Г. РАДЕМАХЕР и О. ТЕПЛИЦ

ЧИСЛА И ФИГУРЫ

БИБЛИОТЕКА МАТЕМАТИЧЕСКОГО КРУЖКА
ВЫПУСК 10

Г. РАДЕМАХЕР и О. ТЕПЛИЦ

ЧИСЛА И ФИГУРЫ

ОПЫТЫ МАТЕМАТИЧЕСКОГО МЫШЛЕНИЯ

ИЗДАНИЕ ТРЕТЬЕ

Перевод с немецкого
В. И. КОНТОВТА

Под редакцией,
с дополнениями и примечаниями
И. М. ЯГЛОМА

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1962

АННОТАЦИЯ

Книга содержит 27 маленьких очерков, посвященных различным вопросам математики. Каждый из них представляет образец изящного и доступного научного исследования; для чтения их не требуется никакой специальной математической подготовки — достаточно знаний, приобретенных в средней школе. Ценность книги состоит в том, что она не только знакомит читателя с материалом, над которым работает наука, но и показывает научные методы в действии. С этой стороны книга представляет исключительное явление в мировой научно-популярной литературе.

Радемахер Ганс, Теплиц Отто.

Числа и фигуры. Опыты математического мышления.

Редактор *И. Е. Морозова.*

Техн. редактор *В. Н. Крючкова.* Корректор *З. В. Автонеева.*

Сдано в набор 3/I 1962 г. Подписано к печати 17/III 1962 г. Бумага $84 \times 108^1/_{32}$. Физ. печ. л. 8,25. Условн. печ. л. 13,53. Уч.-изд. л. 14,52. Тираж 40 000 экз. Цена книги 54 коп. Заказ № 2554.

Государственное издательство физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Первая Образцовая типография имени А. А. Жданова
Московского городского совнархоза.
Москва, Ж-54, Валовая, 28.

СОДЕРЖАНИЕ

ОТ РЕДАКТОРА

Книга видных немецких ученых Ганса Радемахера и Отто Теплица «Числа и фигуры» занимает в ряду научно-популярных сочинений по математике совершенно особое место. Вышедшая первым изданием еще в 1930 г. и затем неоднократно переиздававшаяся и переводившаяся, эта книга вполне может быть включена в число «классических» сочинений, хорошо известных всем, интересующимся вопросами популяризации математики, и оказавших значительное влияние на всю последующую литературу такого рода. Очень большое влияние оказала эта книга, ранее уже дважды издававшаяся на русском языке (в 1936 и в 1938 гг.), и на нашу научно-популярную литературу, в частности на серию книг «Библиотека математического кружка».

Весьма удачной следует признать основную идею авторов — создание своеобразной «математической хрестоматии» из ряда внешне не связанных между собой отрывков, излагающих изолированные вопросы, относящиеся к разным разделам математики. Все эти отрывки в совокупности должны создать у читателя достаточно цельное впечатление, если не о математической науке, то о математическом мышлении, ознакомление с которым является значительно более важной задачей, чем просто ознакомление с математическими фактами. И этой последней цели книга Г. Радемахера и О. Теплица достигает в наилучшей степени — это обеспечивается и очень тщательным подбором тем, весьма элементарных по используемому аппарату, но достаточно глубоких и содержательных по существу затрагиваемых проблем, и продуманным изложением, выделяющим узловые моменты доказательств и подчеркивающим идейную сторону вопроса. Эта книга впервые раскрыла все возможности, заложенные в подобной системе изложения — и за ней последовали многочисленные «математические хрестоматии» сходного рода, ни одна из которых, впрочем, не имела успеха этой книги. Сильно сказалось появление книги Радемахера и Теплица и на деятельности школьных и студенческих математических кружков, культивируя в них изложение разрозненных «математических этюдов» в противоположность кружкам с четко очерченной тематикой.

Настоящее, третье, русское издание книги Г. Радемахера и О. Теплица включается в серию книг «Библиотека математического кружка», рассчитанную на школьников старших классов и студентов младших курсов и предназначенную для использования в математических кружках; можно только пожалеть, что она не составляет первого выпуска этой серии, созданной под заметным влиянием настоящей книги. При подготовке этого издания мы отказались от каких бы то ни было дополнений и изменений в основном тексте, хотя было бы нетрудно указать ряд тем, «напрашивающихся» в эту книгу, а в отдельных случаях можно было бы предложить некоторые усовершенствования в принятом здесь изложении. Зато заново составлены «Примечания и дополнения» к книге, в которых частично использованы и авторские «Примечания и дополнения». Здесь, в частности, имеются многочисленные ссылки на более позднюю литературу, в первую очередь — на другие книги серии «Библиотека математического кружка». Эти «Примечания» вполне могут быть опущены при первоначальном чтении книги; они, однако, окажутся полезными докладчику, выступающему в математическом кружке с сообщением по этой книге, а также руководителю кружка, или читателю, желающему углубить и дополнить содержащийся здесь материал.

И. М. Яглом

ПРЕДИСЛОВИЕ К ПЕРВОМУ НЕМЕЦКОМУ ИЗДАНИЮ

За своеобразной символикой формул, за интегралами и сигмами, словно за высокой стеной, уединилась математика от окружающего ее мира. То, что происходит за этой стеной, остается обычно тайной для непосвященного, и в стремлении ее разгадать он представляет себе бескровный механизм «мертвых чисел», функционирующий по законам внутренней необходимости. Тому же, кто остается за стеной, она зачастую закрывает горизонт, мешает взглянуть на внешний мир; он увлекается возможностью оценивать математические факты собственными мерками и находит тщеславное удовлетворение в том, что в его владения не проникает профан.

Нельзя ли пробить эту непроницаемую стену и открыть математику «непосвященному» и притом так, чтобы это принесло ему переживание собственного активного участия в математическом познании и творчестве? Может быть такое «активное» восприятие математики доступно лишь узкому кругу «математически одаренных»? Конечно, математическим дарованием в собственном смысле слова, т. е. способностью самостоятельно открывать новые математические истины, наделены лишь немногие. Но ведь, например, и музыкальным дарованием, т. е. уменьем создавать музыкальные произведения достаточно высокой ценности, также обладают лишь немногие. Между тем мы встречаем множество людей, понимающих музыку, способных «музицировать», во всяком случае находящих в музыке радость. Нам кажется, что если бы только удалось преодолеть то недоверие, с которым весьма многие под влиянием случайных школьных впечатлений сторонятся всего, связанного с математикой, то людей, склонных «импровизировать» в области несложных произведений математического искусства, оказалось бы не меньше, чем активных любителей музыки.

Задача, которую мы себе ставим, и заключается как раз в том, чтобы показать, что предубеждение против математики исчезает, если только заложенный в основе математического произведения замысел представить в раскрытом виде, не загроможденном ни формулами, ни сложными вычислениями. Мы стремимся дать здесь представление о многообразии всех тех явлений, которые объединяются понятием «математика», представление о математике как таковой, о *внутренней* ее ценности, которой она обладает сама по себе.

Уже много раз пытались говорить о математике, обращаясь к нематематикам. При этом некоторые, стремясь найти доступ к читателю, старались выдвигать на первый план практическое ее значение: указывали на пользу, которую математика может представить в технических или иного рода применениях, иллюстрируя это удобопонятными примерами. Другие писали книги о математических играх и развлечениях; в них встречается много забавного, но они дают лишь карикатуру на то, что собственно следует называть математикой. Наконец, третьи излагали основания математики в плане общих философских значений; читатель, интересующийся чистой математикой, с особым увлечением занялся бы именно такой оценкой математики — с точки зрения теории познания и общего мировоззрения. Нам, однако, и этот подход представляется по существу лишь внешним подходом к математике, стремлением оценивать математику масштабами, лежащими вне ее самой.

В этой книге мы не имеем возможности осветить то влияние, которое излагаемые здесь идеи оказывают на самую математику, рассмотреть те, если можно так выразиться, внутренние приложения, которые одна область математики находит в другой. Иначе говоря, мы вынуждены здесь отказаться от попытки передать нечто весьма существенное для природы математического здания: от раскрытия поразительных внутренних связей, пронизывающих это здание во всех направлениях. В этом пункте наше самоотречение не добровольно, ибо грандиознейшие открытия были сделаны как раз при прокладке таких внутренних связей, при установлении этих широких взаимозависимостей. Однако для ознакомления с ними от читателя потребовалась бы обширная и углубленная подготовка, более того — основательная тренировка, а это уже не входит в наши намерения. Короче говоря, центр тяжести нашего изложения будет лежать *не в фактах*, в которых обычно раскрывается содержание науки перед неспециалистом; а в *типах математических феноменов* (*явлений*), *в методике постановки вопросов*, *в поисках решения поставленных вопросов*. Понимание больших математических произведений, широкообъемлющих теорий требует, конечно, продолжительной подготовки и значительных усилий. Но то же самое относится и к музыке. Человек, впервые попавший на концерт, едва ли сумеет оценить «Искусство фуги» Баха и охватить строение симфонии. Но ведь наряду с грандиозными музыкальными композициями существуют и так называемые «малые формы», в которых таится подчас обаяние большого искусства и гений которых может быть раскрыт каждому. Такие же «малые формы» хотелось нам извлечь здесь и из обширного мира математики: ряд тем, из которых каждая может быть понята сама по себе, каждая несет свою внутреннюю ценность. В лекционном изложении такая тема укладывается в час, причем слушатель не утруждается связью с другими темами или школьной наукой. Теоремы о равенстве треугольников и правила умножения скобок читатель вспомнит постепенно при чтении.

Эстетическая ценность музыкального произведения заключена не в одной только ведущей линии мелодии: легкая вариация основной темы, неожиданная модуляция может сообщить всему произведению новое заострение, однако уловить и ощутить это будет способен лишь тот, кто предварительно прислушается к основной теме. В этом же смысле и наш читатель должен «прислушаться» к основному мотиву проблемы, которым открывается та или иная тема, к ее постановке, к первым простым примерам, которыми эта тема подкрепляется, прежде чем развернется решительный штурм центральной идеи; он должен следовать за ходом мысли несколько более активно, чем это обычно требуется при чтении. Тогда перед ним исчезнут все препятствия к овладению основным замыслом. Ему будет дано как бы соучаствовать в драме «искания — открытия» некоторых великих мыслителей, которые, покидая иногда обширные здания создаваемых ими общих теорий, возводили тот или иной простой математический факт в небольшое законченное произведение искусства, являвшее в себе фрагментарный прообраз «математического».

Г. Радемахер, О. Теплиц

ПРЕДИСЛОВИЕ КО ВТОРОМУ НЕМЕЦКОМУ ИЗДАНИЮ

Второе издание лишь очень немногим отличается от первого. Тема восьмая («Сечения прямого кругового конуса») заменена менее известной темой из комбинаторики[1]). В шестой теме, с любезного разрешения автора, приведено одно еще неопубликованное доказательство, замечательное своей простотой и близкое и по содержанию, и по методу к предмету этой темы.

Некоторые критики ставили нам на вид, что в нашей книге обижены те или иные отделы математики; при этом одни имели в виду алгебру, другие — геометрию. Мы отступили бы от задач этой книги, намеченных в предисловии к ее первому изданию, если бы приняли во внимание эти возражения. Тема сама по себе нигде не была для нас решающей; все зависело от того, можно ли было изложить ее настолько просто без апелляции к чему-либо, известному читателю заранее, и настолько строго, насколько это попытались сделать мы в нижеследующих двадцати семи этюдах. Чтобы дать читателю то представление о математике, к которому мы здесь стремились, совершенно не важно, из каких специальных дисциплин почерпнуты наши темы. Прием, который книга встретила среди читателей и у большинства критиков, свидетельствует, как нам думается, о том, что наши намерения были поняты в общем правильно.

Г. Радемахер, О. Теплиц

---

[1]) В русском издании оставлена также и тема о сечениях прямого кругового конуса. (*Прим. перев.*)

1. РЯД ПРОСТЫХ ЧИСЕЛ

Число 6 равно произведению двух чисел: 2 и 3. Число 7 нельзя разложить подобным образом на два сомножителя. Поэтому 7 называют *простым* числом. Вообще простым или первоначальным числом называется целое положительное число, которое нельзя разложить на два меньших сомножителя. 5 и 3 тоже простые числа; напротив, число 4 не простое, так как $4=2\cdot 2$. Сама двойка также является простым числом. По отношению к 1 обсуждение вопроса о возможности разложения числа на множители теряет смысл. Таким образом, первые простые числа суть

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ 29,\ 31,\ 37,\ldots$$

С первого же взгляда видно, что ряд несколько причудлив; никакого простого закона в его строении непосредственно не обнаруживается.

Любое число можно разлагать на сомножители до тех пор, пока оно не распадется на одни только простые числа. Представив 6 в виде $2\cdot 3$, мы убеждаемся в этом для 6 непосредственно; $30=5\cdot 6$, но в свою очередь $6=2\cdot 3$, поэтому $30=2\cdot 3\cdot 5$ является произведением трех простых сомножителей; $24=3\cdot 8=3\cdot 2\cdot 4=3\cdot 2\cdot 2\cdot 2$ разлагается на четыре сомножителя, из которых один, именно 2, повторяется несколько раз. Ясно, что при допущении таких повторений подобное разложение на простые множители осуществимо для всякого числа. Поэтому простые числа являются в этом смысле как бы основными элементами, из которых построен весь числовой ряд.

В IX книге «Начал» Евклида ставится вопрос: *имеет ли последовательность простых чисел конец?* И там же дается ответ на этот вопрос: доказывается, что *за каждым простым числом может быть указано еще одно, большее простое число, т. е. что ряд простых чисел бесконечен* [1].

Доказательство Евклида необычайно остроумно. Оно основывается на следующем простом замечании. Таблица, получающаяся при умножении последовательных чисел на 3:

$$3,\ 6,\ 9,\ 12,\ 15,\ 18,\ 21,\ 24,\ \ldots,$$

содержит все числа, в которые входит множителем число 3; ни в какое другое число 3 не входит, в частности, оно не входит ни в одно из тех чисел, которые следуют непосредственно за числами этой последовательности, т. е. за кратными числа 3; таковы, например, $19=6\cdot3+1$, $22=7\cdot3+1$ и т. д.

Подобным же образом число 5 не может быть множителем числа, следующего непосредственно за числами, кратными 5, например числа $21=4\cdot5+1$; то же самое будет верно по отношению к 7, 11 и т. д.

Евклид строит числа

$$\begin{aligned}
2\cdot3+1&=7\\
2\cdot3\cdot5+1&=31\\
2\cdot3\cdot5\cdot7+1&=211\\
2\cdot3\cdot5\cdot7\cdot11+1&=2311\\
2\cdot3\cdot5\cdot7\cdot11\cdot13+1&=30\,031 \text{ и т. д.,}
\end{aligned}$$

которые получаются перемножением нескольких первых простых чисел и прибавлением единицы к полученному произведению.

Легко видеть, что получающиеся таким образом числа не могут содержать в качестве множителей тех простых чисел, с помощью которых они сами были построены. Например, последнее из вышенаписанных чисел не делится на 3, так как оно на 1 больше числа, кратного 3; но оно же на 1 больше числа, кратного 5 или любого другого использованного при его образовании простого числа. Ни одно из чисел 2, 3, 5, 7, 11, 13 не может, следовательно, войти в него множителем. Если бы поэтому в противоположность числам 7, 31, 211, 2311, являющимся простыми, число 30 031 было не простым, то можно быть все же уверенным в том, что ни одно из чисел 2, 3, 5, 7, 11, 13 уже не входит в него множителем и что самый меньший его простой множитель должен быть больше 13. И действительно, путем нескольких проб можно обнаружить, что $30\,031=59\cdot509$; оба эти числа — и 59 и 509 — простые числа, бо́льшие 13.

Это рассуждение будет справедливо и в том случае, если мы продолжим процесс образования таких чисел сколь угодно далеко. Пусть p — какое-либо простое число; образуем из всех простых чисел от 2 до p число

$$2\cdot3\cdot5\cdot7\cdot11\ldots p+1=N;$$

тогда ни одно из использованных здесь простых чисел 2, 3, 5, 7,..., p не войдет множителем в N. В таком случае либо само N будет простым числом, и притом значительно превышающим p, либо N будет разлагаться на простые множители, заведомо отличные от чисел 2, 3, 5,..., p, т. е. бо̀льшие чем p. И в том и в другом случае должны существовать простые числа, бо̀льшие p. За каждым простым числом следует, таким образом, еще одно большее. А это и требовалось доказать [2].

Неизвестно, что должно нас больше всего удивлять в этом тексте Евклида: то ли, что греческие математики вообще могли поставить подобный вопрос ради него самого, из внутреннего влечения к математическому мышлению, т. е. по мотивам, не свойственным ни одному из более древних народов и переданным в наследство позднейшим культурам лишь греческой традицией; или то, что они поставили именно **этот** вопрос, столь легко ускользающий от наивного наблюдателя, кажущийся ему праздным и тривиальным, вопрос, вся трудность и глубина которого раскрывается лишь тому, кто безуспешно пытался отыскать простой закон для ряда простых чисел, закон, позволяющий неограниченно продолжать этот ряд. Пожалуй, самым удивительным здесь нужно считать то, что греки сумели обойти отсутствие подобной закономерности тем искусным приемом доказательства, с которым мы только что познакомились.

Ведь доказательство Евклида дает вовсе не **ближайшее** следующее за p простое число, а вообще лишь некоторое число, лежащее обычно весьма далеко от p, например, в качестве простого числа, заведомо превышающего 11, доказательство дает не 13, а 2311; за 13 оно дает простое число 59 и т. д. И действительно, обычно между тем простым числом, которое дает нам доказательство Евклида, и ближайшим простым числом находится много других простых чисел. В этом именно и заключается свидетельство того большого чувства такта, с которым греческий математик, столь мудро себя ограничивая, пролагал путь в таком сложном по своей структуре ряде простых чисел [3].

Чтобы дать здесь несколько более конкретное представление о сложности этой структуры, мы покажем, что в *ряду простых чисел могут быть сколь угодно большие пробелы*; так, например, мы покажем, что среди тысячи следующих друг за другом чисел может не оказаться ни одного простого числа. Мы исходим при этом из соображений, весьма близких к евклидовым [4].

Выше мы заметили, что число $2\cdot3\cdot5+1=31$ не делится ни на одно из простых чисел 2, 3, 5. Воспользуемся теперь тем простым обстоятельством, что сумма двух чисел, делящихся на 3, также делится на 3 и что подобное свойство остается верным для 5, 7, а также и для всякого другого делителя. Отсюда мы заключаем, что ни одно из чисел

$$2\cdot3\cdot5+2=32;\quad 2\cdot3\cdot5+3=33;\quad 2\cdot3\cdot5+4=34;$$
$$2\cdot3\cdot5+5=35;\quad 2\cdot3\cdot5+6=36$$

не может быть простым; ведь число, прибавляемое к 30, делится во всех этих случаях или на 2, или на 3, или на 5, а так как 30 также делится на них, то делится и сумма. Лишь в отношении следующей за ними суммы $2\cdot3\cdot5+7=37$ так рассуждать уже нельзя, и действительно, число 37 не делится ни на 2, ни на 3, ни на 5, и потому оно само является простым числом.

Обозначив теперь первое простое четырехзначное число (именно 1009) через p и образовав тысячу последовательных чисел

$$2\cdot3\ldots p+2,\quad 2\cdot3\ldots p+3,\ldots,\ 2\cdot3\ldots p+1001,$$

применим к ним только что проиллюстрированный метод рассуждения. Так как каждое из чисел 2, 3,..., 1001 делится по крайней мере на одно из простых чисел 2, 3,..., p, произведение же $2\cdot3\ldots p$ делится на в с е эти числа, то каждая из вышенаписанных сумм делится по крайней мере на одно из этих простых чисел; это значит, что ни одна из них не может быть простым числом. Итак, мы нашли *тысячу последовательных чисел, среди которых нет ни одного простого.*

Конечно, нужно зайти довольно далеко в ряду простых чисел, прежде чем встретится такого рода пробел. Но если пойти достаточно далеко, то можно, следуя тому же принципу, найти пробел, охватывающий миллион последовательных чисел, и вообще пробелы сколь угодно большой величины.

Эта вторая проблема, сколь ни близка она к первой как по характеру постановки вопроса, так и по методу доказательства, не встречается ни у кого из греческих математиков. Ее выдвинула новая математика, связав с нею целый круг дальнейших вопросов, доказываемых в большинстве случаев уже далеко не столь просто; из этих вопросов развилась одна из глубочайших, одна из самых волнующих своими нерешенными проблемами областей математического анализа.

Мы остановимся здесь на одном небольшом примере, который поддается исследованию методом Евклида и позволяет составить некоторое представление о том, в каком именно направлении современная математика расширяет проблематику греков. Выше мы рассматривали сначала числа 3, 6, 9,..., получающиеся от умножения последовательности чисел натурального ряда на 3, а затем последовательность следующих за ними чисел 4, 7, 10,...; теперь рассмотрим еще оставшиеся после этого числа

$$2,\ 5,\ 8,\ 11,\ 14,\ 17,\ 20,\ 23,\ldots,$$

т. е. *числа, которые при делении на* 3 *дают в остатке* 2. Покажем, что и *среди одних только этих чисел содержится уже бесчисленное множество простых чисел*, т. е. покажем, что последовательность простых чисел

$$2,\ 5,\ 11,\ 17,\ 23,\ldots$$

также не имеет конца.

Доказательство нуждается в простом предварительном замечании. Если перемножить между собой два каких-либо числа последовательности

$$1,\ 4,\ 7,\ 10,\ 13,\ldots,$$

то получается число, принадлежащее к этой же последовательности. Действительно, все эти числа при делении на 3 дают в остатке единицу и имеют вид $3x+1$, где x целое; если перемножить два таких числа, например $3x+1$ и $3y+1$, то получим:

$$(3x+1)(3y+1)=9xy+3y+3x+1=\\=3(3xy+y+x)+1,$$

т. е. опять число, на единицу большее кратного трем.

Из этого простого замечания следует, что среди простых множителей любого из чисел последовательности 2, 5, 8, 11,... всегда найдется по крайней мере один, принадлежащий к этой же последовательности. Так, например, для $14 = 2 \cdot 7$ таковым является множитель 2, а для $35 = 5 \cdot 7$ — множитель 5. Действительно, ни один из таких простых множителей не может принадлежать к ряду 3, 6, 9,..., в котором имеется лишь одно-единственное простое число 3. С другой стороны, если бы рассматриваемое число состояло из одних только простых множителей, принадлежащих к последовательности 4, 7, 10, 13,..., то, согласно предыдущему замечанию, оно само должно было бы принадлежать к этой же последовательности. Таким образом, чтобы число принадлежало к последовательности 2, 5, 8, 11,..., оно должно содержать по меньшей мере один простой множитель, принадлежащий к этой же самой последовательности.

Мы сможем теперь легко доказать высказанное выше утверждение. Для этого придется только несколько видоизменить доказательство Евклида, а именно вместо выражения

$$2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \ldots p + 1 = N$$

рассматривать выражение

$$2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \ldots p - 1 = M,$$

которое, будучи на единицу меньше кратного 3, принадлежит к последовательности 2, 5, 8, 11,... Число M, так же как и N, не делится ни на одно из простых чисел 2, 3, 5, 7, 11,..., p. Является ли само M простым числом или оно разлагается на несколько простых множителей — в обоих случаях эти простые числа будут больше p. Доказанное выше положение утверждает, что среди этих множителей должен быть по крайней мере один, принадлежащий к последовательности 2, 5, 8, 11,... Таким образом, в этой последовательности заведомо имеются простые числа, превышающие сколь угодно большое простое число p и расположенные, следовательно, сколь угодно далеко [5].

Этим, однако, еще не предрешается вопрос о том, содержит ли также и последовательность 1, 4, 7, 10, 13, ... бесконечно много простых чисел; нет ничего невозможного в предположении, что из всей совокупности простых чисел бесконечно много их приходится на последовательность

2, 5, 8, ... и только конечное число — на последовательность 1, 4, 7, ...; общее их количество при этом по-прежнему остается бесконечно большим. Доказательство того, что и вторая последовательность содержит также бесконечно много простых чисел, требует уже совершенно других методов, представление о которых мы попытаемся дать в одном из следующих этюдов [6]. Нам хотелось лишь показать здесь, что современная математика не только в отдельных частных случаях, но и в проблемах кардинальной важности — как в постановке вопросов, так и в методах их решения — примыкает к греческой. Это должно быть особо подчеркнуто после того, как О. Шпенглер выдвинул свой эффектный, возбудивший большой шум тезис о том, что греческая и современная математика не имеют будто бы между собой ничего общего, что они представляют собой две различные по существу математики [7]. Во всяком случае с помощью столь схематичной формулы невозможно отобразить всю сложность соотношений, имеющих место в действительности.

2. МАРШРУТЫ В СЕТИ КРИВЫХ

Управление городских железных дорог намерено по-новому перегруппировать маршруты, которыми оно обслуживает свою трамвайную сеть. Оно предполагает распределить эти маршруты таким образом, чтобы каждая линия обслуживалась впредь лишь одним- единственным маршрутом; при этом пассажир получает право с одним и тем же билетом менять маршруты и делать столько пересадок, сколько ему нужно, чтобы достигнуть места своего назначения. Задача заключается в том, чтобы определить *наименьшее число маршрутов, необходимое для полного проведения в жизнь этого принципа.*

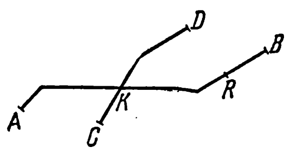


Рис. 1.

Для маленького города, где рельсовая сеть может быть схематически представлена рис. 1, вопрос решается крайне просто. Мы должны либо установить один маршрут по линии AB и другой по линии CD, либо объединить в один маршрут AK и KD, а по линии BKC назначить второй маршрут, либо, наконец, провести один маршрут из A через K в C, а другой — из B через K в D.

Очевидно, что этим исчерпываются все возможности и что при любой из этих комбинаций необходимы два маршрута. Одним-единственным маршрутом обойтись нельзя, как бы мы ни комбинировали наши четыре линии AK, BK, CK и DK. При этом мы исключаем заведомо не отвечающие задаче решения; так, например, нельзя маршрут, ведущий из K, закончить в какой-нибудь точке R, а от R до B назначить другой маршрут, т. е. устроить в R пересадочный пункт, как это иногда бывает необходимо делать в связи с ремонтными работами. При таких условиях нам потребовалось бы, конечно, больше двух маршрутов, и можно было бы даже сколько угодно увеличивать их число подобным образом. Однако задача, как уже было сказано, состоит в том, чтобы обойтись наименьшим числом маршрутов.

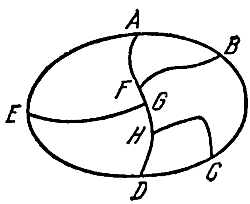


Рис. 2.

В несколько более трудных условиях находится управление, ведающее городской трамвайной сетью, схематически изображенной на рис. 2; впрочем, и эта схема все еще довольно проста. Управление может установить прежде всего окружной или кольцевой маршрут от A через B, C, D, E опять к A, далее, второй — диаметральный маршрут от A через F, G, H к D, и, наконец, к этим двум оно должно, очевидно, присоединить еще три маршрута: EG, BF, CH, и всего, таким образом, ввести пять маршрутов. Впрочем, один из них можно сэкономить: например, маршруты первый и второй можно объединить в один, заставив курсировать вагоны сначала от A через B, C, D, E опять к A и затем далее, от A через F, G, H к D. *Но, может быть, можно обойтись и тремя маршрутами?*

Или рассмотрим еще один пример. По схеме рис. 3 мы можем назначить один маршрут от A через B, C, D, E опять к A и оттуда через F к B. Тогда для обслуживания всей сети нам потребуются еще три маршрута: FC, FD и FE. Но и здесь это число можно сократить, назначив второй маршрут от C через F до D и введя «челнок» на участке EF. В результате мы обходимся здесь тремя маршрутами, причем для двух из них точка F является промежуточной станцией, а для третьего эта точка служит конечным пунктом. *Но нельзя ли обойтись только двумя маршрутами?*

В двух последних случаях (впрочем, все еще довольно простых) мы уже сталкиваемся со сравнительно большим числом возможностей, которые нам нужно разобрать, чтобы с такой же ясностью и полнотой, как для схемы рис. 1, ответить на вопрос, можно ли обойтись меньшим числом маршрутов или нельзя?

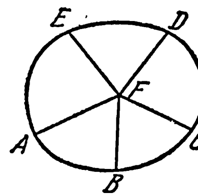


Рис. 3.

И вот уже сквозь внешнее оформление задачи, столь далекой от обычных условий и масштабов трамвайной сети больших городов, намечаются контуры математической проблемы, правда, очень несложной. В схеме рис. 1 мы просто испробовали все возможности и получили искомый ответ (именно два маршрута). Однако такой подход к задаче мы не можем назвать математическим. И если мы прилежно исчерпаем все возможности схемы рис. 2, пока не удостоверимся, что для нее нельзя обойтись меньше, чем четырьмя маршрутами, то и эта работа имеет столь же мало общего с математикой, как, например, терпеливое, пусть даже безошибочное, перемножение семизначных чисел. Математика, или по крайней мере легкое веяние математики, появляется только тогда, когда мы находим решение задачи не путем пересмотра всех случаев и не только для одной частной схемы рис. 2, а путем логических умозаключений сразу для произвольно сложных сетей и кривых.

Идея, к которой сводится решение нашей задачи, чрезвычайно проста. Обратим внимание на то, где должны быть расположены конечные пункты отдельных маршрутов. Прежде всего, совершенно ясно, что они должны быть там, где находится конец одного какого-нибудь отрезка нашей сети, например — для схемы рис. 1 — в точках A, B, C, D. Так как в этой схеме имеется всего четыре конечные точки и так как каждый маршрут имеет самое большее два конечных пункта (именно, когда он некольцевой), то ясно, что между этими четырьмя конечными пунктами должно быть назначено *по меньшей мере два* маршрута. Таким образом, решение, полученное нами выше в результате проб, достигнуто теперь посредством общего рассуждения.

Но в схеме рис. 2 вовсе нет свободных концов. Зато здесь имеются такие точки, как, например, A, в которых сходятся три направления. Если каждый участок должен обслуживаться одним только маршрутом, то очевидно, что в такой точке должен быть по крайней мере один конечный пункт некоторого маршрута. Поскольку в схеме рис. 2 таких точек, как это легко сосчитать, имеется восемь, то здесь должно быть по меньшей мере восемь конечных пунктов маршрутов. Восемь — опять число четное — число концов четырех маршрутов. Таким образом, *для схемы рис. 2 должно быть установлено, самое меньшее, четыре маршрута*; трех маршрутов здесь будет недостаточно. В схеме рис. 3 мы встречаемся с подобным же положением: здесь имеется пять точек пересечения того же типа, что и выше рассмотренные, и, кроме того, шестая точка — F, в которой сходятся не три, а пять направлений и которую мы назовем «точкой пересечения 5-го порядка». Две пары из этих пяти направлений могут принадлежать, как это выяснено выше, каждая одному маршруту, и одно остается лишним, так что точка F обязательно должна быть конечным пунктом. Сказанное имеет силу для любой точки пересечения любого нечетного порядка. Полученное число конечных пунктов, 6, опять четное и требует *по меньшей мере трех маршрутов*, двух будет недостаточно. Теперь ясно, каким образом для любой, сколь угодно сложной сети кривых можно непосредственно вычислить наименьшее возможное число необходимых маршрутов: нужно перебрать все точки пересечения и сосчитать, сколько среди них имеется точек нечетного порядка; тогда наименьшее число необходимых маршрутов будет равно половине числа таких точек.

Продвинувшись так далеко по пути овладения нашей проблемой, мы намерены теперь довести ее решение до конца; мы покажем, что *число точек пересечения нечетного порядка всегда четно и что всегда можно обойтись числом маршрутов, равным половине числа точек нечетного порядка.*

Ясно, что вообще всякую, сколь угодно сложную систему кривых можно обслужить таким количеством маршрутов, что ни один отрезок не будет обслуживаться одновременно двумя маршрутами, — надо только на каждом отдельном участке между парой соседних точек пересечения ввести отдельный маршрут («челнок»), для чего потребуется столько же маршрутов, сколько имеется таких участков.

Но в то же время ясно, что, вообще говоря, мы вводим при этом много лишних маршрутов. Нас же интересует наименьшее возможное число маршрутов. Ясно также, что среди всех возможных систем маршрутов заведомо должны существовать так называемые «оптимальные системы», т. е. системы с наименьшим числом маршрутов, подобно тому, как среди учеников школы непременно должны быть самые младшие, хотя какие именно это ученики и нельзя в точности установить до проверки метрик.

Далее ясно, что система маршрутов ни в коем случае не может быть оптимальной, если в ней имеются излишние конечные пункты, подобные, например, пункту R в схеме рис. 1. Система не может быть оптимальной и тогда, когда в ней не произведены возможные соединения маршрутов и имеются такие точки, в которых смыкаются своими конечными пунктами два различных маршрута; в этом случае можно уменьшить число маршрутов путем попарного их соединения (как, например, в схеме рис. 3 объединялись маршруты CF и DF). Оптимальная система маршрутов не допускает подобных сокращений. Все такие пары маршрутов в ней уже соединены, и точка пересечения нечетного порядка может быть концом не более чем одного маршрута. Точка же пересечения четного порядка вообще не может быть конечным пунктом.

Остается еще выяснить, могут ли входить в состав оптимальных систем замкнутые кольцевые маршруты, вовсе не имеющие конечных пунктов. Подобный кольцевой маршрут нам встретился при рассмотрении схемы рис. 2. Мы сумели там устранить это кольцо, присоединив к нему в пункте A маршрут $DHGFA$ и уменьшив тем самым общее число маршрутов с пяти до четырех. Этот прием можно применить во всех тех случаях, когда на кольцевом маршруте имеется точка пересечения нечетного порядка. Таким образом, в *оптимальной* системе (с точкой пересечения нечетного порядка) такого кольцевого маршрута быть не может, ибо при наличии его число маршрутов допускало бы сокращение.

Если все лежащие на кольце точки пересечения — четного порядка, то можно поступить подобным же образом. Пусть A будет такой точкой пересечения четного порядка (рис. 4) на кольцевом маршруте (имеющем форму восьмерки); дальнейшее направление исходящих из A (пунктирных) ветвей на рисунке не показано и может быть совершенно произвольным. Как уже было выяснено, точка A в оптимальной системе заведомо не является конечным пунктом для пунктирных ветвей, сходящихся в A. Маршрут, ведущий, например, из B в A, находит в этой точке свое продолжение, скажем, в ветви AE. При этих условиях мы можем остановить движение по этому маршруту BAE в точке A, заставить вагон обойти кольцо и только уже после этого позволить ему продолжать путь в направлении AE. Кольцо было бы таким образом включено в маршрут BAE, общее же число маршрутов благодаря этому уменьшилось бы на один, что для оптимальной системы невозможно.

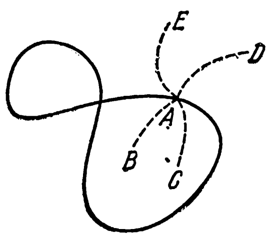


Рис. 4.

Итак, мы пришли к выводу, что в оптимальной системе конечные пункты маршрутов могут быть только в точках пересечения нечетного порядка, и именно по одному конечному пункту на каждую точку пересечения. Мы выяснили далее, что в оптимальной системе кольцевых маршрутов вообще не может быть, за исключением того случая, когда вся сеть и все имеющиеся в сети точки пересечения четного порядка охватываются одним-единственным кольцевым маршрутом. Таким образом, число точек пересечения нечетного порядка равно числу конечных пунктов маршрутов в оптимальной системе. Из этого следует, что число точек пересечения нечетного порядка всегда четно и что можно установить маршруты так, что их число будет равно половине числа точек пересечения нечетного порядка. Если же точки пересечения все четного порядка, то вопрос решается единственным, не имеющим концов, кольцевым маршрутом [8].

3. НЕСКОЛЬКО ЗАДАЧ НА МАКСИМУМ

1. Сравним между собой несколько различных прямоугольников, имеющих одинаковый периметр, равный, скажем, 4 *см* (рис. 5). Продолговатые низкие прямоугольники, ширина которых близка к 2 *см*, имеют незначительную площадь, тем меньшую, чем меньше их высота; точно так же площадь узких высоких прямоугольников тем меньше, чем эти прямоугольники уже. Сравнительно бо́льшую площадь имеют прямоугольники с некоторыми промежуточными пропорциями. Встает вопрос: *какой же именно из всех прямоугольников с периметром 4 см имеет наибольшую площадь?*

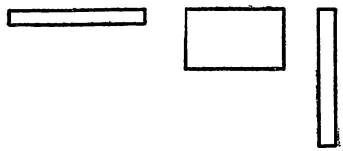


Рис. 5.

Такова типичная схема задачи на максимум. Приведенная задача, вероятно, самая простая и самая древняя из всех задач подобного рода. Именно поэтому на ней лучше, чем на какой-либо другой, можно разъяснить сущность задач на максимум, что мы и сделаем, прежде чем перейти к разбору того вопроса, которому, собственно, посвящена настоящая тема.

В VI книге Евклида [9] эта задача решается следующим приемом, принцип которого мы сохраним полностью, видоизменив несколько лишь форму изложения. Возьмем произвольный прямоугольник $ABCD$ с заданным периметром U и построим, как это показано на рис. 6, квадрат $BEFG$ со стороной $\frac{1}{4}U$ и, следовательно, с тем же самым периметром U; мы утверждаем, что *этот квадрат как раз и представляет собой решение нашей задачи,* т. е. что его площадь больше площади прямоугольника.

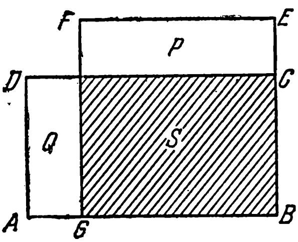


Рис. 6.

При том расположении квадрата и прямоугольника, которое мы видим на рисунке, у них есть общий (заштрихованный) прямоугольник S. Квадрат состоит из этого заштрихованного прямоугольника S и части P; данный прямоугольник $ABCD$ состоит из заштрихованного прямоугольника S и части Q. Но так как полупериметр квадрата $GB+BE$ равен полупериметру прямоугольника $AB+BC$, то $AG+GB+BC=GB+BC+CE$, откуда $AG=CE$, т. е. высота прямоугольника P равна ширине прямоугольника Q. Что же касается другого измерения прямоугольника P, то им служит сторона квадрата, между тем как второе измерение Q есть лишь часть стороны квадрата и, следовательно, короче ее. Но из двух прямоугольников, имеющих по одной равной стороне (рис. 7), тот больше, у которого больше вторая сторона. Таким образом, площадь P больше площади Q и, следовательно, $P+S$ больше $Q+S$, т. е. у квадрата площадь больше, чем у прямоугольника. Сторона прямоугольника Q будет тогда равна стороне квадрата, когда за основной прямоугольник взят квадрат; в этом и только в этом случае мы имеем равенство рассматриваемых

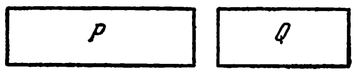


Рис. 7.

площадей. Таким образом, *квадрат* действительно *по площади больше всех других прямоугольников равного с ним периметра* [10].

Если x и y — длины сторон прямоугольника в сантиметрах, то полученный результат показывает, что xy — число, измеряющее в квадратных сантиметрах площадь прямоугольника, — меньше площади квадрата с периметром $x+y+x+y=$ $=2(x+y)$. Сторона такого квадрата равна четвертой части этой величины, т. е. $\frac{1}{2}(x+y)$, а площадь его получится, если $\frac{1}{2}(x+y)$ возвести в квадрат. Переходя с излюбленной греками геометрической формы изложения на язык формул современной математики [1]), получаем, что:

Два любых положительных числа всегда удовлетворяют соотношению

$$xy \leqslant \left(\frac{x+y}{2}\right)^2, \tag{1}$$

или

$$\sqrt{xy} \leqslant \frac{x+y}{2}, \tag{2}$$

[1]) Знак $<$ здесь, как это принято в математике, имеет смысл «меньше чем» (например, $3<5$); знак $\leqslant$ означает «меньше или равно» или, что то же, «не больше».

т. е. *среднее геометрическое двух чисел меньше или равно их среднему арифметическому;* равенство возможно лишь в том случае, когда равны сами эти числа ($x=y$) [11].

2. Теперь ясно, что такое задача на максимум и что мы понимаем под решением ее. Решить такую задачу — это значит указать ее ответ и доказать, что этот ответ при сравнении со всеми другими возможными превосходит их в отношении исследуемого свойства (в данном случае — величины площади).

Обратимся к основной теме этой главы, а именно к задаче *об отыскании среди всех вписанных в данный круг треугольников такого, который имеет наибольшую площадь.* Одно место в «Меноне» Платона позволяет предполагать, что эта задача была, по-видимому, поставлена, если и не разрешена, еще во времена Платона, т. е. за сто лет до появления «Начал» Евклида. Впрочем, нижеследующее доказательство, которое, если судить по его стилю, вполне могло бы быть известно древним, не встречается ни у Евклида, ни в современной ему литературе.

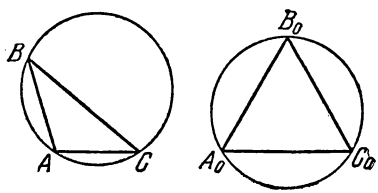


Рис. 8.

Наряду с каким-либо вписанным в наш круг треугольником ABC рассмотрим равносторонний треугольник $A_0B_0C_0$, вписанный в тот же самый или в равный ему круг (рис. 8). Площадь этого треугольника — вполне определенная величина и не зависит от его положения внутри круга. Мы утверждаем, что *равносторонний треугольник имеет бо́льшую площадь, чем всякий другой треугольник, вписанный в такой же круг,* т. е. что этот треугольник представляет собой решение нашей задачи.

Для доказательства будем исходить из того, что вершины равностороннего треугольника делят окружность на три равные дуги; для всякого же другого треугольника получаем три какие-либо, вообще говоря, неравные дуги, также дающие в сумме полную окружность. Заметим, что по крайней мере одна из этих трех дуг должна быть меньше трети полной окружности и хотя бы одна — больше этой трети. Ибо если бы не оказалось ни одной дуги, меньшей трети окружности, то все эти три дуги вместе дали бы сумму, бо́льшую целой окружности, за исключением случая, когда каждая из них была бы в точности равна трети окружности, т. е. случая равностороннего треугольника, который мы предполагаем здесь исключенным. Точно так же мы заключаем, что одна из трех дуг должна превышать треть окружности.

Вопрос о том, больше или меньше трети окружности третья дуга, остается при этом открытым; ничего определенного по этому поводу высказать нельзя, да в этом и нет необходимости.

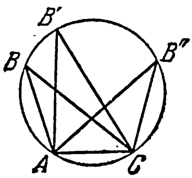


Рис. 9.

Пусть вершины заданного нам произвольного треугольника обозначены таким образом, что дуга AB составляет меньше трети, а дуга BC — больше трети полной окружности. Отложим (рис. 9) на дуге CB, начиная от точки C, дугу CB'', равную AB, так что треугольник CAB'' будет зеркальным отображением треугольника ACB относительно диаметра круга, перпендикулярного AC. Кроме того, от точки A по направлению к B отложим дугу AB', равную трети окружности. Точка B' будет лежать от A во всяком случае дальше, чем B, так как по условию дуга AB меньше трети окружности. С другой стороны, точка B' должна лежать перед B'', т. е. между точками B и B''. В самом деле, если бы она лежала за B'', то дуга AB' была бы больше дуги AB''; но последняя равна своему зеркальному отображению CB, а поскольку, по предположению, CB больше трети полной окружности, то дуга AB'', а следовательно и дуга AB', должна была быть больше трети окружности, между тем как дуга AB' по построению как раз равна трети окружности. Так как точка B' лежит между B и B'', то вершина треугольника ACB' расположена **выше** вершины треугольника ACB, имеющего то же самое основание AC. В силу известной теоремы о площади треугольника (равной половине произведения основания на высоту) площадь треугольника ACB' будет **больше** площади треугольника ACB.

Таким образом, мы нашли новый треугольник ACB', вписанный в тот же самый круг и имеющий площадь бо́льшую, чем первоначально данный нам треугольник; одна сторона этого нового треугольника, именно AB', равна стороне вписанного в тот же круг равностороннего треугольника.

Возможно, что построенный нами треугольник ACB' окажется уже равносторонним. Это произойдет в том случае, если дуга AC, стягиваемая одноименной стороной данного треугольника, в точности равна трети окружности. В этом случае наше доказательство того, что равносторонний треугольник по площади больше данного неравностороннего треугольника, было бы уже закончено. В противном случае примем за основание всей нашей фигуры вместо AC, выступавшего в этой роли до сих пор, сторону AB' (для этого нужно только повернуть всю фигуру против часовой стрелки, пока AB' не примет горизонтального положения, как на рис. 10) и применим по отношению к треугольнику $AB'C$ с основанием AB' те же самые рассуждения, какие мы применили только что по отношению к треугольнику ACB с основанием AC. Для этого от точки B' по направлению к C отложим по окружности дугу $B'C'$, равную трети полной окружности, так что точка C' вместе с A и B образует равносторонний вписанный в круг треугольник, равный треугольнику $A_0B_0C_0$ (рис. 8). По аналогии с предыдущим мы убеждаемся в том, что этот равносторонний треугольник имеет площадь еще бо́льшую, чем треугольник $AB'C$, а следовательно, и бо́льшую, чем первоначально данный треугольник, если только последний сам не равносторонний.

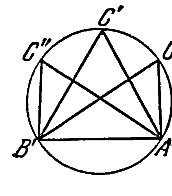


Рис. 10.

Итак, мы приходим к выводу, что *равносторонний треугольник* действительно *имеет бо́льшую площадь, чем всякий другой треугольник, вписанный в тот же самый круг* [12].

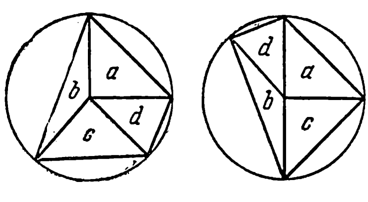


Рис. 11.

3. С помощью подобных же рассуждений можно показать, что *из всех n-угольников, вписанных в один и тот же круг, правильный n-угольник имеет наибо́льшую площадь.*

В дальнейшем нам понадобится следующее чрезвычайно простое предварительное замечание: если дан какой-либо вписанный в круг n-угольник (рис. 11), то в тот же самый круг можно вписать новый n-угольник, стороны которого равны по величине сторонам первого n-угольника, но расположены в иной, произвольно выбранной последовательности [1]). Для этого стóит только разбить круг радиусами, проходящими через вершины данного n-угольника, на n секторов, представить себе, что эти секторы вырезаны, положим, из картонного диска, и тогда легко будет непосредственно убедиться в том, что из этих секторов при расположении их в каком-либо ином, новом порядке можно опять составить тот же самый круговой диск. При этом получается и новый n-угольник. Очевидно, что площадь его от такой перестановки секторов не меняется.

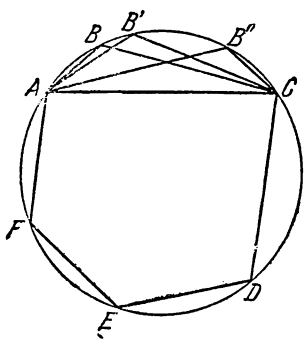


Рис. 12.

Приняв во внимание это замечание, мы можем приступить к доказательству нашего утверждения. Как и в случае треугольника, начнем с утверждения, что одна какая-либо сторона вписанного в круг неправильного n-угольника должна стягивать дугу, меньшую $\frac{1}{n}$ доли полной окружности, а какая-нибудь другая — бóльшую. Однако здесь уже нельзя утверждать, что обе эти стороны расположены рядом. В треугольнике это неизбежно, так как каждая из трех его сторон примыкает к обеим другим, но n-угольник с более чем тремя сторонами этим свойством уже не обладает. Тем не менее, на основании сделанного замечания, если рассматриваемые две стороны не смежны, в тот же самый круг можно вписать новый n-угольник той же самой площади, в котором эти стороны окажутся соседними.

Обозначим (рис. 12) меньшую из этих сторон через AB, бóльшую — через BC. Тогда от точки A по направлению к B можно отложить дугу AB', равную в точности $\frac{1}{n}$ доле целой окружности, и, так же как и прежде, убедиться в том, что

[1]) Этим замечанием мы обязаны устному сообщению Э. Штейница (E. Steinitz).

точка B' должна расположиться между B и ее зеркальным отображением B''. Если поэтому вершину B в данном нам n-угольнике заменить через B', сохранив все другие вершины без изменения, то площадь n-угольника при этом увеличится, а одна сторона станет равной стороне правильного вписанного n-угольника. Точно так же мы поступим последовательно и со всеми остальными $n-1$ сторонами нашего n-угольника, пользуясь в случае надобности нашим предварительным замечанием. В результате мы убедимся, что данный n-угольник меньше по площади того правильного n-угольника, к которому мы должны будем прийти с помощью наших построений, при которых все стороны данного n-угольника, одна за другой, делаются последовательно равными сторонам правильного n-угольника.

Точно таким же способом можно показать, что *из всех n-угольников, описанных около данного круга, правильный имеет наименьшую площадь* [13].

4. НЕСОИЗМЕРИМЫЕ ОТРЕЗКИ И ИРРАЦИОНАЛЬНЫЕ ЧИСЛА

Измерение длин, площадей и объемов является, без сомнения, началом всякой геометрии. Весьма просто измерить отрезок посредством некоторого другого отрезка в том случае, когда этот последний укладывается в первом целое число раз (рис. 13). Если же в результате откладывания меньшего отрезка на большем получается остаток (рис. 14), то пробуют

Рис. 13.

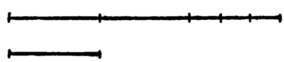


Рис. 14.

сначала, не составляет ли этот остаток $\frac{1}{2}$, $\frac{1}{3}$, или $\frac{2}{3}$, или вообще какую-либо дробную часть меньшего отрезка, служащего здесь мерой длины. Если такую дробь найти удается, то мы получаем, что хотя меньший отрезок и не укладывается целое число раз в бо́льшем, но зато в нем укладывается целое число раз некоторая доля меньшего отрезка. Эта доля, укладывающаяся целое число раз в обоих данных отрезках, и представляет собой их «общую меру».

Самые ранние геометрические исследования, без сомнения, рассматривали такие «общие меры». Если, например, начертить прямоугольник, у которого одна сторона равна 3 *см*, а другая — 4 *см*, то, согласно теореме Пифагора, квадрат диагонали такого прямоугольника равен

$$3^2+4^2=9+16=25 \text{ см}^2,$$

и так как, следовательно, сама диагональ равна 5 *см*, то (рис. 15) меньшая сторона прямоугольника и его диагональ имеют общей мерой отрезок в 1 *см* и отношение их длин равно отношению 3:5.

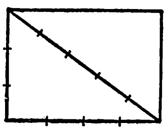


Рис. 15.

Естественно вместо этого прямоугольника взять квадрат и задаться вопросом об отыскании общей меры его стороны и диагонали. При первых попытках откладывания, как показывает рис. 16, такой общей меры не обнаруживается; мы вынуждены прибегать к помощи все более и более мелких дробных долей. Перед нами встает вопрос, найдется ли вообще когда-нибудь при достаточно мелком дроблении отрезков общая мера, или такой общей меры не существует вовсе, и отрезки, значит, должны быть признаны «н е с о и з м е р и м ы м и».

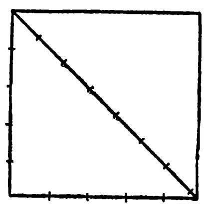


Рис. 16.

В связи с этим возникает и общая проблема: можно ли делить отрезок безгранично на все более и более мелкие части, или этот процесс деления упирается в некоторую границу, когда отрезок распадается на очень большое, но все же конечное число весьма малых «н е д е л и м ы х» — неразложимых далее частиц, обнаруживая, таким образом, «атомистическое строение»? Мы знаем, что поколение греческих ученых, предшествовавшее П л а т о н у, в особенности Д е м о к р и т, учило, что материя построена атомистически. Это, однако, совсем не то же самое, что атомистическое строение отрезка. Можно очень хорошо представить себе безгранично делимый отрезок, на котором материя распределена атомистически. У нас имеется относящийся к тому же времени, т. е. к эпохе около 450 г. до н. э., фрагмент афинского философа Анаксагора, в котором утверждается, что отрезок делим неограниченно. Если подобный обрывок из утерянного учебника или лекции вообще дошел до нас, то произошло это, надо думать, потому, что мы имеем здесь дело не со случайным высказыванием, а с тезисом, который был знаменит в свое время, так как почитался оригинальным и возбуждал большие споры; в те времена человечество, по-видимому, впервые столкнулось с великой проблемой непрерывности.

Отсюда мы можем судить о том громадном впечатлении, которое должно было произвести более глубокое открытие, а именно, *что сторона и диагональ квадрата несоизмеримы.* Дошедшие до нас сообщения приписывают это открытие пифагорейцам—южноиталийскому тайному союзу, о котором, впрочем, мы знаем очень мало достоверного. Легенда говорит, что пифагореец, поведавший об этом открытии людям, поплатился за это гибелью при кораблекрушении. Зато совершенно достоверным является свидетельство Платона: в своих «Законах» он рассказывает, какое сильное впечатление произвело на него это открытие, когда он уже в зрелые годы познакомился с ним [14].

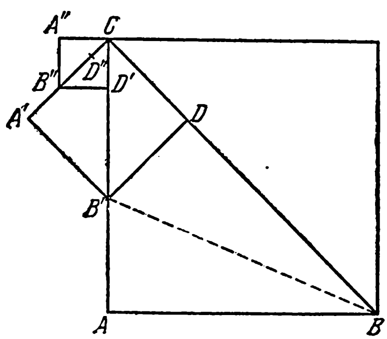


Рис. 17.

Ниже мы приводим два различных доказательства этого открытия, оставляя в стороне очень интересный сам по себе исторический вопрос о том, какое из этих доказательств нужно считать более древним. Второе приводится не только Евклидом, но уже и Аристотелем; первое во всяком случае выражено вполне в духе греческих методов; оно не выходит из круга идей X книги Евклида и является, по-видимому, более ранним [15].

П е р в о м у д о к а з а т е л ь с т в у мы предпошлем одно геометрическое соображение элементарного характера, возникшее в результате тщетных попыток найти общую меру между стороной и диагональю квадрата по способу повторного откладывания. Отложим (рис. 17) на диагонали квадрата от точки B его сторону, что удается сделать один-единственный раз. Из полученной точки D восставим перпендикуляр к BD, который пересечет сторону квадрата AC в точке B', и соединим B' с B. Тогда

$$BA = BD, \quad BB' = BB'$$

и

$$\angle BAB' = \angle BDB',$$

как прямые углы. Отсюда, как известно, следует, что треугольники BAB' и BDB' равны и, значит, $AB' = DB'$, как соответственные стороны в равных треугольниках. Далее, угол $B'CB$, как угол между стороной и диагональю квадрата, равен половине прямого, и, так как угол CDB' по построению прямой, для третьего угла треугольника CDB' остается половина прямого. Этот треугольник является, таким образом, равнобедренным прямоугольным треугольником, и отсюда, в частности, следует, что $DB' = DC$.

Таким образом, мы пока доказали, что

$$AB' = B'D = DC. \tag{1}$$

Восставим теперь перпендикуляр к диагонали CB в точке C, затем отложим на этом перпендикуляре отрезок CA', равный DB', и соединим его конечную точку A' с точкой B'. Тогда фигура $A'B'DC$ будет квадратом, меньшим первоначального; диагональ $B'C$ в нем уже проведена. С этим квадратом поступим так же, как с первоначальным, т. е. отложим его сторону на его диагонали и из полученной точки D' восставим перпендикуляр к диагонали нового квадрата. Этот перпендикуляр пересечет его сторону в точке B''; аналогично предыдущему будем иметь:

$$A'B'' = B''D' = D'C. \tag{2}$$

Очевидно, что этот процесс можно продолжать неограниченно. При этом всякий раз будет получаться новый остаток, меньший предыдущего:

$$CD > CD' > CD'' > CD''' > \ldots, \tag{3}$$

и каждый такой остаток будет представлять собой разность между диагональю и стороной одного из квадратов, последовательно получаемых в процессе построения:

$$CD = CB - AB, \quad CD' = CB' - A'B',$$
$$CD'' = CB'' - A''B'', \ldots \tag{4}$$

Это построение позволяет теперь приступить к самому доказательству, которое мы проведем методом от противного. Предположим, что сторона и диагональ нашего квадрата соизмеримы; пусть, следовательно, существует общая мера этих отрезков, т. е. некоторый отрезок E, являющийся точной дробной частью стороны квадрата, вместе с тем укладывается целое число раз также и в диагонали. Следует принять во внимание, что разность между какими-либо двумя отрезками, являющимися целыми кратными E, точно так же является целым кратным E (рис. 18). Если, таким образом, CB и AB являются целыми кратными E, то, согласно (4), отрезок CD, а вместе с ним и отрезок $A'B'$, являющийся другой стороной того же квадрата, также должны быть целыми кратными E. Но в таком случае на основании (1) и диагональ этого квадрата $CB' = CA - B'A = AB - CD$, как разность между двумя отрезками, кратными E, будет кратна E. После того как это свойство доказано для стороны и диагонали квадрата, обозначенного буквами с одним штрихом, оно по аналогии распространяется и на все последующие квадраты.

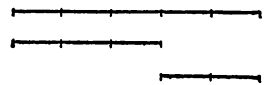


Рис. 18.

Теперь наше доказательство от противного может быть закончено, т. е. приведено к желаемому противоречию, ибо если предположить, что существует общая мера стороны и диагонали данного квадрата, то, с одной стороны, все фигурирующие в (3) отрезки оказываются кратными E, а с другой стороны, соотношения (3) утверждают, что эти отрезки, кратные E, становятся все меньше, причем процесс уменьшения не может ни прекратиться, ни привести к нулю. Но для кратных некоторого определенного отрезка E это невозможно. Действительно, пусть первый стоящий в соотношении (3) отрезок CD в 1000 раз больше E; в таком случае CD' должен превышать E в меньшее, но опять-таки в целое, число раз, т. е. самое большее в 999 раз, и т. д.; следовательно, 1000-й член этой цепи, оставаясь кратным E, должен быть в то же время меньше E, иначе говоря, вопреки доказанному, он должен обратиться в нуль. Таким образом, предположение о соизмеримости стороны и диагонали квадрата приводит нас к противоречию.

В т о р о е д о к а з а т е л ь с т в о много проще; оно требует лишь небольшого арифметического введения, более короткого, чем те длинные рассуждения из элементарной геометрии, которые необходимо было предпослать нашему первому (геометрическому) доказательству.

Это арифметическое введение касается вопроса о четных и нечетных числах. Четным, как известно, называется всякое число, которое в два раза больше какого-либо другого целого числа, т. е. число, имеющее вид $2x$; нечетным называется всякое число, следующее в натуральном ряду чисел непосредственно за четным и имеющее вид $2x+1$. Квадрат нечетного числа есть всегда опять нечетное число, ибо

$$(2x+1)^2=4x^2+4x+1=2(2x^2+2x)+1$$

на единицу больше четного числа, т. е. нечетно.

Отсюда непосредственно получается

1-я лемма. *Если квадрат некоторого числа есть число четное, то и само это число четно.*

Действительно, если бы оно было нечетным, то, согласно только что сказанному, и квадрат его был бы числом нечетным.

2-я лемма. *Квадрат четного числа всегда делится на 4, т. е. представляет собой некоторое учетверенное целое число вида $4g$.*

В самом деле, $2x\cdot 2x=4x\cdot x=4g$.

Само доказательство проведем снова методом от противного. Предположим, что сторона и диагональ квадрата имеют общую меру E и что в диагонали эта общая мера E содержится, например, точно d раз, а в стороне — точно s раз. Применим теорему Пифагора к одному из двух прямоугольных треугольников, на которые квадрат разделяется диагональю. Мы получим:

$$d^2=s^2+s^2,$$

или

$$d^2=2s^2. \tag{5}$$

Далее, мы имеем право предположить, что оба целых числа d и s взаимно простые, ибо если бы эти числа можно было сократить на какой-либо общий множитель, подобно тому, как, например, числа 10 и 16 сократимы до 5 и 8, то это значило бы, что в качестве общей меры взята слишком малая величина и что ее можно увеличить в соответствующее число раз. Положим поэтому, что такое сокращение произведено уже заранее.

В таком случае из (5) следует, что d^2, будучи в два раза больше s^2, есть число четное. 1-я лемма позволяет сделать отсюда вывод, что и d — также число четное. Но тогда s должно быть нечетным, ибо если бы и d и s оба были четными числами, то их можно было бы сократить на 2, а эта возможность противоречит сделанному нами предположению, что эти числа сокращены уже заранее.

С другой стороны, 2-я лемма показывает, что если d — четное число, то d^2 делится на 4, т. е. $d^2 = 4g$; но тогда $2s^2 = 4g$, и следовательно, $s^2 = 2g$. Если квадрат s есть число четное, то опять-таки, в силу 1-й леммы, и s должно быть четным — вопреки тому, что, как мы показали раньше, в случае четного d число s должно быть нечетным. Мы пришли к противоречию со сделанным нами в начале доказательства предположением о соизмеримости стороны и диагонали квадрата.

Существенным моментом в обоих доказательствах является то, что убывающая последовательность целых положительных чисел должна когда-нибудь иметь конец. В первом доказательстве это положение выступает ясно. Во втором оно скрыто в том месте, где идет речь о сокращениях, ибо при выполнении последовательных сокращений пары чисел в той мере, в какой это возможно, оба эти числа постоянно уменьшаются, хотя выше мы пользовались тем, что должна существовать уже несократимая дальше форма. Это заключение, которое повседневно применяется в школьном курсе без всякого критического исследования, опирается в сущности как раз на то соображение, что всякий процесс уменьшения целых положительных чисел должен рано или поздно закончиться.

Переписав формулу (5) в виде

$$\left(\frac{d}{s}\right)^2 = 2,$$

мы можем выразить результат доказательства в следующей форме: не существует такой дроби (такого рационального числа) $x = \frac{d}{s}$, квадрат которой был бы равен 2, или, иными словами, не существует рационального числа x, равного $\sqrt{2}$; $\sqrt{2}$ есть число иррациональное [16].

5. ОДНО МИНИМАЛЬНОЕ СВОЙСТВО ТРЕУГОЛЬНИКА, ОБРАЗОВАННОГО ОСНОВАНИЯМИ ВЫСОТ, ПО Г. ШВАРЦУ

На этот раз мы опять займемся задачей на максимум или, вернее, на минимум. Здесь мы познакомимся с образцом наглядного и вместе с тем очень тонкого математического рассуждения. Оно принадлежит математику Герману Амандусу Шварцу, еще сравнительно недавно преподававшему в Берлине и блестяще проявившему себя как в подобного рода математических миниатюрах, так и в крупных трудах [17].

1. В качестве подготовительного упражнения предпошлем этой задаче другую, чрезвычайно простую задачу, связанную с известным законом отражения света. Этот закон гласит, что если луч света (рис. 19, a), выходящий из точки A,

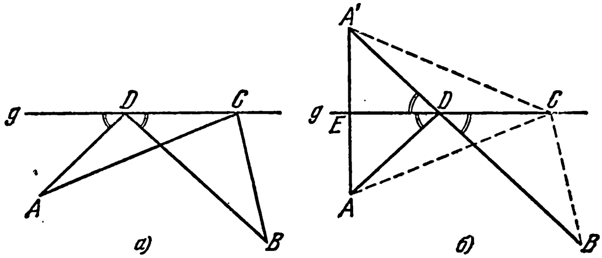


Рис. 19.

отражается от плоского зеркала g и после отражения попадает в точку B, то угол падения луча на прямую g равен углу его отражения.

Мы покажем, что путь, по которому идет при этом луч света, чтобы из точки A после прикосновения к зеркалу g попасть в B, есть наикратчайший, т. е. тот путь, который должен быть выбран пароходом, идущим из пункта A в пункт B с заходом на берег, изображаемый прямою g. Мы не собираемся здесь останавливаться на вопросе, почему луч света, лишенный способности мышления, избирает тот же самый путь, что и одаренный этой способностью капитан парохода; мы хотим лишь установить чисто математический факт, что *путь ADB, совершаемый при условии равенства* *углов падения и отражения, короче всякого другого пути* ACB [18].

Доказательство основано на одном приеме, который с математической точки зрения похож на фокус, но становится вполне естественным, если обратиться к его оптическому истолкованию. Отразим в зеркале g точку A и прямые AC и AD. Рис. 19, *б* повторяет рис. 19, *а* с присоединением этого зеркального отображения. Пусть A' будет зеркальным отображением A, тогда $A'C$ будет зеркальным отображением AC и $A'D$ — отображением AD. Так как $A'C = AC$ и $A'D = AD$, то треугольники EDA и EDA' будут равны, и следовательно, $\angle EDA = \angle EDA'$. Но по предположению $\angle EDA = \angle CDB$. Следовательно, угол CDB играет по отношению к углу EDA' роль вертикального угла, а это значит, что $A'D$ есть прямолинейное продолжение DB.

Ломаная линия ADB равна отрезку $A'DB$, а $ACB = A'CB$. Так как $A'DB$, как только что доказано, есть отрезок, соединяющий точки A' и B, то $A'DB$ короче $A'CB$, а следовательно, и ADB короче ACB, что и требовалось доказать.

2. Задача, которой мы намерены сейчас заняться, состоит в том, чтобы *в данный остроугольный треугольник* ABC *вписать треугольник* UVW *наименьшего периметра* (рис. 20, *а*). Мы утверждаем, что *треугольник* EFG,

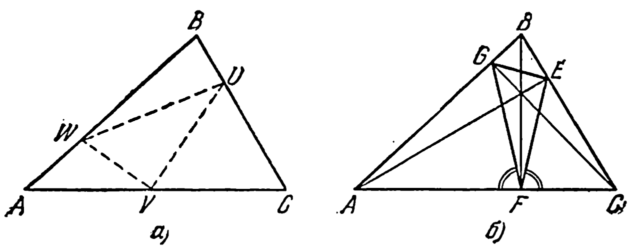


Рис. 20.

вершинами которого являются E, F, G — *основания высот, опущенных из вершин данного треугольника* ABC *на противолежащие стороны* (*рис.* 20, *б*), *имеет меньший периметр, чем всякий другой треугольник* UVW, *вписанный в* ABC.

Докажем сначала одну вспомогательную теорему. А именно: докажем, что *угол* AFG *равен углу* CFE (как в законе отражения света) и что соответствующие равенства имеют место и для углов при вершинах E и G. Для доказательства этой теоремы нам придется вспомнить кое-что из школьной геометрии, а именно так называемую теорему Фалеса (угол, опирающийся на диаметр,— прямой; рис. 21), теорему о вписанных углах (рис. 22) и теорему о высотах треугольника (высоты треугольника пересекаются в одной точке H)

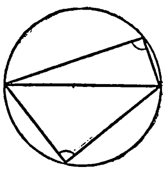


Рис. 21.

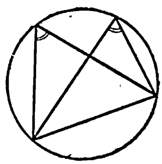


Рис. 22.

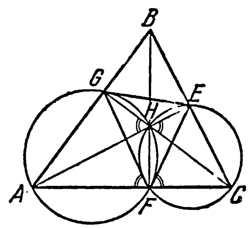


Рис. 23.

Пользуясь этими теоремами, мы найдем из рис. 23, что окружность, построенная на диаметре AH, проходит через G и F, а окружность, построенная на диаметре CH, пройдет через E и F; далее, угол AFG равен углу AHG (как вписанные, опирающиеся на одну и ту же дугу AG), и соответственно $\angle CFE = \angle CHE$. Но углы AHG и CHE равны, как вертикальные, следовательно, и $\angle AFG = \angle CFE$, что и требовалось доказать.

3. После этих предварительных замечаний мы приступаем к доказательству Г. А. Шварца. Отразим треугольник ABC

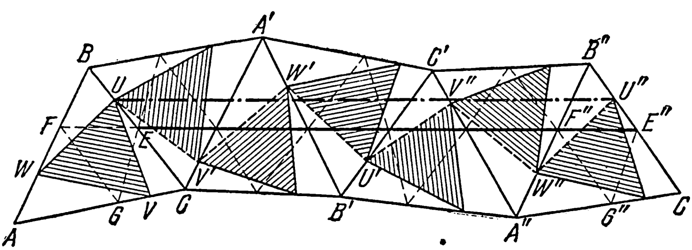


Рис. 24.

(рис. 24) от стороны BC; полученный таким путем треугольник $A'BC$ отразим от стороны CA'; результат этого второго отражения отразим от стороны $A'B'$ и после этого произведем еще три отражения последовательно от сторон $B'C'$, $C'A''$ и $A''B''$.

Обратим прежде всего внимание на то, что, как это непосредственно видно из рисунка, конечное положение $A''B''C''$ может быть получено из начального положения ABC простым параллельным переносом. Чтобы в этом убедиться, посмотрим, что происходит с треугольником после двух первых отражений. Вместо того чтобы подвергать его двукратному отражению, мы можем перевести наш исходный треугольник в третье положение простым вращением его по часовой стрелке в его плоскости вокруг неподвижной вершины C на угол 2γ (α, β, γ — углы, противолежащие соответственно вершинам A, B, C). Точно так же мы могли бы перевести его из этого положения в пятое простым вращением по часовой стрелке вокруг неподвижной точки B' на угол 2β, и, наконец, вращением на угол 2α вокруг неподвижной точки A'' мы привели бы его в положение седьмое, т. е. в конечное положение. В общем мы повернули бы треугольник по часовой стрелке на угол $2\gamma+2\beta+2\alpha$, т. е. (так как в треугольнике сумма углов $\alpha+\beta+\gamma=2d$) на угол, равный $4d$. В результате треугольник совершает полный оборот и принимает то же самое положение, что и вначале, оказываясь лишь передвинутым в своей плоскости параллельно самому себе. Следовательно, BC параллельна $B''C''$.

Проследим теперь за теми изменениями, которые получит при этих последовательных отражениях положение треугольника, образованного основаниями высот. В дальнейшем этот треугольник мы будем обозначать знаком $\triangle$.

На основании доказанной выше леммы мы замечаем прежде всего, что отрезок EG в своем втором положении будет прямолинейным продолжением отрезка FE в первом положении и что точно так же в последующих положениях одна из сторон треугольника $\triangle$ будет последовательно располагаться на продолжении прямой, проходящей через FE. *Поэтому выделенная на рисунке жирным прямая* EE'' будет состоять из шести отрезков, из которых два равны FG, два — GE и два — EF; следовательно, она *равна удвоенному периметру треугольника* $\triangle$.

Проследим точно таким же образом за положениями, которые будет последовательно принимать какой-либо другой треугольник UVW, вписанный в данный треугольник ABC,—

штриховка на чертеже облегчит нам эту задачу. Совершенно так же, как и раньше, мы убедимся, что *выделенная черным пунктиром ломаная линия* $UV'W'U'V''W''U''$, *простирающаяся от точки* U *до* U'', *равна удвоенному периметру треугольника* UVW. В четырехугольнике $EE''U''U$ противоположные стороны UE и $U''E''$ параллельны и равны, как соответственные отрезки в различных положениях треугольника ABC. Следовательно, согласно известной теореме элементарной геометрии, этот четырехугольник есть параллелограмм, а значит, и две его другие противоположные стороны равны: $UU''=EE''$. Таким образом, UU'' также равно двойному периметру треугольника $\triangle$, образованного основаниями высот.

Но, с другой стороны, непосредственно очевидно, что UU'' короче проведенной между теми же конечными точками ломаной, равной, как мы обнаружили, удвоенному периметру треугольника UVW. Следовательно, *периметр треугольника* $\triangle$ *меньше периметра треугольника* UVW, что и требовалось доказать.

Это подлинно математическое доказательство. Предпосылка и утверждение преобразовываются здесь таким образом, что сущность теоремы сразу же становится ясной [19].

6. ТО ЖЕ МИНИМАЛЬНОЕ СВОЙСТВО ТРЕУГОЛЬНИКА ПО Л. ФЕЙЕРУ

1. Предыдущая тема была посвящена доказательству того, что *из всех треугольников, вписанных в данный остроугольный треугольник, наименьшим периметром обладает тот, вершины которого совпадают с основаниями высот данного треугольника.* Если мы теперь для этой же самой теоремы даем еще второе доказательство, то делаем это по той причине, что метод решения излагаемых здесь вопросов для нас важнее, чем их математическое содержание. В только что изложенном доказательстве интересующей нас теоремы мы, следуя Шварцу, во-первых, воспользовались тем основным фактом, что прямая линия представляет собой кратчайшее расстояние между двумя точками, и, во-вторых, применили преобразование фигуры по способу зеркального отражения. Эти два принципа лежат также и в основе второго доказательства. Особый интерес представляет здесь как раз противопоставление различных способов применения этих принципов в обоих случаях. Нижеследующее доказательство принадлежит венгерскому математику Л. Фейеру (L. Fejer.) Он, еще будучи студентом, нашел это доказательство, чрезвычайно понравившееся самому Шварцу [1]).

2. Пусть в данный нам остроугольный треугольник ABC (рис. 25) вписан произвольный треугольник UVW так, что его вершина U лежит на стороне BC, V — на стороне CA и W — на стороне AB.

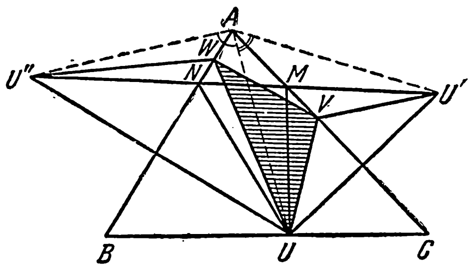


Рис. 25.

Отобразим зеркально вершину U от двух прямых AC и AB Пусть ее зеркальными образами будут соответственно точки U' и U''. Тогда, в силу основных свойств зеркального отображения, отрезок UV равен отрезку $U'V$, а отрезок UW равен отрезку $U''W$. Поэтому периметр треугольника UVW, составленного из отрезков UV, VW и WU, будет равен длине ломаной линии $U'VWU''$.

Если мы оставим точку U на прежнем месте, а точкам V и W придадим другие положения, то точки U' и U'', положения которых определяются одной лишь точкой U и данным треугольником ABC, останутся неподвижными. Это означает, что ломаная линия $U'VWU''$, длина которой всегда равна периметру треугольника UVW, при всех возможных положениях точек V и W остается натянутой между неподвижными концами U' и U''. Но линия, соединяющая точки U' и U'', будет кратчайшей лишь в том случае, если она будет прямой. Следовательно, отрезок $U'U''$ даст нам величину наименьшего периметра, который может иметь вписанный

[1]) Доказательство Фейера не было нигде опубликовано и излагается здесь впервые.

треугольник с фиксированной вершиной U. Обозначим две остальные вершины такого треугольника с минимальным периметром и фиксированной вершиной U через M и N.

3. Поскольку из всех возможных вписанных треугольников с общей вершиной U мы уже отыскали тот, периметр которого имеет наименьшее значение, нам остается теперь сравнить между собою все минимальные треугольники, соответствующие различным положениям вершины U, и выбрать из них тот, периметр которого будет наименьшим; очевидно, что этот треугольник и будет обладать наименьшим периметром сравнительно со всеми другими возможными вписанными треугольниками.

Требуется, следовательно, поместить вершину U так, чтобы отрезок $U'U''$ был наименьшим. Заметим прежде всего, что треугольник $AU'U''$ р а в н о б е д р е н н ы й, ибо две его стороны AU' и AU'', будучи зеркальными отображениями одного и того же отрезка AU, равны: $AU = AU' = AU''$.

К тому же, если длина этих двух сторон треугольника $AU'U''$ зависит от положения точки U на BC, то *величина угла $U'AU''$ от положения U не зависит* и вполне определяется заданным треугольником ABC. Действительно, между углами нашей фигуры существуют следующие соотношения:

$$\angle UAB = \angle U''AB,$$
$$\angle UAC = \angle U'AC;$$

отсюда

$$\angle U''AU = 2\angle UAB$$

и

$$\angle U'AU = 2\angle UAC,$$

следовательно,

$$\angle U'AU + \angle U''AU = 2\angle UAB + 2\angle UAC,$$

или

$$\angle U'AU'' = 2\angle CAB,$$

что и доказывает наше утверждение относительно угла $U'AU''$.

4. Далее, отрезок $U'U''$, который мы должны сделать наименьшим, является основанием равнобедренного треугольника $AU'U''$. Так как угол при вершине этого треугольника от положения U не зависит, то во всех треугольниках $AU'U''$, полученных при всевозможных положениях U, углы при вершине совпадают. Из этих треугольников наименьшее основание имеет тот, у которого боковые стороны наименьшие. Но стороны AU' и AU'', как уже доказано, равны расстоянию AU. Следовательно, мы получим наименьший отрезок $U'U''$, если выберем U так, чтобы расстояние AU было наикратчайшим. Вопрос о наименьшем отрезке $U'U''$ сводится тем самым к задаче о нахождении наименьшего отрезка AU.

Но отрезок AU соединяет точку A с прямой BC. Так как известно, что кратчайшим расстоянием от точки до прямой является перпендикуляр, опущенный из точки на прямую, то отрезок AU должен быть перпендикулярен к BC или, иными словами, должен быть высотой треугольника ABC, опущенной из вершины A.

5. Теперь мы можем построить искомый вписанный треугольник EFG наименьшего периметра (рис. 26). Пусть E — основание перпендикуляра, опущенного из A на BC. Обозначим

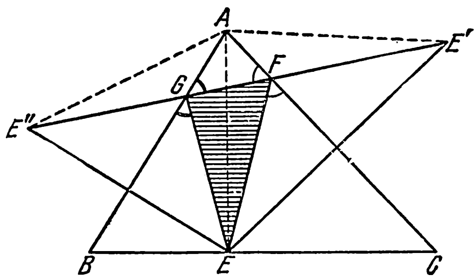


Рис. 26.

буквами E' и E'' зеркальные отображения E, получающиеся при отражении этой точки от AC и AB; тогда отрезок EE' будет равен наименьшему периметру вписанного треугольника. Две другие вершины искомого треугольника определяются точками пересечения F и G прямой $E'E''$ со сторонами AC и AB.

Действительно, всякий отличный от EFG вписанный треугольник UVW должен иметь периметр бóльший, чем треугольник EFG (в чем легко убедиться, проследив еще раз за ходом наших рассуждений). В таком треугольнике, если вершина U не будет совпадать с точкой E, то соответствующий этой вершине отрезок $U'U''$, а следовательно, и периметр треугольника, который не меньше, чем $U'U''$, будет больше отрезка $E'E''$. Если же вершина U совпадает с E, а по крайней мере одна из вершин V или W не совпадает соответственно с точкой

F или G, то вместо прямолинейного отрезка $EFCE''$ мы получим ломаную $E'VWE''$, т. е. и в этом случае периметр UVW будет больше периметра EFG.

6. Задача об отыскании вписанного треугольника с наименьшим периметром имеет, следовательно, только одно решение. Этой однозначностью мы и воспользуемся для некоторых дальнейших выводов. Наше построение минимального треугольника было отнюдь не однородно относительно трех его вершин. Одна вершина E была найдена как основание высоты, опущенной из A. Две другие вершины F и G — без проведения высот из B и C, посредством построения зеркальных отображений точки E. Ясно, что все наши построения, имевшие отправной точкой вершину A, можно было бы начать, например, с вершины B, и тогда вместо того, чтобы отображать, как это мы делали в § 2, точку U от сторон AB и AC, мы отобразили бы точку V от сторон BA и BC и соответственно продолжали бы дальше наши построения. При этом в качестве минимального мы получили бы треугольник, у которого вершина F совпала бы с основанием высоты, опущенной из B. Но, как мы только что установили, существует л и ш ь о д и н минимальный треугольник, и следовательно, построение, имеющее своей отправной точкой вершину B, должно привести к тому же самому треугольнику EFG, что и построение, начатое с точки A. Поскольку то же самое справедливо и относительно точки C, мы заключаем, что в минимальном треугольнике EFG не только E, но также F и G являются основаниями высот, опущенных из противолежащих вершин. Этим исчерпывается доказательство высказанной выше теоремы о минимальных свойствах треугольника, образованного основаниями высот.

7. Но наше доказательство дает нам вместе с тем и коечто бóльшее. Если мы, исходя из единственности решения, приписали точкам F и G свойство, обнаруженное в точке E, то можно, наоборот, свойство, присущее точкам F и G, распространить также и на точку E. На основании свойств зеркального отражения $\angle EFC = \angle E'FC$. Вертикальные углы $E'FC$ и GFA также равны, откуда вытекает, что $\angle EFC = \angle GFA$, или, иными словами, обе проходящие через F стороны минимального треугольника образуют со стороной AC основного треугольника равные углы. Соответствующее равенство углов имеет место и в точке G. Если мы теперь построим точку F как основание опущенной из B высоты, то точку E можно найти путем построения зеркального отражения, причем углы GEB и FEC при точке E будут также равны.

Отвлечемся теперь от минимальных свойств треугольника EFG; мы знаем, на основании п. **6**, что этот треугольник можно однозначно определить как треугольник, образованный основаниями высот.

Сопоставляя это определение с только что приведенными результатами, получаем теорему: *стороны треугольника, вершинами которого являются основания высот остроугольного треугольника ABC, образуют со сторонами этого последнего треугольника попарно равные углы, так что к каждой стороне треугольника ABC примыкает пара равных углов.*

В этой теореме уже не содержится больше никаких высказываний относительно минимальных свойств, и она всецело принадлежит обычной элементарной геометрии, средствами которой и может быть доказана. Такое доказательство мы, собственно, и дали на стр. 38. Для доказательства минимальных свойств треугольника $\triangle$ по способу Шварца как раз и требовалось существенное дополнение из элементарной геометрии в виде теорем, относящихся к учению об окружности. Преимущество доказательства Фейера заключается в том, что оно ведет к цели без всяких других вспомогательных средств, кроме принципов кратчайшего расстояния и зеркального отражения. Помимо этого доказательство Фейера замечательно еще и тем, что оно пользуется только двумя зеркальными отражениями, в то время как в доказательстве Шварца их имеется шесть [[20]].

8. Имеется теорема (в известном смысле сопряженная с теоремой о треугольнике, образованном основаниями высот), гласящая, что *в любом остроугольном треугольнике ABC существует одна и только одна точка P, расстояния от которой до трех вершин треугольника имеют наименьшую сумму. Эта точка P расположена так, что прямые, соединяющие ее с вершинами треугольника, образуют друг с другом углы по* 120°.

Л. Шрутка (L. Schruttka) в специальном сборнике к 50-летнему докторскому юбилею Г. А. Шварца (Берлин, 1914) дал доказательство этой теоремы, намеренно построив его по образцу доказательства самого Шварца, составившего содержание предшествующей темы. Позже кенигсбергский студент Бюкнер (Bückner) нашел изумительно краткое доказательство этой теоремы, которое стоит в таком же отношении к доказательству Шрутки, как доказательство основной теоремы, данное Фейером, к доказательству Шварца. Оно требует всего нескольких строк.

Пусть P — произвольная точка внутри треугольника ABC (рис. 27, *а*, *б*). Повернем треугольник ACP вокруг точки A на 60°; пусть он займет при этом положение $AC'P'$. Поворот при этом должен производиться так, чтобы сторона AC поворачивалась всегда в сторону, внешнюю для треугольника ABC, иначе говоря, чтобы после поворота прямая AC оказалась

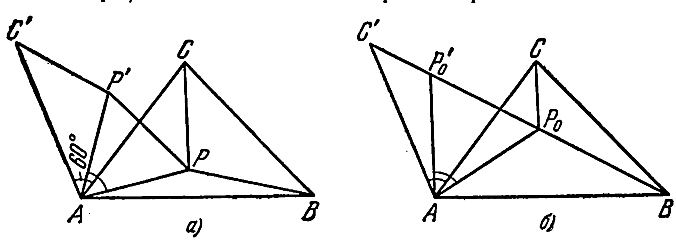


Рис. 27.

расположенной между AC' и AB. Тогда будем иметь $C'P' = CP$, $PP' = AP$ (ибо в треугольнике APP' угол при A равен 60°, и следовательно, этот треугольник не только равнобедренный, но и равносторонний). *Ломаная $BPP'C'$ представляет собою, таким образом, сумму расстояний от точки P до трех вершин треугольника A, B, C.* Точка C' от положения P совершенно не зависит. Каково бы ни было положение точки P в треугольнике, все получающиеся при этом ломаные соединяют точки B и C'. Кратчайшей из всех этих ломаных будет прямолинейный отрезок BC'. Таким образом, чтобы точка P удовлетворяла поставленному условию минимума, она должна лежать на отрезке BC', причем ее положение однозначно определяется тем, что угол AP_0C' должен быть равен 60°. Дополнительный угол AP_0B равен, следовательно, 120°. Построение показывает, что существует только о д н а точка P_0, удовлетворяющая условию задачи. Поэтому, если мы в основу построения возьмем не вершину A, а любую из двух других вершин, то мы придем к той же самой точке P_0. Значит, и углы BP_0C и CP_0A также равны 120° [21].

7. ЭЛЕМЕНТЫ ТЕОРИИ МНОЖЕСТВ

Мы посвящаем эту главу предмету, близко соприкасающемуся с основами нашей науки, не в силу философской важности этих основ, а по той причине, что крайне простые в своей сущности, не требующие никаких предварительных познаний идеи и выводы великого основоположника теории множеств Георга Кантора являют собой образец подлинно математического стиля[1]). Подлинная математика заключается не в нагромождении искусственных вычислительных приемов, а в умении получать нетривиальные результаты путем размышления при минимуме применяемого аппарата.

Каких чисел больше — целых или четных? Где больше точек — на отрезке прямой или на площади квадрата? Из такого рода вопросов исходил Г. Кантор при построении теории множеств. В попытках ответить на эти вопросы здесь важно не допустить логического прыжка. Разумеется, в только что приведенных формулировках эти вопросы лишены точного смысла, и первый знаменательный шаг Кантора состоял как раз в том, что он впервые сообщил им совершенно точный и ясный смысл. Исходя из обычного счета конечного числа предметов, Кантор отметил имеющееся здесь существенное различие, которое грамматически выражается разницей между количественными и порядковыми числительными.

Представьте себе, что вы, находясь в танцевальном зале, задаетесь вопросом, кого больше среди громадного количества присутствующих — мужчин или женщин. Как это можно установить? Один из возможных способов таков: руководитель выстраивает в ряд всех мужчин у одной стены, всех женщин — у другой, пересчитывает оба ряда и сравнивает полученные числа. Много проще другой способ: стóит только начать

---

[1]) Значение теории множеств отнюдь не исчерпывается ее философскими выводами. Устанавливая общие закономерности для явлений и фактов, принадлежащих различным, порой весьма далеко отстоящим одна от другой областям математики, внутренне систематизируя содержание этих разнообразных областей, теория множеств представляет в настоящее время необходимую базу для развития основных разделов математики — алгебры, топологии, теории вероятностей, функционального анализа. Через эти науки устанавливается и связь теории множеств с техникой. (*Прим. перев.*)

танцы, каждый мужчина пригласит даму, и тогда само собой выясняется, кого больше — мужчин или женщин.

Этот принцип попарного сочетания или взаимно однозначного соответствия Кантор и принял за исходный пункт своих построений. Например, если он хочет узнать, больше ли целых чисел, чем четных, то он спрашивает, можно ли или нельзя установить между обоими этими видами чисел такое попарное соответствие, при котором ни одно число не остается без пары. И вот оказывается, что эти числа действительно возможно объединить в пары так, что каждое целое число составит пару с некоторым четным и при этом ни одно не останется лишним. Соответствие это устанавливается так:

$$\begin{array}{l} 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ \ldots \\ 2,\ 4,\ 6,\ 8,\ 10,\ 12,\ \ldots \end{array}$$

Каждое целое число верхнего ряда сочетается здесь с расположенным под ним четным числом нижнего ряда. Ни одно число ни в верхнем, ни в нижнем ряду не остается без соответствующей ему пары. Кажется удивительным, что такое объединение в пары между числами верхнего и нижнего рядов осуществляется без остатка, несмотря на то, что нижний ряд не содержит в себе никаких других чисел, кроме части чисел верхнего ряда.

Здесь, естественно, возникает одно возражение. Бросается в глаза разница между только что разобранным случаем и случаем конечных чисел. Так, в примере с танцевальным залом, о котором у нас шла речь раньше, совершенно безразлично, какую женщину пригласит тот или иной мужчина; как бы ни происходили эти различные выборы, число оставшихся в излишке не будет от них зависеть и будет оставаться неизменно одним и тем же, если только новые лица не войдут в зал или кто-либо из него не выйдет. В примере с целыми и четными числами дело обстоит иначе. Выше мы указали для этих чисел объединение в пары, удавшееся без остатка; можно, однако, наряду с этим указать другое объединение в пары, уже не удовлетворяющее этому условию и нарушающее равновесие в пользу целых чисел. Я могу, например, объединять в пары таким, пожалуй, даже более естественным способом: числу 2 первого ряда поставить в соответствие число 2 второго, 4 сочетать с 4, 6 с 6 и т. д. Тогда все четные числа войдут в состав пар, нечетные же останутся лишними [22]. Существенным в рассуждениях Кантора является именно то, что он отказывается от произвольности в выборе способа сочетания в пары, требуя существования хотя бы одного такого объединения в пары, которое удается без остатка, и в этом случае он говорит, что *оба рассматриваемых множества «имеют одинаковую мощность»*.

Кантор утверждает дальше, что *множество всех рациональных чисел, т. е. всех целых чисел вместе со всеми дробями, имеет мощность, не бо́льшую мощности множества целых чисел.*

Доказательство этого утверждения основано на том, что дроби располагаются не по их величине, а по величине суммы, получающейся при сложении числителя со знаменателем. Сначала берут все дроби (конечно, уже сокращенные), у которых сумма числителя и знаменателя равна 2, — этим свойством обладает только одна дробь $\frac{1}{1}=1$; затем все те, у которых сумма равна 3, а именно $\frac{1}{2}$ и $\frac{2}{1}=2$; далее берут дроби с суммой числителя и знаменателя, равной 4: $\frac{1}{3}$, $\frac{2}{2}$, $\frac{3}{1}$, причем $\frac{2}{2}$ мы отбрасываем, так как это дробь сократимая; затем — с суммой, равной 5: $\frac{1}{4}$, $\frac{2}{3}$, $\frac{3}{2}$, $\frac{4}{1}$; и т. д. В этом же порядке их записывают последовательно под числами натурального ряда

$$\begin{array}{ccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & \dots \\ 1 & \frac{1}{2} & 2 & \frac{1}{3} & 3 & \frac{1}{4} & \frac{2}{3} & \frac{3}{2} & 4 & \frac{1}{5} & 5 & \frac{1}{6} & \dots \end{array}$$

Теперь у нас при попарном сочетании чисел, стоящих одно над другим, получается взаимно однозначное соответствие. Действительно, в нижнем ряду не может отсутствовать ни одно рациональное число, ибо каждое рациональное число имеет определенную сумму числителя и знаменателя и поэтому должно рано или поздно появиться в этом ряду.

Кантор выражает этот поразительный факт, говоря, что множество всех рациональных чисел есть множество «счетное», так как эти числа, объединенные в пары с числами натурального ряда, можно пересчитывать по одному. Иными словами, он называет множество счетным, если оно может быть приведено во взаимно однозначное соответствие с множеством натуральных чисел, т. е. если оно имеет одинаковую с этим последним мощность. Кантор доказывает далее, что некоторые множества, казалось бы значительно более обширные, чем множество рациональных дробей, в действительности имеют также мощность счетного множества. Мы опустим эти рассуждения, требующие более высоких математических познаний и представляющие собой лишь новые иллюстрации к тому, что было только что разъяснено на одном случае.

Наконец, Кантор приходит к тому открытию, которое собственно впервые и сообщило его теории достаточную жизнеспособность: он обнаружил, что существуют множества более мощные, чем множество чисел натурального ряда. А именно он доказывает, что *множество всех точек отрезка имеет более высокую мощность, чем множество чисел натурального ряда.*

Доказательство ведется от противного: приводится к противоречию предположение, что число точек отрезка счетно, т. е. что между точками отрезка, длиной, например, в 1 *м*, и целыми числами существует взаимно однозначное соответствие.

Предположим, что можно каким-то способом пересчитать точки отрезка, записав их в ряд, само собой разумеется, не в порядке их обычного расположения, а в какой-либо другой последовательности, так что можно назвать первую точку, вторую и т. д. Лучше всего численно характеризовать точки нашего метра в сантиметрах, миллиметрах и т. д., отсчитываемых по шкале-эталону. При таком способе середина отрезка характеризуется числом 0,5, и аналогично всякая другая точка характеризуется некоторой десятичной дробью, в которой при полной точности будет, вообще говоря, бесконечно много знаков; например, для точки, лежащей в конце первой трети отрезка, будем иметь дробь 0,3333... Предположение, что множество точек нашего отрезка счетно, переносится теперь на характеризующие их бесконечные десятичные дроби, которые в таком случае можно было бы занумеровать в известном порядке. На первом месте стояла бы какая-нибудь десятичная дробь, имеющая вид 0, ...; на втором — какая-то другая дробь того же вида и т. д. У нас получится при этом картина, аналогичная приведенной ниже, где для наглядности взят определенный пример. Из технических соображений ряд записывается теперь не горизонтально, а вертикально:

1	0,	**3**	5	4	2	0	...
2	0,	6	**1**	7	7	3	...
3	0,	5	5	**5**	4	9	...
4	0,	0	1	0	**0**	7	...
5	0,	2	0	2	0	**6**	...
.	.	.	.	.	.	.	. . .
.	.	.	.	.	.	.	. . .

Теперь нужно показать, что все же возможно найти такую точку, т. е. такую бесконечную десятичную дробь вида 0, . . ., которая не содержится в этом ряде. Мы строим ее следующим образом: на первое место после запятой ставим какую-либо произвольную цифру, отличную от первой цифры первой десятичной дроби последовательности. Мы имеем здесь выбор из 9 различных цифр. Чтобы остановиться на чем-нибудь определенном, возьмем цифру 1. Если первая цифра первой десятичной дроби как раз 1, то примем цифру 2. Можно быть уверенным теперь, что как бы ни выбирали следующие цифры конструируемой десятичной дроби, она во всяком случае будет отличаться от первой дроби написанного выше ряда, ибо, если две десятичные дроби различаются уже в первом знаке, они не могут характеризовать одной и той же точки, даже если бы все остальные их знаки и совпадали. Для второго знака выбираем опять 1, если только второй знак второй дроби не равен также 1; в этом случае мы взяли бы вторую цифру равной 2; иными словами, мы брали бы всякий раз такую цифру, чтобы она была отлична от второй цифры второй десятичной дроби. Тогда конструируемая десятичная дробь будет во всяком случае отличаться не только от первой, но и от второй из десятичных дробей последовательности. Таким же образом мы продолжим дальше. В нашем примере получающаяся десятичная дробь будет начинаться так: 0,12111..., и так как мы можем продолжать ее неограниченно, то тем самым определяется некоторая десятичная дробь, заведомо отличающаяся от всех дробей пересчитанной последовательности. Следовательно, эта последовательность не может заключать в себе все бесконечные десятичные дроби вида 0, . . ., вопреки предположению, сделанному в начале доказательства. Это значит, что множество точек отрезка не является счетным.

Нельзя не упомянуть об одном возражении, которое можно выдвинуть против этого доказательства; оно побудило Кантора облечь свое доказательство в совершенно иную форму, хотя это возражение и можно чрезвычайно легко устранить. Дело в том, что существуют такие бесконечные десятичные дроби, которые содержат, начиная с некоторого знака, одни лишь девятки, как, например, дробь 0,26999 ..., представляющая собой в действительности не что иное, как 0,27000... или, как в подобных случаях принято кратко писать, 0,27. Мы встречаемся здесь, таким образом, с тем неприятным обстоятельством, что две различные десятичные дроби характеризуют одну и ту же точку, в то время как в нашем доказательстве мы пользовались тем, что каждая точка метра характеризуется одной десятичной дробью. Выход из положения, однако, весьма несложен: мы просто исключаем из рассмотрения подобные бесконечные десятичные дроби с девятками. Осложнение в доказательстве могло бы теперь возникнуть лишь в том случае, если бы построенная в результате дробь оказалась составленной из девяток. Но эта возможность в доказательстве предусмотрена и исключена, ибо построенная десятичная дробь составлена лишь из цифр 1 и 2; цифра 9 здесь вовсе не встречается.

Полученный результат позволяет сделать интересный вывод. Так как мы уже знаем, что множество рациональных чисел менее мощно, чем множество всех чисел между 0 и 1, то, значит, в интервале между 0 и 1 обязательно должны быть числа, не принадлежащие к рациональным. Таким образом, существование иррациональных чисел обнаруживается здесь из рассуждений весьма общего характера, совсем иначе, чем при рассмотрении 4-й темы.

Следующий полученный Кантором результат является столь же неожиданным.

Множество точек площади квадрата имеет мощность не бóльшую, чем множество точек стороны квадрата.

Удивительна здесь полная независимость от понятия измерения: рассматриваемый как множество точек одномерный отрезок имеет одинаковую мощность с двумерным квадратом; и можно было бы точно так же показать, что и трехмерный куб имеет не бóльшую мощность.

Так же как и в предыдущем случае, доказательство проводится арифметически. Точки отрезка опять характеризуются бесконечными десятичными дробями вида 0, ..., за исключением дробей с девяткой в периоде. Точки квадрата отмечаются пáрами таких десятичных дробей, а именно, во-первых, дробью, которая измеряет (рис. 28) горизонтальное расстояние x точки от левой стороны квадрата, во-вторых, дробью, измеряющей ее вертикальное расстояние y от нижней стороны квадрата. Взаимно однозначное соответствие между точками квадрата и точками отрезка устанавливается теперь следующим образом: взяв какую-нибудь точку квадрата P, определяют обе характеризующие ее дроби:

$$x = 0,\ a_1a_2a_3 \ldots,\quad y = 0,\ b_1b_2b_3 \ldots;$$

затем образуют из них десятичную дробь

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

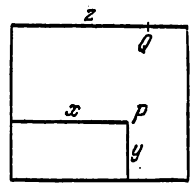


Рис. 28.

в которой цифры обеих первых дробей x и y размещены чередуясь, и, наконец, определяют такую точку Q на отрезке, которая характеризуется числом z. Например, центру квадрата будет соответствовать десятичная дробь $z = 0{,}550000\ldots$, полученная описанным выше способом из значений $x = 0{,}500\ldots$ и $y = 0{,}500\ldots$ Таким способом каждой точке квадрата будет поставлена в соответствие некоторая точка отрезка.

В этом собственно еще нет никакого достижения, то же самое можно было бы осуществить гораздо проще. Если, например, любой точке P квадрата мы поставим в соответствие основание перпендикуляра, опущенного из P на нижнюю сторону квадрата, то этим самым мы каждой точке квадрата поставили бы в соответствие точку нижней стороны. Но при этом точке нижней стороны соответствовала бы не одна точка квадрата, а все бесконечное множество их, заполняющее перпендикуляр с этим основанием. В описанном выше более тонком соответствии такая возможность исключена. Ибо если бы некоторая другая точка P' квадрата с определяющими числами

$$x' = 0,\ a'_1a'_2a'_3 \ldots,\quad y' = 0,\ b'_1b'_2b'_3 \ldots$$

была поставлена в соответствие с той же самой точкой Q верхней стороны квадрата, т. е. с тем же самым определяющим числом z, что и выше, то мы имели бы

$$z = 0,\ a'_1b'_1a'_2b'_2a'_3b'_3 \ldots$$

Но две десятичные дроби могут быть равны только в том случае, если они совпадают во всех своих знаках. Следовательно, все знаки этой дроби должны быть те же самые, что и написанные раньше для дроби z, т. е.

$$a'_1 = a_1, \quad b'_1 = b_1, \quad a'_2 = a_2, \quad b'_2 = b_2, \quad \ldots,$$

а это значит, что все знаки x' совпадают с соответствующими знаками x и все знаки y' — с соответствующими знаками y, вопреки предположению, что P' и P различны. Следовательно, одна и та же точка стороны не может соответствовать различным точкам квадрата.

Соответствие будет полным, если удастся установить, что *всякой* точке отрезка соответствует некоторая точка квадрата. Это действительно так, ибо пусть

$$z = 0,\ c_1c_2c_3c_4c_5c_6 \ldots$$

— определяющее ее число; тогда точка квадрата с определяющими числами

$$x = 0,\ c_1c_3c_5 \ldots, \quad y = 0,\ c_2c_4c_6 \ldots$$

будет соответствовать той точке отрезка, которая определяется числом, получающимся соединением дробей x и y указанным выше способом, а оно совпадает как раз с исходным числом z.

Против этого доказательства также может быть выдвинуто возражение в связи с рядами девяток. Имённо, когда мы находим точку квадрата, сопряженную с той точкой отрезка, которая определяется числом 0,2202020 ..., то характеризующие эту точку числа будут такими:

$$x = 0{,}2000 \ldots, \quad y = 0{,}2222 \ldots$$

С другой стороны, точке отрезка 0,12929292... соответствует точка квадрата с числами

$$x' = 0{,}1999 \ldots, \quad y = 0{,}2222 \ldots$$

Полученный ряд девяток приводит к тому, что две различные точки отрезка сопряжены с одной и той же точкой квадрата.

Эту двойственность можно исправить, если и не совсем очевидным, то во всяком случае чрезвычайно простым приемом. В вышеприведенной форме доказательство неправильно, но оно становится правильным после следующего видоизменения. В том случае, если в десятичной дроби встречаются девятки, их вместе с ближайшими примыкающими к ним соседними знаками объединяют в своего рода «молекулы» таким, например, образом: 0, (1) (2) (92) (92) (92), или, чтобы привести другой пример: 0, (7) (3) (94) (990) (9997) ..., и рассматривают эти молекулы как нераздельные. Если подобная дробь рассматривается как значение z, ее разбивают так, чтобы не раздробить этих молекул:

$$x = 0,\ (7)(94)(9997)\ \ldots,\quad y = 0,\ (3)(990)\ \ldots$$

Соответствие при этом выполнено по правилу, несколько отличающемуся от прежнего, и легко убедиться, что, при условии запрещения рядов с девятками, наше возражение отпадает.

В этом пункте своих исследований Кантор столкнулся с чрезвычайно интересной проблемой, а именно с вопросом: существует ли между множеством натуральных чисел и множеством точек отрезка некоторое промежуточное множество, мощность которого выше мощности первого множества и ниже мощности второго, или такое промежуточное множество не существует?

Эта проблема, не решенная и до сих пор, носит название «проблемы континуума».

Формулировка этой проблемы столь проста, что для ее понимания не требуется даже и тех знаний, которые преобретаются в средней школе,— достаточно лишь понятия о целом числе и отрезке,— и тем не менее ни одна математическая проблема не сопротивлялась столь упорно всем усилиям одолеть ее. В математике совсем нетрудно придумать десятки проблем, охватывающих ряд сложнейших построений и понятий, проблем, над которыми может быть безуспешно будут ломать головы даже талантливые математики. Но очень трудно — и в этом заключается подлинное искусство постановки математических задач — выдвинуть проблему, трактующую о самых простых понятиях и которую при этом нельзя легко решить, которая бы не была тривиальна. Проблема континуума с этой точки зрения, т. е. с точки зрения искусства постановки вопроса, представляет собой блестящее достижение.

Весьма скоро выяснилось, что проблема континуума связана с самыми основами теории множеств, а вместе с тем и всей математики и что исследование этой проблемы предполагает глубокий анализ понятия множества [23]. К такому анализу еще более настоятельно побуждают и возникшие с течением времени парадоксы теории множеств. Рассмотрением этих парадоксов мы и закончим настоящую тему.

Мы уже долго говорили о множествах. Наши множества состояли из каких-то «элементов». Множество точек отрезка, например, имеет своими элементами отдельные точки отрезка; равным образом, отдельные целые числа суть элементы множества целых чисел. Элементы в множестве — это то же самое, что члены в каком-либо обществе или объединении. Иногда членами объединения могут быть юридические лица, являющиеся в свою очередь объединениями; например, «Германский союз математиков» представляет собой не что иное, как объединение различных математических обществ, в него входят: «Общество германских математиков», «Общество содействия математическому образованию» и т. д. Точно так же и среди элементов множества могут быть такие, которые сами являются множествами, например, множество всех счетных множеств состоит лишь из таких элементов, которые сами являются множествами.

На заседании «Германского союза математиков» ни один член «Общества германских математиков» как таковой не имеет права голоса, поскольку последним обладают не отдельные математики, а математические общества; точно так же и число $\frac{1}{5}$, являющееся элементом множества всех рациональных чисел, не может считаться элементом множества всех счетных множеств; в качестве элемента в это более обширное множество входят множества рациональных чисел как целые.

Только хорошо освоившись с этими новыми понятиями, мы сможем поставить своеобразный вопрос: может ли множество содержать само себя в качестве элемента? Для всех тех обычных множеств, с которыми мы до сих пор имели дело, на этот вопрос следует ответить отрицательно. Однако легко убедиться, что должны все же существовать и такие необычные множества. Например, *множество всех вообще мыслимых множеств* принадлежит без сомнения к таковым, поскольку оно само является множеством. Мы видим, что множества такого необыкновенного вида существуют. Будем пока называть множества, которые содержат сами себя как элемент, необыкновенными множествами, остальные — обыкновенными.

Сделаем еще один шаг дальше и рассмотрим *множество всех обыкновенных множеств*. Обозначим его через M и поставим вопрос, обыкновенное ли это множество или необыкновенное? Очевидно, что должно иметь место одно из двух положений: либо это множество обыкновенное, либо необыкновенное. Предположим, что это множество необыкновенное, тогда оно содержится среди своих элементов, т. е. оно встречается среди элементов множества M, которые все представляют собой множества обыкновенные, а это противоречит нашему предположению, что мы имеем дело с необыкновенным множеством. Наоборот, если бы оно было обыкновенным множеством, то оно не встречалось бы среди своих элементов, которые исчерпывают собой все без исключения обыкновенные множества, т. е. оно не могло бы быть обыкновенным множеством, оно должно было бы быть, следовательно, множеством необыкновенным, что опять-таки противоречит предположению. Таким образом, и эта вторая возможность приводит также к противоречию.

Этот парадокс вовсе не является специфическим достоянием теории множеств. Чтобы сделать эту мысль более ясной, приведем шуточную формулировку того же парадокса, в которой слово «множество» даже и не упоминается. Один солдат в полку был назначен для выполнения обязанностей цирюльника; точнее, ему было вменено в обязанность брить в полку всех тех и только тех, кто не бреется сам. Как этот солдат должен поступать с самим собой? Если он будет брить себя, значит он попадет в число тех, кто бреется сам, и тогда он не должен себя брить. Если же он не будет себя брить, то он будет принадлежать к тем, кто сам не бреется и кого как раз он-то и должен брить. Что же, наконец, должен делать этот человек, чтобы в точности выполнить возложенное на него поручение?

Мы имеем здесь дело с чисто логическим парадоксом; теория множеств впервые обнаружила его в явной форме, однако по своей сущности парадокс вовсе с ней не связан. Старая, скучная логика стала интересной. Уже много лет логики и математики в общей напряженной работе стараются освободить логику от ее старых аристотелевских форм [24].

8. СЕЧЕНИЯ ПРЯМОГО КРУГОВОГО КОНУСА

Эллипсом называется геометрическое место точек P, сумма расстояний которых до двух постоянных точек F_1 и F_2 постоянна (рис. 29)

$$PF_1 + PF_2 = s.$$

Эллипс вполне определяется заданием точек F_1 и F_2 и суммой расстояний s; точки F_1 и F_2 называются фокусами эллипса.

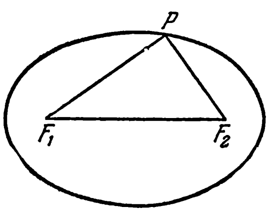


Рис. 29.

Если в точках F_1 и F_2 воткнуть две иголки и к ним прикрепить концы нити длиной s, то острие карандаша, натягивающее нить в точке P, опишет эллипс.

Аполлоний в своем учении о конических сечениях исходил из совершенно другого определения эллипса. Пусть перед нами в горизонтальной плоскости лежит окружность; представим себе, что из ее центра M к ее плоскости восставлен перпендикуляр MS и точка S соединена со всеми точками окружности (рис. 30). Полученная таким образом пространственная фигура есть прямой круговой конус; проведенные нами при этом прямые называются *образующими конуса*. Модель конуса легко приготовить из дерева на токарном станке. Если теперь обыкновенной плоской пилой срезать с него верхнюю часть, то при горизонтальном направлении распиловки мы получим в сечении круг, параллельный кругу основания конуса, а при наклонном направлении распиловки — некоторую продолговатую фигуру.

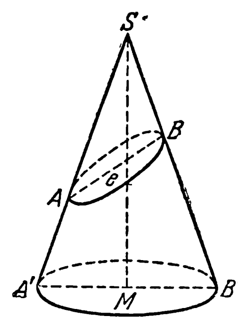


Рис. 30.

Непосредственно видно, что эта продолговатая фигура имеет AB осью симметрии, но то, что она имеет еще и другую, перпендикулярную к AB ось симметрии, непосредственно из чертежа уже не видно. Это свойство представляет собой следствие того важного и поразительного факта, что *фигура, полученная в сечении прямого кругового конуса, есть не что иное, как эллипс* (в определенном выше смысле).

Косвенно этот факт может быть доказан самыми различными способами. Мы даем здесь прямое и чисто геометрическое доказательство, принадлежащее Данделену (Dandelin) [25].

Прежде чем перейти к самому доказательству, сделаем два предварительных замечания:

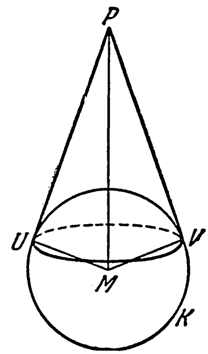


Рис. 31.

1) Из любой точки P (рис. 31) можно провести к данной окружности K две касательные; обе эти касательные имеют одинаковую длину. Представим себе теперь, что полученная таким образом плоская фигура (образованная окружностью и парой касательных) вращается вокруг оси PM; тогда PU (и PV) опишет прямой круговой конус, окружность K образует при этом вращении шар, причем образующие конуса, все одной и той же длины, будут касательными, проведенными из одной точки к этому шару. Мы получаем таким образом теорему: *все касательные, проведенные из одной точки к одному и тому же шару, имеют одну и ту же длину и являются образующими прямого кругового конуса.*

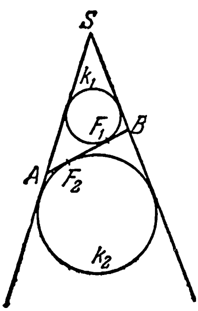


Рис. 32.

2) Пусть в некоторой плоскости (рис. 32) даны угол со сторонами SA и SB и прямая AB, пересекающая эти стороны. Далее, пусть дана окружность, вписанная в треугольник SAB, касающаяся стороны AB в F_1 изнутри, и, кроме того, еще вторая окружность, касающаяся тех же трех прямых, но примыкающая к AB извне в точке F_2. Заставим теперь фигуру, состоящую из угла ASB и двух окружностей (без прямой AB), вращаться вокруг биссектрисы угла ASB, на которой лежат центры обеих окружностей. Проведем через прямую AB плоскость e, перпендикулярную к плоскости SAB; тогда мы будем иметь прямой круговой конус с вершиной в S и двумя вписанными в него шарами, касающимися плоскости e, один сверху в точке F_1, другой снизу в F_2. Таким способом мы опять получаем конус, подобный изображенному на рис. 30, причем мы убедились, что в него можно вписать два шара K_1 и K_2.

Перейдем теперь к самому доказательству. Пусть P — какая-либо точка кривой, образованной пересечением плоскости e с конусом. Докажем, что эта точка принадлежит эллипсу (рис. 33).

Соединяем P с точками F_1 и F_2. Пусть k_1 и k_2 — окружности, по которым шары K_1 и K_2 касаются конуса. Проведем образующую конуса SP, проходящую через P; окружность k_1 она встретит в точке T_1, окружность k_2 — в T_2; в этих точках она, следовательно, касается обоих шаров. Так как, далее, плоскость e касается шара K_1 в точке F_1, то PF_1 также является касательной к шару K_1 в точке F_1. Таким образом, из P проведены две касательные к шару K_1: PT_1 и PF_1. Согласно первому предварительному замечанию, обе они равной длины:

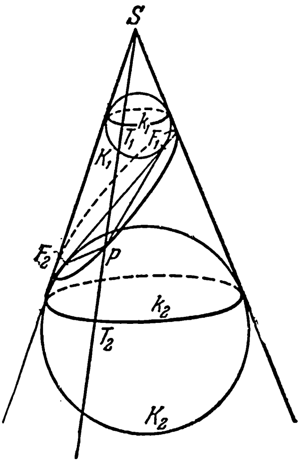


Рис. 33.

$$PF_1 = PT_1,$$

и точно так же

$$PF_2 = PT_2.$$

Отсюда в результате сложения следует, что

$$PF_1 + PF_2 = PT_1 + PT_2.$$

Правая часть последнего равенства не что иное, как T_1T_2 — тот отрезок прямой SP, который заключен между окружностями k_1 и k_2 и который сохраняет постоянную длину, когда точка P пробегает всю кривую сечения. Если обозначить эту постоянную длину T_1T_2 через s, то будем иметь

$$PF_1 + PF_2 = s.$$

Тем самым основное свойство нашей кривой доказано. Подобное же доказательство без каких-либо новых принципиальных затруднений можно провести также для гиперболы и параболы [26].

9. О КОМБИНАТОРНЫХ ЗАДАЧАХ

1. О том типе задач, которым мы сейчас будем заниматься, лучше всего составить себе представление, рассмотрев один частный пример. Пусть нам дано 4 красных, 1 черный и 2 белых шара, которые мы обозначим соответственно буквами

К, К, К, К, Ч, Б, Б.

Предположим, что эти 7 шаров, которые могут отличаться друг от друга только цветом, а в остальном совершенно одинаковы, размещаются в двух ящиках A и B; в обоих этих ящиках есть место как раз для всех 7 шаров. Допустим, что в A помещается не больше 3 шаров, в B не больше 4. Спрашивается: *скольким способами можно разместить эти 7 цветных шаров в ящиках A и B*?

Так как в данном простом случае мы имеем всего два ящика, то достаточно задать содержимое только одного из них, например A. Тогда в ящике B должны уложиться остальные 4 шара. Мы приступаем к своего рода систематическому обзору всех представляющихся при этом возможностей. Прежде всего можно ящик A заполнить исключительно красными шарами, оставив остальные четыре шара для ящика B:

1) в A: К К К, в B: К Ч Б Б.

При этом совершенно безразлично, какие именно из имеющихся у нас 4 красных шаров мы отберем в ящик A: ведь по условию шары отличаются лишь своим цветом, и следовательно, шары одного цвета неразличимы.

Затем можно в ящик A поместить только 2 красных шара, тогда для заполнения A придется взять еще 1 шар, либо черный, либо белый:

2) в A: К К Ч, в B: К К Б Б.
3) в A: К К Б, в B: К К Ч Б.

Если в A положить только 1 красный шар, то относительно двух других можно сразу сказать, что или оба они должны быть белыми (Б Б), или один из них будет белый, другой черный (Б Ч):

4) в A: К Ч Б, в B: К К К Б.
5) в A: К Б Б, в B: К К К Ч.

И, наконец, в A может не быть ни одного красного шара; тогда, чтобы заполнить A, мы должны взять все три имеющихся у нас некрасных шара — 1 черный и 2 белых:

6) в A: Ч Б Б, в B: К К К К.

Мы видим, следовательно, что возможны 6 различных размещений 4 красных, 1 черного и 2 белых шаров в двух ящиках, вмещающих соответственно 3 и 4 шара. Этот вывод мы сформулируем в более краткой форме, которой и будем пользоваться в настоящей главе: число размещений 4, 1, 2 шаров [1]) в ящиках вместимостью в 3 и 4 шара равно 6, или, еще короче, символически:

$$\{4,\ 1,\ 2\,|\,3,\ 4\}_7 = 6.$$

Внутри фигурных скобок слева от вертикальной черты указаны числа шаров различных цветов, справа от черты — меры вместимости ящиков. *Сумма чисел справа от черты должна быть, конечно, равна сумме чисел слева от черты, так как все шары должны заполнять как раз все ящики.* Это общее число размещаемых шаров (в нашем примере 7) мы для ясности указываем еще раз в виде индекса справа от скобок.

2. Перейдем теперь к общему случаю. Пусть нам дано n шаров f различных цветов и g ящиков общей вместимостью в n шаров. Символ

$$Z = \{к,\ ч,\ б,\ \ldots\,|\,a,\ b,\ c,\ \ldots\}_n \tag{1}$$

выражает число размещений этих n шаров, а именно: $к$ — красных, $ч$ — черных, $б$ — белых и т. д., в ящиках вместимостью a, b, c и т. д. Задача заключается в том, чтобы по числам n, $к$, $ч$, $б, \ldots$ и a, b, $c, \ldots$ найти Z. Мы попытаемся решить эту задачу, если и не во всей общности, то по крайней мере на различных примерах и для важнейших частных случаев. Прежде всего, разумеется, совсем необязательно заниматься размещением именно цветных шаров, а не каких-либо иных предметов. Например, выше для обозначения 4 красных, 1 черного и 2 белых шаров мы пользовались буквенными символами К К К К Ч Б Б и затем, вместо того чтобы распределять 7 шаров по двум ящикам A и B, мы размещали эти 7 букв в двух группах по 3 и 4 буквы в каждой.

3. Значение и удобство введенного символа для обозначения числа размещений станут для нас яснее, если мы разберем

---

[1]) Названия цветов можно спокойно опустить: ведь очевидно, что будь шары не красного, черного и белого цветов, а, скажем, желтого, зеленого и синего, — результат получился бы один и тот же.

еще несколько примеров, в которых о цветных шарах не будет больше речи. Впрочем, каждый пример мы попытаемся истолковать так, чтобы для него оказалась пригодной схема размещения цветных шаров.

Зададимся, например, вопросом, *скольким и способами можно разместить n лиц по n местам.* Чтобы яснее уловить в этой задаче связь с проблемой размещения цветных шаров по ящикам, нужно обратить внимание на то, что все эти n лиц различны, поэтому каждому из них можно поставить в соответствие некоторый цвет (его «имя»), и тогда наша задача равносильна такой: сколькими способами n шаров, окрашенных в различные цвета, могут быть размещены по n местам, причем каждому «месту» в соответствии с условиями старой задачи соответствует ящик вместимостью в один шар. В принятых нами символических обозначениях задача сводится к отысканию числа размещений:

$$P_n = \{1, 1, \ldots | 1, 1, \ldots\}_n. \quad (2)$$

Представим себе, что наши ящики (места) перенумерованы, тогда только что сформулированная задача сведется к тому, чтобы расположить в определенном порядке n разноцветных шаров (или n лиц). Всякое такое расположение называется перестановкой, так что мы можем сказать: P_n в формуле (2) представляет собой *число перестановок из n различных элементов*. Вопрос о том, какое числовое значение имеет P_n при данном n или, иными словами, каким образом, зная n, вычислить P_n, мы разберем позже, а сейчас дадим еще несколько примеров применения введенного нами символа.

4. *Сколькими способами можно распределить карты при игре в скат?* Каждый из трех игроков получает по 10 карт, две карты идут в «скат», все 32 карты различны. Очевидно, что здесь мы будем иметь то же самое число распределений, как и в случае, если бы нам нужно было 32 разноцветных шара разместить по ящикам вместимостью в 10, 10, 10 и 2 шара. Искомое число всех возможных распределений карт выразится с помощью введенного нами скобочного символа следующим образом:

$$S = \{1, 1, \ldots | 10, 10, 10, 2\}_{32}. \quad (3)$$

Фактическое вычисление числа S мы опять-таки пока отложим.

5. Следующий пример дает нам так называемая теорема о многочленах, которую мы рассмотрим здесь не во всей ее общности, а лишь для частного случая, например, для трех переменных x, y, z. Требуется вычислить

$$(x+y+z)^n.$$

Это значит, что нам нужно найти произведение n одинаковых сомножителей, из которых каждый равен трехчлену $(x+y+z)$.

Как известно, при умножении скобки на некоторый множитель нужно каждое слагаемое скобки умножить на этот множитель. Применяя это правило, мы замечаем, что получающееся после всех перемножений выражение представляет собой сумму произведений, каждый член которой содержит n множителей, а именно по одному x, y или z из каждой скобки. Если теперь расположить множители в каждом произведении так, чтобы первым была некоторая степень x, вторым — некоторая степень y и, наконец, третьим — некоторая степень z (как известно, в произведении порядок множителей можно изменять), то каждое произведение будет иметь вид $x^a y^b z^c$, где a, b, c — целые числа, такие, что:

$$a+b+c=n, \quad a \geqslant 0, \quad b \geqslant 0, \quad c \geqslant 0. \tag{4}$$

Так как множитель x (точно так же y и z) может быть взят из любой из n скобок, то для данных a, b, c вообще получается несколько произведений $x^a y^b z^c$. Спрашивается, *сколько именно произведений $x^a y^b z^c$ с определенными a, b, c мы получим при вычислении такой степени?*

Каждая из n скобок должна дать в произведении $x^a y^b z^c$ только по одному множителю x, y или z из каждой скобки. Поставим в соответствие с каждой из этих n скобок «ящик», в который будем «бросать» указанные буквы. Тогда только что поставленную задачу можно будет сформулировать иначе: сколькими способами можно разместить a элементов «x» (красные шары), b элементов «y» (черные шары) и c элементов «z» (белые шары) в n ящиках, из которых каждый вмещает по одному элементу?

Искомое количество — обозначим его через $P^{(n)}_{a,b,c}$ — представляет собой, таким образом, число размещений и может быть записано символом

$$P^{(n)}_{a,b,c} = \{a, b, c \mid 1, 1, \ldots, 1\}_n. \tag{5}$$

Итак, при вычислении нашей степени получается по $P^{(n)}_{a,\,b,\,c}$ равных произведений $x^a y^b z^c$, и потому перед каждым произведением $x^a y^b z^c$ должен стоять числовой коэффициент $P^{(n)}_{a,\,b,\,c}$. Так как этот коэффициент получается при возведении в степень «полинома» $(x+y+z)$, то он называется **полиномиальным коэффициентом**.

Рассмотрим для наглядности несколько подробнее один частный случай, например случай $n=4$. При вычислении

$$(x+y+z)^4$$

появляются всевозможные члены вида $x^a y^b z^c$, соответствующие условиям (4) при $n=4$, а именно

$$\begin{array}{llllll} x^4, & y^4 & z^4, & & & \\ x^3y, & x^3z, & xy^3, & y^3z, & xz^3, & yz^3, \\ x^2y^2, & x^2z^2, & y^2z^2, & & & \\ x^2yz, & xy^2z, & xyz^2. & & & \end{array}$$

Умножая на соответствующие полиномиальные коэффициенты и складывая, получаем:

$$\begin{aligned}(x+y+z)^4 = {} & P^{(4)}_{4,\,0,\,0}x^4 + P^{(4)}_{0,\,4,\,0}y^4 + P^{(4)}_{0,\,0,\,4}z^4 + \\ & + P^{(4)}_{3,1,0}\,x^3y + P^{(4)}_{3,0,1}\,x^3z + P^{(4)}_{1,3,0}\,xy^3 + \\ & + P^{(4)}_{0,3,1}\,y^3z + P^{(4)}_{1,0,3}\,xz^3 + P^{(4)}_{0,1,3}\,yz^3 + \\ & + P^{(4)}_{2,2,0}\,x^2y^2 + P^{(4)}_{2,0,2}\,x^2z^2 + P^{(4)}_{0,2,2}\,y^2z^2 + \\ & + P^{(4)}_{2,1,1}\,x^2yz + P^{(4)}_{1,2,1}\,xy^2z + P^{(4)}_{1,1,2}\,xyz^2. \qquad (6)\end{aligned}$$

Но здесь, собственно, дано только строение формулы. Фактическое же вычисление $P^{(4)}_{a,\,b,\,c}$ и более общее — для $P^{(n)}_{a,\,b,\,c}$ на основе символического представления в виде (5) будет дано ниже.

6. В заключение запишем в принятой нами символической форме ответ на такой вопрос: *сколькими способами можно из n различных предметов выбрать k предметов*? Или, как принято говорить, каково *число сочетаний из n элементов по k*? Предполагается, что все n элементов различны (все шары различных цветов $к=1$, $ч=1$, $б=1$, ...). Выбранные шары в числе k мы помещаем в ящик, вмещающий как раз k шаров, остальные $n-k$ шаров откладываем (помещаем в другой ящик). Искомое число C_k^n можно записать так:

$$C_k^n = \{1,\ 1,\ \ldots,\ 1 | k,\ n-k\}_n. \qquad (7)$$

Двойственность «скобочного символа»

7. Введенный нами символ для обозначения числа размещений обладает одним важным свойством, которое окажет нам большую помощь при фактическом вычислении этого числа. А именно всегда имеет место соотношение:

$$\{к,\ ч,\ б,\ \ldots | a,\ b,\ c,\ \ldots\}_n = \{a,\ b,\ c,\ \ldots | к,\ ч,\ б, \ldots\}_n, \qquad (8)$$

т. е. числа, стоящие внутри скобок слева от вертикальной черты, можно переставить с числами, стоящими справа от черты. Если мы вспомним еще раз, что скобки означают число размещений, то в соотношении (8) обнаружим сопоставление двух комбинаторных задач: этим соотношением утверждается, что имеется одинаковое число возможностей как при размещении $к$ — красных, $ч$ — черных и т. д. шаров по ящикам вместимостью a, b и т. д., так и при размещении a — красных, b — черных и т. д. шаров по ящикам вместимостью $к$, $ч$ и т. д. При этом всегда предполагается, что

$$к + ч + \ldots = a + b + \ldots = n.$$

Таким образом, здесь проявляется «двойственность» каждой такой комбинаторной задачи. Утверждение (8) является, однако, почти тривиальным, если вспомнить, что в сущности означает наш символ. Достаточно доказать соотношение (8) на простом примере, так как основная идея этого доказательства обнаруживается уже здесь.

Покажем, например, что

$$\{3,\ 4 | 1,\ 1,\ 5\}_7 = \{1,\ 1,\ 5 | 3,\ 4\}_7. \qquad (9)$$

Обратим сначала внимание на левую часть равенства; из нее видно, что у нас имеется 3 красных и 4 черных шара:

К, К, К, Ч, Ч, Ч, Ч,

размещаемых в двух ящиках A и B вместимостью по 1 шару и ящике C вместимостью в 5 шаров. Одно из подобных размещений можно осуществить таким образом:

$$\underset{A}{|\text{К}|} \quad \underset{B}{|\text{Ч}|} \qquad \underset{C}{|\text{ККЧЧЧ}|}.$$

Но это же самое размещение можно записать и иначе, а именно под каждым шаром указать во второй строке ящик, в котором этот шар помещается

$$\begin{matrix} \text{К} & \text{Ч} & \text{К} & \text{К} & \text{Ч} & \text{Ч} & \text{Ч} \\ A & B & C & C & C & C & \end{matrix}. \tag{10а}$$

Мы приходим, таким образом, к 7 *парам* букв: верхние К и Ч указывают шары, нижние A, B, C — ящики. Поменяем теперь в каждой из этих пар роли партнеров и будем рассматривать A, B, C как обозначения шаров, К и Ч — как обозначения ящиков. Как и прежде, буквы, обозначающие ящики, будем писать в нижнем ряду, а буквы, обозначающие шары, — в верхнем:

$$\begin{matrix} A & C & C & B & C & C & C \\ \text{К} & \text{К} & \text{К} & \text{Ч} & \text{Ч} & \text{Ч} & \text{Ч} \end{matrix} \tag{10б}$$

В (10б) стоят те же самые пары, что и в (10а), только верхние знаки поменялись местами с нижними. Схему (10б) мы должны понимать как размещение одного шара цвета A, одного шара цвета B и пяти шаров цвета C в ящиках К и Ч вместимостью соответственно в 3 и 4 шара. Но это и есть как раз одно из тех размещений, которые мы записываем символом, стоящим в правой части соотношения (9).

Совершенно так же, как мы это установили сейчас для размещений (10а) и (10б), можно установить подобное двойственное соответствие и для каждого из рассматриваемых размещений. При этом одно из этих размещений будет учитываться в левой части равенства (9), другое — в правой. Мы получим, таким образом, два типа (множества) размещений, таких, что если одно из сопоставленных размещений принадлежит к одному типу, то другое — ко второму типу, и наоборот. Но два множества, состоящих из конечного числа элементов, между которыми установлено взаимно однозначное соответствие, имеют одинаковое число элементов. Справедливость равенства (9) тем самым доказана.

При доказательстве соотношения (9) не было надобности в установлении числовой величины исследуемого нами символа размещений. Но путем систематического обзора всех мыслимых размещений легко найти, что

$$\{3,\ 4\,|\,1,\ 1,\ 5\}_7 = \{1,\ 1,\ 5\,|\,3,\ 4\}_7 = 4.$$

Существенным в доказательстве равенства (9) была идея о перестановочности обозначений цветов и ящиков. Но эта же идея без всяких изменений может быть применена и в доказательстве общего равенства (8), которое на этом основании мы будем считать доказанным.

Определение числа размещений в простейших случаях

8. Определение числа размещений при произвольно заданных количествах шаров $к$, $ч$, ... и вместимостях ящиков a, b, ... чрезвычайно кропотливо, и поэтому заниматься им мы здесь не будем. Но в приведенных выше примерах (пп. 3—6) скобочный символ нигде ведь и не встречался во всей его общности; для всех этих примеров характерным было то, что внутри скобок — или справа, или слева от вертикальной черты, или, наконец, и справа, и слева — стояли только единицы, т. е. или все шары были разноцветны, или все ящики вмещали лишь по одному шару, или имели место одновременно оба эти условия. На основании формулы двойственности здесь достаточно рассмотреть лишь один случай, например тот, в котором все шары разноцветны. Итак, мы хотим вычислить

$$\{1,\ 1,\ \ldots | a,\ b,\ c,\ \ldots\}_n,$$

где индекс n означает, как и раньше, что имеется всего n шаров, общая же вместимость всех ящиков также равна n:

$$1+1+\ldots+1=a+b+c+\ldots=n.$$

Так как в дальнейшем слева от вертикальной черты у нас всегда будут стоять единицы, мы позволим себе несколько упростить обозначение нашего символа и записывать его сокращенным способом так:

$$\{1,\ 1,\ \ldots,\ 1\,|\,a,\ b,\ c\ \ldots\}_n=\{a,\ b,\ c,\ \ldots\}_n;$$

никаких предположений о числе ящиков заранее не делается.

Совершенно очевидно, что

$$\{n\}_n=1, \tag{11}$$

ибо в этом случае возможно только одно размещение, при котором все шары попадают в один ящик. Если мы теперь вместо одного ящика N вместимостью в n шаров возьмем два ящика: N_1 вместимостью в $n-1$ и N' вместимостью в 1 шар, то переход к новой задаче можно осуществить, положив в ящик N' любой из n шаров, лежавших в N, а остальные $n-1$ шаров поместив в ящик N_1. Но имеется n возможностей выбрать один шар из ящика N, поэтому

$$\{n-1,\ 1\}_n = n.$$

Точно так же, очевидно, получаем и соотношение

$$a \cdot \{a,\ b,\ c,\ \ldots\}_n = \{a-1,\ 1,\ b,\ c,\ \ldots\}_n. \tag{12}$$

В этом соотношении оба символа опять относятся к n разноцветным шарам. Ящики B, C, ... вместимостью соответственно b, c, ... шаров в обоих случаях тождественны. Только ящик A вместимостью a, имеющийся в левой части, заменен в размещениях правой части двумя ящиками: A_1 вместимостью $a-1$ и A' вместимостью 1. Так как из каждого размещения по ящикам A, B, C, ... можно получить размещение по ящикам A_1, A', B, C, ... путем перемещения любого из a шаров из A в A', что может быть проделано a различными способами (остальные $a-1$ шаров целиком переходят в A_1), то, значит, размещений по ящикам A_1, A', B, C в a раз больше, чем размещений по ящикам A, B, C, ... Это обстоятельство и выражено соотношением (12).

Ящик A_1 можно, конечно, в свою очередь заменить двумя ящиками A_2 и A'' вместимостью соответственно в $a-2$ шара и 1; для этого случая мы получим:

$$(a-1) \cdot \{a-1,\ 1,\ b,\ c,\ \ldots\}_n = \{a-2,\ 1,\ 1,\ b,\ c,\ \ldots\}_n.$$

Подобным же образом мы приходим и к соотношению

$$(a-2) \cdot \{a-2,\ 1,\ 1,\ b,\ c, \ldots\}_n = \{a-3,\ 1,\ 1,\ 1,\ b,\ c, \ldots\}_n,$$

и к целому ряду следующих аналогичных соотношений, последнее из которых будет иметь вид:

$$2 \cdot \{2,\ \underbrace{1,\ 1,\ \ldots,\ 1}_{a-2},\ b,\ c,\ \ldots\}_n = \{\underbrace{1,\ 1,\ \ldots\ 1}_{a},\ b,\ c,\ \ldots\}_n.$$

В правой части этого соотношения разложение a на слагаемые единицы доведено до конца.

Перемножим (12) со всеми последующими, вытекающими из него соотношениями; тогда в обеих частях полученного равенства мы будем иметь общие множители:

$$\{a-1,\ 1,\ b,\ c,\ \ldots\}_n,\ \{a-2,\ 1,\ 1,\ b,\ c,\ \ldots\}_n,\ \ldots$$
$$\ldots,\ \{2,\ \underbrace{1,\ 1,\ \ldots,\ 1}_{a-2},\ b,\ c,\ \ldots\}_n,$$

которые можно отбросить, не нарушив равенства.

Получим:

$$a(a-1)(a-2)\ldots 2\cdot\{a, b, c, \ldots\}_n=\{\underbrace{1, 1, \ldots, 1}_{a}, b, c, \ldots\}_n.$$

В произведении чисел $a(a-1)(a-2)\ \ldots\ 2$ переставим порядок множителей и, воспользовавшись сокращенным обозначением, напишем [1]):

$$a! = 1\cdot 2\cdot 3\ldots(a-1)\,a.$$

Тогда мы будем иметь:

$$a!\,\{a,\ b,\ c,\ \ldots\}_n=\{\underbrace{1,\ 1,\ \ldots,\ 1}_{a},\ b,\ c,\ \ldots\}_n. \tag{13}$$

Тот же самый процесс последовательной замены ящика A все менее и менее емкими может быть произведен и над ящиком B. На этом основании мы можем написать:

$$b!\ \{\underbrace{1,\ 1,\ \ldots,\ 1,}_{a}\ b,\ c,\ \ldots\}=\{\underbrace{1,\ \ldots,\ 1}_{a},\ \underbrace{1,\ \ldots,\ 1}_{b},\ c,\ldots\}_n$$

или, принимая во внимание (13),

$$a!b!\,\{a,\ b,\ c,\ \ldots\}_n=\{\underbrace{1,\ \ldots,\ 1}_{a+b},\ c,\ \ldots\}_n,$$

или, наконец,

$$a!b!c!\ \ldots\ \{a,\ b,\ c,\ \ldots\}_n=\{1,\ 1,\ \ldots,\ 1\}_n, \tag{14}$$

где количество единиц в скобках правой части равно n.

Впрочем, если $a=n$ уже в формуле (13), т. е. если у нас имеется только один ящик A, тогда

$$n!\,\{n\}_n=\{1,\ 1,\ \ldots,\ 1\}_n$$

или, принимая во внимание (11),

$$n!=\{1,\ 1,\ \ldots,\ 1\}_n. \tag{15}$$

---

[1]) $a!$ читается: «a факториал».

При этом мы получаем из (14):

$$\{a, b, c, \ldots\}_n = \{1, 1, \ldots, 1 \mid a, b, c, \ldots\}_n = \frac{n!}{a!b!c!\ldots}. \quad (16)$$

Здесь мы записываем символ размещений без сокращения, в его первоначальном виде.

Этим в сущности и исчерпывается задача определения числового значения символа размещений в той мере, в какой это необходимо для наших примеров.

9. Просмотрев еще раз наши примеры, мы получаем такие частные результаты:

1) Из уравнений (2) и (15) имеем:

$$P_n = \{1, 1, \ldots, 1 \mid 1, 1, \ldots, 1\}_n = \{1, 1, \ldots,\}_n = n!,$$

т. е. n лиц можно разместить по n местам $n!$ различными способами или, иными словами: существует $n!$ перестановок из n различных элементов; $n!$, как видно из приводимой таблички, растет по мере увеличения n очень быстро:

$1! =$	1	$6! =$	720
$2! =$	2	$7! =$	$5\,040$
$3! =$	6	$8! =$	$40\,320$
$4! =$	24	$9! =$	$362\,880$
$5! =$	120	$10! =$	$3\,628\,800$

2) На основании равенств (3) и (16) число возможных размещений при игре в скат равно

$$S = \frac{32!}{10!\,10!\,10!\,2!}.$$

Это очень большое число. Вычисление дает:

$$S = 2\,753\ 294\ 408\ 504\ 640.$$

3) Общий вид полиномиального коэффициента (5), на основании соотношения (16) и свойства двойственности (8), таков:

$$P^{(n)}_{a,b,c} = \frac{n!}{a!\,b!\,c!}\ (a + b + c = n).$$

В частности, в (6) для $n=4$ вычисление полиномиальных коэффициентов дает формулу

$$(x+y+z)^4 = x^4 + y^4 + z^4 + 4x^3y + 4x^3z + 4xy^3 + 4y^3z + \\ + 4xz^3 + 4yz^3 + 6x^2y^2 + 6x^2z^2 + 6y^2z^2 + 12x^2yz + \\ + 12xy^2z + 12xyz^2.$$

4) Из соотношений (7) и (16) число сочетаний из n элементов по k оказывается равным

$$C_k^{(n)} = \frac{n!}{k!\,(n-k)!}.$$

Это и есть то число способов, которыми можно выбрать k предметов из совокупности в n предметов [27].

10. ПРОБЛЕМА ВАРИНГА

Ряд квадратов 1, 4, 9, 16, ... при неограниченном продолжении его становится все более и более разреженным: промежутки между последовательными членами этого ряда постепенно возрастают. В этих промежутках встречаются, однако, такие числа, которые можно рассматривать как суммы двух квадратов, например: $13 = 9 + 4$; $41 = 25 + 16$ и т. д. Но не всякое число может быть представлено в виде суммы двух квадратов; например, если нам дано число 6, то квадратами, которые в сумме дают 6, могли бы быть лишь 1 и 4 — единственные квадраты, меньшие 6. Однако ни $1 + 1$, ни $4 + 4$, ни $1 + 4$ не дают 6. Чтобы получить в сумме 6, нужно сложить три квадрата: $6 = 4 + 1 + 1$. Применение этого способа разложения к числу 7 не приводит к результату даже при использовании трех квадратов; здесь требуются четыре квадрата: $7 = 4 + 1 + 1 + 1$. Для 8 достаточно опять только двух квадратов: $8 = 4 + 4$; число 9 само является квадратом; $10 = 9 + 1$; $11 = 9 + 1 + 1$; $12 = 9 + 1 + 1 + 1 = 4 + 4 + 4$, и т. д.

Эти первые наблюдения дают, казалось бы, основание ожидать, что очень скоро должен наступить момент, когда не будет хватать даже и четырех квадратов, и что по мере дальнейшего движения вперед будет необходимо все большее и большее количество квадратов. Тем более удивительным представляется результат, полученный Ферма (Fermat), одним из величайших, наряду с Декартом, математиков XVII века. Ферма показал, что *любое целое положительное число можно представить в виде суммы самое большее четырех квадратов* [28]. Английский математик XVIII века Э. Варинг (E. Waring) поставил себе задачу найти аналогичные теоремы для кубов, биквадратов и более высоких степеней; отсюда эта проблема и носит название проблемы Варинга.

Кубами являются следующие числа: 1, 8, 27, 64, ... Если попробовать представить некоторые небольшие числа в виде сумм возможно меньшего количества кубов, то окажется, что, например, для 7 требуется семь кубов ($7=1+1+1+1+1+1+1$), для 15 — 8 кубов ($15=8+1+1+1+1+1+1+1$), а для 23 — даже 9 кубов ($23=8+8+1+1+1+1+1+1+1$). Однако прежде чем мы достигнем числа 31, мы встретим третий куб, именно 27, который изменит всю картину: $31=27+1+1+1+1$ есть сумма только пяти кубов. Знаменитый немецкий математик К. Г. Якоби (K. G. Jacobi), желая хотя бы косвенным образом использовать в интересах развития математики вычислителя Дазе (Dahse), поручил ему дальнейшее рассмотрение натурального ряда чисел с точки зрения разложимости его членов на суммы возможно меньшего числа кубов. Это эмпирическое исследование выяснило, что, кроме числа 23, на всем протяжении натурального ряда чисел от 1 до 12 000 конечного пункта вычислений Дазе девяти кубов требует еще только одно число, а именно 239; восьми кубов требуют числа: 15, 22, 50, 114, 167, 175, 186, 212, 231, 238, 303, 364, 420, 428, 454; семи кубов — числа 7, 14, 21, 42, 47, 49, 61, 77, 85, 87, 103, ..., 5306, 5818, 8042; так что, по-видимому, и этот более богатый ряд также постепенно иссякает; дальнейшие эмпирические исследования подтвердили этот факт.

Однако подобного рода эмпирические исследования никогда не смогут доказать ни того, что любое число можно представить в виде суммы, самое бóльшее, 9 кубов, ни того, что любое число, начиная с некоторого места натурального ряда, можно разложить на сумму, самое бóльшее, 8 или даже 7 кубов. Первое из этих двух утверждений было доказано Виферихом (Wieferich) — в то время еще совсем юным математиком; что же касается второго вопроса, то крупнейший немецкий теоретико-числовик Э. Ландау сложными вспомогательными средствами доказал, что, начиная с некоторого определенного числа, можно обойтись восемью кубами.

Для биквадратов (четвертых степеней чисел) мы должны провести подобные же рассуждения. Первыми биквадратами являются числа 1, 16, 81, 256, ... Число 15 является, очевидно, суммой 15 биквадратов: 31 — суммой 16 биквадратов; 47 — 17; 63 — 18; 79 — 19 биквадратов; далее мы встречаем число 81, и картина резко меняется. Здесь возник такой вопрос —

всегда ли, т. е. на всем ли протяжении натурального ряда, можно обойтись 19 биквадратами. Медленно и постепенно приближались математики к решению этого вопроса. Лиувилль (Liouville) доказал, что для разложения любого числа на сумму биквадратов достаточно, самое бо́льшее, 53 биквадратов; постепенно это количество снижалось до 47, 45, 41, 39, 38. Виферих побил рекорд, снизив это число до 37; но до эмпирически предположенного 19 было ещё далеко.

Гильберт в одной из своих знаменитых работ подошел ко всему этому комплексу вопросов с более широкой точки зрения; он не побил ни одного из ранее установленных рекордов, да и не стремился к этому. Он доказал весьма общее предложение о том, что *при разложении любого числа на суммы одинаковых степеней всегда существует некоторое предельное количество слагаемых (зависящее лишь от показателя степени, но не от разлагаемого числа);* выше мы высказали это предложение для квадратов (где предельное количество слагаемых равно 4), для кубов (где это количество равно 9) и для биквадратов (где оно равно 37). Естественно, что число, показывающее, насколько степеней с данным показателем степени можно разложить всякое целое число, увеличивается при увеличении показателя степени [[29]].

Английские математики Гарди (Hardy) и Литтлвуд (Littlewood) подошли к проблеме с совершенно иными методами, чем Гильберт. Читатель получит некоторое представление о необычайной силе этих методов, если узнает, что им удалось показать — и это всего лишь один из полученных названными авторами результатов,— что *всякое число натурального ряда, начиная с определенного места, всегда может быть представлено в виде суммы, самое бо́льшее* 19 *биквадратов.*

Выше мы видели, что числа, требующие для своего разложения 19 биквадратов, встречаются уже в самом начале натурального ряда. Согласно Гарди и Литтлвуду, существует некоторое определенное число N, начиная с которого все числа могут быть представлены как суммы, самое бо́льшее, 19 биквадратов (число N, впрочем, настолько велико, что Гарди и Литтлвуд даже и не брали на себя труда фактически вычислить его). Нужно было бы еще только пересмотреть, все ли числа, не превосходящие N, допускают такое разложение, чтобы показать, что все числа вообще могут быть представлены в виде суммы, самое бо́льшее, 19 биквадратов, и тем самым достигнуть того, что уже сделано по отношению к квадратам и кубам. Впрочем, пока теории не удастся заменить число Гарди N значительно меньшим, такой пересмотр превосходит силы всякого вычислителя [30].

Мы так долго задержались на фактах для того, чтобы дать понятие о том, как по мере накопления конкретного материала, своего рода математических эмпирических результатов, получают оформление математические проблемы.

Попытаемся теперь дать некоторое представление о методах, нашедших применение в этой области и оказавших услугу даже Гильберту в его замечательном доказательстве. Мощный математический аппарат, столь успешно использованный Гарди и Литтлвудом, выходит далеко за рамки настоящей книги и потому не может быть здесь освещен.

Рассмотрим, как мы это всегда делали, сначала более простой случай. Вы, может быть, еще помните заученную когда-то наизусть формулу

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

А если вы ее и забыли, то посредством раскрытия скобок и простого вычисления нетрудно убедиться в том, что она правильна, каковы бы ни были значения чисел a и b. Подобные равенства, справедливые при любых значениях входящих в них букв, математики называют тождествами.

Приведем другое, несколько более сложное тождество:

$$(a^2+b^2)(c^2+d^2)=(ac+bd)^2+(ad-bc)^2, \qquad (1)$$

справедливость которого станет ясной, если вспомнить формулу

$$(x+y)^2=x^2+2xy+y^2$$

и, воспользовавшись ею, произвести указанные в правой части нашего тождества (1) возведения в квадрат:

$$(a^2c^2+2acbd+b^2d^2)+(a^2d^2-2adbc+b^2c^2)=$$
$$=a^2c^2+a^2d^2+b^2c^2+b^2d^2+2abcd-2abcd.$$

Два последних члена здесь взаимно уничтожаются, а остальные можно записать в такой форме:

$$a^2(c^2+d^2)+b^2(c^2+d^2)=(a^2+b^2)(c^2+d^2),$$

— а это и есть как раз левая часть соотношения (1).

Из этого тождества можно извлечь следующее небезынтересное следствие:

Если мы имеем два числа, каждое из которых может быть представлено в виде суммы двух квадратов, то этим же свойством обладает и произведение этих чисел. Например, 13 и 41 — числа как раз такого рода: $13 = 9 + 4$; $41 = 25 + 16$. Пользуясь формулой (1), получаем:

$$533 = 13 \cdot 41 = (3^2 + 2^2)(5^2 + 4^2) = (3 \cdot 5 + 2 \cdot 4)^2 + \\ + (3 \cdot 4 - 2 \cdot 5)^2 = 23^2 + 2^2,$$

и тем самым число 533 действительно представлено в виде суммы двух квадратов; точно таким же образом, пользуясь формулой (1), мы можем поступить и с произведением любых других чисел подобного рода [31].

Проделаем еще одно упражнение. Заметим прежде всего следующую формулу для возведения четырехчлена в квадрат:

$$(x_1 + x_2 + x_3 + x_4)^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2x_1x_2 + 2x_1x_3 + \\ + 2x_2x_3 + 2x_1x_4 + 2x_2x_4 + 2x_3x_4.$$

Великий математик XVIII века Л. Эйлер (L. Euler) нашел следующее тождество:

$$(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) = \\ = (-a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4)^2 + \\ + (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 + \\ + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2, \quad (2)$$

в правильности которого можно без особого труда убедиться, если перемножить скобки и принять во внимание вышеприведенную формулу для четырехчлена. По аналогии с (1) формула (2) показывает, что *произведение двух чисел, каждое из которых можно представить в виде суммы четырех квадратов, само может быть представлено в виде суммы четырех квадратов.* Лагранж (Lagrange) воспользовался этим замечанием для чрезвычайно изящного доказательства теоремы о том, что **любое** число можно представить в виде суммы четырех квадратов. А именно из приведенного замечания прежде всего следует, что стоит доказать справедливость этого утверждения для всякого **простого** числа, как отсюда будет следовать справедливость его и для всех составных чисел.

Но тождество (2) явилось основой и для дальнейшего хода доказательства Лагранжа. Доказательство это, к сожалению, слишком длинно и несколько искусственно, и мы не можем его привести и сделать действительно ясным в узких рамках настоящей книги. Мы вынуждены здесь поэтому ограничиться лишь только что намеченными общими и приблизительными указаниями о характере этого доказательства. В дальнейшем мы будем считать эту теорему установленным фактом, как если бы мы уже выполнили ее доказательство [32].

Опираясь на эту теорему, согласно которой всякое целое положительное число равно сумме четырех квадратов, Лиувилль, видный французский математик первой половины XIX века, доказал, что любое число можно рассматривать как сумму 53 биквадратов [33]. Он исходил при этом также из тождества, а именно из следующего:

$$\begin{aligned}6(x_1^2+x_2^2+x_3^2+x_4^2)^2 &= (x_1+x_2)^4+(x_1+x_3)^4+\\ &+(x_2+x_3)^4+(x_1+x_4)^4+(x_2+x_4)^4+(x_3+x_4)^4+\\ &+(x_1-x_2)^4+(x_1-x_3)^4+(x_2-x_3)^4+(x_1-x_4)^4+\\ &+(x_2-x_4)^4+(x_3-x_4)^4. \qquad (3)\end{aligned}$$

Чтобы установить справедливость этого тождества, нужно при умножениях в правой части принять во внимание, что

$$\left.\begin{array}{l}(x_1+x_2)^4+\\ +(x_1-x_2)^4\end{array}\right\}=\left\{\begin{array}{r}x_1^4+4x_1^3x_2+6x_1^2x_2^2+4x_1x_2^3+x_2^4+\\ +x_1^4-4x_1^3x_2+6x_1^2x_2^2-4x_1x_2^3+x_2^4=\\ =2x_1^4+2x_2^4+12x_1^2x_2^2\end{array}\right.$$

и что подобное же соотношение имеет место для любой другой пары биквадратов, содержащих x с одинаковыми значками. Правая часть тождества (3) после всех этих преобразований будет равна:

$$\begin{aligned}&6(x_1^4+x_2^4+x_3^4+x_4^4)+\\ &\quad+12(x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2+x_1^2x_4^2+x_2^2x_4^2+x_3^2x_4^2),\end{aligned}$$

а левая примет в точности этот же самый вид, если для нее применить вышеприведенную формулу возведения в квадрат четырехчлена. Лиувилль пользуется этим тождеством следующим образом. Пусть n — какое-нибудь целое положительное число; нужно показать, что оно есть сумма, самое бóльшее, 53 биквадратов. Лиувилль начинает с того, что выделяет из данного числа наибольшую часть, кратную числу 6, т. е. представляет данное число в виде $n=6x+y$, где y — остаток, равный одному из чисел 0, 1, 2, 3, 4, 5. Например, число 28 он записывает как $6\cdot4+4$. Затем он обращается к теореме Ферма, применяя ее к числу x; поскольку любое число можно представить в виде суммы, самое бо́льшее, 4 квадратов, то это имеет место также и в отношении x; пусть, например,

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

тогда

$$n=6x+y=6(a^2+b^2+c^2+d^2)+y=\\=6a^2+6b^2+6c^2+6d^2+y.$$

К каждому из чисел a, b, c, d Лиувилль снова применяет теорему Ферма, полагая, например, что

$$a=a_1^2+a_2^2+a_3^2+a_4^2,$$
$$b=b_1^2+b_2^2+b_3^2+b_4^2,$$
$$c=c_1^2+c_2^2+c_3^2+c_4^2,$$
$$d=d_1^2+d_2^2+d_3^2+d_4^2;$$

при этом

$$n=6(a_1^2+a_2^2+a_3^2+a_4^2)^2+\ldots+6(d_1^2+d_2^2+d_3^2+d_4^2)^2+y.$$

Теперь, наконец, применяется тождество (3). Это тождество показывает, что первое слагаемое правой части можно представить в виде суммы 12 биквадратов; то же самое, очевидно, справедливо и по отношению к трем следующим слагаемым правой части; всего пока имеем $4\cdot12=48$ биквадратов; к ним можно еще прибавить столько единиц, сколько их содержится в числе y (бо́льшие биквадраты не входят в число y); этих единиц не больше 5, так как остаток y есть одно из чисел 0, 1, 2, 3, 4, 5. Всего, таким образом, как это и утверждалось, мы имеем, самое бо́льшее, 53 биквадрата [34].

11. О ЗАМКНУТЫХ САМОПЕРЕСЕКАЮЩИХСЯ КРИВЫХ

1. Проведем, начиная из произвольной точки, какую-либо кривую, пересекающую несколько раз самое себя и возвращающуюся в исходную точку. Мы получим замкнутую кривую, вычерчиваемую одним росчерком. Положим, что через точки, в которых кривая пересекает самое себя, она проходит не больше двух раз; такие точки мы будем называть **двойными** точками[1]). Мы будем, следовательно, рассматривать кривые, подобные изображенным на рис. 34 и 35, и исключим из нашего исследования кривые, подобные изображенной на рис. 36 [[35]]. При полном обходе нашей кривой каждая ее двойная точка проходится ровно два раза. Если двойные точки перенумеровать, то та последовательность, в которой они проходятся при обходе кривой, может быть записана с помощью

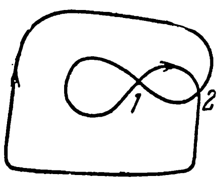


Рис. 34.

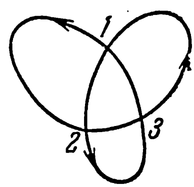


Рис. 35.

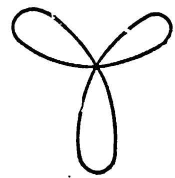


Рис. 36.

соответствующей последовательности чисел — номеров. Например, для кривых, изображенных на рис. 34 и 35, мы получим соответственно следующие последовательности номеров: *1221* и *123123*. Само собой очевидно, что при этом способе обозначения каждый номер, отвечающий той или иной двойной точке, должен встречаться два раза. Гаусс, однако, заметил,

[1]) Хотя факты, сообщаемые в этой теме, имеют много точек соприкосновения с содержанием темы второй, читатель все же хорошо сделает, если будет воспринимать эти факты независимо от изложенного в теме второй, так как задача, исследуемая здесь, имеет совершенно иной смысл. Там была дана сеть линий и ставился вопрос о различных возможностях обхода этой сети. Здесь, напротив, обход в направлении единственного кольцевого маршрута предписан уже заранее. Кроме того, в рассматриваемых нами здесь кривых изучаются только такие точки, которые в смысле темы второй являются точками пересечения четвертого порядка, а здесь названы нами двойными точками. Наконец, двойная точка проходится здесь непременно таким образом, чтобы из четырех сходящихся в точке направлений взаимно противоположные были попарно между собой связаны, ибо двойная точка может быть определена как точка самопересечения; самопересечение же происходит только при указанном способе обхода.

что не всякая произвольная последовательность номеров с двукратным повторением каждого номера может выражать расположение двойных точек на кривой. Для пары двойных точек мы пришли выше к последовательности *1221*, но не существует такой замкнутой кривой с двумя двойными точками, которой соответствовала бы последовательность *1212*, в чем весьма легко убедиться путем проб.

Вообще же имеет место следующая теорема: *двойная точка должна приходиться один раз на место с четным номером, другой раз — на место с нечетным номером*, или, что сводится к тому же: между двумя ее местами или вовсе не встречается других двойных точек, или таковые встретятся четное число раз.

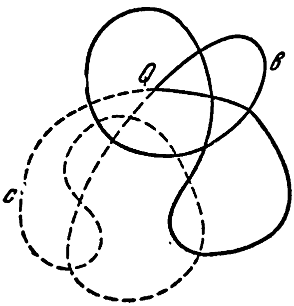


Рис. 37.

Эта теорема исключает возможность последовательности *1212*, в которой между двумя *1* имеется только одно промежуточное место [36].

2. Чтобы доказать эту теорему, возьмем произвольную двойную точку Q нашей кривой a (полная петля, рис. 37). Если, отправившись из Q, мы будем двигаться вдоль кривой a, то мы непременно вернемся опять в Q. Мы пройдем при этом во всяком случае не всю кривую a, а лишь одну часть b (сплошная линия на рис. 37) всей кривой a, ибо из четырех выходящих из Q путей мы используем только два (при выходе из точки Q и при возвращении в нее). Эту часть b мы можем рассматривать как замкнутую в Q кривую, вычерчиваемую одним росчерком (то, что при этом в Q получается угол, никакого значения не имеет, так как углы и без того могут быть на кривой). Продолжая от точки Q движение по кривой a дальше, мы пройдем при этом другую часть c кривой a. Кривая c (пунктирная) сама по себе есть также замкнутая в Q кривая, образующая в этой точке угол. В точке Q обе эти кривые, (сплошная) b и (пунктирная) c, не пересекаются, а только встречаются. Поэтому точка Q, будучи двойной для кривой a, не является таковой ни для b, ни для c. Требуется доказать, что, пробегая b, начиная от точки Q и в эту же точку возвращаясь, мы встретим двойные точки четное число раз[1]). Двойными точками a, лежащими на b, могут быть кроме Q, во-первых, точки пересечения b с самой собой, т. е. двойные точки b, и, во-вторых, точки пересечения b с c.

Каждая двойная точка b при полном обходе b должна быть пройдена ровно два раза; эти двойные точки дадут, следовательно, все вместе четное число переходов через двойные точки. Каждая точка пересечения b с c проходится при полном обходе b только один раз, так как через такую точку проходит только один принадлежащий b отрезок пути (пересекающий его в этой точке путь принадлежит c и потому не может проходиться при обходе b).

Остается, значит, еще доказать, что кривые b и c пересекаются в четном числе точек. Тогда тем самым будет доказано, что при полном обходе b от Q до Q двойные точки a, т. е. отчасти двойные точки b, отчасти же точки пересечения b с c, проходятся четное число раз. Конечно, сама точка Q при этом подсчете во внимание не принимается.

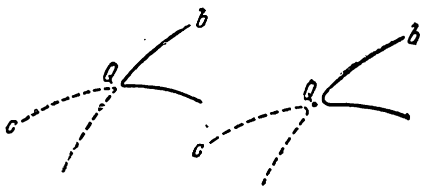


Рис. 38.

3. Если каждую из кривых b и c мы несколько закруглим (рис. 38) около точки Q, то число точек пересечения этих кривых от этого не изменится. Мы будем тогда иметь дело с двумя замкнутыми кривыми b и c, пересекающими друг друга, но нигде не соприкасающимися. Нам нужно показать, что две такие кривые или вовсе не пересекаются, или пересекаются в четном числе точек. С этой целью преобразуем одну из них, например c, однако так, *чтобы не изменить при этом числа точек пересечения ее с другой кривой*, в четности которого мы и должны как раз убедиться. Таким образом, мы сможем устранить все двойные точки c и тем самым получить более ясное представление о взаимном положении кривых b и c. Пусть P есть двойная точка c (рис. 39, a). Кривая c распадается в точке P на две замкнутые кривые d

[1]) Это, конечно, не значит, что встречается четное число двойных точек. На схеме рис. 34 при расположении двойных точек, задаваемом последовательностью *1221*, между *1* и *1* встречается только одна двойная точка *2*, но четное число раз.

и *e* точно так же, как раньше кривая *a* распадалась у нас в точке *Q* на кривые *b* и *c*. Кривые *d* и *e* встречаются в *P*. Припишем всей кривой *c* некоторое определенное направление обхода; тем самым кривым *d* и *e* будут также сообщены определенные направления обхода (рис. 39, *a*). Будем теперь двигаться, начиная от *P*, в установленном направлении по кривой *d*; вернувшись в *P*, присоединим сюда еще обход по *e*, но в направлении, противоположном установленному; в итоге мы придем опять в *P*. Путем изменения направления

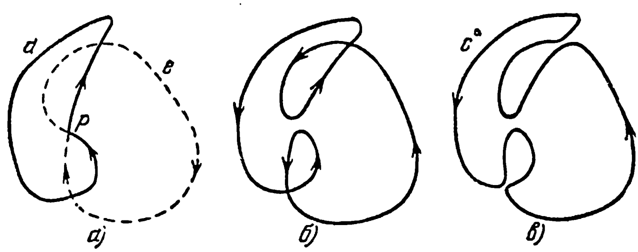


Рис. 39.

обхода по *e* мы избегли скрещения путей в *P*. В результате мы и на этот раз прошли весь путь *c* с одного раза, точку же *P* прошли на двух путях, образующих в *P* углы, но больше в *P* не пересекающихся, а лишь встречающихся.

Раздвинем немного эти пути в точке *P* и закруглим углы; тогда вместо *c* мы получим другую, также замкнутую кривую, которая уже не будет пересекать себя в точке *P* и, следовательно, будет иметь одной двойной точкой меньше, чем первоначальная кривая *c* (рис. 39, *б*). При этом мы предполагаем, что закругления углов около точки *P* выполнены, конечно, настолько близко к *P*, что это не оказывает никакого влияния на другие точки пересечения *c* с самой собой или с какими бы то ни было другими кривыми.

Повторяя этот прием, мы можем таким образом устранить из *c* все двойные точки одну за другой, пока не получим замкнутую кривую c^* (рис. 39, *в*), *свободную от двойных точек* и совпадающую с *c* почти всюду, за исключением мест, близких к двойным точкам *c*, где она несколько расходится с этой кривой.

4. Замкнутая кривая без двойных точек всегда ограничивает, как в этом легко убедиться наглядно, некоторую область, которую мы называем в н у т р е н н е й относительно этой кривой. Остальная часть плоскости, не принадлежащая ни самой кривой, ни внутренней ее области, называется в н е ш н е й относительно кривой. Внутренняя и внешняя области разделяются самой замкнутой кривой. Если замкнутая кривая без двойных точек пересекается в одной точке другой кривой, то эта вторая кривая должна переходить или из внешней области во внутреннюю, или наоборот [37].

Воспользуемся этим, чтобы доказать высказанное в начале п. **3** утверждение, что число точек пересечения кривых b и c четное. Представим себе, что кривая c, как только что было разъяснено, заменена свободной от двойных точек кривой c^*, пересекающейся с кривой b в тех же самых точках, что и кривая c. Может случиться, что b целиком лежит внутри c^*, тогда c^* и b вовсе не будут пересекаться. То же самое будет иметь место и в том случае, если b будет лежать вне c^*. Остается, следовательно, еще тот случай, когда b лежит частью внутри c^*, частью вне ее. Пусть T будет какой-либо точкой b, лежащей вне c^*. Будем теперь двигаться по b из T; где-нибудь кривая b должна впервые вступить внутрь c^*; при этом мы получаем одну точку пересечения. Поскольку кривая c замкнутая и, следовательно, возвращается в T, она должна выйти из внутренней области c^* и вернуться опять во внешнюю, что дает вторую точку пересечения. Независимо от того, сколько раз в дальнейшем b входит внутрь c^* или выходит оттуда, каждый вход во внутреннюю часть области сопряжен непременно с выходом из нее, поскольку мы наш обход по b начали с точки T, лежащей вне c^*, и в конце концов в эту же точку возвращаемся. Следовательно, кривая c^* или вовсе не пересекается кривой b, или пересекается ею в четном числе точек. Согласно вышесказанному, то же самое верно и для кривых c и b. А этим, как мы уже указали в конце п. **2**, доказательство теоремы о распределении двойных точек полностью завершается.

5. Эту теорему, утверждающую, что при обходе замкнутой кривой a каждая двойная точка должна встретиться один раз под четным номером, а другой раз под нечетным, можно истолковать и несколько иначе. Будем рассматривать кривую a как проекцию некоторого п р о с т р а н с т в е н н о г о образа, а каждую двойную точку — как проекцию с к р е щ е-ния двух кусков кривой, из которых один расположен над другим.

При обходе кривой каждое такое скрещение проходится попеременно: один раз сверху, другой раз снизу. Вопрос как раз и заключается в том, всегда ли это положение имеет место. Ведь если, например, мы прошли наше скрещение сверху, то во второй раз для прохождения этого скрещения нам остается лишь один путь — снизу. С другой стороны, после того как мы сделали сначала выбор направления в одной из двойных точек, чередование прохождения сверху и снизу становится определенным в каждом таком скрещении. Не столкнемся ли мы при возвращении к какому-нибудь скрещению Q с тем противоречием, что, отправляясь из Q сверху, мы после попеременного прохода сверху и снизу промежуточных двойных точек попадем в Q опять сверху? Нет, наша теорема как раз и утверждает, что такого противоречия не может случиться. Действительно, отправившись из Q сверху и проходя попеременно снизу и сверху все скрещивающиеся отрезки кривой, мы должны будем попасть в Q снизу, ибо по пути мы должны пройти двойные точки четное число раз (согласно доказанной теореме), именно: в первую точку после Q — *снизу*, во вторую — *сверху* и т. д., в последнюю точку также *сверху* (ввиду того, что число переходов через двойные точки четное), и значит, в Q приходим *снизу*. Но это как раз и есть то, что мы утверждали выше. Новое толкование двойных точек — как пространственных скрещений — приводит нас к заключению, что если мы вышли из Q сверху, то во второй раз мы пройдем Q обязательно снизу.

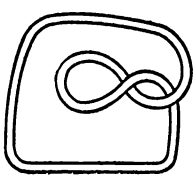


Рис. 40.

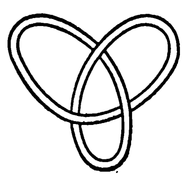


Рис. 41.

В помещенных рядом рис. 40 и 41 кривые, изображенные на рис. 34 и 35, представлены по только что описанному методу как проекции пространственных кривых, причем двойные точки изображены как скрещения. Эти пространственные кривые образуют таким образом «узлы». Узел, в проекции которого проходы сверху и снизу чередуются, называется «альтернирующим» узлом. Строго говоря, рис. 40 представляет не настоящий узел, ибо замкнутую нить, сложенную по схеме рис. 40, можно, очевидно, растянуть в окружность без узлов. Наоборот, нить, свернутая по схеме рис. 41, образует действительный узел, так называемую «трилистниковую» петлю, которую нельзя, не разрывая, вытянуть в окружность.

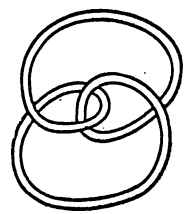


Рис. 42.

В том, что не всякий узел есть непременно альтернирующий, можно убедиться из рис. 42. То обстоятельство, что существуют неальтернирующие узлы, еще раз самым наглядным образом свидетельствует, что наша теорема в обеих ее формулировках никоим образом не тривиальна [18].

12. ОДНОЗНАЧНО ЛИ РАЗЛОЖЕНИЕ ЧИСЛА НА ПРОСТЫЕ СОМНОЖИТЕЛИ?

1. Любое целое число мы можем разлагать на множители до тех пор, пока не придем к неразложимым далее, так называемым «простым» множителям. Например, число 60 можно представить сначала в виде произведения $6 \cdot 10$, а затем 6 разложить на множители 2 и 3, а 10 — на 2 и 5, так что в результате 60 можно представить произведением

$$60 = 2 \cdot 3 \cdot 2 \cdot 5.$$

Эти четыре сомножителя сами уже неразложимы: они являются простыми сомножителями.

То же самое число 60 можно было бы разложить на множители и иначе: $60 = 4 \cdot 15$, далее $4 = 2 \cdot 2$ и $15 = 3 \cdot 5$, откуда получаем:

$$60 = 2 \cdot 2 \cdot 3 \cdot 5,$$

т. е. те же множители в том же количестве, только взятые в иной последовательности. Располагая эти множители в порядке возрастания их величин, мы получим для обоих случаев следующее разложение:

$$60 = 2^2 \cdot 3 \cdot 5.$$

Еще в школе мы привыкли смотреть на все это, как на нечто само собой разумеющееся. Путь, которым мы приходим к разложению на простые множители, допускает много вариантов, но мы чувствуем, что простые множители, получающиеся в окончательном результате нашего разложения, представляют собой как бы основные элементы, из которых путем умножения построены все целые числа, и что при разложении любого числа на множители мы неизбежно должны в конце концов прийти к тем же самым постоянным для данного числа простым множителям.

Это, конечно, правильно. Вопрос заключается лишь в том, так ли все очевидно, как кажётся. Будем говорить конкретно: если нам посчастливилось разложить такое большое число, как 30 031, на два простых множителя $59 \cdot 509$, то откуда, собственно, у нас берется уверенность в том, что нет никаких других простых чисел, входящих множителями в то же самое число 30 031, и что это число не допускает никаких иных способов разложения, при которых мы могли бы обойтись без числа 59 или 509?

В таком предположении скрыто нечто такое, что отвергается нашей интуицией и противоречит нашей уверенности в существовании вполне определенных простых множителей, содержащихся в данном числе. Нижеследующие рассуждения имеют своей целью вскрыть, сколь плохо по существу обоснована эта интуиция, и вместе с тем доказать, что *разложение числа на простые множители действительно однозначно.*

2. Чтобы освободиться от всех соблазнов интуиции, перейдем в совершенно иную, непривычную нам область, в область чисел вида $a+b\sqrt{6}$, где a и b — какие-либо два целых числа. К такого рода числам относятся, например, $12+5\sqrt{6}$ или $\sqrt{6}-2$; ими мы сейчас и займемся. Обыкновенные целые числа отсюда отнюдь не исключаются, наоборот, все они входят в эту область, если положить $b=0$; только что определенная нами числовая область представляет собой лишь расширение совокупности обыкновенных целых положительных и отрицательных чисел.

Оперировать в этой области так же легко, как и в области обыкновенных целых чисел. Из примера

$$(3+\sqrt{6})+(5+\sqrt{6})=8+2\sqrt{6}$$

непосредственно видно, каким образом надо складывать и вычитать такие числа; но и умножать их мы можем научиться столь же легко на нескольких примерах:

$$(3+\sqrt{6})(3-\sqrt{6})=9-6=3,$$
$$(\sqrt{6}+2)(\sqrt{6}-2)=6-4=2,$$
$$(3+\sqrt{6})(\sqrt{6}-2)=3\sqrt{6}-6+6-2\sqrt{6}=\sqrt{6},$$
$$(3-\sqrt{6})(\sqrt{6}+2)=\sqrt{6},$$
$$(3+\sqrt{6})(2+\sqrt{6})=12+5\sqrt{6},$$
$$(3-\sqrt{6})(\sqrt{6}-2)=-12+5\sqrt{6}.$$

Будем надеяться, что этих примеров достаточно для подтверждения того, как легко можно ориентироваться в этой новой для нас области. Школьники, если бы потребовалось обучить их этому исчислению, быстро освоились бы с его несложными правилами.

О делении мы не говорим по весьма понятной причине: точно так же, как и для обыкновенных целых чисел, оно иногда выполнимо, а иногда нет. Нас здесь интересует вопрос о разложении чисел на множители.

Мы исходим из чрезвычайно простого и на первый взгляд даже обидного напоминания, что

$$6=2\cdot 3=\sqrt{6}\cdot\sqrt{6}. \qquad (1)$$

При этом мы сразу же сталкиваемся с тем фактом, что для 6, помимо обычного разложения на множители, оказывается возможно еще одно, совершенно иное разложение. Если мы вспомним поставленный выше вопрос о возможности разложения числа 30 031 на иные простые множители помимо уже известных $59\cdot 509$, то нам может показаться, что с такого рода возможностью мы здесь как раз и встретились. Дело, однако, объясняется очень просто. Хотя числа 2 и 3 суть числа простые в обычном смысле этого слова, т. е. они неразложимы на обыкновенные целые числа, тем не менее они вполне допускают разложение на множители вида $a+b\sqrt{6}$. В этом мы легко можем убедиться из приведенных выше примеров умножения. Поскольку там

$$2=(\sqrt{6}+2)(\sqrt{6}-2), \quad 3=(3+\sqrt{6})(3-\sqrt{6}),$$

то в действительности $6=2\cdot 3$ распадается вовсе не на два множителя, а на четыре:

$$6=(\sqrt{6}+2)(\sqrt{6}-2)(3+\sqrt{6})(3-\sqrt{6}).$$

Два различных разложения, приведенные нами выше, представляют собой не что иное, как лишь преобразования этого четырехчленного разложения в двучленное, причем в первом случае в один множитель были соединены вместе первый и второй, а в другой множитель — третий и четвертый члены, во втором же случае были соединены первый и четвертый и соответственно второй и третий множители. Ясно, что возможно еще и третье соединение множителей, которого мы еще не указывали, а именно первого и третьего и соответственного второго и четвертого. Все необходимые вычисления и для этого случая уже произведены в вышеприведенных примерах; третье разложение

$$6=(12+5\sqrt{6})(-12+5\sqrt{6})$$

может быть легко проверено простым перемножением скобок и непосредственным вычислением.

Нам удалось достаточно хорошо разобраться в этом примере, и было бы нетрудно довести его до конца и показать, что четыре полученных нами множителя числа 6 дальнейшего разложения уже не допускают [1]).

3. Теперь мы перейдем в новую область, а именно в область чисел вида $a+b\sqrt{-6}$, и здесь нам уже не удастся аналогичными рассуждениями получить такой же простой результат. Тем не менее, научившись так легко ориентироваться в области чисел $a+b\sqrt{6}$, мы не встретим особых затруднений при новом изменении нашей области. То обстоятельство, что в нее входит корень из отрицательного числа, не представляет по существу никакого препятствия, ибо вычисления в этой новой области производятся столь же просто и однозначно, как и выше с числами $a+b\sqrt{6}$.

Аналогично предыдущему можно написать:

$$6=2\cdot 3=-\sqrt{-6}\cdot\sqrt{-6}. \qquad (2)$$

[1]) Конечно, только разложения внутри рассматриваемой нами области. Поэтому разложение, подобное, например, $\sqrt{6}=\sqrt{2}\cdot\sqrt{3}$, недопустимо.

Попробуем теперь разложить на множители числа 2, 3 и $\sqrt{-6}$. При этой попытке обнаружится, что такое разложение невозможно. Для упрощения наших рассуждений введем одно вспомогательное понятие. Под нормой числа $a+b\sqrt{-6}$ будем понимать произведение его на число $a-b\sqrt{-6}$, т. е.

$$N(a+b\sqrt{-6})=(a+b\sqrt{-6})(a-b\sqrt{-6})=a^2+6b^2.$$

Иначе говоря, чтобы получить норму числа $a+b\sqrt{-6}$, нужно вместо $\sqrt{-6}$ подставить в это число $-\sqrt{-6}$ и результат перемножить с первоначально данным числом. Следовательно, норма всегда является обыкновенным целым положительным числом.

Имеет место следующая теорема: *норма произведения двух чисел нашей области равна произведению их норм.* Действительно, согласно только что сформулированному правилу:

$$N[(a+b\sqrt{-6})(c+d\sqrt{-6})]=$$
$$=[(a+b\sqrt{-6})(c+d\sqrt{-6})]\cdot[(a-b\sqrt{-6})(c-d\sqrt{-6})],$$

т. е. норма равна первоначально взятому числу (заключенному в первые квадратные скобки), умноженному на выражение, которое получается из этого числа путем замены в нем всюду $\sqrt{-6}$ на $-\sqrt{-6}$ (это выражение заключено во вторые квадратные скобки). Всего, таким образом, мы имеем здесь четыре множителя; изменив порядок их следования, мы можем записать наше выражение в такой форме:

$$(a+b\sqrt{-6})(a-b\sqrt{-6})(c+d\sqrt{-6})(c-d\sqrt{-6}).$$

Согласно же определению понятия нормы, это произведение и представляет собой как раз не что иное, как произведение норм наших чисел:

$$N(a+b\sqrt{-6})\cdot N(c+d\sqrt{-6}).$$

Пусть теперь число 2 разлагается на два множителя нашей области:

$$2=(a+b\sqrt{-6})(c+d\sqrt{-6}).$$

В таком случае его норма

$$N(2)=N(a+b\sqrt{-6})\cdot N(c+d\sqrt{-6}).$$

Представив число 2 в виде $2+0\sqrt{-6}$, соответствующем общему виду чисел рассматриваемой нами области, вычислим его норму:

$$N(2)=N(2+0\sqrt{-6})=$$
$$=(2+0\sqrt{-6})\cdot(2-0\sqrt{-6})=2\cdot 2=4.$$

С другой стороны, имеет место равенство

$$N(a+b\sqrt{-6})\cdot N(c+d\sqrt{-6})=(a^2+6b^2)\cdot(c^2+6d^2),$$

и следовательно, мы получаем для числа 4 разложение

$$(a^2+6b^2)\cdot(c^2+6d^2).$$

Но разложение числа 4 на два обыкновенных целых множителя возможно лишь двумя способами: либо $2\cdot 2$, либо $4\cdot 1$. Однако ни то, ни другое разложение в данном случае не подходит, потому что 2 не может быть представлено в форме x^2+6y^2, где x и y — целые числа; единица же может быть представлена в этой форме, но лишь при $x=1$, $y=0$. Следовательно, 2 можно разложить на два множителя только в том случае, если один из них равен $1+0\sqrt{-6}=1$. Это тривиальное разложение мы столь же мало считаем разложением, как и разложение обыкновенного числа 5 на множители 5 и 1. В области наших чисел мы также будем называть число неразложимым, если оно не допускает никаких других разложений, кроме тех, в которые входят само число и множитель 1.

Этим же способом можно убедиться, что числа 3 и $\sqrt{-6}$ также неразложимы; для этого вместо 4 нам нужно было бы разложить на множители вида x^2+6y^2 в первом случае 9, а во втором 6.

Мы стоим, таким образом, перед тем фактом, что, согласно соотношению (2), число нашей области допускает д в а р а з л и ч н ы х р а з л о ж е н и я на неразложимые множители [39]. Совершенно очевидно поэтому, что если речь идет об обычных целых числах, то нет никаких оснований считать логически само собой разумеющимся, что наряду с одним таким разложением невозможно никакое другое; ибо если бы это было логически очевидным, то оно было бы логически очевидно также и для нашего последнего случая. Если же тем не менее исследуемое нами свойство присуще обыкновенным числам, то это является их особенностью, которая должна быть доказана на основании частных свойств такого рода чисел.

Замечательно, что греческие математики, не зная, вероятно, противоречащих примеров, подобных вышеприведенному, исключительно из соображений логической объективности и ясности инстинктивно почувствовали необходимость доказать однозначность такого рода разложения. У Евклида, впрочем, это утверждение не имеет столь простой современной формулировки. Да и независимо от различия в формулировке мы и самое доказательство поведем несколько иначе, чем оно дано у Евклида.

4. Число 30 кратно 3; оно также кратно 5; его называют поэтому общим кратным 3 и 5. Вообще под общим кратным двух чисел понимают такое число, которое кратно обоим этим числам. Общее кратное существует всегда для любой пары чисел a и b; таковым является, например, их произведение ab. Так, кроме 30, для 3 и 5 общим кратным будет число 15. Само собой разумеется, что 30, т. е. удвоенное 15, и вообще всякое кратное ab является также общим кратным a и b. Два числа, следовательно, имеют всегда бесконечно много общих кратных.

Однако общими кратными двух чисел могут быть не только произведение этих чисел и кратные его. Например, для чисел 10 и 15 общими кратными являются не только числа 150, 300, 450, ..., но, очевидно, и число 30, а вместе с ним и вся последовательность 30, 60, 90, 120, 150, 180, ..., в которой кратные 150 образуют только часть. С другой стороны, число 30 является, очевидно, наименьшим числом, представляющим общее кратное 10 и 15. Вообще ясно, что среди всех общих кратных двух чисел a и b всегда существует одно наименьшее, ибо если испытать все числа до ab, не являются ли они кратными a и b, то должно встретиться какое-то первое, обладающее этим свойством (в крайнем случае первым может оказаться само произведение ab); это число мы назовем наименьшим общим кратным a и b.

Нашей первой задачей будет доказать теорему: *всякое общее кратное двух чисел есть кратное их наименьшего общего кратного.* По отношению к числам $a = 10$, $b = 15$ это должно означать, что последовательность чисел, кратных 30, т. е. 60, 90, 120, ..., исчерпывает все общие кратные чисел 10 и 15; на данном примере, точно так же как и на всяком другом числовом примере, это легко проверить, но нам нужно удостовериться в справедливости этой теоремы для общего случая.

Доказательство основано на простом замечании, что разность двух общих кратных a и b также всегда есть общее кратное a и b. Разность двух кратных a, как это было показано еще в теме 1, есть опять кратное a, разность двух кратных b есть опять кратное b; если же взять разность двух чисел, из которых каждое является кратным как a, так и b, то эта разность тоже должна быть одновременно кратна и a, и b, т. е. она будет общим кратным обоих этих чисел.

Положим теперь, что v есть общее наименьшее кратное чисел a и b, а W — какое-либо общее кратное этих чисел; тогда, согласно предпосланному выше замечанию, разность $W-v$ также должна быть общим кратным a и b, а если из этой разности вычесть v еще раз, то мы опять получим общее кратное чисел a и b; короче говоря, если мы образуем последовательно числа

$$W-v,\quad W-2v,\quad W-3v,\ \ldots,$$

то все они должны быть общими кратными a и b. Так как v является наименьшим из этих общих кратных, то первое из чисел этого ряда $W-v$ непременно положительно. Несколько последующих чисел того же ряда также, возможно, будут положительными.

Но рано или поздно это кончится; в последовательности должно будет встретиться отрицательное число, начиная с которого все дальнейшие числа будут отрицательными. По существу мы имеем здесь дело не с чем иным, как с делением W на v: от W отнимают v столько раз, сколько это возможно (т. е. пока мы получаем положительные остатки), и если деление нацело окажется невозможным, то в результате получится положительный остаток, меньший делителя. Но в нашем случае, если бы это деление нацело оказалось невозможным, то полученный остаток, т. е. последнее из положительных чисел $W-xv$, был бы общим кратным чисел a и b, и притом меньшим числа v, что противоречило бы тому, что v есть наименьшее кратное. Значит, деление нацело должно быть возможно, т. е. W, как это и утверждалось, должно быть кратным v.

5. Понятию общего кратного двух чисел a и b противопоставляется понятие их общего делителя: число t называется общим делителем a и b, если оно входит множителем как в a, так и в b.

Из теоремы п. **4** следует, в частности, что произведение ab, которое всегда заключается среди общих кратных чисел a и b, является кратным v. Докажем сейчас одну лемму, необходимую нам для дальнейшего: *частное от деления произведения двух чисел a и b на их наименьшее общее кратное v, т. е. число*

$$d = \frac{ab}{v},$$

есть всегда общий делитель a и b.

Доказательство очень просто. Из уравнения для d следует, что

$$d\frac{v}{a} = b;$$

так как v кратно a, то число $\frac{v}{a}$ целое, а не дробь, как можно было бы подумать на основании его внешнего вида. Таким образом, b есть целое число, кратное d, или, другими словами, d есть делитель b. Точно так же доказывается, что d есть делитель a, а следовательно, как это и утверждалось нашей леммой, d есть общий делитель a и b.

6. Теперь мы можем доказать решающую теорему, из которой однозначность разложения на простые множители будет следовать в качестве непосредственного вывода. *Если простое число p входит множителем в произведение xy двух чисел x и y, то оно входит множителем или в x, или в y, т. е. во всяком случае в одно из них.*

Доказательство основывается на том, что рассматривают общее наименьшее кратное v обоих чисел p и x. Произведение xy по предположению кратно p и вместе с тем, очевидно, кратно x, т. е. оно является общим кратным p и x, а следовательно, в силу теоремы п. **4**, кратно числу v, что дает нам право положить

$$xy = hv, \tag{3}$$

где h — число целое.

С другой стороны, из теоремы п. **5** следует, что число

$$d = \frac{px}{v} \tag{4}$$

целое и является общим делителем p и x; но делителем простого числа p может быть лишь 1 или p, откуда или $d=p$, или $d=1$. В первом случае $d=p$ есть делитель x, и тогда наша теорема подтверждается, так как p входит в первый множитель произведения xy. Во втором случае, т. е. для $d=1$, мы имеем, в силу соотношения (4), $1=\frac{px}{v}$, т. е. $v=px$, и тем самым, в силу (3), $xy=hpx$; сократив на x, получаем $y=hp$, т. е. в этом случае p входит в y — другой множитель произведения xy. Итак, в обоих случаях оно входит по крайней мере в один из двух множителей.

Отсюда непосредственно следует, что *если простое число входит множителем в произведение нескольких чисел, то оно обязательно входит в одно из них.* Действительно, если, например, оно входит множителем в произведение xyz, то это значит, что оно входит, согласно предыдущей теореме, или в x, или в yz, а в последнем случае оно входит или в y, или в z и, значит, во всяком случае в одно из этих трех чисел.

7. Теорема об однозначности разложения на простые множители следует отсюда непосредственно. Положим, что некоторое число N разложено у нас на простые множители двумя способами:

$$N=p\cdot q\cdot r\cdot s\ldots=P\cdot Q\cdot R\cdot S\ldots$$

В таком случае p, входящее в N, входит и в произведение правой части нашего равенства, а следовательно и в какой-либо простой множитель правой части. Но если одно простое число входит в состав другого, то, в соответствии с понятием простого числа, это может произойти только в том единственном случае, когда оба эти числа тождественны. На этом основании p должно непременно встретиться среди простых множителей правой части; то же самое относится и ко всякому другому простому множителю разложения левой части. А поскольку оба разложения равноправны, отсюда следует, что и каждый простой множитель правого разложения должен встретиться среди левого разложения; короче говоря, оба разложения должны состоять в точности из одних и тех же простых чисел.

Остается лишь выяснить еще, входят ли эти простые числа и в правую, и в левую части с одинаковой кратностью.

Если бы, например, p в левую часть входило a раз, а в правую — A раз, то мы имели бы, с одной стороны,

$$N = p^a q^b r^c \ldots,$$

а с другой стороны,

$$N = p^A q^B r^C,$$

причем при различных A и a одно из них, например A, должно было бы быть больше. Разделим тогда наши равенства на p^a и образуем

$$M = \frac{N}{p^a} = p^{A-a} q^B r^C \ldots = 1 \cdot q^b r^c \ldots$$

Число M оказалось здесь разложенным на простые множители также двумя способами. Но в то время как в правом разложении простой множитель p явно отсутствует, в левом он содержится, если A больше a. Но мы уже показали, что два разложения одного и того же числа должны содержать одни и те же простые множители; это положение должно быть справедливо, в частности, и для числа M, вопреки только что полученному выводу. Значит, a должно быть равно A, и точно так же соответственно должны быть равны и остальные пары показателей $B = b$, $C = c$ и т. д. [40].

Естественно возникает вопрос, почему все эти рассуждения нельзя распространить и на разобранные в п. **3** числа вида $a + b\sqrt{-6}$. В действительности почти все полученные нами выводы переносятся непосредственно и на эти числа. Единственное, что не может быть перенесено, это доказательство теоремы п. 4. В ней, следовательно, и заключается сущность нашей темы.

13. ПРОБЛЕМА ЧЕТЫРЕХ КРАСОК

1. Постановка задачи. А. Кели (A. Cayley) поставил в 1879 г. в Лондонском географическом обществе следующую задачу. Как известно, географические карты печатаются в несколько красок; лучше всего каждую страну печатать особой краской. Но так как многоцветное печатание чрезвычайно дорого, удовлетворяются тем, что в различные краски окрашиваются лишь те страны, которые примыкают одна к другой вдоль какой-либо границы. Географическая карта острова, показанного на рис. 43, *а*, требует при этих условиях т р е х красок: море, как обычно, закрашивается синей краской, и две другие краски используются для двух стран. На рис. 43, *б* показана карта, безусловно требующая ч е т ы р е х красок, ибо здесь все три страны примыкают к морю, и потому ни одна из них не может быть окрашена в тот же цвет, что и море; а поскольку каждая из них соприкасается с двумя другими, все три должны получить разные окраски; всего, значит, требуется четыре краски. Карта, схематически изображенная на рис. 43, *в*, требует также четырех красок, если даже мы не учтем море: средняя страна играет здесь ту же самую роль, какую на рис. 43, *б* играло море. Несколько более сложные карты, как это видно из рис. 44 и 45, можно при вышеуказанном условии окрасить с помощью трех и соответственно четырех красок, обозначенных для краткости буквами *a*, *b*, *c*, *d*.

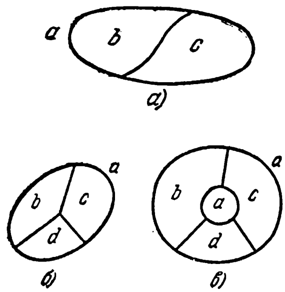


Рис. 43.

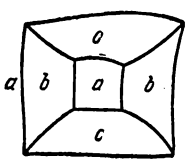


Рис. 44.

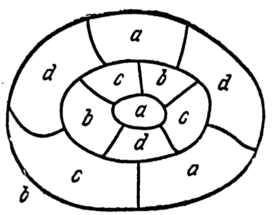


Рис. 45.

Казалось бы, нужно ожидать, что более сложные карты потребуют и бо́льшего числа красок. В действительности, однако, ***никому еще до сих пор не удалось начертить такой, сколь угодно сложной карты, в которой нельзя было бы при указанных условиях обойтись только четырьмя красками.*** С другой стороны, никому также не удалось доказать, что для л ю б о й мыслимой карты достаточно лишь четырех красок. Здесь мы опять сталкиваемся с чрезвычайно простой и понятной без всякой математической подготовки задачей, которую можно в течение нескольких минут растолковать всякому, не имеющему ничего общего с математикой, и которую тем не менее никто еще не сумел решить.

Зато доказано, что *любую карту можно раскрасить пятью красками* — при том, опять-таки, условии, что ни одна пара стран, имеющих общую границу, не получает одинаковой раскраски; однако для стран, соприкасающихся лишь в одной точке (вершине), допускается одна общая краска (как в шахматной доске). Доказательство этого предложения мы и хотим здесь привести.

2. Теорема Эйлера [41]. Основным вспомогательным средством при доказательстве теоремы о пяти красках является общая теорема о количестве вершин e, поверхностей (стран) f и ребер (границ) k в любой географической карте; значение этой теоремы далеко выходит за рамки нашей частной темы. Эта теорема утверждает, что всегда

$$e+f=k+2. \qquad (1)$$

На карте рис. 44, например, имеется 8 вершин (т. е. точек, в которых соприкасаются по крайней мере 3 страны), 6 стран и 12 границ (каждая граница считается от одной вершины до другой, ближайшей), и действительно, мы имеем здесь

$$8+6=12+2.$$

Теорема эта была открыта великим математиком XVIII века Л. Эйлером, но еще за сто лет до Эйлера о ней знал Декарт.

Чтобы доказать ее, представим себе, что наш рисунок изображает вовсе не географическую карту, а систему разграниченных плотинами полей; извне система окружена водой, и нам требуется последовательно снимать эти плотины одну за другой так, чтобы все поля одно за другим оказались орошенными водой (рис. 46). Для осуществления этого вовсе нет необходимости снимать в с е плотины: наоборот, мы условимся оставлять те плотины, с обеих сторон которых поля уже орошены; во всяком случае можно открывать только те плотины, с одной стороны которых поле уже орошено, с другой стороны — еще нет, так что при снятии каждой из них для воды открывается всякий раз новое поле. Ясно, что для орошения имеющихся у нас, кроме внешнего пространства,

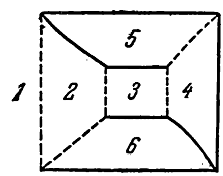


Рис. 46.

$f-1$ полей необходимо снять не более $f-1$ плотин. С другой стороны, ясно также, что процесс снятия плотин должен продолжаться до тех пор, пока у нас не останется ни одного неорошенного поля, т. е. пока у нас не будут орошены все $f-1$ полей. Процесс, следовательно, не может быть закончен, прежде чем не будут сняты $f-1$ плотин. Итак, число подлежащих снятию плотин в точности равно $f-1$.

Рассмотрим теперь систему плотин, оставшихся нетронутыми.

1. *Двигаясь по ним, можно еще сухим путем достигнуть любой вершины, выходя из любой другой.* Ведь первоначально, когда вода находилась лишь вне нашей системы,

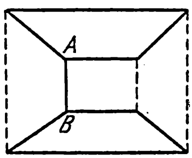


Рис. 47.

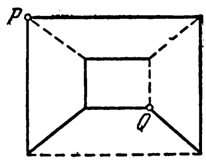


Рис. 48.

это было, конечно, возможно — в противном случае мы имели бы дело не с одной, а с двумя или несколькими независимыми системами полей, с островами, расположенными в окружающей массе воды, и нашу задачу нужно было бы в этом случае решать для каждой такой системы в отдельности. Но если бы в результате одного из последующих снятий какой-либо плотины AB оказалось, что оставшиеся плотины распались на две независимые системы (рис. 47), так что с вершины, расположенной на одной из них, нельзя было бы уже сухим путем достигнуть вершины, расположенной на другой системе, то вода с обеих сторон свободно протекала бы между A и B, а это означало бы, что поля с обеих сторон плотины AB уже были орошены до ее снятия. Мы же условились такие плотины не снимать.

2. *Если из произвольной вершины, например P, послать двумя различными путями двух курьеров для обхода всех плотин и посещения всех вершин, то возможность встречи этих курьеров в какой-либо точке Q исключается.* Ведь если бы вдоль оставшихся плотин от P к Q вело два различных пути, как это указано на рис. 48, то оба эти пути, вместе взятые, ограничили бы некоторую область, со всех

сторон окруженную нетронутыми плотинами и внутрь которой, следовательно, вода никогда не смогла бы проникнуть.

Из данного исходного пункта можно достигнуть любой другой вершины только единственным, вполне определенным путем. При этом, прежде чем попасть в вершину, нужно пройти всякий раз некоторый определенный отрезок (плотину), и, наоборот, по прохождении каждого отрезка (плотины) достигается некоторая вершина — конечный пункт этого отрезка. Следовательно, таких конечных точек существует столько же, сколько остается нетронутых плотин. Так как кроме них в общее число вершин входит еще вершина, служащая исходным пунктом, то количество нетронутых плотин в точности равно $e-1$. Всего, значит, у нас будет $f-1$ снятых и $e-1$ неснятых плотин. Поэтому общее число плотин

$$k=(f-1)+(e-1).$$

3. Доказательству теоремы о пяти красках мы предпошлем еще одно вспомогательное рассуждение: покажем, что для доказательства теоремы в самом общем случае *достаточно доказать нашу теорему для случая карты, в вершинах которой сходятся не более чем три страны.*

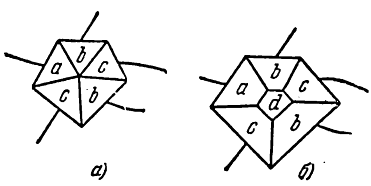


Рис. 49.

Действительно, пусть нам дана карта, в одной из вершин которой сходится больше трех стран (рис. 49, *а*). Начертим рядом вторую карту (рис. 49, *б*), вообще говоря, представляющую собой точную копию первой, но с тем отличием, что в ней около этой вершины образована небольшая новая страна. На этой карте будет, следовательно, еще одна лишняя страна; при этом возникнет также несколько добавочных вершин, в которых сходится по три страны, но зато будет исключена та вершина, в которой сходилось больше трех стран. Поступив точно таким же образом со всеми остальными такими вершинами, можно в конце концов получить карту, в которой таких вершин не будет вовсе. Если бы теперь удалось показать, что всякую карту, в вершинах которой встречается не больше 3 стран, можно окрасить пятью красками, т. е.

что можно обойтись пятью красками в нашей вновь начерченной карте, например, так, как это показано на рис. 49, *б* с помощью букв, то очевидно, что раскраска рис. 49, *а*, удовлетворяющая требованиям задачи, получается непосредственно, если для каждой страны оставить ту же краску, какая была на рис. 49, *б* (так как страны, соприкасающиеся не вдоль границ, а лишь в вершинах, можно по условию окрашивать одинаково), т. е. что пяти красок будет достаточно для любой карты.

4. Доказательство. Пусть на нашей карте имеется f_2 стран с двумя вершинами, f_3 стран с тремя вершинами и т. д. Тогда общее число всех стран f выразится суммой[1])

$$f = f_2 + f_3 + f_4 + \ldots \tag{2}$$

Каждая из f_2 стран, имеющих по 2 вершины, имеет и 2 границы; всего, значит, они дадут $2f_2$ границ; каждая из стран, имеющих по 3 вершины, имеет 3 границы, всего они дадут $3f_3$ границ, и т. д. Пересчитывая таким образом границы стран, мы получим окончательно все границы, но каждую дважды, так как будем учитывать ее для обеих пограничных стран.

Поэтому

$$2k = 2f_2 + 3f_3 + 4f_4 + \ldots \tag{3}$$

Так как мы вправе предположить, что в каждой вершине встречаются 3 страны[2]), то такое же перечисление по странам всех вершин даст нам утроенное их число. Следовательно,

$$3e = 2f_2 + 3f_3 + 4f_4 + \ldots \tag{4}$$

Из сопоставления (3) и (4) следует

$$3e = 2k. \tag{5}$$

Умножив обе части формулы (1) на 6, будем иметь

$$6e + 6f = 6k + 12,$$

---

[1]) Страны без вершин или с одной вершиной, как это легко понять, можно здесь не учитывать.

[2]) Соображения п. 3 дают нам непосредственно лишь право предположить, что в каждой вершине не больше трех стран, но это число не может быть и меньше 3; в противном случае мы имели бы не вершину, а обыкновенную точку на линии.

или, принимая во внимание (5):

$$6e+6f=9e+12,$$

откуда

$$6f=3e+12$$

или, в силу (2) и (4):

$$6(f_2+f_3+f_4+\ldots)=(2f_2+3f_3+4f_4+\ldots)+12.$$

Или, наконец, располагая f_2, f_3, ... в порядке возрастания номеров букв f_2, f_3, f_4, ..., получаем:

$$4f_2+3f_3+2f_4+f_5=12+f_7+2f_8+\ldots \qquad (6)$$

Это дает предварительный результат: *на всякой карте, где в вершинах встречается не больше трех стран, непременно существует страна, имеющая меньше* 6 *вершин.*

Действительно, если бы таковой не оказалось, т. е. если бы $f_2=0$, $f_3=0$, $f_4=0$, $f_5=0$, то левая часть формулы (6) обратилась бы в нуль, что невозможно, так как в правой части ее должно быть по меньшей мере 12. Для карт, изображенных схематически на рис. 43, *а*, *б*, 44, 45, формула (6) дает в правой части как раз 12 (стран с 7, 8 и т. д. вершинами здесь нет), и мы имеем:

рис. 43, *а*: $f_2=3$, $f_3=0$, $f_4=0$, $f_5=0$;
рис. 43, *б*: $f_2=0$, $f_3=4$, $f_4=0$, $f_5=0$;
рис. 44: $f_2=0$, $f_3=0$, $f_4=6$, $f_5=0$;
рис. 45: $f_2=0$, $f_3=0$, $f_4=0$, $f_5=12$.

Во всех этих случаях, следовательно, всякий раз только одно из чисел f_i отлично от нуля. На рис. 43, *в* дело обстоит иначе.

Итак, мы знаем теперь, что на нашей карте существует непременно хотя бы одна страна, имеющая, самое бóльшее, 5 вершин, но может быть и такая, у которой их меньше 5. Рассмотрим по порядку все возможности: двуугольника, треугольника, четырехугольника и пятиугольника.

1. Двуугольник. Для него возможны только две пограничные страны. Представим себе, что одна из его двух границ (на рис. 50 проведенная пунктиром) уничтожена. В измененной таким образом карте будет уже не f стран, а только $f-1$. Допустим, что нам удалось уже окрасить эти $f-1$

стран пятью красками, причем страна, возникшая в результате соединения двух других стран после уничтожения границы, получает окраску a, другая же, соседняя с ней, — окраску b. В таком случае нам ничто не мешает после восстановления уничтоженной границы окрасить нашу исходную страну с 2 вершинами в краску c; действительно, поскольку она не граничит ни с какой другой страной, кроме этих двух, а эти страны окрашены в краски a и b, для окраски ее остаются свободными на выбор краски c, d, e.

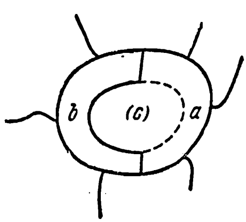


Рис. 50.

Если, следовательно, измененную карту с $f-1$ странами можно было окрасить пятью красками, то это тотчас же осуществимо и для заданной карты с f странами.

2. Треугольник L с тремя смежными странами L_1, L_2, L_3 (рис. 51). Уничтожаем одну из его трех границ. Допустим, что полученная таким путем карта $f-1$ стран уже окрашена пятью красками; тогда при восстановлении уничтоженной границы стране L потребуется дать краску, отличную от тех, которыми окрашены страны L_1, L_2, L_3. Таких свободных красок имеется даже две на выбор.

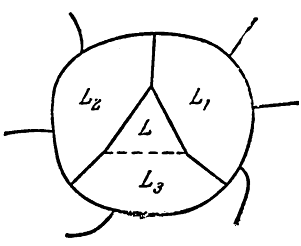


Рис. 51.

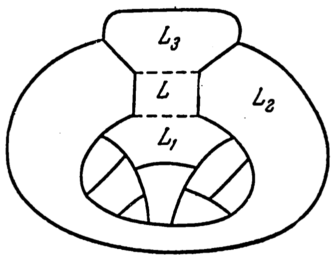


Рис. 52.

3. Четырехугольник. По аналогии очевидно, что здесь для L остается еще одна из пяти красок, так как для пограничных с нею 4 стран может потребоваться, самое большее, четыре различные краски. Однако здесь возникает затруднение иного характера. Одна и та же страна, как, например, L_2 на рис. 52, может примыкать к L вдоль различных

границ. Уничтожив одну из этих границ, мы должны уничтожить и другую, ибо понятие границы несовместимо с тем, чтобы какая-либо страна по этой границе граничила сама с собой. Но в таком случае мы приходим к стране, имеющей форму к о л ь ц а, ограниченного двумя совершенно независимыми границами — внутренней и внешней. Подобные страны мы, не оговаривая того, исключали (уже при доказательстве теоремы Эйлера). Здесь мы также можем их избежать. Ибо

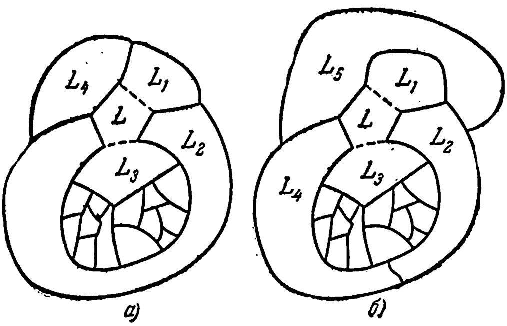


Рис. 53.

если L и L_2 образуют вместе кольцо, то две другие границы L должны принадлежать двум странам L_1 и L_3, разделяемым этим кольцом. Уже в силу этого L_1 и L_3 отличны одна от другой и не могут иметь общей границы, а потому их можно окрасить в один цвет.

Уничтожим границы L с L_1 и L_3. Тогда мы получим карту всего лишь с $f-2$ странами, т. е. с числом стран, меньшим f. Если в ней можно обойтись пятью красками, то их будет достаточно и для первоначальной карты, ибо со страной L граничит лишь 3 смежные страны, из которых две окрашены одинаково. Следовательно, для L остается еще 3 свободные краски.

4. П я т и у г о л ь н и к. В этом случае трудность, возникшая уже в случае четырехугольника, становится еще более серьезной. Может случиться, что две из соседних с L стран совпадают друг с другом (рис. 53, *a*) или примыкают где-нибудь одна к другой (L_2 и L_4 на рис. 53, *a*). В обоих случаях существуют по крайней мере две такие соседние с L страны, как, например, L_1 и L_3, которые не совпадают и не

примыкают друг к другу, ибо их разделяет ленточная цепь, образованная на рис. 53, *а* соединением L с одним только L_2, а на рис. 53, *б* — L с L_2 и L_4 вместе взятыми. Так или иначе, к L примыкают две страны L_1 и L_3, не имеющие общей границы (рис. 54). Уничтожим границы L с обеими этими странами. Мы опять получим карту с $f-2$, т. е. меньшим, чем f, числом стран; если мы сумеем ограничиться для нее пятью красками, то страны $L+L_1+L_3$ получат, например, цвет a, L_2 — цвет b, L_4 — c, L_5 — d, всего, значит, 4 краски. Чтобы окрасить первоначальную карту пятью красками, нам потребуется лишь при восстановлении границ страну L окрасить оставшейся еще пятой краской.

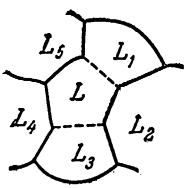


Рис. 54.

Поскольку на **всякой** карте должен быть либо двуугольник, либо треугольник, либо четырехугольник, либо пятиугольник, постольку наш процесс сведе́ния этим завершается, и мы приходим к выводу, что на любой карте можно уничтожить одну или две границы таким образом, что если полученная при этом карта допускает раскраску пятью красками, то это возможно также и для первоначальной карты.

Представим себе теперь, что подобный процесс сведе́ния продолжается до тех пор, пока у нас останется не больше 5 стран. Карту пяти или меньшего числа стран можно, само собой понятно, окрасить пятью красками. Следовательно, карты, рассмотренные выше, в нашем процессе последовательных сведений, а значит, и первоначально данную карту также можно закрасить, обходясь пятью красками [42].

5. Заключение. Во всех наших рассуждениях совершенно не имеет значения, начерчена ли карта на плоском листе, как это мы предполагали до сих пор, или на глобусе. Внешняя страна или окружающий океан заполняет всю остальную поверхность земного шара. Только в том случае, если бы мы не принимали его во внимание, при переходе к глобусу могло бы возникнуть разноречие. Наши рассуждения остаются неизменными и для глобуса; это непосредственно видно из всего их характера. Ведь мы не пользовались здесь теоремами о конгруэнтности или равенстве отрезков и углов, которые на кривой поверхности шара должны были бы получить иной вид; в наших доводах участвовали лишь общие свойства взаимного положения фигур, которые сохраняют свою силу также и на поверхности шара.

Несколько иначе обстояло бы дело, если бы мы пожелали начертить карту на кольце Сатурна (представляя его себе для этой цели твердым телом, а не распыленной массой, каковой оно является в действительности). Здесь можно представить себе карту 7 стран, из которых каждая граничит с 6 остальными; такая карта требует поэтому 7 красок. Однако привести доказательство этого без модели чрезвычайно трудно, и мы принуждены поэтому ограничиться лишь сообщением факта [43]. У нас тотчас же должен возникнуть вопрос: каким образом могло случиться, что наше доказательство достаточности 5 красок потеряло здесь свою силу? Каким образом выводы, оставшиеся правильными при переходе к глобусу, здесь вдруг оказались непригодными?

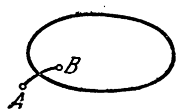


Рис. 55.

Легко указать оба места, в которых происходят разрывы в ходе наших рассуждений. Одно из них лежит в доказательстве теоремы Эйлера, в конце, где, опираясь на рис. 48, мы сделали заключение, что оба пути от P к Q образуют вместе как бы одну целую плотину, замыкающую некоторую часть поверхности, в которую извне вода проникнуть никоим образом не может. Другое лежит в доказательстве основной теоремы, именно там, где разбираются случаи четырехугольника и пятиугольника и где относительно рис. 52, 53, *а* и *б* говорится, что L_1 и L_3 нигде не могут соприкасаться друг с другом, так как в противном случае они должны были бы пересечь сплошное кольцо, образованное из L_2 или соответственно из L_2 и L_4.

В обоих случаях дело сводится к одному и тому же обстоятельству: нельзя попасть с одного берега замкнутой кривой (рис. 55) на другой ее берег, не пересекая кривой. Это положение, всегда справедливое для плоскости и глобуса, перестает быть правильным для кольца Сатурна и вообще для всякой кольцевой поверхности, или, как говорят математики, тора.

Представим себе, например, что кольцо Сатурна (рис. 56) с его внешней, противоположной самому Сатурну стороны обтекает замкнутый поток, и требуется из точки A на берегу этого потока попасть в точку B, лежащую на противоположном берегу, как раз напротив точки A, не пересекая при этом самого потока. Подобная задача совершенно не разрешима в плоскости и на шаре. Здесь, наоборот, стóит только отправиться из точки A через точку C в направлении к Сатурну, затем по внутренней, обращенной к нему стороне спуститься вниз и, пройдя внизу через точку D, подняться вверх, чтобы достигнуть B, не пересекая самого потока.

Случай этот является типичным в том отношении, что показывает, как осторожно в математике нужно обращаться с наглядным представлением, как легко это наглядное представление ведет к скачкам в логической последовательности выводов. Мы видим теперь, насколько подсказанное рис. 48 заключение опиралось на наглядность, и убеждаемся, что в действительности оно должно было бы быть строго логически выведено из особых свойств плоскости и шара, и притом так, чтобы эти выводы для кольцевой поверхности уже не имели силы. Не будем углубляться в дальнейший анализ природы плоскости и шара и применимости по отношению к ним нашего вывода — это вовсе не просто, — а сообщим лишь, что на кольцевой поверхности вместо эйлеровой теоремы имеет место другая:

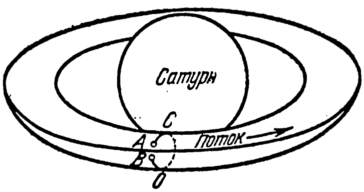


Рис. 56.

$$e+f=k,$$

и что соответствующая теорема о красках в применении к кольцевой поверхности утверждает, что здесь для любой карты достаточно 7 красок. Любопытно, что для более сложной на первый взгляд поверхности — кольцевой — проблема красок окончательно решена, в то время как для более простых случаев — плоскости и шара — вопрос о том, нельзя ли обойтись не 5, а всего 4 красками, остается открытым [44].

14. ПРАВИЛЬНЫЕ МНОГОГРАННИКИ

1. Воспользуемся теперь теоремой Эйлера для разрешения еще одного важного вопроса, хотя и относящегося к совершенно иной области, а именно вопроса о возможности существования правильных многогранников. Под многогранником разумеется тело, ограниченное со всех сторон плоскими гранями; многогранник, согласно определению Евклида, называется **правильным** в том случае, если ограничивающие его грани попарно равны и при этом являются многоугольниками с соответственно равными сторонами и углами (**правильными многоугольниками**). Нам будет удобнее расширить постановку вопроса и положить в основу более общее определение. Назовем **правильным** такой многогранник, все грани которого имеют одинаковое количество вершин и во всех вершинах которого сходится одинаковое количество граней. Величины углов, площади и другие метрические характеристики этих многоугольников в данном случае остаются для нас совершенно безразличными; нас будут интересовать только комбинаторные соотношения.

Пусть многоугольниками, из которых построен наш многогранник, будут только n-угольники (например, треугольники, семиугольники), так что n будет числом вершин каждой грани. Далее, в каждой вершине многогранника пусть сходятся по m граней. В прямолинейном многоугольнике может быть, самое меньшее, 3 угла, и поэтому

$$n \geqslant 3. \tag{1}$$

Так как телесный угол может быть составлен, самое меньшее, из 3 плоских граней, то имеет место также неравенство

$$m \geqslant 3. \tag{2}$$

Через e, f, k обозначим соответственно число вершин, граней и ребер многогранника. Представим себе, что наш многогранник полый, а его грани сделаны из гибкого материала (например, ткани или резины); вообразим далее, что мы станем надувать наш многогранник, пока он не превратится в шарообразное тело неправильной формы. Грани, бывшие ранее плоскими, превращаются при этом в куски кривой шарообразной поверхности, прямолинейные ребра изогнутся в дуги кривых, и если представить себе, что наш шар изображает глобус, то эти линии образуют географическую карту. Число стран на этой географической карте будет равно как раз f — числу граней многогранника, число границ будет равно k — числу ребер многогранника, число вершин e — числу вершин многогранника. Во всех f странах окажется по одинаковому числу n вершин и границ; так что только одно из чисел

$f_2, f_3, \ldots$ будет отлично от нуля и будет поэтому равно общему числу граней f. В каждой вершине сходятся m стран. На рис. 43, *б*, 44, 45 (стр. 96) даны примеры таких географических карт. Мы имеем для них соответственно

$$n=3, \quad m=3; \quad n=4, \quad m=3; \quad n=5, \quad m=3.$$

Посредством этого процесса раздувания многогранника мы приходим к теореме Эйлера:

Соотношение

$$e+f=k+2 \tag{3}$$

имеет место не только для вершин, стран и границ географической карты, но также и для вершин, граней и ребер многогранника.

Исторически именно это второе истолкование и представляет собой ту теорему, которая была сначала открыта Эйлером. Ее поэтому обычно называют теоремой Эйлера о многогранниках. Из формулы (3) мы получим наши главные результаты.

2. Каждая грань правильного в нашем смысле многогранника имеет n вершин и n ребер; все f граней имеют, следовательно, $f \cdot n$ ребер. Если принять во внимание, что каждое ребро принадлежит одновременно двум граням и потому учитывается в этом произведении дважды, то действительное число ребер k определится из равенства

$$f \cdot n = 2k. \tag{4}$$

Аналогично в каждой вершине сходятся m граней и m ребер. Это дает всего $e \cdot m$ ребер; но и здесь также каждое ребро считается дважды в обеих вершинах, к которым оно примыкает своими концами. Поэтому

$$e \cdot m = 2k. \tag{5}$$

Из формулы (3) следует прежде всего, что

$$e+f-k=2,$$

а умножив обе части этого равенства на $2m$, получим:

$$2em+2fm-2km=4m.$$

Заменим здесь, согласно формуле (4), $2k$ через fn, а так как, в силу формул (4) и (5), $fn=em$, то мы можем и em

также заменить на fn. Тогда мы получим соотношение

$$2fn+2fm-fnm=4m,$$

или

$$f(2n+2m-mn)=4m. \tag{6}$$

Так как f и $4m$ — числа положительные, то положительным должно быть и выражение, заключенное в скобки:

$$2n+2m-nm>0. \tag{7а}$$

Формулы (1) и (2) дают нижние границы для n и m, из формулы (7а) мы попытаемся получить теперь для них верхние границы. Прежде всего мы можем записать (7а) в виде

$$nm-2n-2m<0, \tag{7б}$$

изменив в формуле (7а) все знаки на обратные. Сравним левую часть неравенства (7б) с произведением

$$(n-2)(m-2)=nm-2n-2m+4. \tag{8}$$

Последнее отличается от нее лишь слагаемым 4; прибавив поэтому и к правой, и к левой части неравенства (7б) по 4, получим:

$$nm-2n-2m+4<4.$$

Тогда в левой части этого неравенства мы будем иметь как раз выражение (8) и, таким образом, придем к соотношению

$$(n-2)(m-2)<4. \tag{9}$$

Каждый из сомножителей $(n-2)$ и $(m-2)$ этого произведения на основании соотношений (1) и (2) должен быть по крайней мере равен 1. Следовательно, для чисел $n-2$ и $m-2$ мы получили существенное ограничение; они должны быть *целыми положительными числами, произведение которых должно быть меньше* 4.

Но мы можем перечислить непосредственно все произведения, обладающие этим свойством; это будут, очевидно, следующие произведения:

$$1\cdot1;\quad 1\cdot2;\quad 2\cdot1;\quad 1\cdot3;\quad 3\cdot1, \tag{10}$$

т. е. их оказывается всего 5. Все другие произведения двух целых положительных чисел дают в результате по крайней мере 4 или еще больше. *Из того обстоятельства, что мы* *можем только пятью способами образовать произведение двух целых положительных чисел так, чтобы это произведение было меньше* 4, *следует, что может существовать только* 5 *типов правильных многогранников.*

3. В перечисленных выше в (10) произведениях будем считать первый множитель за $n-2$, второй — за $m-2$. Тогда мы получим следующие пять пар значений для n и m:

n	m
3	3
3	4
4	3
3	5
5	3

Вспомнив, какое значение имеют буквы n и m, мы узнаем из этой таблицы, полученной исключительно на основании теоремы Эйлера, что *правильные многогранники, если только они существуют, могут быть ограничены лишь треугольниками, четырехугольниками или пятиугольниками, причем в их вершинах могут сходиться три, четыре или, самое бо́льшее, пять граней.* Из формулы (6) следует, далее, что

$$f=\frac{4m}{2n+2m-nm}.$$

На основании формулы (8) мы можем это соотношение записать в более удобной форме:

$$f=\frac{4m}{4-(n-2)(m-2)}. \tag{11}$$

Пользуясь этим соотношением, выводим из формулы (4), что

$$k=\frac{fn}{2}=\frac{2mn}{4-(n-2)(m-2)}. \tag{12}$$

А на основании формул (5) и (12) получаем:

$$e=\frac{2k}{m}\,\frac{4n}{4-(n-2)(m-2)}. \tag{13}$$

Формулы (11), (12), (13) для каждой из допустимых пар значений n и m дают соответствующие значения f, k, e, сведенные в нижеследующую табличку:

n	m	f	k	e	Многогранник
3	3	4	6	4	тетраэдр
3	4	8	12	6	октаэдр
4	3	6	12	8	гексаэдр (или куб)
3	5	20	30	12	икосаэдр
5	3	12	30	20	додекаэдр

Итак, возможны только 5 *правильных многогранников*; все вместе они называются также платоновыми правильными телами, причем каждый из них, как это видно из таблички, получил наименование от греческих числительных соответственно числу своих граней.

4. Табличка обнаруживает при этом одну особенность. Если взаимно обменять значения n и m, икосаэдр перейдет в додекаэдр, октаэдр — в гексаэдр, а тетраэдр — сам в себя. При этом k останется неизменным, в то время как e и f также обменяются значениями. Но этого можно было ожидать уже из ранее приведенных формул. Прежде всего важнейшие формулы (1), (2) и (9) симметричны относительно n и m, так что всякая допустимая пара значений n и m после взаимного обмена значениями между n и m переходит снова в допустимую пару значений. Далее, (12) показывает, что k зависит от n и m симметрично, а (11) и (13) показывают, что обмен значениями между n и m влечет за собой такой же обмен между e и f.

Геометрически подобная связь также ясна. Стоит только внутри каждой грани выбрать по точке и принять эти точки за вершины нового многогранника, соединив затем ребром две такие точки, расположенные на двух соседних гранях данного многогранника. Тогда на каждой грани данного многогранника расположится как раз по одной вершине нового, каждое ребро старого многогранника скрещивается с одним ребром нового, и, наконец, каждая грань нового многогранника будет окружать одну из вершин старого. Значит, число ребер k в обоих многогранниках будет одно и то же, а числа e и f обменяются своими значениями.

5. Нужно еще обратить внимание на следующее. Наши рассуждения показывают только, что может быть не больше пяти правильных многогранников, ибо только приведенные в таблице значения e, f и k могут соответствовать таким многогранникам, грани которых имеют по равному числу вершин, где сходятся равные количества граней. Других значений e, f и k, кроме приведенных в табличке, быть не может.

Но мы еще не установили, что эти многогранники можно действительно построить. Ведь могло бы случиться, что

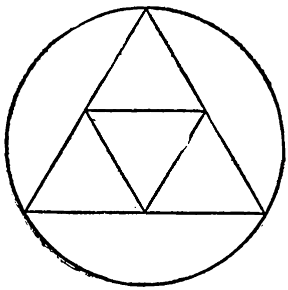


Рис. 57.

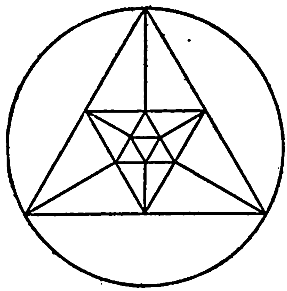


Рис. 58.

правильные многогранники должны быть подчинены еще какимлибо, не учтенным нами дополнительным ограничениям, из-за которых некоторые из перечисленных в табличке видов многогранников должны отпасть. Короче говоря, мы выяснили лишь *необходимые*, но не *достаточные* условия существования правильных многогранников.

Наши рассуждения применимы, собственно, не только к многогранникам, но и к значительно более общему понятию «правильной» географической карты на глобусе. Мы можем даже забыть о том, что такие карты возникали у нас в результате надувания многогранника, ограниченного плоскими гранями. Географические карты, соответствующие пяти приведенным в табличке видам, мы можем действительно построить. Рис. 43, 44 и 45 изображают соответственно виды тетраэдра, гексаэдра и додекаэдра. Виды октаэдра и икосаэдра представлены на рис. 57 и 58.

Следовательно, если отвлечься от тех деформаций, которым мы подвергли наш многогранник, то окажется, что установленные выше необходимые условия существования правильных многогранников являются и достаточными [45].

Мы, однако, еще не доказали, что наши правильные многогранники могут быть правильными также и в более узком, евклидовом или метрическом смысле, т. е. что они могут быть ограничены плоскими правильными конгруэнтными многоугольниками. Доказательство этого положения требует совершенно иных вспомогательных средств, ибо наши исследования касались *только таких свойств, которые остаются неизменными при рассмотренных выше деформациях.* Конгруэнтность же не принадлежит к числу этих свойств. Здесь смогут привести к цели лишь методы метрической геометрии, в которой решающую роль играют равенства отрезков и углов. Мы не будем приводить здесь доказательства того, что перечисленные нами пять правильных многогранников могут быть правильными также и в смысле метрической геометрии. Это доказательство восходит к классической древности и приписывается ученику Платона Теетету. Оно приведено в заключении «Начал» Евклида, в XIII книге [46].

15. ПИФАГОРОВЫ ЧИСЛА И ПОНЯТИЕ О ТЕОРЕМЕ ФЕРМА

1. Согласно известной теореме Пифагора, площадь квадрата, построенного на гипотенузе прямоугольного треугольника, равна сумме площадей квадратов, построенных на его катетах. Если, обратно, имеются три таких отрезка, что квадрат, построенный на одном из них, равен сумме квадратов, построенных на двух остальных, то построенный из этих трех отрезков треугольник будет всегда прямоугольным. Будем рассматривать вместо отрезков измеряющие их величины: a, b, c; тогда величины площадей построенных на них квадратов будут a^2, b^2, c^2, и уравнение $a^2+b^2=c^2$ будет выражать тот факт, что a, b, c представляют собой величины сторон прямоугольного треугольника.

Мы уже видели в 4-й теме что, катеты и гипотенуза равнобедренного прямоугольного треугольника несоизмеримы и уравнение $2a^2=c^2$ не может быть разрешено в целых значениях a и c. Возникает вопрос, существуют ли такие неравнобедренные прямоугольные треугольники, катеты которых были бы соизмеримы с гипотенузой, или, иными словами, существуют ли решения уравнения

$$a^2+b^2=c^2 \tag{1}$$

при целых значениях a, b и c. Простой и всем известный пример, а именно

$$3^2+4^2=5^2, \quad \text{или} \quad 9+16=25$$

показывает, что вопрос решается в положительном смысле.

Существуют ли кроме указанных еще и другие такие «пифагоровы» числа, т. е. целочисленные решения уравнения (1), и если существуют, то каковы они? На этот вопрос мы намерены дать исчерпывающий ответ.

2. Если найдено одно целочисленное решение (a, b, c) уравнения (1), то из него чрезвычайно простым способом, а именно умножением всех значений a, b, c на одно и то же число, можно получить новое решение. Например, из приведенного выше решения 3, 4, 5 тотчас же получается новое 6, 8, 10, так как

$$6^2+8^2=10^2,$$

и вообще $3n$, $4n$, $5n$ при любом целом n. Равным образом решение a, b, c влечет за собой и решение an, bn, cn, так как из $a^2+b^2=c^2$ следует $a^2n^2+b^2n^2=c^2n^2$, т. е. $(an)^2+(bn)^2=(cn)^2$.

Однако этот способ получения новых решений настолько очевиден, что не представляет большого интереса. Мы будем интересоваться только основными решениями, т. е. такими, которые не могут быть получены из более простых решений путем умножения на целое число. Такое основное решение характеризуется тем, что значения a, b, c не имеют ни одного общего делителя, как, например, решение 3, 4, 5.

Чтобы решение a, b, c было основным, необходимо, чтобы числа a, b, c попарно не имели общего делителя: они, как говорят, должны быть попарно взаимно простыми, так как если бы, например, a и b имели общий делитель t, то и любой делитель числа t был бы также их общим делителем. Обозначив через p какое-нибудь простое число, входящее множителем в t (если t — само простое число, то $p=t$), мы могли бы положить:

$$a=pa_1, \quad b=pb_1,$$

и из $a^2+b^2=c^2$ получилось бы, что

$$p^2(a_1^2+b_1^2)=c^2.$$

Но это уравнение показывает, что p^2 должно быть делителем c^2. Отсюда заключаем, что p должно быть делителем c, ибо, согласно теореме о делимости произведения на простое число, c^2 не может делиться на p, если на p не делится каждый из его равных сомножителей c. Поэтому p было бы общим делителем не только a и b, но a, b и c. Таким же образом общий простой делитель a и c или общий простой делитель b и c был бы общим делителем всех трех чисел a, b, c.

3. Итак, мы будем искать только такие решения a, b, c уравнения (1), в которых числа a, b, c попарно взаимно простые. В частности, ни одна пара чисел a, b, c не может быть четной, т. е. делиться на 2. Из трех чисел a, b, c только одно может быть четным. С другой стороны, все три числа не могут быть одновременно и нечетными. Ибо нечетное число $a=2l+1$ дает в квадрате $a^2=4l^2+4l+1=$ $=4(l^2+l)+1$, т. е. опять нечетное число; если a и b нечетны, то нечетны также a^2 и b^2, и следовательно, их сумма a^2+b^2 — число четное, что невозможно при нечетном c.

Остается, следовательно, только та возможность, что из трех чисел a, b, *c два нечетных и одно четное*. Кроме того, легко понять, что c должно быть нечетным; ибо если c четное, т. е. делится на 2, то c^2 должно делиться на 4. При четном c оба других числа a и b должны быть нечетными; если, следовательно,

$$a=2l+1, \quad b=2m+1,$$

то

$$a^2+b^2=(4l^2+4l+1)+(4m^2+4m+1)=$$
$$=4(l^2+l+m^2+m)+2.$$

Число a^2+b^2 хотя и четное, но на 4 не делится, оно дает при делении на 4 остаток 2 и потому не может быть равно c^2, делящемуся на 4.

Итак, единственная остающаяся возможность для основных решений: c — нечетное число, а из двух остальных чисел a и b — одно четное, другое нечетное. Положим, что a нечетно, а b четно; в нашем примере именно так и было: $a=3$, $b=4$, $c=5$.

4. Уравнение (1) мы можем написать в такой форме:

$$b^2 = c^2 - a^2 = (c+a)(c-a). \tag{2}$$

Здесь $(c+a)$ и $(c-a)$ как сумма и разность двух нечетных чисел сами должны быть четными, т. е. иметь общий делитель 2. Других общих делителей у них не может быть, т. е. $\frac{c+a}{2}$ и $\frac{c-a}{2}$ должны быть взаимно простыми числами. В самом деле, если бы t входило множителем в оба эти выражения, т. е. если бы имели место равенства

$$\frac{c+a}{2} = tf, \quad \frac{c-a}{2} = tg,$$

то, складывая и вычитая эти равенства, мы получили бы, что

$$c = t(f+g), \quad a = t(f-g),$$

т. е. t входило бы множителем и в a, и в c, что невозможно, так так, согласно предположению, у a и c нет ни одного общего множителя.

Так как b, $c+a$, $c-a$ — числа четные, то формулу (2) мы можем написать в целых числах таким образом:

$$\left(\frac{b}{2}\right)^2 = \frac{c+a}{2} \cdot \frac{c-a}{2}. \tag{3}$$

Равенством (3) мы разложили квадратное число $\left(\frac{b}{2}\right)^2$ на два взаимно простых множителя $\frac{c+a}{2}$ и $\frac{c-a}{2}$. Отсюда можно заключить — и это является центральным пунктом наших рассуждений,— что $\frac{c+a}{2}$ и $\frac{c-a}{2}$ также должны быть квадратами. И действительно, если мы имеем разложение на простые множители

$$\frac{b}{2} = p^{\alpha} q^{\beta} r^{\gamma} \ldots,$$

где p, q, r — различные простые числа, то

$$\left(\frac{b}{2}\right)^2 = p^{2\alpha} q^{2\beta} r^{2\gamma}.$$

В состав $\frac{c+a}{2}$ и $\frac{c-a}{2}$, с одной стороны, должны войти все простые множители их произведения $\left(\frac{b}{2}\right)$, а с другой стороны, каждое простое число, например p, может войти или только в $\frac{c+a}{2}$, или только в $\frac{c-a}{2}$, так как $\frac{c+a}{2}$ и $\frac{c-a}{2}$ не имеют ни одного общего множителя. Поэтому простые множители $\left(\frac{b}{2}\right)^2$ распределяются между $\frac{c+a}{2}$ и $\frac{c-a}{2}$ таким образом, что каждая степень простого числа из $p^{2\alpha}$, $q^{2\beta}$, $r^{2\gamma}$ входит целиком или в $\frac{c+a}{2}$, или в $\frac{c-a}{2}$. Этим самым устанавливается, что $\frac{c+a}{2}$ и $\frac{c-a}{2}$ содержат только четные степени входящих в них простых множителей и, следовательно, являются квадратами.

5. Мы можем поэтому положить:

$$\frac{c+a}{2}=u^2, \quad \frac{c-a}{2}=v^2, \tag{4а}$$

$$\left(\frac{b}{2}\right)^2=u^2v^2, \tag{4б}$$

где u и v — числа взаимно простые, так как взаимно простыми являются u^2 и v^2.

Из формулы (4б) следует, что

$$b=2uv, \tag{5}$$

а из обоих уравнений (4) получаем сложением и вычитанием

$$c=u^2+v^2, \quad a=u^2-v^2. \tag{6}$$

Так как c и a должны быть нечетными, то из двух чисел u^2 и v^2, а следовательно, и из двух чисел u и v одно должно быть четным, другое нечетным (так как если бы оба были четны или нечетны, то сумма и разность u^2 и v^2 были бы непременно четны). В последующем два числа, из которых одно четное, другое нечетное, мы будем для краткости называть *разнородными*.

Таким образом, мы пришли к выводу, что три взаимно простых числа a, b, c, удовлетворяющих уравнению (1), могут быть с помощью формул (5) и (6) выражены через два взаимно простых *разнородных* числа u и v. В нашем старом примере $a=3$, $b=4$, $c=5$ мы имеем, очевидно,

$$u^2=\frac{5+3}{2}=4, \quad v^2=\frac{5-3}{2}=1,$$

$$u=2, \quad v=1;$$

и действительно:

$$b = 4 = 2 \cdot 2 \cdot 1 = 2 \cdot u \cdot v.$$

6. До сих пор мы исходили из основного решения a, b, c и определяли соответствующие ему числа u и v. Если мы теперь, наоборот, обратим внимание на то, что каждая пара разнородных взаимно простых чисел u и v при $u > v$ дает с помощью формул (5) и (6) основные решения уравнения (1), то мы получим возможность систематически находить все основные решения.

Но мы имеем тождество

$$(u^2 - v^2)^2 + (2uv)^2 = (u^2 + v^2)^2,$$

в справедливости которого легко можно убедиться, раскрыв скобки в обеих частях уравнения. Следовательно, все числа a, b, c, получающиеся из формул (5) и (6), действительно удовлетворяют уравнению (1). Из того, что u и v разнородны, т. е. не могут быть одновременно ни четными, ни нечетными, следует, что a и c, согласно (6), должны быть нечетными, так что a, b, c не могут иметь 2 своим общим множителем. Нечетного общего множителя у них также нет, ибо тогда они имели бы также нечетный простой множитель [1] p, и можно было бы положить:

$$c = pc_1, \quad a = pa_1,$$

так что из (4а) должно было бы следовать соотношение

$$2u^2 = c + a = p(a_1 + c_1),$$
$$2v^2 = c - a = p(c_1 - a_1).$$

Значит, $2u^2$ и $2v^2$ должны были бы также делиться на p. Но так как p нечетно, то этот множитель должен был бы уже содержаться в u^2 и v^2, а это несовместимо с тем обстоятельством, что u и v — взаимно простые числа.

Для иллюстрации вычислим с помощью формул (5) и (6) несколько пифагоровых чисел:

$u=2$,	$v=1$:	$a=3$,	$b=4$,	$c=5$;
$u=3$,	$v=2$:	$a=5$,	$b=12$,	$c=13$;
$u=4$,	$v=1$:	$a=15$,	$b=8$,	$c=17$;
$u=4$,	$v=3$:	$a=7$,	$b=24$,	$c=25$;
$u=5$,	$v=2$:	$a=21$,	$b=20$,	$c=29$;
$u=5$,	$v=4$:	$a=9$,	$b=40$,	$c=41$.

[1]) Всякое простое число, отличное от 2, будет, очевидно, нечетным.

Бросается в глаза, что число b здесь не только четно, но и всюду кратно 4. Что это всегда должно быть так, показывает непосредственно формула $b=2uv$, в которой произведение двух разнородных чисел u и v должно быть четным [47].

7. После того как мы изучили уравнение (1), можно перейти к некоторым обобщениям. Можно, например, задаться вопросом о существовании целочисленных решений уравнений

$$x^3+y^3=z^3, \tag{7а}$$

$$x^4+y^4=z^4, \tag{7б}$$

или в общем виде

$$x^n+y^n=z^n \tag{7в}$$

для любого $n>2$. Согласно знаменитому, правда, до сих пор все еще не доказанному, но также и не опровергнутому утверждению Пьера Ферма (1601—1665), *уравнение* (7в) *не имеет целых положительных решений ни для какого целого* $n>2$ [48]. Для некоторых n, например для всех n, лежащих в промежутке от 3 до 100, утверждение Ферма доказано Куммером (1810—1893) и его учениками [49]. Еще Эйлер (1707—1783) обнаружил, в частности, неразрешимость (в целых числах) уравнений (7а) и (7б). Полученные сведения о пифагоровых числах дадут нам возможность доказать неразрешимость в целых числах уравнения (7б).

Мы докажем, что даже такое уравнение, как

$$x^4+y^4=w^2, \tag{8}$$

не имеет решений в целых числах. А поскольку каждая четвертая степень является квадратом (хотя не всякий квадрат является четвертой степенью), неразрешимость уравнения (8) дает даже больше, чем неразрешимость (7б).

8. Если уравнение (8) написать в виде

$$(x^2)^2+(y^2)^2=w^2, \tag{9}$$

то легко заметить, что оно представляет собой частный случай уравнения (1) при $a=x^2$, $b=y^2$, $c=w$. Так как и здесь речь идет о нахождении только основных решений x, y, w, мы заключаем, как и раньше, что w нечетно и что из чисел x^2 и y^2 одно должно быть четным, другое нечетным, если только решение вообще существует.

Пусть опять $x^2=a$ нечетно, $y^2=b$ четно. Если уравнение (9) имеет решение, то последнее может быть выражено формулами (5) и (6), а это значит, что должны существовать два разнородных взаимно простых числа u и v, такие, что

$$x^2 = u^2 - v^2, \tag{10а}$$

$$y^2 = 2uv, \tag{10б}$$

$$w^2 = u^2 + v^2, \tag{10в}$$

откуда формулу (10а) можно написать еще так:

$$x^2 + v^2 = u^2, \tag{11}$$

что дает новое пифагорово уравнение с взаимно простыми x, v, u. Здесь u играет роль прежнего c и поэтому, согласно условию (3), должно быть нечетным; так как x — нечетное число, то для v остается лишь одна возможность: оно должно быть четным, т. е. выступать в роли прежнего b. Величины x, v, u как основное решение пифагорова уравнения (11) могут быть опять-таки, в силу формул (5) и (6), представлены с помощью новых разнородных взаимно простых чисел u_1 и v_1:

$$x = u_1^2 - v_1^2, \quad v = 2u_1v_1, \quad u = u_1^2 + v_1^2. \tag{12}$$

Теперь вспомним об уравнении (10б). Так как u нечетно, а v четно и, кроме того, u и v — взаимно простые числа, то u и $2v$ также взаимно простые, ибо множитель 2 в состав u не входит. Поэтому, согласно условию (10б), y^2 разлагается на взаимно простые множители u и $2v$. Но по соображениям, приведенным в п. 4, ясно, что разложение квадрата на взаимно простые множители возможно только в том случае, если сами эти множители являются квадратами, т. е. если для нашего случая

$$u = w_1^2, \quad 2v = 4t_1^2, \tag{13}$$

где учтено, что $2v$ — число четное. Вставив эти значения u и v в наши последние уравнения (12), получим:

$$t_1^2 = u_1v_1, \quad w_1^2 = u_1^2 + v_1^2. \tag{14}$$

Здесь u_1 и v_1 — также взаимно простые числа, дающие в произведении квадрат t_1^2; поэтому они сами должны быть квадратами

$$u_1 = x_1^2, \quad v_1 = y_1^2, \tag{15}$$

что после подстановки в уравнение (14) дает равенство

$$x_1^4 + y_1^4 = w_1^2. \tag{16}$$

9. Это последнее уравнение — того же вида, что и исходное уравнение (8). Мы видим, следовательно, что если существует одно основное решение уравнения (8) x, y, w, то имеется еще и другое его решение x_1, y_1, w_1, причем x_1, y_1, w_1 могут быть получены из x, y, w путем определенных операций, а w первого решения должно быть непременно больше w_1 второго решения, ибо, согласно (10в) и первому уравнению, имеем:

$$w = u^2 + v^2 = w_1^4 + v^2 > w_1^4,$$

а следовательно тем более $w > w_1$.

Из этих формул легко вывести противоречие. Ибо точно так же, как в п. 8 из x, y, w мы получили новое решение x_1, y_1, w_1, мы можем путем применения того же процесса из x_1, y_1, w_1 получить решение x_2, y_2, w_2, которое будет обладать аналогичным свойством: $w_1 > w_2$. Новое применение того же процесса дает нам решение x_3, y_3, w_3 с $w_2 > w_3$. Продолжая этот процесс, мы получим убывающую последовательность чисел

$$w > w_1 > w_2 > w_3 > \ldots \tag{17}$$

Все эти числа должны быть целыми и положительными, но таких чисел, меньших w, в убывающей последовательности может быть только конечное число, в силу чего в последовательности (17) должно быть некоторое последнее число, например w_k. Так как, с другой стороны, это последнее число также должно быть решением уравнения (8), то из него, согласно (8), можно получить новое решение x_{k+1}, y_{k+1}, w_{k+1}, причем $w_k > w_{k+1}$, а этот вывод приводит нас к противоречию. Основную идею этого доказательства принято называть, следуя Ферма, методом безграничного спуска (descente infinie[1]). Идея эта заключается в противоречивости того положения, что некоторый процесс [в данном случае повторение операции (8)] доставляет бесконечную последовательность убывающих целых положительных чисел, между тем как убывающий ряд целых положительных чисел может иметь только конечное число членов, ибо в убывающей

[1] Впервые этот метод был применен математиком XIII века Кампаном при доказательстве иррациональности золотого сечения. (*Прим. перев.*)

последовательности за каждым числом может лежать только конечное число целых чисел, а именно только числа $n-1$, $n-2$, ..., 3, 2, 1, т. е., самое большое, $(n-1)$ чисел.

Этот метод безграничного спуска лежит также и в основе доказательства иррациональности $\sqrt{2}$ (тема 4) [50].

16. ЗАМЫКАЮЩАЯ ОКРУЖНОСТЬ ТОЧЕЧНОЙ СОВОКУПНОСТИ

1. В плоскости дано n точек P_1, P_2, ..., P_n. Мы будем называть их *точечной совокупностью*[1]). Рассмотрим теперь все возможные расстояния между произвольными парами точек P_i и P_k точечной совокупности. Среди конечного множества[2]) этих расстояний должно существовать наибольшее. Отметим особо эту величину и назовем ее *диаметром* точечной совокупности, состоящей из n точек.

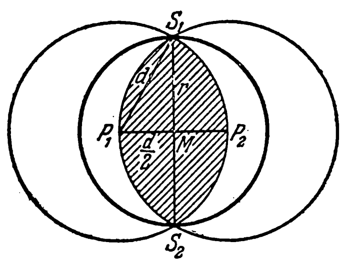


Рис. 59.

Вокруг точечной совокупности диаметра d можно, очевидно, описать окружность радиуса d, внутри и на границе которой будут заключены все ее n точек (рис. 59). Достаточно для этого описать окружность радиуса d из произвольной точки нашей совокупности, например из P_1. Так как всякая другая точка отстоит от P_1 самое бóльшее, на d, наша окружность, кроме своего центра P_1, будет заключать внутри себя и все остальные точки P_2, P_3, ..., P_n.

Но легко построить и окружность меньшего радиуса, которая, однако, также будет содержать все n точек нашей

[1]) Мы сознательно отказываемся здесь от термина «точечное множество», поскольку он применяется также и для обозначения совокупностей бесконечно большого числа точек. Согласно нашей терминологии, точечная совокупность должна содержать только *конечное* число точек.

[2]) Легко сообразить, что для n точек существует $\frac{n(n-1)}{2}$ возможностей образовать различные пары (см. стр. 72).

совокупности. С этой целью отыщем пару точек, расстояние между которыми имеет наибольшее значение d. (Если имеется несколько таких пар точек, отстоящих на d, выбираем любую из этих пар.) Около каждой из точек этой пары — обозначим их через P_1 и P_2 — проведем окружность радиуса d; каждая из этих окружностей, очевидно, пройдет через другую точку пары. Но, как мы уже знаем, все наши n точек лежат, с одной стороны, внутри окружности, описанной из P_1, с другой стороны, все они должны лежать внутри окружности, описанной из P_2. Следовательно, все n точек должны лежать в области, общей обеим окружностям, т. е. внутри заштрихованного двуугольника (рис. 59), образованного дугами окружностей. Около этого двуугольника с вершинами S_1 и S_2 можно в свою очередь описать окружность с диаметром S_1S_2. Чтобы узнать ее радиус r, применяем теорему Пифагора к прямоугольному треугольнику P_1MS_1:

$$r^2 = d^2 - \left(\frac{d}{2}\right)^2 = \frac{3}{4}d^2,$$

откуда

$$r = \frac{d}{2}\sqrt{3}.$$

Круг, содержащий внутри себя или на своей периферии все n точек нашей точечной совокупности H, мы назовем **объемлющим кругом** точечной совокупности. Вместо первоначально указанного объемлющего круга радиуса $r = d$ мы нашли объемлющий круг меньшего радиуса:

$$r = \frac{d}{2}\sqrt{3} = 0{,}866\ldots d.$$

2. Теперь возникает вопрос, нельзя ли это число $r = \frac{d}{2}\sqrt{3}$ заменить еще меньшим. На этот вопрос дает ответ следующая теорема Юнга [51]:

Для любой точечной совокупности диаметра d всегда существует объемлющий круг, радиус которого не превышает

$$\frac{d}{3}\sqrt{3} = 0{,}577\ldots d.$$

Для некоторых точечных совокупностей этот радиус можно еще уменьшить [1]), но существуют точечные совокупности, для которых меньший радиус вообще недостижим. Доказательство теоремы Юнга и составляет задачу настоящей темы.

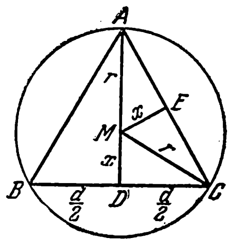


Рис. 60.

Для точечной совокупности, состоящей из вершин равностороннего треугольника со стороной d, мы легко можем указать объемлющую окружность радиуса Юнга $\frac{d}{3}\sqrt{3}$. Такой будет окружность, описанная около данного нам равностороннего треугольника. Ибо, положив в прямоугольном треугольнике ABD (рис. 60) $r+x=h$, мы будем иметь:

$$d^2=h^2+\frac{d^2}{4},$$

$$h^2=\frac{3d^2}{4}, \tag{1}$$

откуда

$$h=\frac{d}{2}\sqrt{3}. \tag{2}$$

Далее из треугольника DCM имеем:

$$x^2+\frac{d^2}{4}=r^2,$$

$$(h-r)^2+\frac{d^2}{4}=r^2,$$

$$h^2-2hr+r^2+\frac{d^2}{4}=r^2,$$

$$h^2+\frac{d^2}{4}=2hr;$$

в силу (1),

$$d^2=2hr,$$

$$r=\frac{d^2}{2h},$$

[1]) Так как две точки, отстоящие одна от другой на расстоянии d, не могут поместиться внутри круга с радиусом, меньшим $\frac{d}{2}$, а в точечной совокупности диаметра d всегда имеется пара точек с расстоянием d, то *радиус объемлющего круга ни в коем случае не может быть меньше* $\frac{d}{2}$.

откуда, принимая во внимание (2), получаем:

$$r=\frac{d^2}{d\sqrt{3}}=\frac{d}{3}\sqrt{3},$$

что и требовалось доказать.

Для равностороннего треугольника совершенно очевидно, что описанная окружность является в то же самое время и наименьшей объемлющей. Мы не будем заниматься этим вопросом сейчас, так как в дальнейшем это обстоятельство выявится само собой.

3. Чтобы доказать теперь теорему Юнга в общем случае для произвольной точечной совокупности диаметра d, начнем с того, что постараемся среди всех возможных объемлющих окружностей нашей точечной совокупности отыскать наименьшую. С этой целью применим ряд процессов, которые должны привести к последовательному уменьшению объемлющей окружности.

I. *Объемлющая окружность K_1, на которой не лежит ни одной точки точечной совокупности H, может быть заменена меньшей окружностью K_2.* Для этого нужно только из центра M окружности K_1 описать окружность K_2, проходящую через наиболее удаленную от M точку точечной совокупности. Эта окружность будет лежать целиком внутри K_1, так как K_1 содержит все точки H, а следовательно, и ту, которая лежит на окружности K_2. При этом очевидно, что K_2 есть также объемлющая окружность H.

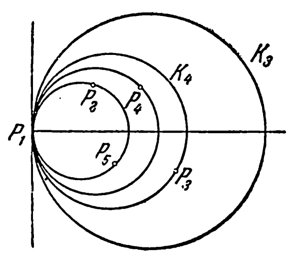


Рис. 61.

II. *Объемлющую окружность K_3, на которой лежит только одна точка точечной совокупности, можно уменьшить* (рис. 61). Пусть P_1 будет точкой H, лежащей на K_3. Опишем все окружности, имеющие общую с K_3 касательную в P_1 и на которых лежит еще по одной точке H. Все эти окружности лежат внутри K_3. Наибольшая из них K_4 (которая должна отличаться от K_3, ибо на K_3 лежит только P_1, а на K_4, кроме P_1, есть еще другая точка H) объемлет все остальные окружности, а вместе с тем и все точки H, причем две из n точек лежат на ней. Эта окружность меньше окружности K_3.

III. Точки H, лежащие на объемлющей окружности, делят эту окружность на дуги; эти дуги не содержат внутри точек H. В дальнейшем мы будем называть такую дугу, на которой нет ни одной точки H, «свободной»; при этом конечными точками свободной дуги могут быть точки H. Мы утверждаем, что *объемлющая окружность K_5, в которой имеется свободная дуга длиной больше полуокружности*[1]), *может быть, заменена меньшей объемлющей окружностью K_6.*

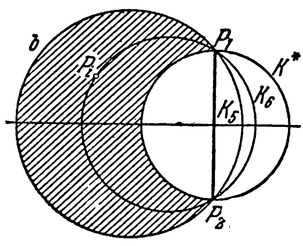


Рис. 62.

Действительно, пусть на окружности K_5 точки P_1 и P_2, принадлежащие H, являются конечными точками дуги b, длина которой больше полуокружности (рис. 62). На отрезке P_1P_2, как на диаметре, построим окружность K^*. Если внутри и на самой этой окружности содержатся все точки H, значит, она уже и есть искомая нами объемлющая окружность, меньшая окружности K_5, ибо в этой последней хорда P_1P_2 стягивает дугу, не равную полуокружности, и потому будет меньше ее диаметра. Если K^* заключает в себе не все точки H, то все не входящие в K^* точки H должны лежать в серповидной (заштрихованной) области между b и K^*, причем на самой дуге b, кроме ее конечных точек P_1 и P_2, нет по условию никакой другой точки H. Для точек, расположенных в заштрихованной серповидной области, можно провести ряд объемлющих окружностей, проходящих также через P_1 и P_2. Внутри окружности K_5 все они проходят по заштрихованной серповидной области, т. е. вне K^*, а та их часть, которая располагается внутри K^*, лежит вне K_5. Пусть K_6 будет одной из этих окружностей, причем такой, что ее дуга в заштрихованной серповидной области отстоит на максимальном расстоянии от хорды P_1P_2. Она объемлет все точки H, так как она содержит все точки H, лежащие внутри заштрихованной серповидной области, и, кроме того, общую часть

[1]) Положения I и II можно, впрочем, рассматривать как частные случаи III, так как и в том, и в другом — свободной является полная окружность.

кругов K_5 и K^*, в которой должны лежать все остальные точки H. Далее, K_6 меньше, чем K_5, так как дуга K_6 лежит между K_5 и K^*, и при этом центр K_6 расположен ближе к хорде P_1P_2, чем центр K_5.

На объемлющей окружности, которую уже нельзя больше уменьшить операциями I, II, III, не может быть, следовательно, свободной дуги, большей полуокружности. На такой объемлющей окружности могут оказаться или только две точки H, являющиеся концевыми точками ее диаметра, или не менее трех точек H так, что определяемые ими свободные дуги меньше полуокружности. Окружность первого рода мы будем называть в дальнейшем объемлющей диаметральной окружностью, а окружность второго рода — объемлющей трехточечной окружностью. Очевидно, что операции I, II, III дают вместе с тем и способ получения объемлющей окружности того или другого вида.

4. Представим себе все окружности, диаметрами которых служат отрезки, соединяющие все возможные пары точек нашей совокупности, а также все окружности, проходящие через каждые три точки совокупности[1]); очевидно, что окружностей обоих видов будет конечное число и, разумеется, не все эти окружности будут объемлющими. Но среди них найдутся и только что определенные диаметральные и трехточечные объемлющие окружности. Отсюда мы тотчас же заключаем, что как тех, так и других может существовать тоже только конечное число.

Отыщем их и путем конечного числа сравнений определим наименьшую из них; назовем ее k. Тогда k будет *наименьшей из всех вообще возможных объемлющих окружностей*, так как она найдена в результате сравнения со всеми объемлющими диаметральными и трехточечными окружностями и на основании положений I — III должна быть меньше всех объемлющих окружностей, не являющихся ни диаметральными, ни трехточечными. Эта окружность определяется вполне однозначно, ибо если бы существовала еще вторая объемлющая окружность k' того же размера, что и k (рис. 63), то совокупность H лежала бы одновременно и в k, и в k', а значит, и в области, общей обоим кругам k и k', т. е. в дуговом

[1]) И тех и других окружностей вместе может быть, самое бо́льшее,
$$\frac{n(n-1)}{1\cdot 2}+\frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}.$$

двуугольнике, около которого можно было бы описать еще меньший объемлющий круг k^*, а это противоречило бы тому условию, что круг k является наименьшим. Такую однозначно определенную наименьшую объемлющую окружность точечной совокупности H мы назовем **замыкающей окружностью** точечной совокупности.

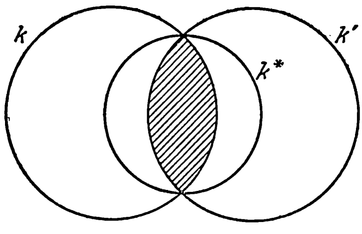


Рис. 63.

В замыкающей окружности k не может быть свободной дуги, большей чем полуокружность,— иначе, в силу положения III, k не могла бы быть наименьшей объемлющей.

5. Мы покажем теперь, что *радиус замыкающей окружности не может быть больше, чем* $\frac{d\sqrt{3}}{3}$. С этой целью отыщем на окружности k такую пару точек H, расстояние между которыми имеет наибольшую величину δ, не превосходящую, конечно, диаметра точечной совокупности d.

Разберем отдельно два случая. Может оказаться, **во-первых**, что две точки H лежат на противоположных концах диаметра окружности k. В таком случае диаметр $2r$ окружности k равен $\delta \leqslant d$, откуда $r \leqslant \frac{d}{2}$, и значит, во всяком случае $r < \frac{d}{3}\sqrt{3}$.

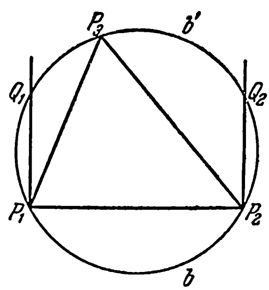


Рис. 64.

Рассмотрим теперь **второй** случай, когда на окружности k нет таких двух точек из H, которые лежали бы на противоположных концах диаметра. Тогда на замыкающей окружности k найдем наибольшую свободную дугу b (если имеется несколько равных, возьмем одну из них) и обозначим ее конечные точки, принадлежащие H, через P_1 и P_2 (рис. 64). Эта дуга b должна быть обязательно **меньше полуокружности**, ибо, как уже сказано выше, свободной дуги, большей, чем полуокружность, на k быть не может, а тот случай, когда b равна как раз полуокружности и, следовательно, точки P_1 и P_2 лежат на одном диаметре k, уже разобран. Стянем дугу b хордой и проведем в концах хорды два перпендикуляра, которые пересекут нашу окружность соответственно в точках Q_1 и Q_2. Точки Q_1 и Q_2 определят на окружности k дугу b', противолежащую b и ей конгруэнтную. Обе эти точки Q_1 и Q_2 не принадлежат H, ибо Q_1 диаметрально противоположна P_2, а Q_2 диаметрально противоположна P_1, а таких пар точек в рассматриваемом случае в H быть не может. Дуга b' не может быть свободной, ибо в противном случае, поскольку ее конечные точки P_1 и P_2 не принадлежат H, она составляла бы часть еще большей свободной дуги. Но это невозможно, так как свободную дугу b мы выбрали наибольшей.

Значит, дуга b' между точками Q_1 и Q_2 содержит по крайней мере одну точку P_3, принадлежащую H. Точки P_1, P_2, P_3 образуют остроугольный треугольник. Действительно, углы при P_1 и P_2 — острые, потому что они меньше построенных при этих точках прямых углов, а угол при P_3 — острый потому, что дуга b, на которую он опирается, меньше полуокружности (меньшей дуге соответствует меньший вписанный угол, а по теореме Фалеса только опирающийся на полуокружность вписанный угол равен прямому).

Из трех дуг, на которые делится окружность k точками P_1, P_2, P_3, должна быть по крайней мере одна дуга, равная или бóльшая трети окружности, которая, однако, ввиду того, что треугольник P_1, P_2, P_3 остроугольный, будет меньше полуокружности. Хорда этой дуги P_iP_j должна быть во всяком случае не меньше хорды третьей части окружности или, другими словами, не меньше стороны s вписанного в круг k правильного треугольника. Вместе с тем, так как наибольшее значение, которое может получить P_iP_j, равно диаметру d, то тем самым установлено, что $s \leqslant d$.

Радиус окружности, описанной около равностороннего треугольника со стороной s, равен (согласно п. 2) $r = \frac{s}{3}\sqrt{3}$. Так как $s \leqslant d$, то мы имеем для радиуса замыкающей окружности k неравенство

$$r \leqslant \frac{d}{3}\sqrt{3},$$

что и требовалось доказать.

6. Теперь нам легко понять, почему верхняя граница $\frac{d\sqrt{3}}{3}$ для радиуса замыкающей окружности точечной совокупности диаметра d не может быть понижена. Действительно, окружность, описанная около равностороннего треугольника со стороной d и имеющая, следовательно, радиус, равный $\frac{d\sqrt{3}}{3}$, является замыкающей окружностью точечной совокупности, состоящей из трех вершин этого треугольника. Ведь, согласно п. 4, замыкающую окружность можно искать или среди диаметральных, или среди трехточечных. Но каждая из

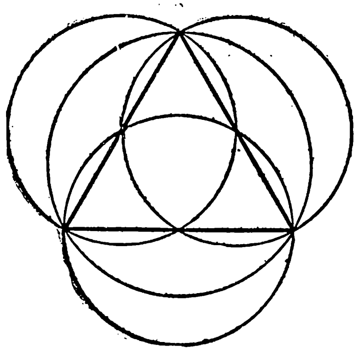


Рис. 65.

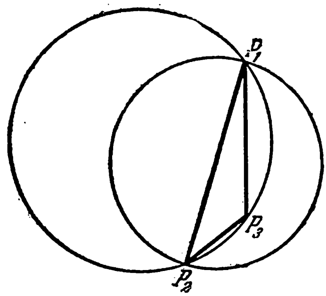


Рис. 66.

трех диаметральных окружностей содержит лишь по две вершины равностороннего треугольника (рис. 65); следовательно, в качестве единственно возможной минимальной объемлющей окружности остается т р е х т о ч е ч н а я, т. е. окружность, описанная около равностороннего треугольника. Впрочем, хотя интуитивно ясно, что окружность, описанная около равностороннего треугольника, является замыкающей для совокупности его вершин, логически это не является вполне само собой разумеющимся, ибо окружность, описанная около тупоугольного треугольника, уже не является замыкающей для точечной совокупности его вершин; замыкающей окружностью здесь будет диаметральная окружность, построенная на большей стороне треугольника, как на диаметре (рис. 66) [[52]].

17. ПРИБЛИЖЕННОЕ ВЫРАЖЕНИЕ ИРРАЦИОНАЛЬНЫХ ЧИСЕЛ ЧЕРЕЗ РАЦИОНАЛЬНЫЕ

Еще древним было известно, что число π, измеряющее площадь круга радиуса 1, приблизительно равно $\frac{22}{7}$, а $\sqrt{2}$ почти равен $\frac{7}{5}$. Но в чем, собственно, заключается ясный математический смысл подобных утверждений? Какое значение имеют слова «приблизительно», «почти», которые, собственно говоря, не принадлежат к словарю точной математики?

1. Если дано какое-либо число w, то можно указать такие дроби или, как говорят математики, такие рациональные числа, которые будут сколь угодно близки к этому числу w. Если, например, известно, что разложение числа π в бесконечную десятичную дробь дает $3{,}14159\ldots$, то для числа π можно написать следующие дроби:

$$3{,}1=\frac{31}{10},\quad 3{,}14=\frac{314}{100},\quad 3{,}141=\frac{3141}{1000},\ \ldots,$$

все менее и менее отличающиеся от π. Первая дробь отличается от π, очевидно, меньше чем на $\frac{1}{10}$ (так как $3{,}2=\frac{32}{10}$ является уже значением с избытком), вторая — менее чем на $\frac{1}{100}$ и т. д. Подобным же образом и всякое другое число, разложение которого в десятичную дробь нам известно, можно выразить с любой степенью точности некоторой дробью.

Это утверждение страдает тем недостатком, что оно слишком тесно связано со случайностью нашей десятеричной системы счисления, выбор которой связан со строением нашего тела, но не с природой самой математики. Мы можем, однако, легко освободиться от этой случайности выбора системы счисления и высказать следующее утверждение, в котором 10, 10^2 и т. д. заменены степенями произвольного числа n:

Теорема 1. *Если w — произвольное число, а n — произвольное целое число, то существует такая рациональная дробь $\frac{m}{n}$ со знаменателем n, которая отличается от w*

меньше чем на $\frac{1}{n}$:

$$0 \leqslant w - \frac{m}{n} < \frac{1}{n}.$$

Если, например, $w = \sqrt{2}$, $n = 5$, то w заключено между 1 и 2 в одном из 5 интервалов между числами

$$1, \frac{6}{5}, \frac{7}{5}, \frac{8}{5}, \frac{9}{5}, 2, \tag{1}$$

каждый из которых равен $\frac{1}{5}$. Без дальнейших вычислений ясно, что то из этих чисел, которое лежит непосредственно перед $\sqrt{2}$, представляет собой дробь со знаменателем 5, отличающуюся от $\sqrt{2}$ меньше чем на $\frac{1}{5}$. Точно так же и при произвольно данном w нужно рассмотреть числа

$$g, g+\frac{1}{n}, g+\frac{2}{n}, \ldots, g+\frac{n-1}{n}, g+1, \tag{2}$$

где g есть ближайшее целое число, меньшее w, и среди них отыскать такое, которое стоит непосредственно перед w или равно ему и, следовательно, отличается от w меньше чем на $\frac{1}{n}$. Пусть это число равно $g+\frac{s}{n}$, тогда будут справедливы неравенства

$$0 \leqslant w - \left(g+\frac{s}{n}\right) < \frac{1}{n}. \tag{3}$$

Этим самым теорема 1 доказана для произвольного w.

Вычислим теперь, между какими числами ряда (1) лежит в действительности значение $\sqrt{2}$. Для удобства вычислений мы освободимся от знаменателей, увеличив все числа ряда (1) в 5 раз, и зададимся вопросом, на какой именно интервал между пятикратными значениями чисел ряда (1), т. е. между числами

$$5, 6, 7, 8, 9, 10,$$

падает пятикратное значение $\sqrt{2}$, т. е. $5\sqrt{2} = \sqrt{25}\sqrt{2} = \sqrt{50}$, или, иными словами, какое целое число является ближайшим к $\sqrt{50}$. Так как $49 < 50 < 64$, то $7 < \sqrt{50} < 8$;

таким образом, искомое число есть 7:

$$0<\left(\sqrt{2}-\frac{7}{5}\right)<\frac{1}{5}. \tag{4}$$

Точно так же и общее доказательство высказанной теоремы будет значительно проще, если мы будем вместо w рассматривать nw и отыскивать наибольшее целое число, не превосходящее nw. Обозначив такое число через m, мы будем иметь:

$$m \leqslant nw < m+1,$$

т. е.

$$0 \leqslant nw - m < 1.$$

Если теперь перейти к n-м долям, то требуемое соотношение получится у нас непосредственно:

$$0 \leqslant w - \frac{m}{n} < \frac{1}{n}.$$

2. В том факте, что в б л и з и чисел $\sqrt{2}$ и π существуют дроби (рациональные числа), нет, следовательно, ничего особенного. Какой же в таком случае смысл имеют такие, например, высказывания, что π почти равно $\frac{22}{7}$, а $\sqrt{2}$ почти равен $\frac{7}{5}$? Смысл здесь в том, что $\frac{7}{5}$ намного ближе к $\sqrt{2}$, чем это обеспечивается только что доказанной теоремой. Действительно, имеет место более точная оценка:

$$\frac{7}{5}<\sqrt{2}<\frac{17}{12},$$

и следовательно, $\sqrt{2}$ отличается от $\frac{7}{5}$ во всяком случае меньше, чем $\frac{17}{12}$ отличается от $\frac{7}{5}$, т. е.

$$\sqrt{2}-\frac{7}{5}<\frac{17}{12}-\frac{7}{5}=\frac{1}{60},$$

между тем как, в силу формулы (4), получается лишь, что эта разность должна быть меньше $\frac{1}{5}$.

Подобное же значение имеет и утверждение Архимеда относительно числа π. Данная им более точная оценка

$$3+\frac{10}{71}<\pi<3+\frac{1}{7}$$

утверждает, что

$$\frac{22}{7}-\pi<\left(3+\frac{10}{70}\right)-\left(3+\frac{10}{71}\right)=\frac{10}{70}-\frac{10}{71}=\frac{10}{70\cdot 71}=\frac{1}{497},$$

тогда как по нашей теореме получаем для такой разности предел, равный $\frac{1}{7}$.

Но и выражение «намного ближе» также не входит в словарь математика. Докажем теперь теорему, вносящую в эти вопросы полную определенность.

Теорема 2. *Если w — иррациональное число, а N — какое-либо целое число, то существует некоторая дробь $\frac{m}{n}$, знаменатель которой не превосходит N и которая отличается от w меньше чем на $\frac{1}{Nn}$. Существует поэтому бесконечно много дробей $\frac{m}{n}$, отличающихся от w меньше чем на $\frac{1}{n^2}$.*

Для двух вышерассмотренных числовых примеров это значит, что $\sqrt{2}-\frac{7}{5}$ меньше, чем $\frac{1}{5^2}=\frac{1}{25}$, и $\frac{22}{7}-\pi$ меньше, чем $\frac{1}{7^2}=\frac{1}{49}$. Хотя в обоих этих случаях действительность далеко превосходит границы, устанавливаемые этой новой теоремой, последняя все же дает значительное уточнение сравнительно с тривиальной первой теоремой.

Для доказательства этой второй теоремы рассмотрим не только Nw и ближайшее к нему меньшее число, но также ряд чисел

$$w,\ 2w,\ 3w,\ \ldots,\ Nw;$$

обозначим ближайшие к ним и меньшие их целые числа соответственно через

$$g_1,\ g_2,\ g_3,\ \ldots,\ g_N,$$

так что

$$0<w-g_1<1,\quad 0<2w-g_2<1,\ \ldots,\ 0<Nw-g_N<1.$$

Лучше всего мы разберем это на примере ($w=\sqrt{2}, N=13$):

$$
\begin{aligned}
\sqrt{2} &= 1{,}414\ldots = 1+0{,}414\ldots\\
2\sqrt{2} &= 2{,}828\ldots = 2+0{,}828\ldots\\
3\sqrt{2} &= 4{,}242\ldots = 4+0{,}242\ldots\\
4\sqrt{2} &= 5{,}656\ldots = 5+0{,}656\ldots\\
5\sqrt{2} &= 7{,}071\ldots = 7+0{,}071\ldots\\
6\sqrt{2} &= 8{,}485\ldots = 8+0{,}485\ldots\\
7\sqrt{2} &= 9{,}899\ldots = 9+0{,}899\ldots\\
8\sqrt{2} &= 11{,}313\ldots = 11+0{,}313\ldots\\
9\sqrt{2} &= 12{,}727\ldots = 12+0{,}727\ldots\\
10\sqrt{2} &= 14{,}142\ldots = 14+0{,}142\ldots\\
11\sqrt{2} &= 15{,}556\ldots = 15+0{,}556\ldots\\
12\sqrt{2} &= 16{,}970\ldots = 16+0{,}970\ldots\\
13\sqrt{2} &= 18{,}384\ldots = 18+0{,}384\ldots
\end{aligned}
$$

Из 13 избытков над ближайшими целыми числами пятый, очевидно, является наименьшим: $5\sqrt{2}-7=0{,}071\ldots$, а двенадцатый наибольшим: $12\sqrt{2}=16+0{,}970\ldots$ или $=17-0{,}030\ldots$, откуда $17-12\sqrt{2}=0{,}030\ldots$

Предположим, что эти 13 избытков, то возрастающие, то убывающие в нашей таблице, будут расположены в порядке возрастания их по величине; тогда в промежутке между числами 0 и 1 мы будем иметь 13 чисел, разбивающих этот промежуток на 14 интервалов. Конечно, эти интервалы не будут равными, и их величина будет несколько отклоняться от $\frac{1}{14}$. Но по крайней мере один из них будет меньше $\frac{1}{14}$, ибо если бы все они были $\geqslant\frac{1}{14}$, то дали бы в сумме, самое меньшее, $\frac{14}{14}=1$, и притом эта сумма не превышала бы 1 лишь в том единственном случае, когда каждый интервал был бы равен $\frac{1}{14}$. Стóит только о д н о м у интервалу быть больше $\frac{1}{14}$, и наша сумма превысит 1. Но если бы все они были равны $\frac{1}{14}$, то и $\sqrt{2}$, как одно из наших 13 чисел, имел бы избыток, кратный $\frac{1}{14}$: $\sqrt{2}=1+\frac{m}{14}$, т. е. $\sqrt{2}$ был бы рациональным числом.

Так как w иррационально по предположению, а об иррациональности $\sqrt{2}$ мы знаем из темы 4 (стр. 35), то можно заключить, что хотя бы один из интервалов будет непременно меньше $\frac{1}{14}$; для нас не имеет никакого значения, между какими кратными $\sqrt{2}$ лежит этот интервал. Пусть $a\sqrt{2}-g_a=r_a$ — нижняя, а $b\sqrt{2}-g_b=r_b$ — верхняя граница этого интервала, так что

$$0<r_b-r_a=(b\sqrt{2}-g_b)-(a\sqrt{2}-g_a)<\frac{1}{14},$$

откуда, следовательно,

$$0<(b-a)\sqrt{2}-(g_b-g_a)<\frac{1}{14}.$$

Так как здесь a и b — произвольные числа ряда 0, 1, ..., 13, то абсолютное значение разности $b-a$ есть также одно из этих чисел, т. е. $-13\leqslant b-a\leqslant 13$. Этим самым мы доказали, что для одного из тринадцати чисел, кратных $\sqrt{2}$, а именно для $(b-a)\sqrt{2}$, или, если $b-a$ отрицательно, для $(a-b)\sqrt{2}$, избыток над ближайшим меньшим целым числом или недостаток до ближайшего большего целого числа меньше $\frac{1}{14}$. Обозначив абсолютное значение разности $b-a$ через n и g_b-g_a — через m, имеем:

$$-\frac{1}{14}<n\sqrt{2}-m<\frac{1}{14}, \quad \text{где} \quad n\leqslant 13,$$

а разделив на n, получим:

$$-\frac{1}{14n}<\sqrt{2}-\frac{m}{n}<\frac{1}{14n}.$$

Если вместо $\sqrt{2}$ мы возьмем w, вместо 13 возьмем N и вместо 14 возьмем $N+1$, то точно таким же образом получим, что существует такое $n\leqslant N$, для которого справедливы неравенства

$$-\frac{1}{(N+1)n}<w-\frac{m}{n}<\frac{1}{(N+1)n}, \tag{5}$$

а это как раз те неравенства, которые мы хотели доказать. Этим самым доказана первая часть нашей теоремы 2.

Так как $n \leqslant N$, то из (5) сразу же получаем:

$$-\frac{1}{n^2} < w - \frac{m}{n} < \frac{1}{n^2}. \tag{6}$$

Таким образом, мы нашли дроби, отличающиеся от w меньше чем на $\frac{1}{n^2}$. В неравенствах (6) число N выпало, но оно было использовано для нахождения дроби $\frac{m}{n}$, знаменатель которой n обладает свойством $n \leqslant N$.

В рассмотренном числовом примере, пока N будет равно одному из чисел 5, 6, 7, ..., 11, число n будет равно 5, а именно

$$0 < 5\sqrt{2} - 7 < \frac{1}{5}, \quad 0 < \sqrt{2} - \frac{7}{5} < \frac{1}{5^2}.$$

Когда же N достигнет значения 12, n станет равно 12:

$$0 < 17 - 12\sqrt{2} < \frac{1}{12}, \quad 0 < \frac{17}{12} - \sqrt{2} < \frac{1}{12^2}.$$

и т. д.

Вторая часть теоремы 2 для данного иррационального числа w утверждает, что существует даже бесконечно много дробей $\frac{m}{n}$, обладающих свойством, выраженным в неравенствах (6). Мы покажем только, что для любой дроби $\frac{m}{n}$, удовлетворяющей неравенствам (6), можно указать дробь $\frac{m'}{n'}$, стоящую к w еще ближе, чем $\frac{m}{n}$, и также удовлетворяющую (6). Так как разность $w - \frac{m}{n}$ не может быть равна нулю (ибо w иррационально, т. е. отличается от любой дроби), то она отличается от нуля на некоторое определенное число. Поэтому существует дробь $\frac{1}{N'}$ с достаточно большим знаменателем, лежащая к нулю ближе, чем $w - \frac{m}{n}$, т. е.

$$0 < \frac{1}{N'} < w - \frac{m}{n} \quad \text{или} \quad 0 < \frac{1}{N'} < \frac{m}{n} - w, \tag{7}$$

смотря по тому, положительна или отрицательна разность $w - \frac{m}{n}$. Для этого же, числа N' (так же, как и для числа N) мы можем, согласно доказанному, найти дробь $\frac{m'}{n'}$, так что при условии $n' \leqslant N'$ будут выполняться неравенства, аналогичные (5):

$$-\frac{1}{(N'+1)\,n'} < w - \frac{m'}{n'} < \frac{1}{(N'+1)\,n'}, \tag{8}$$

и следовательно, тем более мы будем иметь:

$$-\frac{1}{n'^2} < w - \frac{m'}{n'} < \frac{1}{n'^2}.$$

Здесь $\frac{m}{n}$ несомненно отличается от $\frac{m'}{n'}$, ибо, согласно (8),

$$-\frac{1}{N'+1} < w - \frac{m'}{n'} < \frac{1}{N'+1},$$

т. е. $w - \frac{m'}{n'}$ отличается от нуля *меньше* чем на $\frac{1}{N'+1}$, в то время как, в силу (7), $w - \frac{m}{n}$ отличается от нуля *больше* чем на $\frac{1}{N'}$. Поэтому $\frac{m'}{n'}$ ближе к w, чем $\frac{m}{n}$; при этом $\frac{m}{n}$ и $\frac{m'}{n'}$ различны.

3. Из сказанного следует, что *не существует последней дроби* $\frac{m}{n}$, *удовлетворяющей неравенствам* (6): за всякой такой дробью имеются следующие, значит, и бесконечное множество, как это утверждается нашей теоремой 2. В то время как в тривиальной теореме 1 допустим любой знаменатель, здесь допускаются лишь некоторые знаменатели, обладающие свойством (6). Назовем такие знаменатели хорошими. При приближении к $\sqrt{2}$ хорошими знаменателями являются 2, 5, 12 и, как это может быть получено из дальнейших вычислений, 29, 70, 169, ...

Для π хорошим знаменателем является 7, даже еще значительно более хорошим, чем это гарантируется теоремой 2, ибо $\frac{22}{7} - \pi < \frac{1}{497}$, а не только $< \frac{1}{7^2} = \frac{1}{49}$. Возникает вопрос, нет ли лучших знаменателей и для всякого числа w, например таких, чтобы

$$-\frac{1}{n^3} < w - \frac{m}{n} < \frac{1}{n^3}.$$

Мы покажем теперь, что такое улучшение теоремы 2 возможно не для всякого числа w. А именно мы покажем, что, в частности для $w = \sqrt{2}$, *какова бы ни была дробь* $\frac{m}{n}$, *ее отклонение от* $\sqrt{2}$ *всегда превышает* $\frac{1}{3n^2}$.

Рассмотрим сначала случай, когда $\frac{m}{n} > \sqrt{2}$. При этом мы можем считать $\frac{m}{n} < 1{,}55$. Так как $\sqrt{2} < 1{,}45$, всякая дробь $\frac{m}{n}$, превышающая 1,55, отличается от $\sqrt{2}$ больше чем на 0,1, а $\frac{1}{3n^2}$, начиная с $n = 2$, уже меньше 0,1; таким образом, наше утверждение для $\frac{m}{n} > 1{,}55$, наверное, справедливо. Если теперь $\frac{m}{n}$ лежит между $\sqrt{2}$ и 1,55, то

$$\left(\frac{m}{n}\right)^2 - 2 = \frac{m^2 - 2n^2}{n^2} = \frac{g}{n^2},$$

где g — целое положительное число, и значит, во всяком случае

$$\left(\frac{m}{n}\right)^2 - 2 \geqslant \frac{1}{n^2}.$$

Из соотношения $a^2 - b^2 = (a+b)(a-b)$ следует:

$$\left(\frac{m}{n} + \sqrt{2}\right)\left(\frac{m}{n} - \sqrt{2}\right) \geqslant \frac{1}{n^2},$$

и отсюда

$$\frac{m}{n} - \sqrt{2} > \frac{1}{n^2} \cdot \frac{1}{\frac{m}{n} + \sqrt{2}}.$$

Если теперь $\frac{m}{n} < 1{,}55$, то $\frac{m}{n} + \sqrt{2} < 1{,}55 + 1{,}45 = 3$, значит, обратная величина этой суммы меньше $\frac{1}{3}$, и следовательно,

$$\frac{m}{n} - \sqrt{2} > \frac{1}{3n^2},$$

что и требовалось доказать.

Если $0 < \frac{m}{n} < \sqrt{2}$, то, исходя из соотношения

$$2 - \frac{m^2}{n^2} = \frac{g}{n^2} \geqslant \frac{1}{n^2},$$

можно прийти к тому же результату:

$$\sqrt{2} - \frac{m}{n} \geqslant \frac{1}{n^2} \frac{1}{m/n + \sqrt{2}} \geqslant \frac{1}{n^2} \frac{1}{2\sqrt{2}} > \frac{1}{3n^2}.$$

И в том и в другом случае наше утверждение доказано [53].

18. ШАРНИРНЫЕ ПРЯМОЛИНЕЙНО-НАПРАВЛЯЮЩИЕ МЕХАНИЗМЫ

1. Паровая машина Джемса Уатта снабжена замечательным вспомогательным механизмом, так называемым параллелограммом Уатта, схематический вид которого представлен на рис. 67. Прибор состоит из стержней, соединенных между собой подвижными шарнирами C, D, E, F. В точках A и B стержни закреплены, однако с возможностью вращения. Оси всех шарниров нужно представить себе перпендикулярными к плоскости чертежа[1]). В точке F присоединен поршневой шток паровой машины. Механизм должен сообщать концевой точке F поршневого штока прямолинейное движение, необходимое для того, чтобы поршень не шатался и не застревал в цилиндре. Математическое исследование показывает, что хотя точка F шарнирного механизма Уатта не будет двигаться по прямой, тем не менее в достаточно малой области это движение все же *можно рассматривать как приближенно прямолинейное*, так что технические тре-

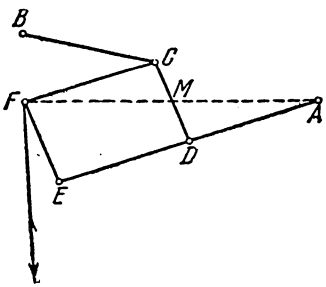


Рис. 67.

[1]) В схематических чертежах шарнирных механизмов неподвижные оси вращения (шарниры) обозначаются обычно черными кружками, подвижные — светлыми кружками; пунктирные линии — вспомогательные, и никаких стержней они не изображают.

бования в отношении прямолинейности движения будут в достаточной мере удовлетворены этим механизмом.

Параллелограмм $CDEF$, благодаря которому механизм получил свое наименование, не является его самой существенной частью. Он служит лишь для того, чтобы увеличить пригодную часть траектории. Существенной частью механизма нужно считать скорее систему стержней $ADCB$, в которой середина M среднего стержня DC при не слишком больших ее удалениях описывает почти прямолинейный путь. Если далее $AD = DE = CF = BC$ и $DC = EF$ (стержень AE в шарнире D не сгибается), то точки A, M и F всегда лежат на одной прямой, в силу подобия треугольников ADM и AEF. Далее, так как AF на том же самом основании вдвое больше AM, то точка F описывает траекторию, подобную траектории точки M, только в удвоенном масштабе. Можно было бы аналогичным образом увеличить траекторию точки M в k раз, если взять параллелограмм таким, чтобы

$$AD : AE = DM : EF = 1 : k.$$

Если, как и раньше, положить здесь $AD = BC$, то середина M стержня CD попрежнему будет совершать приблизительно прямолинейное движение (рис. 68).

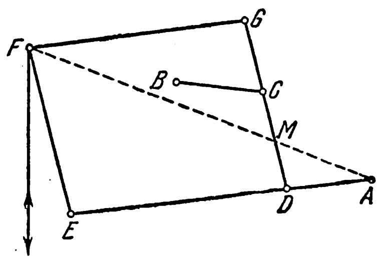


Рис. 68.

При оценке параллелограмма Уатта как прямолинейно-направляющего механизма[1]) дело сводится, таким образом, к исследованию движения точки M. В том, что оно не строго прямолинейное, легко можно убедиться из рассмотрения некоторых крайних положений системы стержней. Но мы здесь не будем входить в детальное изучение этого сложного движения, поскольку для нашей задачи оно значения не имеет.

2. Проблема выпрямления траектории имела большое значение для машиностроения, и ей было посвящено поэтому большое количество работ. Метод, разработанный Уаттом,

---

[1]) Приборы, предназначенные для преобразования различных видов движения в прямолинейное, называются прямолинейно-направляющими, или прямилами. (*Прим. перев.*)

в наши дни уже не находит применения: даже самый поверхностный осмотр наружных частей механизма локомотива убеждает нас в том, что в настоящее время предпочитается иной метод прямления — с помощью крейцкопфа, движущегося между двумя параллелями. Возможно, что неверные представления о значении трения побудили ранних машиностроителей отказаться в свое время от идеи крейцкопфа и прийти вместо того к шарнирному прямилу.

Проблема построения прямолинейно-направляющего механизма заинтересовала не только техников: она привлекла к себе также внимание математиков. Последние, конечно, поставили проблему во всей ее строгости: найти такой шарнирный механизм, в котором один шарнир теоретически точно движется по прямой. Над параллелограммом Уатта и его усовершенствованиями работал великий русский математик П. Л. Чебышев (1821—1894); но задачу отыскания точного прямолинейно-направляющего механизма ему, однако, разрешить не удалось [1]). После многократных тщетных попыток найти решение этой проблемы в математических кругах в середине прошлого столетия было высказано даже сомнение в возможности точного решения проблемы прямления посредством шарнирных механизмов.

Только в 1864 г. французский генерал Поселье (Peaucellier) изобрел аппарат, с помощью которого можно было производить выпрямление, а также и различные другие операции; аппарат был назван инверсором [[54]]. Впоследствии, как это часто случается в истории открытий, удалось найти еще целый ряд других решений проблемы, причем прямая линия получается в них лишь как частный случай обширного класса кривых, построение которых возможно с помощью шарнирных механизмов.

3. Рассмотрим сначала шарнирный механизм Поселье. Он назван инверсором на том основании, что с его помощью можно выполнять инверсию. Инверсия есть частный вид отображения плоскости на себя, сущность которого будет изложена ниже.

---

[1]) Чебышев исследовал соотношения, которым должны удовлетворять размеры частей прибора, чтобы движение некоторой точки его наименее уклонялось от прямой. На основе этого развились его общие исследования о функциях, наименее уклоняющихся от нуля. (*Прим. перев.*)

Под отображением понимается установление некоторого соответствия, в силу которого каждой точке данной фигуры соответствует некоторая другая точка — ее отображение или образ. Простейшими отображениями являются, например: 1) зеркальное отражение от прямой, которым мы неоднократно пользовались в темах 5 и 6; 2) параллельный перенос, при котором любой точке ставится в соответствие в качестве ее образа такая другая точка, которая получается из первой в результате смещения на отрезок определенной длины в определенном направлении; 3) поворот плоскости вокруг точки на заданный угол; 4) растяжение плоскости (гомотетия), при котором любая точка плоскости сдвигается по лучу, выходящему из центра и проходящему через эту точку, так что ее расстояние от центра растяжения увеличивается в некотором определенном, постоянном для данного отображения отношении $1:\lambda$.

Точка-образ не всегда отлична от ее оригинала. При параллельном переносе каждая точка действительно отображается в другую, отличную от нее точку. При вращении же и растяжении центры переходят сами в себя. При зеркальном отражении любая точка той прямой, которая служит зеркалом, также переходит сама в себя.

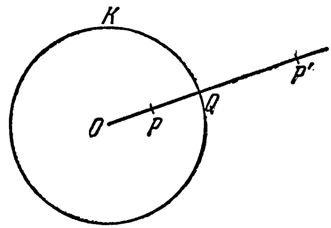


Рис. 69.

Приведенные примеры предназначены для уяснения понятия отображения. Инверсия представляет собой отображение, которое можно охарактеризовать следующим образом. Дан круг с центром (центр инверсии) в точке O (рис. 69); *чтобы для некоторой отличной от O точки P найти образ P', проводят через P луч из O, пересекающий окружность в Q, и точку P' определяют на этом луче таким образом, чтобы выполнялось соотношение*

$$OP:OQ = OQ:OP'.$$

Положим, что радиус данной нам окружности равен a; обозначим далее OP через r, OP' — через r'; тогда характеризующее инверсию соотношение можно записать следующим образом:

$$rr' = a^2.$$

Из этого определения инверсии сразу же заключаем, что образ точки, лежащей внутри круга инверсии, находится вне круга, и обратно; любая точка окружности инверсии является в то же время и своим собственным образом. Поэтому инверсию называют также зеркальным отражением от окружности. С обыкновенным зеркальным отражением инверсия имеет то общее важное свойство, что и здесь и там образ образа есть опять первоначальная точка.

4. Об отображении посредством инверсии отдельной точки по существу уже все сказано в самом определении. Новые вопросы возникают лишь при отображении кривых. Мы лишены здесь возможности предпринимать подробное изучение этих вопросов, но в качестве важнейшего свойства отметим, что при инверсии прямые и окружности переходят в прямые или окружности [55]. В дальнейшем нам потребуется только одно предложение, а именно: *образ окружности, проходящей через центр инверсии, есть прямая*. Эту теорему мы и хотим доказать.

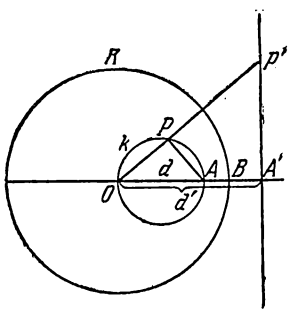


Рис. 70.

Пусть k будет окружность диаметра $OA = d$, проходящая через центр инверсии O, и K — окружность инверсии с радиусом a (рис. 70). Образ A' точки A лежит на продолжении диаметра OA, причем расстояние $OA' = d'$ удовлетворяет соотношению $dd' = a^2$. Докажем, что перпендикуляр, восставленный к OA' в точке A', и есть образ k. Для этого нам нужно показать, что любая проходящая через O прямая пересекает окружность k в точке P, а прямую — в точке P', причем для этих точек имеет место соотношение $OP \cdot OP' = a^2$. Соединив точки A и P вспомогательной прямой AP, мы убедимся в том, что оба прямоугольных треугольника OPA и $OA'P'$ с общим углом при O подобны. Поэтому мы имеем право написать пропорцию

$$OP : OA = OA' : OP',$$

или, положив $OP = r$, $OP' = r'$:

$$rr' = dd'.$$

Но $dd'=a^2$, следовательно,

$$rr'=a^2,$$

а это и есть как раз то соотношение между точкой и ее образом, которое устанавливается инверсией.

Для дальнейшего напомним одну известную теорему из элементарной геометрии (рис. 71). Если из какой-либо точки A провести несколько лучей (секущих) к окружности, то произведение расстояний AP_1 и AP_1' от точки A до точек P_1 и P_1' пересечения луча с окружностью есть величина постоянная для всех лучей, выходящих из A. Если $AP_1=s_1$, $AP_1'=s_1'$, $AP_2=s_2$, $AP_2'=s_2'$, то наша теорема утверждает, что $s_1s_1'=s_2s_2'$. Это следует непосредственно из подобия треугольников AP_1P_2' и AP_2P_1', имеющих общий угол при A и равные углы при P_1' и P_2' (как вписанные и опирающиеся на одну и ту же дугу).

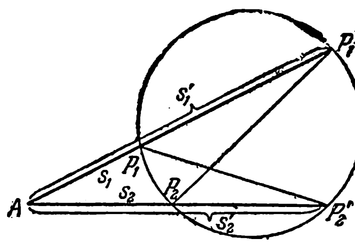


Рис. 71.

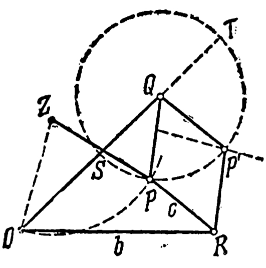


Рис. 72.

5. Приняв во внимание эти предварительные замечания, переходим к рассмотрению инверсора Поселье (рис. 72). Он состоит из шарнирного ромба $PQP'R$ со стороной c, в противоположных вершинах которого Q и R укреплены посредством шарниров два равных стержня длиной b; эти стержни другими своими концами закреплены в неподвижной точке O, вокруг которой они могут вращаться. При этом $b>c$. Весь механизм состоит, таким образом, из 6 стержней: OQ, OR, PQ, PR, $P'Q$, $P'R$. В силу симметрии прибора, точки O, P и P' лежат всегда на одной прямой — оси симметрии прибора. Опишем теперь из Q как из центра окружность радиуса c; она пройдет через P и P'; стержень OQ и его продолжение пересекут окружность в точках S и T. Тогда OPP' и OST

будут секущими, к которым можно будет применить только что приведенную теорему элементарной геометрии. В данном случае эта теорема дает:

$$OP \cdot OP' = OS \cdot OT = (b - c)(b + c) = b^2 - c^2.$$

Следовательно, произведение $OP \cdot OP'$ есть величина постоянная. Положим $b^2 - c^2 = a^2$; тогда a будет катетом прямоугольного треугольника с гипотенузой b и другим катетом c. Но в таком случае уравнение $OP \cdot OP' = a^2$ показывает, что точка P' получена из P путем инверсии из центра O радиусом a.

Этим и объясняется наименование «инверсор». Прибор, следовательно, выполняет операцию инверсии в той части плоскости, которая достижима для точек P и P'. Чтобы воспользоваться им для проведения прямой, стоит только заставить P двигаться по окружности, проходящей через O, тогда, *в силу известных нам теперь свойств инверсии, P' будет совершать прямолинейное движение.* Но осуществить в шарнирном механизме круговое движение точки P очень легко. Для этого нужно лишь с помощью шарнира присоединить в P новый стержень ZP, вращающийся вокруг неподвижного центра Z.

Чтобы круговая траектория точки P проходила через O, длину стержня ZP нужно выбрать равной неизменному расстоянию OZ. Вследствие того, что у нас теперь присоединяется седьмой стержень PZ с неподвижным концом Z, точка P может двигаться только по дуге окружности, проходящей через O, в связи с чем P' должна будет двигаться по прямой (между прочим, перпендикулярной к OZ, см. рис. 72).

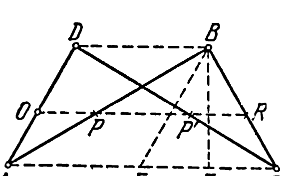


Рис. 73.

6. Были изобретены еще и другие инверсоры, так же, как и инверсор Поселье, решавшие проблему проведения прямой. Инверсор Гарта (Hart) (рис. 73) состоит всего лишь из 4 стержней. Он представляет собой а н т и п а р а л л е л о г р а м м, т. е. шарнирный четырехугольник с равными противоположными сторонами, расположенными, однако, не так, как в параллелограмме, а как показано на рис. 73, с пересечением одной пары противоположных сторон[1]).

На каждом из стержней можно отметить по такой точке O, P, P', R, что при некотором определенном положении антипараллелограмма они будут лежать на прямой, параллельной обеим «диагоналям», AC и BD[2]).

И вот оказывается, что они остаются на прямой, параллельной «диагоналям», и при всяком другом положении антипараллелограмма. Действительно, так как отрезки OP и DB параллельны, то мы имеем:

$$AO:OD=AP:PB. \tag{1}$$

Точно так же

$$AO:OD=CP':P'D, \tag{2}$$

$$AP:PB=CR:RB, \tag{3}$$

$$CR:RB=CP':P'D. \tag{4}$$

Очевидно, что эти соотношения между неизменными длинами стержней остаются в силе при всех возможных движениях шарнирного четырехугольника. Но в таком случае из (1) следует, что при любом положении шарнирного четырехугольника $OP \parallel DB$. На основании (2) $OP' \parallel AC$, а так как $AC \parallel BD$, то также и $OP' \parallel DB$. Но если OP и OP' параллельны одной и той же прямой DB, то точки O, P и P' должны лежать на одной прямой, а именно на прямой, параллельной DB и проходящей через точку O. Аналогичным рассуждением можно установить, что точки R, P', P также лежат на прямой, параллельной DB, проведенной через R. Мы получаем, таким образом, что все четыре точки O, P, P', R всегда лежат на прямой, параллельной DB. Далее, опять-таки вследствие параллельности трех прямых

$$OP:DB=AO:AD,$$

$$OP':AC=DO:DA.$$

---

[1]) Посредством вращений вокруг шарниров стержни можно переводить из одного положения в другое. Промежуточным положением между параллелограммом и антипараллелограммом является такое, при котором все стороны располагаются на одной прямой, причем ABC совпадает с ADC.

[2]) Обе диагонали AC и BD в антипараллелограмме всегда параллельны, ибо треугольники ACB и ACD симметричны и при общем основании AC имеют одну и ту же высоту.

Вместо этих пропорций можно написать:

$$OP \cdot AD = AO \cdot DB,$$
$$OP' \cdot DA = DO \cdot AC.$$

Перемножив правые и левые части этих двух равенств и разделив на AD^2, получим:

$$OP \cdot OP' = \frac{AO \cdot DO}{AD^2} \cdot AC \cdot DB. \tag{5}$$

Стоящие в правой части AO, DO, AD — величины неизменные, определенные раз навсегда масштабом прибора; напротив, AC и DB суть длины диагоналей, очевидно, зависящие от положения антипараллелограмма. И тем не менее *произведение этих переменных величин $AC \cdot DB$ есть величина постоянная.* Действительно, из рисунка, на котором мы проведем еще две вспомогательные прямые $BE \parallel DA$ и $BF \perp AC$, непосредственно видно, что

$$AC \cdot BD = (AF + FC)(AF - FC) = AF^2 - FC^2. \tag{6}$$

Теорема Пифагора в применении к треугольникам AFB и FCB дает нам теперь

$$AF^2 + FB^2 = AB^2,$$
$$FC^2 + FB^2 = CB^2,$$

откуда вычитанием получаем:

$$AF^2 - FC^2 = AB^2 - CB^2.$$

Подставляя это значение разности квадратов $AF^2 - FC^2$ в формулу (6), получаем:

$$AC \cdot BD = AB^2 - CB^2. \tag{7}$$

В правой части этого уравнения стоит величина постоянная. Сопоставляя формулы (5) и (7), имеем:

$$OP \cdot OP' = \frac{AO \cdot DO}{AD^2} \cdot (AB^2 - CB^2). \tag{8}$$

Справа мы здесь имеем величину п о с т о я н н у ю, т. е. зависящую только от масштаба самого прибора. Если теперь закрепить на плоскости точку O, в остальном предоставив антипараллелограмму возможность совершать любые движения, то уравнение (8) будет означать, что точки P и P', лежащие, как уже доказано, на одной прямой с O, получаются инверсией друг из друга, причем квадрат радиуса круга инверсии определится несколько сложным, но допускающим элементарную конструкцию выражением в правой части формулы (8).

Вышеприведенные соображения доказывают, что инверсор Гарта действительно осуществляет инверсию. Если путем присоединения пятого стержня заставить точку P описывать дугу окружности, проходящей через O, то точка P', как мы уже знаем, будет двигаться по прямолинейному отрезку.

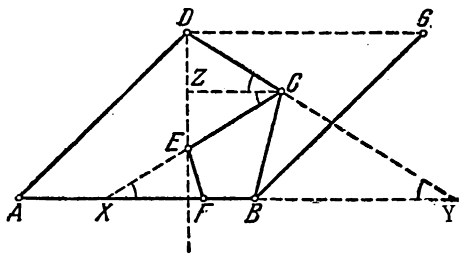


Рис. 74.

7. В настоящее время известны также шарнирные механизмы, осуществляющие движение по прямой и не использующие инверсию. Остроумную конструкцию представляет собой двойной ромбоид Кемпе (Kempe) (рис. 74). Р о м б о и д о м или д е л ь т о и д о м называется такой четырехугольник, в котором имеются две пары равных сторон, причем обе равные стороны в каждой паре примыкают одна к другой. В ромбоиде непременно имеется пара равных углов, именно тех, которые заключены между сторонами, принадлежащими к разным парам. Пусть $ABCD$ и $BCEF$ — два шарнирных ромбоида с одинаковым отношением сторон, причем малые стороны DC и CB большого ромбоида равны бо́льшим сторонам EC и CB малого ромбоида. Стержни располагаются так, что ромбоиды выпуклы. Вершина F малого ромбоида лежит на стороне большого, так что оба ромбоида имеют общий угол при B. Поэтому углы при D, E и B равны. Два выпуклых четырехугольника с пропорциональными сторонами, имеющие по одному равному углу, подобны. При любом движении шарнирного механизма, составленного из таких двух ромбоидов, оба ромбоида $ABCD$ и $BCEF$ остаются всегда п о д о б н ы м и.

Если две противоположные стороны ромбоида продолжить до их пересечения, то в точке пересечения они образуют некоторый угол. Для подобных ромбоидов такие углы EXF и CYB равны. Если через C провести прямую CZ, параллельную AB, то углы DCZ и CYB будут равны, как соответственные; далее, углы ZCE и EXF равны, как внутренние накрест лежащие при параллельных прямых ZC и AB. Значит, все четыре отмеченных на рисунке угла равны. Поэтому в равнобедренном (по условию) треугольнике CDE прямая CZ является биссектрисой угла при его вершине и, значит, должна быть перпендикулярна к его основанию DE. Так как CZ была проведена параллельно AB, то прямая, проходящая через точки D и E, должна быть перпендикулярна к AB.

Сделаем теперь точку D неподвижной, а прямую AB заставим перемещаться параллельно самой себе; тогда перпендикуляр, опущенный из D на AB, будет также неподвижен. Но так как точка E лежит на этом перпендикуляре, то для нее возможно только прямолинейное движение по этому перпендикуляру. Чтобы теперь осуществить необходимое параллельное перемещение AB, присоединим в точке B с помощью шарнира еще один стержень BG, по длине равный AB и AD, т. е. $BG = AB = AD$. Точка G закрепляется так же, как и D, причем DG тоже берется равным AB, так что $ADGB$ есть ромб. Противоположные стороны в ромбе всегда параллельны, следовательно, AB при любом положении механизма параллельна неизменяемой прямой DG. Этим самым и разрешена задача прямолинейного движения точки E.

8. Двойной ромбоид Кемпе допускает еще и другой, особенно изящный способ выпрямления. Для этого его нужно изготовить в двух симметричных экземплярах: $ABCDEF$ и $A'B'CD'E'F'$.

Оба эти экземпляра соединяются таким образом, что вершина C у них получается общая, а стержни DCE' и $D'CE$ — сквозные (рис. 75). Биссектриса CZ угла DCE, как уже доказано, параллельна AB; аналогично и биссектриса CZ' угла $D'CE'$ параллельна $A'B'$. А так как углы DCE и $D'CE'$ по построению вертикальные, то обе биссектрисы CZ и CZ' лежат на одной прямой, параллельной прямым AB и $A'B'$. Следовательно, AB и $A'B'$ параллельны между собой при любом положении шарнирного механизма. Более того, $A'B'$ должна лежать всегда на продолжении AB. Так как параллельность AB и $A'B'$ мы уже доказали, достаточно показать, что точка C отстоит одинаково как от AB, так и от $A'B'$. Для этого мы докажем, что ромбоиды $ABCD$ и $A'B'CD'$ при любом положении механизма симметричны.

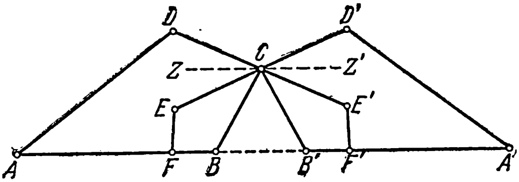


Рис. 75.

Действительно, в обеих половинах прибора углы DCE и $D'CE'$ равны, как вертикальные. Согласно доказанному в п. 7, каждый из этих углов вдвое больше угла между противолежащими сторонами ромбоида. А так как стороны ромбоида также даны, то этими величинами ромбоид вполне определен.

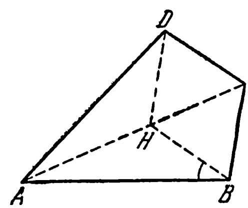


Рис. 76.

В этом можно убедиться следующим вспомогательным построением. Выделим из ромбоида $ABCD$ ромб $BCDH$ (рис. 76). Точки A, H, C, в силу симметрии, лежат на одной прямой[1]).

Так как $HB \| DC$, то угол ABH равен углу DCZ и, следовательно, равен половине угла DCE (рис. 75). Так как сторона AB дана, а BH равна BC, то треугольник ABH полностью определен. Далее, D есть зеркальное отображение B относительно оси AH, а C — четвертая вершина ромба $BCDH$, уже определенного тремя вершинами B, H, D.

Итак, обе половины прибора, схематически изображенного на рис. 75, как это и утверждалось, всегда симметричны, поэтому расстояние точки C до AB равно ее расстоянию до $A'B'$. Значит, точки A, B, A', B' при любом положении механизма лежат на одной прямой. Если закрепить теперь точки A и B неподвижно, то отрезок $A'B'$ сохранит свою подвижность. Он сможет двигаться только по прямой AB взад и вперед.

---

[1]) Этот чертеж совпадает со схемой инверсора Поселье.

Таким образом, мы получили новый метод проведения прямой с помощью шарнирного механизма.

Этот последний прямолинейно-направляющий механизм дает даже нечто бо́льшее сравнительно с предыдущим. Тот сообщает прямолинейное движение только одной точке, с помощью же последнего механизма мы заставляем двигаться по прямой целый отрезок $A'B'$. Так как с отрезком $A'B'$ можно связать неизменно произвольные фигуры, даже целую плоскость, то этот прибор осуществляет параллельное перемещение всей плоскости по самой себе, так что все ее точки проходят взаимно параллельные и конгруэнтные отрезки прямых [56].

19. СОВЕРШЕННЫЕ ЧИСЛА

В конце IX книги своих «Начал» — последней из трех, посвященных арифметике, — Евклид, доказав неограниченность ряда простых чисел способом, воспроизведенным нами в теме 1, переходит затем к так называемым совершенным числам. Название «совершенное число» встречается часто и у Платона, в частности, в том месте «Государства», где он пользуется им в туманных рассуждениях по евгенике в связи с так называемым «брачным» числом.

Современная наука рассматривает эту теорию вместе со всем тем, что присоединили к ней позднейшие математики, как своего рода курьез, оставшийся в стороне от широких путей научной мысли. Если мы все же обращаемся теперь к этой теме, то делаем это потому, что она представляет глубокий интерес с точки зрения методологической; в ней есть искорка, которую, как это мы покажем в следующей лекции, удалось зажечь позднее Эйлеру и которая ныне сверкает ярким пламенем в виде одного из наиболее своеобразных разделов современной математики, а именно теории распределения простых чисел.

Совершенным по Евклиду называется такое число, которое равно сумме всех своих делителей [1]). Число 6, например, есть совершенное число, так как его делители, как это легко проверить, суть 1, 2, 3 и сумма их $1+2+3=6$. Ближайшим следующим совершенным числом, которое можно обнаружить путем проб, является $28=1+2+4+7+14$. Следующее совершенное число встречается уже не так скоро.

[1]) Исключая, конечно, из их состава само число.

Требуется найти все дальнейшие совершенные числа не путем подобных проб, а систематически.

1. Ясно, что ни одно простое число не может быть совершенным. Ведь простое число p делится лишь на 1 и на p, а так как, согласно определению совершенного числа, само число не включается в число его делителей, то сумма всех делителей оказывается здесь всегда равной 1 и потому ни в коем случае не может быть равна p.

2. Исходя из этого чрезвычайно простого факта, будем осторожно продвигаться дальше. Число 9 не простое, но представляет собой квадрат простого числа; оно делится на 1 и на 3 и, кроме этих чисел, никаких других делителей не имеет (в этом легко убедиться, испытав все числа от 1 до 8). А поскольку сумма этих делителей $1+3=4$ много меньше 9, данное число также не является совершенным.

Абсолютно ясно, что подобным же образом и квадрат *всякого* другого простого числа p^2 делится на 1 и на p. Однако тот факт, что двумя этими числами исчерпываются *все* делители числа p^2, уже не представляется возможным проверить для всех p так же, как это мы сделали для 9; удостовериться в этом можно только путем *доказательства*. Это обстоятельство было замечено Евклидом, и он рассмотрел все необходимые вспомогательные теоремы. Мы приводим здесь его доказательство в несколько иной формулировке, основываясь на теореме о единственности разложения на простые множители, о которой была речь в теме 12. Действительно, если бы p^2 делилось на какое-либо иное число, кроме p, его можно было бы разложить иначе: с использованием иных простых множителей, чем в непосредственно данном нам разложении $p^2=p\cdot p$.

3. Вообще говоря, никакая другая более высокая степень простого числа также не может быть числом совершенным. Действительно, все делители p^a суть также степени p, только более низкие, чем p^a:

$$1,\ p,\ p^2,\ \ldots,\ p^{a-1}; \tag{1}$$

их сумма, согласно правилу сложения членов геометрической прогрессии, которое, кстати сказать, Евклид и выводит как раз по этому поводу, равна

$$1+p+p^2+\ldots+p^{a-1}=\frac{p^a-1}{p-1}. \tag{1a}$$

Встречающийся здесь знаменатель $p-1$ для простейшего случая ($p=2$) равен 1, для всякого же p, большего 2, превышает 1; значит, $(p-1)$-я доля разности p^a-1 или равна самой разности p^a-1, или меньше ее, но во всяком случае не может быть больше ее; поэтому сумма всех делителей p^a может быть равна, самое бо́льшее, p^a-1 и, следовательно, обязательно меньше чем p^a. Таким образом, число вида p^a не может быть совершенным.

4. После этих предварительных замечаний уже нетрудно исследовать вопрос, может ли быть совершенным число, разлагающееся на произведение степеней двух простых чисел, как, например, $72=2^3\cdot 3^2$ или, в более общей форме, число вида p^3q^2. Прежде всего непосредственно ясно, что наше число делится на следующие числа:

$$\begin{array}{llll} 1, & p, & p^2, & p^3 \\ q, & qp, & qp^2, & qp^3 \\ q^2, & q^2p, & q^2p^2, & q^2p^3, \end{array} \tag{2}$$

т. е. на все числа вида $p^\alpha q^\beta$, где α изменяется от 0 до 3, а β — от 0 до 2.

На основании теоремы о единственности разложения на простые множители, этим исчерпываются все его делители. Чтобы образовать сумму этих делителей, заметим, что в строках нашей таблицы (2) стоят произведения величин первой строки: во второй строке — на q, в третьей строке — на q^2. Сумма же делителей, помещенных в первой строке, равна сумме геометрической прогрессии $1+p+p^2+p^3$; таким образом, для получения искомой суммы делителей нужно взять сумму этой геометрической прогрессии, затем произведение ее на q и, наконец, произведение ее на q^2, т. е.

$$T=(1+q+q^2)(1+p+p^2+p^3).$$

Аналогично, если взять исходное число $N=p^aq^b$, получаем сумму

$$T=(1+q+\ldots+q^b)(1+p+\ldots+p^a)=$$
$$=\frac{q^{b+1}-1}{q-1}\cdot\frac{p^{a+1}-1}{p-1}. \tag{2a}$$

При этом мы должны обратить внимание на то, что в отличие от предыдущих случаев само N входит здесь в число делителей.

5. Обобщение этих выводов на числа N, состоящие из степеней более чем двух различных простых чисел, не вызывает никаких затруднений. Для $N=p^a q^b r^c \ldots$, кроме таблицы (2), мы получим еще таблицу, члены которой увеличены соответственно в r раз, в r^2 раз и т. д. до r^c раз; следовательно, сумма всех делителей увеличится в $(1+r+\ldots+r^c)$ раз, а для каждого следующего простого множителя будет возникать необходимость в новом соответствующем умножении (число N снова засчитывается в совокупность своих делителей):

$$T=(1+p+\ldots+p^a)(1+q+\ldots+q^b)(1+r+\ldots+r^e)\ldots=\frac{p^{a+1}-1}{p-1}\cdot\frac{q^{b+1}-1}{q-1}\cdot\frac{r^{c+1}-1}{r-1}\ldots \qquad (3)$$

6. Все это в основном имеется у Евклида, однако он рассматривает только число вида $p\cdot 2^b$, т. е. случай двух простых множителей, из которых один $q=2$, а другой входит лишь в степени $a=1$. Для этого частного случая из нашей общей формулы (2а) получаем:

$$T=\frac{2^{b+1}-1}{2-1}\cdot\frac{p^2-1}{p-1}.$$

Число N будет совершенным, если сумма его делителей T, в которую входит и само N, будет равна $N+N$, т. е. $2N$:

$$T=(2^{b+1}-1)(p+1)=2N$$

или, если вставить значение N:

$$(2^{b+1}-1)(p+1)=2\cdot p\cdot 2^b=2^{b+1}\cdot p.$$

Отсюда мы тотчас же получаем, что

$$2^{b+1}(p+1)-(p+1)=2^{b+1}\cdot p,$$

или

$$2^{b+1}=p+1,$$

или, наконец,

$$p=2^{b+1}-1. \qquad (4)$$

Если, следовательно, число p есть не произвольное простое число, а имеет вид $2^{b+1}-1$, то число N есть совершенное. При тех или иных определенных значениях b число $2^{b+1}-1$ не всегда будет простым. Таким образом, вместе с Евклидом мы получили следующий результат:

Число $N=(2^{n+1}-1)2^n$ *есть всегда совершенное число, если* $2^{n+1}-1$ — *число простое* [57].

7. Будем придавать n последовательно значения 1, 2, 3, ... и посмотрим, какие при этом получаются совершенные числа:

$n=1$, $N=(2^2-1)\cdot 2=3\cdot 2=6$;

2, $(2^3-1)\cdot 2^2=7\cdot 4=28$;

3, $(2^4-1)\cdot 2^3=15\cdot 8$ отпадает, так как 15 — число не простое;

4, $(2^5-1)\cdot 2^4=31\cdot 16=496$;

5, $(2^6-1)\cdot 2^5=63\cdot 32$ отпадает, так как 63 — число не простое;

6, $(2^7-1)\cdot 2^6=127\cdot 64=8128$;

. . . .

Мы видим, что полученный Евклидом результат порождает новую проблему: *для какого* n *число* $2^{n+1}-1$ *простое*? Один шаг к решению этой проблемы можно сделать тотчас же: $n+1$ должно быть во всяком случае простым числом. Действительно, если бы $n+1$ было составным числом, например, если бы $n+1=uv$, то мы имели бы

$$2^{n+1}-1=2^{uv}-1=(2^u)^v-1,$$

и, подставляя $x=2^u$ в формулу геометрической прогрессии

$$x^v-1=(x^{v-1}+x^{v-2}+\ldots+x+1)(x-1),$$

мы получили бы, что

$$(2^u)^v-1=[(2^u)^{v-1}+\ldots+(2^u)+1]\,[(2^u)-1],$$

т. е. наше число разлагалось бы на два множителя и, таким образом, не было бы простым числом. Оно может быть простым, следовательно, только в том случае, если $n+1$ — число простое.

При продолжении нашей таблицы ближайшим, следующим за 7 простым числом является 11, поэтому очередное значение, которое может принять n, равно 10. Но является ли при этом число $2^{11}-1$ действительно простым, а число $(2^{11}-1)\cdot 2^{10}$ в силу этого совершенным — нисколько еще этим не предрешается. И действительно, как в этом можно убедиться после многократных проб, оказывается, что $2^{11}-1=$ $=2047=23\cdot 89$. Какие из последующих чисел являются совершенными, имеет ли их последовательность конец или допускает неограниченное продолжение — все это проблемы, к решению которых не смогла подойти и современная наука [58].

8. Вот все, что дает теорема Евклида. О том, что древние имели об этом предмете более обширные сведения, свидетельствуют разнообразнейшие источники. Прежде всего Ямвлих без всяких дальнейших пояснений сообщает о том, что, *кроме указанных Евклидом, никаких других четных совершенных чисел быть не может.* Было ли это доказано древними и каким именно способом — мы не знаем. Доказательство этого предложения, данное Эйлером, опирается целиком на формулу (3), имеющуюся в своей основе уже у Евклида, и осуществляется посредством несложных рассуждений. Мы его здесь вкратце воспроизводим, хотя оно и не имеет никакого значения для основной идеи дальнейшего.

Пусть N — произвольное четное число; в таком случае 2 в какой-либо степени, например 2^n, непременно имеется среди его простых множителей, все же остальные простые множители дают в произведении нечетное число. Число N, значит, можно представить в форме $N=2^n u$, где u — нечетное число. Образуем для этого N по формуле (3) сумму его делителей T; первый множитель в этой формуле будет иметь вид

$$\frac{2^{n+1}-1}{2-1}=2^{n+1}-1.$$

Остальные множители будут иметь такой вид, как если бы вместо суммы T делителей числа N мы искали сумму U делителей u, т. е.

$$T=(2^{n+1}-1)\,U. \tag{5}$$

Чтобы N было совершенным числом, это выражение, в котором N само участвует в качестве делителя, должно дать в сумме $2N$:

$$(2^{n+1}-1)\,U=2N=2\cdot 2^n\cdot u=2^{n+1}\cdot u.$$

Раскрывая скобки, получаем:

$$2^{n+1}\cdot U-U=2^{n+1}\cdot u,$$

или

$$2^{n+1}(U-u)=U=(U-u)+u,$$

или, наконец,

$$(2^{n+1}-1)(U-u)=u. \qquad (6)$$

U есть сумма всех делителей u, включая само u. Следовательно, $U-u$ есть сумма всех делителей u, не считая самого u. Формула (6) говорит, что если N есть число совершенное, то u делится на $U-u$.

Но если $U-u$ само является делителем u и в то же время суммой всех делителей u, оно должно быть единственным делителем u. Среди же делителей u должна быть во всяком случае 1. Значит, $U-u$ должно быть равно 1, а u есть простое число, и тогда формула (6) принимает более простой вид: $u=2^{n+1}-1$, так что

$$N=2^n(2^{n+1}-1).$$

А это значит, что из всех четных чисел совершенными являются только те, которые указаны Евклидом.

9. Здесь мы сталкиваемся со второй проблемой, связанной с нашей темой и еще не преодоленной современной наукой: *существуют ли нечетные совершенные числа?* До сих пор не было найдено ни одного такого числа, и представляется весьма невероятным, чтобы таковые существовали; однако никто еще не смог доказать, что это действительно так.

10. О том, что древним было известно больше, чем приведено у Евклида, позволяет думать одно малоисследованное место в V книге «Законов» Платона. Платон рекомендует там для вновь организуемого государства назначать числа безземельных и собственников таким образом, чтобы эти числа имели возможно больше делителей, например 5040, число делителей которого равно 60—1. Законодатель должен быть настолько свéдущим в арифметике, чтобы суметь установить эти числа сообразно с масштабами государства.

Обратим прежде всего внимание на то, каким образом определяется число делителей какого-либо числа, в частности числа 5040. С этой целью нам нужно будет прибегнуть к тем же рассуждениям, с помощью которых мы научились выше отыскивать сумму делителей некоторого числа. Если, например, все делители числа p^3q^2 представлены таблицей (2), то из нее легко определить как сумму делителей, так и число их, а именно $4 \cdot 3 - 1 = 12 - 1$ (исключая само N): точно так же при $N=p^aq^b$ для числа делителей Z получаем:

$$Z=(a+1)(b+1)-1;$$

наконец, вообще для $N = p^a q^b r^c \ldots$:

$$Z = (a+1)(b+1)(c+1)\ldots - 1.$$

В частности, для $N = 5040 = 2^4 \cdot 3^2 \cdot 5 \cdot 7$ имеем:

$$Z = (4+1)(2+1)(1+1)(1+1) - 1 = 5 \cdot 3 \cdot 2 \cdot 2 - 1 = \\ = 60 - 1.$$

Необычный для Платона способ задания числа 59 в виде разности 60—1, а также указание, что законодатель должен понимать толк в этих вещах, дают нам право предполагать, что и сам Платон хорошо в них разбирался.

11. Связь этих двух вопросов — о сумме и числе делителей некоторого числа — станет еще отчетливей, если мы обратим внимание на то, что оба они являются частными случаями более общей задачи изучения суммы s-х степеней всех делителей N, т. е., например для частного случая $N = 6$, выражения

$$1^s + 2^s + 3^s.$$

Действительно, при $s = 1$ это выражение представляет собой не что иное, как сумму делителей T. При $s = 0$ каждое слагаемое этой суммы обращается в единицу, и таких единиц получается столько же, сколько имеется делителей в N, т. е. при $s = 0$ это выражение дает число делителей. Для $s = 2$ мы получили бы сумму квадратов делителей и т. д. Для любого s можно было бы непосредственно вывести формулу, аналогичную (3); метод рассуждений остается прежним [59]. Ничто не мешает провести этот вывод и для $s = -1$, т. е. для (-1)-х степеней или обратных величин делителей; в нашем примере при $N = 6$ для

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3}.$$

Для такой суммы обратных величин делителей также справедлива формула, аналогичная (3):

$$R = \left(1 + \frac{1}{p} + \ldots + \frac{1}{p^a}\right)\left(1 + \frac{1}{q} + \ldots + \frac{1}{q^b}\right) \times \\ \times \left(1 + \frac{1}{r} + \ldots + \frac{1}{r^c}\right)\ldots$$

В основе всего этого лежит единая идея о возможности получения всех делителей числа посредством таблицы типа (2). Математики древней Греции владели этой идеей и, как об этом позволяет судить указание Платона, сумели оценить ее красоту и важность.

20. ДОКАЗАТЕЛЬСТВО НЕОГРАНИЧЕННОСТИ РЯДА ПРОСТЫХ ЧИСЕЛ ПО ЭЙЛЕРУ

Эта же основная идея послужила Эйлеру и для того, чтобы дать еще второе доказательство неограниченности ряда простых чисел, в дополнение к тому, с которым мы познакомились в теме 1 и которое у Евклида, кстати сказать, непосредственно предшествует учению о совершенных числах.

Предпошлем доказательству два простых замечания:

Рис. 77.

1. Пусть AB — отрезок длиной 2 *м* (рис. 77). Будем двигаться по нему от A до его середины M, затем от M до середины M_1 остающейся половины MB, затем от M_1 до середины M_2 остатка M_1B и т. д.; приближаясь таким образом все время к B, мы всегда будем оставаться перед B. При этом мы будем получать последовательно отрезки в 1 *м*, $\frac{1}{2}$ *м*, $\frac{1}{4}$ *м* и т. д.; все они вместе будут меньше 2 *м*:

$$1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^n}<2.$$

Подобное же неравенство мы получим также и в том случае, если отрезки будут уменьшаться не в отношении 1:2, а в каком-либо ином отношении $x<1$, а именно, согласно теореме о сумме геометрической прогрессии,

$$1+x+x^2+\ldots+x^n=\frac{1-x^{n+1}}{1-x}=\frac{1}{1-x}-\frac{x^{n+1}}{1-x};$$

если $x<1$, то вторая дробь правой части положительна и, следовательно, ее вычитание уменьшает результат, т. е., иными словами,

$$1+x+x^2+\ldots+x^n<\frac{1}{1-x}.$$

В частности, если p — какое-либо простое число, то всегда $\frac{1}{p}<1$, и поэтому полученное неравенство наверное удовлетворяется, если в нем вместо x подставить значение $\frac{1}{p}$. Отсюда получаем:

Если p — какое-либо простое число, n — целое число, то всегда имеет место неравенство

$$1+\frac{1}{p}+\frac{1}{p^2}+\ldots+\frac{1}{p^n}<\frac{1}{1-1/p}=\frac{p}{p-1}. \qquad (1)$$

2. Положим:

$$A_m=1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2^m}.$$

Тогда, очевидно:

$$A_2=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}>1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)=1+\frac{2}{2},$$

$$A_3=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}>$$

$$>A_2+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)>1+\frac{3}{2},$$

$$A_4=1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{16}>A_3+\left(\frac{1}{16}+\ldots+\frac{1}{16}\right)>$$

$$>1+\frac{4}{2},$$

и вообще

$$A_m=1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{2^m}>1+\frac{m}{2}. \qquad (2)$$

Чем больше m, тем больше будет становиться и A_m. Достаточно, например, взять $m=1998$, чтобы $\frac{m}{2}$ достигло значения 999, и таким образом A_m превзошло 1000; точно так же легко можно было бы определить такое значение m, при котором A_m будет заведомо больше миллиона, и т. д.[1]).

[1]) Так как степени 2 встречаются в ряде натуральных чисел как угодно далеко, то из неравенства (2), между прочим, следует, что выражение $1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}$, где n не обязательно есть степень 2, при возрастании n неограниченно возрастает, или, как говорят математики, *бесконечный ряд* $1+\frac{1}{2}+\frac{1}{3}+\ldots$ *расходится*,

3. Пусть p — какое-либо простое число. Тогда для всех простых чисел до p включительно мы можем написать, согласно формуле (1):

$$1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^n}<\frac{2}{2-1},$$
$$1+\frac{1}{3}+\frac{1}{3^2}+\ldots+\frac{1}{3^n}<\frac{3}{3-1},$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$1+\frac{1}{p}+\frac{1}{p^2}+\ldots+\frac{1}{p^n}<\frac{p}{p-1},$$

и образовать произведение R_p всех этих сумм. Это произведение будет меньше произведения M_p чисел, находящихся в правых частях наших неравенств:

$$R_p<M_p=\frac{2}{1}\cdot\frac{3}{2}\cdot\frac{5}{4}\cdot\frac{7}{6}\cdots\frac{p}{p-1}.$$

Но произведение R_p мы научились уже вычислять в предыдущей теме: R_p равно сумме всех слагаемых вида

$$\frac{1}{2^\alpha}\cdot\frac{1}{3^\beta}\cdots\frac{1}{p^\mu},$$

где α, β, ..., μ пробегают ряд значений от 1 до n; иными словами, оно равно сумме величин, обратных делителям числа

$$N=2^n\cdot 3^n\cdot 5^n\ldots p^n.$$

Таким образом, R_p представляет собой сумму дробей, числители которых все равны 1, а знаменатели пробегают ряд чисел, составленных из простых множителей, взятых в степенях, не превышающих n.

Например,

$$\left(1+\frac{1}{2}+\frac{1}{2^2}\right)\left(1+\frac{1}{3}+\frac{1}{3^2}\right)=$$
$$=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{6}+\frac{1}{9}+\frac{1}{12}+\frac{1}{18}+\frac{1}{36},$$

хотя отдельные слагаемые этого ряда становятся все меньше и меньше.

Это интересное обстоятельство, так же как и идея сходимости бесконечного ряда, не имеет, однако, никакого отношения к рассматриваемому вопросу и могло бы только запутать наши, сами по себе простые, рассуждения.

т. е. равно сумме чисел, обратных всем делителям числа $36 = 2^2 3^2$.

4. Теперь мы приступаем, собственно, к доказательству. Пусть m будет каким-либо положительным целым числом; тогда выражение A_m содержит обратные величины всех целых чисел до 2^m включительно:

$$A_m = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{2^m}.$$

Представим себе все простые множители, входящие в состав чисел ряда

$$1,\ 2,\ 3,\ 4,\ \ldots,\ 2^m. \tag{3}$$

Пусть наибольшее из этих простых чисел будет q. В таком случае числа ряда (3) будут составлены исключительно из простых множителей $2, 3, 5, \ldots, q$, причем ни одно из них не войдет в степени, большей m. Действительно, если бы первое из этих чисел, 2, вошло в состав какого-либо числа из ряда (3) в степени, большей m, то само это число должно было бы быть больше 2^m, т. е. выйти за пределы рассматриваемого ряда. По отношению к следующим, бо́льшим простым числам $3, 5, \ldots$ это тем более невозможно. Таким образом, *все числа ряда* (3) *должны содержаться среди делителей числа* $2^m \cdot 3^m \cdot 5^m \ldots q^m$.

Образуем теперь, согласно п. **3**, произведение R_p не для произвольного простого числа p, а для только что указанного простого числа q, взяв m вместо n. Тогда выражение R_q будет содержать в себе в с е слагаемые A_m и, кроме того, много других. В этом легко убедиться из приведенного в конце п. **3** числового примера: сюда входят в с е слагаемые A_2, т. е. $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}$, и, кроме того, много других. Таким образом, во всяком случае $A_m < R_q$. Далее, согласно п. **3**, $R_q < M_q$. Наконец, формула (2) дает $1 + \frac{m}{2} < A_m$, и мы получаем окончательно, что

$$1 + \frac{m}{2} < \frac{2}{1} \cdot \frac{3}{2} \cdot \frac{5}{4} \cdots \frac{q}{q-1}. \tag{4}$$

Напомним еще раз, что m — п р о и з в о л ь н о е целое положительное число, а q — наибольший из простых множителей, входящих в числа ряда (3). Левую часть неравенства (4) можно, разумеется, сделать сколь угодно большой: для этого нужно лишь взять достаточно большое m. Но если бы существовало только некоторое определенное конечное число простых чисел, то q не могло бы превзойти последнее из этих простых чисел, и тогда правая часть неравенства (4) не могла бы неограниченно возрастать; тем самым мы пришли бы к противоречию. Поэтому последовательность простых чисел не может оборваться [60].

Это доказательство гораздо сложнее приведенного в теме 1. Но оно может быть распространено на весьма широкий круг подобных же, более трудных проблем, из которых лишь немногие были упомянуты в конце первой темы. Кроме того, на этом доказательстве основано все учение о распределении простых чисел — одна из наиболее трудных, наиболее актуальных и обширных по количеству выдвигаемых ею проблем областей современной математики.

Намеченная здесь основная идея, сущность которой восходит к древним, является руководящей во всей этой области.

21. ПРИНЦИПИАЛЬНЫЕ ОСНОВЫ ЗАДАЧ НА МАКСИМУМ

Мы уже не раз, а именно в темах третьей, пятой и шестой, занимались задачами на максимум, и в каждой из этих тем имели возможность предложить вниманию читателя тот или иной небольшой математический шедевр. Сейчас, совершенно независимо от изложенного раньше, мы хотим выяснить некоторые принципиальные моменты, лежащие в основе решения задач на максимум.

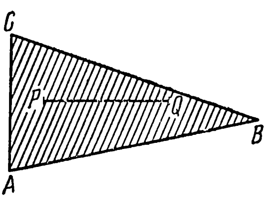


Рис. 78.

Займемся задачей на максимум несколько иного типа, чем те, которые мы решали раньше; крайняя ее простота как раз и позволит нам особенно удобно раскрыть в ней принципиальную сторону вопроса. *Дан треугольник* (рис. 78) (представим себе, например, что он вырезан из картона). *Спрашивается, какие две точки P, Q этого куска плоскости (включая и его края) дальше всего отстоят одна от другой?* Угадать здесь решение совсем легко: это будут концы самой длинной стороны треугольника. Но как это доказать? Для всех подобного рода доказательств существует один общий простой рецепт, быстро ведущий к результату, которым мы, однако, не пользовались в предшествующих главах. Рассуждают так: если одна из данных точек P, Q, например P, лежит в н у т р и треугольника, то расстояние между ними не будет наибольшим, так как на продолжении отрезка PQ существует еще точка P_1, удаленная от Q больше, чем P, и все же еще лежащая внутри треугольника. Если же обе точки расположены на сторонах треугольника, но так, что одна из них, например P, не совпадает с какой-либо вершиной треугольника, то точно так же

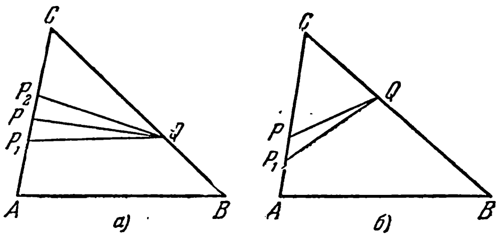


Рис. 79.

по соседству с точкой P на той же самой стороне треугольника можно найти некоторую точку P_1, отстоящую от Q дальше, чем P. В этом легко убедиться путем элементарного рассуждения как для случая, когда отрезок PQ перпендикулярен к той стороне треугольника, на которой лежит P (рис. 79, *а*), так и для того случая, когда он к ней не перпендикулярен (рис. 79, *б*). Максимум возможен, следовательно, лишь в том случае, когда обе точки P и Q являются вершинами треугольника. Поэтому отрезок PQ должен быть стороной треугольника, и притом, естественно, наибольшей.

Руководствуясь этим же самым принципом, мы могли бы распространить полученный результат и на n-угольник: *чтобы найти две наиболее удаленные друг от друга точки на поверхности n-угольника, нужно выбрать наибольшее из всех возможных расстояний между его вершинами.* При этом совсем не обязательно, чтобы n-угольник был непременно выпуклым; например, в четырехугольнике (рис. 80) с входящим углом парой наиболее удаленных точек будут две вершины острых углов.

Этот же принцип мог бы внести значительные упрощения и в решение ранее разобранных задач. Чтобы показать, например, что из всех вписанных в круг треугольников (рис. 81) равносторонний имеет наибольшую площадь, нужно было бы взять какой-либо неравносторонний треугольник, т. е. имеющий две неравные стороны (такими неравными сторонами могут быть AC и BC), построить на AB как на основании равнобедренный

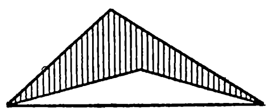


Рис. 80.

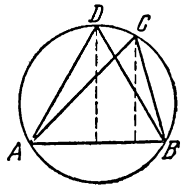


Рис. 81.

вписанный в тот же круг треугольник ABD с высотою большей, чем в треугольнике ABC, и тогда легко было бы убедиться в том, что площадь его должна быть также больше площади ABC. Таким образом, ни один неравносторонний треугольник не может дать максимума, искомый треугольник должен быть равносторонним. Столь же кратки были бы наши рассуждения и в задаче о треугольнике, образованном основаниями высот (темы 5 и 6).

Почему же мы тогда потратили столько труда на решение задачи иным способом и почему мы не удовлетворены только что найденным решением разобранной здесь задачи? Да потому, что в нем содержится одна очень грубая логическая ошибка, в течение двух столетий остававшаяся незамеченной математиками и впервые вскрытая лишь во второй половине XIX века Вейерштрассом.

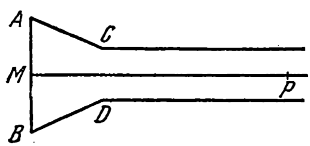


Рис. 82.

Ошибка эта станет очевидна, если мы попробуем применить наш метод к следующей чрезвычайно простой фигуре (рис. 82). Ясно, что для этой фигуры, простирающейся в бесконечность, нельзя указать такую пару точек, которые были бы удалены друг от друга на максимальное расстояние; чем дальше, например, отодвигается точка P вправо, тем больше будет расстояние MP. Тем не менее только что описанный нами метод рассуждений применим и к этой фигуре. Где бы ни лежали точки P и Q, внутри фигуры или на ее периметре, даже и тогда, когда P и Q находятся в каких-либо двух из четырех вершин фигуры, можно, очевидно, найти такую соседнюю с P точку P_1, которая отстоит от Q еще дальше, за исключением одного-единственного случая, когда PQ есть отрезок AB. Мы пользуемся здесь тем же самым принципом; для всех возможных точечных пар, за исключением A, B, отыскиваем такую соседнюю пару, для которой расстояние становится бóльшим; ни одна пара, кроме A, B, не может дать максимума. Отсюда, в строгом соответствии с предыдущим доказательством, мы должны были бы заключить, что этим максимумом является AB [61].

Очевидно, что по форме этот вывод вполне аналогичен предыдущему. Но здесь он приводит к неверному результату. Откуда же нам известно, что он дает надежный результат в предыдущих случаях?

Легко указать, где именно скрыта ошибка. Мы совершенно правильно показали, что никакая другая пара точек треугольника, кроме концов наибольшей стороны его, не может дать максимума. Но откуда нам, собственно, известно, что расстояние между этими двумя точками действительно больше, чем между какими-либо другими точками треугольника? Если бы мы знали, что эта задача имеет решение в том смысле, что для нее существует максимум, тогда, конечно, мы могли бы логически заключить, что указанная пара точек и есть это решение, ибо другого выбора у нас нет. Но указание на то, что задача вообще имеет решение, отсутствует, и именно это обстоятельство повлекло за собой бессмысленный вывод в отношении фигуры, изображенной на рис. 82; здесь как раз ошибочно предположение, что максимум существует, хотя все остальные выводы и правильны [62].

Теперь ясно, что доказательства, в предыдущих темах на первый взгляд чересчур обстоятельные, представляют не только эстетическую, но и логическую ценность. Для задачи же, приведенной в этой теме (рис. 78), остается еще только дать заключительное доказательство, которое, однако, будет неприложимо по отношению к фигуре, изображенной на рис. 82.

Предпошлем доказательству весьма элементарное замечание. Если мы имеем (рис. 83) две параллельные прямые g и h и прямую, к ним перпендикулярную, то отрезок PQ перпендикуляра между параллелями всегда короче любого отрезка UV, соединяющего какую-либо точку U одной параллельной прямой с какой-либо точкой V другой параллели, и тем более короче отрезка $U'V'$, соединяющего точку U', лежащую левее g, с точкой V', лежащей правее h.

Мы не будем доказывать нашу теорему о максимальном расстоянии между двумя точками специально для треугольника, так как ее можно доказать без особых затруднений сразу для любого n-угольника.

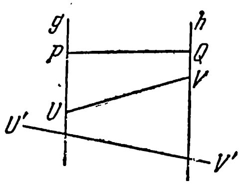


Рис. 83.

Пусть P и Q — какие-либо две точки (рис. 84) внутри или на периметре этого n-угольника. Восставим в обеих этих точках перпендикуляры g и h к прямой PQ, их соединяющей. Они вырежут из всей фигуры параллельную полосу. Так как на прямой g лежит по крайней мере одна точка P этого n-угольника, то слева от g или в крайнем случае на самой g должна лежать какая-либо из вершин n-угольника, например A; точно так же какая-нибудь вершина, например B, должна лежать справа от h или на самой h. Таким образом, расстояние AB будет не меньше, чем PQ; следовательно, и наибольшее из всех возможных расстояний между вершинами n-угольника будет также не меньше PQ, где точки P и Q произвольны; значит, оно и будет максимумом [63].

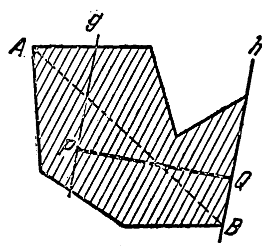


Рис. 84.

Следующая тема посвящена задаче на максимум более трудной, чем все до сих пор разобранные. В этой задаче требуется отыскать фигуру (прямолинейную или криволинейную), имеющую наибольшую площадь при данном периметре, и мы докажем, что никакая другая фигура, кроме круга, этого максимума дать не может. Мы не будем приводить доказательства того, что площадь круга действительно больше площади всякой другой фигуры того же периметра. То обстоятельство, что круг есть криволинейная фигура, вносит такие большие осложнения, что нам потребуется много труда и искусства даже для того, чтобы доказать наше более скромное утверждение.

22. ФИГУРА, ИМЕЮЩАЯ НАИБОЛЬШУЮ ПЛОЩАДЬ ПРИ ДАННОМ ПЕРИМЕТРЕ (ЧЕТЫРЕХШАРНИРНЫЙ МЕТОД ШТЕЙНЕРА)

Почему мыльные пузыри принимают форму шара? Потому, что, будучи сжаты окружающей средой, они стремятся в силу сцепления образовать при неизменном объеме (дальнейшее сжатие воздуха здесь не играет роли) возможно более толстую поверхностную пленку, или, иначе говоря, потому, что они сумели разрешить вопрос о том, *какое тело при данном объеме имеет наименьшую поверхность.*

Задачу, которая оказалась под силу любому мыльному пузырю, любой капле воды, попытаемся теперь решить и мы,

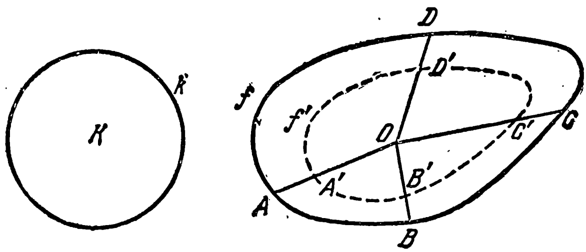


Рис. 85.

хотя бы и в значительно более скромных масштабах, не в пространстве, а всего лишь в двух измерениях. Иначе говоря: *из всех замкнутых кривых, имеющих данную площадь, мы постараемся отыскать такую, периметр которой имеет наименьшую величину* [64].

1. Поставленный настоящей темой вопрос об отыскании фигуры наибольшей *площади* при данном периметре звучит несколько иначе, чем только что сформулированная задача о фигуре наименьшего *периметра* при данной площади, однако различие здесь только кажущееся.

Действительно, предположим, что наибольшую площадь при данном периметре имеет не круг K (рис. 85) с периметром k, а какая-либо иная фигура F с периметром f, т. е. предположим, что площадь этой фигуры больше площади круга с тем же самым периметром: $F > K$ при $f = k$. В таком случае мы могли бы изобразить фигуру F в уменьшенном масштабе, для чего достаточно сократить в одном и том же отношении все лучи OA, OB, OC, OD, идущие из какой-либо точки O внутри фигуры F к ее периметру. Выберем этот масштаб так, чтобы получившаяся в результате сокращения фигура F' с периметром f' имела площадь, равную площади K; так как при этом периметр фигуры, очевидно, сокращается, то мы имели бы $f' < k$. Таким образом, оказалось бы, что периметр фигуры F' меньше периметра круга, имеющего ту же самую площадь, что и F'. Но если будет доказано, что это невозможно, то тем самым решается в пользу круга и та задача, которая поставлена в заголовке. Эти выводы можно провести точно так же и в обратном порядке. Обе формулировки, следовательно, совершенно эквивалентны.

2. Мы только что неявным образом воспользовались следующей теоремой о периметре и площади плоской фигуры:

I. *Если плоскую фигуру уменьшить подобно (точнее— гомотетично) в отношении* $1:r$, *то периметры первоначальной и уменьшенной фигур будут относиться, как* $1:r$, *площади же, как* $1:r^2$.

Соберем теперь здесь все остальные, необходимые нам в дальнейшем теоремы о периметре и площади. Мы не будем останавливаться на их доказательствах, иначе пришлось бы заняться исследованием самих этих понятий, что не входит сейчас в нашу задачу. Если бы мы провели все доказательства (а это потребовало бы обширного и систематического анализа), то убедились бы в справедливости теорем, занумерованных в дальнейшем римскими цифрами. Во всяком случае мы будем пользоваться только этими, отмеченными римскими цифрами теоремами и, таким образом, имеем перед собой ясно поставленную задачу.

II. *Если одна фигура является частью другой фигуры, то площадь этой части меньше площади всей фигуры* (рис. 86).

Для периметра аналогичная теорема не имеет места, в чем можно непосредственно убедиться из рис. 87: если контур внутренней фигуры достаточно извилист, периметр ее наверное больше периметра внешней кривой. Но теорема будет справедлива, если ограничиться рассмотрением лишь в ы п у к л ы х фигур.

Фигура называется выпуклой, если отрезок, соединяющий две какие-либо ее точки, целиком принадлежит этой *фигуре*. На рис. 88 показана *не* выпуклая фигура: P и Q — точки, лежащие внутри этой фигуры, отрезок же PQ не весь принадлежит ей.

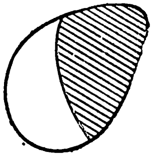


Рис. 86.

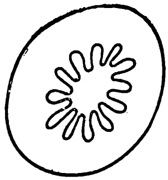


Рис. 87.

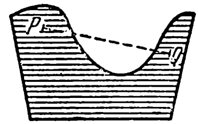


Рис. 88.

Имеют место следующие теоремы:

III. *Если одна выпуклая фигура объемлет другую выпуклую, то периметр первой больше периметра второй.*

IV. *Всякая невыпуклая фигура может быть дополнена до выпуклой, так что площадь фигуры при этом увеличивается, а периметр уменьшается* (рис. 89).

Эти четыре факта, приведенные нами без доказательств, интуитивно вполне очевидны, и если здесь могут быть какие-либо сомнения, то разве лишь в отношении III.

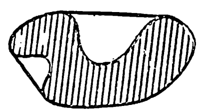


Рис. 89.

3. Теперь мы вступаем в круг идей Штейнера, изложение которых и составляет задачу настоящей темы. Начнем с утверждения, что *кривая, имеющая при данном периметре наибольшую площадь, должна быть во всяком случае выпуклой*. Выводится это непосредственно из IV и I. Действительно, если данную фигуру дополнить в соответствии с IV до выпуклой, то периметр ее при этом уменьшится, площадь же увеличится. Увеличим полученную фигуру подобно в таком масштабе, чтобы новый периметр равнялся по величине первоначальному; тогда, на основании I, площадь опять увеличится. Это значит, что для каждой невыпуклой фигуры можно построить выпуклую с тем же самым периметром, но большей площадью. Следовательно, выпуклая фигура никогда не может дать максимума площади при данном периметре.

4. Мы будем поэтому иметь дело только с выпуклыми кривыми. Прежде всего покажем, что *для любой выпуклой* *фигуры данного периметра существует выпуклая фигура того же периметра, не меньшей площади, обладающая осью симметрии.*

Пусть P — точка на кривой (рис. 90); отыщем на этой же кривой такую точку Q, которая делит полный периметр, отсчитываемый от точки P, на две равные части. Хорда PQ делит данную фигуру на две части; обратим внимание на ту из них, которая имеет бо́льшую площадь, и отразим ее зеркально от прямой PQ. Эта часть вместе со своим зеркальным отражением образует фигуру, площадь которой больше площади данной фигуры при той же величине периметра. Этим самым, следовательно, исходная фигура заменена б о́ л ь ш е й, она, значит, не давала максимума площади при данном периметре. В случае, если обе части фигуры равны, мы также производим аналогичную замену одной части фигуры зеркальным отражением второй части. В этом случае наше зеркальное отражение дает нам новую фигуру той же самой площади и того же периметра, но имеющую то преимущество перед исходной, что в ней есть ось симметрии PQ.

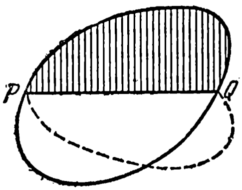


Рис. 90.

Для этого последнего случая мы можем тут же сделать и некоторые другие выводы. Если исходной фигурой был круг, то PQ будет одним из его диаметров, и новая фигура останется тем же самым кругом. Но и обратно: новая фигура будет кругом только в том случае, если та половина старой фигуры, которую мы зеркально отразили, была полукругом. Может случиться, однако, что одна половина фигуры была полукругом, другая же половина имела иную форму (рис. 91); в этом случае фигура, сохранив прежние периметр и площадь, была бы кругом, между тем как исходная не была им; есть, однако, возможность избежать этого и сделать верной также и обратную теорему. Ведь мы располагаем свободным выбором, какую из двух равновеликих полуфигур подвергнуть зеркальному отражению; установим поэтому раз и навсегда, что мы будем брать ту половину, которая не является полукругом.

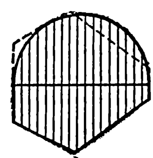


Рис. 91.

Теперь мы можем подвинуться дальше в доказательстве высказанного выше положения. Если новая фигура получится не выпуклой, то путем дополнения ее можно сделать выпуклой и, увеличив в соответствующем масштабе, вернуть ей прежний периметр; ось симметрии при этом сохранится; таким образом, наша фигура может быть, как мы будем говорить, «улучшена».

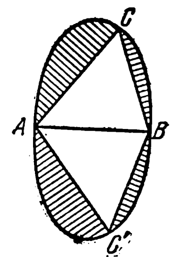


Рис. 92.

5. Итак, нам известно следующее: если исходная фигура не круг, то либо ее можно «улучшить» (и тогда, значит, она не дает максимума), либо можно найти выпуклую фигуру с теми же самыми площадью и периметром, которая также не будет кругом, но будет обладать осью симметрии.

Мы хотим теперь показать, что из выпуклой, симметричной относительно оси фигуры, не являющейся кругом, всегда можно получить другую фигуру того же периметра, с площадью не только не меньшей, но непременно большей. Если это удастся показать, то это будет означать, что *всякую фигуру, не являющуюся кругом, всегда можно «улучшить»*.

Заметим прежде всего, что на кривой, ограничивающей нашу фигуру (рис. 92), должна существовать точка C, из которой ось симметрии AB видна не под прямым углом. Ибо если бы AB из каждой точки кривой была видна под прямым углом, то эта кривая, в силу теоремы Фалеса или, если угодно, теоремы о вписанных углах (см. стр. 38, рис. 22), должна была бы быть окружностью с диаметром AB; по условию же эта кривая как раз н е является окружностью.

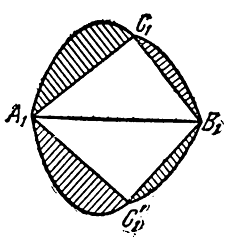


Рис. 93.

Рядом с треугольником ABC построим теперь другой треугольник, прямоугольный (рис. 93), с катетами $A_1C_1 = AC$ и $B_1C_1 = BC$; отразим его зеркально от A_1B_1 и к четырем сторонам полученного таким образом четырехугольника $A_1C_1B_1C_1'$ присоединим четыре заштрихованных сегмента, тождественных соответствующим сегментам рис. 92. Всю эту операцию можно представить себе иначе, а именно считать, что стороны четырехугольника $ACBC'$ на рис. 92 суть стержни, соединенные в вершинах шарнирами, а четыре сегмента неизменно прикреплены к этим стержням и что шарниры вращаются, пока при C и C' не образуются прямые углы; отсюда и название — четырехшарнирный метод Штейнера.

Периметр полученной таким образом фигуры остался прежним; он состоит из тех же самых дуг кривых. Площади сегментов, разумеется, также остались неизменными. Все дело, следовательно, сводится к сравнению четырехугольников или треугольников ABC и $A_1B_1C_1$, составляющих, в силу симметрии, половины этих четырехугольников. Площадь треугольника по известной теореме равна половине произведения основания на высоту; значит, площадь треугольника $A_1B_1C_1$ выразится произведением $\frac{1}{2} \cdot A_1C_1 \cdot B_1C_1$. Высота треугольника ACB, опущенная из вершины B на AC, будет непременно меньше BC, так как угол при C не прямой, и следовательно, площадь этого треугольника будет меньше, чем $\frac{1}{2} AC \cdot BC$. Но $A_1C_1 = AC$, $B_1C_1 = BC$; отсюда получается, что площадь ACB непременно меньше площади $A_1C_1B_1$, и значит, новая фигура имеет бо́льшую площадь, чем первоначальная.

Этим все сделано: для любой фигуры, не являющейся кругом, можно построить другую фигуру с тем же самым периметром, но с бо́льшей площадью. Следовательно, ни одна фигура, кроме круга, не может дать максимума площади при данном периметре.

Почему этим еще не разрешается вся проблема и почему приведенных доводов недостаточно для того, чтобы утверждать, что круг действительно превышает по площади все другие фигуры того же периметра, — об этом более подробно рассказано в предыдущей теме. Полное разрешение этой задачи, нуждающейся еще в некоторых заключительных рассуждениях, было бы, пожалуй, осуществимо и здесь, но для этого потребовалась бы некоторая математическая техника, доступная лишь тому, кто систематически в ней упражняется, а этого наша книга от читателя не требует [[65]].

23. ПЕРИОДИЧЕСКИЕ ДЕСЯТИЧНЫЕ ДРОБИ

1. При обращении обыкновенных дробей в десятичные мы встречаемся, как известно, с тремя различными случаями, о характере которых дают представление следующие типичные примеры:

I. $\frac{1}{5}=0,2,\qquad \frac{3}{40}=0,075;$

II. $\frac{4}{9}=0,4444...,\qquad \frac{1}{7}=0,142857142857...;$

III. $\frac{1}{6}=0,1666...,\qquad \frac{7}{30}=0,2333...$

Наиболее простыми здесь являются примеры первой строки. Если вспомнить, что десятичные дроби суть дроби, у которых знаменателями служат степени десяти, то приведенные в примерах I дроби можно записать еще следующим образом:

$$\frac{1}{5}=\frac{2}{10},\quad \frac{3}{40}=\frac{75}{1000}.$$

Эти равенства не заключают в себе ничего особенного. Они утверждают лишь, что данные обыкновенные дроби можно представить в таком виде, что знаменателями их будут степени 10. Возможно это для всех тех дробей, у которых знаменатель входит множителем в некоторую достаточно высокую степень 10. Простым отличительным признаком такого знаменателя является то, что он не содержит никаких других простых множителей, кроме 2 и 5. Действительно, число $2^\alpha \cdot 5^\beta$ входит множителем в число 10^γ, где γ больше обоих чисел α и β. Следовательно, всякую дробь со знаменателем $2^\alpha \cdot 5^\beta$ всегда можно превратить в десятичную дробь со знаменателем 10^γ.

2. Если в знаменатель данной нам дроби входит, кроме того, еще некоторый другой множитель k, не делящийся ни на 2, ни на 5, то такую дробь превратить описанным способом в десятичную будет, очевидно, уже нельзя; действительно, если допустить, что

$$\frac{1}{2^\alpha 5^\beta k}=\frac{a}{10^\delta}=\frac{a}{2^\delta 5^\delta},$$

то мы должны были бы заключить, что

$$2^{\delta-\alpha}5^{\delta-\beta}=ak,$$

откуда k (если оно не равно 1) должно было бы делиться на

2 или на 5, так как в силу единственности разложения на простые множители в ak, а следовательно также и в k, не может войти никаких других простых множителей, кроме 2 и 5.

Подобного рода примеры приведены в строках II и III. В этих случаях говорят, что дроби $\frac{4}{9}$, $\frac{1}{7}$, $\frac{1}{6}$ и $\frac{7}{30}$ можно превратить в бесконечные десятичные дроби. Но, в сущности говоря, такая дробь вовсе не есть десятичная дробь. Действительно, что является знаменателем такой дроби? Так как в бесконечной десятичной дроби нет последнего знака, то никакая сколь угодно высокая степень десяти не может послужить достаточно большим знаменателем для такой дроби, и значит, никакая несократимая дробь, знаменатель которой делится, например, на 3, как мы уже выяснили, не может быть представлена в виде десятичной дроби.

Под бесконечной десятичной дробью разумеется, следовательно, десятичная дробь в несколько обобщенном смысле слова. Нас интересует здесь только формальная сторона этого нового математического образа. Процесс обращения обыкновенной дроби в десятичную не обрывается, но дает все новые и новые десятичные знаки. Получаемые при этом бесконечные десятичные дроби являются периодическими, т. е. в последовательности десятичных знаков обнаруживаются (начиная с определенного знака) повторения некоторой группы цифр.

Так как существует всего лишь 10 цифр, то некоторые из этих цифр должны, конечно, повторяться в любой бесконечной десятичной дроби; однако неправильно было бы думать, что по этой причине всякая бесконечная десятичная дробь будет дробью периодической. Пример бесконечной, но не периодической десятичной дроби представляет собой дробь

$$0{,}101001000100001\ldots,$$

составленная из нулей и единиц так, что *за n-й единицей следуют всякий раз n нулей.*

3. Примеры дробей, в знаменатели которых кроме 2 и 5 входят еще и другие простые множители, приведены в строках II и III. Относительно таких дробей мы хотим доказать, что они разлагаются в периодические десятичные дроби. Известно, что десятичная дробь получается из соответствующей простой в результате повторного выполнения операции деления, например:

$$1:7=0{,}142857\ldots$$

$$\begin{array}{r|l}
10 & 7 \\ \cline{2-2}
 & 142857\ldots \\
7 & \\ \hline
30 & \\
28 & \\ \hline
20 & \\
14 & \\ \hline
60 & \\
56 & \\ \hline
40 & \\
35 & \\ \hline
50 & \\
49 & \\ \hline
1 &
\end{array}$$

Как только в этом примере мы получаем в остатке 1, вычислительный процесс, который был начат нами также с 1, возобновляется сначала, а последовательности остатков 1, 3, 2, 6, 4, 5 и частных 1, 4, 2, 8, 5, 7 *повторяются в том же порядке* всякий раз заново, и значит, полученная нами десятичная дробь будет иметь п е р и о д 142857. Так как и в дальнейшем нам неоднократно придется иметь дело с подобного рода последовательными делениями, мы введем для них сокращенное обозначение. Мы будем писать, например,

$$\begin{array}{c} 1:7=0{,}\overline{142857}\ldots, \\ \scriptstyle 1\;3\;2\;6\;4\;5\;1 \end{array}$$

указывая мелким шрифтом остатки, относящиеся к соответствующим частным; период бесконечной десятичной дроби мы будем отмечать горизонтальной чертой сверху.

Другим примером может служить дробь

$$\begin{array}{c} 3:41=0{,}\overline{07317}\ldots \\ \scriptstyle 3\;\;30\;13\;7\;29\;3 \end{array}$$

Процесс деления повторяется здесь при появлении остатка 3, так как именно с 3 он и был начат. Ч а с т н ы е могут повторяться при этом и раньше, как, например, в данном случае цифра 7, однако окончательным критерием для повторения всего процесса деления и, следовательно, замыкания периода является повторение о с т а т к а. При последовательных делениях обязательно должен появиться остаток, уже встречавшийся ранее (числитель данной нам обыкновенной дроби причисляется к «остаткам»). Ведь существует только конечное число возможных остатков, а именно при разложении дроби $\frac{a}{b}$, т. е. при делении a на b, возможны остатки: $1, 2, 3, \ldots$ $\ldots, b-1$, т. е. число возможных остатков на 1 меньше знаменателя. Остаток 0 сюда не входит, так как он означает возможность выполнять деление нáцело, и следовательно, приводит к конечной десятичной дроби, что возможно лишь в том случае, когда наш знаменатель составлен из простых множителей 2 и 5. Следовательно, *дробь* $\frac{a}{b}$ *должна дать при обращении периодическую дробь, причем период ее должен содержать, самое большее,* $b-1$ *знаков.* В дроби $\frac{1}{7}$ достигнута именно эта максимальная длина периода; в дроби $\frac{3}{41}$, как показывает наш пример, действительная длина периода значительно ниже возможной. Другой пример максимальной длины периода представляет дробь $\frac{1}{17}$:

$$\frac{1}{17}=0,\overline{0\,5\,8\,8\,2\,3\,5\,2\,9\,4\,1\,1\,7\,6\,4\,7},$$

1 10 15 14 4 6 9 5 16 7 2 3 13 11 8 12 1

с длиной периода, равной $b-1=17-1=16$.

4. Будем сейчас рассматривать только такие дроби, знаменатели которых суть числа, взаимно простые с 10, т. е. не содержат множителей 2 и 5. Для этого случая мы сможем определить максимальное число знаков периода дроби $\frac{a}{b}$ несколько точнее.

Будем предполагать, что дробь $\frac{a}{b}$ уже заранее с о к р а щ е н а, следовательно, a и b — числа взаимно простые. В таком случае при делении в остатках могут появиться только числа взаимно простые с b. Пусть одним из таких взаимно простых с b остатков будет r. Ближайшая очередная операция в вычислениях будет состоять в делении $10r$ на b; пусть получающееся при этом частное будет q. Значит, b в $10r$

будет содержаться q раз, причем может еще получиться остаток r_1:

$$10r = qb + r_1,$$

или

$$r_1 = 10r - qb. \tag{1}$$

Тогда и r_1 должно быть взаимно простым с b, так как ни один простой множитель b не входит ни в 10, ни в r, а значит, и ни в $10r$, ни в $10r - qb$. Следовательно, за каждым взаимно простым с b остатком r следует подобный же остаток r_1.

Так как при этом с самого начала мы исходили из остатка a, взаимно простого с b, значит, и в дальнейшем все возникающие остатки будут взаимно простыми с b. Отсюда мы можем заключить, что *период дроби* $\frac{a}{b}$ *может иметь, самое большее, столько же знаков, сколько может быть остатков, взаимно простых с* b.

Количество остатков, взаимно простых с некоторым числом b, само по себе представляет интересную величину. В теории чисел его обозначают обычно через $\varphi(b)$. Например, $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$, $\varphi(5) = 4$, $\varphi(6) = 2$, $\varphi(7) = 6$, $\varphi(8) = 4$. Так как, в частности, для простого числа p все $p - 1$ предшествующих чисел суть взаимно простые с ним, то справедливо соотношение

$$\varphi(p) = p - 1. \tag{2}$$

Резюмируя, мы можем сказать: *период дроби* $\frac{a}{b}$ *при* b, *взаимно простом с* 10, *содержит, самое бо́льшее,* $\varphi(b)$ *знаков* [66].

5. Мы доказали, что при обращении обыкновенной дроби в бесконечную десятичную период может возникнуть и в том случае, когда в b содержится хотя бы один простой множитель, отличный от 2 и 5. Из конечного числа возможных остатков один обязательно должен повториться. На этом основании мы можем уже утверждать, что период должен быть, однако мы еще не можем указать, где он должен начаться. В примерах II период начинается сейчас же после запятой, в примерах III — лишь через несколько знаков после запятой. При этом знаменатели в примерах II отличаются тем свойством, что в них нет ни одного множителя, общего с 10.

Мы покажем теперь, что период десятичной дроби $\frac{a}{b}$ начинается всегда тотчас же после запятой в том случае, если b и 10 — числа **взаимно простые**. Для этого нам нужно показать, что первый повторяющийся остаток тождественно равен первому остатку, т. е. числителю. Действительно, пусть у нас оказались равными два остатка: $r_m = r_n$; в таком случае и предшествующие им остатки (если таковые имеются) также должны быть равны: $r_{m-1} = r_{n-1}$. Чтобы убедиться в этом, напишем разность двух выражений:

$$10r_{m-1} = q_{m-1} \cdot b + r_m;$$

$$10r_{n-1} = q_{n-1} \cdot b + r_n,$$

и, положив в ней $r_m = r_n$, получим:

$$10(r_{m-1} - r_{n-1}) = (q_{m-1} - q_{n-1})\, b.$$

Значит, $10(r_{m-1} - r_{n-1})$ делится на b. Но так как b — число, взаимно простое с 10, то на b должна делиться разность $r_{m-1} - r_{n-1}$. Эта разность, следовательно, должна встретиться в ряду

$$0, \quad \pm b, \quad \pm 2b, \quad \pm 3b, \ldots$$

Но так как, с другой стороны, r_{m-1} и r_{n-1} являются остатками и, значит, должны быть меньше b, то и абсолютное значение их разности должно быть меньше b. В нашем ряду мы находим для этой разности только одно возможное значение

$$r_{m-1} - r_{n-1} = 0,$$

откуда

$$r_{m-1} = r_{n-1}.$$

Период, таким образом, должен начаться непосредственно после запятой.

6. Если длину периода обозначить через λ, то при обращении дроби $\frac{a}{b}$ в десятичную мы получим после λ делений в остатке снова a. Но ведь при каждом делении мы приписываем к делимому справа нуль, поэтому, получая в остатке a, мы делим собственно не a, а $a \cdot 10^\lambda$. Следовательно, $a \cdot 10^\lambda - a$ делится на b. Но так как

$$a \cdot 10^\lambda - a = a(10^\lambda - 1),$$

а a — число, взаимно простое с b, то выражение $10^\lambda - 1$ должно делиться на b. При этом λ является наименьшим числом, для которого $10^\lambda - 1$ делится на b. Ведь период замыкается уже при первом появлении остатка a, и следовательно, λ является наименьшим числом, для которого $a \cdot 10^\lambda$ при делении на b дает остаток a. Мы имеем, таким образом, теорему:

I. *Длина периода b дроби $\frac{a}{b}$ есть наименьшее число λ, при котором выражение $10^\lambda - 1$ делится на b.*

В этой теореме содержатся два утверждения, заслуживающих того, чтобы обратить на них особое внимание. Во-первых, здесь утверждается, что для данного b, взаимно простого с 10, можно указать такое число λ, что выражение $10^\lambda - 1$ будет делиться на b. Существование такого числа λ само по себе не очевидно и в данном случае основывается на том факте, что λ обозначает здесь число знаков периода, существование которого было в своем месте (п. **3**) доказано. Во-вторых, здесь получается, что λ зависит только от b, но не от a. *Все несократимые дроби $\frac{a}{b}$ с одинаковым знаменателем b имеют одинаковую длину периода* λ, зависимость которой от b мы запишем в виде $\lambda = \lambda(b)$.

7. Рассмотрим теперь совместно несколько разложений $\frac{a}{b}$ в десятичные дроби при различных a, но с одним и тем же постоянным b. Мы уже имели разложение

$$1:7 = 0,\overline{\underset{1}{1}\,\underset{3}{4}\,\underset{2}{2}\,\underset{6}{8}\,\underset{4}{5}\,\underset{5}{7}}\,\underset{1}{\ldots}$$

Для сравнения вычислим

$$2:7 = 0,\overline{\underset{2}{2}\,\underset{6}{8}\,\underset{4}{5}\,\underset{5}{7}\,\underset{1}{1}\,\underset{3}{4}}\,\underset{2}{\ldots}$$

Период 285714, соответствующий дроби $\frac{2}{7}$, получается из периода 142857, соответствующего дроби $\frac{1}{7}$, путем простой циклической перестановки цифр. Это станет сразу же ясным, если обратить внимание на то, что остаток 2, которым начинается разложение $\frac{2}{7}$, встречается также и среди остатков разложения $\frac{1}{7}$, а вследствие совпадения этих остатков должны совпадать и последующие операции деления. Среди остатков разложения $\frac{1}{7}$ встречаются точно так же и остатки 3, 4, 5, 6. И действительно, здесь должны появиться все шесть остатков, возможных при делении на 7, ибо период разложения $\frac{1}{7}$ имеет 6 знаков. Если представить себе, что порядок следования цифр в периоде циклический, т. е. вслед за последней цифрой следует начальная, и каждой цифре периода поставлен в соответствие остаток, из которого она получается в результате деления:

остатки	1	3	2	6	4	5
частные	1	4	2	8	5	7

то из этой таблицы можно тотчас же увидеть, что разложение $\frac{6}{7}$ имеет период, начинающийся с 8 и, следовательно, имеющий вид 857142, откуда $\frac{6}{7}=0{,}857142$, и т. д.

Периодичность обнаруживается, таким образом, не только в появлении через определенные промежутки одинаковых частных, но и в чередовании остатков. Поэтому мы будем теперь говорить не о периодах вообще, а *о периодах частных и периодах остатков.*

8. Не всегда период разложения $\frac{a}{b}$ достигает той максимальной длины $\varphi(b)$, которой он, согласно нашим исследованиям, не может превзойти. Для $\frac{1}{7}$ и $\frac{1}{17}$, как показывают наши примеры, получается максимальная длина периода, для $\frac{3}{41}$ — не максимальная. Какова длина периода в действительности — вопрос трудный, и решается он в каждом конкретном случае путем непосредственного вычисления. Тем не менее мы можем относительно *действительной длины периода* $\lambda(b)$ дать кое-какие указания, более определенные, чем уже найденное неравенство $\lambda(b) \leqslant \varphi(b)$.

Конкретизируем наши рассуждения на новом примере $\frac{1}{21}$, для которого $\lambda(21) < \varphi(21)$. Число остатков, взаимно простых с 21, т. е. $\varphi(21)$, равно 12, как это легко проверить. Однако мы имеем:

$$1:21=0,\overline{0\,4\,7\,6\,1\,9}\ldots,$$
$$\scriptstyle 1\quad 10\;16\;13\;4\;19\;1$$

откуда $\lambda(21)=6$. В процессе деления встречаются, следовательно, не все 12, взаимно простых с 21, остатков, а только 6, которые мы вместе с соответствующими частными соберем в таблице:

остатки	1	10	16	13	4	19
частные	0	4	7	6	1	9

(A)

Из этой таблицы видно, что

$$\frac{10}{21}=0,\overline{476190}\ldots,\quad \frac{4}{21}=0,\overline{190476}\ldots$$

Остаток же 2 здесь не встречается, и для разложения $\frac{2}{21}$ мы должны получить нечто новое. Действительно:

$$2:21=0,\overline{0\,9\,5\,2\,3\,8}\ldots,$$
$$\scriptstyle 2\quad 20\;11\;5\;8\;17\;2$$

что дает таблицу:

остатки	2	20	11	5	8	17
частные	0	9	5	2	3	8

(B)

Здесь новым является не только один остаток 2, но и все остальные. Можно было и заранее предвидеть, что ни один из старых остатков не появится в новом разложении. Ведь каждый остаток однозначно определяет ход разложения, в частности, любой остаток таблицы (A) влечет за собой весь шестичленный период остатков (A). Если бы, следовательно, в периоде остатков (B) содержался хотя бы один остаток из (A), в нем должны были бы содержаться и все остальные остатки из (A), и, так как в нем есть место только для 6 членов, никаких иных остатков в нем не могло бы быть, что не соответствует действительности, так как остаток 2 не встречается в таблице (A). В обеих же таблицах (A) и (B) содержатся теперь как раз все $\varphi(21)=12$ остатков, которые могут быть числителями несократимых правильных дробей со знаменателем 21.

9. В построении таблиц (А) и (В) для знаменателя 21 заложена некоторая общая идея, на которой мы хотим остановиться.

Если разложение $\frac{1}{b}$ имеет максимальную длину периода, $\lambda(b)=\varphi(b)$, то получается только одна таблица периода, как, например, для $\frac{1}{7}$ в п. **7** этой темы.

Если, напротив, $\lambda(b)<\varphi(b)$, как для $b=21$, то при разложении $\frac{1}{b}$ могут появиться только $\lambda(b)$ остатков, из которых мы построили таблицу (А). В этой таблице, значит, содержатся не все $\varphi(b)$ взаимно простых с b остатков, а только $\lambda(b)$, т. е. меньше $\varphi(b)$. Пусть r будет остатком, не входящим в таблицу (А). Разложив дробь $\frac{r}{b}$, мы получим период, в котором будет также $\lambda(b)$ членов, как это мы знаем хотя бы из примеров п. **6**. Получающиеся при этом остатки и частные мы выделим в новую таблицу (В); в этой новой таблице будет не содержащийся в таблице (А) остаток r, и все другие ее остатки также будут отличаться от остатков таблицы (А), ибо наличие любого остатка из (А) влекло бы и наличие всех других остатков из (А), и это исключило бы возможность появления r.

Таблицы (А) и (В) вместе содержат $2\lambda(b)$ различных взаимно простых с b остатков. Этими остатками либо исчерпываются все возможные остатки, т. е. $2\lambda(b)=\varphi(b)$, либо, кроме этих $2\lambda(b)$, существуют еще другие остатки.

Пусть s будет одним из таких остатков, не входящих ни в (А), ни в (В); разложим дробь $\frac{s}{b}$ и построим соответствующую таблицу (С), в которой будет содержаться $\lambda(b)$ новых остатков, не входящих ни в (А), ни в (В). Всего, таким образом, у нас получилось $3\lambda(b)$ различных взаимно простых с b остатков. Если $3\lambda(b)=\varphi(b)$, то возможные остатки исчерпаны. В противном случае повторим весь процесс еще раз и будем строить одну таблицу за другой, пока не будет исчерпан весь запас $\varphi(b)$ остатков. Существенным здесь является то, что один новый остаток влечет за собой всякий раз $\lambda-1$ других остатков.

По окончании такого построения все $\varphi(b)$ взаимно простых с b остатков будут разнесены, скажем, по k таблицам, причем ни один из этих остатков не входит одновременно больше чем в одну таблицу. Таким образом,

$$\varphi(b) = k\lambda(b), \tag{3}$$

и тем самым мы получим теорему:

Длина периода $\lambda(b)$ *есть делитель* $\varphi(b)$[1]).

Для простого числа p, как уже было замечено выше, $\varphi(p) = p - 1$. В частности, мы получаем, что длина периода $\lambda(p)$ дроби $\frac{a}{p}$ для простого числа p есть делитель разности $p - 1$; с примерами, иллюстрирующими этот закон, мы уже встречались в наших вычислениях: $\lambda(3) = 1$, $\lambda(7) = 6$, $\lambda(17) = 16$, $\lambda(41) = 5$.

Распределение остатков по различным периодам можно проследить еще на примере знаменателя 39. Имеем:

$$\begin{array}{l} 1:39 = 0,\overline{0\ 2\ 5\ 6\ 4\ 1}\ldots \\ \scriptstyle 1 \quad 10\ 22\ 25\ 16\ 4\ 1 \end{array} \tag{A}$$

Теперь спрашивается: все ли взаимно простые с 39 остатки вошли сюда? Сразу же видно, что здесь не хватает остатка 2; им мы пользуемся для построения следующей таблицы периодов:

$$\begin{array}{l} 2:39 = 0,\overline{0\ 5\ 1\ 2\ 8\ 2}\ldots \\ \scriptstyle 2 \quad 20\ 5\ 11\ 32\ 8\ 2 \end{array} \tag{B}$$

Мы имеем здесь 6 новых остатков. Наименьший взаимно простой с 39 остаток, не входящий ни в (A), ни в (B), есть 7. Исходя из 7, мы получаем новую таблицу остатков и частных:

$$\begin{array}{l} 7:39 = 0,\overline{1\ 7\ 9\ 4\ 8\ 7}\ldots \\ \scriptstyle 7 \quad 31\ 37\ 19\ 34\ 28\ 7 \end{array} \tag{C}$$

Среди 18 входящих в (A), (B), (C) остатков отсутствует ряд остатков, взаимно простых с 39, из которых 14 наименьший. Это дает нам основание построить новую таблицу:

$$\begin{array}{l} 14:39 = 0,\overline{3\ 5\ 8\ 9\ 7\ 4}\ldots \\ \scriptstyle 14 \quad 23\ 35\ 38\ 29\ 17\ 14 \end{array} \tag{D}$$

Теперь у нас имеется 24 различных взаимно простых с 39 остатка, и этим в действительности исчерпываются все возможные остатки, ибо в промежутке между 1 и 39 общие делители с 39 имеют только те числа, которые делятся или

---

[1]) Несобственный делитель, т. е. само $\varphi(b)$, не исключается из этого числа.

на 3, или на 13; на 3 делится третья часть всех чисел в промежутке между 1 и 39, т. е. 13 чисел. Кроме того, имеется два числа, именно 13 и 26, делящихся на 13 (39 мы уже засчитали среди чисел, делящихся на 3). Всего, следовательно, в промежутке между 1 и 39 находятся 15 чисел, имеющих общие делители с 39. Значит, в этом промежутке остается $39-15=24$ числа, взаимно простых с числом 39. Итак, мы имеем $\varphi(39)=24$, $\lambda(39)=6$, т. е. $\varphi(39)=4\lambda(39)$. Мы имеем здесь 4 периода остатков и частных.

10. Из полученных нами сведений о длине периода выведем еще одно важное заключение; для этого нам придется воспользоваться следующей простой леммой:

Если x и k — целые положительные числа, то x^k-1 делится на $x-1$.

Это утверждение станет тотчас же очевидным, как только мы вспомним о геометрической прогрессии $1+x+x^2+\ldots$ $\ldots+x^{k-1}$, для суммирования которой мы просто умножаем ее на $x-1$:

$$(1+x+x^2+\ldots+x^{k-1})(x-1)=x^k-1,$$

причем при приведении подобных членов после умножения все члены, кроме крайних, взаимно уничтожаются.

Положим теперь $x=10^{\lambda(b)}$, тогда из только что приведенной леммы будет следовать, что $10^{\lambda(b)}-1$ есть делитель разности $10^{k\lambda(b)}-1$.

Выберем, в частности, k из равенства (3). Тогда получим, что

$$10^{k\lambda(b)}-1=10^{\varphi(b)}-1,$$

и значит, $10^{\lambda(b)}-1$ будет делителем $10^{\varphi(b)}-1$. По теореме п. 6 b есть делитель $10^{\lambda(b)}-1$; отсюда мы получаем, что b есть также делитель $10^{\varphi(b)}-1$. Мы приходим, таким образом, к следующей теореме:

Если b есть число, взаимно простое с 10, то $10^{\varphi(b)}-1$ делится на b.

Эта теорема уже не связана с вопросом о десятичных дробях; величина $\varphi(b)$ имеет совершенно независимое от этого вопроса значение. Равенство (2) дает, кроме того, в частности, и следующий факт:

Если p есть простое число, не входящее множителем в 10, то $10^{p-1}-1$ делится на p.

То обстоятельство, что в обеих этих теоремах фигурирует число 10, совершенно несущественно. Оно вошло сюда лишь случайно по той простой причине, что употребляемая нами цифровая система построена на основании 10. В результате исследования иной цифровой системы с произвольным основанием g и изучения соответствующих g-адических дробей мы пришли бы теми же самыми путями к аналогичным выводам:

II. *Если b есть число, взаимно простое с g, то $g^{\varphi(b)}-1$ делится на b*; и в частности:

III. *Если простое число p не является делителем g, то $g^{p-1}-1$ делится на p.*

Здесь мы пришли к теореме, далеко выходящей за пределы нашей специальной темы о периодических десятичных дробях и представляющей собой одну из основных теорем классической теории чисел. Теорему III называют по имени математика, впервые ее доказавшего, теоремой Ферма и именно «малой теоремой Ферма», в отличие от еще не доказанной большой, о которой мы упоминали в теме 15. Теорема II представляет собой обобщение малой теоремы Ферма, данное Эйлером.

Эти теоремы можно иллюстрировать несколькими примерами:

$p=5$:

$$\begin{aligned}2^{5-1}-1&=15=3\cdot 5,\\ 3^{5-1}-1&=80=16\cdot 5,\\ 4^{5-1}-1&=255=51\cdot 5;\end{aligned}$$

$p=7$:

$$\begin{aligned}2^{7-1}-1&=63=9\cdot 7,\\ 3^{7-1}-1&=728=104\cdot 7,\\ 5^{7-1}-1&=15\,624=2332\cdot 7,\\ 10^{7-1}-1&=999\,999=142\,857\cdot 7;\end{aligned}$$

$b=6,\ \varphi(b)=2$:

$$\begin{aligned}5^2-1&=24=4\cdot 6,\\ 7^2-1&=48=8\cdot 6;\end{aligned}$$

$b=9,\ \varphi(9)=6$:

$$\begin{aligned}2^6-1&=63=7\cdot 9,\\ 4^6-1&=4095=455\cdot 9,\\ 5^6-1&=15\,624=1736\cdot 9;\end{aligned}$$

$b=10,\ \varphi(10)=4$:

$$\begin{aligned} 3^4-1 &= 80 = 8\cdot 10,\\ 7^4-1 &= 2400 = 240\cdot 10,\\ 9^4-1 &= 6560 = 656\cdot 10. \end{aligned}$$

11. После этого отступления мы возвращаемся к периодическим десятичным дробям. Мы уже установили, что разложение несократимой дроби $\frac{a}{b}$ со знаменателем, взаимно простым с 10, приводит к периодической десятичной дроби, период которой начинается тотчас же после запятой.

Если, наоборот, нам дана периодическая десятичная дробь, перед нами возникает задача узнать, из какой обыкновенной дроби она получена. Пусть длина периода P данной нам дроби есть λ; будем читать λ-значный период как целое λ-значное число, записанное в десятичной системе. Например, для $\frac{1}{7}=0,\overline{142857}$ период $P=142\,857$. Тогда легко можно найти дробь $\frac{a}{b}$, из которой получена данная нам десятичная дробь с λ-значным периодом P. Действительно, при делении a на b через λ знаков должен появиться вновь остаток a, значит, $a\cdot 10^\lambda - a$ должно делиться на b, причем частным является период P, заканчивающийся как раз при появлении остатка a, т. е.

$$a(10^\lambda-1)=bP;$$

отсюда

$$\frac{a}{b}=\frac{P}{10^\lambda-1}. \tag{4}$$

Дробь, стоящая в правой части, в общем случае может быть сократимой и приводиться к виду $\frac{a}{b}$. Так как числа 2 и 5 являются делителями числа 10^λ, они не могут быть делителями числа $10^\lambda-1$ и, следовательно, не могут быть также делителями b. Этим самым мы установили, что *всякая* так называемая *чистая периодическая дробь*, т. е. такая, период которой начинается тотчас же после запятой, *возникает из такой обыкновенной дроби* $\frac{a}{b}$, *знаменатель которой b есть число, взаимно простое с* 10. Обратную теорему мы уже нашли раньше.

12. До сих пор мы рассматривали только чистые периодические дроби. Но в п. 1 под рубрикой III мы привели примеры и таких дробей, у которых между запятой и периодом имеются также и другие цифры. Такие десятичные дроби называются смешанными периодическими. Так как конечные десятичные дроби возникают из таких простых дробей, знаменатели которых составлены из простых множителей 2 и 5, чистые периодические — из таких, в знаменатели которых множители 2 и 5 не входят, то для смешанных периодических дробей остается третья возможность: в знаменатель соответствующих им простых дробей входят множители, общие с 10, т. е. 2 и 5, и, кроме того, еще другие простые множители, отличные от 2 и 5. Об этом третьем случае мы только упомянем: он не представляет ничего нового, что заслуживало бы подробного исследования.

13. После этих основных фактов остановимся еще на одном, скорее любопытном, чем важном свойстве периодов бесконечных десятичных дробей. Период дроби $\frac{1}{7}$ состоит из 6 цифр: 142857. Разобьем его на две половины и сложим их:

$$142+857=999.$$

Период $\frac{1}{17}$ состоит из 16 цифр, именно 0588235294117647. Поступим с ним точно так же:

$$05882352+94117647=99999999.$$

Для периода $\frac{1}{11}$, т. е. для 09 имеем:

$$0+9=9.$$

Мы покажем теперь, что для периодов дробей $\frac{a}{p}$, *знаменатель которых есть число простое*, это свойство суммы обеих половин периода сохранится во всех случаях, когда вообще период можно разбить пополам, т. е. когда он состоит из четного числа знаков.

Итак, пусть длина периода нашей дроби λ есть число четное: $\lambda=2l$. Обозначим период через P, а его половины через A и B. При этом P мы будем читать как λ-значное число, A и B — как l-значные. Тогда, принимая во внимание зависимость значения цифры от ее места, будем иметь:

$$P=A\cdot 10^l+B.$$

Мы знаем, что обыкновенную дробь $\frac{a}{p}$ можно вычислить по ее периоду P, пользуясь равенством (4):

$$\frac{a}{p}=\frac{P}{10^\lambda-1}=\frac{A\cdot 10^l+B}{10^\lambda-1}. \qquad (5)$$

Мы взяли $\lambda=2l$, так что

$$10^\lambda-1=10^{2l}-1=(10^l-1)(10^l+1). \qquad (6)$$

Из (5) следует, что знаменатель p можно увеличить до величины $10^\lambda-1$, т. е. p есть делитель $10^\lambda-1$. Это, впрочем, мы могли бы заключить также непосредственно из теоремы п. **6**. Если P входит множителем в $10^\lambda-1=(10^l-1)(10^l+1)$, то, будучи числом простым, оно должно входить множителем по крайней мере в один из множителей 10^l-1 или 10^l+1. Число 10^l-1 делиться на p, очевидно, не может, ибо l меньше λ, длина же периода λ, согласно теореме I (п. **6**), есть *наименьшее* число, при котором 10^l-1 делится на p. Следовательно, *число* $10^\lambda+1$ *должно делиться на* p. Из (5) и (6) получаем

$$\frac{a}{p}=\frac{A\cdot 10^l+B}{(10^l-1)\,(10^l+1)}.$$

Умножив обе части этого соотношения на (10^l+1), найдем:

$$\frac{a\,(10^l+1)}{p}=\frac{A\cdot 10^l+B}{10^l-1}.$$

В левой части здесь стоит целое число, так как 10^l+1, как мы только что установили, делится на p. Следовательно, и выражение

$$\frac{A\cdot 10^l+B}{10^l-1}$$

также является целым числом. Но

$$\frac{A\cdot 10^l+B}{10^l-1}=\frac{A\,(10^l-1)+A+B}{10^l-1}=A+\frac{A+B}{10^l-1},$$

и так как A — целое число, то и

$$\frac{A+B}{10^l-1}=h \tag{7}$$

также есть целое число. Мы утверждаем, что h должно быть равно 1. Действительно, A состоит из l цифр и имеет наибольшее значение в том случае, когда все эти l цифр принимают наибольшие значения, т. е. 9. Но число, состоящее из l девяток, можно записать в виде 10^l-1. Поэтому $A \leqslant 10^l-1$. Точно так же и l-значное число $B \leqslant 10^l-1$. На этом основании мы имеем:

$$A+B \leqslant 2(10^l-1).$$

Равенства, однако, здесь не может быть, и наше соотношение должно быть записано более точно:

$$A+B<2(10^l-1). \tag{8}$$

Ибо если бы $A+B=2(10^l-1)$, то *обе* величины A и B должны были бы принять наибольшие значения, и у нас получилось бы

$$A=10^l-1, \quad B=10^l-1.$$

Но в таком случае и A, и B состояли бы из l девяток, а весь период был бы составлен из $2l$ одинаковых цифр 9. Но это бессмысленно, ибо в таком случае уже одна цифра 9 составляла бы период, и этот последний был бы однозначным, между тем как по предположению он имеет четное число знаков. Отсюда мы заключаем, что справедливо неравенство (8). Из (7) находим:

$$A+B=h(10^l-1) \tag{9}$$

где h есть целое (положительное) число, которое, согласно (8), меньше 2. Остается, значит, одна возможность:

$$h=1,$$

и следовательно,

$$A+B=10^l-1,$$

т. е. $A+B$ дает l-значное число, состоящее из одних девяток, что и требовалось доказать [67].

24. ОБ ОДНОМ ХАРАКТЕРИСТИЧЕСКОМ СВОЙСТВЕ ОКРУЖНОСТИ

Когда идет дождь, бывает сыро, но когда сыро, не обязательно, чтобы шел дождь. Этим примером пользуются обычно для того, чтобы разъяснить детям разницу между прямым предложением и обратным. Эта разница, столь очевидная в такой формулировке, нечетко соблюдается в обыденной жизни. Она становится совершенно ясной людям, как только довести ее до их сознания, и, однако, часто видишь остроумных юристов, которые бессознательно путают эти вещи в ежедневной практике; у ораторов такое смешение прямых и обратных предложений часто встречается как искусный метод жонглирования высказываниями противника; обращая их, они делают их тем самым смешными, многочисленная же аудитория обычно не замечает этого обращения. Нам, математикам, это прекрасно известно из преподавательской практики, ибо первое, что мы должны воспитать в начинающем студенте, — это умение избегать всяких, даже подсознательных, ошибок в этом отношении.

Для математика-исследователя осознанный переход от прямой теоремы к обратной является одним из плодотворнейших принципов. Эта и следующая темы посвящены тем путям, которыми этот принцип приводит нас к новым теоремам и новым понятиям; в остальном же эти темы совершенно независимы одна от другой, и читать их можно отдельно.

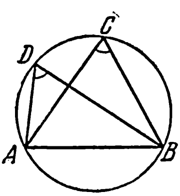


Рис. 94.

1. Начнем с простого примера математической теоремы и ее обращения. Вы еще, конечно, не забыли теорему о вписанных углах (рис. 94): *все вписанные углы, опирающиеся на одну и ту же хорду, равны.* Для нас существенным является то, что здесь и теорема, обратная этой, также справедлива.

Геометрическое место всех точек, из которых данный отрезок AB виден под одним и тем же углом, есть окружность [1]).

[1]) Если D лежит на нижней дуге AB, нужно брать угол, образуемый продолжением за точку D луча AD с DB (иначе мы будем иметь дело с дополнительным углом) [[68]].

Школьное преподавание далеко не всегда освещает то обстоятельство, что здесь заложен центральный пункт всего учения об окружности. Ведь если вместе с этой теоремой справедлива и ей обратная, то это значит, что указанное свойство вписанных углов является характеристическим для окружности, что оно пригодно для определения этой кривой, и это определение можно положить в основу учения об окружности вместо обычного определения, согласно которому окружность есть геометрическое место точек, равноудаленных от некоторой данной точки — ее центра. И по существу все самые интересные теоремы учения об окружности можно доказать лишь после теоремы о вписанных углах, и только основываясь на ней, а вовсе не на классическом определении окружности. В этом смысле наша теорема является краеугольным камнем всего учения об окружности, подлинным определением окружности.

Рис. 95.

2. После этих предварительных соображений обратимся к другому свойству окружности и покажем, что оно также является характеристическим для этой кривой. Предпошлём дальнейшему маленькое замечание. Школа приучила нас, по вполне, впрочем, уважительным причинам, понимать под «углом» наклон двух прямых линий. Но ничто нам здесь не мешает рассматривать также и наклон двух кривых линий; если угодно, его можно истолковать как угол, образуемый двумя касательными к обеим этим кривым в вершине S, и тем самым свести новое понятие к обычному понятию угла между прямыми (рис. 95).

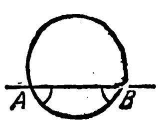


Рис. 96.

Окружность обладает, очевидно, тем свойством, что любая ее хорда образует с нею пару равных, разумеется, криволинейных углов. Эту теорему мы хотим обратить. Мы хотим выяснить, *является ли кривая* (рис. 96), *которая с любой своей хордой образует равные криволинейные углы, всегда окружностью, или существуют еще и другие кривые, обладающие тем же самым свойством?* Мы покажем, что и это свойство вполне характеризует окружность.

Представим себе какую-нибудь замкнутую кривую (рис. 97), которая образует равные углы с любой хордой, соединяющей две произвольные точки этой кривой. Пусть A, B, C — какие-либо точки этой кривой, α, β, γ — углы между кривой и хордой, и $\angle A$, $\angle B$, $\angle C$ — углы между хордами. Помимо трех хорд проведем через эти же точки также три касательные к кривой. Тогда, согласно условиям доказываемой нами теоремы, между полученными при этом углами должны существовать равенства, выраженные в обозначениях рис. 97.

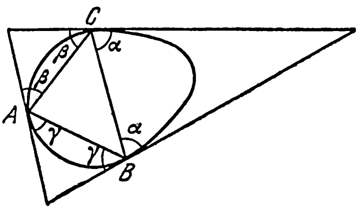


Рис. 97.

С другой стороны, три угла, имеющих вершиной точку A, в сумме должны дать два прямых; то же самое относится к тройкам углов при B и при C. Иначе говоря, мы должны иметь три уравнения:

$$\angle A + \beta + \gamma = \pi,$$
$$\alpha + \angle B + \gamma = \pi,$$
$$\alpha + \beta + \angle C = \pi.$$

Складывая их, получаем:

$$(\angle A + \angle B + \angle C) + 2(\alpha + \beta + \gamma) = 3\pi.$$

Применяя теперь к треугольнику ABC теорему о сумме углов треугольника, имеем:

$$\angle A + \angle B + \angle C = \pi,$$

откуда следует:

$$2(\alpha + \beta + \gamma) = 2\pi$$

или

$$\alpha + \beta + \gamma = \pi.$$

Если сопоставить это последнее равенство с ранее полученными, например с равенством

$$\angle A + \beta + \gamma = \pi,$$

то мы будем иметь $\angle A = \alpha$; точно так же $\angle B = \beta$, $\angle C = \gamma$.

Из этих предварительных рассмотрений нам понадобится лишь результат $\angle C = \gamma$. Обратимся теперь к самому доказательству. Пусть D — какая-либо другая точка кривой; тогда в отношении треугольника ABD мы сможем сделать те же самые заключения, что и в отношении ABC. В треугольнике ABD и на чертеже, аналогичном рис. 97, точки A и B, а вместе с ними и касательные в этих точках остались прежними, равно как и угол γ. Следовательно, $\angle D$ при точке D будет также равен γ, и мы, значит, будем иметь $\angle D = \angle C$. Иными словами, хорда AB будет видна из D под тем же углом, что и из C. Отсюда на основании обратной теоремы о вписанных углах, о которой шла речь выше, D, а значит и любая другая точка нашей кривой, лежит на окружности, определяемой точками A, B, C. Следовательно, эта кривая есть действительно окружность.

В следующей теме мы познакомимся с другим свойством окружности, которое, в противоположность только что рассмотренному, присуще не только окружности, но также и целому ряду других кривых; принцип обращения в этом случае приведет к образованию новых понятий — к замечательному классу кривых, объединяемых этим общим свойством.

25. КРИВЫЕ ПОСТОЯННОЙ ШИРИНЫ

1. Окружность определяется как кривая, все точки которой отстоят на одно и то же постоянное расстояние от некоторой данной точки — центра. Это свойство окружности находит себе непосредственное практическое применение, например в колесе обыкновенной повозки. Так как все спицы имеют равную длину, то втулка при вращении остается на постоянной высоте; этим обеспечивается горизонтальность движения повозки. Для горизонтального перемещения некоторых тяжелых грузов, например ящиков или плит, пользуются иногда не колесами, укрепленными на оси, а более примитивным приспособлением — цилиндрическими катками или валиками (рис. 98). Если поперечное сечение катков круглое, то груз при передвижении остается на постоянной высоте по отношению к пути.

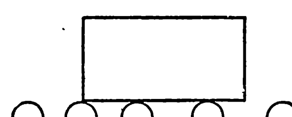


Рис. 98.

Но в то время как обод колеса **обязательно** должен быть очерчен по **окружности**, а его втулка должна приходиться в **центр** этой окружности, ибо всякая другая форма колеса вызвала бы дополнительные перемещения повозки то вверх, то вниз, — для катков, как это ни удивительно, круговая форма сечения вовсе не обязательна: они могут выполнять свое назначение и при некоторых иных формах поперечного сечения, так как центр сечения никакой особой роли здесь не играет. В этом случае скорей имеет значение то обстоятельство, что обе параллельные касательные к окружности удалены одна от другой на одно и то же неизменное расстояние, как бы мы ни вращали круг между ними. Окружность имеет постоянную ширину во всех направлениях, она представляет собой *кривую постоянной ширины*. Можно было бы подумать, что это свойство исчерпывающе характеризует окружность и, подобно свойству, рассмотренному нами в предыдущей главе, присуще исключительно окружности. Однако, вопреки всяким ожиданиям, в действительности дело обстоит иначе, а именно существует множество кривых постоянной ширины, которые не являются окружностями. С некоторыми из них мы познакомимся в дальнейшем.

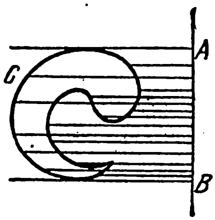


Рис. 99.

2. Ширина какой-либо замкнутой кривой C в заданном направлении определяется следующим образом. Все точки кривой C проектируют ортогонально на любую прямую заданного направления (рис. 99). Совокупность проекций всех этих точек заполняет некоторый отрезок AB прямой, величина которого и выражает ширину данной кривой в этом направлении.

Каждый из двух крайних проектирующих лучей, проходящих через A и B, обладает тем свойством, что вся кривая располагается по одну сторону такого луча, имея, однако, с ним хотя бы одну общую точку. Прямые, обладающие таким свойством в отношении некоторой кривой, называются ее опорными прямыми.

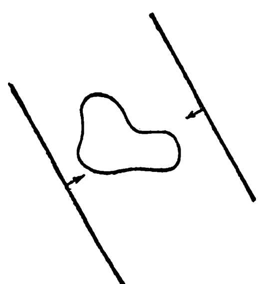


Рис. 100.

Для замкнутой кривой в любом направлении существуют две опорные прямые. Убедиться в этом можно или из рис. 99, повторив для различных направлений выполненные на нем построения, или проведя (рис. 100) две произвольные параллельные прямые заданного направления так, чтобы кривая расположилась между ними, а затем взаимно сближая их, пока они не коснутся кривой.

Нужно, однако, при этом заметить, что понятие опорной прямой не совпадает с понятием касательной. На рис. 101, *а*, например, прямая t, касательная к кривой в точке T, не является ее опорной прямой. На рис. 101, *б* опорная прямая s не является касательной.

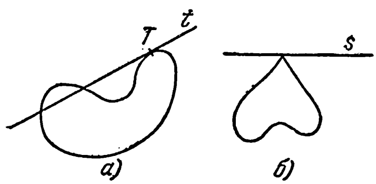


Рис. 101.

Для кривой постоянной ширины b *две любые параллельные опорные прямые удалены одна от другой на одно и то же постоянное для всех направлений расстояние* b. Если к такой кривой провести две пары параллельных опорных прямых (рис. 102), то получающийся при этом параллелограмм будет ромбом. Если эти пары взаимно перпендикулярны, то ромб будет прямоугольный и, следовательно, обратится в квадрат. Стороной этого квадрата будет b — ширина кривой. Поэтому все квадраты, описанные вокруг кривой постоянной ширины, конгруэнтны. Это свойство можно сделать наглядным посредством модели, состоящей из кривой постоянной ширины, вырезанной, например, из картона, и квадратной рамки. Если сторона этой рамки равна ширине кривой, то рамка подходит к кривой при всяком ее положении, т. е. эту рамку можно вращать вокруг кривой постоянной ширины так, что она всегда будет плотно охватывать эту кривую. Обратное соотношение также имеет место: любую кривую постоянной ширины можно вращать внутри квадрата так, что она постоянно будет прилегать к его сторонам, и всякая кривая, обладающая этим свойством, должна быть, очевидно, кривой постоянной ширины [[69]].

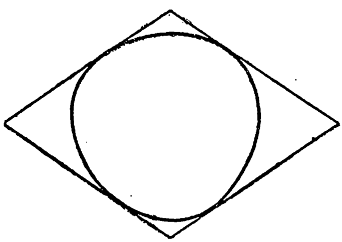


Рис. 102.

3. Простейшая кривая постоянной ширины, не являющаяся окружностью, — это равносторонний треугольник, составленный из дуг окружностей таким образом, что каждая вершина его совпадает с центром противолежащей дуги (рис. 103). Все три дуги имеют один и тот же радиус, который и представляет собой постоянную ширину b кривой. Ведь из двух параллельных опорных прямых либо одна непременно проходит через вершину, а другая касательна к противолежащей дуге, либо обе проходят через вершины, и тогда они обе в этих вершинах касательны к дугам. И в том и в другом случае расстояние между двумя параллельными опорными прямыми равно длине радиуса, перпендикулярного к касательной в точке касания.

Из всех кривых постоянной ширины этот дуговой треугольник первым был замечен в технике. Механик Рело (Reuleaux) в своей классификации механизмов установил, что эта фигура допускает вращение внутри квадрата при постоянном соприкосновении с его сторонами, а это свойство характерно для всех кривых постоянной ширины.

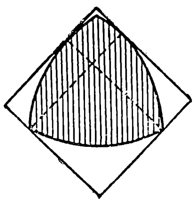


Рис. 103.

4. Лежащая в основе многоугольника Рело конструктивная идея, согласно которой дуги окружностей равного радиуса располагаются так, что каждой вершине противолежит некоторая описанная из этой вершины дуга, может быть реализована самыми разнообразными способами. Некоторую точку B плоскости фиксируют как вершину, и из нее радиусом, равным b, описывают дугу окружности. На этой дуге берут две новые точки A и C. Дуга радиуса b, описанная из C, пройдет через B, так как $BC=b$, согласно только что выполненному построению. Некоторую точку D на этой дуге принимают за новую вершину. Дуга радиуса b, описанная из D, пройдет через C. Чтобы возникающую в таком построении фигуру сделать замкнутой, новую вершину E нужно взять на этой дуге не произвольно, а так, чтобы она одновременно лежала и на дуге, описанной из точки A радиусом b и проходящей через B; иначе говоря, эта вершина должна определяться пересечением двух дуг. Таким путем получается пятиугольный дуговой многоугольник $ADBEC$ постоянной ширины (рис. 104, *а*). Путем повторения аналогичных построений можно получить многоугольники и с большим числом вершин, также являющиеся кривыми постоянной ширины. На рис. 104, *б* показан подобного рода семиугольник. Из того, что против каждой вершины лежит дуга радиуса b с центром в этой вершине, непосредственно следует, что в результате описанного построения должны получаться многоугольники постоянной

ширины. Для некоторых дальнейших применений соединим радиусами каждую вершину этого дугового многоугольника с двумя вершинами противолежащей дуги. В результате внутри нашего дугового многоугольника образуется прямолинейный многоугольник с равными взаимно пересекающимися сторонами. В каждой вершине стороны этого многоугольника образуют центральный угол противолежащей дуги.

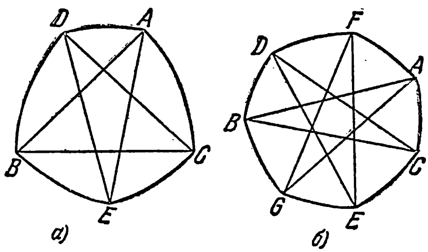
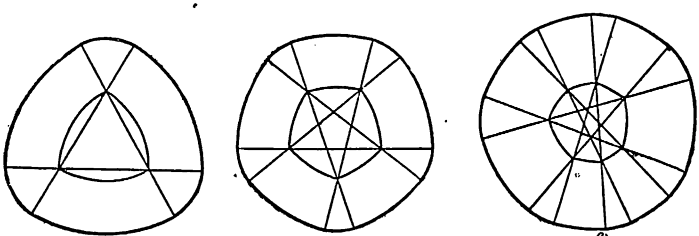


Рис. 104.

Построенные описанным приемом дуговые многоугольники всегда имеют нечетное число сторон. Действительно, отметим какую-либо вершину и противолежащую ей сторону и будем, начиная с этой вершины, обходить дуговой многоугольник вдоль его периметра. Мы пройдем при этом сначала одну сторону, потом вершину, потом опять сторону и снова вершину и т. д., пока не достигнем вершины, непосредственно предшествующей отмеченной ранее стороне. Всего на этом пути между отмеченной вершиной и отмеченной стороной окажется столько же вершин, сколько и сторон; положим, что тех и других будет по n. Но этот переход от отмеченной вершины к отмеченной стороне можно совершить и иначе, начиная движение от отмеченной вершины в противоположном первоначальному направлении. При этом мы также пройдем n сторон и n вершин, так как против каждой стороны первого маршрута лежит вершина второго маршрута, против каждой вершины первого — сторона второго. К этому нужно присоединить еще отмеченную вершину и отмеченную сторону; в общем, таким образом, у нас получится $2n+1$ вершин и столько же сторон.

5. Исходя из рассмотренных выше дуговых многоугольников, имеющих вершины, можно построить также и кривые постоянной ширины без вершин. С этой целью построим кривую, параллельную данному дуговому многоугольнику и расположенную вне его на расстоянии d (рис. 105, *а*, *б*, *в*). Это легко выполнить с помощью начерченного нами диагонального многоугольника. Для этого просто увеличивают радиус каждой дуги на одну и ту же величину d, оставляя прежний центр. Вершина первоначального дугового многоугольника, которую можно рассматривать как дугу радиуса, равного нулю, заменяется в новой кривой дугой радиуса d. Возникающая таким образом кривая составляется из нечетного числа дуг



а) б) в)

Рис. 105.

двух различных радиусов. Две дуги различных радиусов имеют одну общую точку (соответствующую вершине основной кривой).

Наконец, можно прийти и к более общим типам дуговых многоугольников постоянной ширины, если допустить, что число различных радиусов больше двух, и сохранить принцип построения только что начерченных фигур, состоящий в том, чтобы противолежащие дуги были описаны из одного центра и опирались на одинаковые центральные углы (рис. 106).

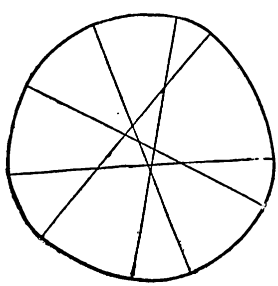


Рис. 106.

Приведенный способ построения дает нам бесчисленное множество кривых постоянной ширины. Все эти кривые отличаются, однако, той особенностью, что они составлены исключительно из дуг окружностей. Во избежание недоразумений здесь же отметим, что существуют и такие кривые постоянной ширины, в которых даже самая ничтожная дуга не принадлежит окружности.

6. После этих отдельных примеров кривых постоянной ширины мы приведем несколько общих теорем, относящихся ко всем кривым этого типа. Во всех наших примерах встречаются только *выпуклые* кривые, т. е. такие, которые имеют лишь две общие точки с секущей прямой. Для простоты мы и в дальнейшем будем рассматривать только выпуклые кривые, имея их в виду даже в тех случаях, когда свойство выпуклости и не будет специально оговорено. Дадим точное определение:

Выпуклой кривой называется кривая, ограничивающая выпуклую область; выпуклая же область характеризуется тем, что если две любые, принадлежащие этой области, точки соединить отрезком, то весь этот отрезок будет также лежать в этой области [70].

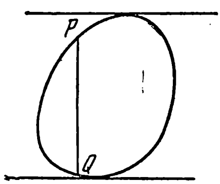


Рис. 107.

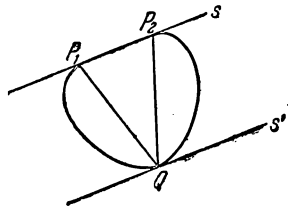


Рис. 108.

Примерами выпуклых областей являются квадрат, круг, треугольник, эллипс, а также вышеприведенные области постоянной ширины. Опорная прямая выпуклой области имеет с ее границей или только одну общую точку или целый общий отрезок. Но для кривых постоянной ширины имеет место следующая теорема:

I. *Кривая постоянной ширины с каждой из своих опорных прямых имеет лишь по одной общей точке.*

Чтобы это доказать, сделаем следующее простое замечание:

II. *Расстояние между двумя точками кривой постоянной ширины b не может превышать b.*

Действительно, если P и Q — точки кривой (рис. 107), то отрезок PQ должен быть расположен между парой опорных прямых, перпендикулярных к этому отрезку, и следовательно, расстояние между этими опорными прямыми не может быть меньше PQ. Так как это расстояние равно b, то отрезок PQ не может превзойти b, что и требовалось доказать.

Чтобы доказать теорему I, допустим обратное: на опорной прямой s лежат две точки P_1 и P_2 кривой (рис. 108). С противоположной стороны кривой проведем вторую опорную прямую s', параллельную данной. Пусть на этой прямой s' лежит точка Q кривой. Расстояние между s и s' пусть будет опять b.

Так как в треугольнике P_1QP_2 не может быть двух прямых углов, то отрезки P_1Q и P_2Q не могут быть одновременно перпендикулярны к s. Поэтому один из них должен быть больше b, что противоречит теореме II. Таким образом, допущение, что наша кривая имеет две общие точки с опорной прямой, ведет к противоречию, и теорема I доказана.

Если мы еще раз воспользуемся тем фактом, что только перпендикуляр к обеим опорным прямым имеет длину b, все же другие линии, соединяющие опорные прямые, больше b, то тотчас же получим следующую теорему:

III. *Отрезок, соединяющий точки касания двух параллельных опорных прямых к кривой постоянной ширины, перпендикулярен к опорным прямым.*

7. Если из какой-либо точки кривой постоянной ширины b описать окружность радиуса b, то, согласно теореме II, эта окружность охватит всю кривую. Покажем теперь, что не вся кривая будет лежать внутри окружности и что по крайней мере одна точка этой кривой лежит на окружности.

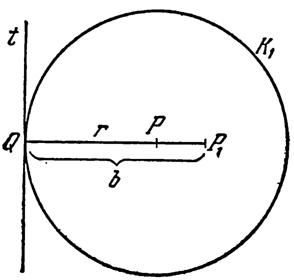


Рис. 109.

Пусть P будет какая-либо точка кривой C постоянной ширины b. Опишем из P как из центра окружность столь большого радиуса, чтобы внутри окружности лежала вся кривая, за исключением по крайней мере одной точки Q (рис. 109), которая лежала бы на окружности. Назовем построенную таким образом окружность K_1. Радиус ее r не может превышать по величине b, так как окружность K радиуса b, описанная концентрически из того же центра P, уже содержит кривую внутри себя. Таким образом, окружность K_1 можно получить из окружности K, если только они не равны, уменьшая последнюю. Касательная t к окружности K_1 в точке Q является опорной прямой не только для окружности, но также и для кривой C, так как она проходит через точку Q кривой C, а заключенная внутри K_1 кривая C лежит по одну сторону ее.

Параллельная t опорная прямая s на другой стороне C отстоит от t на расстояние b, ибо, по предположению, кривая C — постоянной ширины b и, согласно теореме III, точка касания P_1 прямой s лежит на перпендикуляре, восставленном из точки Q к t. Когда $r = b$, точки P и P_1 совпадают; если $r < b$, то P лежит между Q и P_1. Последнее же невозможно, так как тогда три точки Q, P, P_1 должны принадлежать кривой C и лежать на одной прямой. Между тем выпуклая кривая имеет с секущей прямой лишь две общие точки. Прямая могла бы иметь с выпуклой кривой более двух общих точек лишь в том случае, если бы она была опорной прямой. Но, согласно теореме I, кривая постоянной ширины имеет с опорной прямой только одну общую точку. Поэтому P_1 должна совпасть с P, т. е. $r = b$.

Точка P была выбрана нами на C произвольно, и тем не менее мы все же могли построить проходящую через нее опорную прямую s для C. Поэтому мы имеем право высказать такую теорему:

IV. *Через любую точку кривой постоянной ширины проходит по крайней мере одна опорная прямая* [71].

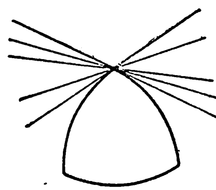


Рис. 110.

На кривой постоянной ширины могут быть такие точки, через которые проходит больше одной опорной прямой. Такие точки мы называем **вершинами**. Кривые с вершинами встречались многократно в наших примерах. Так как все прямые, лежащие внутри угла между двумя опорными прямыми вершины, очевидно, также являются опорными прямыми, то это значит, что через вершину выпуклой кривой проходит целый **пучок** опорных прямых (рис. 110). Среди этих опорных прямых имеются две крайние, ограничивающие пучок.

Если P — произвольная точка C, то по теореме IV мы можем провести через нее опорную прямую s к C. Восставим в точке P перпендикуляр к s; он пересечет C в противолежащей точке Q, причем PQ имеет длину b. Окружность, описанная из Q как из центра радиусом b, объемлет b и касается s. Эти свойства мы резюмируем в нижеследующей теореме:

V. *Через каждую точку P кривой постоянной ширины проходит окружность радиуса b, объемлющая кривую,* *причем опорная прямая, проведенная к кривой через P, касается этой окружности в точке P.*

8. Следующая теорема также устанавливает связь между кривой постоянной ширины и окружностью.

VI. *Радиус окружности, имеющей не меньше трех общих точек с кривой постоянной ширины b, не превышает b.* То, что радиус такой окружности действительно может быть равен b, показывает дуговой треугольник Рело. Стóит только одну из его трех дуг дополнить до полной окружности, и мы увидим, что эта окружность, у которой имеется бесконечно много общих точек с кривой постоянной ширины, имеет радиус b.

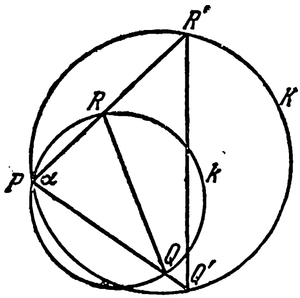


Рис. 111.

Доказательство. Дано, что окружность k имеет три общие точки P, Q, R с кривой C постоянной ширины b (рис. 111).

Из трех углов треугольника PQR по крайней мере один не меньше двух других, т. е. или он больше обоих, или он равен одному и больше другого, или, наконец, он равен двум другим. Пусть этот угол лежит при точке P, обозначим его через α. Проведем теперь через P к данной кривой C опорную прямую и окружность K радиуса b, касающуюся этой прямой в точке P и объемлющую нашу кривую C. Точки Q и R будут, следовательно, лежать или внутри, или на окружности. Но если обе точки Q и R лежат на K, то K и k должны совпадать, так как через три точки P, Q и R можно провести только одну окружность. В этом случае все доказано.

В противном случае продолжим PQ и PR до их пересечения с K в точках Q' и R' (рис. 111). Докажем, что $Q'R'$ больше QR.

Прежде всего, если Q совпадает с Q' (рис. 112, *а*), то точка R' должна быть отлична от R, так как тот случай, когда Q и R вместе лежат на K, уже разобран раньше[1]). Тогда угол $QRR' = \delta$ есть внешний угол треугольника PQR.

[1]) Рис. 112, *а* и *б* нужно рассматривать как особо выделенные детали рис. 111.

Но по известной теореме элементарной геометрии внешний угол больше любого внутреннего угла треугольника, с ним не смежного. В нашем случае, следовательно, $\delta > \alpha$. Далее, β есть внешний угол треугольника QRR' и, следовательно, $\beta > \beta'$. Но в треугольнике PQR угол α был выбран так, что $\alpha \geqslant \beta$. Соединяя эти неравенства вместе, имеем $\delta > \alpha \geqslant \beta > \beta'$; отсюда следует, что $\delta > \beta'$.

В треугольнике QRR' сторона QR' лежит против бо́льшего угла, чем QR, а отсюда, на основании хорошо известной теоремы элементарной геометрии, следует, что $QR' > QR$.

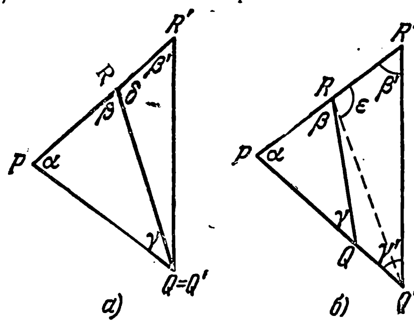


Рис. 112.

Если различными являются не только точки R и R', но также Q и Q' (рис. 112, *б*), то мы замечаем сначала, что $\beta + \gamma = \beta' + \gamma'$, так как каждая из этих пар углов в сумме равна $180° - \alpha$, в силу теоремы о сумме углов треугольника. Поэтому неравенства $\beta' > \beta$ и $\gamma' > \gamma$ несовместимы. Пусть, например, не выполняется первое из этих неравенств (как на рис. 112, *б*), т. е. $\beta' \leqslant \beta$. В таком случае проведем в четырехугольнике $QQ'R'R$ ту из диагоналей, которая не делит угла β', а именно $Q'R$ (в случае $\gamma' \leqslant \gamma$ нужно провести, напротив, QR'). Совершенно так же, как и в случае с рис. 112, *а*, доказываем сначала, что $Q'R > QR$. Обозначим угол $Q'RR'$ через ε. Для треугольника $PQ'R$ угол ε есть внешний, и потому $\varepsilon > \alpha$. Так как мы уже имеем $\alpha \geqslant \beta$ и $\beta \geqslant \beta'$, то $\varepsilon > \beta'$. Значит, в треугольнике $Q'R'R$ сторона $Q'R'$ лежит против бо́льшего угла, чем сторона $Q'R$, откуда следует, что $Q'R' > Q'R$. Так как мы уже установили, что $Q'R > QR$, то получаем $Q'R' > QR$.

Это неравенство, следовательно, всегда имеет место по отношению к рис. 111, к которому мы теперь и возвращаемся. В искомой окружности k угол α является вписанным углом, опирающимся на хорду QR; в окружности K тот же угол α опирается на хорду $Q'R'$. Поэтому обе эти хорды $Q'R'$ и QR соответствуют одному и тому же центральному углу 2α.

Наложив эти углы один на другой, мы получим рис. 113, из которого непосредственно видно, что названные хорды относятся как радиусы соответствующих окружностей. Но поскольку $Q'R' > QR$, то радиус b окружности K должен быть больше радиуса окружности k, что и требовалось доказать.

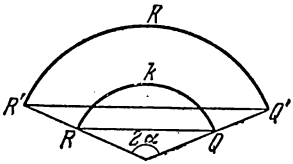


Рис. 113.

9. Простейшая, отличная от окружности кривая постоянной ширины, а именно многоугольник Рело, имеет вершины. Следующая теорема показывает, что одно отличительное свойство вершин этого многоугольника выделяет его среди других кривых постоянной ширины.

VII. *Угол при вершине кривой постоянной ширины не может быть меньше чем* 120°. *Единственной кривой постоянной ширины, имеющей угол при вершине в* 120°, *является треугольник Рело, у которого и остальные два угла при вершинах имеют также по* 120°.

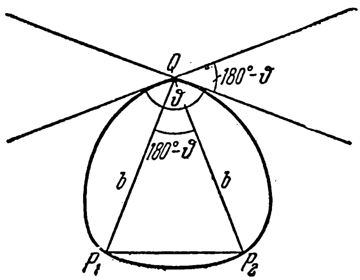


Рис. 114.

Доказательство. Угол при вершине мы изменяем как угол между крайними опорными прямыми пучка при вершине. Если (рис. 114) угол при вершине Q равен ϑ, то для пучка опорных прямых остается отверстие в $180° - \vartheta$. Перпендикуляры, восставленные в Q к каждой из этих опорных прямых пучка, образуют также пучок с тем же самым отверстием $180° - \vartheta$, которое измеряется углом P_1QP_2. Каждый перпендикуляр от точки Q до противолежащей точки кривой имеет длину b (согласно теореме III). Следовательно, против вершины Q на кривой лежит дуга P_1P_2 круга радиуса b, соответствующая центральному углу $180° - \vartheta$. Длина хорды P_1P_2, стягивающей эту дугу, на основании теоремы I не может превзойти ширины кривой b. Поэтому в равнобедренном треугольнике QP_1P_2 с боковыми сторонами, равными b, основание P_1P_2 не может быть больше боковых сторон QP_1 и QP_2. Следовательно, угол при вершине Q в этом треугольнике может быть равен, самое бо́льшее, 60°. Но этот угол P_1QP_2, как уже было показано, равен $180° - \vartheta$, поэтому мы имеем $180° - \vartheta \leqslant 60°$, откуда $\vartheta \geqslant 120°$. Так как угол ϑ соответствует произвольно выбранной вершине Q, то первая часть нашей теоремы доказана.

Если теперь угол при вершине $\vartheta = 120°$, то угол P_1QP_2 равен 60°, и равнобедренный треугольник QP_1P_2 становится при этом равносторонним (рис. 115). Тогда длина P_1P_2 в точности равна b. Так как эта длина в то же время равна ширине кривой, то обе перпендикулярные к P_1P_2 опорные прямые s_1 и s_2 должны проходить через точки P_1 и P_2. Отсюда мы убеждаемся в том, что обе эти точки также должны быть вершинами нашей кривой. Ведь уже установлено, что отрезок кривой P_1P_2 есть дуга окружности. В точке P_1 опорной прямой является не только s_1, но также и касательная t_1 (в точке P_1) к дуге окружности. Эти опорные прямые s_1 и t_1 образуют, как легко заметить, внутренний угол в 120°. Значит, они должны быть крайними опорными прямыми пучка, проходящего через P_1, ибо угол при вершине не может быть меньше 120°. Значит, угол при вершине P_1, а также и при вершине P_2 в точности равен 120°. В таком случае точки P_1 и P_2 обладают теми же свойствами, что и вершина Q. Против них лежат дуги окружностей радиуса b, соответствующие центральным углам по 60°. Но тем самым мы и получили как раз дуговой треугольник Рело, и таким образом, доказали и вторую часть теоремы [72].

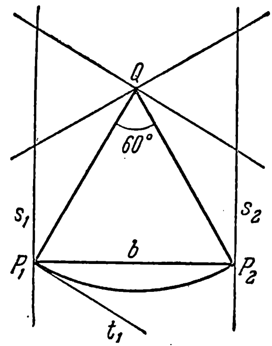


Рис. 115.

10. До сих пор мы построили всего лишь несколько примеров кривых постоянной ширины. Но свойства, установленные в теоремах I—VII, относятся ко всем вообще кривым постоянной ширины, хотя в этих теоремах и не сказано ничего об их существовании. Сейчас мы хотим указать совершенно общий метод построения кривых постоянной ширины. Благодаря ему мы получим представление о всех возможных кривых указанного типа. Теорема V дает особо важное свойство этих кривых. Кривые постоянной ширины достаточно охарактеризованы этим свойством в том смысле, что половину такой кривой между двумя противолежащими точками можно задать произвольно, если только она удовлетворяет условиям теоремы V. Точнее, мы утверждаем следующее:

VIII. *Если мы имеем выпуклую*[1]) *криволинейную дугу* Γ *с хордой длины* b, *обладающую такими свойствами, что*:

1) *вся она лежит между двумя перпендикулярами, восставленными к хорде в концах этой хорды, и*

2) *она объемлется целиком всякой окружностью радиуса* b, *которая проходит через какую-либо ее точку, касательна в этой точке к опорной прямой для* Γ *и расположена по ту же сторону от этой опорной прямой, что и данная дуга* Γ,

то такую дугу можно дополнить до кривой постоянной ширины.

11. Для доказательства возможности получения кривой постоянной ширины из дуги Γ мы находим целесообразным пойти обходным путем и дать это доказательство сначала для о б л а с т и постоянной ширины; тогда искомая к р и в а я постоянной ширины получится как граница области.

Одно предварительное замечание. Если дано несколько областей, то часть области, общую всем данным областям, называют п е р е с е ч е н и е м этих областей. Например, пересечение двух кругов есть дуговой двуугольник (рис. 116).

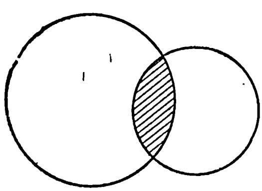


Рис. 116.

Пересечение произвольного множества выпуклых областей само непременно выпуклое.

Для этого достаточно показать, что отрезок, соединяющий две точки, принадлежащие пересечению, сам целиком принадлежит пересечению. Но это очевидно. Действительно, если обе точки P и Q принадлежат пересечению, то это значит, что они принадлежат одновременно всем областям данного множества их. Но выпуклость каждой области влечет за собой то, что, кроме точек Q и P, каждой из них должен принадлежать и весь отрезок PQ. Следовательно, отрезок PQ входит

[1]) То есть дугу, ограничивающую вместе со стягивающей ее хордой выпуклую область.

во все выпуклые области множества, а поэтому он принадлежит и пересечению всех этих выпуклых областей. Для этих рассуждений совершенно безразлично, содержит ли это множество конечное или бесконечное количество выпуклых областей[1]).

12. Пусть нам дан отрезок кривой Г, удовлетворяющий условиям теоремы VIII (рис. 117); хорда AB имеет длину b; вместе с кривой Г она ограничивает выпуклую область G_1; прямые s и t, перпендикулярные к AB и проходящие соответственно через точки A и B, являются опорными прямыми для G_1; всякая окружность радиуса b, касающаяся в некоторой точке Г опорной прямой для Г, проведенной через эту точку, объемлет всю кривую Г целиком.

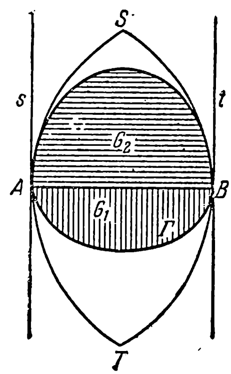


Рис. 117.

Присоединим к области G_1 треугольную область ABS, ограниченную хордой AB, дугой AS окружности с центром в B и дугой BS окружности с центром в A. Обозначим этот дуговой треугольник через G_2. Обе выпуклые области G_1 и G_2 вместе образуют выпуклую область G. Эта последняя ограничена кривой Г и обеими дугами окружностей AS и BS.

Рассмотрим теперь совокупность всех кругов радиуса b, центры которых лежат на Г. Область G и это бесконечное множество кругов имеют выпуклое пересечение D (на рис. 117 оно заштриховано). Мы утверждаем, что это пересечение D представляет собой область постоянной ширины b, причем отрезок кривой Г составляет часть его границы.

Прежде всего пересечение кругов радиуса b с центрами на Г образует область. Проведем два круга радиуса b с центрами в точках A и B. По предположению Г лежит в каждом из этих кругов. Общая часть этих двух кругов, т. е.

[1]) Заметим, что в последнем случае пересечение областей может не оказаться областью, а состоять, например, из одной точки или отрезка, что формально не противоречит высказанному положению, так как и одну точку, и отрезок можно считать выпуклыми. Но для наших целей эти случаи непригодны, поэтому здесь рассматриваются только такие бесконечные множества областей, пересечение которых дает тоже область. (*Прим. ред.*)

дуговой двуугольник $SATBS$, разбивается хордой AB на две части, причем Γ целиком лежит в нижней части TAB. Очевидно, что расстояние между двумя любыми точками TAB не больше, чем b; и, в частности, две точки из Γ отстоят друг от друга также не больше, чем на b. Дуга Γ входит, таким образом, в каждый из рассматриваемых кругов, а следовательно и в выпуклую область D.

Так как D — область выпуклая, то вместе с Γ к ней принадлежит также и любая хорда Γ. Следовательно, G_1 составляет часть D.

Теперь покажем, что наибольшее возможное расстояние между точками области D есть b, и так как D входит в G, т. е. частью в G_1, частью в G_2, то речь здесь может идти или о двух точках из G_1, расстояние между которыми по только что доказанному не может превышать b, или о двух точках D из G_2, для которых справедливо то же самое, или, наконец, о двух точках D, из которых одна, P_1, лежит в G_1, другая, P_2, лежит в G_2. Такие две точки также не могут отстоять одна от другой дальше, чем на b. Действительно, соединим точки P_1 и P_2 прямой и продолжим ее до пересечения с Γ в точке P; все эти три точки будут лежать на прямой в такой последовательности: P, P_1, P_2. Круг радиуса b, описанный из P, содержит D, а значит и эти три точки; следовательно, P_1 и P_2 лежат на одном радиусе длиной b, а потому расстояние между ними не может превысить этой величины.

Таким образом, ширина полученной нами области ни в каком направлении не может превзойти величины b. Докажем, что в любом направлении эта ширина в точности равна b. В направлении AB ширина b была задана. Рассмотрим другое произвольное направление и проведем обе перпендикулярные к этому направлению опорные прямые s_1 и s_2 к D. Одна из них s_1 имеет общую с Γ точку Q. Восставим в этой точке Q перпендикуляр к s длиной b. Его конец обозначим через M и докажем, что M будет принадлежать к D. Для этого нужно доказать, что M лежит в G и во всех кругах радиуса b, описанных вокруг точек Γ. Это будет следовать из того, что M отстоит от точек Γ, самое бóльшее, на b. Действительно, круг радиуса b, описанный из M, касающийся s_1 в G, объемлет по условию весь этот отрезок кривой Γ; это и значит, что любая точка Γ удалена от M, самое бóльшее, на b. Так как, в частности, A и B удалены от M также не больше чем на b, то M должна лежать в дуговом двуугольнике $SATBS$ и именно в той его половине, которая противолежит Γ, т. е. в G_2, и точка M, таким образом, лежит в области G. Этим самым доказано, что M лежит во всех тех областях, пересечение которых есть D. Следовательно, M лежит в D.

Так как QM перпендикулярна к s_1 и s_2, а точки Q и M принадлежат D, то опорные прямые s_1 и s_2 отстоят одна от другой по крайней мере на $QM = b$. Большего расстояния между ними быть не может, так как в противном случае точки касания к ним были бы удалены одна от другой больше, чем на b, а между тем в области D, как уже показано, не может быть таких точек.

Этим самым мы доказали теорему VIII. Легко, впрочем, видеть, что дугу кривой Γ не только можно дополнить до другой кривой постоянной ширины, но что она однозначно определяет эту кривую.

13. В заключение приведем без доказательства еще одно замечательное свойство наших кривых.

Все кривые одной и той же постоянной ширины b имеют один и тот же постоянный периметр, равный, следовательно, периметру окружности диаметра b.

Для дуговых многоугольников, описанных в пп. **4** и **5**, это легко можно было бы доказать на основании подобия дуг окружностей, соответствующих одному и тому же центральному углу. Однако для общего вида кривых постоянной ширины это доказательство выходит из круга наших исследований, так как оно требует уточнения понятия длины кривой [[73]].

26. НЕОБХОДИМОСТЬ ЦИРКУЛЯ В ПОСТРОЕНИЯХ ЭЛЕМЕНТАРНОЙ ГЕОМЕТРИИ

1. Элементарные геометрические задачи на построение в том виде, в каком они встречаются еще у Евклида, разрешаются с помощью линейки и циркуля [[74]]; принято даже выделять элементарную геометрию в рамках общей геометрии именно по этому признаку, т. е. по возможности построения геометрических образов с помощью циркуля и линейки. Но оба эти инструмента играют не совсем точно разграниченные роли: многие задачи можно решать на выбор тем или другим инструментом. Согласно исследованиям Маскерони (Mascheroni) и значительно более ранним работам датского математика Мора (Mohr), можно совсем обойтись без линейки и *все построения, выполнимые с помощью циркуля и линейки, осуществлять одним лишь циркулем* [1]) [[75]].

Якоб Штейнер (Jacob Steiner) достиг весьма многого в противоположном направлении, показав, что все построения элементарной геометрии можно выполнить *одной лишь линейкой, если только предварительно начерчена некоторая постоянная окружность и дан ее центр* [2]) [[76]]. Легко доказать, и это будет сделано в дальнейшем, что одной такой окружностью (без центра) обойтись еще нельзя. Мы покажем именно, что постоянной окружности, *центр которой неизвестен,* недостаточно для того, чтобы все последующие построения можно было выполнить одной линейкой. *Двух непересекающихся* окружностей с неизвестными центрами также еще недостаточно. Но известно, что две пересекающиеся окружности без центров или три пересекающиеся окружности без центров могут заменить в построениях Штейнера окружность с центром (см. стр. 264—265).

2. Так как по способу Штейнера все построения элементарной геометрии могут быть выполнены линейкой, если только дана окружность с центром, то нужно показать, что *ни в окружности, данной нам без центра, ни в двух непересекающихся окружностях нельзя найти центры с помощью одной лишь линейки* [[77]]. Нам требуется, значит, дать два доказательства невозможности. Мы это сделаем, исходя из предположения, что центр или центры могут быть найдены с помощью одной лишь линейки. Придя при этом предположении к противоречию, мы тем самым докажем невозможность такого построения. По существу это доказательство от противного будет опираться на принцип отображения.

Прежде всего заметим, что предположение о невозможности с помощью одной линейки найти центры в двух окружностях является более сильным и уже заключает в себе недо-

---

[1]) При этом прямая линия (которая, разумеется, не может быть начерчена без помощи линейки) задается парой точек; пользуясь циркулем, возможно также найти сколь угодно много точек прямой.

[2]) Впервые это утверждение было сделано Понселе (Poncelet, Traité des propriétés projectives, Paris, 1822). Еще более ранним предшественником Штейнера в этого рода исследованиях нужно считать Ламберта (Freie Perspective, 1744). (*Прим. перев.*)

статочность одной-единственной окружности без центра. Мы начнем, однако, с задачи, относящейся к единственной окружности. Являясь геометрически более простой, она скорее позволит нам вскрыть принципиальную сторону наших рассуждений.

3. Допустим, что для данной начерченной окружности мы сумели каким-то способом, пользуясь одной лишь линейкой, построить центр. Мы проводили, следовательно, какие-то прямые линии, пересекавшиеся с окружностью или между собой, и какие-то точки пересечений соединяли опять прямыми линиями. Так как точку при этом можно фиксировать только прямыми, на которых она лежит, то и центр окружности может быть найден лишь как точка пересечения каких-то прямых. Построенная нами фигура должна была бы поэтому состоять из данной окружности и нескольких прямых линий, из которых две пересекаются в искомом центре.

Исследуем теперь особое отображение этой фигуры, при котором *окружность преобразуется опять в окружность, каждая прямая линия — в прямую линию и каждая точка пересечения — опять в точку пересечения соответствующих линий.* Отображений, удовлетворяющих этим условиям, существует, конечно, очень много; таковым, например, является подобное увеличение или уменьшение фигуры. Для наших целей, однако, отображение подобия как раз непригодно, но мы укажем такое отображение, которое, хотя и переводит окружность в окружность и всякую прямую в прямую, тем не менее совершенно искажает фигуру и прежде всего *центр окружности преобразует в точку, не являющуюся уже центром образа окружности.* Если мы сумеем найти такое отображение, наша теорема будет доказана, ибо действительно, сколь бы сильно отображенная фигура ни отличалась от первоначальной, обе эти фигуры вполне равноправны в отношении выполняемых в них построений. Каждую операцию, производимую в первоначальной фигуре, например проведение прямой, нахождение точки пересечения или соединение двух точек пересечения прямой, мы можем выполнить и притом в той же самой последовательности и в отображенной фигуре, так как и для окружности, и для каждой прямой, и для каждой точки пересечения в первоначальной фигуре мы будем иметь соответствующую в отображенной. Но так как, согласно предположенным свойствам нашего отображения, центр первоначальной окружности не отображается в центр отображенной окружности, то построение в отображенной фигуре не может привести к цели: прямым, пересекающимся на первоначальной фигуре в центре окружности, соответствуют отображенные прямые, точка пересечения которых не совпадает с центром отображенной окружности. Поэтому, хотя все необходимые построения и были бы шаг за шагом выполнены в отображенной фигуре, они не привели бы к нахождению центра окружности. Но это противоречит самому смыслу метода построения. Следовательно, такого метода и не может быть: *построение центра окружности, данной без центра, невыполнимо с помощью одной лишь линейки.*

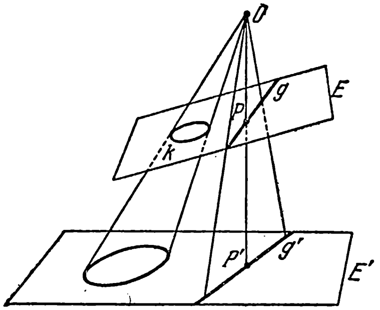


Рис. 118.

Для случая д в у х окружностей доказательство проводится совершенно аналогично.

4. Теперь остается только указать отображение, обладающее указанными свойствами. Мы получим такое отображение с помощью пространственных соображений, а именно с помощью ц е н т р а л ь н о й п р о е к ц и и. Возьмем (рис. 118) какую-нибудь точку O вне плоскости E нашей фигуры и из этой точки O через каждую точку P плоскости E проведем луч. Он пересечет плоскость отображений или плоскость проекций E' в некоторой точке P' — образе точки P. Аналогично точке P на плоскость E' может быть отображена и вся фигура, лежащая в E. Это отображение можно рассматривать как тень, отбрасываемую фигурой в E на плоскость E' от светящейся точки O. При этом очевидно, что любая прямая g отобразится опять в прямую g'. Действительно, совокупность всех проектирующих лучей из O, проходящих через отдельные точки, расположена в плоскости (проектирующей плоскости), определенной точкой O и прямой g; эта плоскость в пересечении с E' должна дать прямую g'.

Но проекция окружности, вообще говоря, уже не будет окружностью. Совокупность лучей, направленных из O к точкам окружности k, образует конус и при этом в общем случае —

косой круговой конус. Прямым называется такой круговой конус, в котором прямая, соединяющая вершину O с центром круга M, перпендикулярна к плоскости круга; все остальные круговые конусы называются косыми. Плоскость проекций E' в пересечении с конусом дает коническое сечение (см. тему 8), которое, как известно, вообще не есть круг. Но для нашей цели окружность необходимо отобразить опять в окружность. Это может быть достигнуто в двух частных случаях.

Первый случай тривиальный. Он имеет место, когда плоскости E и E' параллельны. Тогда отображение, полученное путем проекции, будет, очевидно, отображением подобия, увеличением или уменьшением в зависимости от того, лежит ли E' дальше или ближе от O, чем E. Этот случай, естественно, не годится для нашей цели, так как никакого искажения фигуры здесь не получается и, в частности, центр окружности переходит опять в центр окружности, чего мы как раз и хотим избежать.

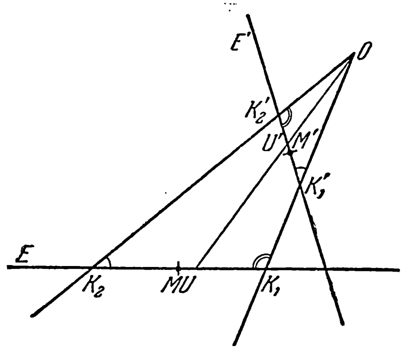


Рис. 119.

Другой случай связан с одной теоремой, которую мы приведем пока без доказательства, отложив его до конца этой главы, чтобы не прерывать здесь хода наших рассуждений. Плоскость, перпендикулярная к плоскости основания косого конуса (рис. 119) и проходящая через центр круга этого основания и через вершину конуса O, есть плоскость симметрии косого конуса. В ней расположены наименьшая и наибольшая образующие конуса OK_1 и OK_2. Эта плоскость симметрии совмещена с плоскостью нашего рисунка. Круг основания проектируется на эту плоскость диаметром K_1K_2, ибо плоскость круга перпендикулярна к плоскости рисунка. Любая плоскость, параллельная плоскости этого круга, пересекает конус, очевидно, опять по окружности. Относительно плоскости, пересекающей обе крайние образующие OK_1 и OK_2 в точках K_1' и K_2' так, что $\angle OK_1'K_2' = \angle OK_2K_1$ и, следовательно, $\angle OK_2'K_1' = \angle OK_1K_2$ (ибо сумма всех углов в треугольнике равна двум прямым), мы будем говорить, что она пересекает конус по сопряженному сечению. Тогда наша вспомогательная теорема, доказательство которой мы откладываем до конца темы, будет гласить:

Сопряженное сечение косого кругового конуса есть круг.

Так как все плоскости, параллельные круговому сечению, образуют в пересечении с конусом также круги, то на основании этой теоремы мы имеем в косом круговом конусе два параллельных семейства круговых сечений, ничем не отличающихся друг от друга.

Проекция из точки O лежащего в E круга K_1K_2 на круг $K_1'K_2'$ в E' — как раз такая, которая была нужна для нашего доказательства: как мы сейчас увидим, она не преобразует центр M круга K_1K_2 в центр M' круга $K_1'K_2'$. Прежде всего заметим, что треугольники K_1OK_2 и $K_1'OK_2'$ имеют общую биссектрису угла при вершине O. Так как биссектриса в каждом из упомянутых треугольников делит противоположные стороны K_1K_2 и $K_1'K_2'$ в. отношении двух боковых сторон, а эти последние по условию (ведь наш конус косой) не равны, то эта биссектриса не может пройти ни через середину M стороны K_1K_2, ни через середину M' стороны $K_1'K_2'$. При этом точки M и M' должны расположиться по разные стороны от биссектрисы. Чтобы в этом убедиться, отобразим зеркально треугольник $OK_1'K_2'$ от биссектрисы OU'. Так как этот треугольник подобен треугольнику OK_1K_2, то в результате зеркального отображения он займет такое положение, что $K_2'K_1'$ ляжет параллельно K_1K_2. Но при этом середины этих сторон M' и M расположатся по одну сторону от биссектрисы (так как последняя делит отрезки K_1K_2 и $K_2'K_1'$ в одном и том же отношении); значит, до отображения они лежали по разные стороны от нее. Отсюда следует, что в первоначальном положении точки M и M' не могут быть получены одна из другой с помощью проектирования.

5. Если принять во внимание, что теорему относительно сопряженных сечений косого кругового конуса мы обещали доказать отдельно в конце главы, то в остальном доказательство того, что с помощью линейки невозможно найти центр в данной окружности, нужно считать законченным. Резюмируем его для ясности: фигура в плоскости E, состоящая из окружности K_1K_2 и нескольких прямых линий, проектируется из точки O на плоскость E', причем прямые переходят в прямые, а окружность K_1K_2 — в окружность $K_1'K_2'$. Центр же M окружности K_1K_2 не проектируется при этом в центр M' окружности $K_1'K_2'$. Построение из одних прямых линий, определявшее центр окружности в E, не могло бы дать центр окружности в E'. Поэтому искомое построение невозможно вообще [78].

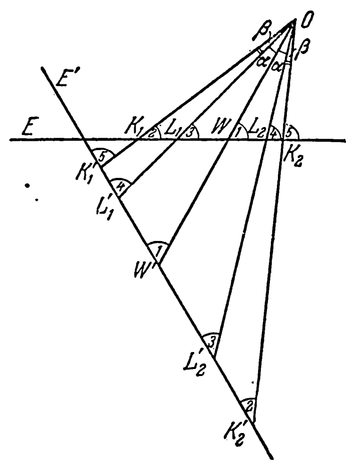


Рис. 120.

6. Более труден выбор отображения при д в у х заданных окружностях. Проектируя из центра O, мы получаем все время два косых конуса, и нужно сделать так, чтобы плоскость E' пересекала о б а по сопряженным сечениям.

Будем различать два случая: Во-первых, *одна окружность может лежать целиком внутри другой*. Выберем плоскость чертежа, перпендикулярную к E и проходящую через оба центра M и N обеих окружностей. Диаметры окружностей изобразятся тогда (рис. 120) отрезками K_1K_2 и L_1L_2. Если бы нам удалось центр проекций O выбрать так, чтобы исходящие из него биссектрисы углов K_1OK_2 и L_1OL_2 совпали, то мы могли бы сказать: сопряженные круговые сечения в обоих конусах расположены параллельно одно другому и потому могут быть взяты в одной и той же плоскости E'. Для этого нужно, чтобы биссектриса OWW' встречала плоскость E' под тем же самым углом, что и E, только противоположно ориентированным. Тогда, как это непосредственно видно из рис. 120, одинаково отмеченные углы окажутся равными, причем $\angle L_1OW = \angle L_2OW$ и $\angle K_1OW = \angle K_2OW$ по построению. Таким образом, плоскости E и E' действительно образуют сопряженные сечения. Две расположенные одна внутри другой окружности K_1K_2 и L_1L_2 спроектируются в две окружности $K_1'K_2'$ и $L_1'L_2'$, также лежащие одна внутри другой.

Далее, так же как и выше, всякое построение для нахождения центра окружности с помощью линейки приводит к абсурду. Построение, выполненное с помощью одних прямых в плоскости E, отображается в подобное же построение в E'. Но так как хотя бы один из двух центров M и N не переходит при этом в центр окружности, то рассматриваемое построение не приводит к результату в плоскости E'. А поскольку обе плоскости E и E' совершенно равноправны, то такого построения не может быть вообще.

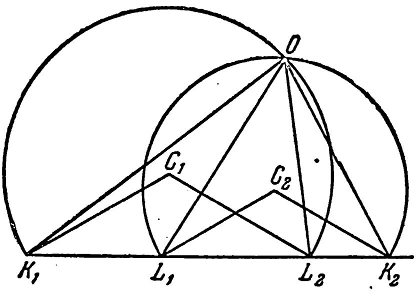


Рис. 121.

Теперь нужно только суметь выбрать вершину O таким образом, чтобы было выполнено наше требование о совпадении биссектрис в двух треугольниках OK_1K_2 и OL_1L_2. Приняв во внимание, что $\angle K_1OL_2 = \angle K_2OL_1$ (это равенство получается в результате сложения равенств: $\angle K_1OW = \angle K_2OW$ и $\angle L_2OW = \angle L_1OW$), будем поступать следующим образом: возьмем произвольный угол $\delta = \angle K_1OL_2 = \angle K_2OL_1$ и построим на K_1L_2 и K_2L_1, как на хордах, окружности так, чтобы углы $\angle K_1OL_2$ и $\angle K_2OL_1$ были вписаны в эти окружности и опирались соответственно на хорды K_1L_2 и K_2L_1 (рис. 121)[1]. Эти две дуги обязательно должны пересечься, так как их хорды перекрываются. Обозначим точку пересечения дуг через O. Мы имеем тогда $\angle K_1OL_2 = \angle K_2OL_1$ и, следовательно, $\angle K_1OL_1 = \angle K_2OL_2$. Биссектриса угла L_1OL_2 делит, следовательно, пополам также и угол K_1OK_2, что нам и нужно было получить [79].

7. Остается разобрать еще случай, когда *обе данные окружности с неизвестными центрами расположены одна вне другой*. Здесь мы снова получаем два отдельно расположенных

[1]) Если центры этих окружностей суть C_1 и C_2, то центральные углы $K_1C_1L_2$ и $L_1C_2K_2$ также должны быть равны. Отсюда для построения рис. 121 получаем

$$\angle C_1K_1L_2 = \angle C_2L_1K_2 = 90^\circ - \delta.$$

проектирующих конуса, и при каком бы то ни было положении O в треугольниках OK_1K_2 и OL_1L_2 общую биссектрису найти невозможно. Мы должны тогда прибегнуть к помощи дополнительного конуса (рис. 122). Если плоскость E' пересекает дополнительный к OL_1L_2 конус $OL_1'L_2'$ и в пересечении с конусом OK_1K_2 дает сечение $K_1'K_2'$, сопряженное с плоскостью E, то: во-первых, биссектриса OV угла L_1OL_2 пересекает E под тем же самым углом, под которым продолжение этой прямой — биссектриса OV' угла $L_1'OL_2'$ — пересекает плоскость E', и, во-вторых, биссектриса OU угла K_1OK_2 пересекает плоскости E и E' под равными углами.

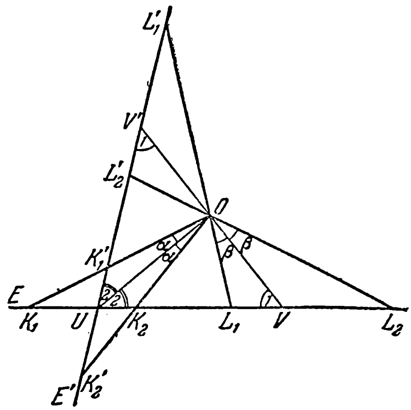


Рис. 122.

Ориентировка плоскостей и прямых показана на рис. 122. Из положения VOV' следует, что треугольник UVV' равнобедренный. Из равенства углов, образуемых OU с E и E' получается, далее, что UO есть биссектриса угла при вершине U в равнобедренном треугольнике UVV'. Так как в равнобедренном треугольнике биссектриса угла при вершине есть в то же время и высота, т. е. перпендикулярна основанию (VV'), то углы VOU и $V'OU$ прямые. При обозначениях рис. 122 мы имеем, кроме того:

$$\angle K_1OL_1 = 90° + \alpha - \beta,$$
$$\angle K_2OL_2 = 90° - \alpha + \beta,$$

откуда

$$\angle K_1OL_1 + \angle K_2OL_2 = 180°. \qquad (1)$$

Точка O должна быть, следовательно, расположена так, чтобы для углов при O имело место соотношение (1). Но тогда биссектрисы этих углов должны быть взаимно перпендикулярны, ибо

$$2\angle UOV = 2\angle UOK_2 + 2\angle K_2OL_1 + 2\angle L_1OV =$$
$$= (\angle K_1OU + \angle UOK_2 + \angle K_2OL_1) + (\angle K_2OL_1 +$$
$$+ \angle L_1OV + \angle VOL_2) = \angle K_1OL_1 + \angle K_2OL_1 = 180^\circ,$$

откуда

$$\angle UOV = 90^\circ.$$

Для этого плоскость E' должна быть расположена именно так, как показано на рисунке, т. е. с обоими конусами она должна давать сечения, сопряженные относительно E. Проекция плоскости E на E' из точки O дает опять то необходимое отображение, из которого мы, как и раньше, делаем вывод, что построение центров обеих окружностей K_1K_2 и L_1L_2 с помощью одних лишь прямых невозможно.

Чтобы найти теперь требуемое задачей положение O, мы можем из двух углов $\angle K_1OL_1 = \varphi$ и $\angle K_2OL_2 = \psi$ один выбрать произвольно, позаботившись лишь о том, чтобы их сумма была равна 180°. Описываем стягиваемые хордами K_1L_1 и K_2L_2 дуги окружности со вписанными углами φ и ψ. Так как K_1L_1 и K_2L_2 по предположению пересекаются, то и две эти дуги должны пересекаться. Точка их пересечения O и будет обладать требуемыми свойствами.

Этим заканчивается доказательство теоремы о том, что двух непересекающихся окружностей с неизвестными центрами недостаточно для выполнимости всех построений элементарной геометрии с помощью одной линейки [80].

Дополнение

В п. 4 мы сформулировали теорему о сопряженных сечениях косого кругового конуса и в дальнейшем ею неоднократно пользовались, отложив, однако, ее доказательство. Сейчас мы это доказательство приводим.

Представим себе, что косой круговой конус, как и раньше, начерчен у нас в разрезе (рис. 123). Плоскость чертежа проходит через вершину O конуса и перпендикулярна к кругу основания, который она пересекает по диаметру K_1K_2. При этих условиях, очевидно, плоскость чертежа будет плоскостью симметрии косого конуса. Наша задача заключается в том, чтобы указать еще другую плоскость симметрии.

Около треугольника OK_1K_2 опишем окружность и проведем в нем биссектрису угла K_1OK_2, которая при продолжении пересечет окружность в некоторой точке M. Так как углы K_1OM и K_2OM равны, должны быть также равны и дуги K_1M и K_2M как соответствующие равным вписанным углам. Перпендикуляр, опущенный из центра C окружности на хорду K_1K_2, разделит эту хорду пополам и поэтому при продолжении встретит окружность также в точке M. Если заставить окружность вращаться вокруг оси MCM', то она опишет поверхность вращения — шар. Окружность K_1K_2 основания конуса и его вершина O лежат на поверхности этого шара, следовательно, конус вписан в шар. Точка M при этом равноудалена от всех точек основной окружности.

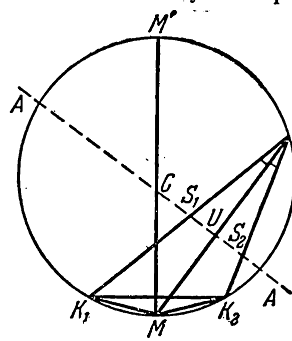


Рис. 123.

Прямая OM играет в конусе особую роль, в силу которой ее можно считать осью конуса. А именно, если через эту ось провести плоскость, то она в пересечении с конусом дает треугольник (рис. 124) с вершиной в O и двумя другими вершинами H_1 и H_2 на окружности основания; поверхность шара она пересечет по окружности, описанной около треугольника OH_1H_2, причем точка M будет лежать на этой окружности, в результате чего мы получим рисунок, вполне аналогичный рис. 123. Ось OM должна быть к тому же и биссектрисой угла H_1OH_2, так как, в силу положения точки M относительно окружности основания конуса, обе хорды H_1M и H_2M должны быть равны, вследствие чего должны быть равны и опирающиеся на них вписанные углы H_1OM и H_2OM. Мы приходим, таким образом, к следующему результату: *ось косого конуса обладает тем свойством, что всякая проходящая через нее плоскость дает в сечении с конусом треугольник, в котором эта ось является биссектрисой.*

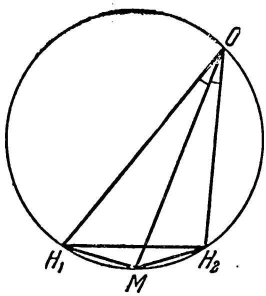


Рис. 124.

Пересечем теперь конус плоскостью A, перпендикулярной к оси конуса; эта плоскость будет перпендикулярна также и к плоскости рисунка, в которой расположена эта ось. След этой секущей плоскости на рис. 123 обозначен через AA. В пересечении плоскости A с поверхностью косого конуса получится некоторое сечение. Если плоскость A принять за плоскость рисунка, то это сечение будет иметь вид, изображенный на рис. 125. Ось конуса, следовательно, перпендикулярна к этой новой плоскости чертежа. Старая плоскость чертежа в пересечении с новой A дает след ZZ. Пусть произвольная проходящая через ось плоскость B дает в сечении с A след BB, на котором точки T_1, T_2 являются точками конуса. Так как в треугольнике T_1OT_2 ось UO перпендикулярна к T_1T_2 и, на основании выше сформулированного свойства, делит угол T_1OT_2 пополам [1]), то T_1U должно быть равно T_2U. Точка U, значит, является центром симметрии конического сечения, лежащего в плоскости A [2]).

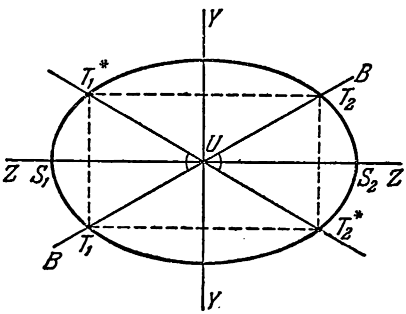


Рис. 125.

Так как старая плоскость чертежа (рис. 123) была плоскостью симметрии косого конуса, то след ее в новой плоскости должен быть осью симметрии полученного нами сечения (рис. 125), а в таком случае в этом сечении должны существовать зеркальные отображения точек T_1 и T_2 относительно оси ZZ, т. е. точки T_1^* и T_2^*.

Теперь легко убедиться в том, что прямая YY, перпендикулярная к ZZ, должна быть также осью симметрии нашей фигуры, ибо четыре точки T_1, T_2^*, T_2, T_1^* образуют прямоугольник, серединой которого является точка U (в силу того,

[1]) На рис. 125 это, естественно, не показано, так как точка O не лежит в плоскости рисунка: ее нужно представить себе на перпендикуляре к этой плоскости над точкой U.

[2]) Более подробный анализ мог бы показать, что это сечение есть эллипс. Но в данном случае это нас не интересует.

что $T_1U = T_2U$, будем иметь также $T_1^*U = T_2^*U$); так как вертикальные углы $\angle T_1UT_1^*$ и $\angle T_2UT_2^*$ равны, то равны также и треугольники $T_1^*T_1U$ и $T_2^*T_2U$, а значит, $T_1T_1^* = T_2T_2^*$. Далее, отрезки $T_1T_1^*$ и $T_2T_2^*$, будучи оба перпендикулярны к ZZ, параллельны, следовательно, $T_1T_2^*T_2T_1^*$ есть параллелограмм. Но диагонали этого параллелограмма T_1T_2 и $T_1^*T_2^*$ равны, поэтому он представляет собой прямоугольник. Подобное рассуждение применимо, конечно, не только для нашей частной фигуры, — можно высказать совершенно общую теорему:

Фигура, имеющая ось симметрии и центр симметрии, имеет также и вторую ось симметрии, которая проходит через центр симметрии перпендикулярно первой оси.

В применении к косому конусу это означает, что плоскость, проходящая через его ось и перпендикулярная к плоскостям чертежей рис. 123 и 125, представляет собой вторую плоскость симметрии косого конуса, ибо в этой плоскости лежит прямая yy, которую нужно вообразить себе перпендикулярной к плоскости чертежа рис. 123 в точке U.

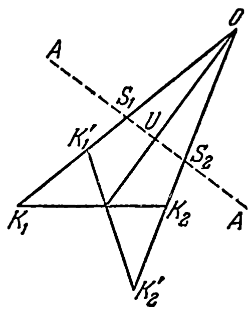


Рис. 126.

Если отобразить теперь конус зеркально относительно этой второй оси симметрии, то образ конуса совпадает с ним самим; это обстоятельство и есть выраженное другими словами свойство симметрии. Но при этом отображении окружность K_1K_2, расположенная на конусе, переходит в свое отображение $K_1'K_2'$, которое, конечно, также является окружностью и также должно лежать на конусе (рис. 126). Эта окружность и есть сечение, сопряженное с сечением K_1K_2. Плоскость этого сечения занимает относительно оси конуса положение, которое требуется определением сопряженного сечения. Вообще говоря, окружность K_1K_2 не совпадает со своим зеркальным отображением; это может случиться только в том случае, когда плоскость окружности перпендикулярна к оси конуса. Но тогда, вопреки условию, мы имели бы дело не с косым, а с прямым круговым конусом. Этим самым теорема о сопряженных сечениях косого конуса, которой мы пользовались выше, доказана.

27. ОБ ОДНОМ СВОЙСТВЕ ЧИСЛА 30

Как 10, так и 21 — числа не простые, но среди множителей, на которые разлагаются эти числа, нет ни одного такого, который содержался бы одновременно в обоих этих числах. Поэтому числа 10 и 21 называют **взаимно простыми**. Числа 6 и 10 нельзя назвать взаимно простыми, так как у них есть общий множитель 2.

Число 10 — взаимно простое со следующими меньшими его числами в пределах от 1 до 9: 3,7,9; из них 9 — число не простое, хотя оно и взаимно простое с 10. Иначе обстоит дело с числом 12. Здесь из ряда чисел в пределах от 1 до 11 взаимно простыми с 12 являются: 5, 7, 11, т. е. только простые числа.

Кроме числа 12, тем же самым свойством, как это легко обнаружить путем несложной проверки, обладают и числа

$$3,\ 4,\ 6,\ 8,\ 12,\ 18,\ 24,\ 30.$$

Путем проб никаких других чисел этого типа обнаружить не удается. Возникает вопрос: *является ли* 30 *последним числом, обладающим тем свойством, что все взаимно простые с ним и меньшие его числа являются простыми?* Докажем, что это именно так.

Начнем с одного чрезвычайно простого замечания. Начиная с 4, любое число N рассматриваемого вида должно делиться на 2, ибо если бы оно было нечетным, оно было бы взаимно простым с числом 4, не являющимся простым. Точно так же число N, большее 9, должно делиться по тем же самым соображениям на 3, а так как оно, сверх того, делится на 2, значит, оно делится и на $2\cdot 3 = 6$. Продолжая подобные рассуждения дальше, получим такую таблицу:

Начиная с 4, число N должно делиться на . . .	$2 = 2$
Начиная с 9, число N должно делиться на . . .	$2\cdot 3 = 6$
Начиная с 25, число N должно делиться на . .	$2\cdot 3\cdot 5 = 30$
Начиная с 49, число N должно делиться на . . .	$2\cdot 3\cdot 5\cdot 7 = 210$
Начиная с 121, число N должно делиться на . .	$2\cdot 3\cdot 5\cdot 7\cdot 11 = 2310.$

Между 4 и 9 остается, значит, исследовать только 4, 6, 8, между 9 и 25 — лишь 12, 18, 24, между 25 и 49 — только 30 (60 — ближайшее следующее число, делящееся на 30, уже превышает 49), между 49 и 121 — ни одного числа, так как 210 уже больше 121 (30 было еще меньше 49). Подобным

же образом можно было бы продолжить наши рассуждения дальше, и мы немедленно доказали бы высказанное выше утверждение, если бы знали, что обнаруженная нами закономерность останется в силе и в дальнейшем, т. е. если 13^2 будет меньше произведения $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11$, если 17^2 будет меньше произведения $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13$, и вообще, если при $n > 4$

$$p_{n+1}^2 < p_1 p_2 p_3 \ldots p_n, \qquad (1)$$

где через $p_1, p_2, p_3, \ldots$, как в теме 1, обозначены последовательные простые числа 2, 3, 5, 7, 11, 13, 17, ...

Доказательство Евклида, приведенное в теме 1, дает

$$p_{n+1} < p_1 p_2 \ldots p_n.$$

Нам же нужно доказать нечто большее, а именно, что

$$p_{n+1} < \sqrt{p_1 p_2 \ldots p_n}.$$

С другой стороны, доказывая неравенство (1), мы тем самым решаем проблему, поставленную в этой теме. Поэтому мы и займемся сейчас лишь доказательством соотношения (1), которое и само по себе представляет достаточный интерес.

В действительности удовлетворяется гораздо большее, чем требует неравенство (1): ближайшее простое число, следующее за $p_5 = 11$, не только меньше чем $\sqrt{2 \cdot 3 \cdot 5 \cdot 7 \cdot 11} = \sqrt{2310} = 48{,}06\ldots$, оно равно всего-нáвсего 13, и т. д., но при чрезвычайных неправильностях ряда простых чисел необычайно трудно высказать что-либо общее для всего ряда. Весьма сложными вспомогательными средствами Чебышев показал, что следующее за p_n простое число $p_{n+1} < 2p_n$. Это, конечно, значительно сильнее того, что утверждается соотношением (1). Но возникает вопрос, нельзя ли значительно более слабое неравенство (1) доказать просто и притом элементарными средствами.

Одному студенту из Мюнстера, Бонзе, в 1907 г. удалось под руководством своего учителя М. Дена прийти к чрезвычайно остроумному выводу, который не только не нуждается в тех вспомогательных средствах анализа и бесконечных процессах, которыми пользуется Чебышев, но и вообще не требует почти никаких математических познаний [81].

1. Основная идея доказательства непосредственно связана с теоремой Евклида, приведенной нами в теме 1. Только вместо того, чтобы из n простых чисел составлять выражение

$$N = p_1 p_2 \ldots p_n + 1 \text{ или } M = p_1 p_2 \ldots p_n - 1,$$

мы возьмем из всего ряда этих чисел $p_1 p_2 \ldots p_n$ лишь i членов и вместо одного выражения M или N образуем из них нижеследующие p_i выражений:

$$\begin{aligned}
M_1 &= p_1 p_2 \ldots p_{i-1} \cdot 1 - 1, \\
M_2 &= p_1 p_2 \ldots p_{i-1} \cdot 2 - 1, \\
M_3 &= p_1 p_2 \ldots p_{i-1} \cdot 3 - 1, \\
M_4 &= p_1 p_2 \ldots p_{i-1} \cdot 4 - 1, \\
& \ldots\ldots\ldots\ldots \\
M_{p_i} &= p_1 p_2 \ldots p_{i-1} \cdot p_i - 1.
\end{aligned}$$

Относительно этих выражений так же, как и относительно выражений, рассмотренных Евклидом, можно высказать такие утверждения:

а) *ни одно из этих выражений не делится ни на одно из чисел* $p_1, p_2, \ldots, p_{i-1}$;

б) *из всех этих выражений не больше, чем одно, может делиться на* p_i. Действительно, если бы какие-либо два из этих выражений, например $p_1 \ldots p_{i-1} x - 1$ и $p_1 \ldots p_{i-1} y - 1$, делились на p_i, то делилась бы на p_i и разность этих выражений $p_1 \ldots p_{i-1}(x - y)$, а так как ни один из первых $i - 1$ множителей на p_i не делится, на него должна была бы делиться разность $x - y$. Но x и y суть какие-то из чисел ряда $1, 2, 3, 4, \ldots, p_i$, и разность двух таких чисел, даже когда меньшее из них равно 1, а большее p_i, равна всего лишь $p_i - 1$ и, следовательно, во всяком случае меньше p_i. Таким образом, поскольку меньшее число не может иметь множителем большее число, разность $x - y$ не делится на p_i;

в) точно так же *только одно из этих выражений может делиться на* p_{i+1}, *только одно из них может делиться на* p_{i+2} *и т. д., только одно из них может делиться на* p_n.

Мы приходим к первому важному заключению: если число чисел $p_i, p_{i+1}, \ldots, p_n$ меньше, чем число выражений $M_1, \ldots, M_{p_i}$, т. е., иными словами, если

$$n - i + 1 < p_i, \tag{2}$$

то среди выражений M имеется по крайней мере одно, *не делящееся ни на одно из чисел* $p_i, p_{i+1}, \ldots, p_n$, так как каждое из простых чисел p_i, p_{i+1}, ..., p_n входит множителем, самое бо́льшее, лишь в одно из выражений M_i, ..., M_{p_i}. В силу утверждения а), в это выражение не может также входить множителем ни одно из чисел $p_1, \ldots, p_{i-1}$, и мы приходим к выводу, что оно *не делится* вообще *ни на одно из первых n простых чисел*. Пусть это будет выражение M_h. Теперь доказательство в точности следует образцу евклидовой методики. Выражение M_h или само должно быть простым числом, и в таком случае оно обязательно должно быть больше p_n, или оно разлагается на несколько простых множителей, среди которых, однако, не может встретиться ни одного множителя из ряда $p_i, \ldots, p_n$ и которые все, следовательно, должны быть больше p_n. Но как бы то ни было, должно существовать некоторое простое число, бо́льшее p_n, входящее множителем в M_h или равное ему, и следовательно, не большее M_h. Так как из всех выражений M наибольшим является M_{p_i}, то очевидно, что наше число во всяком случае не превышает M_{p_i}. Мы не утверждаем при этом, что определенное таким образом число есть как раз ближайшее, следующее за p_n, т. е. p_{n+1}. Но если это и не так, то оно во всяком случае еще больше чем p_{n+1}, и потому p_{n+1} тем более не может превышать M_{p_i}:

$$p_{n+1} \leqslant p_1 \ldots p_{i-1} p_i - 1 < p_1 \ldots p_i.$$

Предварительный результат исследования сводится, таким образом, к следующему: *если имеет место соотношение* (2), *то справедливо неравенство*:

$$p_{n+1} < p_1 \ldots p_i. \tag{3}$$

2. Результат наших исследований представляет собой, очевидно, у с и л е н и е того, что дает доказательство Евклида:

$$p_{n+1} < p_1 \ldots p_n.$$

Так как i меньше n, то правая часть неравенства (3) уменьшена по сравнению с неравенством Евклида. Спрашивается лишь, как велико это уменьшение. Действительно, условие (2) не позволяет нам брать для i произвольно малые значения, i должно быть настолько велико, чтобы количество чисел $p_i, \ldots, p_n$, равное $n - i + 1$, было меньше p_i, т. е. первого из них. Постараемся получить представление о значении этого условия на одном простом примере.

Пусть $n=5$: это значит, что мы будем иметь дело только с 5 первыми простыми числами 2, 3, 5, 7, 11. Выберем для i значение 2, тогда $p_i=3$, и группа чисел $p_i, \ldots, p_n$ будет состоять из чисел 3, 5, 7, 11, причем число их $n-i+1=$ $=5-2+1=4$ будет не меньше $p_i=3$, т. е. первого из них. Оказалось, таким образом, что мы выбрали для i чересчур низкое значение. Возьмем для i значение на 1 большее: $i=3$, тогда $p_i=5$ и чисел $p_i, \ldots, p_n$ будет у нас уже только 3, именно (5, 7, 11), и это количество 3 действительно меньше первого из этих трех чисел, именно 5. То что всякое большее значение i тем более должно удовлетворять требованию (2) — не нуждается ни в каких пояснениях.

Положение, которое мы хотим теперь доказать, формулируется следующим образом: *если выбрать для i возможно меньшее значение, удовлетворяющее условию* (2), *то имеет место соотношение:*

$$p_1 \ldots p_i < p_{i+1} \ldots p_n. \tag{4}$$

Это положение верно при $n=5$, ибо, как это видно из непосредственного вычисления, $2 \cdot 3 \cdot 5 < 7 \cdot 11$. Чтобы выяснить, остается ли оно справедливым и дальше, нам нужно представить себе яснее, каким образом происходит выбор наилучшего i при возрастающем n.

Если от $n=5$ мы перейдем к $n=6$, т. е. будем рассматривать первые 6 простых чисел, то очевидно, что для p_i мы можем сохранить прежнее значение 5. Действительно, $n-i+1$ было меньше чем 5 на 2 единицы. Если теперь к числам группы присоединить еще число 13, величина $n-i+1$ возрастает до 4, но все же останется меньше чем $p_i=5$. Только когда мы перейдем к $n=7$ и наша группа будет заключать в себе 5 чисел (присоединится число 17), это число окажется не меньше $p_i=5$. Здесь, следовательно, значение i должно быть увеличено на 1, и тогда $p_i=7$. При этом группа может заключать до 6 чисел, т. е. числа $7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23$. Выбор $p_i=7$ достаточен, следовательно, для ближайших следующих трех шагов, т. е. для $h=7, 8, 9$; далее мы должны взять $p_i=11$ и, так как на этот раз ближайшее следующее за 7 простое число больше его не на 2, а на 4 единицы, этот выбор будет достаточен не для 3 ближайших следующих шагов, а уже для 5. Лучше всего понять эти соотношения из нижеприведенной таблицы, в которой p_i

выделены жирным шрифтом:

$n=5$: 2, 3, **5**, 7, 11
$n=6$: 2, 3, **5**, 7, 11, 13
$n=7$: 2, 3, 5, **7**, 11, 13, 17
$n=8$: 2, 3, 5, **7**, 11, 13, 17, 19
$n=9$: 2, 3, 5, **7**, 11, 13, 17, 19, 23
$n=10$: 2, 3, 5, 7, **11**, 13, 17, 19, 23, 29
$n=11$: 2, 3, 5, 7, **11**, 13, 17, 19, 23, 29, 31
$n=12$: 2, 3, 5, 7, **11**, 13, 17, 19, 23, 29, 31, 37
$n=13$: 2, 3, 5, 7, **11**, 13, 17, 19, 23, 29, 31, 37, 41
$n=14$: 2, 3, 5, 7, **11**, 13, 17, 19, 23, 29, 31, 37, 41, 43
$n=15$: 2, 3, 5, 7, 11, **13**, 17, 19, 23, 29, 31, 37, 41, 43, 47
. . . .

Выяснив, каким образом происходит выбор наилучшего i при возрастающем h, проверим теперь справедливость высказанного выше положения (4). То что оно остается верным и для $n=6$, не вызывает сомнений, ибо если

$$2\cdot 3\cdot 5<7\cdot 11, \tag{5}$$

то и

$$2\cdot 3\cdot 5<7\cdot 11\cdot 13, \tag{6}$$

так как при переходе от $n=5$ к $n=6$ i осталось прежним.

При переходе от $n=6$ к $n=7$ положение вещей меняется. В левой части здесь появляется множитель 7, в правой — 17. Необходимо выяснить, справедливо ли неравенство

$$2\cdot 3\cdot 5\cdot 7<11\cdot 13\cdot 17. \tag{7}$$

Решить это непосредственно на основании неравенства (6) без вычислений мы не можем. Неравенство (6) останется, однако, верным, если слева присоединить к нему множитель 7, а справа 17; мы получаем тогда неравенство

$$2\cdot 3\cdot 5\cdot 7<7\cdot 11\cdot 13\cdot 17,$$

в котором справа имеется нежелательный множитель 7. Если мы примем за основу не (6), а (5), то неравенство (7) сможем получить из него без всяких вычислений. Для этого в левую часть неравенства (5) внесем множители $7\cdot 7$, а в правую — $13\cdot 17$; тогда получим $2\cdot 3\cdot 5\cdot 7\cdot 7<7\cdot 11\cdot 13\cdot 17$. Сократив на 7, находим искомое неравенство (7).

Эти рассуждения были связаны с переходом от $n=6$ к $n=7$, но так как на всех этапах нашего доказательства мы обошлись без каких бы то ни было вычислительных операций, то ясно, что это доказательство пригодно в равной мере и для всякого последующего перехода. Основывается оно на том простом факте, что при возрастании n число i, как это ясно видно из нашей таблицы, никогда не подвергается увеличению два раза последовательно в двух соседних переходах,— всякий раз при этом происходит «передышка», охватывающая в действительности не только два, но и три (а часто и больше) перехода. Соотношение, аналогичное (7), имеет место и при всяком i.

3. Умножив обе части неравенства (4) на $p_1 \ldots p_i$, получим $(p_1 \ldots p_i)^2 < p_1 \ldots p_n$, или $p_1 \ldots p_i < \sqrt{p_1 \ldots p_n}$. Сопоставив этот результат с неравенством (3), получим соотношение (1), которое нам и требовалось доказать.

4. Нужно еще добавить, что этот же метод позволяет доказать, что для $n > 5$ имеет место более сильное неравенство: $p_{n+1} < \sqrt[3]{p_1 \ldots p_n}$. Это следует довольно просто из того, что «передышка», на которой основывался вывод п. **2,** охватывает всегда по крайней мере три перехода вместо использованных двух, что позволяет распространить наш вывод и на кубический корень. В то же время абсолютно ясно, что распространять его непосредственно таким же образом дальше нельзя.

Читатель согласится, конечно, с тем, что все это доказательство требует не большей математической подготовки, чем тема первая, т. е. собственно ничтожную долю из обычной школьной математики, и апеллирует лишь к голому рассудку. Тем яснее оно показывает, сколь внутренне богатой, сколь трудной может быть математика не только в силу многообразия отдельных областей, сочетаемых и надстраиваемых ею для достижения единой цели, но и в тех случаях, когда, отрекаясь порой от этого многообразия, исходя из самых простых данных (как в разобранном случае — из понятия простого числа), она создает построения, поражающие своей глубиной. Если эта последняя тема и превосходит по трудности предыдущие, то тем яснее и с тем большей чистотой звучит в ней лейтмотив всей книги.

ДОПОЛНЕНИЯ И ПРИМЕЧАНИЯ

[1] У Евклида (см. «Начала» Евклида, книги VII—X, М.—Л.. Гостехиздат, 1949, книга IX, предложение 20, стр. 89) теорема формулируется дословно следующим образом: *первых* (*простых*) *чисел существует больше всякого предложенного количества первых чисел*. Для греческой математики характерно отсутствие в этой формулировке термина «бесконечность», недопустимого, по мнению древних, в точной науке.

[2] Доказательство Евклида является первым доказательством неограниченности ряда простых чисел; в настоящее время известно очень много разных доказательств этого. Одно, более позднее доказательство будет рассмотрено ниже в этой книге (см. тему 20). Здесь мы рассмотрим два других, весьма простых доказательства, отличных как от доказательства Евклида, так и от излагаемого в теме 20 доказательства Эйлера.

Доказательство первое. Составим следующий ряд чисел:

$$2+1=3,\quad 2^2+1=5,\quad 2^4+1=17,\quad 2^8+1=257,$$
$$2^{16}+1=65\,537,\ \ldots,\ 2^{(2^n)}+1,\ \ldots$$

Мы утверждаем, что *каждые два числа этого ряда являются взаимно простыми*. В самом деле, рассмотрим два числа $2^{(2^k)}+1$ и $2^{(2^l)}+1$, где $k>l$. Поскольку $(2^{(2^k)}+1)-2=2^{(2^k)}-1=(2^{(2^l)})^{(2^{k-l})}-(1)^{(2^{k-l})}$ делится на $2^{(2^l)}-1$ («разность четных степеней двух количеств делится на сумму оснований»), то число $2^{(2^k)}+1$ при делении на $2^{(2^l)}+1$ дает в остатке 2, что и доказывает наше утверждение. А из того, что существует бесконечная последовательность целых чисел, каждые два из которых взаимно просты, сразу же вытекает, что *различных простых чисел существует бесконечно много* — ибо если мы имели бы всего N разных простых чисел, то нельзя было бы указать более N попарно взаимно простых целых положительных чисел (см. далее тему 12; мы здесь ограничиваемся целыми числами, отличными от 1, поскольку по отношению к единице вообще нельзя ставить вопрос о том, является ли это число взаимно простым с другим).

Доказательство второе. Предположим, что существует всего n различных простых чисел 2, 3, 5, 7, 11, ..., p_n; подсчитаем, сколько, в таком случае, может существовать натуральных чисел, не превосходящих заданного числа N. Выпишем прежде всего все степени чисел 2, 3, 5, 7, 11, ..., p_n, не превосходящие N:

$$1=2^0,\ 2,\ 2^2,\ 2^3,\ 2^4,\ \ldots,\ 2^{r_1};$$
$$1=3^0,\ 3,\ 3^2,\ 3^3,\ 3^4,\ \ldots,\ 3^{r_2};$$
$$1=5^0,\ 5,\ 5^2,\ 5^3,\ 5^4,\ \ldots,\ 5^{r_3};$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$1=p_n^0,\ p_n,\ p_n^2,\ p_n^3,\ p_n^4,\ \ldots\ p_n^{r_n}.$$

При этом, так как $2^{r_1}\leqslant N$, $r_1\log 2\leqslant\log N$ (логарифмы берутся по любому основанию), и следовательно:

$$r_1\leqslant\frac{\log N}{\log 2};$$

аналогично этому получаем

$$r_2\leqslant\frac{\log N}{\log 3},\ r_3\leqslant\frac{\log N}{\log 3},\ \ldots,\ r_n\leqslant\frac{\log N}{\log p_n}.$$

Итак, мы пока имеем не более $\frac{\log N}{\log 2}+1$ степеней 2, $\frac{\log N}{\log 3}+1$ степеней 3, ..., $\frac{\log N}{\log p_n}+1$ степеней числа p_n, не превосходящих N. Если мы образуем какие угодно произведения этих степеней, то число этих произведений не превзойдет

$$\left(\frac{\log N}{\log 2}+1\right)\left(\frac{\log N}{\log 3}+1\right)\left(\frac{\log N}{\log 5}+1\right)\cdots\left(\frac{\log N}{\log p_n}+1\right)$$

(поскольку можно $\frac{\log N}{\log 2}+1$ способами выбирать первый член произведения, $\frac{\log N}{\log 3}+1$ способами — второй член произведения и т. д.). Так как из всех множителей наибольшим является первый, то последнее выражение заведомо меньше, чем

$$\left(\frac{\log N}{\log 2}+1\right)^n;$$

тем более оно меньше, чем

$$\left(2\,\frac{\log N}{\log 2}\right)^n=\left(\frac{2}{\log 2}\right)^n\cdot(\log N)^n.$$

Но известно, что *логарифмическая функция* (при любом основании системы логарифмов, большем 1) *растет медленнее любой степенной функции: при любых* $c>0$, $k>0$ *и* $a>1$ *найдется такое значение* N, *что при* $x>M$

$$\log_a x<cx^k$$

(см., например, С. И. Новоселов, Специальный курс элементарной алгебры, М., Сов. наука, 1958, § 108). Отсюда следует, в частности, что при достаточно большом N

$$\log N < \frac{\log 2}{2} N^{1/n}; \quad \left(\frac{2}{\log 2}\right)^n (\log N)^n < N.$$

Таким образом, если бы различных простых чисел имелось бы лишь конечное число n, то этих чисел, их степеней и произведений их степеней не хватило бы на то, чтобы породить все натуральные числа, откуда и вытекает бесконечность множества простых чисел.

[3] Доказательство Евклида оставляет открытым вопрос о том, сколь часто встречаются простые числа в ряду натуральных чисел, другими словами, какую величину будет иметь n-е по величине простое число p_n (где n — заданный номер). Из него следует лишь, что *при* $n > 1$

$$p_n < 2^{(2^n-1)}.$$

В самом деле, очевидно, $p_1 = 2$, $p_2 < 2^2 = 4$, $p_3 < 2^4 = 16$ и т. д., и если предположить, что неравенства такого типа вплоть до $p_k < 2^{(2^k-1)}$ имеют место, то в силу рассуждения Евклида получим:

$$p_{k+1} \leqslant p_1 p_2 \ldots p_k + 1 < 2^{1+2+\ldots+2^{k-1}} + 1 = 2^{2^k-1} + 1 < 2^{(2^k)},$$

что, в силу принципа математической индукции, доказывает справедливость нашего утверждения. Однако неравенство $p_n < 2^{(2^n-1)}$, из которого следует, что *число* $\pi(N)$ *простых чисел, не превосходящих данного* N, *больше* $\log_2 \log_2 N + 1$, является очень грубой оценкой для величины $\pi(N)$. Более точные оценки этой величины были получены в ряде классических работ, обзор которых имеется в любом курсе теории чисел (см., например, И. В. Арнольд, Теория чисел, М., Учпедгиз, 1939, гл. III, или А. А. Бухштаб, Теория чисел, М., Учпедгиз, 1960, гл. 35).

[4] Доказательство существования пробелов в ряду простых чисел дано здесь по Л. Кронекеру. См. L. Kronecker «Vorlesungen über Zahlentheorie», I, Leipzig, 1901, стр. 68.

[5] Подобными же элементарными приемами можно доказать существование бесконечного множества простых чисел в некоторых других аналогичных последовательностях, например в последовательностях 3, 7, 11, 15, ..., $4n-1$, ... или 5, 11, 17, 23, 29, ..., $6n-1$ ...; доказательства здесь почти не отличаются от приведенного в тексте и могут быть предоставлены читателю.

[6] Доказательство того, что последовательность чисел 1, 4, 7, 10, 13, ..., $3n+1$, ... или последовательность 1, 5, 9, 13, 19, 21,... ..., $4n+1$, ... содержит бесконечно много простых чисел значительно сложнее разобранных в тексте доказательств; кроме той идеи, которую использовал еще Евклид, здесь необходимо использовать и некоторые другие соображения (см. по этому поводу, например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы элементарной математики, ч. 1, М., Физматгиз,

1959, задача 254 б), А. А. Бухштаб, Теория чисел, М., Учпедгиз, 1960, стр. 349, или Г. Хассе, Лекции по теории чисел, М., ИЛ, 1953, стр. 210 и след.). Аналогично этим последним доказательствам можно убедиться в существовании бесконечного числа простых членов прогрессии 1, 11, 21, 31, 41, ..., $10n+1$, ... (см. Д. О. Шклярский и др., Избранные задачи ..., ч. 1, задача 254 в)) или произвольной прогрессии 1, $2p+1$, $4p+1$, ..., $2kp+1$, ..., где число p простое, и некоторых других прогрессий.

Что же касается до общей теоремы, утверждающей, что *во всякой арифметической прогрессии, первый член которой не имеет общих делителей с разностью прогрессии* (ясно, что если это условие не выполнено, то все члены прогрессии будут делиться на первый член a_1), *имеется бесконечно много простых чисел*, то она была первоначально доказана известным немецким математиком П. Дирихле в его знаменитой работе 1837 г. (см. также П. Г. Лежен Дирихле, Лекции по теории чисел, М—Л., ОНТИ, 1936, стр. 296 и след.). Доказательство Дирихле не является элементарным, и в течении долгих лет не было видно никаких элементарных подходов к доказательству этой замечательной теоремы. Элементарное доказательство теоремы Дирихле было впервые получено в 1949 г. (через 112 лет после Дирихле!) видным датским математиком А. Сельбергом; в настоящее время имеется целый ряд таких доказательств (см., например, Э. Трост, Простые числа, М., Физматгиз, 1959, стр. 88 и след.).

Весьма сложным является вопрос об оценке количества простых чисел, не превосходящих заданной величины, содержащихся в арифметической прогрессии. Этой старой проблеме было посвящено много глубоких исследований; обзор полученных в этом направлении результатов можно найти в курсах теории чисел (см., например, упомянутые выше курсы: И. В. Арнольд, Теория чисел, гл. III, и А. А. Бухштаб, Теория чисел, гл. 36). Другой вопрос, связанный с теоремой Дирихле, заключается в определении по первому члену a и разности d прогрессии наименьшего принадлежащего ей простого числа. В этом отношении наибольший интерес представляет глубокий результат ленинградского математика Ю. В. Линника, доказавшего в 1944 г. *существование такого универсального показателя k, что во всякой арифметической прогрессии с разностью d и взаимно простым с d членом a содержится простое число, не превосходящее d^k*; однако доказательство Линника не позволяет, к сожалению, ничего сказать о величине показателя k.

Теорему Дирихле можно также сформулировать следующим образом: *всякая линейная функция $y=dx+a$ с целочисленными взаимно простыми коэффициентами d* (разность прогрессии) *и a при x, пробегающем все натуральные значения* ($x=1, 2, 3, \ldots$) *бесконечное число раз, принимает простые значения*. Аналогично этому можно поставить вопрос о том, принимает ли конечное или бесконечное число раз простые значения квадратичная функция $y=cx^2+bx+a$ с целыми взаимно простыми коэффициентами a, b, c или какая-либо еще более сложная функция натурального аргумента x. Однако ни одна задача подобного рода, отличная от задачи о простых числах в арифметических прогрессиях, до сих пор не решена, и мы не имеем никаких подходов к решению таких задач. В частности остается без ответа старый вопрос о том, конечно ли или бесконечно число простых чисел, превышающих полный квадрат на единицу (т. е. принимает ли конечное или бесконечное число раз простые значения функция $y = x^2 + 1$ натурального аргумента x).

В известном смысле обратный характер имеет задача о нахождении таких многочленов $y = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n$ с целыми взаимно простыми коэффициентами $a_1, a_2, \ldots, a_n$ или иных функций натурального аргумента x, *которые принимали бы лишь простые значения* (или принимали бы составные значения не более чем конечное число раз; это условие несущественно отличается от предыдущего, поскольку если функция $y = f(x)$ при всех натуральных x, больших N, принимает лишь простые значения, то функция $y_1 = f(x + N)$ принимает простые значения уже при всех натуральных x). Еще Эйлер указал многочлен $x^2 - 81x + 1681$, принимающий простые значения при всех натуральных x, где $1 \leqslant x \leqslant 80$. Однако Эйлер же доказал, что это обстоятельство имеет случайный характер — каждый многочлен с целочисленными коэффициентами при бесконечно многих натуральных значениях аргумента принимает значения, являющиеся составными числами (ср. Д. О. Шклярский и др., Избранные задачи и теоремы ..., ч. 1, задача 219; И. В. Арнольд, Теория чисел, стр. 94; А. А. Бухштаб, Теория чисел, стр. 34). Не оправдалось также и предположение знаменитого французского математика XVII века П. Ферма, считавшего, что все числа $2^{(2^n)} + 1$, где n — натуральное число, являются простыми — составным оказалось уже число $2^{32} + 1$. За последние годы было доказано существование целого ряда функций $f(x)$, принимающих при всех натуральных значениях x лишь простые значения (или даже все последовательные простые значения), — так, в 1947 г. Миллс доказал существование такого (не целого!) числа A, что при всех натуральных x число $[A^{(3^x)}]$, где прямые скобки обозначают целую часть числа, является простым, а в 1951 г. Нивен установил даже, что для всякого $c > {}^8/_3$ существует такое (зависящее, разумеется, от c) число A, что все числа $[A^{(c^x)}]$, где x — натуральное число, являются простыми; в 1951 г. Райт доказал существование такого числа A, что все числа $[2^A]$, $[2^{(2^A)}]$, $[2^{(2^{(2^A)})}]$ и т. д. являются простыми (см. Э. Трост, Простые числа, стр. 50); в 1952 г. известный польский математик В. Серпинский доказал существование такого числа A, что n-е простое число p_n можно записать формулой $p_n = [10^{(2^n)} \cdot A] - 10^{(2^{n-1})} \cdot [10^{(2^{n-1})} \cdot A]$, и одновременно сходную формулу получил датский математик Банг и т. д. Однако все эти результаты являются неэффективными и не позволяют обнаруживать новые простые числа: во всех них фигурирует некоторое число A, которое не может быть указано точно (или нахождение которого само требует знания всех простых чисел).

[7] Замечание относительно взглядов Шпенглера развито более подробно О. Теплицем в его статье «Математика и античность» в журнале Die Antike 1, 1925, стр. 175—203. Точку зрения о том, что истоками современной математики являются достижения (а частично — и неудачи) греческих ученых, аргументированно защищает также известный немецкий математик В. Бляшке в статье «Грече-

ская и наглядная геометрия», см. сборник «Математическое просвещение», вып. 2, 1957, стр. 111—130; вып. 3, 1958, стр. 101—138.

[8] Это решение задачи на минимум (о нахождении системы, содержащей наименьшее возможное число маршрутов) является неэффективным, т. е. не указывает сразу оптимальную систему маршрутов; оно существенно опирается на существование оптимальной системы. Связанные с этим вопросы обсуждаются в теме 21 (стр. 164). Отметим только, что примененный здесь метод не связан ни с каким затруднением ввиду того, что общее количество возможных систем маршрутов, где каждый маршрут состоит из целого числа отрезков, заключенных между двумя точками пересечения (т. е. система не содержит лишних «пересадочных пунктов» вроде точки R на рис. 1), является конечным, и поэтому ясно, что среди этих систем найдется одна или несколько, содержащих наименьшее число маршрутов (см. сравнение с задачей о самом младшем ученике в тексте книги).

Заметим еще, что благодаря интерпретации нашей задачи на примере сети городских железных дорог может возникнуть представление, что сеть кривых непременно расположена в одной плоскости. Однако по существу вопроса здесь это совсем не обязательно; все сказанное сохраняет силу также и для пространственной сети кривых. В этом отношении тема 11, имеющая некоторое внешнее сходство с настоящей, в своей основе совершенно отлична от нее; гораздо ближе к этой теме известная задача о характеристике таких сетей линий, которые могут быть начерчены одним росчерком пера, без отрыва пера от бумаги и без обведения одной и той же линии дважды (см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы элементарной математики, ч. 2, М., Гостехиздат, 1952, задача 38).

[9] См. «Начала» Евклида, книги I—VI, М.—Л., Гостехиздат, 1948, книга VI, предложение 27, стр. 208.

[10] Отсюда следует также, что *периметр квадрата меньше периметра любого, отличного от квадрата прямоугольника той же площади* (ср. ниже стр. 169—170).

[11] Обратно из неравенства (2) (для которого известно очень много различных доказательств) следует наша геометрическая теорема.

Аналогично этому теорему о среднем арифметическом и среднем геометрическом трех количеств:

$$\sqrt[3]{xyz} \leqslant \frac{x+y+z}{3}$$

(см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы элементарной математики, ч. 1, М., Физматгиз, 1959, задача 263) можно сформулировать как утверждение о том, что *из всех прямоугольных параллелепипедов с заданной суммой длин ребер наибольший объем имеет куб.*

[12] Все наши выводы распространяются и на треугольники, вписанные в эллипс. Роль правильного треугольника играет в этом случае такой, для которого касательная к эллипсу в вершине треугольника параллельна стороне, противолежащей этой вершине.

[13] Широкий обзор подобных свойств правильных многоугольников приведен в § 3, гл. I книги Л. Фейеш Тот, Расположения на плоскости, на сфере и в пространстве, М., Физматгиз, 1958. См. также литературу, указанную в примечаниях к теме 22.

[14] Высказывания Анаксагора см. у Дильса, Фрагменты досократиков (Diels «Fragmente der Vorsokratiker» 1, № 46, B3, 2-е изд., стр. 314_{16-19}). Высказывания Платона см. в его «Законах» (VII, 819 d_5 — $820c_9$).

[15] Первое доказательство не встречается ни у кого из греческих математиков, хотя оно по своему типу должно было бы быть им известно и, вероятно, в действительности было известно. Второе имеется у Евклида (см. «Начала» Евклида, книги VII—X, М.—Л., Гостехиздат, 1949, книга X, предложение 9; следствие, стр. 111); одно указание у Аристотеля в его «Analytica priora» ($41a_{26}$ и $50a_{37}$) дает основание полагать, что это доказательство было известно задолго до Евклида. Иногда математический отрывок в начале «Теетета» Платона понимают в том смысле, что сначала для доказательства иррациональности $\sqrt{2}$ греки пользовались первым методом и лишь впоследствии (Теетет) нашли второй метод.

[16] Обсуждение вопроса о несоизмеримости диагонали квадрата с его стороной и об иррациональных числах имеется в статье Г. М. Фихтенгольца, Иррациональные числа в средней школе, сб. Математическое просвещение, вып. 2, 1957.

[17] Г. А. Шварц (H. A. Schwarz, 1843—1921) — известный немецкий математик, специалист по математическому анализу и теории функций; имеет также крупные заслуги в областях, связанных с так называемой «изопериметрической проблемой» (см. ниже тему 22) и с ее обобщениями.

[18] Это связано с общим принципом Ферма (играющим в физике чрезвычайно важную роль), согласно которому луч света всегда выбирает такой путь, чтобы достигнуть конечного пункта в кратчайший срок. Также и закон преломления света подчиняется этому принципу (см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы..., ч. 2, задача 148 и примечание к ней).

[19] См. Г. А. Шварц, Собрание сочинений (H. A. Schwarz, Ges. Abh., II, стр. 344—345); также Я. Штейнер, Сочинения (J. Steiner «Werke», II, стр. 728—729); Штейнер высказывает эту теорему одновременно как для плоских, так и для сферических треугольников.

Требование, чтобы треугольник был остроугольным, очевидно, необходимо, так как только в таком треугольнике треугольник, образованный основаниями высот, целиком лежит внутри исходного треугольника. В тупоугольном треугольнике треугольник Δ, образованный основаниями высот, не был бы вписанным в собственном смысле этого слова (в прямоугольном треугольнике он выродился бы в отрезок). В доказательстве Шварца это требование выражается в том, что треугольник, образованный основаниями высот, уже предполагается вписанным, что существенно используется в дальнейших рассуждениях. Если же треугольник является прямоугольным или тупоугольным, то наименьший по периметру вписанный «треугольник» является вырожденным — он представляет собой дважды взятую высоту треугольника, опущенную из наибольшего угла (рис. 127; периметр любого вписанного в ABC треугольника больше дважды взятой высоты CG, но существуют такие вписанные в ABC треугольники, периметр которых сколь угодно близок к величине $2CG$). По этому поводу см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы..., ч. 2, задача 144 в), в решении которой содержится полный разбор всех случаев, которые могут представиться при отыскании треугольника наименьшего возможного периметра, вершины которого принадлежат трем заданным прямым (даже не обязательно образующим треугольник).

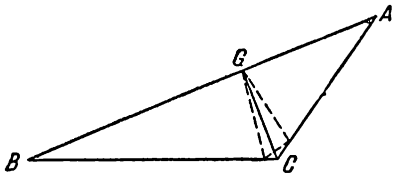


Рис. 127.

Интересно отметить также различие между этой задачей и аналогичной задачей об отыскании четырехугольника наименьшего возможного периметра, вписанного в данный четырехугольник. К решению этой задачи можно подойти теми же методами, которые были применены Г. А. Шварцем для нахождения треугольника наименьшего возможного периметра, однако полученный результат будет совсем другим. В самом деле, рассмотрим четырехугольник $ABCD$ и вписанный в него четырехугольник $PQRS$ (рис. 128); отразим затем наш чертеж последовательно от сторон AB (причем $ABCD$ перейдет в $ABC'D'$), BC'(причем $ABC'D'$ перейдет в $A'BC'D''$), $C'D''$ (причем $A'BC'D''$ перейдет в $A''B'C'D''$) и $D''A''$ (причем $A''B'C'D''$ перейдет в $A''B''C''D''$). Ясно, что если нам надо *найти вписанный в $ABCD$ четырехугольник наименьшего возможного периметра, одна вершина которого совпадает с заданной точкой P*, то для решения этой задачи необходимо соединить точку P стороны AD с отвечающей ей точкой P''' стороны $A''B''$; если прямая PP''' пересечет стороны AB, BC' и $C'D''$ полученных четырехугольников, то она определит искомый вписанный четырехугольник $PQRS$ (см. рис. 128); если же PP'' не пересечет всех этих отрезков, то искомый «четырехугольник» будет вырожденным. Для того же, чтобы найти *четырехугольник абсолютно наименьшего периметра, вписанный в $ABCD$*, нам надо еще выбрать точку P таким образом, чтобы отвечающий ей отрезок PP''' был наименьшим.

Но если $PQRS$ — четырехугольник, доставляющий решение вспомогательной задачи (с заданной точкой P), то стороны его образуют одинаковые углы со сторонами AB, BC и CD четырехугольника (на рис. 128 $\angle PQA = \angle BQR' = \angle BQR$; $\angle QRB = \angle SRC$, ибо $\angle QR'B = \angle S''R'C' = \angle S'R'C'$; $\angle RSC = \angle PSD$, ибо $\angle R'S''C' = \angle P'''S''D'' = \angle P''S''D''$). А так как искомый четырехугольник $PQRS$ является также наименьшим по периметру из всех вписанных четырехугольников с фиксированной вершиной Q, то стороны его в силу тех же соображений образуют одинаковые углы со сторонами BC, CD и DA исходного четырехугольника. Таким образом, мы заключаем, что если только решение задачи существует, то (в обозначениях рис. 128) $\angle APQ = \angle DPS$, $\angle AQP = \angle BQR$, $\angle BRQ = \angle CRS$ и

$\angle CSR = \angle DSP$; эти четыре пары равных углов мы обозначим цифрами *1, 2, 3* и *4*.

Теперь остается только заметить, что $\angle A = 180° - \angle 1 - \angle 2$, $\angle B = 180° - \angle 2 - \angle 3$, $\angle C = 180° - \angle 3 - \angle 4$ и $\angle D = 180° - \angle 4 - \angle 1$, откуда следует, что

$$\angle A + \angle C = \angle B + \angle D = 360° - (\angle 1 + \angle 2 + \angle 3 + \angle 4).$$

Таким образом, мы видим, *что если только задача о нахождении вписанного в $ABCD$ четырехугольника наименьшего возможного периметра имеет решение* (т. е. искомый четырехугольник $PQRS$ не

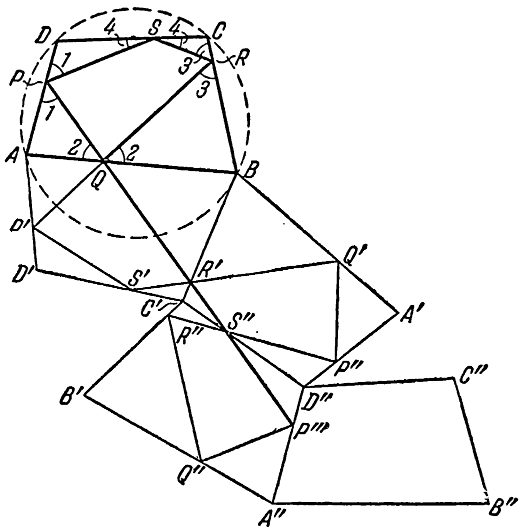


Рис. 128.

является вырожденным), *то* суммы противоположных углов четырехугольника $ABCD$ равны, т. е. *вокруг $ABCD$ можно описать окружность.*

Заметим еще, что, как нетрудно усмотреть непосредственно из рис. 128, четырехугольник $BC'D''A'$ получается из четырехугольника $BCDA$ поворотом вокруг вершины B на угол $2\angle B$; четырехугольник $D''A''B''C''$ получается из четырехугольника $D''A'C'B$ поворотом вокруг точки D'' на угол $2\angle D$. Но в нашем случае сумма $2\angle B + 2\angle D$ равна 360°; отсюда вытекает, что четырехугольник $A''B''C''D''$ расположен параллельно $ABCD$ и, в частности, $A''D'' \parallel AD$. Последнее означает, что *независимо от выбора точки P на стороне AD отрезок PP''' будет иметь одну и ту же длину*; таким образом, исходя из любой точки P стороны AD (такой, что отвечающий ей отрезок PP'''

пересекает отрезки AB, BC' и $C'D''$), мы придем к вписанному четырехугольнику $PQRS$ одного и того же периметра. Таким образом можно утверждать, что *если только существует хоть один невырожденный четырехугольник $PQRS$, вписанный в $ABCD$ и имеющий периметр, не превышающий периметра всех других вписанных в $ABCD$ четырехугольников, то подобных четырехугольников имеется бесконечно много* (рис. 129).

Аналогично этому можно показать, что задача *вписать в данный n-угольник n-угольник наименьшего возможного периметра при n* нечетном будет в достаточно широком классе случаев иметь единственное невырожденное решение; при четном же n она всегда будет невозможна (т. е. будет приводить лишь к вырожденным многоугольникам) или неопределенна (иметь бесконечно много решений).

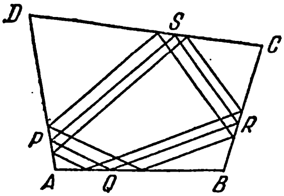


Рис. 129.

[20] Остроугольность треугольника здесь связана с тем, что угол $U'AU''$ меньше двух прямых, и точки пересечения M, N прямой $U'U''$ с AC и AB лежат на самих этих сторонах (а не на их продолжениях). Точно так же основание E высоты, опущенной из A, лежит на стороне BC, а не на ее продолжении, потому что треугольник остроугольный.

Серьезное преимущество доказательства Фейера перед доказательством Шварца состоит в том, что оно остается верным и для неевклидовой геометрии, в которой доказательство Шварца неприменимо, так как в последнем используется теорема о равенстве суммы углов треугольника двум прямым, и понятие параллельности употребляется в смысле евклидовой геометрии. Доказательство Фейера применимо, в частности, также и для сферических треугольников.

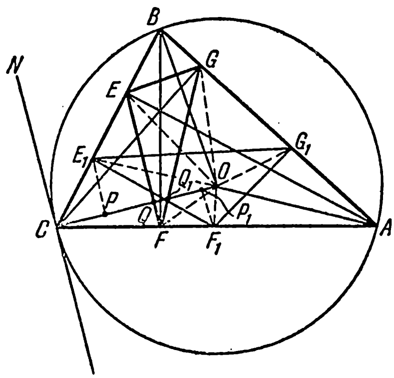


Рис. 130.

Приведем в заключение еще одно изящное доказательство теоремы о треугольнике наименьшего возможного периметра, вписанном в остроугольный треугольник, основанное на совсем другой идее. Докажем прежде всего, что *стороны треугольника, вершинами которого служат основания высот треугольника, перпендикулярны радиусам описанной окружности, проведенным в вершины треугольника.* В самом деле, если E и F — основания высот AE и BF треугольника ABC (рис. 130), то $\triangle AEC \backsim \triangle BFC$ (по двум равным углам) и, следовательно, $\frac{CF}{CE}=\frac{CB}{CA}$; но отсюда вытекает, что $\triangle CEF \backsim \triangle CAB$ и $\angle CEF=\angle CAB$. С другой стороны, если MN — касательная к описанной вокруг треугольника окружности в вершине C, то $\angle NCB=\angle A$ (оба эти угла измеряются половиной дуги BC окружности) и, значит, $FE \| MN \perp OC$ (где O — центр описанной окружности); аналогично доказывается, что $FG \perp OA$, $GE \perp OB$.

Сравним теперь треугольник EFG с другим вписанным в треугольник ABC треугольником $E_1F_1G_1$. Заметим, что так как диагонали EF и OC четырехугольника $OECF$ взаимно-перпендикулярны, то

$$S_{OECF}=\frac{1}{2}OC\cdot EF=\frac{1}{2}R\cdot EF$$

(R — радиус описанной окружности; ясно, что $S_{\triangle OEC}=\frac{1}{2}EQ\cdot R$, $S_{\triangle OFE}=\frac{1}{2}FQ\cdot R$, где Q — точка пересечения EF и OC). С другой стороны,

$$S_{OE_1CF_1}\leqslant\frac{1}{2}OC\cdot E_1F_1=\frac{1}{2}R\cdot E_1F_1$$

(ибо если E_1P и F_1P_1 — высоты треугольников OE_1C и OF_1C, Q_1 — точка пересечения E_1F_1 и OC, то $S_{\triangle OE_1C}=\frac{1}{2}E_1P\cdot R\leqslant\frac{1}{2}E_1Q_1\cdot R$, $S_{\triangle OF_1C}=\frac{1}{2}F_1P_1\cdot R\leqslant\frac{1}{2}F_1Q_1\cdot R$). Учитывая подобные же соотношения, относящиеся к четырехугольникам $OFAG$ и OF_1AG_1, $OGBE$ и OG_1BE_1, получаем:

$$S_{\triangle ABC}=S_{OECF}+S_{OFAG}+S_{OGBE}=$$
$$=\frac{1}{2}R\cdot EF+\frac{1}{2}R\cdot FG+\frac{1}{2}R\cdot GE=\frac{1}{2}R\cdot(EF+FG+GE)$$

и

$$S_{\triangle ABC}=S_{OG_1CF_1}+S_{OF_1AG_1}+S_{OG_1BE_1}\leqslant\frac{1}{2}R\cdot E_1F_1+\frac{1}{2}\cdot F_1G_1+$$
$$+\frac{1}{2}R\cdot G_1E_1=\frac{1}{2}R\cdot(E_1F_1+F_1G_1+G_1E_1).$$

Отсюда и следует, что

$$EF+FG+GA\leqslant E_1F_1+F_1G_1+G_1E_1,$$

т. е. что *периметр треугольника EFG — наименьший возможный*. То что треугольник EFG доставляет единственное решение интересующей нас задачи, следует из того, что у отличного от EFG вписанного треугольника $E_1F_1G_1$ хоть одна сторона будет не перпендикулярна соответствующему радиусу описанной окружности (ибо в ABC нельзя вписать двух треугольников с параллельными сторонами, см. рис. 131).

Остроугольность треугольника используется в этом доказательстве в виде. предположения, что центр O описанной окружности лежит в н у т р и ABC. Последнее доказательство удобно еще и тем, что оно может быть использовано для обобщения исходной задачи (см., например, И. М. Я г л о м, Геометрические преобразования, I, М., Гостехиздат, 1955, задача **102**).

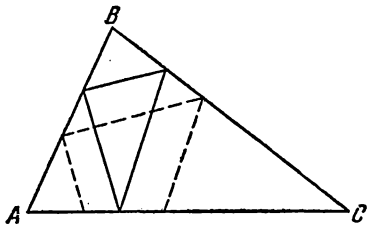


Рис. 131.

[21] Остроугольность треугольника здесь используется в том обстоятельстве, что точка P лежит внутри треугольника ABC. Но для этого остроугольность вовсе не необходима; достаточно, чтобы все углы треугольника ABC были меньше 120°. Таким образом, мы заключаем, что *если все углы треугольника ABC меньше* 120°, *то точка, сумма расстояний от которой до трех вершин является наименьшей — это та, из которой все стороны треугольника видны под одинаковыми углами* (равными 120°); *в противоположном случае искомой точкой явится вершина тупого угла* (см., например, И. М. Я г л о м, Геометрические преобразования, I, решение задачи **103**).

Заметим еще, что для этой задачи также существует решение, близкое к тому доказательству теоремы о вписанном треугольнике, образованном основаниями высот, которое было приведено в примечании [20] (и даже два подобных доказательства, в известном смысле двойственных друг другу). По этому поводу см. И. М. Я г л о м, Геометрические преобразования, I, задачи **104** а) и **104** б) (еще одно решение той же задачи составляет содержание задачи **105** б) той же книги). Эти доказательства также позволяют обобщить поставленную задачу (задачу **106** указанной книги И. М. Яглома).

Интересно обратить внимание еще на одно обстоятельство. В случае, если один угол треугольника ABC превосходит 120°, окружности, построенные на сторонах треугольника и вмещающие углы по 120° (иными словами, окружности, описанные вокруг построенных на сторонах треугольника вне его правильных треугольников), пересекаются во в н е ш н е й по отношению к треугольнику точке P; в этом случае решением задачи о точке с наименьшей суммой расстояний до вершин треугольника будет являться не эта точка P, а вершина C тупого угла. Однако точка P также доставляет решение интересной задачи на минимум — для нее достигает *наименьшего возможного значения сумма расстояний до двух вершин треугольника, уменьшенная на расстояние от этой точки до третьей вершины* (вершины тупого угла). Доказательство этого можно получить, используя буквально любой прием, приводящий к решению основной задачи; мы очень рекомендуем читателю постараться самостоятельно это сделать. При этом *если угол C треугольника не превосходит* 120°, *то выражение $MA + MB - MC$ достигает минимума для точки M,* *совпадающей с вершиной угла C*; таким образом, из четырех задач о нахождении точки M, для которой достигает минимума одно из выражений

$$MA+MB+MC, \quad MA+MB-MC,$$
$$MA+MC-MB \text{ или } MB+MC-MA,$$

всегда решение одной доставляется точкой P, а трех остальных — вершинами треугольника.

[22] Возможно также сопоставлять каждому числу верхнего ряда в четыре раза большее его число, т. е. спаривать 1 с 4, 2 с 8, 3 с 12 и т. д.; при этом все нечетные числа войдут в состав пар, а четные будут использованы лишь частично, из чего может составиться представление о том, что четных чисел имеется больше, чем всех чисел.

Эта зависимость возможности установления взаимно однозначного соответствия от способа сочетания пар иногда формулируется в виде требования о том, чтобы в случае (реально, разумеется, невозможном) наличия бесконечно большого числа гостей пришедшие особенно внимательно следили за тем, чтобы одеть при уходе именно свои калоши: в противоположность случаю конечного числа гостей здесь иначе можно представить себе такую ситуацию, при которой уже часть гостей заберет все калоши и для других ничего не останется, несмотря на то, что каждый гость пришел в одной паре калош и ушел также, взяв единственную пару (здесь можно мысленно также представить себе случай, когда все гости ушли домой в калошах и, несмотря на это, ряд пар калош — даже, если угодно, бесконечно большое число пар — остались хозяину дома).

[23] Исследования специалистов по математической логике (в первую очередь крупнейшего логика К. Гёделя) привели к убеждению, что (во всяком случае, в большинстве принятых аксиоматик теории множеств) проблема континуума неразрешима, т. е. не может быть выведена с помощью цепочки дедуктивных умозаключений из основных предложений теории множеств.

[24] Элементы теории множеств входят в излагающийся на математических отделениях университетов и techинститутов курс теории функций вещественной переменной; с их изложения начинаются все учебники этой дисциплины (см., например, Н. Н. Лузин, Теория функций действительного переменного, М., Учпедгиз, 1948, гл. I; И. П. Натансон, Теория функций вещественной переменной, М.—Л., Гостехиздат, 1950, гл. I; П. С. Александров, Введение в общую теорию множеств и функций, М.— Л., Гостехиздат, 1948, гл. 1). Более специальный характер имеет монография: Ф. Хаусдорф, Теория множеств, М.—Л., ОНТИ, 1937.

[25] Доказательство Данделена можно найти в «Новых мемуарах Королевской Академии наук и литературы в Брюсселе» (Nouveaux mémoires de l'Académie Royale des Sciences et Belles Lettres de Bruxelles), Tome II, 1822, стр. 171—173, в статье Mémoire sur quelques propriétes remarquables de la focale parabolique par M-r G. Dandelin, officier du génie, au service de S. M. le Roi des Pays Bas, lu à la séance du 1-er Avril 1822 («Мемуар о некоторых замечательных свойствах параболических полюсов, представленный на заседании 1-го апреля 1822 г. г-ном Ж. Данделеном»).

[26] См., например, Б. Н. Делоне и Д. А. Райков, Аналитическая геометрия, ч. I, М.—Л., Гостехиздат, 1948, стр. 342—345.

[27] Применение в комбинаторике еще более общего определения числа размещений принадлежит известному норвежскому математику Вигго Бруну (Viggo Brun) (оно называется у Бруна «функцией размещения»). См. Нетто, Учебник комбинаторики (Netto «Lehrbuch der Kombinatorik», 2-е изд., осуществленное В. Бруном и Т. Сколемом, гл. 14, Leipzig, 1926). Тот частный вид функции размещения, которым мы занимались в тексте, Брун рассматривал отдельно в статье Sur les éléments de l'analyse combinatoire (L'Enseignement mathematique 29, 1930, стр. 231—237). Ср. также Дж. Г. Кемени и др., Введение в конечную математику, М., ИЛ, 1962, §§ 4—7 гл. III.

[28] Баше де Мезириак (Bachet de Méziriac) в своем издании Диофанта обратил внимание на то, что в одном месте у Диофанта предполагается, что любое число может быть представлено в виде суммы четырех квадратов. Ферма в одном из своих писем дал набросок доказательства этого предложения; Эйлер и Лагранж доказали его впоследствии способом, намеченным нами в тексте.

[29] Гильбертовское решение проблемы Варинга является чрезвычайно сложным; оно использует мощные методы современной теории чисел, весьма далекие от тех, какими владеет элементарная математика. В 1943 г. ленинградский математик Ю. В. Линник нашел новое доказательство той же самой теоремы, все еще очень сложное, но по методам совершенно элементарное, не требующее никаких знаний по математике, выходящих за пределы, сообщаемые в средней школе. См. по этому поводу А. Я. Хинчин, Три жемчужины теории чисел, М.—Л., Гостехиздат, 1948.

[30] Варинг в «Meditationes algebraicae» (Cambridge, 1782) только высказал предположение, что возможно обойтись девятью кубами, девятнадцатью биквадратами и т. д. для высших степеней.

Подробности относительно работ Якоби можно найти в собрании его сочинений (C. Jacobi, Werke, 6, стр. 322). Работы Вифериха (Wieferich) напечатаны в Math. Ann. **66**, 1909, стр. 95—101 и 106—108. Далее см. Гильберт (Hilbert), Math. Ann. **67**, 1909, стр. 281—300; Харди (Hardy) и Литлвуд (Littlewood), Math. Zeitschr. **23**, 1925, стр. 1—37.

[31] Тождество (1) может быть использовано для полного решения вопроса о том, какие целые числа могут быть разложены на сумму двух полных квадратов. А именно оказывается, что *целое число N в том и только в том случае может быть разложено в сумму двух квадратов, если в разложение этого числа на простые множители все простые делители, дающие при делении на* 4 *остаток* 3 (в том числе и делитель 3, если N делится на 3), *входят в четных степенях*. См. по этому поводу, например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы элементарной математики, ч. 1, задача 247.

[32] Относительно доказательства этой теоремы см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы..., ч. 1, задача 249, или Л. Г. Шнирельман, Простые числа, М.—Л., Гостехиздат, 1940, § 5.

Несколько неожиданным образом более сложной оказывается задача о полной характеристике тех чисел, которые можно разложить в сумму трех квадратов. Оказывается, что такое представление возможно для *всех чисел, кроме чисел, имеющих вид* $4^n \cdot (8k-1)$, *где n и k целые*; однако в то время как доказательство утверждения, что *числа вида* $4^n \cdot (8k-1)$ *непредставимы в виде суммы трех кубов*, совершенно элементарно (см. задачу 250 из указанной выше книги Д. О. Шклярского и др.), установление того, что *все остальные целые числа можно представить в таком виде*, требует очень серьезных усилий.

[33] Сам Лиувилль сообщил приведенное у нас доказательство только устно на лекции в Collège de France; опубликовано оно было знаменитым Лебегом (A. Lebesgue) в «Exercises d'analyse numérique». См. по этому вопросу P. Bachmann «Niedere Zahlentheorie», II, 1910, стр. 329.

[34] Проблема Варинга связана с определением некоторой функции $N(k)$ натурального аргумента n: *$N(k)$ есть наименьшее целое число, такое, что всякое натуральное число n можно представить в виде суммы не более чем N k-х степеней целых чисел*. Как отмечалось в тексте, $N(2)=4$, $N(3)=9$; предполагают также, что $N(4)=19$, хотя до сих пор никому не удалось этого строго доказать. Полное определение функции $N(k)$ представляет собой интересную задачу, на пути решения которой математикам встретились весьма существенные трудности, хотя к настоящему времени в этом отношении достигнуты уже довольно существенные успехи.

Заметим прежде всего, что *число $N(k)$ не может быть меньше величины* $2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$, где прямые скобки, как обычно, обозначают целую часть числа. В самом деле, рассмотрим число $n = 2^k \cdot \left[\left(\frac{3}{2}\right)^k\right] - 1$. Ясно, что целое число 3^k больше, чем n $\left(\text{ибо } \left[\frac{3^k}{2^k}\right] < \frac{3^k}{2^k} \text{ и уже } 2^k \cdot \left[\frac{3^k}{2^k}\right] < 3^k\right)$; поэтому в представлении числа n в виде суммы k-х степеней целых чисел могут фигурировать лишь степени чисел 2 и 1. При этом так как $n < \left[\left(\frac{3}{2}\right)^k\right] \cdot 2^k$, то число степеней 2 в этом представлении не может превосходить $\left[\left(\frac{3}{2}\right)^k\right] - 1$; поэтому подобное представление должно иметь вид

$$\underbrace{2^k + 2^k + \ldots + 2^k}_{\left[\left(\frac{3}{2}\right)^k\right] - 1 \text{ раз}} + \underbrace{1^k + 1^k + \ldots + 1^k}_{n - \left(\left[\left(\frac{3}{2}\right)^k\right] - 1\right) \cdot 2^k = 2^k - 1 \text{ раз}}$$

(здесь фигурируют $\left(\left[\left(\frac{3}{2}\right)^k\right] - 1\right) + (2^k - 1) = 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$ степеней целых чисел), или содержать еще больше слагаемых (последнее будет иметь место, если число степеней 2 в представлении числа n в виде суммы k-х степеней будет меньше $\left[\left(\frac{3}{2}\right)^k\right] - 1$, а число степеней 1 — больше 2^k-1). Отсюда и следует, что

$$N(k) \geqslant 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2.$$

Нетрудно проверить, что

$$2^2 + \left[\left(\frac{3}{2}\right)^2\right] - 2 = 4 + \left[\frac{9}{4}\right] - 2 = 4 + 2 - 2 = 4;$$
$$2^3 + \left[\left(\frac{3}{2}\right)^3\right] - 2 = 8 + \left[\frac{27}{8}\right] - 2 = 8 + 3 - 2 = 9;$$
$$2^4 + \left[\left(\frac{3}{2}\right)^4\right] - 2 = 16 + \left[\frac{81}{16}\right] - 2 = 16 + 5 - 2 = 19$$

— таким образом мы знаем, что для $k=2$ и $k=3$ величина $N(k)$ равна $2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$, а для $k=4$ — вернее всего, равна этой величине. Существует предположение, что $N(k) = 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$ *для всех натуральных* $k \geqslant 2$. Однако пока установлено лишь, что $N(k) = 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$ для всех $k \geqslant 6$ и таких, что $3^k - (2^k - 1) \cdot \left[\left(\frac{3}{2}\right)^k\right] \leqslant 2^k$; при этом до сих пор не обнаружено ни одного числа k, такого, что это неравенство не выполняется, так что можно предполагать, что таких чисел вовсе не существует. В частности, значения $k=6$, 7 и 8 удовлетворяют выписанному выше неравенству, откуда получается точное значение величин $N(k)$ для этих k: $N(6)=73$, $N(7)=143$, $N(8)=276$. Точные значения $N(4)$ и $N(5)$ до сих пор не найдены: для $N(4)$ доказано лишь, что это число не превышает 21 (что достаточно близко к значению $N(4)=19$, даваемому формулой $N(k) = 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$); для числа $N(5)$ установлено, что оно не превосходит 40 (в то время как формула $N(k) = 2^k + \left[\left(\frac{3}{2}\right)^k\right] - 2$ дает $N(5)=37$).

Еще менее удовлетворительными являются результаты, касающиеся оценки величины $M(k)$ — *наименьшего целого числа, такого, что всякое натуральное число n, начиная с некоторого* (быть может — очень большого) *значения n, можно представить в виде суммы $N(k)$ k-х степеней целых чисел.* Для $k=2$ величины $N(2)$ и $M(2)$ совпадают: $N(2)=M(2)=4$ (последнее следует из того, что существуют сколь угодно большие числа вида $4^n \cdot (8k-1)$, непредставимые в виде суммы трех квадратов целых чисел). Однако уже для $k=3$ величины $N(3)$ и $M(3)$ оказываются различными: $N(3)=9$, однако девять кубов необходимо для представления всего двух целых чисел 23 и 239, так что уже для всех $n > 239$ удается обойтись восемью кубами. Неравенство $M(k) \leqslant 8$ доказал, как отмечалось в тексте, Э. Ландау. В 1942 г. Ю. В. Линник доказал, что $M(3) \leqslant 7$; с другой стороны, нетрудно показать, что существуют сколь угодно большие целые числа, непредставимые в виде суммы 5 кубов целых чисел. Таким образом, мы имеем неравенства $M(3)>5$ и $M(3)\leqslant 7$, досадный разрыв между которыми пока не заполнен; существует предположение, что $M(3)=6$. Аналогично этому удалось доказать, что $M(4)\leqslant 17$ (а не только, что $M(4)\leqslant 19$, как показали в свое время Харди и Литлвуд); с другой стороны, нетрудно проверить, что $M(4)>14$ (ни одно число вида $16k-1$ не может быть представлено в виде суммы 14 четвертых степеней целых чисел). Здесь мы также пока имеем двойное неравенство $14<M(4)\leqslant 17$, не позволяющее точно оценить величину $M(4)$; существует предположение, что $M(4)=16$. Для произвольных значений k легко показать лишь, что $M(k)\geqslant k+1$. Значительно больших трудов стоят оценки сверху величины $M(k)$, которые удалось получить лишь с использованием весьма тонких методов современной теории чисел (ср. И. В. Арнольд, Теория чисел, стр. 135—136; А. А. Бухштаб, Теория чисел, стр. 298 и 306).

Отметим еще, что аналогично проблеме Варинга можно поставить и задачу *о представлении рациональных чисел суммами степеней рациональных чисел*. В общем виде эта задача является весьма сложной, однако в некоторых частных случаях она может быть решена сравнительно просто. Так, например, можно показать, что *всякое положительное рациональное число может быть представлено в виде суммы кубов трех положительных рациональных чисел*; с другой стороны, не каждое рациональное число можно представить в виде суммы двух кубов рациональных чисел, так что этот результат не может быть улучшен (см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы ..., ч. 1, задача 252 и примечание к этой задаче).

В заключение укажем, что проблема Варинга представляет собой не единственное обобщение классического результата о представлении каждого целого числа в виде суммы четырех квадратов целых чисел. Квадраты («квадратные числа») связаны с геометрической фигурой — квадратом; эта связь выражается в том, что построенные (см. рис. 132) квадраты из точек будут содержать последовательно

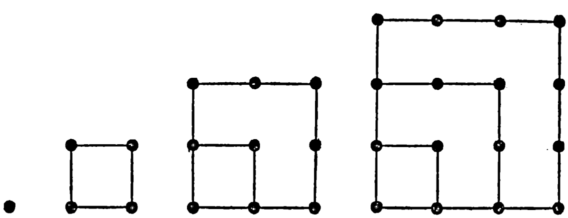


Рис. 132.

1, 4, 9, 16, ... , n^2, ... точек. Аналогично этому, если мы будем строить из точек правильные треугольники (рис. 133), то мы получим, последовательно, 1, 3, 6, 10, ... , $\frac{n(n+1)}{2}$, ... точек. Если мы будем таким же образом образовывать правильные пятиугольники (рис. 134), то получим 1, 5, 12, 22, ..., $\frac{n(3n-1)}{2}$, ... точек. Этот процесс определяет «треугольные числа», «пятиугольные числа» и другие типы так называемых «фигурных чисел». Нетрудно видеть, что определяемое на подобном пути «k-угольное число» представляет собой сумму членов арифметической прогрессии с первым членом 1 и разностью $k-2$; таким образом, n-е k-угольное число $A_n^{(k)}$ равно

$$1+(k-2)+2(k-2)+\ldots+(n-1)(k-2)=\frac{n[(k-2)n-(k-4)]}{2}.$$

П. Ферма поставил задачу о представлении всех натуральных чисел в виде суммы треугольных чисел или пятиугольных чисел, или

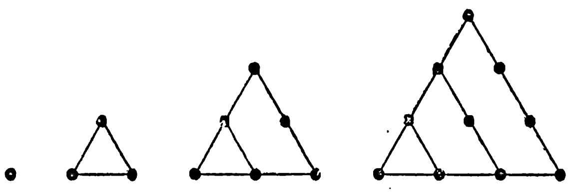


Рис. 133.

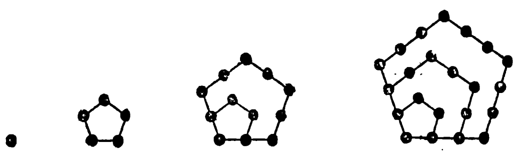


Рис. 134.

вообще k-угольных чисел, где k фиксировано; при $k=4$ мы приходим к задаче о представлении числа суммами квадратов. Эта общая задача оказалась много проще проблемы Варинга; здесь имеет место следующая чрезвычайно изящная теорема, известная еще П. Ферма: *каждое натуральное число может быть представлено в виде суммы k или меньшего числа k-угольных чисел* (впервые доказательство этой теоремы было опубликовано О. Коши в 1815 г.).

[85] Нетрудно понять, что всякая кривая типа, изображенного на рис. 36 (кривая, содержащая «тройные» точки или точки еще более высокой кратности), может быть небольшой деформацией отдельных ее участков превращена в кривую, обладающую лишь двойными точками. Это замечание открывает путь к перенесению доказываемой теоремы и на кривые типа изображенной на рис. 36; мы, однако, здесь на этом не останавливаемся.

[86] См. сочинения Гаусса (Gauss' Werke 8, стр. 272, 282—286). Доказательство этой теоремы имеется у Ю. Секефальви-Надь

(Julius v. Sz. Nagy, Mathem. Zeitschr. 26 стр. 579—592, особенно 580 — 581); однако она сформулирована и доказана у него несколько иначе, чем в тексте книги.

[37] То, что замкнутая кривая без двойных точек разделяет плоскость на две части — внутреннюю и внешнюю, вовсе не очевидно. Теорема эта к тому же справедлива не для всех поверхностей: она неверна, например, на поверхности баранки — на торе, образованном вращением окружности вокруг не пересекающей ее прямой, расположенной в той же плоскости (рис. 135). В теореме, таким образом, должно быть оговорено, что она относится к плоскости, и в такой форме она нуждается в доказательстве (см., например, Р. Курант и Г. Роббинс, Что такое математика, М.—Л., Гостехиздат, 1947, стр. 326—328 и 353—355, где эта теорема обсуждается и доказывается для многоугольников). Доказательство было впервые дано известным французским математиком К. Жорданом (C. Jordan), по имени которого эту теорему называют теоремою Жордана, а кривую, обладающую указанным свойством, жордановой кривой, а после Жордана — и многими другими учеными (см., в частности, Э. И. Вольперт, Элементарное доказательство теоремы Жордана, Успехи математических наук 5, № 5, 1950, стр. 168—172, и А. Ф. Филиппов, Элементарное доказательство теоремы Жордана, там же, стр. 173—176).

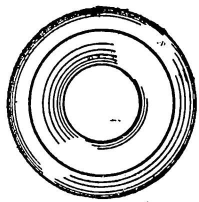


Рис. 135.

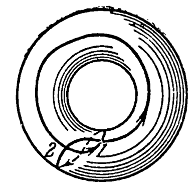


Рис. 136.

Так как на торе теорема Жордана не имеет места, то вся наша аргументация для него отпадает. Действительно, на торе можно указать кривую с двумя двойными точками, расположенными в последовательности *1212*, как это показано на рис. 136.

[38] Об альтернирующих узлах см. прежде всего Tait, Transactions of the Edinburgh Philos. Society 28, стр. 145 (1879). Изображенный на рис. 42 узел можно, впрочем, посредством деформации без разрыва нити преобразовать в трилистниковую петлю (Kleeblattschlinge), т. е. в альтернирующий узел. Но не всякий узел можно деформировать в альтернирующий. Банквиц (C. Bankwitz) дал пример узла, который нельзя деформировать в альтернирующий [Math. Ann. 103, стр. 161 (1930)].

Вопросы, близкие к содержанию тем 2 и 11, исследуются у Петерсена [J. Petersen «Die Theorie der regulären Graphs», Acta Mathematica 15, 1891, стр. 193—220].

[39] Вот еще один, даже более простой пример того же рода. Рассмотрим множество всевозможных четных чисел; при этом мы придем к разложениям

$$180 = 6 \cdot 30 = 18 \cdot 10,$$

где ни одно из чисел 6, 30, 18, 10 нельзя уже дальше разложить в произведение четных множителей (т. е. эти числа в рассматриваемой области являются «простыми»).

[40] См. «Начала» Евклида, книги VII—X, М.—Л., Гостехиздат, 1949, предложение 14 книги IX, стр. 83.

Вот еще одно несложное доказательство той же теоремы. Пусть существуют числа, разложимые на простые множители двумя существенно различными способами; N — наименьшее из них:

$$N = p_1 p_2 \ldots p_k = q_1 q_2 \ldots q_l.$$

Простые множители в обоих разложениях расположим в возрастающем порядке: $p_1 \leqslant p_2 \leqslant \ldots \leqslant p_k$, $q_1 \leqslant q_2 \leqslant \ldots \leqslant q_l$, кроме того, положим, что через p_1 обозначен меньший из двух первых множителей в обоих разложениях: $p_1 < q_1$. (Ясно, что $p_1 \neq q_1$, ибо иначе число $\frac{N}{p_1} = p_2 p_3 \cdot p_k = q_2 q_3 \ldots q_l$ доставляло бы пример меньшего N числа, неоднозначно разлагающегося на множители.) Заметим теперь, что

$$p_1 q_2 q_3 \ldots q_l < q_1 q_2 q_3 \ldots q_l = N,$$

и составим разность

$$M = N - p_1 q_2 q_3 \ldots q_l,$$

которую также можно разложить на множители двумя способами:

$$M = p_1 p_2 p_3 \ldots p_k - p_1 q_2 q_3 \ldots q_l = p_1 (p_2 p_3 \ldots p_k - q_2 q_3 \ldots q_l)$$

и

$$M = q_1 q_2 q_3 \ldots q_l - p_1 q_2 q_3 \ldots q_l = (q_1 - p_1) q_2 q_3 \ldots q_l.$$

Эти два разложения можно продолжить до разложений числа M на простые множители, разложив на простые множители разности $p_2 p_3 \ldots p_k - q_2 q_3 \ldots q_l$ и $q_1 - p_1$ (отметим, что обе эти разности разлагаются на простые множители уже однозначно, поскольку они меньше N — наименьшего целого числа, разлагающегося на простые множители неоднозначно).

Таким образом, мы приходим к двум разложениям числа M на простые множители. Мы утверждаем, что эти разложения являются существенно разными. В самом деле, в первое из них входит простой множитель p_1; во второе же p_1 входить не может, ибо все простые числа $q_2, q_3, \ldots, q_l$ больше p_1, а разность $q_1 - p_1$ не делится на p_1, ибо иначе q_1 должно было бы делиться на p_1. Полученное противоречие со сделанным предположением об отсутствии меньших N целых чисел, разлагающихся на простые множители неоднозначно, и доказывает ошибочность нашего исходного предположения о существовании чисел, имеющих два разных разложения на простые множители.

[41] Свойства геометрических образов, не изменяющиеся при непрерывных деформациях, составляют предмет особой математической дисциплины — топологии; в этой дисциплине теорема Эйлера играет выдающуюся роль. См. по этому поводу В. Г. Болтянский и В. А. Ефремович, Очерк основных идей топологии, сборник «Математическое просвещение», вып. 2, 1957, стр. 3—34, особенно стр. 16—34.

[42] Проблеме четырех красок посвящен 1-й раздел книги: Е. Б. Дынкин и В. А. Успенский, Математические беседы, М.—Л., Гостехиздат, 1952, а также большая часть § 2 книги: Л. И. Головин

и И. М. Яглом, Индукция в геометрии, М., Физматгиз, 1961. В обеих этих книгах имеется значительный материал, относящийся, в частности, к характеристике карт, которые можно окрасить двумя, тремя и четырьмя красками.

[43] См., впрочем, Г. Штейнгауз, Математический калейдоскоп, М.—Л., Гостехиздат, 1949, тема 117.

[44] В последнее время немецкому математику Г. Рингелю удалось решить проблему о закраске географических карт для всех отличных от сферы замкнутых поверхностей: так, например, оказалось, что *на поверхности «кренделя»* (рис. 137) *каждую карту можно*

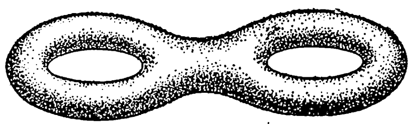


Рис. 137.

закрасить 8 *красками так, чтобы никакие две соседние страны* (т. е. страны, имеющие общую границу) *не были закрашены одним цветом, и можно указать карту из* 8 *стран, каждые две из которых являются соседними* (ясно, что такую карту нельзя закрасить менее чем 8 красками с соблюдением условий задачи). См. по этому поводу G. Ringel, Färbungsprobleme auf Flächen und Graphen, Berlin, 1959 (краткий обзор результатов Рингеля приведен также в заметке И. М. Яглома в сборнике «Математическое просвещение», вып. 3, 1958, стр. 265—266). Таким образом, можно сказать, что в отношении рассматриваемой здесь задачи наша земля устроена наиболее невыгодным образом в том смысле, что для нее решение вопроса о раскраске географических карт составляет наиболее сложную проблему.

[45] Наши исследования о правильных многогранниках являются чисто топологическими (ср. примечание [41] на стр. 252); при этом в основе представления о возможности деформации правильного многогранника в шар лежит специальная топологическая предпосылка. Мы пришли бы к совершенно иным результатам, если бы вместо описанного в тексте стали изучать, например, многогранник, топологически тождественный кольцевой поверхности (тору) (см. тему 13, стр. 105—106), т. е. «многогранник с дыркой». При этом топологическое значение наших рассуждений выступило бы еще отчетливей. А именно оказывается, что на торе существует *бесконечно много* «правильных» в топологическом смысле карт, из которых ни одна, однако, не может послужить к образованию правильного многогранника в смысле метрической геометрии. Действительно, теорема Эйлера для тора имеет вид (см. тему 13, стр. 106)

$$e-k+f=0. \qquad (3^*)$$

С помощью, конечно, и здесь справедливых уравнений (4) и (5) мы получаем вместо (6):

$$f(2n+2m-nm)=0. \qquad (6^*)$$

Так как $f \neq 0$, то

$$2n + 2m - nm = 0, \qquad (7^*)$$

или

$$(n-2)(m-2) = 4. \qquad (9^*)$$

Таким образом, вместо неравенства (9) мы приходим здесь к равенству. Разложение числа 4 на два множителя дает для $(n-2)\times\times(m-2)$ следующие возможности: $1\cdot4$, $2\cdot2$, $4\cdot1$, откуда для n и m получаем такую таблицу возможных значений:

n	3	4	6
m	6	4	3

Эти значения удовлетворяют уравнению (9*), а следовательно и (7*). Уравнение (6*) справедливо при любом значении f, ибо $f\cdot0=0$ всегда правильно. Поэтому в разбираемом нами случае, зная n и m, нельзя вычислить f, а значит и e, и k. И действительно, для любой из трех возможных пар значений n и m существует бесконечно много значений f, e и k, а значит и «правильных многогранников».

Возьмем, например, такую пару значений: $n=4$, $m=4$, и образуем какой-либо прямоугольник из ab квадратов (где a и b — произвольные целые числа, бо́льшие единицы). Представим себе, что этот прямоугольник (рис. 138, *а*) свернут в цилиндр (рис. 138, *б*), а последний согнут в тор (рис. 138, *в*). Тогда мы убедимся, что на торе

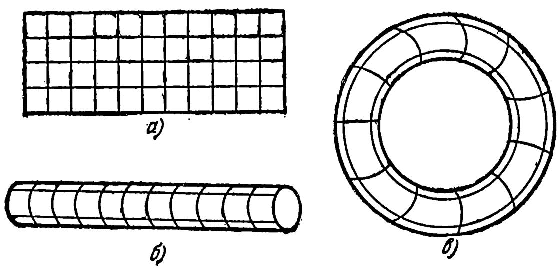


Рис. 138.

действительно наложена правильная карта: в каждой вершине соприкасается по $m=4$ страны и каждая страна имеет $n=4$ границы. Мы имеем здесь $f=a\cdot b$ и, в чем легко убедиться, также $e=ab$, $k=2ab$, причем эти соотношения справедливы для всех значений a и b (при условии $a>1$, $b>1$) и потому дают нам бесконечно много возможных значений для f, e, k. Значит, *существует бесконечно много правильных в топологическом смысле многогранников типа тора.* (Несколько труднее проверить наложимость на поверхность тора правильных карт с парой значений $n=6$, $m=3$. При этом для f, e, k получается, между прочим, система значений: $f=7$, $e=14$, $k=21$, упомянутая в теме 13, п. 5 в связи с задачей о красках в применении к поверхности тора.)

Ни один из этих многогранников нельзя сделать *правильным в смысле метрической геометрии*. Действительно, при $n=3$, $m=4$ четырехугольные боковые грани такого многогранника должны быть квадратами. Но четыре квадрата вокруг одной точки располагаются в одной плоскости и не образуют никакого телесного угла. Присоединяя к ним по тем же правилам сколь угодно много новых квадратов, мы никогда не получим поверхности, ограничивающей тело, а будем лишь продолжать одну и ту же плоскость. Точно так же при $n=6$, $m=3$ три правильных шестиугольника вокруг одной точки располагаются в одной плоскости. При $n=3$, $m=6$ шесть равносторонних треугольников располагаются в одной плоскости и не образуют телесного угла.

Резюмируя, заметим поэтому, что:

во-первых, среди многогранников, топологически эквивалентных шару, имеется лишь пять топологически правильных типов, причем это является особенностью именно такого рода многогранников, тогда как топологически правильных многогранников типа тора существует *бесконечно много*;

во-вторых, топологическая правильность обычно вовсе не влечет за собой правильность метрическую. Для многогранников вида тора, например, такой связи не существует. Если все же пять топологически правильных многогранников типа шара можно построить и метрически правильными, то это возможно лишь на основании дальнейших свойств этого вида поверхности, изучаемых методами метрической геометрии.

Подробный разбор затронутого здесь вопроса о правильных многогранниках разных топологических типов дан в работе В. А. Ефремович, Правильные многогранники, диссертация на соискание степени кандидата физ.-мат. наук, М., МГУ, 1947 (см. также В. А. Ефремович, Правильные многогранники, Доклады Академии наук СССР 57, № 3, 1947, стр. 223—226).

[46] См. «Начала» Евклида, книги XI — XV, М.—Л., Гостехиздат, 1950, книга XIII, а также Ж. Адамар, Элементарная геометрия, т. 2, М., Учпедгиз, 1958, стр. 233—234, или Д. И. Перепелкин, Курс элементарной геометрии, ч. 2, М.—Л., Гостехиздат, 1949, стр. 282 — 288.

Обобщением задачи об отыскании всех правильных многогранников служит задача о так называемых полуправильных многогранниках. По этому поводу см. указанную книгу Д. И. Перепелкина, стр. 289 — 298. Полное перечисление всех типов (архимедовых) полуправильных многогранников имеется в книге Л. А. Люстерник, Выпуклые фигуры и многогранники, М., Гостехиздат, 1956.

[47] Вот еще один вывод формул (5) — (6) для пифагоровых чисел. Заметим прежде всего, что вопрос о целых решениях уравнения $a^2+b^2=c^2$ (a, b, c — неизвестные!) равносилен, очевидно, вопросу о разложении числа 1 на сумму квадратов двух рациональных чисел: $1=x^2+y^2$ (связь между этими двумя задачами дается равенствами: $x=\frac{a}{c}$, $y=\frac{b}{c}$), или вопросу о перечислении всех рациональных точек (точек с рациональными координатами) единичной окружности $x^2+y^2=1$ (рис. 139). Соединим какую-либо рациональную точку P (x, y) окружности с точкой Q $(0, -1)$ пересечения окружности с осью y. Уравнение прямой QP, отсекающей на оси y отрезок

$OQ = -1$, можно записать в виде $y = kx - 1$, где угловой коэффициент k рационален (ибо $k = \frac{y+1}{x}$, где x и y — рациональные координаты точки P). С другой стороны, каждая прямая $y = kx - 1$,

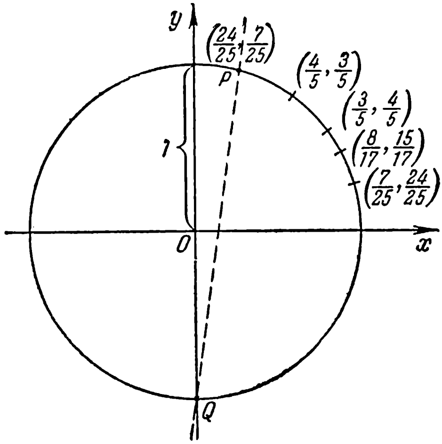


Рис. 139.

где k рационально, пересекает окружность в рациональных точках, координаты которых находятся из системы уравнений

$$x^2 + y^2 = 1, \quad y = kx - 1,$$

имеющей, как нетрудно проверить, решения

$$x_1 = 0, \quad y_1 = -1 \text{ (точка } Q) \quad \text{и} \quad x_2 = \frac{2k}{k^2+1}, \quad y_2 = \frac{k^2-1}{k^2+1} \text{ (точка } P).$$

Таким образом, мы убеждаемся, что *все рациональные точки окружности* $x^2 + y^2 = 1$ *находятся по формулам*

$$x = \frac{2k}{k^2+1}, \quad y = \frac{k^2-1}{k^2+1},$$

где $k = \frac{u}{v}$ рационально, или, что то же самое, — по формулам

$$x = \frac{2uv}{u+v^2}, \quad y = \frac{u^2-v^2}{u^2+v^2}.$$

Но последние формулы непосредственно приводят к формулам (5) — (6).

Формулы (5) — (6) могут быть положены в основу целого ряда более сложных задач, состоящих в отыскании прямоугольных треугольников со сторонами, выраженными целыми числами (пифагоровых треугольников), удовлетворяющих тем или иным дополнительным условиям (например, условию о том, чтобы две стороны треугольника выражались последовательными целыми числами или чтобы одна сторона являлась полным квадратом). Самой знаменитой из задач такого рода является (очень трудная!) задача П. Ферма об отыскании таких пифагоровых чисел, что *длина гипотенузы c и сумма длин катетов $a+b$ являются полными квадратами*, эта задача имеет бесконечно много решений, самым маленьким из которых является следующее:

$$a = 4\,565\,486\,027\,761, \quad b = 1\,061\,652\,293\,520,$$
$$c = 4\,687\,298\,610\,289$$

(здесь $c = (2\,165\,017)^2$ и $a+b = (2\,372\,159)^2$).

Широкий обзор вопроса о пифагоровых числах содержит книга: В. Серпинский, Пифагоровы треугольники, М., Учпедгиз, 1959, в которой рассматривается также целый ряд задач, родственных рассматриваемой здесь задаче; см. также В. Литцман, Теорема Пифагора, М., Физматгиз, 1960, § 7.

[48] Относительно сформулированной здесь знаменитой проблемы Ферма и ее истории см. книгу: А. Я. Хинчин, Великая теорема Ферма, М.—Л., ОНТИ, 1932, а также указанную в предыдущем примечании книгу В. Литцмана, § 8.

[49] В настоящее время американские математики, основываясь на методах Куммера и используя современные электронные вычислительные машины, доказали утверждение Ферма для всех $n > 2$ до $n = 4002$ включительно.

[50] Отметим, что уже доказательство предположения Ферма для $n = 3$ является заметно более сложным, чем проведенное в тексте, поскольку в этом случае не удается ограничиться методом безграничного спуска (доказательство предположения Ферма для $n = 3$ имеется в указанной в примечании [48] книге А. Я. Хинчина; см. также Л. Г. Шнирельман, Простые числа, М.—Л., Гостехиздат, 1940, § 4).

[51] В своей статье Über die kleinste Kugel, die eine räumliche Figur einschliesst, Crelle's Journal, **123**, 1901, стр. 241—257, Юнг (H. W. E. Jung), решая аналогичную задачу в пространстве n измерений, нашел, что наименьший допустимый радиус r n-мерного замыкающего шара для n-мерной точечной совокупности диаметра d может быть выражен такой формулой:

$$r = d\sqrt{\frac{n}{2(n+1)}}.$$

При $n = 2$ это выражение дает результат, полученный в тексте.

[52] По поводу затронутого здесь круга вопросов см. книгу: И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, М.—Л., Гостехиздат, 1951; в частности, задачи **16**, **17**, **19** б), **31**, **35** а), **67** и **68** этой книги. В решениях этих задач излагается несколько различных доказательств теоремы Юнга, из которых, в частности, следует, что наше предположение о том, что рассматриваемая совокупность точек — конечная, не является существенным — *для любого конечного и бесконечного множества точек плоскости* (эти точки могут образовывать некоторую линию или фигуру), *такого, что расстояние между каждыми двумя точками не превосходит d, можно указать*

окружность радиуса $\frac{d\sqrt{3}}{3}$, *содержащую все эти точки.* Можно отметить еще, что окружность радиуса $\frac{d\sqrt{3}}{3}$ можно даже заменить вписанным в нее правильным шестиугольником (правильным шестиугольником со стороной $\frac{d\sqrt{3}}{3}$) — *для каждой такой совокупности точек можно указать содержащий это множество правильный шестиугольник со стороной* $\frac{d\sqrt{3}}{3}$ (см. задачу 31 книги И. М. Яглома и В. Г. Болтянского). В этой связи можно также поставить *задачу о нахождении плоской фигуры самой малой площади, внутрь которой можно поместить любое множество точек, такое, что расстояние между каждыми двумя точками не превосходит d*; эта интересная задача до сих пор не решена.

[53] Существует замечательная теорема известного немецкого математика А. Гурвица, согласно которой *для каждого иррационального числа* w *существует бесчисленное множество таких дробей* $\frac{m}{n}$ (их также можно назвать «хорошими»), *что*

$$-\frac{1}{cn^2} < w - \frac{m}{n} < \frac{1}{cn^2}, \text{ где } c = \sqrt{5} = 2{,}236\ldots;$$

однако для любого $c > \sqrt{5}$ *можно указать такие* («плохие») *иррациональные числа* w, *что подобное обстоятельство уже не будет иметь места.* Доказательство этой теоремы имеется, например, в статье: А. Я. Хинчин, Элементы теории чисел, Энциклопедия элементарной математики, т. 1, М.—Л., Гостехиздат, 1951, стр. 332 — 334.

Заметим еще, что вопрос о характере приближения иррациональных чисел рациональными дробями весьма тесно связан с алгебраической природой этих чисел; так, например, корни уравнений 3-й степени, вообще говоря, лучше приближаются рациональными дробями, чем корни квадратных уравнений, а так называемые «трансцендентные» числа, не являющиеся корнями никаких алгебраических уравнений с целыми коэффициентами (к числу таких чисел принадлежит, например, число π), — лучше чем корни каких бы то ни было уравнений; на этом пути было впервые доказано само существование трансцендентных чисел. См. по этому поводу гл. IV указанной выше статьи А. Я. Хинчина или Р. Курант и Г. Роббинс, Что такое математика, М.—Л., Гостехиздат, 1947, стр. 156 — 162.

[54] Инверсор Поселье был независимо открыт также петербургским (ленинградским) студентом Липкиным, занявшимся этой задачей по предложению П. Л. Чебышева. Лишь после того, как Чебышев рассказал об открытии Липкина своим французским друзьям, те обратили внимание на идентичную (и более раннюю) работу Поселье, первоначально прошедшую совершенно незамеченной.

[55] Относительно инверсии см., например, Ж. Адамар, Элементарная геометрия, ч. I, М., Учпедгиз, 1958, гл. V дополнений к 3-й книге; Д. И. Перепелкин, Курс элементарной геометрии, М.—Л., Гостехиздат, 1948, §§ 87 — 91; И. М. Яглом, Геометрические преобразования, II, М., Гостехиздат, 1956, гл. II части 3-й.

[56] См. G. Koenigs «Leçons de Cinématique», Paris, 1897, стр. 273 — 285. Здесь приведены и другие, но более сложные методы получения прямолинейного движения.

[57] См. «Начала» Евклида, книги VII — X, М.—Л., Гостехиздат, 1949, книга IX, предложение 36 (стр. 98): *Если от единицы откладывается сколько угодно последовательно* [пропорциональных] *чисел в двойном отношении до тех пор, пока вся* [их] *совокупность сложенная не сделается первым* [простым числом; другими словами, Евклид предлагает образовывать сумму $1+2+4+8+\ldots+2^{n-1}=$ $=2^n-1$ с тем, чтобы эта сумма явилась простым числом] *и вся совокупность, умноженная на последнее* [число], *произведет что-то, то возникающее* [число] *будет совершенным.*

[58] В настоящее время известно 18 совершенных чисел; самым большим из них является число $2^{3216}\cdot(2^{3217}-1)$, содержащее около 2000 десятичных знаков. Это число (как и ряд предшествующих ему четных совершенных чисел) было обнаружено с помощью электронной вычислительной машины. См. по этому поводу заметку И. Я. Депмана в разделе «Новости математической науки» сборника «Математическое просвещение», вып. 6, 1961, стр. 324—327.

[59] Ср. И. М. Виноградов, Основы теории чисел, М., Гостехиздат, 1953, § 2 гл. II.

[60] Приведенное здесь доказательство неограниченности ряда простых чисел основано на следующей формуле Эйлера, по существу доказанной в тексте книги:

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{n}+\ldots=$$
$$=\frac{2}{2-1}\cdot\frac{3}{3-1}\cdot\frac{5}{5-1}\cdot\frac{7}{7-1}\cdot\frac{11}{11-1}\cdot\ldots,$$

где произведение в правой части равенства распространено на все простые числа. Так как в левой части этого равенства стоит бесконечно большая величина, ибо ряд $1+\frac{1}{2}+\frac{1}{3}+\ldots$ $\ldots+\frac{1}{n}+\ldots$ расходится (при желании сделать доказательство совершенно строгим без использования учения о бесконечных рядах и произведениях приходится рассуждать так, как это сделано в тексте книги), то число простых чисел должно быть бесконечным (иначе справа стояло бы конечное рациональное число). Аналогично этому можно утверждать, что при любом k

$$1+\frac{1}{2^k}+\frac{1}{3^k}+\frac{1}{4^k}+\ldots+\frac{1}{n^k}+\ldots=$$
$$=\frac{2^k}{2^k-1}\cdot\frac{3^k}{3^k-1}\cdot\frac{5^k}{5^k-1}\cdot\frac{7^k}{7^k-1}\cdot\frac{11^k}{11^k-1}\cdot\ldots;$$

именно эта формула, также установленная Эйлером, играет весьма важную роль при изучении законов распределения простых чисел.

Заметим теперь, что, как было показано еще Эйлером,

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\ldots+\frac{1}{n^2}+\ldots=\frac{\pi^2}{6},$$

где π — отношение длины окружности к диаметру (см., например, А. М. Яглом и И. М. Яглом, Неэлементарные задачи в элементарном изложении, М., Гостехиздат, 1954, задача 143); отсюда и из иррациональности числа π^2 также следует неограниченность ряда простых чисел (ибо, иначе, иррациональное число $\frac{\pi^2}{6}$ было бы равно рациональному числу

$$\frac{2^2}{2^2-1}\cdot\frac{3^2}{3^2-1}\cdot\frac{5^2}{5^2-1}\cdot\frac{7^2}{7^2-1}\cdot\frac{11^2}{11^2-1}\cdot\ldots\cdot\frac{p^2}{p^2-1},$$

где p — последнее простое число).

[61] Вот еще один выразительный пример: рассуждая, как в тексте, мы могли бы из того, что любое натуральное число $n>1$ можно увеличить, возведя его в квадрат, а 1 так увеличить не удается, сделать нелепый вывод о том, что *единица представляет собой наибольшее натуральное число.*

[62] Широкое обсуждение всего этого круга вопросов имеется в гл. VII книги: Р. Курант и Г. Роббинс, Что такое математика, М.—Л., Гостехиздат, 1947.

[63] Более полное обсуждение случаев, в которых можно без риска сделать ошибку пользоваться рассуждениями типа тех, с которых начинается эта тема, имеется в дополнении 1 к книге: И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, М.—Л., Гостехиздат, 1951.

[64] В комментарии Симплиция к сочинению Аристотеля («De coelo», берлинское академическое издание, VII, стр. 412) говорится, что Архимед и Зенодор доказали, — а известно это было еще до Аристотеля, — что круг среди плоских фигур того же периметра имеет наибольшую площадь, а шар среди тел, имеющих ту же величину поверхности, — наибольший объем.

[65] Рассматриваемой в этой теме так называемой «изопериметрической задаче», различным ее доказательствам (в том числе и свободным от указанного в тексте недостатка), следствиям и обобщениям посвящена специальная книжка: Д. А. Крыжановский, Изопериметры, М., Физматгиз, 1959. В конце этой книги указана также другая, относящаяся к этой теме литература, из которой здесь особенно отметим гл. VII неоднократно уже упоминавшейся книги: Р. Курант и Г. Роббинс, Что такое математика; гл. X, VIII, а частично и XI книги: Д. Пойя, Математика и правдоподобные рассуждения, М., ИЛ, 1957; § 5 (а частично и § 6) также уже упоминавшейся книги: И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, а также § 10 историко-математической статьи В. Бляшке, Греческая и наглядная геометрия, сборник «Математическое просвещение», вып. 3, 1958, и совершенно отличную от этой темы по своим методам книгу: Л. А. Люстерник, Кратчайшие линии, М., Гостехиздат, 1955, особенно гл. V.

[66] Число $\varphi(n)$ зависит от n и в теории чисел называется функцией Эйлера. Очевидно, что $\varphi(n)$ можно было бы определить и как число правильных несократимых дробей со знаменателем n.

Если p, q, r, ... — все различные простые делители n, то вообще

$$\varphi(n) = n\frac{p-1}{p}\cdot\frac{q-1}{q}\cdot\frac{r-1}{r}\cdots$$

Доказательство этой теоремы можно найти в любом учебнике элементарной теории чисел (например, И. В. Арнольд, Теория чисел, М., Учпедгиз, 1939, § 23; А. А. Бухштаб, Теория чисел, М., Учпедгиз, 1960, гл. 10); см. также А. М. Яглом и И. М. Яглом, Неэлементарные задачи в элементарном изложении, М., Гостехиздат, 1954, задача 12.

[67] В связи с затронутыми здесь вопросами см., например, книгу: А. Леман, От периодических десятичных дробей к теории чисел (A. Leman «Von periodischen Dezimalbruch zur Zahlentheorie» B. G. Teubner Mathematisch-physikalische Bibliothek, № 19). В своих «Disquisitiones arithmeticae» (art. 312—318) Гаусс также уделял внимание периодическим дробям, однако при изложении их он широко пользуется данными теории чисел, развитыми им в предшествующих главах того же сочинения. См. также § 46 уже упоминавшейся книги: И. В. Арнольд, Теория чисел.

[68] Проходящую через точки A и B окружность можно определить как геометрическое место таких точек M, что направленный угол между прямыми MA и MB, т. е. угол, принимаемый положительным или отрицательным, в зависимости от того, противоположно ли направление вращения от MA к MB направлению вращения часовой стрелки, или совпадает с этим направлением. [Относительно направленных углов см., например, написанные Д. И. Перепелкиным решения задач к книге: Ж. Адамар, Элементарная геометрия, ч. I, М., Учпедгиз, 1957, стр. 488—489.]

[69] Это определение кривых постоянной ширины приводит к естественной мысли рассмотреть кривые, которые могут свободно вращаться, скажем, в правильном треугольнике, соприкасаясь все время со всеми сторонами треугольника (так называемые Δ-кривые, имеющие весьма много общего с кривыми постоянной ширины), или в правильном n-угольнике, где n задано, или в неправильном n-угольнике. Относительно возникающих здесь вопросов см. И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, М.— Л., Гостехиздат, 1951, § 8.

[70] Условие выпуклости кривой постоянной ширины, строго говоря, излишне. Действительно, можно доказать, что всякая жорданова кривая (см. стр. 250—251) постоянной ширины должна быть выпуклой; однако в связи с трудностями, с которыми связано точное описание понятия жордановой кривой, это доказательство неизбежно будет не особенно простым.

[71] Выраженное в теореме IV свойство кривой постоянной ширины справедливо и может быть доказано для всех выпуклых кривых (см. И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, § 1).

[72] Треугольник Рело в ряде отношений является «крайней» кривой постоянной ширины и обладает рядом «максимальных» и «ми-

нимальных» свойств. Так, например, можно доказать, из всех кривых данной постоянной ширины треугольник Рело имеет наибольшую замыкающую окружность (ср. тему 16) или ограничивает наименьшую площадь (см. И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, задачи 89 и 97, 91).

[73] Относительно доказательства этой теоремы (так называемая теорема Барбье) см., например, Л. А. Люстерник, Выпуклые фигуры и многогранники, М., Гостехиздат, 1956, § 7, и И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, решения задач 88, 92 и 96. § 7 последней книги содержит довольно большой материал по теории кривых постоянной ширины, в значительной степени дополняющий изложение настоящей темы.

[74] Относительно точного смысла этого утверждения см., например, Д. И. Перепелкин, Геометрические построения в средней школе, М., Учпедгиз, 1953, § 2 гл. I.

[75] В предисловии к своей книге «Геометрия циркуля» Лоренцо Маскерони [Lorenzo Mascheroni (1750—1800), «La geometria del compasso», Pavia, 1797] говорит, что он занялся построениями с помощью одного лишь циркуля, исходя первоначально из практических потребностей. По различным причинам построения, осуществляемые циркулем, значительно точнее построений линейкой — обстоятельство, которым, как это было известно Маскерони, пользуются астрономы при точной разметке делительных кругов своих инструментов. В итоге глубокого изучения этого рода построений Маскерони пришел впоследствии к тому, что все задачи Евклида могут быть решены одним лишь циркулем. Свою книгу Маскерони снабжает посвящением Наполеону, в котором он восхваляет освободителя Верхней Италии. Со своей стороны, Наполеон по возвращении из Италии (в речи 10 декабря 1797 г.) привлек внимание французских академиков к неизвестной еще тогда во Франции «Геометрии циркуля».

С 1797 г. Маскерони был членом международной метрической комиссии в Париже.

О датском математике Георге Море (Georg Mohr) [родился в 1640 г. в Копенгагене, умер в 1697 г. в Кислингсвальде около Герлица, где он в качестве гостя провел два последних года своей жизни у своего друга, известного математика Вальтера Чирнгауза (Walter v. Tsschirnhaus)] мы знаем лишь по его книге «Датский Евклид» («Euclides Danicus», Амстердам, 1672), случайно обнаруженной в 1928 г. известным датским геометром Ельмслевом в букинистическом магазине в Копенгагене; первая часть этой книги посвящена упомянутой задаче. [Впрочем, в последнее время появилось сообщение о нахождении еще одной книги Г. Мора.] В настоящее время «Датский Евклид» переиздан (Kobenhavn, 1928) Датским королевским научным обществом и снабжен немецким переводом Ю. Паля (Julius Pàl).

Относительно доказательства результата Мора — Маскерони см., например, Р. Курант и Г. Роббинс, Что такое математика, М.—Л., Гостехиздат, 1947, стр. 214—219; И. М. Яглом, Геометрические преобразования, II, М., Гостехиздат, 1956, стр. 209—214; Б. И. Аргунов и М. В. Балк, Геометрические построения на плоскости, М., Учпедгиз, 1957, стр. 228—238; А. Адлер, Теория геометрических построений, Л., Учпедгиз, 1940, гл. III.

[76] Якоб Штейнер (Jacob Steiner, 1796—1863), родом швейцарец, начал свою деятельность, поощрявшуюся вначале братьями Вильгельмом и Александром Гумбольдтами, в Берлине, где в течение ряда лет преподавал в средней школе, терпя известные неудобства, связанные с отсутствием у него высшего образования; с 1834 г. занял должность профессора университета и был избран в члены Академии наук. Биографический очерк о нем см., например, в книге: Д. А. Крыжановский, Изопериметры, М., Физматгиз, 1959, и в книге: Я. Штейнер, «Геометрические построения, выполняемые с помощью прямой и постоянного круга», М., Учпедгиз, 1939; последняя книга специально посвящена рассматриваемой тут задаче. Относительно доказательства сформулированного результата Штейнера [точнее было бы сказать — Понселе — Штейнера; Жан Виктор Понселе (1788—1867) — выдающийся французский геометр, создатель проективной геометрии] см. кроме названной книги Штейнера также указанные выше книги Р. Куранта и Г. Роббинса, стр. 272—274; И. М. Яглома, стр. 123—127; Б. И. Аргунова и М. Б. Балка, стр. 238—246; А. Адлера, гл. II.

[77] Вопрос о возможности определения центра окружности с помощью одной лишь линейки впервые был поставлен выдающимся немецким математиком конца прошлого и начала этого века Давидом Гильбертом (D. Hilbert); им же был предложен и метод решения этого вопроса. Дальнейшее развитие эти идеи получили в работе его ученика Д. Кауера (Detlef Cauer, Über die Konstruktion des Mittelpunktes eines Kreises mit dem Lineal, Math. Ann. **73**, 1912, стр. 90—94, и **74**, 1913, стр. 462—464).

[78] Можно даже доказать, что окружность S можно так спроектировать в другую окружность S', что центр O окружности S перейдет в произвольно заданную внутри S' точку O'. См. по этому поводу И. М. Яглом, Геометрические преобразования, II, § 1, гл. 1 третьей части, где содержится несколько иное изложение рассматриваемых здесь вопросов.

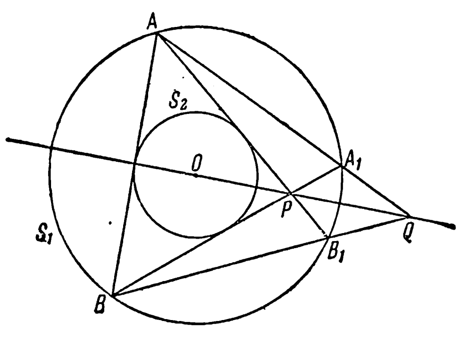


Рис. 140.

[79] Это рассуждение предполагает, что рассматриваемые окружности не являются концентрическими. Действительно, в противном случае отрезки K_1K_2 и L_1L_2 имели бы общую середину W и точка O попала бы на перпендикуляр к плоскости $K_1L_1L_2K_2$ в точке W; наши конусы были бы не косыми, а прямыми и сопряженные сечения не отличались бы от исходных круговых сечений. Это обстоятельство не является случайным: для двух концентрических окружностей S_1 и S_2 нельзя указать центрального про-

ектирования, переводящего их снова в окружности S_1' и S_2', а их общий центр — в точку, не являющуюся центром S_1' (или S_2'). Это проще всего установить, показав, что *общий центр двух окружностей S_1 и S_2, про которые известно, что они являются концентрическими, можно найти с помощью одной линейки.* Соответствующее построение изображено на рис. 140, где AB и AB_1 — касательные, проведенные

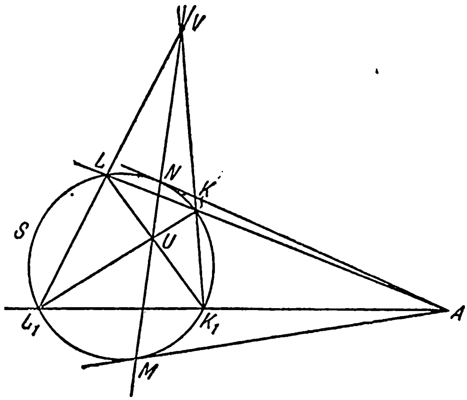


Рис. 141.

к окружности S_2 из точки A окружности S_1 (на вопросе о построении этих касательных с помощью одной линейки мы ниже остановимся отдельно); BA_1 — касательная к S_2, проведенная из точки B; прямая PQ, соединяющая точки пересечения AB_1 и BA_1, AA_1 и BB_1, проходит через общий центр S_1 и S_2, поскольку она является осью симметрии рис. 140; аналогично можно построить еще один диаметр P_1Q_1 окружностей S_1 и S_2, пересечение которого с PQ и определит искомый центр O.

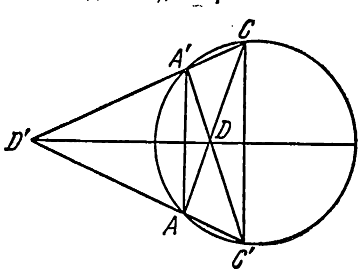


Рис. 142.

Построение одной линейкой касательных AM и AN к окружности S_1, проходящих через заданную точку A, изображено на рис. 141; здесь AKL и AK_1L_1 — произвольные секущие S; U и V — точки пересечения KL_1 и LK_1, KK_1 и LL_1; M и N — точки пересечения UV с окружностью. Относительно доказательства правильности этого построения см., например, Адамар, Элементарная геометрия, ч. 1, М., Учпедгиз, 1958, гл. IV дополнений к третьей книге; И. М. Яглом, Геометрические преобразования, II, задача 153.

[80] Оба разобранных нами случая относятся к непересекающимся окружностям. Попытка найти отображение посредством центральной проекции, также и для двух пересекающихся окружностей, не приводит к цели. Такое отображение свидетельствовало бы о невозможности построения центра окружности с помощью одной лишь линейки, между тем как в случае двух пересекающихся окружностей, как об этом было упомянуто в конце п. 1, такое построение выполнимо. Рис. 142 показывает, как, имея две параллельные хорды AA' и CC', можно найти диаметр: прямая DD', в силу симметрии, проходит через центр. С помощью второй пары параллельных хорд можно, следовательно, найти второй диаметр, который в пересечении с первым даст искомый центр. Дело теперь, значит, сводится к тому, чтобы с помощью одной лишь линейки построить две параллельные хорды.

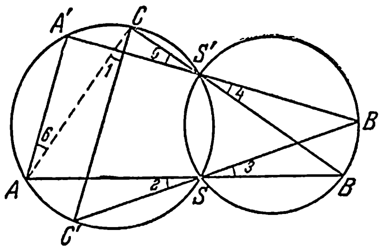


Рис. 143.

Чрезвычайно простой способ такого построения показан на рис. 143. Пусть хорда AA' проведена в одной из данных нам окружностей произвольно. Проведем из A прямые ASB и $BS'C$ и соответственно из A' — прямые $A'S'B'$ и $B'S'C'$. Найденные таким путем точки C и C' будут концами хорды CC', параллельной AA'. В параллельности этих хорд можно убедиться непосредственно, рассмотрев углы, обозначенные на рисунке цифрами *1—6*. Углы *1* и *2* равны как вписанные, опирающиеся на одну и ту же дугу; по той же причине $\angle 3 = \angle 4$, $\angle 5 = \angle 6$. Углы *2* и *3* и соответственно *4* и *5* равны как вертикальные. В силу равенства всех этих углов мы получаем, что и $\angle 1 = \angle 6$. А так как эти углы являются внутренними накрестлежащими при прямой AC, то мы заключаем, что прямые AA' и CC' параллельны.

Построение с помощью одной линейки центров трех непересекающихся кругов заметно сложнее изложенных и предполагает знание более трудных геометрических теорий. Такое построение можно найти, например, у Кауера (см. примечание [77]).

[81] Работа Бонзе (Bonse) помещена в Archiv d. Math. u. Phys. (3), 12, 1907, стр. 292—295. М. Денн — известный немецкий математик. См. также Р. Ремак (R. Remak), там же (3), 15, 1908, стр. 186—193. В настоящее время наука ушла действительно гораздо дальше утверждения, что $p_{n+1} < 2p_n$, но все же недостаточно далеко, чтобы решить, скажем, вопрос о том, всегда ли между двумя последовательными квадратами, например 100 и 121 или между 121 и 144, дожно лежать по крайней мере одно простое число.

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
«ФИЗМАТГИЗ»

БИБЛИОТЕКА МАТЕМАТИЧЕСКОГО КРУЖКА

Вып. 1. **Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом**, Избранные задачи и теоремы элементарной математики, ч. 1. Арифметика и алгебра.

Вып. 2. **Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом**, Избранные задачи и теоремы элементарной математики, ч. 2. Геометрия (планиметрия).

Вып. 3. **Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом**, Избранные задачи и теоремы элементарной математики, ч. 3. Геометрия (стереометрия).

Вып. 4. **И. М. Яглом и В. Г. Болтянский**, Выпуклые фигуры.

Вып. 5. **А. М. Яглом и И. М. Яглом**, Неэлементарные задачи в элементарном изложении.

Вып. 6. **Е. Б. Дынкин и В. А. Успенский**, Математические беседы.

Вып. 7. **И. М. Яглом**, Геометрические преобразования, I. Движения и преобразования подобия.

Вып. 8. **И. М. Яглом**, Геометрические преобразования, II. Линейные и круговые преобразования.

Вып. 9. **М. Б. Балк**, Геометрические приложения понятия о центре тяжести.

Вып. 10. **Г. Радемахер и О. Теплиц**, Числа и фигуры. Опыты математического мышления.

Цена 54 коп.