти до любого натурального числа n. А множество многочленов степени n с рациональными коэффициентами, очевидно, эквивалентно множеству упорядоченных наборов из n рациональных чисел и, следовательно, счетно. (Читатель сам рассмотрит случай, когда старший коэффициент у многочлена степени n равен нулю, т. е. этот многочлен имеет степень, меньшую n.)

14. Рассмотрим одну из букв Т и выберем три точки с рациональными координатами так, как это показано на рисунке 462, которые назовем «сопровождающими» эту букву Т. Нетрудно заметить, что одна и та же тройка точек не может быть сопровождающей для двух (и более) непересекающихся букв Т, поэтому непересекающихся букв Т не больше, чем троек точек с рациональными координатами. Но каждая тройка является набором из 6 рациональных чисел и, как показано в задаче 13, их множество счетно.

Рис. 461

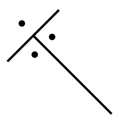


Рис. 462

15. В задаче 13 было доказано, что количество многочленов каждой степени n с рациональными коэффициентами, а подавно и с целыми коэффициентами, счетно. Теперь легко пересчитать все корни таких многочленов. Сначала считаем корни первого, потом второго и т. д. Разумеется, одни и те же корни могут встречаться у разных многочленов, но мы можем не учитывать корень, если он уже встречался ранее. Таким образом, получаем, что количество алгебраических чисел счетно.

16. Рассмотрим числа, составленные из трех цифр, которые могут начинаться и с нуля, например, 311, 045, 960. Всякому такому числу x сопоставим число $999 - x$; полученное число также составлено из трех цифр и может начинаться с нуля. Для приведенных примеров это будут числа 688, 954 и 039. Это соответствие взаимно однозначно. Если сумма цифр первоначального числа равнялась n, то сумма цифр полученного из него числа равняется $27 - n$. Теперь сопоставим каждому счастливому билету билет, у которого последние три цифры заменены по указанному выше правилу. Таких билетов будет столько же, сколько и счастливых, а если сумма первых (как и последних) трех цифр счастливого билета равнялась n, то у полученного билета сумма всех цифр равна $n + (27 - n) = 27$.

17. Следует заметить, что при указанных операциях разность между количеством букв О и Д не меняется. Поэтому слова ОДД и ДДО, имеющие в качестве этих разностей -1 и $+1$, не могут быть получены одно из другого описанными операциями. Нетрудно заметить, что различных слов будет столько же, сколько различных разностей, т. е. бесконечно много, но счетное количество.

18. Нетрудно заметить, что внутри каждого треугольника есть хотя бы одна точка с рациональными координатами (на самом деле их там бесконечно много), и эта точка лежит внутри только этого треугольника. Значит, треугольников не больше, чем рациональных точек на плоскости, и поэтому их количество конечно или счетно.

19. Рассмотрим куб с вершинами в точках $(0; 0; 0)$, $(0; 0; 1)$, $(0; 1; 0)$, $(1; 0; 0)$, $(0; 1; 1)$, $(1; 0; 1)$, $(1; 1; 0)$ и $(1; 1; 1)$. Пусть точка M этого куба имеет координаты $x = 0{,}a_1a_2a_3...$, $y = 0{,}b_1b_2b_3...$ и $z = 0{,}c_1c_2c_3...$. Сопоставим точке M точку N отрезка $[0; 1]$, имеющую координаты $0{,}a_1b_1c_1a_2b_2c_2a_3b_3c_3...$. Это соответствие, очевидно, взаимно однозначно.

20. Предположим, что мы установили взаимно однозначное соответствие между элементами m множества $\boldsymbol{M}$ и всеми его подмножествами: $m \leftrightarrow \boldsymbol{N}_m$. Построим теперь следующее подмножество $\boldsymbol{N}$ множества $\boldsymbol{M}$: элемент x мы включим в $\boldsymbol{N}$, если $x \notin \boldsymbol{N}_x$, а если $x \in \boldsymbol{N}_x$, то его мы не будем включать в $\boldsymbol{N}$.

Легко понять, что подмножеству $\boldsymbol{N}$ при нашем соответствии не поставлен в соответствие ни один элемент множества $\boldsymbol{M}$, следовательно, предположение о наличии взаимно однозначного соответствия неверно. В то же время каждый элемент множества $\boldsymbol{M}$ есть его подмножество, поэтому подмножеств не меньше, чем элементов. Суммируя сказанное, получаем, что утверждение задачи справедливо.

21. Это — множество чисел, делящихся на 10.

22. Это — множество квадратов.

23. По условию 20% жителей Монреаля не знают французского языка, но знают английский, а 30% не знают английского, но знают французский. Поэтому 50% знают лишь один язык и 50% жителей знают оба языка.

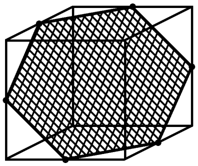


Рис. 463

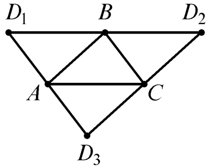


Рис. 464

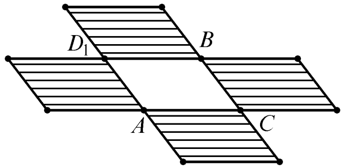


Рис. 465

24. Такая плоскость проходит через центр куба и середины двух соседних ребер; при этом она будет проходить через середины еще четырех ребер куба (см. рис. 463).

25. Все корни уравнения расположены на отрезке [–10; 10], поскольку $|10 \sin x| \le 10$. Из рассмотрения графиков функций $y = x$ и $y = 10 \sin x$ нетрудно получить, что имеется ровно 7 корней.

26. Заметим, что в сумме $|\boldsymbol{A}| + |\boldsymbol{B}| + |\boldsymbol{C}|$ элементы, принадлежащие ровно двум множествам, учитываются дважды, а принадлежащие всем трем — трижды. В каждом из множеств $\boldsymbol{A} \cap \boldsymbol{B}$, $\boldsymbol{A} \cap \boldsymbol{C}$ и $\boldsymbol{B} \cap \boldsymbol{C}$ содержится множество $\boldsymbol{A} \cap \boldsymbol{B} \cap \boldsymbol{C}$, поэтому первые шесть членов выражения справа дадут сумму мощностей множеств $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{C}$ без мощности их пересечения. Для выполнения равенства следует учесть это пересечение.

27. Известно, что если в произвольном четырехугольнике последовательно соединить середины соседних сторон, то полученный четырехугольник будет параллелограммом. Поэтому середина четвертой стороны многоугольника, указанного в задаче, должна дополнять данные три точки до вершин параллелограмма, что можно сделать тремя способами (рис. 464). Рассмотрим возможные расположения вершин исходного четырехугольника для одного из вариантов. Отметим, что диагонали исходного четырехугольника параллельны сторонам параллелограмма и вдвое больше их (по свойству средней линии треугольника). Из соображений выпуклости вершины четырехугольника должны лежать в параллелограммах, конгруэнтных данному и примыкающих к нему по стороне (рис. 465). Если теперь выбрать произвольную точку в одном из этих четырех параллелограммов в качестве вершины исходного четырехугольника, то остальные вершины однозначно восстанавливаются по заданным серединам сторон.

Соответствующие фигуры получаются и для двух других вариантов. Окончательный ответ будет объединением этих трех множеств (рис. 466).

28. Проведем через точку C прямую и отметим на ней проекции точек A и B — точки M и N. Так как угол AMC — прямой, то точка M лежит на окружности с диаметром AC; аналогично точка N лежит на окружности с

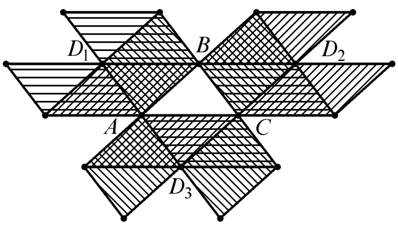


Рис. 466

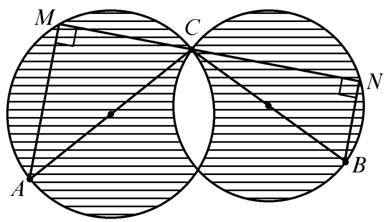


Рис. 467

диаметром BC, а точки отрезка AB проктируются на отрезок между точками M и N. Теперь нетрудно увидеть, что множество проекций всех точек, лежащих внутри отрезка AB, есть объединение двух указанных кругов без их пересечения, а точки A и B проектируются в границы этих кругов (рис. 467).

29. Рассмотрим множество философов, являющихся математиками. Пусть их число равно n, тогда математиков $7n$, а философов $10n$. Значит, философов больше, чем математиков.

30. Так как $\boldsymbol{A} \cap \boldsymbol{B} = \boldsymbol{C}$, то $(\boldsymbol{A} \cap \boldsymbol{B}) \cup \boldsymbol{C} = \boldsymbol{C} \cup \boldsymbol{C} = \boldsymbol{C}$, причем $\boldsymbol{A} \supset \boldsymbol{C}$ и $\boldsymbol{B} \supset \boldsymbol{C}$. Поэтому $\boldsymbol{A} \cup \boldsymbol{B} \supset \boldsymbol{C}$ и $(\boldsymbol{A} \cup \boldsymbol{B}) \cap \boldsymbol{C} = \boldsymbol{C}$. Итак, левая и правая части равенства выражают множество $\boldsymbol{C}$; следовательно, равенство справедливо.

31. Пусть N — число учеников, тогда мальчиков $\frac{2N}{5}$, и столько же участников кружка. Мальчики составляют $\frac{3}{5}$ кружковцев, т. е. их $\frac{6N}{25}$; так как $\frac{6N}{25} : \frac{2N}{5} = 0{,}6$, то 60% мальчиков посещают кружок. При решении полезно нарисовать диаграмму Эйлера—Венна.

32. Слева и справа записано множество элементов, не принадлежащих ни одному из множеств $\boldsymbol{A}$, $\boldsymbol{B}$ и $\boldsymbol{C}$.

33. Множество учеников класса разбивается на 4 подмножества: $\boldsymbol{M}_{\text{р}}$ — мальчики, решившие задачу, $\boldsymbol{M}_{\text{н}}$ — мальчики, не решившие ее, $\boldsymbol{Д}_{\text{р}}$ — девочки, решившие задачу, и $\boldsymbol{Д}_{\text{н}}$ — девочки, не решившие ее (рис. 468). Так как количество решивших равно $\boldsymbol{M}_{\text{р}} + \boldsymbol{Д}_{\text{р}}$, а количество девочек равно $\boldsymbol{Д}_{\text{р}} + \boldsymbol{Д}_{\text{н}}$, то (поскольку $\boldsymbol{M}_{\text{р}} = \boldsymbol{Д}_{\text{н}}$) эти количества равны.

34. Пусть в этом турнире в n партиях выиграли черные, тогда участник, выигравший k партий черными, выиграл $n - k$ партий белыми, а всего он выиграл $k + (n - k) = n$ партий.

35. Нарисуем диаграмму Эйлера—Венна в виде трех кругов, соответствующих изучающим каждый из языков, и начнем ее заполнять, начиная с конца списка в условиях задачи. Получим картину на рис. 469. Видно, что только французский язык изучают 30 человек, а ни одного языка — 20.

36. Если к каждой точке окружности, лежащей в плоскости, приложить одним концом отрезки одной и той же длины, перпендикулярные плоскости (причем все отрезки лежат в одном полупространстве, см. рис. 470), то получится боковая поверхность цилиндра. Множество точек этой поверхности описывается парами $(X; Y)$, где X — точка на окружности, Y — точка на отрезке.

37. Заметим, что окружность — согнутый в кольцо отрезок. В задаче 36 боковую поверхность цилиндра можно рассматривать как согнутый в трубку

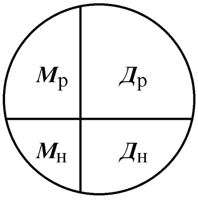


Рис. 468

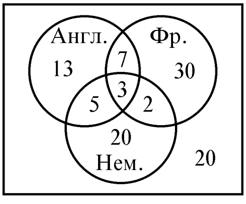


Рис. 469

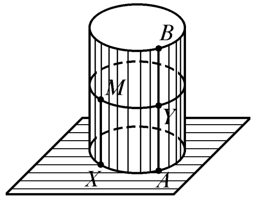


Рис. 470

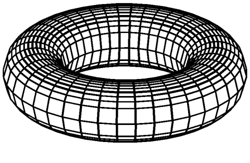


Рис. 471

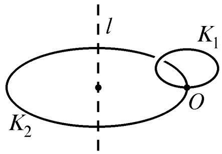


Рис. 472

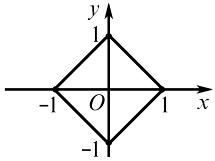


Рис. 473

квадрат, у которого склеены две противоположные стороны. Чтобы получить произведение двух окружностей, нужно дополнительно склеить основания цилиндра. При этом боковая поверхность цилиндра изогнется и примет форму бублика (рис. 471). Такая фигура называется *тором*. Его можно также получить, вращая окружность K_1 вокруг прямой l, находящейся в ее плоскости и не пересекающей эту окружность (рис. 472). Полученную фигуру можно рассматривать как произведение окружности K_1 и окружности K_2, получающейся при вращении одной из точек $O \in K_1$ вокруг прямой l.

38. Пусть сторона квадрата содержит n музыкантов. После перестроения количество музыкантов в шеренге уменьшилось на x и потому имеем соотношение $n^2 = (n - x)(n + 5)$. Из него находим, что $n = \dfrac{5x}{5 - x}$. При целых положительных значениях x это выражение будет целым лишь при $x = 4$, откуда $n = 20$, а количество музыкантов равно 400.

39. Поскольку ладья пробивает всю горизонталь, а горизонталей имеется 8, то с соблюдением условий задачи ладей можно поставить не больше восьми. Найдем число способов расставить 8 ладей. На первой вертикали можно поставить ладью в любую из восьми клеток. После этого на второй вертикали останутся непробитыми 7 клеток, в каждую из которых можно поставить вторую ладью. Получаем $8 \cdot 7 = 56$ возможностей. Третью ладью можно поставить уже шестью способами, что дает $56 \cdot 6 = 336$ способов, и так далее. Всего получаем $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 40320$ способов.

40. Это множество — объединение сторон квадрата, изображенного на рис. 473.

41. а) Нет, поскольку не у всех живущих людей есть дочь.

б) Да, поскольку у всех людей есть (или была) мать.

в) Да, поскольку он определен для каждого человека.

42. Если элемент x принадлежит $\boldsymbol{A} \cap \boldsymbol{B}$, т. е. и $\boldsymbol{A}$, и $\boldsymbol{B}$, то $f(x) = 1$ и $g(x) = 1$, поэтому $h(x) = 1$; если же x не принадлежит $\boldsymbol{A} \cap \boldsymbol{B}$, то он не принадлежит хотя бы одному из множеств $\boldsymbol{A}$ и $\boldsymbol{B}$, поэтому или $f(x) = 0$, или $g(x) = 0$ (или и то, и другое). Отсюда $h(x) = 0$. Таким образом, $h(x)$ — характеристическая функция множества $\boldsymbol{A} \cap \boldsymbol{B}$.

Для множества $\boldsymbol{A} \cup \boldsymbol{B}$ характеристической функцией является функция $k(x) = f(x) + g(x) - f(x)g(x)$. Действительно, для элементов x множества $\boldsymbol{U}$, не принадлежащих $\boldsymbol{A} \cup \boldsymbol{B}$, обе функции $f(x)$ и $g(x)$ равны нулю, поэтому и $k(x) = 0$; для элементов, принадлежащих только $\boldsymbol{A}$ или только $\boldsymbol{B}$, это будет 1, так же как и для элементов, принадлежащих $\boldsymbol{A} \cap \boldsymbol{B}$.

43. При $x \geq 0$ будет $\Theta(x) = 1$, поэтому $x(2\Theta(x) - 1) = x$, а для $x < 0$ будет $\Theta(x) = 0$, поэтому $x(2\Theta(x) - 1) = -x$, а это и есть определение функции $|x|$.

44. а) Нет, поскольку некоторым участникам может сопоставляться несколько номеров.

б) Да, при условии, что участнику олимпиады, не решившему ни одной задачи, сопоставляется число 0.

45. Заметим, что должно быть $x^2 - 16 \geq 0$, но $x^2 - 16 \leq 9$. Из первого неравенства находим $|x| \geq 4$, а из второго $|x| \leq 5$. Ответ: $[-5; -4] \cup [4; 5]$.

46. Эти функции не являются наложениями, поскольку каждая из них не может принимать отрицательные значения и значения, большие, скажем, 3000. Функция из упражнения 41в не может быть вложением, так как многие люди родились в один и тот же год. Функция из упражнения 44б может оказаться вложением, если количество участников олимпиады не больше 16, так как сумма номеров решенных задач может изменяться от 0 до 15.

47. Это отображение не является наложением, поскольку, например, $f^{-1}(-5) = \varnothing$. Оно не является и вложением, так как $f(-x) = f(x)$.

48. Отображение $f(x) = x^2$ не является ни тем, ни другим, поскольку $f(x)$ всегда неотрицательно и $f(x) = f(-x)$. Отображение $f(x) = x^3$ является и вложением, и наложением, поскольку каждому числу y однозначно соответствует число $\sqrt[3]{y}$. Отображение, задаваемое функцией $y = 2^x$, не является наложением, поскольку $2^x > 0$ для всех x, но является вложением, поскольку 2^x — монотонно возрастающая функция, и поэтому различным значениям x соответствуют различные значения 2^x.

49. При $x = 4$ имеем $f(x) = \sqrt{3}$ и далее значения функции $f(x)$ уменьшаются при увеличении x до значения $f(5) = 0$. Поэтому значения на полуинтервале $(\sqrt{3}; 2]$ она не принимает, т. е. прообраз каждой из этих точек — пустое множество. Значение 1 функция принимает для $x = \pm 2\sqrt{5}$, поэтому в силу четности функции прообразом отрезка $[1; 2]$ является объединение двух отрезков $[-2\sqrt{5}; -4] \cup [4; 2\sqrt{5}]$.

50. График этой функции дан на рисунке 474.

51. Да, существует и совпадает с функцией f.

52. Для непрерывной функции необходимым и достаточным условием наличия обратной функции является ее монотонность. Функция $x^3 - x^2$ монотонно возрастает на луче $(-\infty; 0]$, монотонно убывает на отрезке $[0; \frac{2}{3}]$ и вновь монотонно возрастает на луче $[\frac{2}{3}; +\infty)$. На этих множествах $f(x)$ имеет обратную, если рассматривать отображение каждого множества на его образ.

53. Множество значений функции $x + [x]$ не содержит ни одного из чисел y, у которых $|y|$ — нечетное число, поэтому этим y не сопоставляется никакого значения x; следовательно, у этой монотонно возрастающей, но разрывной функции не существует обратной. На интервале $(100; 101)$ функция $x + [x]$ совпадает с функцией $x + 100$ и $f^{-1}(x) = x - 100$.

54. Если бы такой отрезок существовал, то его точки находились бы во взаимно однозначном соответствии с множеством из двух элементов 0 и 1, что невозможно, так как точек отрезка бесконечно много.

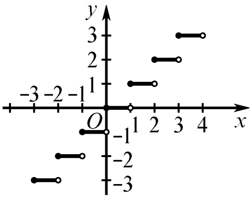


Рис. 474

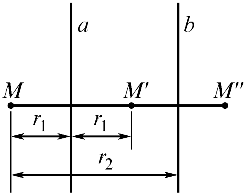


Рис. 475

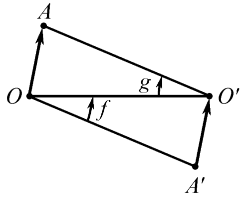


Рис. 476

55. Нет, неверно. Например, $f(x) = x$, $g(x) = 1 - x$ имеют обратные, а $h(x) = f(x) + g(x) = 1$ обратной не имеет.

56. В случае, изображенном на рис. 475, точка M, отстоящая от прямой a на расстоянии r_1 и от прямой b на расстоянии r_2, после первой симметрии переходит в точку M', отстоящую от M на расстоянии $2r_1$, а точка M' после второй симметрии переходит в точку M'', отстоящую от M' на расстоянии $2(r_2 - 2r_1)$ и от точки M на расстоянии $2r_1 + 2(r_2 - 2r_1) = 2(r_2 - r_1)$. Но $r_2 - r_1$ есть расстояние между прямыми a и b. То, что перенос происходит перпендикулярно этим прямым, следует из определения симметрии. При другом расположении точки M рассуждение аналогично.

57. Пусть A и A' — такие точки, что $g(O) = A$, $f(A') = O'$ (рис. 476). Тогда $g(f(O)) = A$ и $g(f(A') = O'$. Поэтому $g \circ f$ есть параллельный перенос на вектор $\vec{p} = \overrightarrow{OA} = \overrightarrow{A'O'}$.

58. См. рис. 477.

59. Преобразуем уравнение:

$$x^4 - 2x^3 - 4x^2 + 5x + 6 = (x^2 - x)^2 - 5(x^2 - x) + 6 = 0.$$

Заменяя $x^2 - x$ на y, получаем уравнение $y^2 - 5y + 6 = 0$, корнями которого являются $y_1 = 2$, $y_2 = 3$. Далее решаем два квадратных уравнения $x^2 - x = 2$ и $x^2 - x = 3$ и получаем четыре корня: $x_1 = -1$, $x_2 = 2$, $x_3 = \dfrac{1 + \sqrt{13}}{2}$, $x_4 = \dfrac{1 - \sqrt{13}}{2}$.

60. Сделаем замену $|x| = y$; получаем уравнение $y^2 + py + q = 0$. Теперь нас будет интересовать количество положительных и нулевых корней этого многочлена. Если дискриминант $p^2 - 4q$ отрицателен, то уравнение вообще не имеет действительных корней. Если $p^2 - 4q = 0$, то имеем два равных корня,

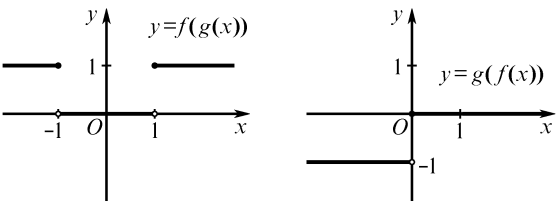


Рис. 477

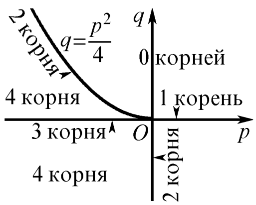


Рис. 478

знак которых противоположен знаку числа p. Поэтому для $p > 0$ нет положительных корней, для $p < 0$ есть два совпадающих положительных корня. При $p = 0$ получаем, что и $q = 0$, уравнение принимает вид $y^2 = 0$ и имеет два нулевых корня.

В случае $p^2 - 4q > 0$ при $q > 0$ корни имеют одинаковые знаки, причем положительные при $p < 0$ и отрицательные при $p > 0$, а при $q < 0$ корни имеют разные знаки. При $q = 0$ один из корней нулевой, а второй имеет знак, противоположный знаку p.

Вернемся к переменной $|x| = y$. Это уравнение при заданном y имеет два корня, если $y > 0$, один, если $y = 0$, и ни одного, если $y < 0$.

Окончательный результат изображен на рисунке 478.

61. Нет, не может, так как не выполняется условие транзитивности. Например, числа 10 и 15 имеют общий делитель 5 (больший единицы), числа 15 и 21 имеют общий делитель 3, а числа 10 и 21 не имеют общего делителя (большего единицы).

62. Свойства рефлексивности, симметрии и транзитивности легко проверяются. В класс с множителем 1 попадают все простые числа. В класс с множителем 7 — все числа, равные произведению 7 на простое число, не превосходящее 7, т. е. 14, 21, 35, 49. В класс с множителем 12 попадает лишь число 24, так как если некоторое число представляется в виде $12k$, где $k > 2$, то оно представляется и в виде $(6k)\cdot 2$, а так как $k > 2$, то $6k > 12$ и 12 не является наибольшим делителем, отличным от самого числа.

63. Нет, так как ромбы содержатся в классе параллелограммов.

64. У треугольников является, а у четырехугольников — нет, поскольку не для каждого четырехугольника существует вписанная в него окружность.

65. Это отношение не транзитивно (для доказательства достаточно рассмотреть три города с числом жителей 5 000, 10 000 и 15 000).

66. Да, будет.

67. Может не быть, например, если команда A выиграла у B, B выиграла у C, а C выиграла у A.

68. Рассмотрим следующую таблицу результатов, поставленных судьями I, II и III претендентакм A, B, C, D и E:

Судьи	**1 место**	**2 место**	**3 место**	**4 место**	**5 место**
I	A	B	C	D	E
II	C	D	E	A	B
III	E	A	B	C	D

Видно, что $A > B > C > D > E > A$, что показывает отсутствие максимального (а также минимального) элемента и транзитивности.

69. В соответствии с определением для каждых двух различных точек M_1 и M_2 строго определено условие $M_1 > M_2$ или $M_1 < M_2$. При этом $M_1 = M_2$, если это одна и та же точка. Это множество является цепью.

70. Предположим противное — что после перестроения по росту найдется восьмиклассник, больший стоящего за ним девятиклассника. Пусть он стоит на k-м месте, тогда все восьмиклассники с меньшими номерами будут не меньше его и поэтому больше k-го девятиклассника и всех, стоящих за ним. Таким образом, этому девятикласснику и остальным, меньшим его, не хватит «маленьких» восьмиклассников, чтобы поставить их впереди себя.

71. При $n = 1$ получаем $1 + 3 = 4 = (1 + 1)^2$. Предположим, что равенство выполняется для $n = k$, и докажем, что в этом случае оно выполняется и для $n = k + 1$. При $n = k + 1$ мы имеем сумму $1 + 3 + 5 + ... + (2k + 1) + (2k + 3)$. По предположению индукции сумма всех слагаемых, кроме последнего, равна $(k + 1)^2$, поэтому вся сумма равна $(k + 1)^2 + (2k + 3) = k^2 + 2k + 1 + 2k + 3 =$ $= k^2 + 4k + 4 = (k + 2)^2$. Утверждение доказано.

72. При $n = 1$ получаем $1^2 = \frac{1 \cdot 2 \cdot 3}{6} = 1$. Условие выполнено. Пусть оно справедливо при $n = k$. Докажем, что оно справедливо тогда и при $n = k + 1$. Действительно,

$$1^2 + 2^2 + ... + k^2 + (k + 1)^2 = \frac{k(k + 1)(2k + 1)}{6} + (k + 1)^2 = \frac{(k + 1)(2k^2 + k + 6k + 6)}{6} =$$
$$= \frac{(k + 1)(k + 2)(2k + 3)}{6} = \frac{(k + 1)((k + 1) + 1)(2(k + 1) + 1)}{6}.$$

73. Как и в упражнении 72, очевидно, что $1^3 = 1$, и из предположения, что $1^3 + 2^3 + ... + k^3 = \frac{k^2(k + 1)^2}{4}$, следует заключить справедливость равенства $1^3 + 2^3 + ... + k^3 + (k + 1)^3 = \frac{(k + 1)^2(k + 2)^2}{4}$. Но по предположению левая часть этого равенства может быть переписана как

$$\frac{k^2(k + 1)^2}{4} + (k + 1)^3 = \frac{(k + 1)^2(k^2 + 4k + 4)}{4} = \frac{(k + 1)^2(k + 2)^2}{4}.$$

Утверждение доказано.

74. Дело в том, что указанный переход от n к $n + 1$ неверен при $n = 1$.

75. При $n = 2000$ выражение слева равно 7 999 999 996, а выражение справа равно 4 000 006 000, и неравенство выполняется Пусть оно выполняется для $n = k \ge 2000$; докажем, что оно справедливо и для $n = k + 1$. При переходе от k к $k + 1$ левая часть увеличивается на $(k + 1)^3 - 4 - (k^3 - 4) = 3k^2 + 3k + 1$, а правая на $1000(k + 1)^2 + 3(k + 1) - 1000k^2 - 3k = 2000k + 1003$, но $3k^2 \ge 6000k$ при $k \ge 2000$, а $3k + 1 > 6000$. Так как $6000k > 2000k$, а $6000 > 1003$, то отсюда следует, что левая часть увеличивается больше чем правая, поэтому неравенство также справедливо для $n = k + 1$. Используя принцип математической индукции, получаем, что указанное неравенство справедливо для любого $n \ge 2000$.

Глава II

76. В таком числе первой цифрой может быть любая из 8 цифр (кроме 0 и 7), второй, третьей и т. д. цифрами могут быть любые 9 цифр (кроме 7), поэтому количество чисел, не содержащих цифру 7, равно $8 \cdot 9^6 = 4\,251\,528$, что меньше половины всех семизначных чисел.

77. Указанную запись можно восстановить 3^{12} способами, поскольку каждый из 12 элементов может принимать 3 значения. Содержательным является вариант «ПРОСТОЙ ШИФР».

78. «Сколько граней у шестигранного карандаша?». Заметим, что если он не очинен, то восемь (две — с торцов).

79. В записи участвуют пять элементов, являющихся цифрами (которых 10, что дает 10^5 вариантов), и три элемента — буквы, из которых естественно исключить «й», «ё» и «ы». Остается 27 букв, что дает 27^3 вариантов. Окончательно получаем $27^3 \cdot 10^5 = 1\,968\,300\,000$ вариантов.

80. Количество зерен в задаче равно $2^0 + 2^1 + 2^2 + ... + 2^{63} = 2^{64} - 1$, а количество вариантов раскладки зерен равно 2^{64}, т. е. на единицу больше.

81. $2002 = 1024 + 512 + 256 + 128 + 64 + 16 + 2 =$

$$= 2^{10} + 2^9 + 2^8 + 2^7 + 2^6 + 2^4 + 2^1 = 11\,111\,010\,010^{(2)}.$$

82. Заметим, что 3^n при делении на 2 дает в остатке 1 для любого $n = 0$, 1, 2, ... Поэтому остаток при делении числа на 2 совпадает с остатком при делении на 2 суммы цифр этого числа, записанного в троичной системе счисления.

Нетрудно аналогичным образом показать, что число записанное в системе счисления с основанием k, делится на $k - 1$, если сумма его цифр делится на $k - 1$.

83. В двоичной системе счисления имеем $1101^{(2)} + 11011^{(2)} = 101000^{(2)}$, что соответствует числу $2^5 + 2^3 = 32 + 8 = 40$.

84. Нетрудно заметить, что О равняется нулю. Действительно, обозначим основание системы счисления через p, тогда если О не равно нулю, то $2 \cdot \text{О} = p + \text{О}$, откуда $\text{О} = p$, но цифры p быть не может. Рассмотрим третий разряд. Ясно, что К = И + 1. Поэтому $2\text{Т} = p + \text{И}$. Из четвертого разряда получаем, что $\text{К} + \text{И} = p$, а из пятого $\text{К} + 1 = \text{Т}$. Полученная система четырех уравнений с четырьмя неизвестными имеет единственное решение: $p = 7$, И = 3, К = 4, Т = 5, О = 0. Итак, имеет место семиричная система счисления.

85. Обозначим основание системы счисления через p, тогда это равенство может быть записано как $(p + 3)^2 = p^2 + 7p + 1$. Из этого уравнения получаем, что $p = 8$.

86. Поскольку различных оценок четыре, то искомое число равно числу перестановок четырех элементов, т. е. 4! = 24.

87. Любому из указанных чисел соответствует выбор пяти элементов из десяти (цифр). Таким образом, искомое количество равно $\binom{10}{5} = 252$ (см. п. 26).

88. Искомое число равно числу размещений трех элементов из 25. Оно равно $25 \cdot 24 \cdot 23 = 13\,800$.

89. Количество вариантов равно $10! = 3\,628\,800$.

90. Число флагов равно числу размещений трех элементов из шести. Оно равно $6 \cdot 5 \cdot 4 = 120$.

91. Так как $\binom{n}{k} = \dfrac{n!}{k!\,(n-k)!}$, а $\binom{n}{n-k} = \dfrac{n!}{(n-k)!\,k!}$, то эти числа равны.

92. Это число равно количеству способов выбрать пять вопросов из 10 и равно $\binom{10}{5} = 252$.

93. В слове МАТЕМАТИКА 10 букв, из них три буквы А, две буквы М и две буквы Т, оставшиеся три буквы входят по одному разу; поэтому количество различных слов из этих букв равно $\dfrac{10!}{3!\,2!\,2!\,1!\,1!\,1!} = 151\,200$.

94. Зафиксируем одну из бусин, тогда оставшиеся 19 бусин можно расположить 19! способами. Однако, перевернув на другую сторону ожерелье, мы получаем еще одно расположение бусин, поэтому количество способов вдвое меньше, т. е. 19!/2.

95. Да, играли. Заметим, что при семи участниках играется $\binom{7}{2} = 21$ партия, при восьми $\binom{8}{2} = 28$ партий, а при девяти $\binom{9}{2} = 36$ партий. Так как было сыграно больше 21 партии, то участников было больше семи, а так как между остальными участниками было сыграно меньше 28 партий, то кроме Саши и Жени было меньше восьми шахматистов.

Число участников не могло быть равно 8, так как в этом случае были бы сыграны все 28 партий, т. е. Саша и Женя доиграли бы до конца, вопреки условию. Значит, было 9 участников; все, кроме Саши и Жени, сыграли между собой 21 партию, Женя и Саша сыграли 7 партий: одну между собой и еще по 3 с другими участниками.

96. Прежде всего возьмем по одному пирожному каждого сорта. Теперь к ним надо добавить шесть пирожных любых сортов (из четырех возможных). Представим эти пирожные в виде 6 белых шаров и добавим к ним три черных шара. Теперь разместим эти 9 шаров в одну линию. Черные шары разбивают белые на четыре группы, причем в каждой группе может быть от 0 до 6 шаров. Это соответствует выбору 6 пирожных четырех сортов. Различных выборов столько же, сколькими способами можно разместить 3 черных шара на 9 возможных местах. А это число равно $\binom{9}{3} = \frac{9 \cdot 8 \cdot 7}{1 \cdot 2 \cdot 3} = 84$.

97. Применяя указанный в предыдущей задаче прием, в этом случае получаем, что количество составов оркестров равно $\binom{6}{2} = \frac{6 \cdot 5}{2} = 15$.

98. Решение этой задачи дается сочетаниями с повторениями. Формула записывается в виде $\binom{n+k-1}{n-1}$, где k — число предметов в наборе, а n — количество видов предметов. Количество способов купить 10 акций семи компаний равно $\binom{16}{6} = 8008$.

99. В этом случае применима формула сочетаний с повторениями при $k = 7$ и $n = 4$. Получаем, что число способов наградить министра семью орденами равно $\binom{10}{3} = 120$.

100. Здесь применима та же формула при $n = 12$ и $k = 6$. Получаем, что число вариантов равно $\binom{17}{11} = \binom{17}{6} = 12\,376$.

101. $(2x + a)^4 = 16x^4 + 32x^3a + 24x^2a^2 + 8xa^3 + a^4$.

102. В разложении бинома $(x + 2)^{10}$ коэффициент при x^k равен $\binom{10}{k} 2^{10-k}$, а при x^{k+1} он равен $\binom{10}{k+1} 2^{9-k}$. Их отношение равно $\frac{2(k+1)}{10-k}$. Это выражение больше 1, если $k \geq \frac{8}{3}$, т. е. при увеличении на 1 показателя степени коэффициент при x^k уменьшается, если степень больше, чем $\frac{8}{3}$, и увеличивается, если степень меньше, чем $\frac{8}{3}$. Отсюда следует, что максимальный коэффициент будет при $k = 3$, он равен $\binom{10}{3}2^7 = 15\,360$.

103. Коэффициент при $x^2y^3z^2$ равен $\binom{5}{2}\binom{7}{2} = 210$.

104. Выражения слева и справа равны коэффициенту при x^n в выражении $(1 + x)^{2n} = (1 + x)^n(1 + x)^n$.

105. Строение таких треугольников Паскаля можно понять из следующего построения для $n \le 15$:

$$\begin{array}{c}
\mathbf{1}\\
\mathbf{1\ 1}\\
\mathbf{1\ 0\ 1}\\
\mathbf{1\ 1\ 1\ 1}\\
\mathbf{1\ 0\ 0\ 0\ 1}\\
\mathbf{1\ 1\ 0\ 0\ 1\ 1}\\
\mathbf{1\ 0\ 1\ 0\ 1\ 0\ 1}\\
\mathbf{1\ 1\ 1\ 1\ 1\ 1\ 1\ 1}\\
\mathbf{1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1}\\
\mathbf{1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1}\\
\mathbf{1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 1}\\
\mathbf{1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1}\\
\mathbf{1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1}\\
\mathbf{1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1}\\
\mathbf{1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1}\\
\mathbf{1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1}
\end{array}$$

Видно, что 1-я, 2-я, 4-я, 8-я, ... строки состоят из одних единиц, а под каждой из этих строк находится перевернутый треугольник, состоящий из одних нулей.

106. Производящая функция для геометрической прогрессии a_0, a_0q, a_0q^2, ... равна $a_0 + a_0qx + a_0q^2x^2 + ... = \frac{a_0}{1 - qx}$.

107. Рассмотрим производящие функции $(1 + x)^n$ и $(1 + x)^m$ и перемножим их: $(1 + x)^n(1 + x)^m = (1 + x)^{n+m}$. Теперь рассмотрим коэффициент при x^s слева и справа. Справа он равен $\binom{n+m}{s}$, а слева равен

$$\binom{n}{0}\binom{m}{s} + \binom{n}{1}\binom{m}{s-1} + ... + \binom{n}{s}\binom{m}{0}.$$

108. Производящая функция для натурального ряда имеет вид $1 + 2x + 3x^2 + ... + nx^{n-1} + ...$. Эта функция является производной от функции $1 + x + x^2 + x^3 + ... + x^n + ...$, которая равна $\frac{1}{1-x}$ (вопросы сходимости мы здесь не рассматриваем). Ее производная равна $\frac{1}{(1-x)^2}$.

109. Рассмотрим многочлен $(1+x)(1+x)^n = (1+x)^{n+1}$. Коэффициент при x^k равен $\binom{n+1}{k}$, а раскрывая биномы слева, получим, что он равен $\binom{n}{k}+\binom{n}{k-1}$.

110. Используя результат задачи 108

$$1 + 2x + 3x^2 + \ldots + nx^{n-1} + (n+1)x^n + \ldots = \frac{1}{(1-x)^2},$$

запишем, что $1 - 2x + 3x^2 - \ldots + n(-1)^{n-1}x^{n-1} + (n+1)(-1)^n x^n + \ldots = \dfrac{1}{(1+x)^2}$ и потому $\dfrac{1}{(1-x^2)^2} = \dfrac{1}{(1-x)^2}\cdot\dfrac{1}{(1+x)^2} = 1 + 2x^2 + 3x^4 + \ldots + nx^{2(n-1)} + (n+1)x^{2n} + \ldots$.

Здесь коэффициент при x^{2n} равен $(n+1)$, а если перемножить разложения $\dfrac{1}{(1-x)^2}$ и $\dfrac{1}{(1+x)^2}$, то получим

$$1 + [1\cdot(-2) + 2\cdot 1]x + [1\cdot 3 + 2\cdot(-2) + 3\cdot 1]x^2 + \ldots +$$
$$+ [1\cdot(-1)^n(n+1) + 2\cdot(-1)^{n-1}n + \ldots + (n+1)\,(-1)^0\cdot 1]\,x^n + \ldots$$

Коэффициент при x^{2n} равен $1\cdot(2n+1) - 2\cdot 2n + 3\cdot(2n-1) - \ldots + (2n+1)\cdot 1$.

111. Так как каждый из 5 человек может иметь не больше 4-х знакомых, то в случае, если все они имели по различному количеству знакомых, один должен иметь 4-х знакомых, а один — не иметь знакомых, что не может выполняться одновременно.

112. Да, можно. Если бы это оказалось невозможным, то среди любых 20 учеников оказались бы ученики моложе 14 лет. Выбрав 20 самых старших учеников, которым вместе по предположению меньше 280 лет, мы получим, что остальные 11 учеников моложе 14 лет и им вместе меньше $14\cdot 11 = 154$ лет, а всем 31 ученикам меньше 434 лет. Получим противоречие.

113. Если бы в каждый из 31 дней было получено не больше 5 писем, то общее количество писем не превзошло бы 155.

114. Заметим, что каждый меньший равносторонний треугольник может покрыть не больше одной вершины первоначального; следовательно, таких треугольников должно быть не меньше трех.

115. Допустим, что есть такие двузначные числа $a_1, a_2, \ldots, a_{50}$, что никакие два из них в сумме не дают 100. Среди этих чисел имеется не более девяти чисел, больших 90. Тогда все числа $100 - a_1$, $100 - a_2$, ..., $100 - a_{50}$, кроме, может быть, девяти, также двузначные (т. е. они больше 9 и меньше 100), причем они различны и не равны никаким из чисел $a_1, a_2, \ldots, a_{50}$. Мы получили $50 + 50 - 9 = 91$ двузначное число. Противоречие.

116. Количество способов вытащить две карты из 36 равно $\binom{36}{2} = 630$. Вытащить две бубны можно $\binom{9}{2} = 36$ способами, а способов вытащить две карты просто одинаковой масти в четыре раза больше, т. е. 144.

117. Событие $A \cup B$: обе карты — трефы, либо обе карты — дамы. Событие $A \cap B$ — невозможное, так как в колоде нет двух трефовых дам. Событие $A \cap C$: обе карты трефовые, причем одна из них — трефовый туз.

118. Выпишем, сколько косточек домино имеют 0, 1, 2, ..., 12 очков:

Очки	0	1	2	3	4	5	6	7	8	9	10	11	12
Количество костей	1	1	2	2	3	3	4	3	3	2	2	1	1

Пятнадцать очков получится, если суммы очков на косточках будут 3 и 12, 4 и 11, 5 и 10, 6 и 9, 7 и 8 очков. Количество таких случаев равно $2 \cdot 1 + 3 \cdot 1 + 3 \cdot 2 + 4 \cdot 2 + 3 \cdot 3 = 28$.

119. $A_1 \cup [(A_2 \cup A_3) \cap A_4]$.

120. $A_2 \cap A_3$.

121. Имеется 36 элементарных событий, сумма очков равна 7 в шести из них (1 + 6, 2 + 5, 3 + 4, 4 + 3, 5 + 2, 6 + 1). Вероятность равна $\frac{6}{36} = \frac{1}{6}$.

122. Имеется $\binom{10}{2} = 45$ способов вытащить два шара, из них $\binom{3}{2} = 3$ способа вытащить два красных шара, столько же — два синих и $\binom{4}{2} = 6$ — два зеленых. Всего $3 + 3 + 6 = 12$ способов вытащить два шара одинакового цвета. Ответ: $p = \frac{12}{45} = \frac{4}{15}$.

123. В комплекте осталась $28 - 7 = 21$ кость. Число способов выбрать пару из них равно $\binom{21}{2} = 210$. Интересующие нас пары определяются одним (из семи) совпадающим числом очков на половине кости и двумя (из оставшихся шести) несовпадающими, их число $7 \cdot \binom{6}{2} = 105$. Вероятность того, что мы возьмем одну из них, равна $\frac{105}{210} = \frac{1}{2}$.

124. Вероятность проиграть первую партию и выиграть остальные три равна $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}$, такой же результат будет еще в трех случаях: проиграна вторая партия, третья или четвертая. Итак, вероятность выиграть ровно три партии из четырех равна $\frac{4}{16} = \frac{1}{4}$. Вероятность проиграть *первые* три партии из восьми и выиграть остальные 5 равна $\frac{1}{2^8} = \frac{1}{256}$. Но число способов выбрать эти три партии из восьми равно $\binom{8}{3} = 56$, поэтому вероятность такого события равна $\frac{56}{256} = \frac{7}{32}$, что меньше $\frac{1}{4}$; следовательно, первое событие вероятнее.

125. Цифры на номере автомашины образуют числа от 0000 до 9999 — всего 10 000 номеров. Две одинаковые цифры можно выбрать десятью способами, а места их расположения $\binom{4}{2} = 6$ способами. Самую левую из оставшихся двух цифр можно выбрать 9-ю способами, а последнюю — 8-ю. Всего

$10 \cdot 6 \cdot 9 \cdot 8 = 4320$ способов. Вероятность встретить машину с таким номером равна 0,432.

126. Пусть k — коэффициент пропорциональности, тогда вероятности выпадения граней равны k, $2k$, $3k$, $4k$, $5k$ и $6k$. Их сумма равна $21k$ и должна быть равной 1, поэтому $k = \frac{1}{21}$; отсюда вероятность выпадения трех очков равна $\frac{3}{21} = \frac{1}{7}$.

127. В этом слове шесть букв, три из них — гласные, поэтому вероятность выбрать гласную букву равна 1/2.

128. После того, как сядет за стол один из указанных двух человек, для второго остается 9 мест, два из которых находятся рядом с первым, поэтому вероятность указанного события равна $\frac{2}{9}$.

129. Есть $\binom{49}{6} = 10\,487\,862$ способов выбрать 6 номеров из 49, поэтому вероятность угадать 6 номеров равна $\frac{1}{10\,487\,862} = 0{,}000\,000\,009\,53...$

130. а) Иметь среди шести карт все четыре масти можно двумя способами: либо три карты одной масти и по одной остальных, либо по две карты двух мастей и по одной из остальных двух. В первом случае имеем 4 возможности выбрать масть, представленную тремя картами, и $\binom{9}{3} = 84$ возможности выбрать эти три карты из карт этой масти, а для остальных карт по 9 возможностей появления карты каждой масти. Итого $4 \cdot \binom{9}{3} \cdot 9 \cdot 9 \cdot 9 = 244\,944$ возможностей. Во втором случае имеется $\binom{4}{2} = 6$ возможностей выбрать пару мастей, каждая из которых будут представлена двумя картами, и для каждой из этих мастей есть $\binom{9}{2} = 36$ возможностей выбора пары карт из 9. Для каждой из других двух карт опять есть по 9 вариантов выбора. Всего получается $6 \cdot 36 \cdot 36 \cdot 9 \cdot 9 = 629\,856$. Итого $244\,944 + 629\,856 = 874\,800$ возможностей. Так как имеется $\binom{36}{6} = 1\,947\,792$ возможностей получить 6 карт, то вероятность получить карты всех мастей равна $\frac{874\,800}{1\,947\,792} = 0{,}449...$.

б) Зафиксируем какую-либо одну масть (это можно сделать четырьмя способами) и вынем карты этой масти из колоды. Останется 27 карт. Иметь среди шести карт (из этих 27) все три оставшиеся масти можно тремя способами (по числу карт одной масти): $4 + 1 + 1$, $3 + 2 + 1$ и $2 + 2 + 2$. Рассуждая как и в случае а), находим, что имеется (при зафиксированной вынутой масти)

$$3 \cdot \binom{9}{4} \cdot \binom{9}{1} \cdot \binom{9}{1} + 3 \cdot \binom{9}{3} \cdot 2 \cdot \binom{9}{2} \cdot \binom{9}{1} + \binom{9}{2} \cdot \binom{9}{2} \cdot \binom{9}{2} = 240\,570$$

возможностей, а всего $240\,570 \cdot 4 = 962\,280$ возможностей. Учитывая, что имеется $\binom{36}{6}$ возможностей выбрать произвольно шесть карт, получаем, что вероятность получить шесть карт, среди которых нет ровно одной масти, равна $\frac{962\,280}{1\,947\,792} = 0{,}494...$.

131. Решим эту задачу, используя формулу Байеса. Обозначим через $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, ..., $\boldsymbol{A}_6$ события, заключающиеся в том, что письмо лежит соответственно в первом, втором, ..., шестом ящике, и через $\boldsymbol{A}_7$, что письма нет в столе. Событие $\boldsymbol{B}$ состоит в том, что в первых четырех ящиках письма нет. Тогда

$$P(\boldsymbol{A}_5/\boldsymbol{B}) = \frac{P(\boldsymbol{B}/\boldsymbol{A}_5)\cdot P(\boldsymbol{A}_5)}{\sum\limits_{i=1}^{7} P(\boldsymbol{B}/\boldsymbol{A}_i)\cdot P(\boldsymbol{A}_i)}.$$

Так как $P(\boldsymbol{A}_5) = \frac{1}{2}$, а $P(\boldsymbol{A}_i) = \frac{1}{2}\cdot\frac{1}{6} = \frac{1}{12}$ при $i = 1, 2, ..., 6$, то $P(\boldsymbol{B}/\boldsymbol{A}_i)$ равно нулю для $i = 1, 2, 3, 4$, а $P(\boldsymbol{B}/\boldsymbol{A}_7) = P(\boldsymbol{B}/\boldsymbol{A}_6) = P(\boldsymbol{B}/\boldsymbol{A}_5) = 1$. Подставляя эти значения в формулу, получаем $P(\boldsymbol{A}_5/\boldsymbol{B}) = \dfrac{\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{12}} = \dfrac{1}{8}$.

132. В этой задаче события $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$ — принадлежность ученика первому «А», первому «Б» и первому «В» классу. Событие $\boldsymbol{B}$ — встреченный ученик является девочкой. Тогда по формуле Байеса

$$P(\boldsymbol{A}_1/\boldsymbol{B}) = \frac{P(\boldsymbol{B}/\boldsymbol{A}_1)\cdot P(\boldsymbol{A}_1)}{P(\boldsymbol{B}/\boldsymbol{A}_1)\cdot P(\boldsymbol{A}_1) + P(\boldsymbol{B}/\boldsymbol{A}_2)\cdot P(\boldsymbol{A}_2) + P(\boldsymbol{B}/\boldsymbol{A}_3)\cdot P(\boldsymbol{A}_3)}.$$

Из условий задачи следует, что $P(\boldsymbol{B}/\boldsymbol{A}_1) = \frac{10}{20} = \frac{1}{2}$, $P(\boldsymbol{B}/\boldsymbol{A}_2) = \frac{8}{21}$, $P(\boldsymbol{B}/\boldsymbol{A}_3) = \frac{12}{19}$. Всего в этих классах учится 60 человек, причем 20 в 1«А», 21 в 1«Б» и 19 в 1«В», поэтому $P(\boldsymbol{A}_1) = \frac{20}{60}$, $P(\boldsymbol{A}_2) = \frac{21}{60}$ и $P(\boldsymbol{A}_3) = \frac{19}{60}$. Подставляем эти величины в формулу Байеса и получаем: $P(\boldsymbol{A}_1/\boldsymbol{B}) = \dfrac{\frac{1}{2}\cdot\frac{20}{60}}{\frac{1}{2}\cdot\frac{20}{60}+\frac{8}{21}\cdot\frac{21}{60}+\frac{12}{19}\cdot\frac{19}{60}} = \dfrac{10}{30} = \dfrac{1}{3}$.

Тот же результат можно получить и проще, заметив, что всего в первых классах 30 девочек, 10 из которых учатся в 1«А».

133. Обозначим через $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$ события, состоящие в том, что промахнулся соответственно первый, второй и третий стрелок. Через $\boldsymbol{B}$ обозначим событие, состоящее в том, что попали только двое. Тогда искомая вероятность есть

$$P(\boldsymbol{A}_3/\boldsymbol{B}) = \frac{P(\boldsymbol{B}/\boldsymbol{A}_3)\cdot P(\boldsymbol{A}_3)}{P(\boldsymbol{B}/\boldsymbol{A}_1)\cdot P(\boldsymbol{A}_1) + P(\boldsymbol{B}/\boldsymbol{A}_2)\cdot P(\boldsymbol{A}_2) + P(\boldsymbol{B}/\boldsymbol{A}_3)\cdot P(\boldsymbol{A}_3)}.$$

Далее получаем:

$$P(\boldsymbol{B}/\boldsymbol{A}_1) = 0{,}1\cdot 0{,}8\cdot 0{,}7,\ P(\boldsymbol{B}/\boldsymbol{A}_2) = 0{,}9\cdot 0{,}2\cdot 0{,}7,\ P(\boldsymbol{B}/\boldsymbol{A}_3) = 0{,}9\cdot 0{,}8\cdot 0{,}3;$$

$$P(A_3/B) = \frac{0{,}9 \cdot 0{,}8 \cdot 0{,}3 \cdot 0{,}3}{0{,}1 \cdot 0{,}8 \cdot 0{,}7 \cdot 0{,}1 + 0{,}9 \cdot 0{,}2 \cdot 0{,}7 \cdot 0{,}2 + 0{,}9 \cdot 0{,}8 \cdot 0{,}3 \cdot 0{,}3} =$$
$$= \frac{648}{956} = \frac{162}{239} = 0{,}6778...$$

134. Обозначим через A_1, A_2, A_3, A_4 события, состоящие в том, что младший брат взял соответственно 0, 1, 2, 3 конфеты «Каракум». Через B обозначим событие, состоящее в том, что старший взял конфету «Каракум». Тогда по формуле полной вероятности

$$P(B) = P(B/A_0)\cdot P(A_0) + P(B/A_1)\cdot P(A_1) + P(B/A_2)\cdot P(A_2) + P(B/A_3)\cdot P(A_3).$$

Вычислим $P(A_i)$ для $i = 0, 1, 2, 3$. Всего взять три конфеты можно $\binom{15}{3} = 455$ способами, из них:

0 «Каракумов» и 3 «Мишки» $\binom{5}{0}\cdot\binom{10}{3} = 120$ способами;

1 «Каракум» и 2 «Мишки» $\binom{5}{1}\cdot\binom{10}{2} = 225$ способами;

2 «Каракума» и 1 «Мишку» $\binom{5}{2}\cdot\binom{10}{1} = 100$ способами;

3 «Каракума» и 0 «Мишек» $\binom{5}{3}\cdot\binom{10}{0} = 10$ способами.

Отсюда $P(A_0) = \frac{120}{455} = \frac{24}{91}$, $P(A_1) = \frac{225}{455} = \frac{45}{91}$, $P(A_2) = \frac{100}{455} = \frac{20}{91}$, $P(A_3) = \frac{10}{455} = \frac{2}{91}$. Найдем $P(B/A_i)$. В случае события A_0 остается 5 «Каракумов» из 12 конфет, в случае A_1 — 4 «Каракума», в случае A_2 — 3 «Каракума», в случае A_3 — 2 «Каракума», поэтому

$$P(B/A_0) = \frac{5}{12},\ P(B/A_1) = \frac{4}{12},\ P(B/A_2) = \frac{3}{12},\ P(B/A_3) = \frac{2}{12}$$

и $P(B) = \frac{5}{12}\cdot\frac{24}{91} + \frac{4}{12}\cdot\frac{45}{91} + \frac{3}{12}\cdot\frac{20}{91} + \frac{2}{12}\cdot\frac{2}{91} = \frac{364}{12\cdot 91} = \frac{1}{3}$.

Заметим, что та же вероятность 1/3 получится и в том случае, если старший брат будет брать конфету первым. В самом деле, поскольку мы не знаем, каков результат выборки, осуществленной младшим братом, то событие B можно считать независимым от A_i.

135. Пусть событие A_1 — сумма очков на кости при двух бросаниях больше 10, A_2 — эта сумма не превосходит 10. Событие B — при бросаниях один раз выпала шестерка. Тогда $P(A_1/B) = \frac{P(B/A_1)\cdot P(A_1)}{P(B/A_1)\cdot P(A_1) + P(B/A_2)\cdot P(A_2)}$.
Сумма очков может быть больше десяти в трех случаях (6 + 6, 6 + 5, 5 + 6) из 36 возможных, поэтому $P(A_1) = \frac{1}{12}$, $P(A_2) = \frac{11}{12}$. Поскольку из трех случаев лишь в двух присутствует одна шестерка, $P(B/A_1) = \frac{2}{3}$, а $P(B/A_2) = \frac{10}{33}$, так как среди 33 оставшихся случаев в 10 присутствует одна шестерка. Получаем

$$P(A_1/B) = \frac{\frac{2}{3} \cdot \frac{1}{12}}{\frac{2}{3} \cdot \frac{1}{12} + \frac{10}{33} \cdot \frac{11}{12}} = \frac{1}{6}.$$

Если же первой выпала шестерка, то, чтобы сумма была больше 10, второй должна быть пятерка или шестерка, вероятность этого $\frac{1}{3}$.

136. Вероятность того, что из n пущенных торпед ни одна не попадет в цель, равна $\left(\frac{2}{3}\right)^n$. Эта величина должна быть меньше, чем 0,1, т. е. $\left(\frac{2}{3}\right)^n < 0{,}1$, или $n \lg \frac{2}{3} < -1$. Но $\lg \frac{2}{3} \approx -0{,}1760912$, и потому n должно быть больше $\frac{-1}{\lg \frac{2}{3}} \approx 5{,}6788$, а так как n — целое число, то $n \geq 6$.

137. Здесь мы имеем схему Бернулли с вероятностями p и q, равными $\frac{1}{2}$, поэтому вероятность правильного ответа равна $P(7) + P(8) + P(9) + P(10) =$ $= \binom{10}{7} \cdot \frac{1}{2^{10}} + \binom{10}{8} \cdot \frac{1}{2^{10}} + \binom{10}{9} \cdot \frac{1}{2^{10}} + \binom{10}{10} \cdot \frac{1}{2^{10}} = \frac{176}{1024} = \frac{11}{64} \approx 0{,}1718$.

138. Вероятность выхода из строя обоих моторов двухмоторного самолета равна p^2, а вероятность выхода из строя трех или четырех моторов у четырехмоторного самолета равна $4p^3(1-p) + p^4$. Условие большей безопасности двухмоторного самолета запишется так: $p^2 < 4p^3(1-p) + p^4$. Считая, что $p \neq 0$ (в случае $p = 0$ оба самолета одинаково безопасны), получаем $1 < 4p(1-p) + p^2$, или $3p^2 - 4p + 1 < 0$. Это может быть лишь при $1 > p > \frac{1}{3}$.

139. Можно считать, что булочка состоит из 100 частей, каждая из которых может быть тестом или изюминкой, поэтому воспользуемся распределением Пуассона с $\lambda = 3$. Вероятность отсутствия изюминок равна $P(0) = e^{-\lambda} = e^{-3} = 0{,}0497... \approx 0{,}05$.

140. Вероятность того, что у всех людей в выборке из n человек кровь не принадлежит указанной группе, равна $(0{,}999)^n$. Найдем n, при котором $(0{,}999)^n < 0{,}5$. Это будет если $n \lg 0{,}999 < \lg 0{,}5$, или $-0{,}0004345n < -0{,}693147$, откуда $n > 1595{,}2754$, т. е. $n \geq 1596$.

141. Так как вероятность выпадения любого количества очков от 1 до 6 одна и та же и равна 1/6, то математическое ожидание числа очков равно $\frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{7}{2} = 3{,}5$.

Дисперсия равна $\frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - 3{,}5^2 = \frac{35}{12}$.

142. В задаче 126 получено, что вероятность выпадения грани с k очками в этой ситуации равна $k/21$, поэтому математическое ожидание числа очков при бросании такого кубика равно $\frac{1}{21}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{13}{3}$. Дисперсия равна $\frac{1}{21}(1^3+2^3+3^3+4^3+5^3+6^3)-\left(\frac{13}{3}\right)^2=21-\left(\frac{13}{3}\right)^2=\frac{20}{9}=2{,}22...$, т. е. эта дисперсия *меньше*, чем для кубика с несмещенным центром тяжести (задача 141).

143. Игра описывается схемой Бернулли с $p=\frac{1}{6}$, $q=\frac{5}{6}$ и тремя испытаниями. Вероятности 0, 1, 2 и 3 выпадений заданного числа очков соответственно равны $\left(\frac{5}{6}\right)^3$, $3\cdot\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)$, $3\cdot\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^2$ и $\left(\frac{1}{6}\right)^3$, т. е. $\frac{125}{216}$, $\frac{75}{216}$, $\frac{15}{216}$ и $\frac{1}{216}$. Математическое ожидание величины выигрыша равно

$$(-1)\cdot\frac{125}{216}+1\cdot\frac{75}{216}+2\cdot\frac{15}{216}+3\cdot\frac{1}{216}=\frac{-17}{216}\approx -0{,}079 \text{ (долларов).}$$

144. Числовой ответ на этот вопрос зависит от величины выплат за угаданное количество номеров и стоимость игральной карточки. Если обозначим эту стоимость через N, а величину выигрыша при угадывании p чисел из k выигрышных при n возможных номерах (в игре «Спортлото» есть два варианта: $k=5$, $n=36$ и $k=6$, $n=49$) — через A_p, то искомое математическое ожидание равно $\sum\limits_{p=0}^{k} A_p \frac{\binom{k}{p}\binom{n-k}{k-p}}{\binom{n}{k}}-N$.

145. Вероятность выпадения герба лишь при k-м бросании равна $\left(\frac{1}{2}\right)^k$, поэтому математическое ожидание числа бросаний равно $\sum\limits_{k=1}^{\infty}\frac{k}{2^k}$. Чтобы вычислить эту сумму, заметим, что при $|x|>1$ будет $\sum\limits_{k=1}^{\infty}\frac{1}{x^k}=\frac{1}{x-1}$. Беря производную обеих частей равенства, получаем $-\sum\limits_{k=1}^{\infty}\frac{k}{x^{k+1}}=\frac{-1}{(x-1)^2}$, следовательно,

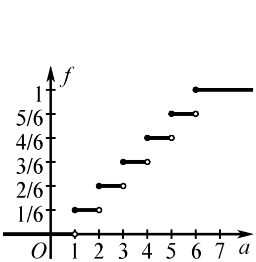


Рис. 479

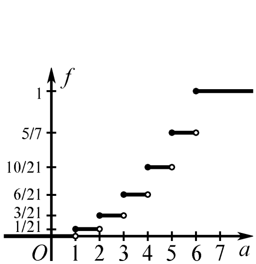


Рис. 480

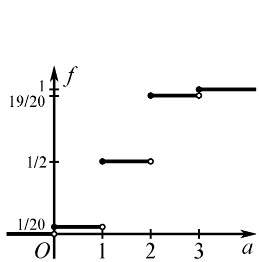


Рис. 481

$\sum_{k=1}^{\infty} \frac{k}{x^k} = \frac{x}{(x-1)^2}$; при $x = 2$ получаем $\sum_{k=1}^{\infty} \frac{k}{2^k} = 2$. Итак, математическое ожидание числа бросаний монеты до первого появления герба равно 2.

146. График изображен на рис. 479.

147. См рис. 480.

148. См рис. 481.

149. Вероятность попадания в область вне интервала $(-3; 3)$ равна 0,0026.

150. Так как функция $y = a \operatorname{arctg} x + b$ монотонная, то для ее возрастания необходимо, чтобы a было больше нуля. Кроме того, необходимо, чтобы $y \to 0$ при $x \to -\infty$ и $y \to 1$ при $x \to \infty$.

Так как $\operatorname{arctg}(-\infty) = \frac{-\pi}{2}$, а $\operatorname{arctg}(+\infty) = \frac{\pi}{2}$, то должны выполняться условия $\frac{-\pi}{2} a + b = 0$ и $\frac{\pi}{2} a + b = 1$, откуда $a = \frac{1}{\pi}$, $b = \frac{1}{2}$, т. е. $y = \frac{1}{\pi} \operatorname{arctg} x + \frac{1}{2}$.

151. На шахматной доске поровну черных и белых клеток Каждая доминошка закрывает одну черную и одну белую клетку, поэтому не могут остаться две клетки одинакового цвета.

152. Заметим, что сумма четного количества нечетных чисел — четное число, поэтому сумма 10 купюр не может равняться 25 рублям.

153. В этом ряду нечетное количество нечетных чисел, и при любой расстановке знаков «+» и «–» получится нечетное число, а нуль — четное число.

154. Если бы каждое звено ломаной пересекалось ровно с одним из остальных звеньев, то звенья разбились бы на пары пересекающихся звеньев, а 9 — число нечетное.

155. Заметим, что в этом случае количество арбузов в двух корзинах, между которыми стоит четное число других корзин, должно отличаться на нечетное число арбузов, а если стоит нечетное число корзин, то на четное. Но так как корзины стоят по кругу и их 99, то с одной стороны между двумя корзинами стоит четное число корзин, а с другой (все остальные) — нечетное.

156. Заметим, что необходимое количество информации для выделения радиоактивного шарика равно $\log_2 30 = 4{,}906...$, а каждое измерение дает один бит информации, поэтому необходимо провести 5 измерений. Это можно сделать, разбивая шарики последовательно на две группы, равные по числу шариков или отличающиеся на 1.

В случае двух радиоактивных шариков количество возможных пар равно $\binom{30}{2} = 435$, а $\log_2 435 \approx 8{,}76$. Значит, необходимо произвести не менее 9 измерений. Однако, разделив шарики на две группы по 15 шариков и получив, что в одной группе есть радиоактивный шарик (хотя бы один), необходимо проверить, есть ли радиоактивный шарик во второй группе. Если он там тоже есть, то для поиска обоих шариков понадобится, очевидно, 10 измерений.

157. За 9 попыток. Сначала открываем книгу между 500 и 501 страницами. Если там только первоначальный шрифт, то открываем между 750 и 751, а если уже другой, то между 250 и 251 страницами. Продолжая так же делить пополам ту часть книги, где имеются оба шрифта, не более чем за 9 попыток раскроем книгу на том месте, где начинаются ответы.

Отметим, что $1000 < 1024 = 2^{10}$, а мы одновременно видим две страницы.

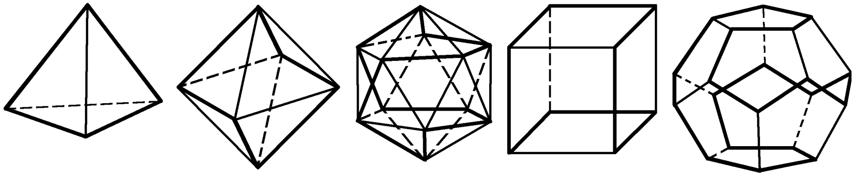


Рис. 482

158. Отметим, что $128 < 200 < 256$, т. е. $2^7 < 200 < 2^8$, поэтому за 8 взвешиваний можно найти массу предмета. Сначала кладем на чашку весов 128 г; если предмет легче, то берем 64 г, а если тяжелее, то 128 + 64 = 192 г, и т. д.

159. Так как $2^9 < 1000 < 2^{10}$, то понадобится 10 гирек, например, с массами 1, 2, 4, 8, 16, 32, 64, 128, 256, и 512 граммов.

160. Поскольку после каждой игры выбывает один участник, то будет проведено 49 игр. Количество туров — число k — наименьшая степень числа 2, при которой $2^k > 50$, откуда $k = 6$.

161. Всего существует 5! = 120 расположений грузиков по массе. Так как $\log_2 120 = 6{,}9068...$, то необходимо не менее 7 взвешиваний, поскольку каждое взвешивание двух грузиков уменьшает число вариантов ровно вдвое. Семи взвешиваний оказывается достаточно. Для этого сначала взвешиваем одну пару грузиков и еще одну пару других грузиков. Потом пятый грузик с одним из них. За оставшиеся 4 взвешивания можно упорядочить грузики, уменьшая каждый раз количество вариантов вдвое.

162. Да, можно. Положим на чашки весов по 4 монеты. Либо весы окажутся в равновесии, либо перетянет одна из чашек. В первом случае либо монеты на чашках настоящие, либо там лежит по одной фальшивой монете. Положим монеты, лежавшие на одной из чашек, на весы по две на каждой чашке. Если весы в равновесии, то все исследованные 8 монет настоящие, а фальшивые могут быть лишь среди оставшихся двух монет. Сравнив третьим взвешиванием их с двумя настоящими, мы определим, являются ли они настоящими, а если нет, то тяжелее или легче они настоящих. Если при втором взвешивании перетянет одна из чашек, то среди этих четырех монет одна фальшивая, а 2 монеты, которые мы еще не взвешивали, настоящие. Третьим взвешиванием сравниваем их с монетами на одной из чашек и определяем, тяжелее фальшивые монеты или легче, чем настоящие.

Рассмотрим второй случай, когда при первом взвешивании перетянула одна из чашек. В этом случае на одной из чашек монеты настоящие, а на другой есть одна или две фальшивых монеты. Разложим 4 монеты с перетянувшей чашки весов по 2 на чашки весов. Если одна из чашек весов перетянет, то фальшивая монета находится на этой чашке и она более тяжелая, чем настоящая. Если же весы окажутся в равновесии, то либо там две фальшивых монеты по одной на каждой из чашек, либо они настоящие, а фальшивые легче настоящих. Осталось сравнить третьим взвешиванием две монеты, лежавшие на одной чашке.

163. Возьмем из первого мешочка одну монету, из второго — две, и т. д., из десятого все десять и положим их на весы. Настоящие монеты весили бы 55 граммов. Номер мешочка с фальшивыми монетами равен разности между 55 и количеством граммов, показанных весами.

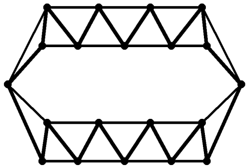


Рис. 483

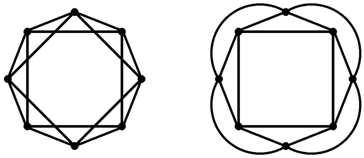


Рис. 484

164. Да, можно. Аналогично предыдущему, возьмем из первого мешочка 0 монет, из второго 1, и т. д., из десятого 9. Тогда разность между 45 и показаниями весов будет на единицу меньше номера мешка с фальшивыми монетами.

165. Только при $n \geq 9$.

166. Заметим, что количество посланных открыток равно $3 \cdot 7 = 21$ — нечетному числу. Если бы каждый получил открытку именно от тех, кому он посылал сам, то открыток было бы четное число. Эта задача сводится к вопросу о существовании графа с 7 вершинами и индексом каждой вершины, равным 3.

167. Количество сыгранных каждой командой матчей может принимать значения от 0 до 29, но одновременно значения 0 и 29 никакие две команды иметь не могут, поэтому возможных значений 29, а команд 30, и по принципу Дирихле найдутся две команды, сыгравшие одинаковое количество матчей.

168. Так как вершин 5 и имеется вершина с индексом 4, то она связана ребрами со всеми остальными вершинами, но есть и вершина с индексом 0, т. е. не связанная ни с одной. Противоречие.

169. Так как количество ребер графа равно половине суммы индексов его вершин, то количество нечетных слагаемых в этой сумме должно быть четно.

170. Задача аналогична задаче 167. Индекс вершины графа с n вершинами не превосходит $n - 1$, поэтому он может принимать значения от 0 до $n - 1$. Если все значения индексов, кроме двух, различны, то их $n - 1$, и потому в их число должно входить либо число 0, либо число $n - 1$.

171. Утверждение очевидно.

172. Нет, у него больше двух вершин с нечетными индексами.

173. Четные индексы вершин имеет только граф октаэдра, поэтому только он уникурсален. Все пять графов гамильтоновы (см. рис. 482).

174. Да, верно.

175. Да, гамильтонов путь указан на рисунке 483.

176. Это частный случай задачи 180.

177. Занумеруем последовательно вершины семиугольника цифрами от 1 до 7. Теперь выбросим отрезки 1—7, 6—7, 2—3, 4—5, а объединение ребер 5—7 и 7—2 будем считать одним ребром, соединяющим вершины 5 и 2. Полученный граф (содержащийся в исходном графе) изоморфен графу «домики и колодцы» (2, 3, 6 — домики, 1, 4, 5 —колодцы).

178. См. рис. 484.

179. Если в правильном n-угольнике провести наименьшие диагонали, то при четном n граф плоский, при нечетном — не плоский.

180. Рассмотрим три соседние вершины и занумеруем их цифрами 1, 5 и 3. Диаметрально противоположные им вершины занумеруем, соответственно, цифрами 4, 2 и 6. Выбросим теперь все остальные вершины и ребра, кроме выходящих из этих вершин. Кроме того, из цепочки ребер, соединяющих по

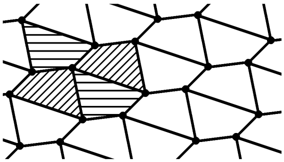


Рис. 485

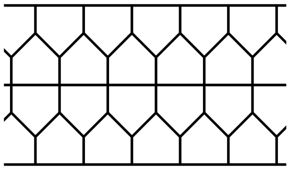


Рис. 486

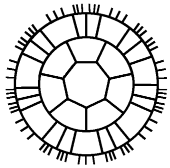


Рис. 487

периметру вершины 1 и 6, составим одно ребро, и из цепочки ребер, соединяющих вершины 3 и 4, также составим ребро. Полученный граф из шести вершин будет графом «домики и колодцы».

181. Если граф состоит из n компонент, то $В - Р + Г = n + 1$.

182. Обозначим количество треугольников через x. Подсчитаем число ребер графа. У каждого треугольника 3 ребра, из них n ребер принадлежат периметру многоугольника, тогда $3x - n$ ребер лежат внутри, причем каждое из них посчитано дважды, следовательно, внутри находится лишь $\frac{3x-n}{2}$ ребер, а всего имеется $\frac{3x-n}{2} + n = \frac{3x+n}{2}$ ребер. Воспользуемся формулой Декарта—Эйлера, заменив $Г$ на $x + 1$: $n + m - \frac{3x+n}{2} + x + 1 = 2$. Отсюда $x = n + 2m - 2$.

183. Так как индексы всех вершин равны трем, то $3В = 2Р$. Подставив значение $В = \frac{2}{3}Р$ в формулу Эйлера, получим, что $Г - \frac{Р}{3} = 2$. Поскольку у каждой грани не более 5 ребер, то удвоенное количество ребер не больше $5Г$, т. е. $Р \le \frac{5}{2}$, отсюда $Г - \frac{5}{2 \cdot 3}Г \le 2$ и потому $Г \le 12$. Двенадцать пятиугольных граней имеет додекаэдр, а его проекция (как на рис. 277) имеет 11 пятиугольных граней.

184. $В - Г + Р = 0$.

185. $В - Г + Р = n - 1$, где n — число компонент графа.

186. У тетраэдра — тетраэдр, у куба — октаэдр, у октаэдра — куб, у икосаэдра — додекаэдр, у додекаэдра — икосаэдр.

187. См. рис. 485 и рис. 296.

188. Да, см. рис. 486.

189. Нет. Допустим, что замостить плоскость конгруэнтными выпуклыми семиугольниками возможно. Подсчитаем сумму углов у n семиугольников, пересекающихся с кругом достаточно большого радиуса R (число n этих семиугольников примерно равно $\frac{2\pi R}{S}$, где S — площадь семиугольника). Эта сумма углов равна $5\pi n$. С другой стороны, в каждой вершине сходится не менее 3 ребер, т. е. она принадлежит не менее чем трем семиугольникам. Поэтому число вершин у взятых семиугольников не более $\frac{7}{3}n$, а сумма углов в вершинах не больше $\frac{2\pi n \cdot 7}{3} = \frac{14\pi n}{3} < 5\pi n$. Правда, мы здесь не учитываем,

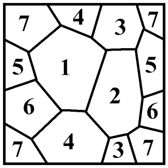


Рис. 488

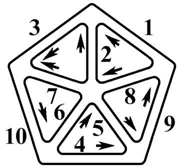


Рис. 489

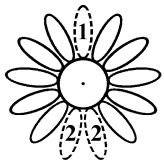


Рис. 490

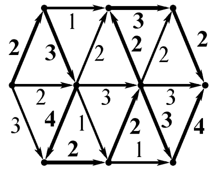


Рис. 491

что у крайних семиугольников (пересекающихся с границей взятого круга) в вершинах может сходиться только два ребра. Но число таких семиугольников примерно равно kR (где k — некоторое число), т. е. примерно равно $k'\sqrt{n}$ и потому сумма углов равна $\frac{14\pi n}{3} + q\sqrt{n}$, что при большом n (т. е. для большого радиуса R) все равно меньше $5\pi n$. Полученное противоречие и дает решение.

190. Да, см. рис. 487 (построение облегчает сеть концентрических окружностей).

191. Доказательство проведем по индукции. Для $n = 1$ это очевидно. Пусть это верно для $n = k$, докажем, что это верно для $n = k + 1$. Рассмотрим карту, образованую $k + 1$ прямой. Уберем одну из них и раскрасим получившуюся карту в два цвета (это возможно согласно предположению индукции). Теперь восстановим убранную прямую и перекрасим одну полуплоскость в противоположные цвета, а другую оставим в прежней раскраске. Полученная раскраска будет правильной.

192. Та же идея, что и в задаче 191.

193. Пусть карта из n окружностей с диаметрами раскрашена в три краски, которые просто занумеруем: 1, 2, 3. Добавим еще одну окружность с диаметром. Окраску вне окружности оставим прежней, в первом полукруге сделаем замену красок $1 \to 2$, $2 \to 3$, $3 \to 1$, а во втором $1 \to 3$, $3 \to 2$, $2 \to 1$. Полученная раскраска правильная. Осталось применить метод математической индукции.

194. Обозначим цвета карты A цифрами 1, 2, 3, 4. Теперь с одной стороны от проведенной прямой l оставим раскраску карты B прежней, а с другой сдвинем все цвета по кругу ($1 \to 2$, $2 \to 3$, $3 \to 4$, $4 \to 1$). Тогда страны карты B, имеющие общей границей отрезок, не лежащий на прямой l, находятся по одну сторону от l, их цвета не изменились или оба сдвинулись в одну сторону и потому различны. Страны же карты B, граничащие по отрезку прямой l, раньше имели один цвет, а после сдвига цветов также имеют различные цвета.

195. Нужная карта указана на рис. 488, где изображена развертка тора на плоскость. Боковые стороны квадрата, как и горизонтальные его стороны, считаются склеенными.

196. Да, можно. Предположим, что для некоторого турнира это оказалось невозможным. Возьмем цепочку максимальной длины. По предположению, в нее войдут не все команды. Рассмотрим одну из команд, не вошедших в цепочку. Если она выиграла у первой команды, то ее можно поставить впереди этой команды, в результате цепочка станет длиннее. Если она ей проиграла, то найдем команду с наибольшим номером, у которой она выиграла, и поставим нашу команду в цепочку перед этой командой. Цепочка вновь удлинится. Если она проиграла всем командам из цепочки, то нашу команду можно поставить последней в цепочке, вновь удлинив ее.

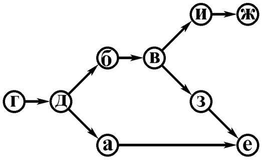


Рис. 492

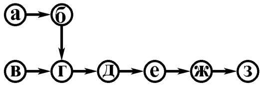


Рис. 493

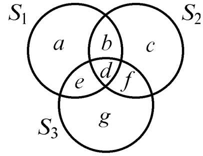


Рис. 494

197. См. рис. 489.

198. Рассмотрим команды A и B, одержавшие одинаковое число побед. Пусть A выиграла у B, покажем, что найдется команда C, которая проиграла B, но выиграла у A. Действительно, если всякая команда, проигравшая B, проиграла и A, то у команды A больше побед, чем у команды B, что противоречит условию.

199. Да, возможно, по маршруту Меркурий — Плутон — Венера — Марс — Земля.

200. Задача эквивалентна задаче 196.

201. Выигрывает начинающий. Он ставит крестик на центральное поле. Противник вынужден ставить нолик на одно из полей. В ответ ставится крестик на поле, симметричное тому, на которое уже был поставлен нолик. Дальнейшие ходы легко проследить.

202. Выигрывает начинающий. Он берет первым ходом одну спичку, а затем всякий раз берет разность между числом 4 и количеством спичек, взятых противником его последним ходом.

203. Предположим, что у игрока, делающего ход вторым, есть выигрышная стратегия. Тогда начинающий делает два хода конем: на некоторое поле и обратно. В результате получается начальная позиция, но уже с первым ходом противника.

204. Выигрывает второй. После хода первого игрока он отрывает один или два лепестка так, чтобы оставшиеся лепестки разбились на две одинаковые группы (см. рис. 490), а далее он отрывает лепестки, симметричные тем, которые оторвал первый.

205. Проигрывает второй. Начинающий соединяет две диаметрально противоположные точки, а далее делает ходы, симметричные ходам второго относительно этого диаметра.

206. См. рис. 491. Время равно 20, причем критических путей — *два*.

208. См. рис. 492.

209. Достаточно двух компьютеров. Вычисления займут 50 секунд. Первый вычисляет $A \to D \to E \to F \to X_1$, а второй $B \to C \to H \to G \to X_2$.

210. См. рис. 493.

Глава III

211. Если марсиан не существует, то у *каждого* из них три глаза, т. е. ответ утвердительный.

212. Множество истинности этого высказывания состоит из чисел 3, 5, 6 и всех целых чисел, больших семи.

213. $(\exists\, x)\ (ax = b)$. Множество истинности этого высказывания: $a \neq 0$, b — любое, а также $a = 0$, $b = 0$.

214. $(\exists\, x>0)\,(ax=b)$. Множество истинности этого высказывания состоит из всех пар $(a; b)$, удовлетворяющих одному из следующих условий: 1) $a>0$, $b>0$; 2) $a<0$, $b<0$; 3) $a=0$, $b=0$.

215. $(\exists\, n\in \boldsymbol{Z})\,(\forall\, m\in \boldsymbol{Z})\left((m\neq 0)\Rightarrow\left(\dfrac{n}{m}\in \boldsymbol{Z}\right)\right)$. Высказывание верно. Таким числом является число ноль.

216. $(\forall\, n\in \boldsymbol{Z})\left(\dfrac{n+(n+1)+(n+2)}{3}\in \boldsymbol{Z}\right)$.

217. $(\forall\, \triangle ABC)$ $(\exists!$ т. $O)$ $(d(O,$ пр. $AB)=d(O,$ пр. $BC)=d(O,$ пр. $AC))$, т. е. расстояния от точки O до прямых AB, BC, AC одинаковы.

218. $(\forall\, r\in \boldsymbol{N}, 0\le r<12)\,(\exists\, p\in \boldsymbol{R}, 0\le p<1)\left(\left\{\dfrac{r+p}{12}\right\}=p\right)$ (здесь $\{\}$ — дробная часть числа, $r+p$ — время в часах). Это высказывание ложно, так как в течение часа от 0 ч 01 мин до 1 ч 01 мин стрелки ни разу не совпадут.

219. $(\forall\, a, b, c\in \boldsymbol{Z})\left(\left(\dfrac{a+b}{2}\in \boldsymbol{Z}\right)\vee\left(\dfrac{a+c}{2}\in \boldsymbol{Z}\right)\vee\left(\dfrac{b+c}{2}\in \boldsymbol{Z}\right)\right)$ — истинно.

220. Рассмотрим такие точки x_1 и x_2, что $x_1=-x_2$, $x_2>1$, а $\dfrac{x_2}{3}$ больше любого из чисел $|a|$, $|b|$, $|c|$. Тогда $y(x_2)>0$, а $y(x_1)<0$. По теореме Больцано многочлен обращается в нуль в некоторой точке отрезка $[x_1; x_2]$.

221. «Есть что-то лучше плохой погоды».

222. «Некоторому овощу — не свое время».

223. «Существует сто попарно неравных двузначных чисел» (это неверно).

224. «Если в треугольнике квадрат длины одной из сторон равен сумме квадратов длин двух других его сторон, то этот треугольник — прямоугольный».

225. Если непрерывная функция на отрезке $[a; b]$ обращается в нуль в некоторой точке этого отрезка, то она принимает в концах отрезка $[a; b]$ значения разных знаков.

226. Да, поскольку обе теоремы верны.

227. Обратная теорема: «Если диагонали четырехугольника перпендикулярны, то этот четырехугольник — ромб» (неверна).

Противоположная теорема: «Если четырехугольник — не ромб, то его диагонали не перпендикулярны» (неверна).

Теорема, обратная противоположной: «Если диагонали четырехугольника не перпендикулярны, то этот четырехугольник — не ромб» (верна).

228. Обратная теорема: «Если диагонали параллелограмма равны, то этот параллелограмм является прямоугольником».

Противоположная теорема: «Если параллелограмм не является прямоугольником, то его диагонали не равны».

Теорема, обратная противоположной: «Если диагонали параллелограмма не равны, то он не является прямоугольником».

Все эти предложения истинны, т. е. равенство диагоналей параллелограмма является необходимым и достаточным условием для того, чтобы он был прямоугольником.

229. Условие, заключающееся в том, что целое число оканчивается на нуль, является достаточным, но не необходимым для делимости этого числа на 5.

Обратное утверждение: «Если целое число делится на 5, то оно оканчивается на нуль».

Противоположное утверждение: «Если целое число не оканчивается на нуль, то оно не делится на 5».

Обратное к противоположному: «Если целое число не делится на 5, то оно не оканчивается на нуль».

230. Это утверждение верно (поскольку если два треугольника конгруэнтны, то их соответствующие элементы также конгруэнтны).

231. Это высказывание истинно.

232. Первое истинно, второе ложно.

233. Всегда.

234. Верны все три высказывания.

235. Если ***A*** ложно, то ложны и оба высказывания в скобках, поэтому ложна и их конъюнкция, а если ***A*** — истинно, то истинно либо ***B***, либо $\neg$ ***B***, и поэтому истинно одно из высказываний в скобках, а следовательно, и их конъюнкция.

236. Да, будет, так как для введенного расстояния выполнены все три аксиомы.

237. а) Да, поскольку это обычное расстояние. б) Да, если брать кратчайшую из этих дуг.

238. Да, будет. Выполнение двух первых аксиом очевидно. Покажем, что выполнена и аксиома треугольника. Пусть имеются три круга S_1, S_2, S_3 (см. рис. 494). Обозначим через a, b, c, d, e, f, g площади кусков плоскости, на которые эти круги ее разбивают. Тогда

$$\rho(S_1, S_2) = a + c + e + f,\ \rho(S_2, S_3) = b + c + e + g,\ \rho(S_1, S_3) = a + b + f + g.$$

Отсюда $\rho(S_1, S_2) + \rho(S_2, S_3) = a + b + 2c + 2e + f + g$, что не меньше, чем $\rho(S_1, S_3)$, и равняется ему, если $c = 0$ и $e = 0$.

239. Да, образуют.

240. Да, образуют.

241. Нет, поскольку сумма двух нечетных чисел четна.

242. Нет, поскольку композиция поворота на 60° и последующей симметрии не совпадает с композицией симметрии относительно той же оси и последующего поворота на 60°.

243. Нет, так как не для всякого нечетного числа существует обратное в этом множестве.

244. Нет, поскольку при $a > 2$ равенство $a + 0 = a$ не выполняется.

245. Нет, так как не существует нуля.

Предметный указатель