Входящий в «Новую школьную энциклопедию полутом «ЧИСЛА И ФИГУРЫ» дает школьникам, их учителям и студентам представление о математике. Его особенностью является то, что он построен как сборник отдельных статей, посвященных различным математическим проблемам или методам; каждая из статей представляет собой весьма глубоко ведущий и в своем роде законченный очерк, а все вместе они образуют калейдоскопичную панораму науки и ее истории. Каждая статья содержит сведения, которые понятны и интересны даже людям, очень далеким от математических проблем.

Материал организован в шесть разделов: «Арифметика», «Геометрия», «Алгебра», «Математический анализ», «Комбинаторика» и «История математики». Каждому разделу предшествует вводная статья. Наиболее полно представлена арифметика, а наименее полно — математический анализ. Это сделано сознательно: хороших учебников по математическому анализу очень много, а вот таких глубоких и доступных для школьника изложений теории чисел до сих пор не было. Малая теорема Ферма, числа Фибоначчи, задачи о представлениях чисел в виде суммы квадратов, уравнения Пелля, цепные дроби и квадратичный закон взаимности изложены весьма подробно и глубоко, что поможет кардинально улучшить преподавание теории чисел.

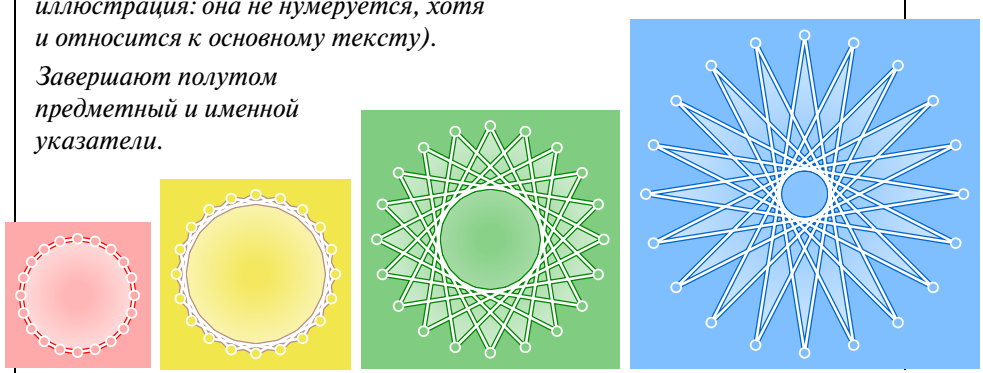
Все разделы книги проиллюстрированы специально для этого издания художником-математиком. Подчеркнем, что все чертежи (в том числе и стереометрические) абсолютно точные: астроида, геликоид, другие кривые и поверхности, многогранники, сечения и проекции тел, вписанные сферы выглядят в точности так, как нарисовано в этой книге. При всем старании ни один школьник не в состоянии нарисовать столь точно и столь много красивых чертежей.

Последний раздел полутома — «История математики» — открывается своеобразной хронологией развития науки. За вводной статьей следует несколько биографических очерков, где жизнеописание великих математиков сочетается с изложением некоторых их идей и открытий.

Основная информационная единица тома — тематическая статья. Различные статьи в зависимости от характера темы и подробности ее рассмотрения занимают объем от одного до девяти разворотов. Основной текст каждой статьи расположен на широкой колонке. На более узких боковых колонках (боковинах) представлена дополнительная информация.

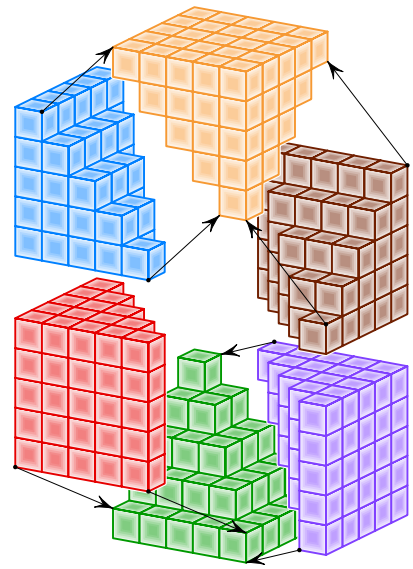
Особо оговорим нумерацию иллюстраций: номера есть лишь у рисунков, относящиеся к основному тексту, непронумерованные иллюстрации относятся к текстам боковин (за исключением случаев, когда на развороте встречается всего одна иллюстрация: она не нумеруется, хотя и относится к основному тексту).

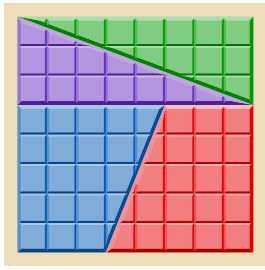
Завершают полутом предметный и именной указатели.

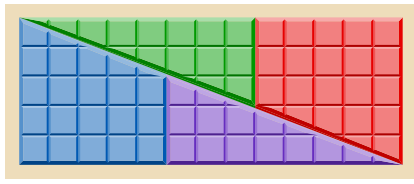


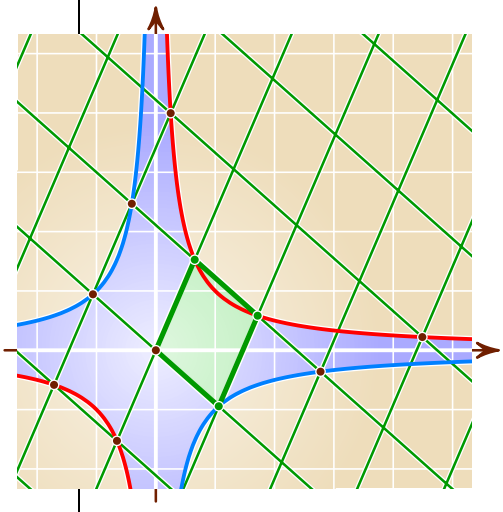
Содержание

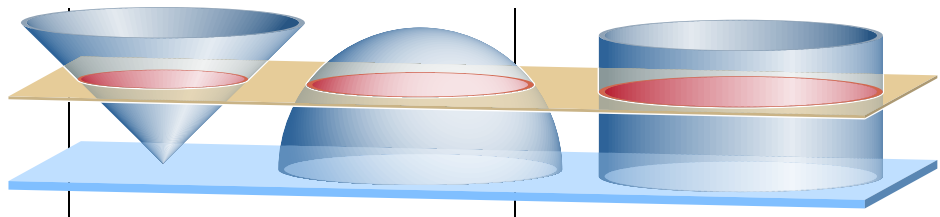
Арифметика



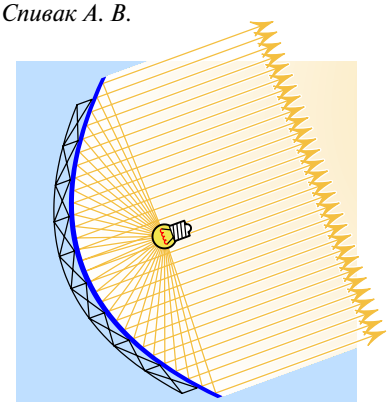


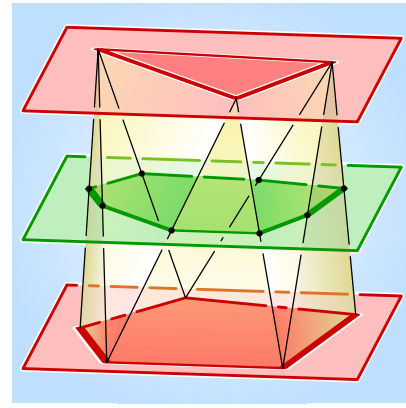


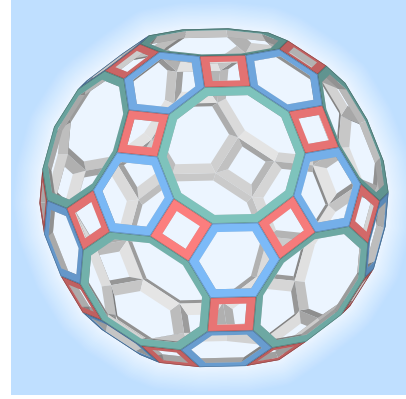


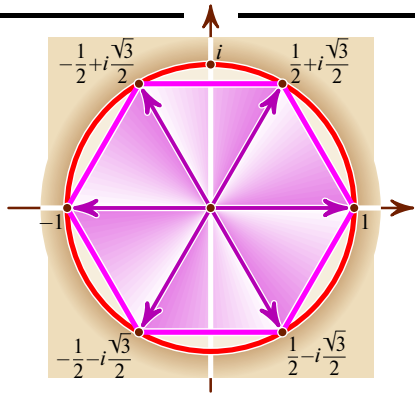


Геометрия

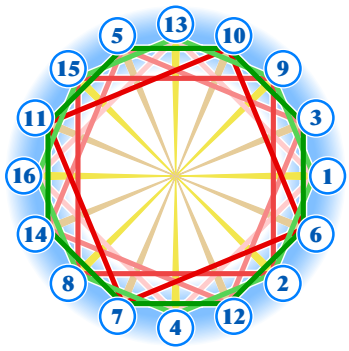


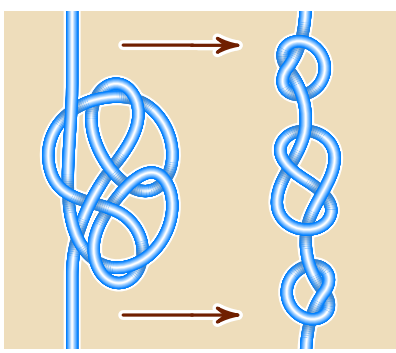


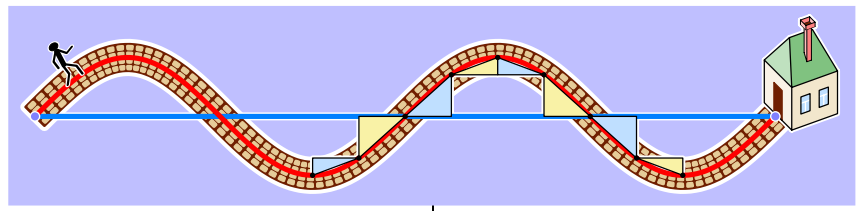




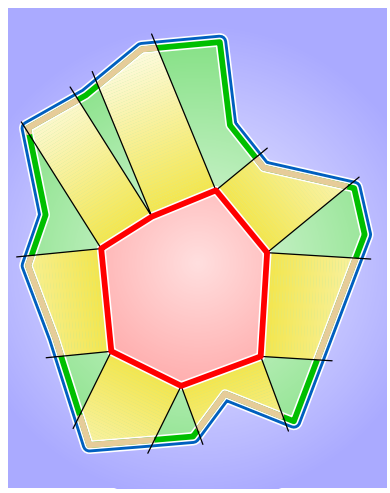
Алгебра



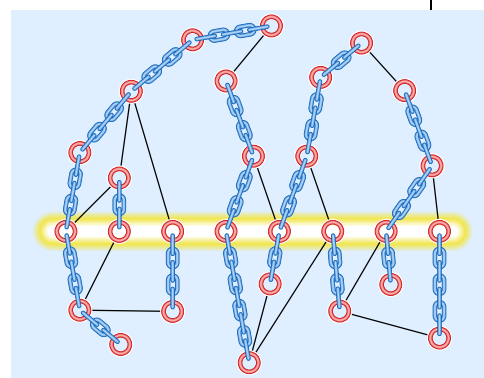


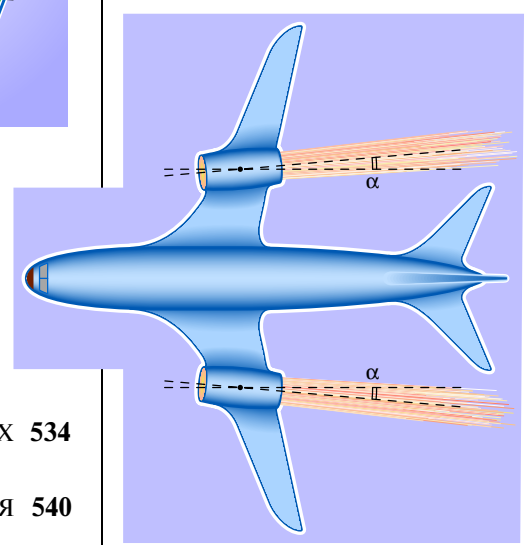


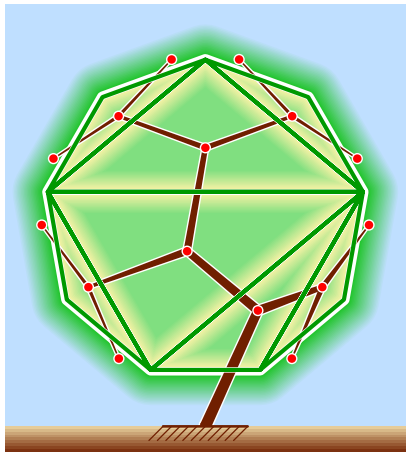
Математический анализ



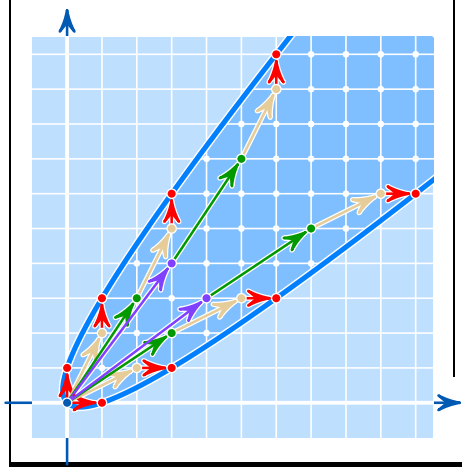
Комбинаторика

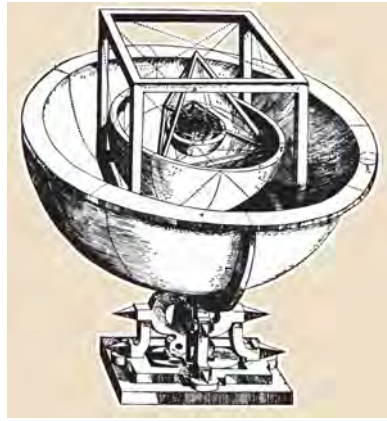


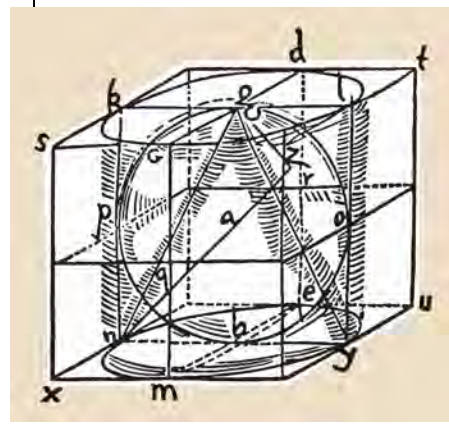




История математики







АРИФМЕТИКА

Традиционный урок алгебры отличается от урока геометрии тем, что на одном вычисляют, а на другом доказывают. В обязательной школьной программе не выделена отдельно теория чисел: элементы арифметики включены в алгебру. Математическая логика тем более не является обязательным предметом.

Но математика давным-давно преодолела древнюю традицию, когда доказательствами снабжали в основном геометрические теоремы, а алгебраические и арифметические методы и закономерности излагали в виде рецептов-алгоритмов, считая, что пользователь скорее всего не нуждается в обосновании, поскольку ему нужен лишь практический результат (а если нуждается — пусть сам догадывается, почему метод работает!).

Арифметика приобрела свою систему аксиом в конце XIX в. — между прочим, практически одновременно с геометрией («Основания геометрии» Д. Гильберта опубликованы в 1899 г.). Основные понятия: натуральное число, следование одного непосредственно за другим и 1 — начальное число натурального ряда.

Аксиомы Дж. Пеано (1858—1932).

1) 1 — *натуральное число*.

2) *Для каждого числа a («число» будет означать пока «натуральное число») существует последующее число a^+*.

3) *Всегда $a^+ \neq 1$, то есть нет числа с последующим числом* 1.

4) *Из $a^+ = b^+$ следует $a = b$, то есть каждое число либо вовсе не является последующим ни для какого числа, либо является последующим точно для одного числа*.

5) (Принцип индукции.) *Каждое множество натуральных чисел, которое содержит число* 1 *и вместе с каждым своим числом a содержит последующее число a^+, содержит все натуральные числа*.

На пятой аксиоме основан метод доказательства с помощью индукции. Чтобы доказать, что некоторым свойством S обладают все числа, доказывают сначала, что им обладает число 1 (база индукции), а затем доказывают его для произвольного числа n^+ при «индуктивном предположении», что число n свойством S уже обладает. В силу аксиомы 5 множество чисел, обладающих свойством S, содержит все числа.

Определение сложения. Каждой паре чисел x, y можно единственным образом сопоставить натуральное число, обозначаемое через $x+y$, чтобы оказались выполненными следующие два условия:

6) $x+1=x^+$ *для каждого x*;

7) $x+y^+=(x+y)^+$ *для каждого x и для каждого y*.

Теорема 1. $(a+b)+c=a+(b+c)$ (*ассоциативность сложения*).

Доказательство — индукция по c. (Над знаками равенства стоят номера примененных аксиом.) **База:** $(a+b)+1 \overset{6}{=} (a+b)^+ \overset{7}{=} a+b^+ \overset{6}{=} a+(b+1)$. **Переход:**

$$(a+b)+c^+ \overset{7}{=} ((a+b)+c)^+ = (a+(b+c))^+ \overset{7}{=} a+(b+c)^+ \overset{7}{=} a+(b+c^+).$$

Теорема 2. $a+1=1+a$.

Доказательство — индукция по a. **База** тривиальна: $1+1=1+1$. **Переход:**

$$a^+ + 1 \overset{6}{=} (a+1)+1 = a+(1+1) \overset{6}{=} a+1^+ \overset{7}{=} (a+1)^+ = (1+a)^+ \overset{7}{=} 1+a^+.$$

Теорема 3. $a+b=b+a$ (*коммутативность сложения*).

Доказательство — индукция по b. База — это утверждение теоремы 2. **Переход:**

$$a+b^+ \overset{7}{=} (a+b)^+ = (b+a)^+ \overset{7}{=} b+a^+ \overset{6}{=} b+(a+1) = b+(1+a) = (b+1)+a \overset{6}{=} b^+ + a.$$

Доказательства следующих утверждений тоже проводятся по индукции.

Теорема 4. *Из равенства $a+b=a+c$ следует $b=c$*.

Определение умножения. Каждой паре двух чисел x, y можно единственным образом сопоставить натуральное число, обозначаемое через $x \cdot y$ или через xy, так, чтобы были выполнены следующие условия:

8) $x \cdot 1 = x$,

9) $x \cdot y^+ = x \cdot y + x$ *для каждого x и для каждого y*.

Теорема 5. $ab \cdot c = a \cdot bc$ (*ассоциативность умножения*).

Теорема 6. $a \cdot b = b \cdot a$ (*коммутативность умножения*).

Теорема 7. $a \cdot (b+c) = a \cdot b + a \cdot c$ (*дистрибутивность*).

Теорема 8. *Из $ab = ac$ следует $b = c$*.

Далее можно ввести понятие степени и доказать свойства степеней, определить понятия меньше и больше и вообще построить арифметику столь же педантично, как на основе аксиом строят геометрию.

Может показаться, что таким образом мы будем изучать узкий круг вопросов. Любое математическое рассуждение (да и не только математическое, а вообще любое рассуждение!) — это текст некоторого языка. У этого языка есть некоторые правила, в соответствии с которыми высказывания признают верными или ошибочными. Тексты не так уж сложно — компьютер с этим справляется! — закодировать последовательностями нулей и единиц. И тогда правила рассуждений станут некоторыми правилами для преобразований чисел. Преобразования эти, конечно,

довольно замысловаты. Но это именно преобразования чисел! Поэтому с полным правом можно сказать, что арифметика включает все остальные вопросы математики.

К. Гёдель (1906—1978) доказал, что множество истинных арифметических теорем является неперечислимым. Таким образом, какой бы мощный компьютер с какой бы умной программой мы ни поставили выводить теоремы арифметики, он или ошибется и выдаст за истину ложное утверждение, или же пропустит бесконечно много истин.

Арифметика богата простыми с виду, но чрезвычайно трудными задачами. Например, $2=5-3=7-5=13-11=19-17=31-29$ — разности соседних простых чисел довольно часто равны 2. Конечно или бесконечно множество таких пар простых чисел-«близнецов»? Никто не знает. Верно ли, что если a, b, c — натуральные числа, причем числа a и b взаимно просты, то уравнение $ap-bq=c$ имеет решение в простых числах p и q? Никто не знает. Конечно или бесконечно множество простых чисел вида n^2+1? Никто не знает!

Каждое четное число (кроме числа 2), насколько удается посчитать на компьютере, является суммой двух простых чисел: $4=2+2$, $6=3+3$, $8=5+3$, ... Но всегда ли это верно или же существует очень большое четное число, не представимое в виде суммы двух простых? Никто не знает. (Правда, в 1937 г. И. М. Виноградов (1891—1983) придумал метод оценивать тригонометрические суммы, при помощи которого удалось доказать, что каждое достаточно большое нечетное число является суммой трех простых чисел. Но окончательно задача еще не решена.)

Легко заметить, что $1+2+3=6$ и $1+2+4+7+14=28$. Много ли еще существует совершенных чисел — чисел n, сумма делителей которых, отличных от n, равна самому числу n, то есть чисел, удовлетворяющих равенству $\sigma(n)=2n$? Л. Эйлер (1707—1783) доказал, что всякое совершенное четное число имеет вид $2^{p-1}(2^p-1)$, где p и 2^p-1 — простые числа. Но до сих пор никто не знает, конечно или бесконечно множество простых чисел вида 2^p-1 — и тем самым конечно или бесконечно множество четных совершенных чисел. Еще интереснее, что не найдено ни одного нечетного совершенного числа — и не доказано, что его не существует!

Тут стоит сделать одно предупреждение. Попытки поиска «голыми руками» или даже попытки компьютерных вычислений *ничего* не дадут: совершенными числами (как и всеми остальными знаменитыми задачами) занимались математики самого высокого уровня. Никакие дилетантские усилия в науке ничего, кроме разочарования, дать не могут.

Яркий тому пример — великая теорема П. Ферма: $x^n+y^n\neq z^n$, где n, x, y и z — натуральные числа, $n>2$. Она казалась столь же недоступно сложной, как задача о совершенных числах. Но дала математике очень много: во многом ради ее решения развивали теорию алгебраических чисел, а решена она была в 1995 г. Э. Уайлсом при помощи глубоко развитой теории эллиптических кривых.

Одна из интересных глав арифметики — теория распределения простых чисел. Через $\pi(x)$ обозначают количество простых чисел, не превосходящих числа x. В «Началах» Евклида доказано, что $\pi(x)\to\infty$ при $x\to\infty$, то есть что множество простых чисел бесконечно. Доказательство в высшей степени красивое. Предположим противное: множество простых чисел конечно и $p_1=2$, $p_2=3$, $p_3=5$, ..., p_n — полный их список. Рассмотрим число $q=p_1p_2\ldots p_n+1$. Оно при делении на любое из простых чисел p_1, p_2, ..., p_n дает в остатке единицу и поэтому не делится ни на одно из них. Получили противоречие: число q больше любого простого числа и поэтому должно быть составным, то есть разлагаться в произведение простых чисел. Но число q не делится ни на какое простое число!

П. Л. Чебышёв (1821—1894) доказал, что отношение $\dfrac{\pi(x)}{x/\ln x}$ при $x>2$ ограничено снизу и сверху числами $\ln\sqrt{2}$ и $\ln 4$; причем если предел $\lim\limits_{x\to\infty}\dfrac{\pi(x)\ln x}{x}$ существует, то он равен 1. Изучая функцию

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}=\prod_{p\text{ простое}}\left(1-\frac{1}{p^s}\right)^{-1},$$

где $s>1$, и ее аналитическое продолжение, Б. Риман (1826—1866) высказал гипотезу, что

$$\pi(x)=\int_2^x\frac{dt}{\ln t}+O(\sqrt{x}\ln x).$$

Эта гипотеза до сих пор не доказана, хотя в 1896 г. Ш. Ж. де ла Валле Пуссен (1866—1862) и Ж. Адамар (1865—1963) доказали, что $\zeta(s)\neq 0$ при $\operatorname{Re} s\geqslant 1$ и вывели отсюда равенство $\lim\limits_{x\to\infty}\dfrac{\pi(x)\ln x}{x}=1$.

Свойства алгебраических и трансцендентных чисел, количество точек с целыми координатами внутри круга $x^2+y^2\leqslant r^2$ и внутри более сложных фигур, проблема Варинга представления чисел в виде суммы некоторого фиксированного количества n-х степеней неотрицательных целых чисел — этими и многими другими сложными задачами, в том числе теснейшим образом связанными с практикой, занимается теория чисел. Приглашаем познакомиться с некоторыми ее интересными и доступными школьнику результатами по нашим статьям.

Как считали в старину и как писали цифры?

Вот как выглядит запись числа 1917 — года, когда, по словам профессора Д. Е. Меньшова, московские математики начали заниматься тригонометрическими рядами, — в разных системах нумерации.

Все числа мы привыкли записывать с помощью десяти знаков — цифр 0, 1, 2, 3, 4, 5, 6, 7, 8 и 9. Например, число, состоящее из четырех сотен, четырех десятков и четырех единиц, мы записываем так: 444. Одна и та же цифра «4» обозначает количество единиц, если она стоит на последнем месте, количество десятков — если на предпоследнем, количество сотен — на предпредпоследнем. Такой принцип записи чисел называют позиционным. Он появился тысячелетие назад и оказался самым удобным среди всех прочих принципов счисления.

Устный счет. Всегда ли люди пользовались позиционным принципом? Обратимся к устному счету. Для цифр мы употребляем специальные названия: «нуль», «один», «два», . . . , «девять». Для следующего числа есть новое слово «десять»; мы не говорим «один-нуль», хотя и записываем его с помощью единицы и нуля: 10.

Названия чисел от 11 до 99, как правило, составлены из названий первых чисел: «одиннадцать» (один-на-десять), «тридцать один» (три-десять-один). Для 100 мы употребляем новое слово — «сто». Все наименования чисел от 101 до 999 опять составные, а для 1000 вводится новое слово — «тысяча». Далее появляются слова «миллион», «миллиард», «триллион». Как видите, по мере роста чисел возрастает и количество используемых слов. Следовательно, наш способ наименования чисел не является позиционным. Он сохранил следы каких-то более старых нумераций.

В одной из таких старых нумераций — римской — есть специальные знаки для единицы (I), пяти (V), десяти (X), пятидесяти (L), ста (C), пятисот (D) и тысячи (M). Остальные числа записывают при помощи этих символов. Например, III — запись числа 3 (I+I+I), IV — числа 4 (V−I), XCI — числа 91 (C−X+I). Число 444 в римской системе счисления записывается в виде CDXLIV; здесь четыре единицы записаны символами IV, четыре десятка — XL, а четыре сотни — CD.

С числами, записанными в римской системе нумерации, очень трудно производить арифметические действия. Попробуйте умножить 444 на 36, если оба числа обозначены римскими цифрами! Сами римляне пользовались для арифметических операций специальной счетной доской — абаком.

Есть еще один существенный недостаток римской системы нумерации: она не дает удобного способа записи больших чисел. Например, чтобы записать число 1 000 000, надо либо 1000 раз повторить знак M, либо ввести новый символ. Причина в том, что римская система не является позиционной. Знак V, например, означает только пять единиц, а не пять десятков и не пять сотен.

Римская нумерация весьма древнего происхождения. Начертание цифр было заимствовано у более ранних обитателей Италии — этрусков. Этрусский знак сотни ⊕ обратился сначала в Ⓒ, затем в C. Этрусское 50, писавшееся как ↓, обратилось сначала в ⊥, затем в ⊥ и затем, наконец, в L. Десятку они обозначали крестом + или ×, римляне оставили лишь вторую форму. Пятерку этруски писали как ∧ или ∨ — половина знака для десятки. (Идею «деления символа» использовали и ацтеки: число 400 они обозначали 𝍠, 300 — 𝍡, 200 — 𝍢, 100 — 𝍣.) ■

Римская нумерация не является строго десятичной. В ней сохранились следы другого основания — пяти: есть специальные знаки для пяти, пятидесяти и пятисот.

1	I	11	XI	30	XXX	400	CD
2	II	12	XII	40	XL	500	D
3	III	13	XIII	50	L	600	DC
4	IV	14	XIV	60	LX	700	DCC
5	V	15	XV	70	LXX	800	DCCC
6	VI	16	XVI	80	LXXX	900	CM
7	VII	17	XVII	90	XC	1000	M
8	VIII	18	XVIII	100	C	2000	MM
9	IX	19	XIX	200	CC	3000	MMM
10	X	20	XX	300	CCC	4000	MMMM

В нашем устном счете имеются некоторые черты, напоминающие римскую систему. Так, мы тоже используем сложение, образуя числительные от 11 до 19. Но начиная с 20 мы используем и умножение, чего нет в римской системе: «двадцать» означает «дважды десять», тридцать — «трижды десять». В русском языке сохранились и следы нумерации с основанием 40, которой пользовались наши предки. Действительно, для этого числа используем новое, несоставное название — «сорок». Известны выражения: «сорок сороков церквей», «сорок сороков черных соболей». О том, что число 40 когда-то играло особую роль при счете, говорят и некоторые связанные с ним поверья. Так, сорок первого медведя считали роковым для охотника. (А в связи с тем, что некогда была распространена двенадцатеричная система счисления, и сейчас некоторые считают число 13 несчастливым.)

Во французском языке сохранились следы нумерации с основанием 20; число 80 читается: «quatre-vingts» — «четыре-двадцать», число 90 — «quatre-vingt-dix» — «четыре-двадцать-десять». Следы двадцатеричной системы сохранились в английском и голландском языках, следы пятеричной — в скандинавских языках.

Итак, устная речь показывает, что наши предки пользовались непозиционной нумерацией, причем основаниями, кроме десяти, служили и другие числа. ■

Счет у первобытных народов. Еще недавно существовали народы, в языке которых были названия только двух чисел: «один» и «два». Но это не значит, что представители этих народов не могли сосчитать до трех. У туземцев островов, расположенных в Торресовом проливе (отделяющем Новую Гвинею от Австралии), единственными числительными являлись «урапун» (один) и «окоза» (два). Островитяне считали так: «окоза-урапун» (три), «окоза-окоза» (четыре), «окоза-окоза-урапун» (пять) и «окоза-окоза-окоза» (шесть). О числах начиная с семи туземцы говорили «много». Таким образом, они освоили лишь несколько первых натуральных чисел. Кстати, пословицы говорят, что именно так дело обстояло и у наших предков. Мы говорим: «у семи нянек дитя без глаза», «семь бед — один ответ», «семеро одного не ждут», «семь раз отмерь, один раз отрежь». Очевидно, слово «семь» употреблено в смысле «много»: если нянек слишком много — за ребенком нет присмотра, много бед — один ответ.

Но вернемся к нашему рассказу. Очень рано у людей появилась необходимость сообщать друг другу о том, что такое-то число предметов должно быть доставлено через столько-то дней или что племя должно выставить такое-то число воинов. Вот как, по рассказу русского путешественника Н. Н. Миклухо-Маклая, поступали на Новой Гвинее: «Излюбленный способ счета состоит в том, что папуас загибает один за другим пальцы руки, причем издает определенный звук, например «бе-бе». Досчитав до пяти, он говорит «ибон-бе» (рука). Затем он загибает пальцы другой руки, повторяя «бе-бе», пока не доходит до «ибон-али» (две руки). Затем он считает дальше, приговаривая «бе-бе», пока не доходит до «самба-бе» и «самба-али» (одна нога, две ноги). Если нужно считать дальше, папуас пользуется пальцами рук и ног кого-то другого».

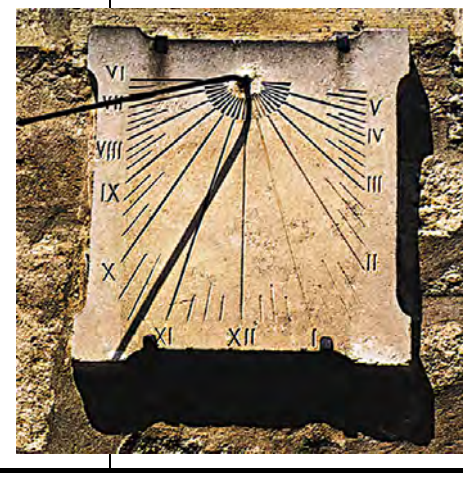
Итак, предметы при счете сопоставлялись пальцам рук и ног. При переговорах папуасу достаточно было сказать, например, что он дошел до третьего пальца правой ноги. Островитяне Торресова пролива для такого счета употребляли не только пальцы, но и другие части тела (запястье, локоть, плечо), всегда в одном и том же порядке. Так они могли считать до 33.

Необязательно считать десятками. Можно вести счет двойками, тройками или даже дюжинами. Возьмем, например, в качестве основания число 2, а в остальном будем поступать точно так же, как мы это делаем в привычной (десятичной) системе. Понадобятся всего две цифры: 0 и 1. Тогда число два запишется как 10, три — как 11, четыре — 100, семь — 111, тридцать пять — 100 011 (проверьте!).

Для записи чисел в троичной системе нужны три цифры: 0, 1 и 2. Число три запишется как 10, четыре — 11, тридцать пять — 1022.

Можно считать не только двойками, тройками или десятками, но и дюжинами: сервизы обычно составляют из 12 чашек, 12 блюдец, 12 тарелок, а комплекты мебели — из 12 стульев или кресел. Существует даже специальное название для дюжины дюжин — гросс.

О широком распространении двенадцатеричной системы свидетельствуют такие факты: мы делим год на 12 месяцев, а сутки — на 24 часа, причем в повседневной жизни часы считаем только до 12, а затем начинаем счет заново («час дня», «два часа дня»). Число 12 часто встречается также в сказках и легендах (двенадцатиглавый змей, двенадцать братьев-разбойников), что свидетельствует о древнем происхождении двенадцатеричной системы. ■



Одни из первых солнечных часов имели шкалу с римскими цифрами.

1 •	11 ◇•	30 P◇
2 ••	12 ◇••	40 PP
3 ∴	13 ◇∴	50 PP◇
4 ⁘	14 ◇⁘	60 PPP
5 ⁘•	15 ◇⁘•	70 PPP◇
6 ⁘|•	16 ◇⁘|•	80 PPPP
7 ⁘|••	17 ◇⁘|••	90 PPPP◇
8 ⁘|∴	18 ◇⁘|∴	100 𓍢
9 ⁘|⁘	19 ◇⁘|⁘	200 𓍢
10 ◇	20 P	300 𓍣
		400 𓍣
		500 𓍣𓍢
		1000 𓍤
		8000 𓆐

Обозначение чисел у ацтеков (Мексика) в XI—XVI вв.

Хорошей иллюстрацией к такому способу счета служат обозначения чисел, принятые в XI—XVI вв. индейцами племени ацтеков (Мексика): единицу они обозначали точкой, двойку — двумя точками и так далее до пяти. В запись числа 6 входила вертикальная черта, отделявшая пять первых точек от шестой. Ясно, что счет вели группами по 5 предметов. Черта отделяла одну такую группу от другой, причем сама черта никакого числа не обозначала. ■

Египетское умножение. Знакомясь с римской системой, мы убедились, что умножать числа, записанные в непозиционной системе, очень неудобно. Как же считали древние египтяне? Умножение и деление они производили при помощи последовательного удвоения чисел (то есть использовали идею двоичной системы счисления). Пусть, например, надо умножить 19 на 37. Египтяне последовательно удваивали число 37, записывая в правый столбец результаты удвоения, а в левый — соответствующие степени двойки:

1	37
2	74
4	148
8	296
16	592

Удваивали до тех пор, пока не оказывалось, что из чисел левого столбца можно составить множитель. В нашем примере: 19 = = 1+2+16. Затем складывали числа, стоящие в соответствующих строках справа. В нашем примере: 37 + 74 + 592 = 703. Так получали произведение. ■

Суть этого способа заключается в том, что для нахождения количества элементов множества его элементы сопоставляли с частями тела, а иногда и просто с палочками. Разумеется, наиболее удобными «инструментами» являются пальцы, вследствие чего предметы при пересчете чаще всего группировали по 5, по 10 или по 20. Это и объясняет, что чаще всего основанием системы счисления служило число 10 (по числу пальцев на обеих руках), реже — 5 или 20.

Со временем хозяйство становилось все более сложным и обширным, так что счет на пальцах перестал удовлетворять людей. Они привыкли при счете располагать предметы устойчивыми группами по 2, по 10 или по 12. Появились специальные слова для обозначения таких совокупностей. Так, у туземцев Флориды слово «на-куа» означало 10 яиц, «на-банара» — 10 корзин. Но слово «на», которое, казалось бы, соответствует числу 10, отдельно не употреблялось. То же можно было наблюдать на островах Фиджи и на Соломоновых островах, где имелись специальные знаки для 100 челноков, 100 кокосовых орехов, 1000 орехов, а отвлеченных чисел не было. Числа были именованными, это еще «числа-совокупности» конкретных предметов.

С течением времени такими устойчивыми «числами-совокупностями» начали обозначать не только данные предметы, но и другие, похожие на них. Например, «числа-совокупности», обозначающие определенное количество орехов, могли впоследствии использовать для счета любых круглых предметов. Это привело к тому, что во многих языках первобытных народов образовалось несколько рядов числительных: одни употреблялись только для счета людей, другие — для круглых предметов, третьи — для продолговатых и так далее. Например, у чишмиенов (Британская Колумбия, ныне одна из канадских провинций) имелось 7 видов числительных, каждый из которых употребляли для счета предметов определенного вида. У большинства народов числа, которыми считали деньги, постепенно вытеснили все остальные. Они-то и стали теми универсальными числами, которые позволили считать любые предметы.

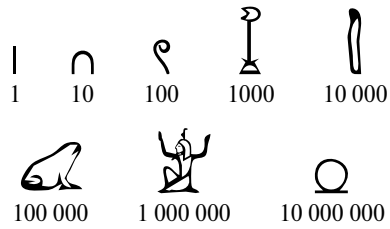
Так образовались числа, которым соответствовали «числа-совокупности»: если счет велся десятками, то появились названия для десяти, десяти десятков, десяти сотен. Кроме того, особые названия получали, как правило, все числа от 1 до 10. Что же касается чисел 11, 12, . . . , 19, 21 и так далее, то они составлялись из основных при помощи тех операций, которые первоначально производили над пересчитываемыми предметами. Так, на языке кламатов (Северная Америка), а также племен Британской Колумбии при обозначении составных чисел использовались специальные глаголы. Например, индеец говорил: «На дважды десять плодов я кладу сверху шесть» — и это обозначало 26 плодов. Такая фраза полностью соответствует фактическому пересчету: индейцы располагали 10 предметов в ряд, с 11-го начинался новый ряд и так далее. Постепенно эти двигательные операции перешли в арифметические.

Основной операцией для образования составных числительных было сложение, но наряду с этим употреблялось и вычитание, а иногда и умножение. Например, в русском языке, как уже упоминалось, для образования числительных употребляются и сложение, и умножение (27 — два × десять + семь). Так происходило освоение натурального ряда.

Посмотрим теперь, какими были первые записи чисел и как люди оперировали с ними. ■

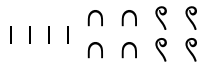
Египетская нумерация. Одна из древнейших нумераций — египетская. До нас дошли надписи, сохранившиеся внутри пирамид, на плитах и обелисках. Они состоят из картинок-иероглифов, которые изображают птиц, зверей, людей, части человеческого тела (глаза, ноги) и различные неодушевленные предметы. Такой способ письма характерен для ранних ступеней культуры. (Забавно, но когда компьютеры перешли из лабораторий ученых к массовому пользователю, произошла реанимация иероглифов: в популярной системе Windows пользователь видит много картинок и мало текста.) Подобные письмена были у обитателей Центральной Америки — индейцев племени майя, в Перу. Расшифровка их представляет огромные трудности, так часто неизвестны ни звучания слов, ни значения иероглифов. Казалось бы, задача неразрешима. И все-таки многие надписи уже прочитаны! Сначала были разгаданы письмена древних египтян, затем вавилонская клинопись. В 30-х гг. XX в. были прочитаны долго не поддававшиеся расшифровке хеттские надписи.

Сохранились лишь два математических папируса, позволяющих судить о том, как считали древние египтяне. Один из них хранится в Британском музее в Лондоне, другой — в Музее изобразительных искусств им. А. С. Пушкина в Москве. Для записи чисел древние египтяне использовали иероглифы, означавшие единицу, десять, сто, тысячу, десять тысяч, сто тысяч (лягушка), миллион (человек с поднятыми руками) и десять миллионов:



Полагают, что иероглиф для сотни изображает измерительную веревку, для тысячи — цветок лотоса, для десяти тысяч — поднятый вверх палец, а для десяти миллионов — Вселенную. Остальные числа составлялись из основных с помощью операции сложения. При этом запись производилась не слева направо, а справа налево. Например, число 15 записывали так:

А число 444 писали так:



Как видите, древнеегипетская нумерация похожа на римскую, только при записи чисел не использовали вычитание. ■

Греческий алфавит и алфавитное обозначение чисел (ионийская система).

Буква	Название	Числовые значения		
Α α	альфа	$\overline{\alpha}=1$	$\overline{,\alpha}=1000$	$\overset{\alpha}{\text{M}}=\alpha^{\text{M}}=\text{M}\alpha=\alpha.=10\,000$
Β β	бета	$\overline{\beta}=2$	$\overline{,\beta}=2000$	$\overset{\beta}{\text{M}}=\beta^{\text{M}}=\text{M}\beta=\beta.=20\,000$
Γ γ	гамма	$\overline{\gamma}=3$	$\overline{,\gamma}=3000$	$\overset{\gamma}{\text{M}}=\gamma^{\text{M}}=\text{M}\gamma=\gamma.=30\,000$
Δ δ	дельта	$\overline{\delta}=4$	$\overline{,\delta}=4000$	$\overset{\delta}{\text{M}}=\delta^{\text{M}}=\text{M}\delta=\delta.=40\,000$
Ε ε	эпсилон	$\overline{\varepsilon}=5$	$\overline{,\varepsilon}=5000$	$\overset{\varepsilon}{\text{M}}=\varepsilon^{\text{M}}=\text{M}\varepsilon=\varepsilon.=50\,000$
Ϝ ϛ	дигамма*)	$\overline{\varsigma}=6$	$\overline{,\varsigma}=6000$	$\overset{\varsigma}{\text{M}}=\varsigma^{\text{M}}=\text{M}\varsigma=\varsigma.=60\,000$
Ζ ζ	дзета	$\overline{\zeta}=7$	$\overline{,\zeta}=7000$	$\overset{\zeta}{\text{M}}=\zeta^{\text{M}}=\text{M}\zeta=\zeta.=70\,000$
Η η	эта	$\overline{\eta}=8$	$\overline{,\eta}=8000$	$\overset{\eta}{\text{M}}=\eta^{\text{M}}=\text{M}\eta=\eta.=80\,000$
Θ θ ϑ	тета	$\overline{\theta}=9$	$\overline{,\theta}=9000$	$\overset{\theta}{\text{M}}=\theta^{\text{M}}=\text{M}\theta=\theta.=90\,000$
Ι ι	йота	$\overline{\imath}=10$		
Κ κ	каппа	$\overline{\kappa}=20$		
Λ λ	ламбда	$\overline{\lambda}=30$		
Μ μ	мю	$\overline{\mu}=40$		
Ν ν	ню	$\overline{\nu}=50$		
Ξ ξ	кси	$\overline{\xi}=60$		
Ο ο	омикрон	$\overline{o}=70$		
Π π	пи	$\overline{\pi}=80$		
Ϙ ϙ	коппа*)	$\overline{\varsigma}=90$		
Ρ ρ	ро	$\overline{\rho}=100$		
Σ σ ς	сигма	$\overline{\sigma}=200$		
Τ τ	тау	$\overline{\tau}=300$		
Υ υ	ипсилон	$\overline{\upsilon}=400$		
Φ φ ϕ	фи	$\overline{\varphi}=500$		
Χ χ	хи	$\overline{\chi}=600$		
Ψ ψ	пси	$\overline{\psi}=700$		
Ω ω	омега	$\overline{\omega}=800$		
ϡ	сампи*)	$\overline{\text{ϡ}}=900$		

*) Поскольку в обычном греческом алфавите только 24 буквы, для числовых обозначений были использованы еще и эти три старые буквы.

Ионийская нумерация. По мере развития торговли и ремесел недостатки непозиционных нумераций становились все чувствительнее, и в Малой Азии, где были древнегреческие колонии, в середине V в. до нашей эры появилась система счисления нового типа — ионийская алфавитная нумерация. Числа обозначали при помощи букв алфавита, над которыми ставили черточки: первые девять букв обозначали числа от 1 до 9, следующие девять — числа 10, 20, . . . , 90, следующие девять — числа 100, 200, . . . , 900. Таким образом можно было записать любое число от 1 до 999. Например, число 444 в ионийской нумерации записывали так: $\overline{\nu\mu\delta}$. Числа 1000, 2000, . . . , 9000 греки обозначали теми же буквами, что и числа 1, 2, . . . , 9, но ставили косую черту слева внизу. Для числа 10 000 использовали знак $\overset{\alpha}{\text{M}}$ — это число называли мириадой; две мириады, то есть 20 000, обозначали так: $\overset{\beta}{\text{M}}$. При этом $\overset{\beta}{\text{M}}$ можно было записать еще как β^{M} или как $\text{M}\beta$. Если знак M записывали позади цифры-буквы, то он часто заменялся точкой. Например, 43 458 записывалось так: $\overline{\delta.\gamma\upsilon\nu\eta}$.

Так можно было обозначить все числа вплоть до

$$\overline{,\theta\text{ϡ}\varsigma\theta.,\theta\text{ϡ}\varsigma\theta}=99\,999\,999,$$

то есть до 10^8-1. Более высокие десятичные разряды уже не могли быть записаны в ионийской нумерации и не имели названий в древнегреческом языке. ■

Славянская система алфавитного обозначения чисел.

«Малые» числа						«Великие» числа	
1	а	21	ка	1000	ҁа тысяча	10^6	а тьма
2	в	22	кв	2000	ҁв	10^{12}	а легион
3	г	23	кг	3000	ҁг	10^{24}	а леондр
4	д	24	кд	4000	ҁд	10^{48}	а ворон
5	є	30	л	10 000	а тьма	10^{49}	а колода
6	ѕ	40	м	20 000	в		
7	з	50	н	100 000	а легион		
8	и	60	ѯ	200 000	в		
9	ѳ	70	о	1 000 000	а леондр		
10	і	80	п	2 000 000	в		
11	аі	90	ч или ҷ				
12	ві	100	р				
13	гі	200	с				
14	ді	300	т				
15	єі	400	у				
16	ѕі	500	ф				
17	зі	600	х				
18	иі	700	ѱ				
19	ѳі	800	ѡ				
20	к	900	ц				

Славянское алфавитное обозначение чисел возникло в X в. Его введение приписывают Кириллу. Система была построена по образцу ионийской, причем числовые значения получили лишь те буквы, которые соответствовали буквам греческого алфавита. Так, например, буква «буки» (**б**) не имела числового значения: значение 2 имела буква «веди» (**в**), так как она соответствовала букве β греческого алфавита, а «буки» греческого прототипа не имела. Буква «фита» (**ѳ**), стоявшая на предпоследнем месте в славянском алфавите, имела числовое значение 9, поскольку соответствующая ей греческая буква θ занимала 9-е место.

Сравните ионийскую и славянскую записи числа 444:

$\overline{\text{νμδ}}$ и $\overline{\text{ϡмд}}$. ■

Славянская нумерация. Алфавитные системы были, кроме ионийцев, у древних евреев, финикийцев, армян, грузин и других народов. Алфавитная нумерация была принята и в Древней Руси. Над буквами, обозначающими числа, ставился специальный знак — титло. Это делалось, чтобы отличить их от обычных слов. Интересно отметить, что хотя в славянской нумерации, как и греческой, запись числа шла слева направо, от высших разрядов к низшим, но для чисел от 11 до 19 делалось исключение: сначала писали единицы, а затем знак для 10.

Удобны ли алфавитные системы? Запишем в славянской нумерации число 444:

$\overline{\text{ϡмд}}$

Запись получилась не длиннее десятичной. Это объясняется тем, что в славянской нумерации используются 27 цифр, тогда как в египетской при обозначениях чисел, меньших 1000, использованы лишь три цифры. Алфавитные нумерации имели крупный недостаток: с их помощью нельзя обозначать сколь угодно большие числа. Они удобны только для записи чисел до 1000. Правда, славяне, как и греки, умели записывать и большие числа, но для этого к алфавитной системе добавляли новые обозначения. Числа 1000, 2000, . . . , 9000 они записывали теми же буквами, что 1, 2, . . . , 9, только слева внизу ставили специальный знак ҁ.

Число 10 000 обозначали той же буквой, что и 1, только без титла, но обводили кружком. Называли это число «тьмой». Отсюда, между прочим, произошло выражение «тьма народу». Далее, 10 тем, или 100 000 — «легион». Для обозначения легионов вокруг первых девяти цифр ставили окружность из точек. 10 легионов составляли новую единицу — леондр. Для обозначения леондров соответствующие числа заключали в окружность из черточек. Эти обозначения можно рассматривать как зачатки позиционной системы, так как для обозначения единиц разных разрядов применяли одни и те же символы, к которым добавлялись знаки для определения разряда.

Числа, меньшие десяти миллионов, называли «малыми». Кроме них, в славянской нумерации были «большие», «великие» числа, в которых словом «тьма» обозначали уже миллион. Тьму тем (то есть 10^{12}) называли легионом, легион легионов (10^{24}) — леондром, леондр леондров (10^{48}) — вороном (вороны обозначали буквой в кружке из крестиков), наконец, число 10^{49} называли «колодой». В рукописи XVII в. говорилось: «И более сего несть человеческому уму разумевати», то есть для больших чисел в славянской нумерации названий не было.

Алфавитные нумерации были малопригодны для оперирования с большими числами. В ходе истории эти системы уступили место позиционным. Рудименты алфавитных нумераций сохранились в нашем обиходе и по сей день. Так, мы часто нумеруем пункты буквами алфавита. Правда, буквы служат только для обозначения последовательности, а не количества. Арифметических операций над такими буквами мы не производим. ■

Позиционные системы. Первая из известных нам позиционных систем счисления — шестидесятеричная вавилонская система, возникшая примерно за 2500—2000 лет до нашей эры. Основанием ее служило число 60.

Следовательно, в ней должно было быть 60 цифр, а таблица умножения состояла из $C_{60}^2 = \frac{60 \cdot 59}{2} = 1770$ произведений. Как же вавилоняне запоминали свои цифры и — тем более — чудовищную таблицу умножения?

Вавилоняне записывали числа от 1 до 59 при помощи десятичной системы, применяя принцип сложения и всего две цифры: прямой клин (𒁹) для обозначения единицы и лежачий (𒌋) — для 10. Число 32, например, писали так: 𒌋𒌋𒌋𒁹𒁹. Эти записи и служили цифрами. Число 60 обозначали тем же знаком, что и 1, то есть 𒁹. Так же обозначали и 3600, и 60^3, и все другие степени числа 60. Например, число 92 записывали в виде 𒁹𒌋𒌋𒌋𒁹𒁹.

Таким образом, «цифры» — числа от 1 до 59 — вавилоняне записывали десятичной непозиционной системой, а число в целом — позиционной с основанием 60.

Нумерация вавилонян имела важную особенность: в ней не было знака для нуля! Если был изображен прямой клин 𒁹, то без дополнительных сведений нельзя было определить, какое число записано: 1, 60, 3600 или какая-то другая степень числа 60. Запись числа 92, приведенная выше, могла обозначать не только $92 = 60 + 32$, но и $3600 + 32 = 3632$. Она могла также обозначать $1\frac{32}{60}$ или $1\frac{32}{3600}$ и так далее.

Таким образом, для определения значения числа нужны были еще дополнительные сведения. Впоследствии вавилоняне ввели специальный символ 𒑚 для обозначения пропущенного шестидесятеричного разряда. Например, число 3632 нужно было писать так: 𒁹𒑚𒌋𒌋𒁹𒁹. В конце числа этот символ обычно не ставили.

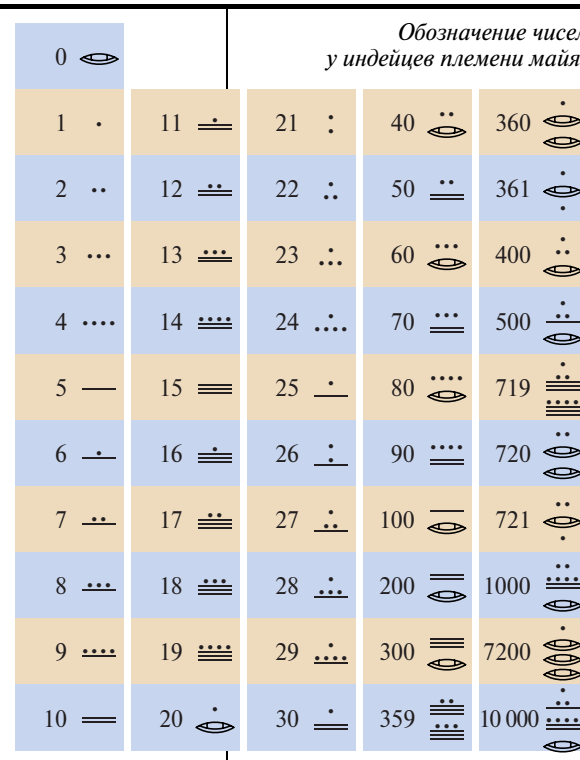
Таблицу умножения вавилоняне никогда не запоминали — это было практически невозможно. Они использовали в вычислениях готовые таблицы умножения.

Шестидесятеричная система сыграла большую роль в развитии математики и астрономии. Следы ее сохранились до наших дней. Так, мы до сих пор делим час на 60 минут, а минуту — на 60 секунд. Точно так же, следуя примеру вавилонян, окружность мы делим на 360 градусов.

В начале нашей эры индейцы племени майя, которые жили на полуострове Юкатан в Центральной Америке, пользовались другой позиционной системой — с основанием 20. Свои цифры индейцы майя, как и вавилоняне, записывали, пользуясь принципом сложения. Единицу они обозначали точкой, а пятерку — горизонтальной чертой, но у майя уже был знак для нуля. Он напоминал полузакрытый глаз. Например, число 20 майя записывали при помощи знака для единицы и внизу знака для нуля (числа писали не в строчку, а в столбец).

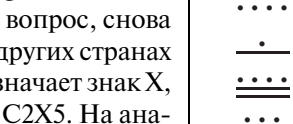
Десятичная система впервые сложилась в Индии не позднее VI в. нашей эры. Здесь же позднее был введен привычный нам символ для нуля. ■

Обозначение чисел у индейцев племени майя.



Основанием системы майя служило число 20, но было одно исключение. 19 цифр записывались посредством точек и черт. Число 20 являлось единицей нового разряда, которая называлась уинальс. Однако единицу следующего разряда образовывали не 20, а 18 уинальсов, называемые туном. Это — единственное отступление от двадцатеричного принципа в системе майя. Оно объясняется тем, что майя делили год на 18 месяцев по 20 дней в каждом плюс еще пять дней. 20 тунов составляли катун, 20 катунов — единицу пятого разряда — цикл, и, наконец, 20 циклов составляли большой цикл. Таким образом, единицами различных разрядов в системе майя были 1, 20, $20 \cdot 18$, $20^2 \cdot 18$, $20^3 \cdot 18$, . . .

В своих календарных и хронологических расчетах майя оперировали очень большими числами. Наибольшее число, найденное в их документах, есть



$$4 \cdot (18 \cdot 20^4) + 6 \cdot (18 \cdot 20^3) + 14 \cdot (18 \cdot 20^2) + 13 \cdot (18 \cdot 20) + 15 \cdot 20 + 1 = 12\,489\,781.$$ ■

Позиционные системы возникли — скорее всего независимо одна от другой — у майя, у вавилонян, у индийцев. Что же привело людей к этому замечательному изобретению? Чтобы ответить на этот вопрос, снова обратимся к истории. В Древнем Китае, Индии и в некоторых других странах существовала следующая система. Пусть, скажем, десятки обозначает знак X, а сотни — C. Тогда запись числа 325 выглядит примерно так: 3C2X5. На аналогичной системе основаны современные счеты: одно и то же количество косточек означает число десятков, сотен, тысяч и так далее в зависимости от того, в каком ряду расположены эти косточки. Такой способ использовали

«Псаммит». Великий математик, механик и инженер древности Архимед (III в. до нашей эры) посвятил целое сочинение наименованиям больших чисел. В сказках, например, встречаются «неразрешимые» задачи: сосчитать звезды на небе, капли в море или песчинки на Земле. Архимед показал, что такие задачи можно решить. Свое сочинение он так и назвал: «Псаммит» («Исчисление песка»). В нем он построил систему, в которой имелись числа, не только превосходящие количество песчинок в его родной Сицилии, но и во всей Вселенной. (Что понимали греки под Вселенной? Архимед, вслед за Аристархом Самосским, полагал, что в центре Вселенной находится Солнце, а Земля и другие планеты вращаются вокруг него. Вселенная же имеет форму сферы, на поверхности которой расположены неподвижные звезды. Это была первая гелиоцентрическая система мира.)

Для подсчета количества песчинок Архимед должен был хотя бы приблизительно определить диаметры Вселенной и песчинки, а затем найти отношение объемов. Опираясь на данные астрономии своего времени и на собственные исследования, он получил примерно 10^{63} песчинок. До Архимеда не было средств ни для записи, ни для наименования такого большого числа. Архимед поступил так: все числа, меньшие мириады мириад, то есть числа от 1 до 10^8-1, объединил в первую октаду и назвал их первыми числами. Число 10^8 служит единицей второй октады, в которую входят все числа от 10^8 до $10^{2\cdot8}-1$. Это «вторые числа». Аналогично, число $10^{2\cdot8}$ является единицей третьей октады, а числа от $10^{2\cdot8}$ до $10^{3\cdot8}-1$ — «третьи». Продолжая это построение, можно дойти до мириадо-мириадной октады, содержащей числа от $10^{8\cdot(10^8-1)}$ до $10^{8\cdot10^8}-1$. Все эти числа Архимед объединил в первый период. Число $10^{8\cdot10^8}$ служит единицей первой октады второго периода и так далее. Этим способом можно дойти до последнего числа последней октады мириадо-мириадного периода. Здесь Архимед останавливается: число песчинок во Вселенной содержится уже в восьмой октаде первого периода! ■

и при счете «числами-совокупностями». Так, йорубы (одно из африканских племен), считая раковины-каури (служившие им деньгами), раскладывали их в кучки по 20 раковин в каждой, затем 20 таких кучек объединяли в одну большую кучу и так далее. При таком способе используется то, что с кучами можно поступать так же, как и с отдельными раковинами. Рассказывают, что так считали стада в Южной Африке: один из африканцев считал отдельных животных, второй — число десятков, сосчитанных первым, а третий — число десятков, сосчитанных вторым, то есть число сотен.

Следующим шагом к позиционному принципу был пропуск названий разрядов при письме (подобно тому как мы говорим «три двадцать», а не «три рубля двадцать копеек»). При записи больших чисел часто оказывался нужен символ для обозначения нуля. ■

Как появился нуль? Уже вавилоняне употребляли межразрядовый знак. Начиная со II в. до нашей эры греческие ученые знакомились с многовековыми астрономическими наблюдениями вавилонян. Вместе с их вычислительными таблицами они переняли и вавилонскую систему счисления, но числа от 1 до 59 записывали не клиньями, а в своей, алфавитной нумерации. Самое замечательное — то, что для обозначения пропущенного шестидесятеричного разряда греческие астрономы начали употреблять символ ◯ (первая буква греческого слова οὐδέν — ничто). Этот знак, видимо, и был прообразом современного нуля.

Индийцы между II и VI в. познакомились с греческой астрономией. Одновременно они должны были познакомиться с шестидесятеричной системой и греческим круглым нулем. Индийцы, наверное, и сделали завершающий шаг в создании нашей нумерации.

Индийская нумерация была занесена в Европу арабами в X—XIII вв. (отсюда и название «арабские цифры»). Преимущества позиционной системы были столь очевидны, что ее сразу стали применять итальянские купцы. Тогда же Леонардо Пизанский (Фибоначчи) выступил сторонником новой системы. В Германии, Франции, Англии до конца XIV в. новую систему почти не употребляли. И хотя вплоть до XVIII в. в официальных бумагах разрешалось применять только римские цифры, к концу XVI — началу XVII в. позиционная индийская система одержала решительную победу: ее стали применять не только купцы, но и ученые.

В России в старину использовали не римскую, а славянскую систему, у которой есть много преимуществ по сравнению с римской. Но и в России индийская система быстро вошла в употребление: во всех известных математических рукописях XVII в. применяли десятичную позиционную систему счисления. При Петре I индийские цифры вытеснили на монетах славянские, а позднее славянские цифры вообще исчезли из обихода. ■

Буквенные обозначения для неизвестных и знаки арифметических операций появились уже у греческого математика Диофанта Александрийского. Современная символика создана в XIV—XVII вв.: в конце XV в. итальянец Л. Пачоли и француз Н. Шюке для сложения и вычитания использовали знаки $\tilde{\mathrm{p}}$ (от латинского plus) и $\tilde{\mathrm{m}}$ (minus), а немецкие математики ввели современные обозначения + и −.

В XVI в. использовалась смешанная запись, содержащая и слова, и математические знаки. Так, уравнение $x^3+5x=12$ Дж. Кардано (1545) записал бы в виде

1. cubus $\tilde{\mathrm{p}}$. 5. positionibus æquantur 12

(cubus — куб, positio — неизвестная, aequantur — равно). Француз Ф. Виет

(1591) записал бы его как

$$1C+5N,\ \text{æquatur}\ 12$$

(C — cubus — куб, N — numerus — число). Но уже в 1631 г. англичанин Т. Гарриот использовал бы для записи этого уравнения вполне понятный для нас вид

$$aaa+5\cdot a=12.$$

Р. Декарт в 1637 г. придал алгебраическим выражениям полностью современный вид. Он изображал неизвестные последними буквами латинского алфавита (z, y, x, ...), а известные величины и параметры — начальными (a, b, c, ...). Постепенно принимали нынешний вид показатели степеней и корней. Современное обозначение $\sqrt{\ }$ представляет собой слитную запись модифицированной первой буквы латинского слова radix (корень) и черты, ограничивающей выражение, из которого извлекают корень. ■

Производная (derivative) обязана своим именем Г. В. Лейбницу (1667). Он предложил и обозначения dx, dy, $\frac{dx}{dy}$. Сто лет спустя Ж. Лагранж предложил удобные записи $y'=\frac{dy}{dx}$ и $dy=y'\,dx$. Термин «дифференциал» (differentia) появился в 1704 г. в универсальном словаре Дж. Харриса «Lexicon technicum».

Привычное обозначение $\frac{\partial u}{\partial x}$ для частной производной u'_x ввел А. Лежандр в 1786 г. Правда, это обозначение ему почему-то не понравилось, и его не использовали аж до 1841 г., когда его возродил К. Якоби.

Лейбниц начал употреблять в качестве знака интеграла omn. (от латинского omnia — всеобщее). Он ввел и знак $\int$ как стилизованную первую букву слова summa. Слово «интеграл» в печати впервые использовал Якоб Бернулли в 1690 г. На изобретение этого символа претендовал и Иоганн Бернулли. К слову сказать, семья Бернулли — три поколения — внесла огромный вклад в науку: математику, физику, химию. Сейчас уже никого, кроме специалистов-историков, не интересует, кто именно из них какое открытие сделал.

Запись $\int_a^b$ введена Ж. Фурье в 1822 г., а обозначение $\oint$ для интеграла по контуру ввел в 1917 г. А. Зоммерфельд.

Знак предела lim. (с точкой) предложил С. Люилье в 1786 г., а принятое теперь $\lim\limits_{x\to a}$ — заслуга Г. Харди (1908). ■

Знак	Значение	Кем введен	Когда введен (г.)
	Отношения		
$=$	равно	Р. Рекорд	1557
$<,>$	меньше, больше	Т. Гарриот	1631
$\equiv$	сравнимо	К. Гаусс	1801
$\parallel$	параллельно	У. Уотред	1677
$\perp$	перпендикулярно	П. Эригон	1634
	Величины		
∞	бесконечность	Дж. Валлис	1655
π	отношение длины окружности к диаметру	У. Джонс Л. Эйлер	1706 1736
i	корень квадратный из -1 (мнимая единица)	Л. Эйлер	1777
	Операции		
$\times$	умножить	У. Уотред	1631
$\cdot$	умножить	Г. Лейбниц	1698
:	делить	Г. Лейбниц	1684
	Функции		
a^n	степень	Р. Декарт	1637
$\sqrt{\ }$, $\sqrt[n]{\ }$	корни	Х. Рудольф А. Жирар	1525 1629
log	логарифм	И. Кеплер	1624
sin	синус	Б. Кавальери	1632
cos	косинус	Л. Эйлер	1748
tg	тангенс	Л. Эйлер	1753
arcsin	арксинус	Ж. Лагранж	1772
$\sum$	сумма	Л. Эйлер	1755
φx	функция	И. Бернулли	1718
$f(x)$	функция	Л. Эйлер	1734
!	факториал	Х. Крамп	1808

Выпросился остаться одну ночку; от одной ночки две ночки, от двух ночек две недели, от двух месяцев два года, а от двух годов жил тридцать лет.

(*Народная присказка.*) ■

Капитан Джонатан
Переплыл океан —
И в пути пеликана
Поймал капитан.

Пеликан Джонатана
Снес яйцо — и нежданно
Стало у капитана
Целых два пеликана.

И второй пеликан
Снес яйцо, как ни странно:
Стало у Джонатана
Целых три пеликана.

Будет род пеликана
прибывать беспрестанно,
Если только омлет
не спасет капитана!

Робер Деснос (*1900—1945*), *французский поэт* ■

1

— Взгляни на этого математика, — сказал логик. — Он замечает, что первые 99 чисел меньше сотни, и отсюда с помощью того, что он называет индукцией, заключает, что любые числа меньше сотни.

— Физик верит, — сказал математик, — что 60 делится на все числа. Он замечает, что 60 делится на 1, 2, 3, 4, 5 и 6. Он проверяет несколько других чисел, например, 10, 20 и 30, взятых, как он говорит, наугад. Так как 60 делится на них, он считает экспериментальные данные достаточными.

— Да, но взгляни на инженера, — возразил физик. — Он подозревает, что все нечетные числа простые. Во всяком случае, 1 можно рассматривать как простое число, доказывает он. Затем идут 3, 5 и 7 — все, несомненно, простые. Затем идет 9 — досадный случай; по-видимому, 9 не является простым числом, но 11 и 13, конечно, простые. Возвращаясь к 9, говорит он, заключаем, что 9 должно быть ошибкой эксперимента. ■

Индукция

Индукция — один из важнейших способов рассуждения, применяемых в математике. Суть этого метода в том, что для доказательства некоторого утверждения A_n, где $n = 1, 2, 3, \ldots$, сначала доказывают его для $n = 1$ (соответствующее утверждение называют базой индукции), а затем для каждого натурального n в предположении, что A_n истинно, доказывают истинность утверждения A_{n+1} (индукционный переход).

Выстроим в ряд костяшки домино (рис. 1). Толкнем первую — она, падая, повалит вторую, та — третью и так далее. Представьте, что доминошки изображают утверждения $A_1, A_2, A_3, \ldots, A_{100}, A_{101}, \ldots$, а падение доминошки означает доказательство соответствующего утверждения. Тогда «толкнуть первую доминошку» — значит доказать, что утверждение A_1 истинно; а то, что каждая доминошка, падая, валит следующую, означает, что при любом k из утверждения A_k следует A_{k+1}. Цепь доказательств, начавшись с первого утверждения, прокатится по всему ряду: она дойдет и до сотого, и до тысячного, и вообще до любого натурального числа. ■

Придумаем 10 различных натуральных чисел, сумма которых кратна каждому из них. В задаче нет параметра n, поэтому поставим более общую задачу: для любого натурального числа $n > 2$ придумать n различных натуральных чисел, сумма которых кратна каждому из них.

Для $n = 3$ годятся числа 1, 2 и 3: их сумма равна 6 и кратна любому из них. Это база индукции — утверждение, с которого начинается цепь рассуждений.

Теперь выполним индукционный переход: научимся от ряда из n чисел переходить к $n+1$ числам. Если сумма $a_1 + a_2 + \ldots + a_n$ кратна любому из слагаемых, то можно добавить к числам $a_1, a_2, \ldots, a_n$ их сумму $a_1 + a_2 + \ldots + a_n$. Таким образом из трех чисел 1, 2, 3 получим четыре числа 1, 2, 3, 6, а из них — пять чисел 1, 2, 3, 6, 12. Так можно действовать и дальше, увеличивая и увеличивая ряд чисел. В частности, для $n = 10$ получим: 1, 2, 3, 6, 12, 24, 48, 96, 192, 384. (Разумеется, нельзя утверждать, что найденный пример единственный: например, отправляясь от равенства

$$\frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} = 1,$$

находим шесть чисел 40, 30, 24, 20, 5 и 1, сумма которых равна 120 и кратна любому из них. Известным уже способом из этих шести чисел можно получить десять, добавив числа 120, 240, 480 и 960.) ■

Три пирата захватили корабль с разнообразнейшим добром. Каждый уверен, что он бы поделил добычу на равные части, но остальные ему не доверяют. Если бы пиратов было двое, то выйти из положения было бы легко: один делит добычу на две части, а другой берет ту, которая ему кажется большей. Сможете организовать раздел добычи, чтобы ни один из троих пиратов не чувствовал себя обделенным? Учтите: добыча настолько разнородна и вкусы пиратов настолько несхожи, что объективного способа сравнения отдельных частей не существует!

Пусть первый пират разделит добычу на три, по его мнению, равные части, а второй и третий укажут те части, которые им кажутся бóльшими. Если они укажут на разные части, то каждый берет ту часть, которую он считает большей, а первый берет оставшуюся — ему все равно!

Если же они укажут на одну часть, пусть поделят ее между собой. Затем второго и третьего попросим указать на ту из оставшихся частей, которая кажется большей. Если они покажут на одну и ту же часть, то вновь делят ее между собой, а первый берет оставшуюся часть. Если же они укажут на разные части, пусть каждый из них делит понравившуюся часть с первым пиратом. Все честно! ■

А если пиратов не 3, а больше? Нам поможет индукция. Предположим, что n пиратов придумали способ справедливого раздела добычи. Пусть их стало на одного больше. Разделим всю добычу между n пиратами и затем предложим каждому из них разделить свою долю на $n+1$ равных частей (по его мнению, равных). Пусть теперь $(n+1)$-й пират возьмет у каждого из них по одной части. У каждого из n пиратов останется, по его мнению, $n/(n+1)$ его прежней доли; прежняя доля составляла, по его мнению, не менее $1/n$ всей добычи; $\frac{n}{n+1}\cdot\frac{1}{n}=\frac{1}{n+1}$.

Не сможет жаловаться и $(n+1)$-й пират, так как он взял у каждого из своих товарищей не менее $1/(n+1)$ доли (по его мнению). (Вообразите, как это выглядит для 15 пиратов. Придется делить все на 14 персон, а перед этим — на 13, на 12, ..., начать же придется с удивительного для непривычных к высокой абстракции злобных пиратов раздела на двоих! Если вам дадут довести эту процедуру до конца — все (но только в самый последний момент!) станут довольны. ■

Менее опасен для вас другой способ. Усадите пиратов вокруг круглого стола и предложите первому взять долю добычи. Пусть второй, если ему кажется, что первый взял слишком много, уменьшит ее до справедливой. (Если второй считает, что первый взял не больше положенного, пусть он ничего не трогает.) Затем пусть то же сделает третий, четвертый и так далее. Возьмет эту долю (полностью выбывая из дележа) пират, последним ее коснувшийся. ■

На кольцевой дороге стоят несколько одинаковых автомашин. Если бы весь бензин из их баков слили в одну, то она смогла бы проехать по всей кольцевой дороге. Докажем, что хотя бы одна из этих машин может объехать все кольцо по часовой стрелке, забирая по пути бензин у остальных машин. База — случай 1 машины — тривиальна. Предположим, что утверждение верно для n машин. Рассмотрим случай $n+1$ машин. Хотя бы у одной машины A бензина хватит, чтобы доехать до ближайшей (по часовой стрелке) машины B: иначе суммарное количество бензина было бы явно недостаточно. Перельем весь бензин из B в A и исключим машину B из рассмотрения. Общее количество бензина не изменилось, а число машин уменьшилось. По предположению индукции, хотя бы одна машина может объехать кольцо по часовой стрелке, забирая по пути бензин у остальных машин. Эта же машина может, очевидно, объехать кольцо и в начальной ситуации. ■

Выясним, на какое наибольшее число частей могут разделить плоскость 15 прямых. Нарисовать 15 прямых нетрудно (рис. 2). А вот разобраться, сколько там частей, вовсе не легко: некоторые прямые пересекаются за пределами чертежа, а мелкие части почти не различимы. Можно, конечно, выйти на стадион с 15 веревками и провести «исследование на местности», но вдруг веревки запутаются? И куда мы пойдем, если вместо 15 прямых будет 150?

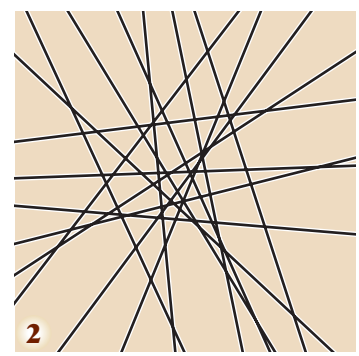
Огастес де Морган (1806—1871), сын полковника английских войск в Индии, первый президент Лондонского королевского математического общества (основанного в 1865 г.), один из создателей математической логики. Его имя носят формулы $\overline{A\cup B}=\overline{A}\cap\overline{B}$ и $\overline{A\cap B}=\overline{A}\cup\overline{B}$, где черта обозначает переход от множества к его дополнению. Ввел термин «математическая индукция» в 1838 г. ■

Докажите по индукции следующие утверждения.

1) Двузначное число 12 делится на 4, трехзначное число 112 — на 8, четырехзначное число 2112 — на 16. Вообще, для любого натурального числа n существует составленное из цифр 1 и 2 число, делящееся на 2^n.

2) Любую дробь m/n, где m, n — натуральные числа, $1<m<n$, можно представить в виде суммы нескольких дробей вида $1/q$, таких, что знаменатель каждой следующей дроби делится на знаменатель предыдущей. (Например, $\frac{3}{43}=\frac{1}{15}+\frac{1}{330}+\frac{1}{14\,190}$.)

3) Квадрат нельзя разрезать на n квадратов тогда и только тогда, когда $n=2$, 3 или 5. ■



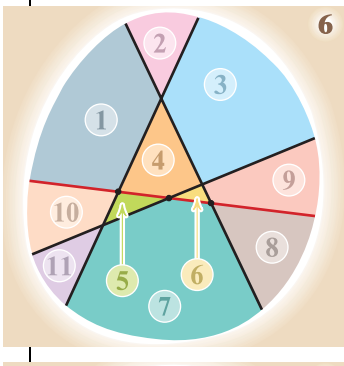
Лучше решим задачу Я. Штейнера об n прямых — найдем формулу для количества $f(n)$ частей, на которые разбивают плоскость n прямых общего положения (прямые общего положения — это прямые, никакие две из которых не параллельны и никакие три не проходят через одну точку). Для $n=1$, 2, и 3 ответы очевидны (рис. 3—5): $f(1)=2$, $f(2)=4$ и $f(3)=7$. Чуть сложнее увидеть, что $f(4)=11$ (рис. 6) и $f(5)=16$ (рис. 7). Запишем найденные значения в таблицу:

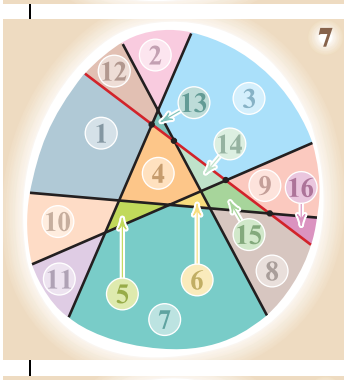
n	1	2	3	4	5
$f(n)$	2	4	7	11	16

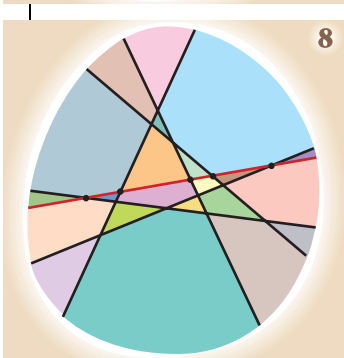
Видите закономерность? $4=2+2$, $7=3+4$, $11=4+7$, $16=5+11$ — числа правого столбца равны сумме соседей слева и сверху. Формулой это можно записать так:

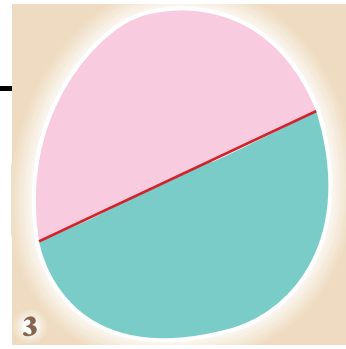
$$f(n)=n+f(n-1). \qquad (*)$$

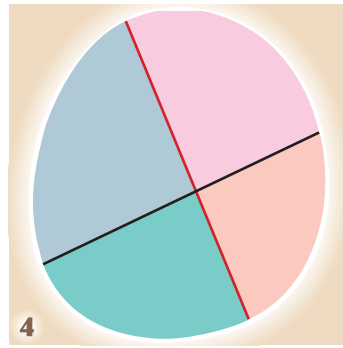
Чтобы доказать эту формулу, посмотрим, что происходит, когда к $n-1$ прямым добавляем еще одну. Например, на рисунке 8 к пяти прямым добавлена шестая (красная) прямая. Она пересекает не все части, на которые пять прямых разрезают плоскость, а только шесть частей. Это не случайно: поскольку шестая прямая пересекает каждую из других прямых, всего возникает пять точек пересечения, которые делят прямую на шесть частей: два луча и четыре отрезка. Каждый из отрезков и лучей, на которые шестую прямую делят точки ее пересечения с пятью другими, является следом от разрезания некоторой части на две. Значит, $f(6)=f(5)+6$. Аналогично обстоит дело и в общем случае: при добавлении n-й прямой к $n-1$ прямым количество частей увеличивается на n (как вы помните, среди прямых не должно быть параллельных и никакие три не должны проходить через одну точку). Формула $(*)$ доказана. Теперь легко найти $f(15)=121$, продолжив таблицу. ■

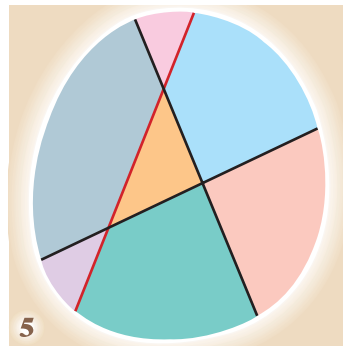












$f(150)=11\,326$ можно вычислить так же, но нельзя ли побыстрее? Можно! Кроме рекуррентной (выражающей следующее значение через предыдущие) формулы $(*)$, есть явная формула

$$f(n)=\frac{n(n+1)}{2}+1. \qquad (**)$$

Докажем ее по индукции.

Три дамы A, B и C сидят с испачканными лицами и смеются. Внезапно A соображает: «Почему B не понимает, что C смеется над ней? О, боже! Они смеются надо мной». ■

Ехали в вагоне поезда мудрецы. Поезд то и дело нырял в тоннели. Собрались все в коридоре, в открытые окна глядят — не наглядятся. Вдруг грохот, дым, пыль! Проехали тоннель — входит проводник. «Тут кое-кто испачкался, — говорит. — В поезде воды нет. Но сейчас остановки пойдут, можно будет выйти помыться».

А мудрецы в вагоне собрались как на подбор: столь же умные, сколь ленивые. Никто зря мыться не пойдет, если не знает наверняка, что испачкался. И у соседей не спросит, чистое у него лицо или грязное — зачем напрасно людей тревожить и самому беспокоиться? — проще сообразить.

Как поступят мудрецы? Оказывается, если у n из них испачканы лица, то на n-й остановке все эти n мудрецов выйдут из поезда мыться!

Докажем утверждение по индукции. При $n=1$ утверждение очевидно. Докажем, что если оно верно для некоторого n, то оно верно и для $n+1$. Пусть лица испачканы у $n+1$ мудрецов. Тогда один из них видит вокруг n грязных лиц и рассуждает так: «Мое лицо или чистое, или грязное. В первом случае все n мудрецов с грязными лицами выйдут умываться на n-й остановке. Поскольку первый случай возможен, мне не следует выходить ни на n-й остановке, ни раньше: если я чистый, это было бы непростительной тратой сил!

Если же мое лицо грязное, то каждый из n других перемазавшихся мудрецов видит n испачканных лиц. Никто из них не пойдет умываться на n-й остановке.

Итак, если они не пойдут мыться на n-й остановке, то я — грязный. Дождусь n-й остановки. Если на ней никто не пойдет умываться, выйду на следующей — $(n+1)$-й остановке!»

Так же рассуждают все мудрецы с грязными лицами. Следовательно, на $(n+1)$-й остановке все они пойдут умываться, что и требовалось доказать.

Изменим слегка наш рассказ. Зная, что в вагоне едут мудрецы, и увидав, что многие из них ис-

пачкались, проводник решает сократить свое объявление. «Зачем говорить, что кое-кто испачкался, — думает он, — когда они сами это видят!» И пропускает первую фразу объявления. Можно ли по-прежнему утверждать, что если лица испачканы у n человек, то на n-й остановке они пойдут мыться?

Или так: пусть мудрецы и без проводника знают, что в поезде нет воды и будут стоянки, на которых можно умыться. Кроме того, пусть испачкался больше чем один человек. Тогда в объявлении проводника нет как будто ничего нового ни для кого из мудрецов! Что же — если бы проводник не приходил — пошли бы n испачкавшихся мыться на n-й остановке? Велик соблазн ответить утвердительно: раз ничего не изменилось в условии задачи, то не должен измениться и ответ! И все же здравый смысл подсказывает, что без объявления проводника мыться, пожалуй, никто не пойдет! И потом — что значит: «ничего не изменилось в условии»? В одном варианте проводник вообще не приходил, а в другом — приходил и что-то сказал! Что же он сказал? Одними восклицаниями этот вопрос, по-видимому, не решишь.

Рассмотрим простейший случай: $n=2$, лица испачканы у двоих, они видят друг друга. Представьте себя на месте одного из них — и увидите проблему: без объявления проводника испачканный мудрец не знает, знает ли другой испачканный, что в вагоне не все чистые. Мы нащупали разгадку: сказанное проводником действительно не было новостью для мудрецов, но каждый видел, что все слушали объявление и поэтому не только любой мудрец знает, что кто-то испачкался, но и любой мудрец знает, что любой мудрец знает, что кто-то испачкался.

Та же ситуация и для n мудрецов, только слова «любой мудрец знает, что» надо повторить не дважды, а n раз. Сравните: «Иль думал, что я думала, что думал он: я сплю» (С. Маршак). ■

Смешливая дама догадалась, что у нее грязное лицо, хотя никакой проводник не делал никаких объявлений. Роль проводника сыграл несдержанный смех ее соседок. Степенные мудрецы, конечно, не позволили бы себе ничего подобного! ■

База очевидна: $f(1)=2=\dfrac{1\cdot(1+1)}{2}+1$.

Переход состоит в том, что если $f(n-1)=\dfrac{(n-1)n}{2}+1$, то

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

Заметьте, как легко! Правда, возникает вопрос: как можно было догадаться, что для $f(n)$ выполнена именно формула $(**)$? ■

Сумма $S(n)=1+2+3+\ldots+n$ первых n натуральных чисел встречалась всякому, кто интересуется математикой. Эта сумма удовлетворяет соотношению

$$S(n)-S(n-1)=n.$$

Для $S(n)$ есть явная формула:

$$S(n)=\frac{n(n+1)}{2}, \qquad (***)$$

вывести которую можно при помощи индукции, а можно и без индукции, записав сумму $S(n)$ два раза: сначала расположив слагаемые по возрастанию, а потом в обратном порядке, по убыванию:

$$S(n)=1+\ \ 2\ \ +\ \ 3\ \ +\ldots+(n-2)+(n-1)+n,$$
$$S(n)=n+(n-1)+(n-2)+\ldots+\ \ 3\ \ +\ \ 2\ \ +1,$$

и сложив эти равенства почленно:

$$2S(n)=(n+1)+(n+1)+(n+1)+\ldots+(n+1)+(n+1)+(n+1)=n(n+1).$$

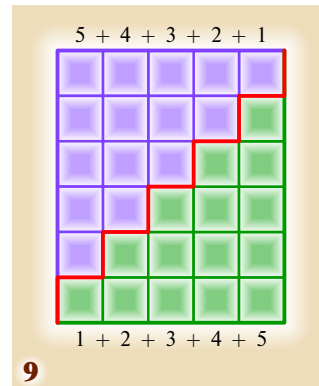
Разделив на 2, получим формулу для $S(n)$. (На рисунке 9 это рассуждение проиллюстрировано геометрически: для $n=5$ показано, как прямоугольник размером $n\times(n+1)$ можно разбить на две равные фигуры, каждая из которых состоит из $1+2+3+\ldots+n$ клеток.) ■

В следующем примере главное — догадаться, какую формулу надо доказывать по индукции. Давайте найдем явную формулу для суммы кубов первых n натуральных чисел. Начинаем:

$$1^3=1,$$
$$1^3+2^3=9,$$
$$1^3+2^3+3^3=36,$$
$$1^3+2^3+3^3+4^3=100,$$
$$1^3+2^3+3^3+4^3+5^3=225,$$
$$1^3+2^3+3^3+4^3+5^3+6^3=441.$$

Что же это за числа: 1, 9, 36, 100, 225, 441? Это квадраты! И не просто квадраты, а квадраты сумм первых n натуральных чисел:

$$1=1^2,$$
$$9=(1+2)^2,$$
$$36=(1+2+3)^2,$$
$$100=(1+2+3+4)^2,$$
$$225=(1+2+3+4+5)^2,$$
$$441=(1+2+3+4+5+6)^2.$$



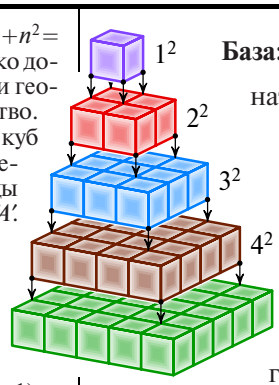
9

Возникает гипотеза: сумма кубов первых n натуральных чисел равна квадрату суммы этих чисел. Вспомнив формулу $(***)$, мы можем сформулировать гипотезу в более удобном для применения индукции виде:

$$1^3+2^3+3^3+\ldots+n^3=\left(\frac{n(n+1)}{2}\right)^2.$$

Докажем ее.

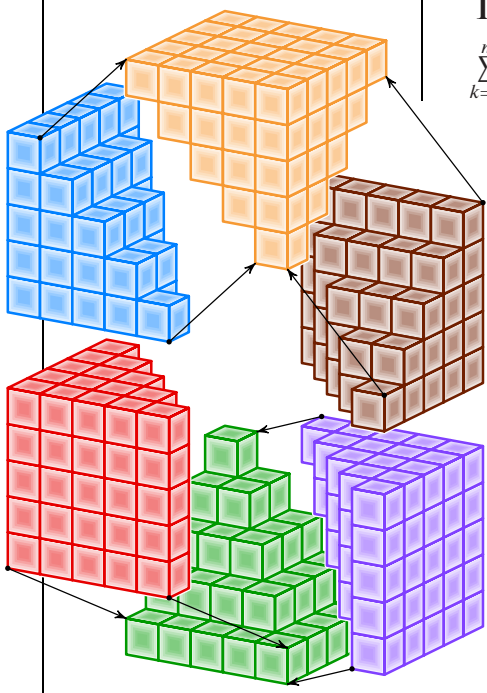
Формулу $1^2+2^2+3^2+\ldots+n^2=n(n+1)(2n+1)/6$ легко доказать по индукции. Есть и геометрическое доказательство. Воспользуемся тем, что куб $ABCDA'B'C'D'$ можно разрезать на три равные пирамиды $ABCC'B'$, $ACDD'C'$ и $AB'C'D'A'$. Очевидно, что из $1^2+2^2+3^2+\ldots+n^2$ кубиков можно сложить ступенчатую пирамиду, из трех таких пирамид — почти куб, из двух почти кубов — параллелепипед $n\times(n+1)\times(2n+1)$. ■



База: $1^3=\left(\frac{1(1+1)}{2}\right)^2$. **Переход:** предположим, что для некоторого натурального числа n формула верна. Тогда

$$(1^3+2^3+3^3+\ldots+n^3)+(n+1)^3=\left(\frac{n(n+1)}{2}\right)^2+(n+1)^3=$$
$$=\frac{(n+1)^2}{4}(n^2+4n+4)=\frac{(n+1)^2(n+2)^2}{4}=\left(\frac{(n+1)(n+2)}{2}\right)^2.$$

Понимаете, что произошло? Сумму кубов первых $n+1$ натуральных чисел мы представили как сумму суммы кубов n чисел и числа $(n+1)^3$. А дальше — всего лишь алгебраические преобразования! ■



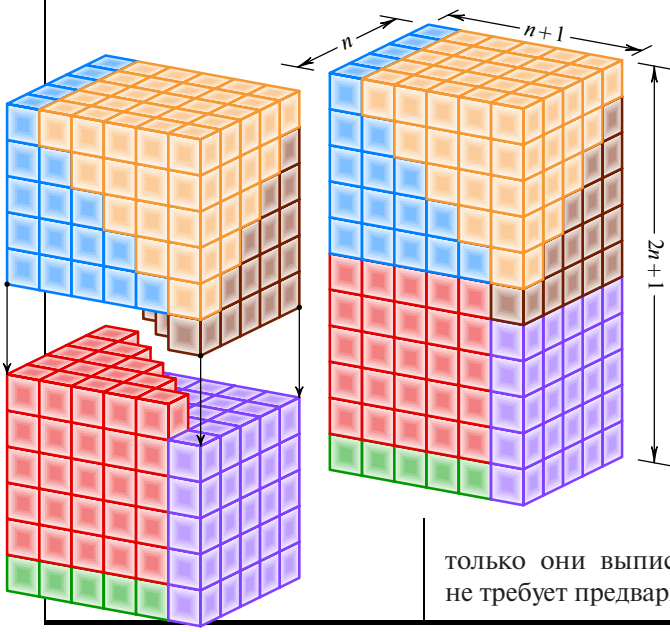
Тождество $3k(k+1)=k(k+1)(k+2)-(k-1)k(k+1)$ позволяет найти

$$\sum_{k=1}^{n}k(k+1)=\frac{1}{3}(1\cdot2\cdot3-0\cdot1\cdot2+2\cdot3\cdot4-1\cdot2\cdot3+3\cdot4\cdot5-2\cdot3\cdot4+\ldots+(n-1)n(n+1)-$$
$$-(n-2)(n-1)n+n(n+1)(n+2)-(n-1)n(n+1))=\frac{n(n+1)(n+2)}{3}.$$ ■

Чему равна сумма четвертых степеней? А сумма пятых степеней? Помогают следующие легко запоминаемые формулы: $\sum\limits_{k=1}^{n}k=\frac{n(n+1)}{2}$, $\sum\limits_{k=1}^{n}k(k+1)=\frac{n(n+1)(n+2)}{3}$, $\sum\limits_{k=1}^{n}k(k+1)(k+2)=\frac{n(n+1)(n+2)(n+3)}{4}$, и вообще, $\sum\limits_{k=1}^{n}k^{\underline{s}}=\frac{n^{\underline{s+1}}}{s+1}$, где $k^{\underline{s}}$ — это произведение $k\cdot(k+1)\cdot\ldots\cdot(k+s-1)$.

Теперь сумму четвертых степеней вычислить несложно: поскольку $k^4=k(k+1)(k+2)(k+3)-6k^3-11k^2-6k$, имеем:

$$\sum_{k=1}^{n}k^4=\sum_{k=1}^{n}k(k+1)(k+2)(k+3)-6\sum_{k=1}^{n}k^3-11\sum_{k=1}^{n}k^2-6\sum_{k=1}^{n}k=$$
$$=\frac{n(n+1)(n+2)(n+3)(n+4)}{5}-6\cdot\frac{n^2(n+1)^2}{4}-11\cdot\frac{n(n+1)(2n+1)}{6}-6\cdot\frac{n(n+1)}{2}.$$ ■



Воспользовавшись формулой $\frac{1}{k(k+1)}=\frac{1}{k}-\frac{1}{k+1}$, получаем

$$\sum_{k=1}^{n}\frac{1}{k(k+1)}=\frac{1}{1\cdot2}+\frac{1}{2\cdot3}+\frac{1}{3\cdot4}+\ldots+\frac{1}{(n-1)n}+\frac{1}{n(n+1)}=$$
$$=1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}+\ldots+\frac{1}{n-1}-\frac{1}{n}+\frac{1}{n}-\frac{1}{n+1}=$$
$$=1-\frac{1}{n+1}=\frac{n}{n+1}.$$

А равенство $\frac{1}{k(k+1)(k+1)}=\frac{1}{2}\left(\frac{1}{k(k+1)}-\frac{1}{(k+1)(k+2)}\right)$ помогает найти сумму $\sum\limits_{k=1}^{n}\frac{1}{k(k+1)(k+2)}$. Разумеется, формулы $\sum\limits_{k=1}^{n}\frac{1}{k(k+1)}=\frac{n}{n+1}$ и $\sum\limits_{k=1}^{n}\frac{1}{k(k+1)(k+2)}=\frac{1}{2}\times$ $\times\left(\frac{1}{2}-\frac{1}{(n+1)(n+2)}\right)$ легко доказать по индукции, как только они выписаны. Изложенный выше способ замечателен тем, что не требует предварительного знания ответа. ■

Одна из интересных особенностей индукции состоит в том, что более точное утверждение иногда легче доказать, чем более слабое. Например, пусть $P_n = \frac{1\cdot3\cdot5\cdot\ldots\cdot(2n-1)}{2\cdot4\cdot6\cdot\ldots\cdot(2n)}$. Докажем неравенство $P_n < \frac{1}{\sqrt{n}}$.

База: $P_1 = \frac{1}{2} < \frac{1}{\sqrt{1}}$. Теперь попытаемся провести **индукционный переход,** то есть из неравенства $P_n < \frac{1}{\sqrt{n}}$ попробуем вывести неравенство $P_{n+1} < \frac{1}{\sqrt{n+1}}$. Очевидно,

$$P_{n+1} = P_n \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{n}} \cdot \frac{2n+1}{2n+2}.$$

Если бы правая часть не превосходила $\frac{1}{\sqrt{n+1}}$, то индукционный переход состоялся бы. Но увы: возводя интересующие нас выражения в квадрат, получаем: $\frac{(2n+1)^2}{4n(n+1)^2}$ и $\frac{1}{n+1}$. Приведя к общему знаменателю, приходим к сравнению чисел $4n^2+4n+1$ и $4n(n+1)$. Очевидно, первое больше, а нам хотелось обратного!

Переход не состоялся. Это не значит, что неравенство $P_n < \frac{1}{\sqrt{n}}$ неверно. Не годится лишь наш метод доказательства. Докажем более точное неравенство $P_n < \frac{1}{\sqrt{n+1}}$.

База индукции: $P_1 = \frac{1}{2} < \frac{1}{\sqrt{2}}$. **Переход:**

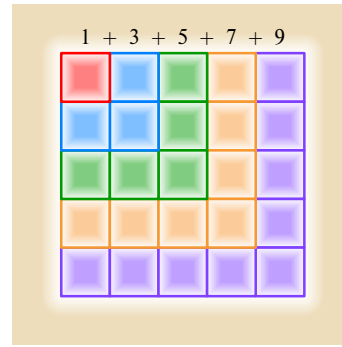
$$P_{n+1} = P_n \cdot \frac{2n+1}{2n+2} < \frac{1}{\sqrt{n+1}} \cdot \frac{2n+1}{2n+2}.$$

Неравенство $\frac{2n+1}{2\sqrt{n+1}(n+1)} < \frac{1}{\sqrt{n+2}}$ после возведения обеих частей в квадрат и перехода к общему знаменателю превращается в неравенство $(n+2)\times$ $\times(4n^2+4n+1) < 4(n+1)^3$. Раскроем скобки: $4n^3+8n^2+4n^2+8n+n+2 <$ $< 4n^3+12n^2+12n+4$, то есть $0 < 3n+2$. Получилось! Причина в том, что в доказательстве более сильного утверждения мы опирались на более сильное предположение индукции. На первый взгляд это удивительно, но так часто бывает: чем точнее утверждение, тем легче его доказать! ■

Довольно точные оценки величины P_n можно получить и без индукции. А именно, $P_n^2 = \frac{1\cdot3}{2^2}\cdot\frac{3\cdot5}{4^2}\cdot\ldots\cdot\frac{(2n-1)\cdot(2n+1)}{(2n)^2}\cdot\frac{1}{2n+1}$. Оставляя первый множитель $\frac{3}{4}$ без изменения, заменяя последний множитель $\frac{1}{2n+1}$ на большее число $\frac{1}{2n}$ и заменяя все другие множители на 1, в силу неравенства $(2k-1)(2k+1) < (2k)^2$ получаем неравенство $P_n^2 < \frac{3}{4}\cdot\frac{1}{2n}$, откуда $P_n < \sqrt{\frac{3}{8n}}$. Аналогично можно получить оценку снизу:

$$P_n^2 = \frac{1}{2}\cdot\frac{3^2}{2\cdot4}\cdot\frac{5^2}{4\cdot6}\cdot\ldots\cdot\frac{(2n-1)^2}{(2n-2)(2n)}\cdot\frac{1}{2n} > \frac{1}{2}\cdot\frac{1}{2n},$$

откуда $P_n > \frac{1}{2\sqrt{n}}$. В статье «Решето Иосифа Флавия» доказана формула Валлиса: $\lim\limits_{n\to\infty} P_n\sqrt{n} = \frac{1}{\sqrt{\pi}} = 0{,}5642\ldots$ ■



Сумма первых n нечетных чисел равна n^2. На рисунке это равенство проиллюстрировано для $n=5$. ■

Каждая из сумм 1, $3+5$, $7+9+11$, $13+15+17+19$, ... равна кубу числа слагаемых. ■

Число $n!$ при любом натуральном n представимо в виде произведения двух натуральных чисел, различающихся не более чем вдвое. **База** — два равенства: $1! = 1\cdot1$ и $2! = 1\cdot2$. **Индукционный переход** проведем не от n к $n+1$, а от n к $n+2$. Пусть $n! = ab$, где a, b — натуральные числа, $1 \leqslant a \leqslant b \leqslant 2a$. Тогда $(n+2)! = a\times$ $\times(n+2)\cdot b(n+1)$, причем

$$\frac{a}{b}\cdot\frac{n+2}{n+1} < 1\cdot2 = 2$$

и $\frac{b}{a}\cdot\frac{n+1}{n+2} < 2\cdot1 = 2$. ■

Через любые n точек можно провести прямую линию.

Доказательство. Применим индукцию. При $n=1$ и $n=2$ все правильно.

Осталось доказать это для бóльших значений n. Допустим, что утверждение верно при некотором $n=k$, и покажем, что в этом случае оно сохранит силу и при $n=k+1$. Итак, пусть произвольно заданы $k+1$ точек M_1, M_2, ..., M_k, M_{k+1}. В силу предположения индукции, через k точек M_1, M_2, ..., M_k проходит некоторая прямая l. В силу того же предположения через k точек $M_2, \ldots, M_k, M_{k+1}$ также проходит некоторая прямая l'.

Эти две прямые имеют по крайней мере две общие точки M_2 и M_k. Но две точки определяют единственную прямую. Поэтому прямые l и l' совпадают. Следовательно, прямая l, проходящая через точки M_1, M_2, ..., M_k, проходит и через точку M_{k+1}. ■

Алгоритм Евклида и линейные диофантовы уравнения

Алгоритм Евклида — это способ отыскания наибольшего общего делителя натуральных чисел. Он тесно связан с алгоритмом разложения чисел в цепные дроби и с решением линейных уравнений в целых числах.

Основная идея алгоритма Евклида — формула

$$\text{НОД}(a, b) = \text{НОД}(a - bq, b),$$

которая верна для любых целых чисел a, b, q. (НОД — первые буквы слов «наибольший общий делитель».) Докажем ее. С одной стороны, всякий общий делитель d чисел a и b очевидным образом является и делителем числа $a - bq$; с другой стороны, всякий общий делитель чисел $a - bq$ и b является и делителем числа $a = (a - bq) + bq$. Поэтому множество общих делителей чисел a и b совпадает с множеством общих делителей чисел $a - bq$ и b. А если совпадают множества, то совпадают и их наибольшие элементы. Равенство доказано. ■

Вычислим НОД(6059; 663). Для этого разделим 6069 на 663 с остатком:

$$6069 = 663 \cdot 9 + 102.$$

Следовательно, НОД(6059; 663) = НОД(663; 102). Поскольку $663 = 102 \cdot 6 + 51$, то НОД(663; 102) = НОД(51; 102). Поскольку 102 делится на 51, получаем ответ: НОД(6059; 663) = НОД(663; 102) = НОД(51; 102) = 51.

Часто для краткости опускают буквы НОД, считая, что круглые скобки обозначают наибольший общий делитель чисел (не обязательно двух; в скобки можно заключить любое множество чисел). Наименьшее общее кратное НОК[a, b] тоже часто пишут без букв НОК.

Итак, алгоритм Евклида — это довольно быстро работающий метод нахождения наибольшего общего делителя двух чисел. Если даны два числа a и b, причем $a > b > 0$, то сначала делим a на b:

$$a = bq_1 + r_1,$$

где q_1 — неполное частное (которое не используется в дальнейших вычислениях), r_1 — остаток, $r_1 < b$. Затем делим число b на r_1 и находим неполное частное q_2 и (только и интересующий нас) остаток r_2. Далее, делим число r_1 на r_2, при этом получаем остаток r_3, меньший, чем r_2, и так далее, пока какое-нибудь число r_{n-1} не разделится на r_n нацело, без остатка (то есть $r_{n+1} = 0$). Последний ненулевой остаток r_n и есть искомый наибольший делитель чисел a и b:

$$\text{НОД}(a; b) = (b; r_1) = (r_1; r_2) = \ldots = (r_{n-2}; r_{n-1}) = (r_{n-1}; r_n) = (r_n; 0) = r_n. \blacksquare$$

Линейные уравнения. Примеры. Алгоритм Евклида тесно связан с решением уравнений вида

$$ax - by = c$$

в целых числах, где a, b и c — данные целые числа, x и y — неизвестные.

```
_20052005 | 43
 172      |466325
 ---
 _285
  258
  ---
 _272
  258
  ---
 _140
  129
  ---
 _110
   86
   ---
 _245
  215
  ---
   30
```

Деление с остатком. При делении «уголком» числа $a = 20\,052\,005$ на $b = 43$ получаем частное 466 325 и остаток 30. Это можно записать формулой

$$20\,052\,005 = 43 \cdot 466\,325 + 30.$$

Если $a = bq + r$, где q, r — целые числа, причем $0 \leqslant r < b$, то число q называют неполным частным, а r — остатком.

Доказать возможность деления с остатком несложно. Достаточно заметить, что любое число a либо само есть кратное числа b, либо лежит между двумя последовательными кратными числа b, то есть

$$bq < a < b(q + 1).$$

В первом случае $r = 0$. Во втором случае

$$0 < a - bq < b,$$

так что число $r = a - bq$ удовлетворяет условиям $0 < r < b$.

Деление с остатком не только возможно, но и производится единственным способом: если

$$a = bq_1 + r_1 = bq_2 + r_2,$$

где q_1, q_2, r_1 и r_2 — целые числа, причем $0 \leqslant r_1 < b$, $0 \leqslant r_2 < b$, то $q_1 = q_2$ и $r_1 = r_2$. В самом деле, имеем

$$b(q_1 - q_2) = r_2 - r_1.$$

Очевидно, $-b < r_1 - r_2 < b$. Единственным числом, которое больше числа $-b$, меньше числа b и нацело делится на b (а делится потому, что оно равно произведению числа b на $q_1 - q_2$), является число 0. Значит, $r_2 - r_1 = 0$ и $q_1 = q_2$. ■

Особенно просто вычислять НОД, пользуясь двоичной системой счисления. Если двоичные записи двух чисел оканчиваются нулями, то они оба четные и мы используем формулу

$$\text{НОД}(2a, 2b) = 2\,\text{НОД}(a, b).$$

Если одно число четное, а другое нечетное, используем формулу

$$\text{НОД}(2a, 2b+1) = \text{НОД}(a, 2b+1).$$

А если оба нечетные, используем формулу

$$\text{НОД}(2a+1, 2b+1) = \text{НОД}(2a-2b, 2b+1) = \text{НОД}(a-b, 2b+1).$$

Вычислим, например, НОД чисел 5622 и 1858. В двоичной системе эти числа выглядят так: $\overline{1010111110110}$ и $\overline{11101000010}$. Имеем:

$$\begin{aligned}
(5622,\ 1858) &= (\overline{1010111110110},\ \overline{11101000010}) \\
2(2811,\ 929) &= 2(\overline{101011111011},\ \overline{1110100001}) \\
2(1882,\ 929) &= 2(\overline{11101011010},\ \overline{1110100001}) \\
2(941,\ 929) &= 2(\overline{1110101101},\ \overline{1110100001}) \\
2(12,\ 929) &= 2(\overline{1100},\ \overline{1110100001}) \\
2(3,\ 929) &= 2(\overline{11},\ \overline{1110100001}) \\
2(3,\ 926) &= 2(\overline{11},\ \overline{1110011110}) \\
2(3,\ 463) &= 2(\overline{11},\ \overline{111001111}) \\
2(3,\ 460) &= 2(\overline{11},\ \overline{111001100}) \\
2(3,\ 115) &= 2(\overline{11},\ \overline{1110011}) \\
2(3,\ 112) &= 2(\overline{11},\ \overline{1110000}) \\
2(3,\ 7) &= 2(\overline{11},\ \overline{111}) \\
2(3,\ 4) &= 2(\overline{11},\ \overline{100}) \\
2(3,\ 1) &= 2(\overline{11},\ \overline{1})
\end{aligned}$$

Значит, $\text{НОД}(5622; 1858) = 2 \times \text{НОД}(3; 1) = 2$. ■

Пусть a и b — длины двух отрезков, $a > b$. Отложим b на a столько раз, сколько возможно; получим остаток r_1. Отложим r_1 на b сколько возможно раз, получим остаток r_2 и так далее. Если, откладывая очередной отрезок длины r_n, мы не получим остатка (то есть $r_{n+1} = 0$), то отрезок длины r_n и есть наибольшая общая мера отрезков длин a и b.

Если числа a и b целые, то все остатки $r_1, r_2, \ldots$ целые неотрицательные. В силу неравенств

$$b > r_1 > r_2 > \ldots$$

процесс рано или поздно закончится. Последнее ненулевое число r_n и есть $\text{НОД}(a; b)$. ■

Не любое уравнение такого вида имеет решения в целых числах. Например, равенство

$$2x - 246y = 345$$

для целых чисел x и y невозможно, поскольку левая часть делится на 2, а правая — не делится.

Рассмотрим уравнение

$$69x - 91y = 1996.$$

Скорее всего, решения у него есть, но как их найти? Можно пытаться угадать пару чисел $(x; y)$, но вдруг не повезет? Быстрый и удобный способ дает алгоритм Евклида. Перепишем уравнение в виде

$$69x - 69y - 22y = 1996.$$

Обозначив $z = x - y$, получим

$$69z - 22y = 1996.$$

Один из коэффициентов полученного уравнения (69) остался от исходного уравнения, а другой (22) меньше, чем коэффициент исходного (91). Причина в том, что 22 — остаток от деления 91 на 69. Продолжим:

$$3z + 66z - 22y = 1996.$$

Обозначив $t = 3z - y$, придем к уравнению

$$3z + 22t = 1996.$$

Его решение легко угадать: $t = 1$, $z = (1996 - 22)/3 = 658$. Теперь легко вернуться к искомым x, y — достаточно посчитать сначала $y = 3z - t = 3 \cdot 658 - 1 = 1971$, потом — $x = z + y = 1971 + 658 = 2629$.

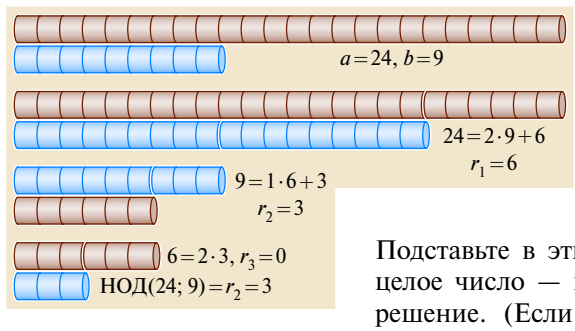
Впрочем, можно было обойтись и совсем без угадывания, переписав уравнение в виде

$$3z + 21t + t = 1996,$$

обозначив $m = z + 7t$ и, таким образом, сведя дело к уравнению

$$3m + t = 1996.$$

Если вместо m подставить любое целое число, то получим (тоже целое!) $t = 1996 - 3m$. Это позволяет нам найти $z = m - 7t = m - 7(1996 - 3m) = 22m - 13\,972$. Далее, $y = 3z - t = 3(22m - 13\,972) - (1996 - 3m) = 69m - 43\,912$. Наконец, $x = z + y = (22m - 13\,972) + (69m - 43\,912) = 91m - 57\,884$.



Мы нашли общее решение уравнения в целых числах:

$$\begin{cases} x = 91m - 57\,884, \\ y = 69m - 43\,912. \end{cases}$$

Подставьте в эти формулы вместо m любое целое число — получите некоторое частное решение. (Если боитесь, что мы допустили арифметическую ошибку, проверьте тождество $69 \cdot (91m - 57\,884) - 91 \times (69m - 43\,912) = 1996$.) В найденном виде представимо любое решение $(x; y)$ интересующего нас уравнения. ■

Точки с целыми координатами на прямой. Прямая на плоскости задается, как известно, уравнением первой степени. Любое такое уравнение можно привести к виду $ax-by=c$. Рассмотрим примеры.

$2x=5$. Уравнение задает прямую, параллельную оси ординат (вертикальную прямую). Точек с целыми координатами на этой прямой нет, поскольку число $5/2$ — не целое.

$y=x$. Это — биссектриса первого («северо-восточного») и третьего («юго-западного») квадрантов. Точек с целыми координатами бесконечно много. Они расположены на равных расстояниях друг от друга. Каждому целому x соответствует целое $y=x$.

$4x=-5y$. Прямая проходит через начало координат. Числа 4 и -5 взаимно просты. Для того чтобы $-5y$ делилось на 4, необходимо и достаточно, чтобы y делилось на 4, то есть $y=4t$, где $t\in\mathbb{Z}$. Подставляя в уравнение, получаем $4x=-5\cdot 4t$, то есть $x=-5t$. Значит, целочисленных точек бесконечно много. Они расположены на равных расстояниях и описываются формулой $(x;\,y)=(-5t;\,4t)$.

$2y-3x=6$. Выразив y через x, получим: $y=\frac{3}{2}x+3$. Если x нечетно, то y — не целое. Если же $x=2t$, то $y=3t+3$. Следовательно, на исследуемой прямой лежит бесконечно много точек с целыми координатами:

$$\begin{cases} x=2t,\\ y=3t+3, \end{cases}$$

где t — любое целое число.

$5x-7y=2$. Легко подобрать пару целых чисел $(x;\,y)=(-1;\,-1)$. Немного подумав, найдем еще одну пару: $x=6$, $y=4$. Вообще, если увеличить x на 7, а y на 5, то выражение $5x-7y$ не изменит своего значения:

$$5(x+7)-7(y+5)=5x+5\cdot 7-7y-7\cdot 5=5x-7y.$$

Поэтому, зная одну пару целых чисел $(x,\,y)$, удовлетворяющих уравнению $5x-7y=2$, мы можем указать бесконечно много других пар:

$$\begin{cases} x=7t-1,\\ y=5t-1. \end{cases}$$

Проверка того, что все эти пары удовлетворяют уравнению, не составляет труда:

$$5(7t-1)-7(5t-1)=35t-5-35t+7=2.$$

Никаких других целочисленных решений исследуемое уравнение не имеет. Доказывать это можно разными способами. Например, запишем

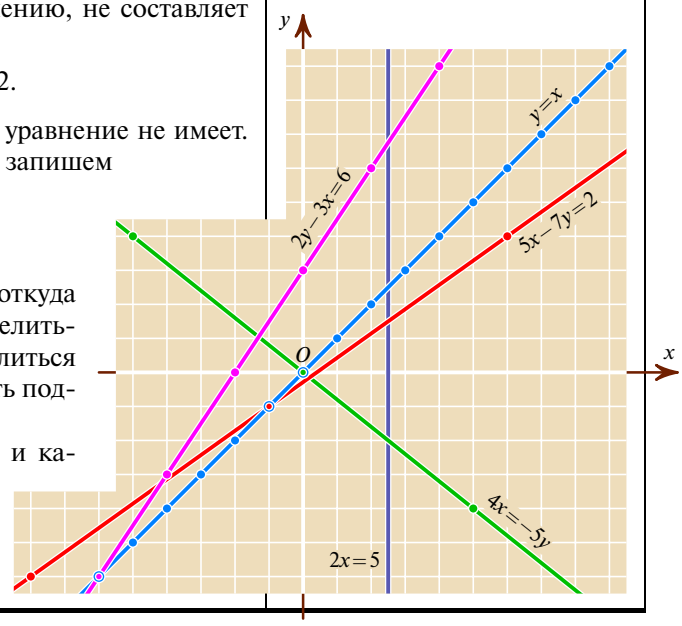
$$\begin{cases} 5x-7y=2,\\ 5\cdot(-1)-7\cdot(-1)=2 \end{cases}$$

и приравняем левые части: $5x-7y=5\cdot(-1)-7\cdot(-1)$, откуда $5(x+1)=7(y+1)$. Теперь ясно, что левая часть должна делиться на 7. Следовательно, число $x+1$ должно нацело делиться на 7, то есть $x+1=7t$, где t — целое. Осталось выполнить подстановку $5\cdot 7t=7(y+1)$ и получить $y=5t-1$.

Разумеется, можно было решать уравнение $5x-7y=2$ и каким-нибудь другим способом. Рассмотренный нами способ интересен тем, что при помощи него можно решить уравнение $ax-by=c$ не только при $(a;\,b;\,c)=(5;\,7;\,2)$, но и в общем случае. ■

Мухаммед ал-Хорезми (780—847) родился и получил образование в Хорезме, а затем переселился в Багдад — крупнейший в то время центр интеллектуальной жизни Средней Азии. Среди предков ал-Хорезми были зороастрийские жрецы. Под его руководством ученые «Дома мудрости» — сегодня мы назвали бы его академией наук — переводили труды Платона, Аристотеля, Евклида, Птолемея, Гиппократа, проводили астрономические наблюдения, проверяли и уточняли данные, приведенные в греческих и индийских сочинениях. По инициативе ал-Хорезми проводились геодезические работы по измерению длины градуса земного меридиана.

Ал-Хорезми написал обширную географическую «Книгу картин Земли», а также книги «О построении астролябии» и «О солнечных часах». Наибольшую славу ему принесли сочинения по арифметике и алгебре. В трактате «Об индийском счете» ал-Хорезми впервые на арабском Востоке ввел десятичную позиционную систему счисления: девять индийских цифр и «маленький кружок, показывающий, что разряд пуст». Ал-Хорезми подробно объяснил, как складывать, вычитать, умножать, делить, как извлекать квадратные корни, причем не только с натуральными числами, но и с дробями, как с обыкновенными, так и шестидесятеричными.



Астролябия — популярный в Средние века на Востоке и в Европе астрономический инструмент.

Трактат «Об индийском счете» был переведен на латынь и оказал огромное влияние на развитие математики. Имя ал-Хорезми в латинизированной форме «алгоризм» или «алгоритм» стало обозначать в Европе десятичную позиционную систему счисления. (Позже, в XVII в., под влиянием Лейбница слово «алгоритм» приобрело более широкий смысл.)

Алгебраический трактат ал-Хорезми называется «Краткая книга восполнения и противопоставления» (по-арабски — «Китаб мухтасар ал-джабр ва-л-мукабала»). Первое из двух основных действий — восполнение (ал-джабр) — это перенос с другим знаком члена из одной части уравнения в другую. Именно от «ал-джабр» произошло современное слово «алгебра». Второе действие — противопоставление (ал-мукабала) — это сокращение равных членов в обеих частях уравнения.

В первой части «Китаб мухтасар ал-джабр ва-л-мукабала» изложена теория линейных и квадратных уравнений. Во второй она применена к решению конкретных задач — хозяйственных, торговых и юридических. Во введении к трактату ал-Хорезми писал: «Я составил краткую книгу об исчислении алгебры и алмукабалы, заключающую в себе простые и сложные вопросы арифметики, ибо это необходимо людям при разделе наследства, составлении завещания, разделе имущества и в судебных делах, в торговых и всевозможных сделках, а также при измерении земель, проведении каналов, геометрии и прочих разновидностях подобных дел». ■

Линейные уравнения. Общий случай.

Пусть при помощи алгоритма Евклида или любым другим способом на прямой $ax-by=c$ мы нашли одну точку с целыми координатами. Найдем все целочисленные точки этой прямой.

Теорема. ***Если числа a, b взаимно простые, а x_0 и y_0 — целые числа, удовлетворяющие равенству $ax_0-by_0=c$, то все пары чисел x, y, удовлетворяющие равенству $ax-by=c$, описываются формулами***

$$x=x_0+bt, \qquad y=y_0+at,$$

где t — целое.

Доказательство. В одну сторону утверждение очевидно:

$$a(x_0+bt)-b(y_0+at)=ax_0+abt-by_0-abt=ax_0-by_0=c.$$

В другую сторону рассуждение чуть сложнее:

$$ax-by=c=ax_0-by_0,$$
$$a(x-x_0)=b(y-y_0).$$

Поскольку числа a и b взаимно просты и $a(x-x_0)$ делится на b, то $x-x_0$ делится на b, то есть $x-x_0=bt$ для некоторого целого t. При этом

$$abt=b(y-y_0),$$

откуда $y=y_0+at$, что и требовалось доказать.

Следующая теорема — необходимое и достаточное условие разрешимости линейного уравнения в целых числах.

Теорема. ***Уравнение***

$$ax-by=c,$$

где a, b, c — данные целые числа, x, y — неизвестные, имеет решения в целых числах тогда и только тогда, когда число c делится на наибольший общий делитель чисел a и b.

Одно доказательство этой теоремы непосредственно вытекает из изложенного выше способа решения с помощью алгоритма Евклида. (Подумайте, как именно!) Второй — не менее замечательный — способ доказательства основан на следующей лемме.

Лемма. ***Пусть a и b — целые числа, хотя бы одно из которых не равно 0. Обозначим буквой m наименьшее натуральное число, представимое в виде***

$$m=ax-by,$$

где x, y — целые числа. Тогда m = НОД(a; b).

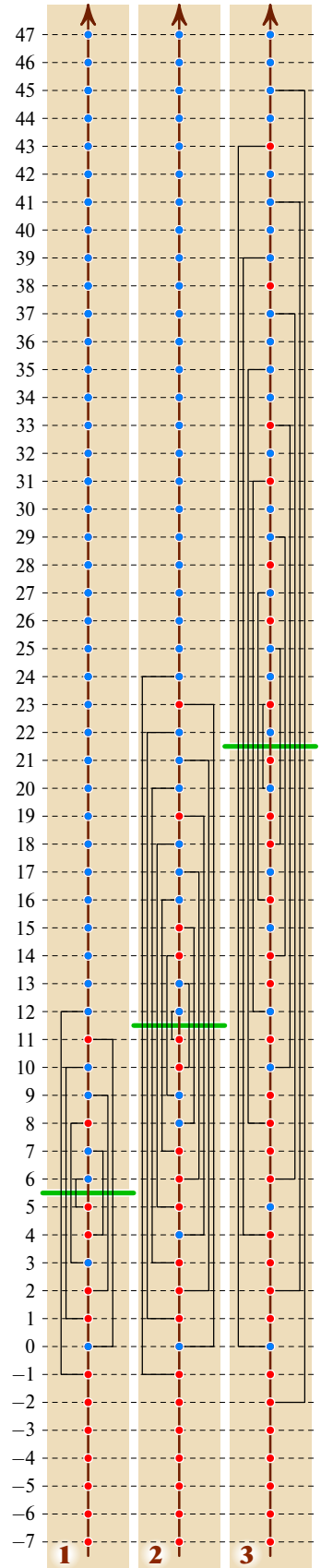
Доказательство. Обозначим $d=\text{НОД}(a;b)$. Тогда любое число вида $ax-by$, в том числе и m, делится на d. Осталось доказать, что не только m делится на d, но и d — на m. Для этого докажем, что как число a, так и число b, делится на m. Рассуждаем «от противного». Пусть, например, a не делится на m. Разделим a на m с остатком:

$$a=mq+r,$$

где $0<r<m$. Поскольку число m представимо в виде $m=aX-bY$, где X, Y — целые, то

$$r=a-mq=a-(aX-bY)q=a(1-Xq)+bYq.$$

Таким образом, число r представимо в виде $r=ax-by$. Но $0<r<m$, а m, как помните, — наименьшее натуральное число, представимое в таком виде. Лемма доказана. Утверждение теоремы следует из нее. ■



Рассмотрим взаимно простые натуральные числа a и b. Буквой M обозначим множество целых чисел, представимых в виде $ax+by$, где x и y — целые неотрицательные числа.

Для $a=3$ и $b=7$ синим цветом на рисунке 1 на числовой прямой изображены числа, принадлежащие множеству M, а красным — не принадлежащие. При симметрии относительно числа 5,5 красные числа переходят в синие, а синие — в красные. То же явление видим на рисунке 2 для $a=4$ и $b=9$ (центр симметрии — 11,5) и на рисунке 3 для $a=5$, $b=12$ (центр симметрии — 21,5).

Теорема. ***Синие точки симметричны красным относительно точки*** $c=\dfrac{ab-a-b}{2}$.

Доказательство. Предположим сначала, что некоторая синяя точка $m=ax_1+by_1$, где x и y — неотрицательные целые числа, симметрична относительно точки c синей точке $n=ax_2+by_2$, где x_2 и y_2 — тоже целые неотрицательные числа. Тогда $c=\dfrac{m+n}{2}$, то есть

$$ab-a-b=a(x_1+x_2)+b(y_1+y_2).$$

Обозначив $x=x_1+x_2$ и $y=y_1+y_2$, получаем

$$ab=a(x+1)+b(y+1).$$

Поскольку числа ab и $a(x+1)$ делятся на a, то и число $b(y+1)$ делится на a. Значит, $y+1$ делится на a и поэтому $y+1\geqslant a$, откуда получаем противоречие:

$$ab=a(x+1)+b(y+1)>b(y+1)\geqslant ab.$$

Теперь предположим, что относительно точки c симметричны две красные точки m и n. Как мы только что видели, это означает, что

$$m+n=ab-a-b.$$

Поскольку числа a и b взаимно просты, то всякое целое число представимо в виде $ax+by$. Поскольку для любого целого t имеем

$$ax+by=a(x+bt)+b(y-bt),$$

то можно считать, что $0\leqslant x<b$. Таким образом,

$$\begin{cases} m=ax_1+by_1,\\ n=ax_2+by_2, \end{cases}$$

где $0\leqslant x_1,\ x_2<b$, а числа y_1 и y_2 отрицательные. Складывая эти равенства, получаем

$$ab-a-b=a(x_1+x_2)+b(y_1+y_2),$$

то есть

$$ab=a(x_1+x_2+1)+b(y_1+y_2+1).$$

Из последнего равенства следует делимость числа x_1+x_2+1 на b. Поскольку x_1 и x_2 меньше числа b, то сумма x_1+x_2+1 меньше $2b$. Следовательно, $x_1+x_2+1=b$. Значит,

$$ab=ab+b(y_1+y_2+1),$$

что противоречит неравенствам $y_1\leqslant -1$ и $y_2\leqslant -1$. ■

Китайская теорема об остатках. Какое число при делении на 3 дает остаток 2, при делении на 5 — остаток 3, а при делении на 7 — остаток 2? Эта задача о решении системы сравнений

$$\begin{cases} x\equiv 2 \pmod 3,\\ x\equiv 3 \pmod 5,\\ x\equiv 2 \pmod 7 \end{cases}$$

была известна уже в Древнем Китае. Сунь-цзы (между II и VI в.) и более полно Цинь Цзю-шао (XIII в.) дают изложенное на примерах описание алго-

ритма решения таких задач. В точности эта задача есть и в «Книге об абаке» итальянского математика Леонардо Пизанского (Фибоначчи) (1202). Ответ получить несложно:

$$x \equiv 23 \pmod{105}.$$

(Проверьте!)

Но нас интересует не эта частная задача, а общий случай. По определению, числа, дающие некоторый остаток r при делении на натуральное число m, имеют вид $mq+r$, где q — целое. Как известно, последовательность

$$r,\ r+m,\ r+2m,\ r+3m,\ \ldots$$

называют арифметической прогрессией с разностью m.

Можно указать арифметические прогрессии из натуральных чисел, пересечение которых пусто. Например, поскольку ни одно число не является одновременно четным и нечетным, пусто пересечение прогрессии 2, 4, 6, . . . с прогрессией 1, 3, 5, . . .

Оказывается, если разности m и n двух арифметических прогрессий, члены которых — натуральные числа, являются взаимно простыми числами, то пересечение прогрессий — арифметическая прогрессия (и разность прогрессии-пересечения равна mn).

Китайская теорема об остатках. ***Если a и b — целые числа, m и n — взаимно простые натуральные числа, то пересечение арифметической прогрессии***

$$a,\ a+m,\ a+2m,\ a+3m,\ \ldots$$

с арифметической прогрессией

$$b,\ b+n,\ b+2n,\ b+3n,\ \ldots$$

является арифметической прогрессией с разностью mn.

Здесь натуральные числа упомянуты по той только причине, что арифметическая прогрессия «направлена в одну сторону». Если вместо прогрессий рассматривать бесконечные в обе стороны последовательности вида

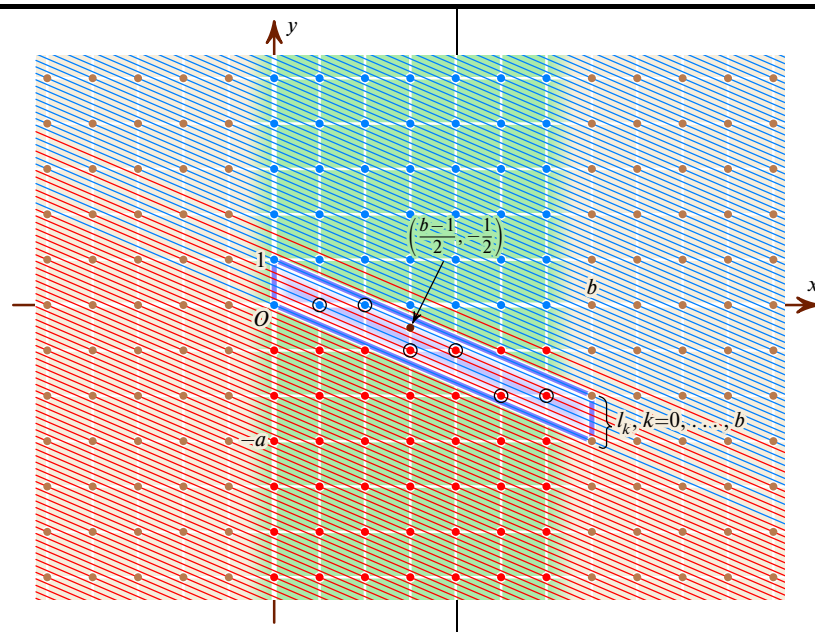
$$\ldots,\ a-2m,\ a-m,\ a,\ a+m,\ a+2m,\ a+3m,\ \ldots,$$

в которых соседи отличаются на m, то пересечение такой последовательности (а такие последовательности настолько важны, что получили название *классы вычетов по модулю m*) с последовательностью

$$\ldots,\ b-2n,\ b-n,\ b,\ b+n,\ b+2n,\ b+3n,\ \ldots,$$

будет (при условии НОД$(m, n)=1$) последовательностью того же вида, но с разностью mn (то есть пересечение класса вычетов по модулю m с классом вычетов по модулю n — класс вычетов по модулю mn).

Другими словами, если знать остаток a от деления целого числа x на m и остаток b от деления числа x на n, то можно, причем единственным образом, найти остаток от деления x на mn. Доказательство китайской теоремы об остатках читатель легко проведет самостоятельно. ■



На рисунке для $a=3$ и $b=7$ нарисованы прямые l_n, заданные уравнениями $ax+by=n$, и выделена полоса, заданная неравенствами $0 \leqslant x < b$. Каждая прямая l_n пересекает полосу в одной целочисленной точке. При симметрии относительно точки $\left(\frac{b-1}{2}; -\frac{1}{2}\right)$ полоса переходит в себя, причем синие точки переходят в красные и наоборот. При этом

$$a \cdot \frac{b-1}{2} + b \cdot \left(-\frac{1}{2}\right) = \frac{ab-a-b}{2} = c. \blacksquare$$

Этот же рисунок помогает доказать и существование решения уравнения $ax+by=n$ для любого целого n при условии НОД$(a; b)=1$. Рассмотрим параллелограмм с вершинами $(0; 0)$, $(0; 1)$, $(b; 1-a)$ и $(b; -a)$. (Длины двух меньших его сторон равны 1, а большие стороны лежат на прямых l_0 и l_b.) Внутри этого параллелограмма (не в вершинах) лежит ровно $b-1$ целых точек: на каждой из прямых $x=1, x=2, \ldots$ $\ldots, x=b-1$ — по одной. Этот параллелограмм пересекают как раз столько же — $b-1$ — прямых: $l_1, l_2, \ldots, l_{b-1}$. Поскольку более одной целочисленной точки, абсциссы которых отличаются менее чем на b, прямая l_n ни при каком n иметь не может, то все $b-1$ точек лежат по одной точке на каждой прямой. В частности, есть целочисленная точка и на прямой l_1. ■

Основная теорема арифметики

Множители любого натурального числа, представленного в виде произведения простых чисел, можно располагать в каком угодно порядке. На этом свобода заканчивается: разложение единственно с точностью до порядка множителей.

Если произведение ab двух натуральных чисел a и b кратно числу 3, то хотя бы один из множителей делится на 3 без остатка. Этот факт настолько привычен, что многие даже не задумываются о его доказательстве. Между тем — если, конечно, рассматривать делимость не только на 3, но и на любое простое число, — это одно из важнейших свойств натуральных чисел.

Основная теорема арифметики. ***Любое натуральное число либо равно 1, либо простое (то есть имеет ровно два натуральных делителя — 1 и само себя), либо единственным с точностью до порядка множителей способом разложимо в произведение простых чисел.***

Мы рассмотрим три доказательства. Первое использует идею из «Начал» Евклида. Второе изложено в «Арифметических исследованиях» К. Ф. Гаусса, изданных в 1801 г. Третье на рубеже XIX и XX вв. придумал Э. Цермело.

Строго говоря, в «Началах» основная теорема арифметики нигде явно не сформулирована. Но все необходимое для доказательства там есть. «Все необходимое» — это алгоритм Евклида нахождения наибольшего общего делителя двух чисел и (легко получаемое при изучении работы этого алгоритма) утверждение: для любых взаимно простых целых чисел m и n существуют такие целые x и y, что $mx-ny=1$.

Основная лемма. ***Если произведение двух натуральных чисел a и b делится на простое число p, то хотя бы одно из чисел a и b делится на p.***

Доказательство. Пусть a не делится на p. Тогда числа a и p взаимно просты, поэтому для некоторых целых чисел x и y выполнено равенство

$$ax-py=1.$$

Следовательно,

$$b=abx-pby=p\left(\frac{ab}{p}\cdot x-by\right).$$

Поскольку число $\frac{ab}{p}$ целое, то b делится на p. Основная лемма доказана. ■

Доказать основную теорему арифметики теперь очень легко. Возможность разложения очевидна: берем любое натуральное число и, если оно разложимо, разлагаем на два множителя. Если получили неразложимые сомножители — хорошо. Если хотя бы один из полученных сомножителей можно разложить — разлагаем его! И так действуем до тех пор, пока можно. Бесконечно долго этот процесс не продлится: на каждом шаге числа уменьшаются, а бесконечно долго уменьшать натуральное число, оставаясь во множестве натуральных чисел, невозможно. ■

Леонтий Филиппович Магницкий (1669—1739) — выдающийся русский математик и педагог — родился в селе Осташковское Тверской губернии. По некоторым источникам, окончил Славяно-греко-латинскую академию в Москве, но его имя не встречается в списках ее учеников.

В начале XVIII в. Россия остро нуждалась в специалистах, и для их подготовки открывались учебные заведения, в которых математика была одним из основных предметов, за преподаванием которых следил лично царь. Первой по указу от 14 января 1701 г. была основана Школа математических и навигацких наук. С момента ее создания в ней работал Л. Ф. Магницкий. Он преподавал арифметику, геометрию, тригонометрию и мореходные науки, а с 1716 г. до конца жизни руководил школой.

В 1703 г. Магницкий совместно с другими преподавателями опубликовал «Таблицы логарифмов синусов, тангенсов и секансов», а в 1722 г. — «Таблицы горизонтальных северных и южных широт».

В 1703 г. Магницкий издал свою знаменитую «Арифметику». Кроме правил арифметических действий, элементов алгебры, геометрии и тригонометрии, в ней приведены расчеты, связанные с коммерцией, картографией, навигацией. Магницкий ввел современную форму записи чисел вместо распространенной ранее в России алфавитной, а старый счет на тьмы, легионы, леодры заменил на принятый в Европе счет миллионами, биллионами, триллионами. Он обогатил русскую терминологию словами «множитель», «произведение», «делимое», «делитель», «среднее пропорциональное», «пропорция», «прогрессия», «извлечение корня».

Многие задачи «Арифметики» даны с подробными решениями: Магницкий предназначал книгу для широкого круга читателей. Он написал ее «ради обучения мудролюбивых российских отроков и всякого чина и возраста людей». До середины XVIII в. она не имела конкурентов среди учебников по математике. «Арифметику» Леонтия Магницкого и «Грамматику» Мелетия Смотрицкого назвал «вратами своей учености» М. В. Ломоносов. ■

Рассмотрим множество натуральных чисел, оканчивающихся цифрой 1, и будем интересоваться лишь разложениями на множители, тоже оканчивающиеся на 1. Числа 11, 21, 31, 41, . . . , 111 разложить не удастся, а $121=11\cdot 11$. Очевидно,

$$(3\cdot 7)\cdot(13\cdot 17)=(3\cdot 17)\cdot(13\cdot 7),$$

то есть

$$21\cdot 221=51\cdot 91.$$

Числа 21, 221, 51 и 91 не представимы в виде произведения отличных от 1 и оканчивающихся на 1 натуральных чисел. Таким образом, не случайно в доказательствах основной теоремы арифметики использованы не только операции умножения и деления, но и сложения и вычитания. ■

Эрнст Цермело (1871—1953)— немецкий математик, создавший вместе с Френкелем в 1921—1922 гг. систему аксиом теории множеств. Именем Цермело часто называют одну из наименее очевидных аксиом теории — аксиому выбора, согласно которой для любого множества M существует определенная на множестве всех его непустых подмножеств функция f, любое значение $f(A)$ которой принадлежит подмножеству A множества M. Из аксиомы выбора следует, что любое множество можно вполне упорядочить. ■

Фрагмент книги «Арифметика» Л. Ф. Магницкого (1-е издание, 1703 г.).

Нетривиальную часть основной теоремы арифметики — однозначность разложения — получаем из основной леммы. А именно, пусть число

$$N=p_1p_2\ldots p_r=q_1q_2\ldots q_s$$

разложено на неразложимые и отличные от 1 множители двумя способами. Тогда произведение $q_1q_2\ldots q_s$ кратно числу p_1. В силу основной леммы хотя бы один из множителей $q_1, q_2, \ldots, q_s$ кратен p_1. (Подумайте, почему можно пользоваться утверждением для s множителей, хотя лемму мы доказали только для двух!) Если некоторое неразложимое натуральное число q_k делится на p_1, то $q_k=p_1$ и обе части равенства можно сократить на p_1.
«Уничтожив» p_1, точно так же поступим с $p_2, \ldots, p_r$. Так и получится, что множители левой части разложения числа N могут отличаться от множителей правой части разве лишь порядком, в котором они записаны. ■

Гаусс доказывает существование разложения на простые множители точно так же, как Евклид, а вот основную лемму доказывает иначе. Пусть произведение некоторых двух натуральных чисел a и b, не делящихся на простое число p, делится на p. Зафиксируем числа a и p, а из всех возможных натуральных чисел b, для которых ab делится на p, выберем *наименьшее*. Очевидно, $b<p$ и число p, будучи простым, не делится на b. Поэтому число p расположено между некоторыми кратными числа b, то есть

$$mb<p<(m+1)b$$

для некоторого натурального m. Обозначим $c=p-mb$. Очевидно, $0<c<b$ и $ac=ap-abm$ делится на p, так что мы получили противоречие: на роль (наименьшего!) числа b претендует число $c<b$. ■

Цермело доказывает единственность разложения следующим образом. Предположим, что существует натуральное число N, которое можно существенно разными способами представить в виде произведения простых чисел:

$$N=p_1p_2\ldots p_r=q_1q_2\ldots q_s.$$

Пусть N — *наименьшее* из таких чисел. Тогда ни одно из чисел $p_1, \ldots, p_r$ не равно ни одному из чисел $q_1, q_2, \ldots, q_s$ (в противном случае мы сократили бы обе части равенства на общий множитель, получив меньшее число). Обозначим $P=p_2\ldots p_r$ и $Q=q_2\ldots q_s$. Тогда

$$N=p_1P=q_1Q.$$

Не ограничивая общности, можно считать, что $p_1<q_1$. При этом $P>Q$ и, значит, $p_1Q<N$. Рассмотрим число

$$M=N-p_1Q.$$

Поскольку $M<N$, число M разлагается на простые множители единственным образом. Но

$$p_1(P-Q)=M=(q_1-p_1)\,q_2\ldots q_s.$$

Левая часть равенства содержит простой множитель p_1. Поскольку ни одно из чисел $q_2, \ldots, q_s$ не равно p_1, то разность q_1-p_1 делится на p_1; следовательно, q_1 делится на p_1, что противоречит простоте числа q_1 и тому, что $p_1\neq q_1$. ■

Ряды Фарея

Выписав в порядке возрастания несократимые правильные дроби, знаменатели которых не превосходят некоторого заданного числа n, мы получаем n-й ряд Фарея. В 1816 г. француз О. Л. Коши доказал две подмеченные Дж. Фареем закономерности.

- *Если* $\frac{a}{b} < \frac{c}{d}$ *— две последовательные дроби ряда Фарея, то*

$$bc - ad = 1.$$

- *Если* $\frac{a}{b} < \frac{c}{d} < \frac{e}{f}$ *— три последовательные дроби ряда Фарея, то*

$$\frac{c}{d} = \frac{a+e}{b+f}.$$

Дроби, равноотстоящие от краев ряда Фарея, имеют одинаковые знаменатели. Например, в 5-м ряде симметрично расположены относительно 1/2 следующие пары дробей: 0/1 и 1/1, 1/5 и 4/5, 1/4 и 3/4, 1/3 и 2/3, 2/5 и 3/5. Более того, симметрично расположенные дроби дополняют одна другую до единицы (то есть их сумма равна числу 1). Причина очевидна: неравенства $x < y$ и $1-x > 1-y$ равносильны.

6-й ряд Фарея отличается от 5-го тем, что добавляются две дроби: 1/6 и 5/6. Все другие правильные дроби со знаменателем 6 сократимы: $2/6 = 1/3$, $3/6 = 1/2$ и $4/6 = 2/3$. А вот в 7-м ряде появляются сразу 6 новых дробей. Каждая из них — медианта дробей, между которыми она заключена:

$$\frac{1}{7} = \frac{0+1}{1+6}, \quad \frac{2}{7} = \frac{1+1}{4+3}, \quad \frac{3}{7} = \frac{2+1}{5+2},$$
$$\frac{4}{7} = \frac{1+3}{2+5}, \quad \frac{5}{7} = \frac{2+3}{3+4}, \quad \frac{6}{7} = \frac{5+1}{6+1}. \blacksquare$$

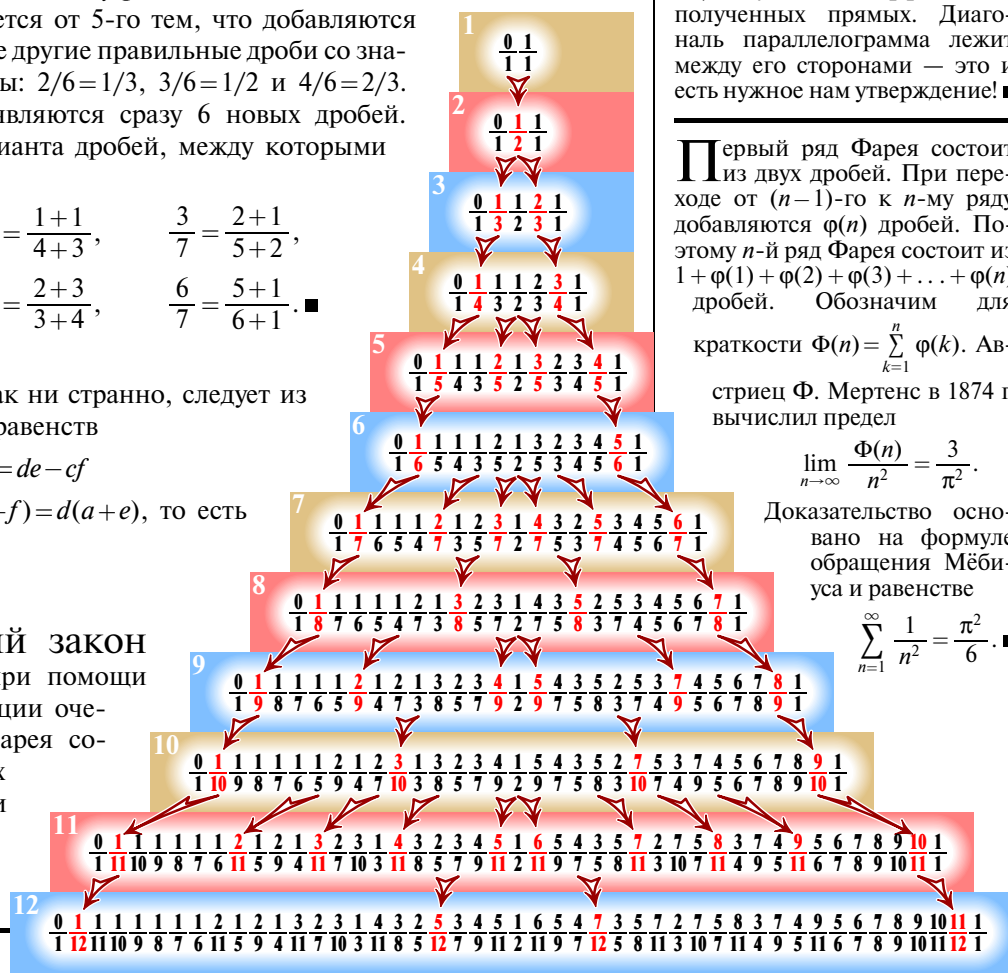
Второй закон, как ни странно, следует из первого. А именно, из равенств

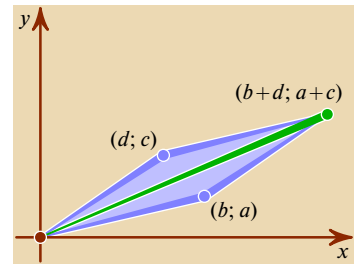
$$bc - ad = 1 = de - cf$$

следует равенство $c(b+f) = d(a+e)$, то есть $\frac{c}{d} = \frac{a+e}{b+f}$. $\blacksquare$

Доказать первый закон Фарея проще всего при помощи индукции. База индукции очевидна: первый ряд Фарея состоит всего лишь из двух дробей 0/1 и 1/1, при этом

$$1 \cdot 1 - 0 \cdot 1 = 1.$$





Медианта дробей $\frac{a}{b}$ и $\frac{c}{d}$ — это дробь $\frac{a+c}{b+d}$, числитель которой равен сумме числителей, а знаменатель — сумме знаменателей.

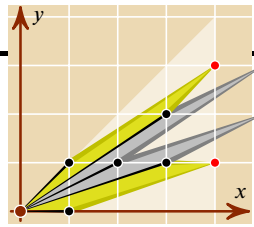
Если числа a, b, c и d положительные и $\frac{a}{b} < \frac{c}{d}$, то медианта $\frac{a+c}{b+d}$ расположена между дробями a/b и c/d. Это легко доказать алгебраически. Но есть и красивое геометрическое доказательство. Рассмотрим точки $(b; a)$ и $(d; c)$ и соединим их с началом координат. Тогда a/b и c/d — угловые коэффициенты полученных прямых. Диагональ параллелограмма лежит между его сторонами — это и есть нужное нам утверждение! $\blacksquare$

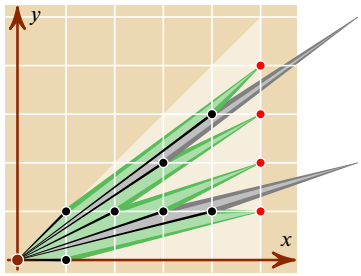
Первый ряд Фарея состоит из двух дробей. При переходе от $(n-1)$-го к n-му ряду добавляются $\varphi(n)$ дробей. Поэтому n-й ряд Фарея состоит из $1 + \varphi(1) + \varphi(2) + \varphi(3) + \ldots + \varphi(n)$ дробей. Обозначим для краткости $\Phi(n) = \sum_{k=1}^{n} \varphi(k)$. Австриец Ф. Мертенс в 1874 г. вычислил предел

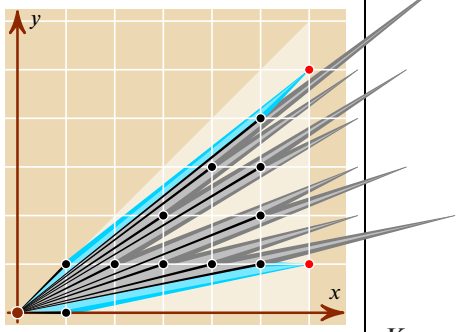
$$\lim_{n\to\infty} \frac{\Phi(n)}{n^2} = \frac{3}{\pi^2}.$$

Доказательство основано на формуле обращения Мёбиуса и равенстве

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







На верхнем рисунке показано, как из изображающих третий ряд Фарея черных точек (1; 0), (3; 1), (2; 1), (3; 2) и (1; 1) получить две красные точки (4; 1) и (4; 3), изображающие дроби 1/4 и 3/4 соответственно. На следующих двух рисунках изображен переход от 4-го к 5-му и от 5-го к 6-му ряду. Процедура очевидна: для каждых двух соседних дробей a/b и c/d строим параллелограмм с вершинами $O(0;0)$, $A(b;a)$, $B(b+d;a+c)$ и $C(d;c)$. Часть параллелограммов каждый раз «вылезают за границу» и поэтому закрашены серым цветом.

Параллелограмм $OABC$ не содержит внутри ни одной точки с целыми координатами, да и на границе их только 4 (вершины O, A, B и C). Поскольку площадь каждого параллелограмма равна $bc-ad=1$, то чем длиннее становятся его стороны, тем он «тоньше».

Чтобы изобразить для первых 12 рядов Фарея «все параллелограммы сразу», использована косоугольная система координат: величина угла между осями не 90°, а 150°. ■

Предположим, что для некоторого натурального n закон верен для $(n-1)$-го ряда Фарея. Чтобы получить из $(n-1)$-го ряда n-й, мы должны добавить дроби вида m/n, где $1\leqslant m\leqslant n$ и НОД$(m, n)=1$. Рассмотрим одну из таких дробей m/n и обозначим через a/b и c/d ближайшие к ней слева и справа дроби $(n-1)$-го ряда:

$$\frac{a}{b}<\frac{m}{n}<\frac{c}{d}.$$

Теперь — внимание:

$$\frac{1}{bd}=\frac{bc-ad}{bd}=\frac{c}{d}-\frac{a}{b}=$$

$$=\frac{c}{d}-\frac{m}{n}+\frac{m}{n}-\frac{a}{b}=\frac{cn-dm}{dn}+\frac{bm-an}{bn}\geqslant$$

$$\geqslant\frac{1}{dn}+\frac{1}{bn}=\frac{b+d}{bdn},\qquad(*)$$

следовательно, $b+d\leqslant n$. Если бы это неравенство было строгим, то дробь $\dfrac{a+c}{b+d}$ — медианта дробей a/b и c/d — лежала бы между ними и принадлежала бы $(n-1)$-му ряду Фарея. Но никаких дробей между a/b и c/d в $(n-1)$-м ряде Фарея нет. Следовательно, $b+d=n$, а неравенство в формуле $(*)$ — на самом деле равенство. Значит,

$$cn-dm=1\quad\text{и}\quad bm-an=1.$$

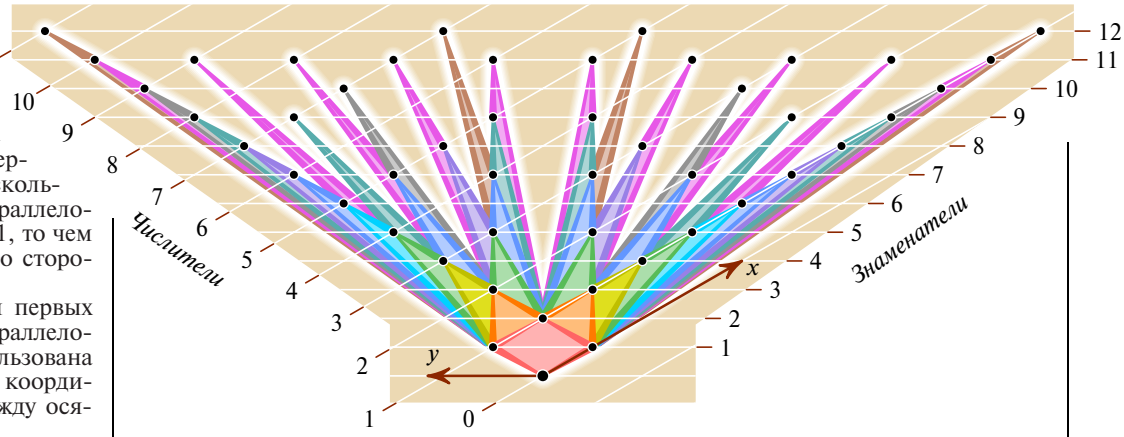
Казалось бы, первый закон Фарея доказан: последние две формулы — это как раз нужные формулы для n-го ряда Фарея. Но. . . почему мы уверены, что в n-м ряде между дробями a/b и c/d расположена лишь одна дробь со знаменателем n? А вот почему: приравняв левые части последних двух равенств, получаем

$$cn-dm=bm-an,$$

откуда $m(b+d)=n(a+c)$. Поскольку $n=b+d$, то $m=a+c$. Значит, при переходе от $(n-1)$-го к n-му ряду Фарея между дробями a/c и b/d вставляется одна дробь

$$\frac{m}{n}=\frac{a+c}{b+d}.$$

Доказательство завершено! ■



Периодические дроби

Какова длина периода десятичного представления дроби $1/7^{77}$? А длина периода суммы двух бесконечных десятичных периодических дробей, длины периодов которых равны 6 и 12? Мы расскажем о связи между обыкновенными дробями и периодическими десятичными дробями настолько подробно, что читатель сам легко ответит на эти и многие другие вопросы, и даже докажем по ходу дела малую теорему Ферма.

Обыкновенная дробь — это число, составленное из целого количества долей единицы. Дробь записывают в виде $\frac{m}{n}$ или m/n, где числитель m — целое число, а знаменатель n — натуральное число. Для получения дроби $\frac{m}{n}$ надо разделить единицу на n равных частей и взять m таких частей. Величина дроби не изменится, если ее числитель и знаменатель умножить на одно и то же натуральное число. Благодаря этому любые две дроби $\frac{a}{b}$ и $\frac{c}{d}$ можно привести к общему знаменателю bd, заменив их на $\frac{ad}{bd}$ и $\frac{bc}{bd}$ соответственно.

Если числитель и знаменатель дроби имеют отличный от единицы общий делитель, то дробь можно сократить. Поэтому всякую дробь можно представить в несократимом виде, то есть в виде дроби, числитель и знаменатель которой — взаимно простые числа. Например, $\frac{120}{344}$ — сократимая дробь $\left(\frac{120}{344}=\frac{15\cdot 8}{43\cdot 8}=\frac{15}{43}\right)$, а $\frac{15}{43}$ — равная ей несократимая дробь.

Дробь $\frac{m}{n}$ называют *правильной*, если $0\leqslant m<n$. Всякую дробь можно единственным образом представить в виде суммы целого числа $\left[\frac{m}{n}\right]$ (целой части дроби $\frac{m}{n}$) и правильной дроби $\left\{\frac{m}{n}\right\}$ (дробной части). Например,

$$\frac{91}{17}=\frac{5\cdot 17+6}{17}=5+\frac{6}{17},$$

так что

$$\left[\frac{91}{17}\right]=5,\qquad \left\{\frac{91}{17}\right\}=\frac{6}{17}.$$

Сумму и разность дробей с одинаковыми знаменателями определяют по правилам:

$$\frac{a}{n}\pm\frac{b}{n}=\frac{a\pm b}{n}.$$

Чтобы сложить или вычесть дроби k/l и m/n с разными знаменателями, их предварительно приводят к общему знаменателю. Обычно в качестве него берут наименьшее общее кратное НОК$[l, n]$ чисел l и n. ■

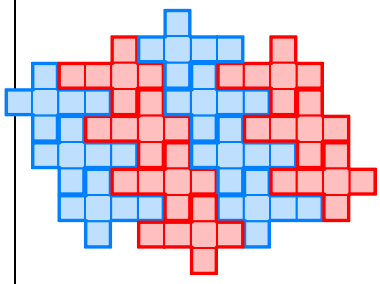
Как записать обыкновенную дробь m/n в десятичной системе счисления? Если n — степень двойки, степень пятерки или произведение степеней двойки и пятерки, то ответ — конечная десятичная дробь. Например,

$$\frac{13}{64}=\frac{13\cdot 15625}{64\cdot 15625}=$$
$$=\frac{203125}{1000000}=0{,}203125;$$
$$\frac{3}{25}=\frac{3\cdot 4}{25\cdot 4}=0{,}12;$$
$$\frac{3}{40}=\frac{3\cdot 25}{40\cdot 25}=\frac{75}{1000}=0{,}075.$$

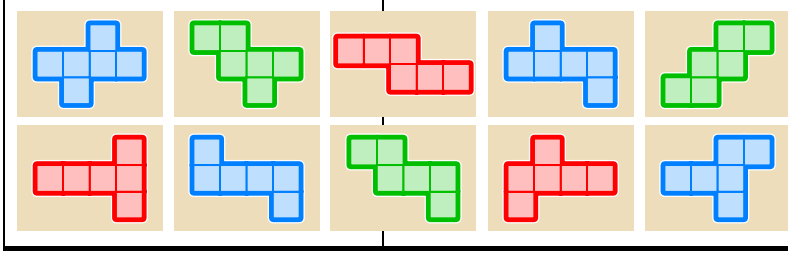
Хотя число 35 не является произведением степеней двойки и пятерки, дробь 7/35 представима в виде конечной десятичной дроби:

$$\frac{7}{35}=\frac{1}{5}=0{,}2.$$

Но если дробь m/n несократима и при этом хотя бы один из простых делителей числа n отличен от 2 и 5, то m/n нельзя представить в виде конечной десятичной дроби. (Действительно, если $m/n=a/10^b$, то $10^b m=an$; рассмотрев любой отличный от 2 и 5 простой делитель p числа n, приходим к противоречию: an кратно p, а равное ему число $10^b m$ — не кратно.) ■



На рисунке показано, как копиями развертки куба можно периодически замостить плоскость. Проверьте, что копиями любой из десяти других разверток куба тоже можно периодически замостить плоскость. (*Решение — на с. 384.*) ■



Десятичные дроби, то есть дроби, знаменатели которых — степени числа 10, предложил использовать нидерландский ученый и инженер С. Стевин (1548—1620). Складывать, вычитать и сравнивать их легче, чем обыкновенные дроби. Десятичные дроби обычно пишут без знаменателя: например, $\frac{5\,481\,475}{10\,000}=548{,}1475$ и $\frac{23}{1\,000}=0{,}023$.

Переводить дроби из обыкновенных в десятичные можно делением «уголком». Например, разделим 3 на 7. Целая часть равна 0. Чтобы получить первую цифру после запятой, разделим 30 на 7. Получим частное 4 и остаток 2. Разделив 20 на 7, получаем частное 2 и остаток 6. Следующий шаг — деление 60 на 7 — дает частное 8 и остаток 4. Далее,

$$40=5\cdot 7+5,$$
$$50=7\cdot 7+1,$$
$$10=1\cdot 7+3.$$

Мы вернулись к задаче деления 3 на 7; произошло зацикливание: если продолжим деление, то опять получим частное 4 и остаток 2, затем будем делить 20 на 7, и так далее:

$$\frac{3}{7}=0{,}428571\,428571\,428571\ldots$$

Обычно этот результат записывают короче:

$$\frac{3}{7}=0{,}(428571),$$

то есть заключают повторяющуюся группу цифр в скобки и говорят: «428571 в периоде».

Если повторяющаяся группа цифр (период) расположена непосредственно после запятой, а целая часть равна нулю, то десятичную дробь называют *чисто периодической*; в противном случае говорят, что дробь имеет *предпериод*, и называют ее *смешанной периодической*. (Существуют и непериодические дроби, например десятичная дробь 0,1010010001 . . . , где количество нулей между единицами все время увеличивается на 1. Именно такие дроби изображают иррациональные числа.)

Теорема 1. ***Десятичное представление дроби m/n, где m, n — натуральные числа, $m<n$, — является периодической дробью, длина наименьшего периода которой не превосходит $n-1$.***

Доказательство. Чтобы получить первую цифру после запятой, мы приписываем к m нуль (то есть умножаем m на 10) и делим (с остатком) полученное число на n. Вообще весь процесс деления уголком — повторяемое вновь и вновь умножение очередного остатка на 10 и деление (с остатком) на n.

Если на каком-то шаге получится нулевой остаток, то дробь — конечная. Конечную дробь, приписав к ней справа бесконечно много нулей, естественно считать периодической с периодом длины 1. По условию, $1\leqslant n-1$, так что в этом случае утверждение теоремы выполнено.

Если же процесс деления никогда не закончится, то будут получаться только ненулевые остатки, то есть числа от 1 до $n-1$. Значит, не позже чем на $(n-1)$-м шаге остаток повторится. С этого момента процесс деления зациклится, что и требовалось доказать. ■

Арбузная пошлина. Наступило утро, и городские ворота со скрипом распахнулись. Караваны, груженные драгоценными индийскими тканями, прекрасной медной и серебряной посудой, хорасанскими коврами и другими товарами, двинулись в Бухару. Вслед за караваном богатого багдадского купца в ворота въехала и арба дехканина Али. На арбе сидели двое: хозяин и Ходжа Насреддин, недавно выручивший Али из неприятной истории с поливом посевов. В благодарность за помощь Али чуть не силой уговорил Насреддина принять от него часть урожая, и теперь они вместе везли арбузы на базар: 104 арбуза Али и 17 — Насреддин.

— Стойте! — закричал стражник, и арба остановилась. — По какому делу едете в благородную Бухару? Чем торгуете? Почему не заплатили деловую пошлину? Вижу, у вас тут арбузы? Сейчас же платите арбузную пошлину! Арба скрипит, словно ее не смазывали целый год! Платите пошлину за скрип! Поворачивайте оглобли! Просто так мы не пустим непонятно кого!

— Мы. . . — начал Али.

— Так уж и быть,— не давая ответить, продолжил стражник,— вижу, деньги у вас вряд ли найдутся. Чего уж по доброте не сделаешь, помогу вам. Платите пошлину арбузами!

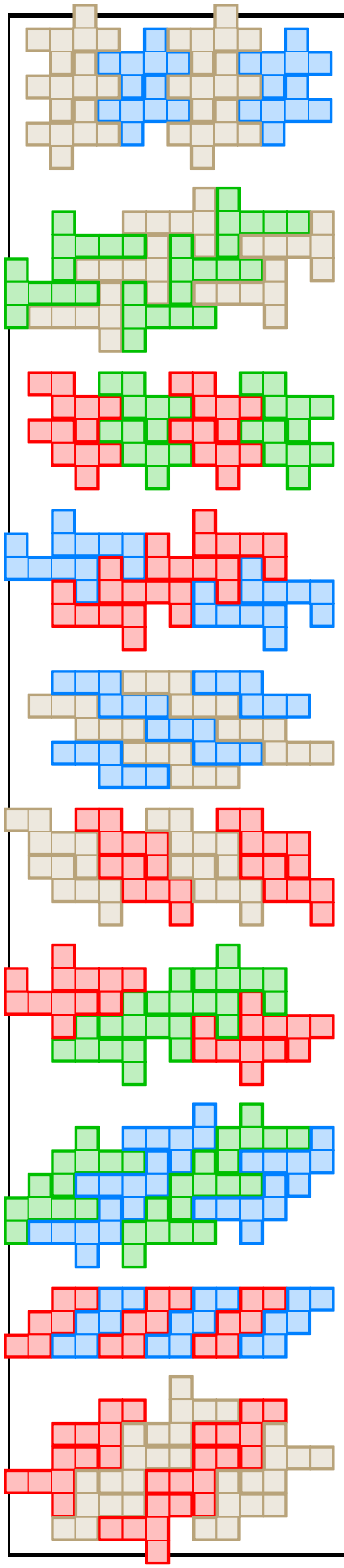
Чуть поторговавшись, стражник и Насреддин пришли к соглашению:

— Значит, я должен тебе 3 арбуза и 1 таньга, — сказал Насреддин. — Правда, у меня нет ни одной монеты. . . Но мой друг Али должен 19 арбузов без 1 таньга. Значит, если мы дадим 22 арбуза, то будем в расчете! По рукам?

— Эх, чего только по доброте не сделаешь! Ограничимся арбузной пошлиной! Все равно больше с вас ничего не возьмешь! — проворчал стражник. — Разгружайте арбузы, так уж и быть!

Вечером, распродав товар, Ходжа и Али зашли в чайхану. Пузатый чайханщик принес каждому по чайнику и поставил на столик пиалы. И не успели они и два часа посидеть, как к ним подсел бородатый старик очень важного вида. Присмотревшись, Насреддин узнал его: это был знаменитый звездочет и мудрец Гуссейн Гуслия, главный математик эмира

(*Продолжение на с. 385.*)



От периодической десятичной дроби к обыкновенной.

Пусть $x=0{,}11111\ldots$ Тогда $10x=1{,}1111\ldots$, откуда $10x=1+x$, то есть $x=\frac{1}{9}$. Мы получили замечательный результат:

$$0{,}11111\ldots=\frac{1}{9}.$$

Это равенство не приближенное, а *точное*: бесконечная десятичная периодическая дробь 0,(1) является в точности тем же самым числом, что и обыкновенная дробь $\frac{1}{9}$. (Между прочим, равенство $0{,}999\ldots=1$ тоже абсолютно точное!) Далее, пусть

$$y=0{,}1733\ 1733\ 1733\ 1733\ldots$$

Тогда $10\,000y=1733{,}1733\ 1733\ 1733\ldots$, откуда

$$10\,000y=1733+0{,}1733\ 1733\ 1733\ 1733\ldots=1733+y.$$

Из уравнения $10\,000y=1733+y$ находим $9999y=1733$, то есть $y=\frac{1733}{9999}$.

Если провести вычисления не для частных примеров, как это сделали мы, а в общем виде, то можно установить следующее правило:

Чисто периодическая правильная дробь равна такой обыкновенной дроби, в числителе которой — период, а в знаменателе — число $10^r-1=\underbrace{9\ldots9}_{r}$ ***, где*** r ***— длина периода.*** ■

Предпериод. Если делить «уголком» 3 на 14, то зацикливание произойдет не сразу:

$$\frac{3}{14}=0{,}2(142857).$$

Период, заметьте, такой же, как у дроби 1/7. Это легко объяснить:

$$\frac{3}{14}=\frac{30}{14}:10=\frac{15}{7}:10=\left(2+\frac{1}{7}\right):10,$$

а делить на 10 очень легко: достаточно перенести запятую на одну позицию. В общем случае выделим в знаменателе степени двойки и пятерки, то есть запишем дробь в виде $\frac{m}{2^a5^bk}$, где a, b — неотрицательные целые числа, k — натуральное число, не кратное ни 2, ни 5. Обозначим наибольшее из чисел a, b буквой c. Поскольку

$$\frac{m}{2^a5^bk}=\frac{m\cdot2^c\cdot5^c}{2^a5^bk}:10^c=\frac{m\cdot2^{c-a}5^{c-b}}{k}:10^c,$$

для решения вопроса о длинах периодов десятичных дробей достаточно изучить дроби со знаменателями, не кратными ни 2, ни 5. ■

Числа вида 99 . . . 9. Взглянем еще раз на равенства

$$\frac{1}{7}=0{,}(142\,857)\quad\text{и}\quad\frac{1}{13}=0{,}(076\,923).$$

Заметьте: $142\,857\cdot7=999\,999$ и $76\,923\cdot13=999\,999$. Это не случайность: в правиле преобразования чисто периодической дроби в обыкновенную тоже фигурирует число $10^r-1=\underbrace{9\ldots9}_{r}$.

Лемма 1. ***Для всякого натурального числа*** k***, не кратного ни 2, ни 5, существует такое натуральное число*** r***, что разность*** 10^r-1 ***кратна*** k***.***

Доказательство. Рассмотрим k чисел: 9, 99, 999, . . . , $\underbrace{99\ldots9}_{k}$. Докажем, что хотя бы одно из них кратно k. Предположим противное: пусть ни одно из них не кратно k. Поскольку количество ненулевых остатков от деления на k равно $k-1$, какие-то два из k рассматриваемых чисел дают одинаковые остатки при делении на k. Разность этих чисел нацело делится на k и представляет из себя несколько девяток, после которых написано несколько нулей:

$$\underbrace{99\ldots9}_{r+s}-\underbrace{99\ldots9}_{s}=\underbrace{99\ldots9}_{r}\underbrace{00\ldots0}_{s}.$$

Поскольку k взаимно просто с 10, из делимости числа $\underbrace{99\ldots9}_{r}\underbrace{00\ldots0}_{s}$ на k следует, что число $\underbrace{99\ldots9}_{r}$ нацело делится на k.

Это доказательство можно изложить и следующим образом. Рассмотрим числа 1, 10, 10^2, . . . , 10^{k-1}. Ни одно из них не кратно k. Поскольку количество ненулевых остатков от деления на k равно $k-1$, какие-то два из k рассматриваемых чисел дают одинаковые остатки при делении на k. Разность этих чисел $10^{r+s}-10^s$, где $0\leqslant s<r+s<k$, нацело делится на k.

Из делимости числа $10^{r+s}-10^s=10^s(10^r-1)$ на k и из взаимной простоты чисел 10 и k следует, что 10^r-1 кратно k, то есть $10^r-1=kt$, где t — натуральное число. (Например, для $k=7$ можно взять $k=6$; при этом $t=(10^6-1)/7=142\,857$.)

Теорема 2. ***Если m, k — взаимно простые натуральные числа, причем k взаимно просто с 10 и $m<k$, то десятичное представление дроби m/k является чисто периодической дробью. Длина ее наименьшего периода — это такое наименьшее натуральное число r, что 10^r-1 кратно k.***

Доказательство. По лемме 1, $10^r-1=kt$ для некоторых натуральных чисел r и t. Следовательно,

$$\frac{m}{k}=\frac{mt}{kt}=mt\cdot\frac{1}{10^r-1}.$$

Воспользовавшись равенством $\dfrac{1}{10^r-1}=0,(\underbrace{0\ldots0}_{r-1}1)$, получаем:

$$\frac{m}{k}=mt\cdot0,(\underbrace{0\ldots0}_{r-1}1).$$

Поскольку $m<k$, то $mt<kt<10^r$, так что произведение числа mt на число $0,(\underbrace{0\ldots0}_{r-1}1)$ — это периодическая дробь, длина периода которой равна r, а период — десятичная запись числа mt, возможно дополненная слева необходимым количеством нулей.

Нам осталось только понять, почему наименьшему возможному числу r соответствует наименьший возможный период. Это очевидно из правила перевода периодической десятичной дроби в обыкновенную.

Следствие теоремы 2. ***Длиной наименьшего периода десятичного представления несократимой дроби m/n, где $n=2^a5^bk$, a, $b\geqslant0$ и НОД$(k, 10)=1$, является такое наименьшее натуральное число r, что 10^r-1 кратно k.***

Следствие следствия теоремы 2. ***Длина наименьшего периода десятичного представления несократимой дроби m/n зависит только от знаменателя n, а не от числителя m.*** ■

бухарского. К старости он стал слаб глазами и не узнал Насреддина, доставившего ему в свое время немало хлопот.

«Сейчас я поражу мудростью этих невежественных людей, — думал Гуссейн, — они расскажут другим, слух дойдет до эмира, и он станет ценить мою ученость еще больше».

— Я слышал ваш разговор со стражниками у бухарских ворот, — начал он, поглаживая длинную бороду. — Но я не расслышал, сколько стоят арбузы и какова пошлина за провоз арбуза в город. Но я — великий ученый, я знаю наизусть великую книгу ал-Хорезми «Ал-джабр ва-л-мукабала», полную мудрости аллаха и недоступную невежественным умам. Я могу назвать цену арбуза, не ходя на базар и никого не спрашивая!

— Что ж, назови,— сказал Ходжа Насреддин.

— 11 таньга! — провозгласил звездочет. — Я великий мудрец эмира, Гуссейн Гуслия, вы должны признать мою несравненную великую ученость. . .

— Не угадал, о великий Гуссейн Гуслия, — перебил его Насреддин.

— Как это «не угадал»! — возмутился старик. — Ты, конечно, ничего не поймешь, но я все равно скажу. Слушай: если арбуз стоит x таньга, а за его провоз берут y таньга, то цена 19 арбузов на 1 таньга больше, чем налог, который Али уплатил за свои 104 арбуза:

$$19x=104y+1.$$

А ты за свои 17 арбузов отдал 3 арбуза, заплатив меньше, чем надо было, на 1 таньга:

$$3x=17y-1.$$

Теперь, пользуясь наукой несравненного ал-Хорезми, углубляться в которую нет нужды, ибо ее могут понять только истинные бухарцы, я умножаю первое уравнение на 3, второе на 19 и получаю:

$$\begin{aligned}3(104y+1)&=3\cdot19x=\\&=19\cdot3x=19(17y-1),\\312y+3&=323y-19,\\11y&=22,\quad y=2,\\3x&=17\cdot2-1,\quad x=11.\end{aligned}$$

Гуссейн Гуслия размахивал пергаментом с расчетами, презрительно поглядывая на Насреддина.

(*Окончание на с. 387.*)

Рассмотрим следующие разложения:
1/7=0,(142857), 4/7=0,(571428),
2/7=0,(285714), 5/7=0,(714285),
3/7=0,(428571), 6/7=0,(857142).
Периоды этих шести дробей начинаются сразу после запятой и получаются друг из друга циклическим сдвигом.
Возьмем вместо 7 число 41. Очевидно, 1/41=0,(02439). «Прокрутим» период:

$$0,(24390)=10/41,$$
$$0,(43902)=10\cdot 0,(24390)-2=$$
$$=100/41-2=18/41,$$
$$0,(39024)=10\cdot 0,(43902)-4=$$
$$=180/41-4=16/41,$$
$$0,(90243)=10\cdot 0,(39024)-3=$$
$$=160/41-3=37/41.$$

Получили цикл из пяти чисел: 1, 10, 18, 16, 37. Каждое из этих чисел — остаток от деления удесятеренного предыдущего на 41. Начав с 2/41=0,(04878), мы получили бы другой цикл: 20/41=0,(48780), 36/41=0,(87804), 32/41=0,(78048), 33/41=0,(80487). Всего для $p=41$ получаем 8 циклов, по 5 дробей в каждом. ■

Если натуральное число n взаимно просто с 10 и отлично от 1, то все правильные несократимые дроби со знаменателем n разбиваются на циклы по $L(n)$ дробей в каждом цикле. Значит, количество таких дробей кратно $L(n)$. В частности, если p — простое число, то все дроби m/p, где $1\leqslant m<p$, — несократимые, откуда и следует обнаруженная юным Гауссом закономерность. ■

Количество правильных несократимых дробей со знаменателем n обозначают через $\varphi(n)$. Для любого простого числа p, очевидно, $\varphi(p)=p-1$.
Поскольку $\varphi(n)$ правильных несократимых дробей можно разбить на циклы по $L(n)$ дробей в каждом цикле, то $\varphi(n)$ кратно числу $L(n)$.
Следствие. ***Если n — натуральное число, взаимно простое с числом 10, то $10^{\varphi(n)}-1$ кратно n.***
Если бы мы рассматривали не десятичную систему счисления, а систему счисления с основанием a, где a — отличное от единицы натуральное число, то аналогично получили бы утверждение, называемое теоремой Эйлера.
Если n — натуральное число, взаимно простое с целым числом a, где $a>1$, то $a^{\varphi(n)}-1$ кратно n. ■

Обозначим через $L(n)$ длину наименьшего периода десятичного представления дроби $1/n$. В силу следствия теоремы 2, если $n=2^a5^bk$, где $a, b\geqslant 0$ и НОД$(k, 10)=1$, то

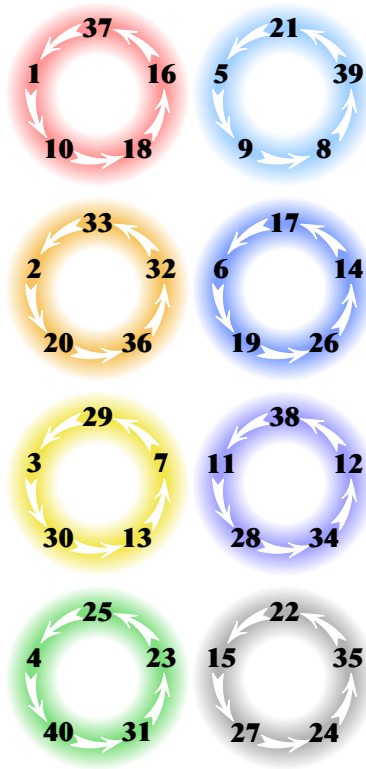
$$L(n)=L(k)=r,$$

где r — это наименьшее натуральное число, для которого 10^r-1 кратно k.

n	1	2	3	4	5	6	7	8	9	10	11	12	13
$L(n)$	1	1	1	1	1	1	6	1	1	1	2	1	6

n	14	15	16	17	18	19	20	21	22	23	24	25	26
$L(n)$	6	1	1	16	1	9	1	6	2	22	1	1	6

Функция L определена на всем множестве натуральных чисел (периодом конечной десятичной дроби естественно считать число 1), но интерес представляют только числа, взаимно простые с числом 10.



Очевидно, если число 10^r-1 кратно каждому из двух взаимно простых натуральных чисел m и n, то 10^r-1 кратно и произведению mn. Следовательно, верна следующая теорема.

Теорема 3. ***Если m, n — взаимно простые натуральные числа, то***

$$L(mn)=\text{НОК}[L(m),L(n)].$$ ■

К. Ф. Гаусс, будучи гимназистом, обращал дроби вида $1/p$, где p — простое число, отличное от 2 и 5, в бесконечные десятичные дроби: в каждом случае он с поразительным терпением ожидал, когда знаки начнут повторяться. Ему хотелось понять, как зависит длина периода такой дроби от p.
Выписывание полного периода, скажем, для $p=47$ — утомительное занятие (46 знаков!). Однако Гаусс не терял надежды и продолжал вычисления: он выписал периоды для всех простых чисел $p<1000$. Главная закономерность, которую он обнаружил, состоит в том, что длина $L(p)$ наименьшего периода такой дроби является делителем числа $p-1$, иногда совпадая с ним. А именно, $L(p)=p-1$ для $p=7$, 17, 23, 47, 59, 61, 97, 109, 113, 131, 149, 167, 179, 181, 193 и некоторых других чисел. (Конечно или бесконечно множество чисел, для которых $L(p)=p-1$, по сей день неизвестно.) ■

p	3	7	11	13	17	19	23	29	31	37	41	43	47
$L(p)$	1	6	2	6	16	18	22	28	15	3	5	21	46

Число 111 делится на 3. Далее, число

$$111\,111\,111 = 111 \cdot 1\,001\,001$$

делится на 9 как произведение двух чисел, каждое из которых делится на 3. (Впрочем, можно было воспользоваться признаком делимости на 9.) Записываемое 27 единицами число

$$111\,111\,111\,111\,111\,111\,111\,111\,111 = 111\,111\,111 \cdot 1\,000\,000\,001\,000\,000\,001$$

делится на 27 как произведение числа, кратного 9, и числа, кратного 3. И вообще, равенство

$$\underbrace{111\ldots11}_{3^n}\underbrace{111\ldots11}_{3^n}\underbrace{111\ldots11}_{3^n} = \underbrace{111\ldots11}_{3^n} \cdot 1\underbrace{00\ldots00}_{3^n-1}1\underbrace{00\ldots00}_{3^n-1}1$$

показывает, что если число, записываемое 3^n единицами, кратно 3^n, то число, записываемое 3^{n+1} единицами, кратно 3^{n+1}.

Таким образом по индукции можно доказать, что число $\underbrace{11\ldots11}_{3^n}$ кратно числу 3^n, то есть число $\underbrace{99\ldots99}_{3^n}$ кратно числу 3^{n+2}, откуда $L(3^{n+2}) \leqslant 3^n$. А если мы еще заметим, что использованные нами числа $1\underbrace{00\ldots00}_{3^n-1}1\underbrace{00\ldots00}_{3^n-1}1$ делятся только на 3, но не на 9, то получим точный результат: $L(3^{n+2}) = 3^n$.

Теорема 4. ***Если p^k — наивысшая степень простого числа p, которой кратно число $10^{L(p)} - 1$, то для любого неотрицательного целого числа m верна формула $L(p^{k+m}) = p^m\, L(p)$. (Например, $L(3^{m+2}) = 3^m$ и $L(7^{m+1}) = 6 \cdot 7^m$.)***

Лемма 2. ***Если $a = 1 + px$, где p — простое число, $p > 2$, x — целое число, то сумма $a^{p-1} + a^{p-2} + \ldots + a + 1$ кратна p, но не кратна p^2.***

Доказательство леммы. Легко понять (рассуждая по индукции или применив бином Ньютона), что при делении на p^2 числа a, a^2, …, a^{p-2} и a^{p-1} дают такие же остатки, как $1 + px$, $1 + 2px$, …, $1 + (p-2)px$ и $1 + (p-1)px$. Следовательно,

$$1 + a + a^2 + \ldots + a^{p-2} + a^{p-1} \equiv$$
$$\equiv 1 + (1 + px) + (1 + 2px) + \ldots + (1 + (p-2)px) + (1 + (p-1)px) =$$
$$= p + px\frac{p(p-1)x}{2} \equiv p \pmod{p^2}.$$

Лемма доказана.

Докажем теорему по индукции. База ($m = 0$) очевидна, а индукционный переход выполняем при помощи леммы. Обозначим $r = p^{k+m}\, L(p)$ и рассмотрим разложение на множители:

$$10^{qr} - 1 = (10^r - 1)((10^r)^{q-1} + (10^r)^{q-2} + \ldots + 10^r + 1),$$

где q — натуральное число. Первый множитель правой части этого равенства кратен p^{k+m} и не кратен p^{k+m+1}.

Каждое слагаемое второго множителя дает остаток 1 при делении на p. Поэтому если q не кратно p, то второй множитель не кратен p. Если же $q = p$, то второй множитель, в силу леммы, кратен p, но не кратен p^2. Таким образом, в разложение числа $10^{pr} - 1$ на простые множители число p входит в $(k + m + 1)$-й степени. Теорема доказана. ■

— Мудрость твоя велика, — ответил Ходжа Насреддин. — Но, как сказал один умный человек, математика — это мельница, которая перемалывает то, что кладут на ее жернова. А ты вместо зерна бросил семена полыни, и доброй муки у тебя не вышло.

— Как это? — возмутился звездочет. — Как можешь ты судить о верности моего решения, ты, не знающий наизусть стихи Корана? Ты всего лишь дехканин, подобный невежеством своему ишаку!

— Скажи мне, о Гуссейн Гуслия, — ответил Насреддин, — зачем Али платить пошлину за те 19 арбузов, которые он отдал стражникам? Ведь их-то он не повез на базар! И я не обязан платить за 3 арбуза! Так что

$$\begin{cases} 19x = (104 - 19)y + 1, \\ 3x = (17 - 3)y - 1. \end{cases}$$

Теперь ты можешь привлечь ту достойную восхищения науку, в которой, как ты думаешь, тебе нет равных, и убедиться, что арбуз стоит 9 таньга.

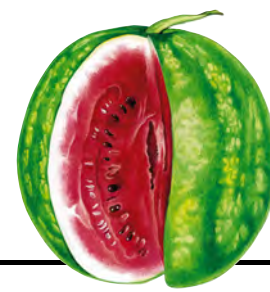
Гуссейн Гуслия погрузился в размышления и обнаружил, что незнакомец прав.

— Возможно, я и погорячился, — неохотно признал он. — Должен сказать, что твои рассуждения достойны самого Ходжи Насреддина. И твое нахальство тоже. Это же надо — заплатить налог не со всего товара, а только с того, что останется после того, как налог будет заплачен!

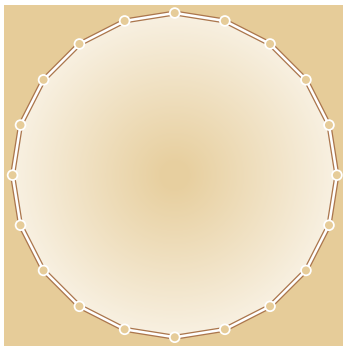
Чайханщик схватился за живот, тихо постанывая: «Ой, умру!» Захохотал Али. Гуссейн Гуслия некоторое вермя смотрел на них с недоумением, потом запричитал:

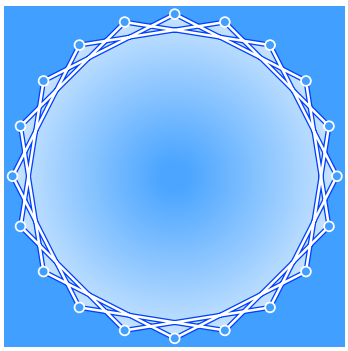
— Так это ты снова явился в Бухару, чтобы посмеяться над моими сединами! Чтоб тебя забрал шайтан, чтоб тебе не знать покоя на том и этом свете, чтоб…

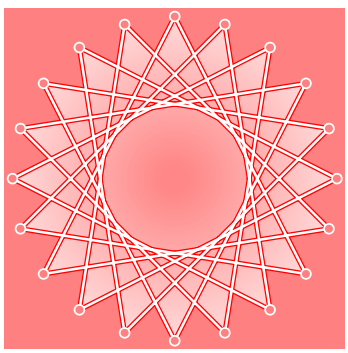
— Да нет, почтенный, — смиренно ответил Ходжа Насреддин. — Я всего лишь приехал продать арбузы. ■

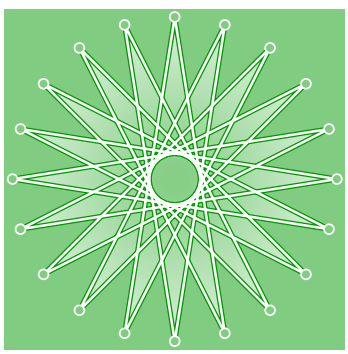


Разделим окружность n точками на n равных частей. Замкнутых n-звенных ломаных с вершинами во всех этих точках, длины всех звеньев которых равны, ровно $\varphi(n)/2$ штук. (Ломаные, получающиеся одна из другой поворотом, не различаем. На рисунках изображены все такие ломаные для $n=20$.) ■









Функция Эйлера

В 1763 г. Л. Эйлер ввел обозначение $\varphi(n)$ *для количества остатков, взаимно простых с* n. *Например,* $\varphi(1)=1$, $\varphi(4)=2$, $\varphi(12)=4$.

Если p простое, то $\varphi(p)=p-1$. Легко вычислить и $\varphi(p^m)$, где m — натуральное число. В самом деле, выпишем все p^m возможных остатков: 0, 1, 2, . . . , p^m-1. Из них кратны p в точности остатки 0, p, $2p$, . . . , p^m-p. Значит,

$$\varphi(p^m)=p^m-p^{m-1}=p^m\left(1-\frac{1}{p}\right). \blacksquare$$

Вычислим $\varphi(1000)$ — количество чисел первой тысячи, которые не кратны ни 2, ни 5. Сначала из 1000 вычтем 500 — именно столько в первой тысяче четных чисел. Не забудем вычесть и $1000:5=200$. А еще заметим, что числа, оканчивающиеся цифрой 0, кратны и 2, и 5. Таких чисел 100 штук; каждое из них мы учитывали оба раза, а надо было — только один раз! Поэтому

$$\varphi(1000)=1000-500-200+100=400.$$

Вычислим $\varphi(300)$. Пусть выписаны все числа от 1 до 300. Вычеркнем 150 четных чисел, 100 чисел, кратных 3, и 60 чисел, кратных 5. Некоторые числа вычеркнуты дважды или даже трижды. Начнем восстанавливать справедливость: к числу $300-150-100-60$ прибавим 50 (количество чисел, кратных $2\cdot3=6$), а также $300:(2\cdot5)=30$ и $20:(3\cdot5)=15$. Но и этого мало: каждое из десяти чисел, кратных $2\cdot3\cdot5=30$, мы сначала трижды выбросили (как кратное 2, 3, 5), а затем трижды возвратили (как кратное 6, 10, 15). Выбросим же их:

$$\varphi(300)=300-150-100-60+50+30+20-10=80. \blacksquare$$

Для числа $n=p_1^{a_1}p_2^{a_2}\cdot\ldots\cdot p_s^{a_s}$, где $p_1, p_2, \ldots, p_s$ — различные простые числа, $a_1, a_2, \ldots, a_s$ — натуральные числа, аналогично получаем при помощи формулы включений-исключений

$$\varphi(n)=n-\frac{n}{p_1}-\frac{n}{p_2}-\ldots-\frac{n}{p_s}+\frac{n}{p_1p_2}+\ldots+(-1)^s\frac{n}{p_1p_2\ldots p_s}=$$
$$=n\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\ldots\left(1-\frac{1}{p_s}\right). \qquad (*)$$
■

Есть и другой способ вывода формулы $(*)$.

Теорема. ***Функция Эйлера мультипликативна, то есть*** $\varphi(mn)=\varphi(m)\varphi(n)$ ***для любых взаимно простых натуральных чисел*** m ***и*** n.

Следствие. $\varphi(n)=\varphi(p_1^{a_1})\varphi(p_2^{a_2})\cdot\ldots\cdot\varphi(p_s^{a_s})=(p_1^{a_1}-p_1^{a_1-1})(p_2^{a_2}-p_2^{a_2-1})\cdot\ldots\cdot(p_s^{a_s}-p_s^{a_s-1})$.

Доказательство теоремы. Рассмотрим числа вида $mx+ny$, где $0\leqslant x<n$ и $0\leqslant y<m$. Запишем остатки от деления их на mn в виде таблицы размером $m\times n$. Например, при $m=21$ и $n=10$ получаем таблицу 1. Все числа таблицы разные. В самом деле, если бы какие-то два числа совпали, то было бы выполнено сравнение

$$mx_1+ny_1\equiv mx_2+ny_2 \pmod{mn},$$

где $0\leqslant x_1, x_2<n$ и $0\leqslant y_1, y_2<m$. Переходя от сравнения по модулю mn к сравнению по модулю m, получаем

$$mx_1+ny_1\equiv mx_2+ny_2 \pmod{m},$$

откуда $ny_1\equiv ny_2 \pmod{m}$. Вследствие взаимной простоты чисел m и n получаем

$y_1 \equiv y_2 \pmod m$. Поскольку $0 \leqslant y_1, y_2 < m$, то $y_1 = y_2$. Аналогично сравнение по модулю n приводит к равенству $x_1 = x_2$.

Итак, все mn чисел таблицы разные. Но возможных остатков от деления на mn ровно столько же, сколько чисел в таблице! Значит, для любого числа $d = 0, 1, \ldots$ $\ldots, mn-1$ существует и единственна такая пара целых чисел x, y, что $0 \leqslant x < n$, $0 \leqslant y < m$ и $d \equiv mx + ny \pmod{mn}$.

В таблице 1 четные числа вместе с числами, кратными 5, образуют шесть столбцов, а числа, кратные 3 или 7, — девять строк. Это не случайно:

$$\text{НОД}(mx+ny, m) = \text{НОД}(ny, m) = \text{НОД}(y, m),$$

аналогично, $\text{НОД}(mx+ny, n) = \text{НОД}(x, n)$. По этой причине в рассматриваемой таблице числа, взаимно простые с m, расположены в $\varphi(m)$ строках (тех, где число y взаимно просто с m); а числа, взаимно простые с n, образуют $\varphi(n)$ столбцов.

Теперь доказательство теоремы не составляет труда: чтобы d было взаимно просто с mn, необходимо и достаточно, чтобы d было взаимно просто с числами m и n. Такие числа d лежат на пересечении $\varphi(m)$ строк с $\varphi(n)$ столбцами; в таблице 1 у этих чисел светлый фон. Всего получаем «решетку» из $\varphi(m)\,\varphi(n)$ чисел, что и требовалось доказать. ■

y \ x	0	1	2	3	4	5	6	7	8	9
0	0	21	42	63	84	105	126	147	168	189
1	10	31	52	73	94	115	136	157	178	199
2	20	41	62	83	104	125	146	167	188	209
3	30	51	72	93	114	135	156	177	198	9
4	40	61	82	103	124	145	166	187	208	19
5	50	71	92	113	134	155	176	197	8	29
6	60	81	102	123	144	165	186	207	18	39
7	70	91	112	133	154	175	196	7	28	49
8	80	101	122	143	164	185	206	17	38	59
9	90	111	132	153	174	195	6	27	48	69
10	100	121	142	163	184	205	16	37	58	79
11	110	131	152	173	194	5	26	47	68	89
12	120	141	162	183	204	15	36	57	78	99
13	130	151	172	193	4	25	46	67	88	109
14	140	161	182	203	14	35	56	77	98	119
15	150	171	192	3	24	45	66	87	108	129
16	160	181	202	13	34	55	76	97	118	139
17	170	191	2	23	44	65	86	107	128	149
18	180	201	12	33	54	75	96	117	138	159
19	190	1	22	43	64	85	106	127	148	169
20	200	11	32	53	74	95	116	137	158	179

Таблица 1.

Третий способ доказательства тоже использует таблицу из m строк и n столбцов (табл. 2). Очевидно, числа, взаимно простые с n, заполняют собой $\varphi(n)$ столбцов таблицы. Остатки от деления на m всех m чисел любого столбца таблицы различны. В каждом столбце присутствует ровно $\varphi(m)$ чисел, взаимно простых с m.

Таблица 2.

0	1	2	. . .	$n-1$
n	$n+1$	$n+2$	. . .	$2n-1$
$2n$	$2n+1$	$2n+2$	. . .	$3n-1$
. . . . .				
$(m-1)n$	$(m-1)n+1$	$(m-1)n+2$	. . .	$mn-1$

Наконец, поскольку число взаимно просто с mn тогда и только тогда, когда оно взаимно просто с n (такие числа лежат в $\varphi(n)$ столбцах) и взаимно просто с m (в каждом столбце таких чисел $\varphi(m)$), то $\varphi(mn) = \varphi(m)\,\varphi(n)$. ■

Сумма значений функции Эйлера. Рассмотрим 100 дробей: $1/100, 2/100, \ldots, 100/100$. Если каждую из них привести к несократимому виду, то получим $\varphi(100) = 40$ дробей со знаменателем 100, $\varphi(50) = 20$ дробей со знаменателем 50, и так далее: для каждого делителя d числа 100 получим $\varphi(d)$ дробей со знаменателем d. (Почему? Потому что $\varphi(d)$ — это количество несократимых правильных дробей со знаменателем d.) Мы получили равенство

$$100 = \varphi(100) + \varphi(50) + \varphi(25) + \varphi(20) + \varphi(10) + \varphi(5) + \varphi(4) + \varphi(2) + \varphi(1).$$

Если бы мы рассмотрели не дроби со знаменателем 100, а дроби со знаменателем n, то точно так же доказали бы, что для любого натурального числа n сумма значений функции Эйлера $\varphi(d)$ по всем делителям d числа n равна n:

$$n = \sum_{n \vdots d} \varphi(d). \qquad (**)$$

■

Формулу $(*)$ можно записать в виде

$$\varphi(n) = \sum_{n \vdots d} \mu(d)\,\frac{n}{d}, \qquad (***)$$

где μ — функция Мёбиуса, равная $(-1)^k$ для произведения любых k различных простых чисел, 1 для 1 и 0 для любого натурального числа, делящегося на квадрат простого числа.

Формулы $(**)$ и $(***)$ связаны теснее, чем кажется на первый взгляд: если не только для функции Эйлера, но для любой мультипликативной функции f определить функцию g формулой

$$g(n) = \sum_{n \vdots d} f(d),$$

то для любого натурального m будет верна формула обращения Мёбиуса

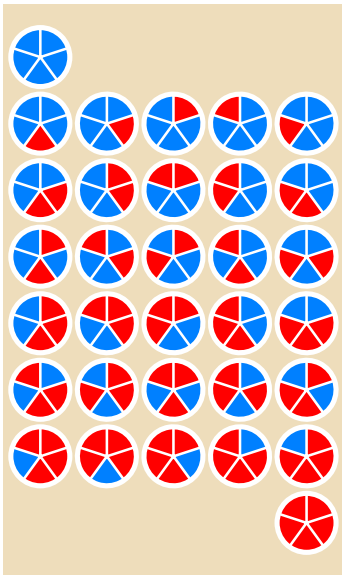
$$f(n) = \sum_{n \vdots d} \mu(d)\, g\left(\frac{n}{d}\right).$$

Доказательство основано на равенстве

$$\sum_{n \vdots d} \mu(d) = \begin{cases} 1, & \text{если } n = 1, \\ 0, & \text{если } n > 1. \end{cases}$$

■

На рисунке изображены все 32 способа раскраски в два цвета круга, который разделен на 5 равных секторов. Среди них выделяются два способа — когда весь круг синий и когда он весь красный. А остальные разбиты на 6 групп по 5 раскрасок, получающихся одна из другой поворотом.



Спросим себя, сколькими способами можно раскрасить a разными красками круг, разбитый на p одинаковых секторов, где p — простое число. (Каждый сектор окрашиваем одной краской; необязательно использовать все краски; две раскраски, совпадающие при повороте круга, считаем одинаковыми.)
Очевидно, можно все секторы покрасить одной краской. Таких способов столько же, сколько красок, то есть a способов.
А вот из любой другой раскраски поворотами можно получить p разных раскрасок (считая и саму эту раскраску: она получается поворотом на 0°). Значит, ответ таков:

$$a+\frac{a^p-a}{p}.$$

Поскольку количество способов не бывает дробным, число a^p-a обязано делиться на p.
Рассуждая аналогично, для случая p^2 секторов получаем $a+{}$ $+(a^p-a)/p+(a^{p^2}-a^p)/p^2$ вариантов, а для pq секторов, где $p\neq q$ — простые числа, — $a+(a^p-a)/p+{}$ $+(a^q-a)/q+(a^{pq}-a^p-a^q+a)/(pq)$ способов. ■

Малая теорема Ферма

Теорема, открытая советником парламента г. Тулуза (Франция) П. Ферма (1601—1665) в 1640 г., формулируется очень коротко: если p — простое число, a — целое число, то a^p-a кратно p. Сразу и не видно, почему такое скромное с виду утверждение столь важно. Тем не менее оно заслуживает величайшего внимания.

Частные случаи. Из любых двух последовательных целых чисел a и $a+1$ одно четное, а другое нечетное. Поэтому произведение $a(a+1)=a^2+a$ четно при любом целом a.
Делимость числа a^2+a на 2 можно доказать и по-другому, разобрав два случая:
- если a четно, то a^2 тоже четно, а сумма двух четных чисел a и a^2 четна;
- если a нечетно, то a^2 тоже нечетно, а сумма двух нечетных чисел a и a^2 четна.

Вот так доказывают замечательное свойство многочлена a^2+a. Впрочем, при $p=2$ в малой теореме Ферма фигурирует другой многочлен: $a^2-a=(a-1)a$. Все его значения в целых точках — четные числа (докажите!).
Многочлен a^3-a тоже легко разложить на множители:

$$a^3-a=a(a^2-1)=a(a-1)(a+1).$$

Получили произведение трех последовательных целых чисел: $a-1$, a и $a+1$. Как мы уже знаем, оно четно. Поскольку из любых трех последовательных чисел одно кратно 3, произведение кратно 3 (и, значит, даже кратно 6). ■

Многочлен a^4-a при $a=2$ и $a=3$ принимает значения $2^4-2=14$ и $3^4-3=78$. Конечно, эти значения четны, но никакого общего делителя, кроме 2 (и 1), у них нет. Не повезло! Впрочем, число 4 составное, а малая теорема Ферма говорит только о многочленах вида a^p-a, где p — простое число. ■

Пусть $p=5$. Вычислим несколько значений многочлена a^5-a. При $a=\pm1$ и при $a=0$ получаем ноль. Смотрим дальше: $2^5-2=30$, $3^5-3=240$, $4^5-4=1020$, $5^5-5=3120$, $6^5-6=7770$, . . . Все эти значения кратны числу 30. Поскольку $30=2\cdot3\cdot5$, доказательство делимости на 30 распадается на три части: во-первых, надо доказать, что a^5-a кратно 2; во-вторых, кратно 3; в-третьих, — 5.
Первая часть очевидна: числа a^5 и a либо оба четны, либо оба нечетны. Не вызывает затруднений и вторая часть:

$$a^5-a=a(a^4-1)=a(a^2-1)(a^2+1)=(a-1)a(a+1)(a^2+1),$$

произведение трех последовательных чисел всегда кратно 3.
Чуть сложнее третья часть. Нет, конечно, из пяти последовательных целых чисел обязательно одно кратно 5, так что произведение $(a-2)(a-1)a(a+1)\times{}$ $\times(a+2)$ кратно 5. Но $a^2+1\neq(a-2)(a+2)$.
Как же быть? Самый бесхитростный способ — перебрать все подряд остатки от деления на 5: любое целое число при делении на 5 дает в остатке 0, 1, 2, 3 или 4. Если остаток равен 0, то кратен 5 второй множитель произведения $(a-1)a(a+1)(a^2+1)$. Если остаток равен 1 или 4, то кратен 5 первый или третий множитель. Если же остаток равен 2 или 3, то в дело вступает четвертый множитель. (Для тех, кто еще не привык работать с остатками, объясним: если $a=5b+2$, то есть если a дает остаток 2 при делении на 5, то $a^2+1=$ $=(5b+2)^2+1=5(5b^2+4b+1)$. Аналогично можно рассмотреть случай $a=5b+3$.)

Есть и другой способ:

$$a^2+1=(a-2)(a+2)+5,$$

значит, если нас интересуют только остатки от деления на 5, то a^2+1 можно-таки заменить на $(a-2)(a+2)$. Формулой это записывают так:

$$a^2+1\equiv(a-2)(a+2)\pmod 5.$$

Предложенное в 1801 г. К. Ф. Гауссом обозначение «$\equiv$» («сравнимо») еще не раз будет использовано нами. По определению, a сравнимо с b по модулю n, если $a-b$ кратно n, то есть $a-b=kn$, где k — целое число. ■

Обозначение $a\equiv b\pmod n$ оказалось удачным потому, что свойства сравнений похожи на свойства обычных равенств. Сравнения можно складывать: если $a\equiv b\pmod n$ и $c\equiv d\pmod n$, то $a+c\equiv b+d\pmod n$. В самом деле, по определению, $a=b+kn$ и $c=d+ln$, где k, l — целые числа. Значит,

$$a+c=(b+kn)+(d+ln)=b+d+(k+l)n.$$

Аналогично формулы

$$a-c=(b+kn)-(d+ln)=b-d+(k-l)n,$$
$$ac=(b+kn)(d+ln)=bd+knd+bln+kln^2=bd+(kd+bl+kln)n$$

позволяют утверждать, что сравнения можно вычитать и умножать. Коли можно умножать, то можно и возводить в степень: если $a\equiv b\pmod n$, то для любого натурального числа m верно сравнение $a^m\equiv b^m\pmod n$.
Сокращать сравнения надо с осторожностью: $6\equiv 36\pmod{10}$, но $1\not\equiv 6\pmod{10}$. ■

Продолжим изучение многочленов вида a^p-a: докажем, что при любом целом a число a^7-a кратно 7. Как всегда, можно рассмотреть все 7 остатков от деления на 7, а именно, $0^7-0=0$, $1^7-1=0$, $2^7-2=126=7\cdot 18$, ..., $6^7-6=279\,930=7\cdot 39\,990$. (Можно и чуточку сэкономить: поскольку любое целое число представимо в виде $a=7b$, $7b\pm 1$, $7b\pm 2$ или $7b\pm 3$, очевидно, при проверке малой теоремы Ферма для $p=7$ можно ограничиться рассмотрением случаев $a=0$, 1, 2 и 3.)
Но бездумная проверка не может научить нас ничему интересному. Лучше рассмотрим разложение на множители:

$$a^7-a=a(a^6-1)=a(a^3-1)(a^3+1)=a(a-1)(a^2+a+1)(a+1)(a^2-a+1).$$

Поскольку $a^2+a+1=(a^2+a-6)+7\equiv a^2+a-6=(a-2)(a+3)$ и $a^2-a+1\equiv$ $\equiv a^2-a-6=(a+2)(a-3)\pmod 7$, имеем

$$a^7-a\equiv a(a-1)(a-2)(a+3)(a+1)(a+2)(a-3)\pmod 7.$$

Произведение семи последовательных целых чисел кратно 7.
Теперь рассмотрим число $p=11$. Очевидно,

$$a^{11}-a=a(a^{10}-1)=a(a^5-1)(a^5+1)=$$
$$=a(a-1)(a^4+a^3+a^2+a+1)(a+1)(a^4-a^3+a^2-a+1).$$

Тут не так-то просто догадаться, как быть дальше. Но полный перебор всех 11 остатков все еще возможен. И когда мы его выполним, окажется, что значения многочлена $a^4+a^3+a^2+a+1$ кратны 11 при $a\equiv 3, 4, 5$ или $9\pmod{11}$, а значения многочлена $a^4-a^3+a^2-a+1$ кратны 11 при $a\equiv 2$, 6, 7 или $8\pmod{11}$.
Между прочим, если мы раскроем скобки в произведении $(a-3)(a-4)\times$ $\times(a-5)(a-9)$, то получим

$$(a^2-7a+12)(a^2-14a+45)\equiv(a^2+4a+1)(a^2-3a+1)=a^4+a^3-10a^2+a+1\equiv$$
$$\equiv a^4+a^3+a^2+a+1\pmod{11}.$$

Аналогично можно проверить, что $(a-2)(a-6)(a-7)(a-8)\equiv a^4-a^3+a^2-a+$ $+1\pmod{11}$.

Докажите следующие утверждения. (*Указания на с. 393.*)

1. Наибольший общий делитель чисел вида a^7-a равен 42, а чисел вида a^9-a — 30. (Заметьте: 30 не кратно 9, впрочем, число 9 не простое, а составное.)
2. Произведение любых четырех последовательных целых чисел кратно 24.
3. Произведение любых пяти последовательных целых чисел кратно 120.
4. a^5-5a^3+4a при всяком целом a кратно 120.
5. Для любого натурального a число a^5 оканчивается на ту же цифру, что и a.
6. m^5n-mn^5 кратно 30 при любых целых m и n.
7. Если число k не кратно ни 2, ни 3, ни 5, то k^4-1 кратно 240.
8. Если n четно, то наибольший общий делитель чисел a^n-a, где a пробегает множество целых чисел, равен 2.
9. $m^{61}n-mn^{61}$ кратно 56 786 730 при любых целых m и n.
10. Если у шестизначного числа, кратного 7, стереть первую цифру и записать ее после последней, то полученное число тоже кратно 7. (Например, из 632 387 и 200 004 получаем 323 876 и 42, которые тоже кратны 7.)
11. Если p — простое число, отличное от 2, 3 и 5, то число, записанное $p-1$ единицами, кратно p. (Например, 111 111 кратно 7.)
12. Если k не кратно 3, то k^3 при делении на 9 дает остаток 1 или 8, а k^{81} при делении на 243 дает остаток 1 или 242.
13. Если $a^3+b^3+c^3$ кратно 9, то хотя бы одно из целых чисел a, b, c кратно 3.
14. Сумма квадратов трех целых чисел кратна 7 тогда и только тогда, когда сумма четвертых степеней этих чисел кратна 7.
15. Десятичная запись числа 7^{9999} оканчивается цифрами 143.
16. Если целое число a взаимно просто с натуральным числом $n>1$, то сравнение $ax\equiv b\pmod n$ равносильно сравнению $x\equiv$ $\equiv a^{\varphi(n)-1}b\pmod n$.
17. Если n — нечетное натуральное число, то $2^{n!}-1$ кратно n.
18. Сумма $1^n+2^n+\ldots+(n-1)^n$ кратна n тогда и только тогда, когда n нечетно.
19. Для любого натурального числа s существует кратное ему натуральное число n, сумма цифр которого равна s. ■

Малую теорему Ферма легко доказать по индукции при помощи бинома Ньютона. Мы сделаем это для натуральных чисел a, оставив случай отрицательных чисел читателю.

Пусть сначала $p=3$. База индукции: $1^3-1=0$ кратно 3. Переход: если для некоторого числа a уже доказали, что a^3-a кратно 3, то $(a+1)^3-(a+1)=a^3+3a^2+3a+1-(a+1)\equiv a^3+1-a-1=a^3-a\equiv 0 \pmod 3$.

Аналогично для $p=5$: база очевидна ($1^5-1\equiv 0 \pmod 5$), а для перехода используем формулу $(a+1)^5=a^5+5a^4+10a^3+10a^2+5a+1$. Видите, коэффициенты при a^4, a^3, a^2 и a кратны 5. Поэтому

$$(a+1)^5\equiv a^5+1 \pmod 5,$$

откуда и следует возможность индукционного перехода:

$$(a+1)^5-(a+1)\equiv a^5+1-a-1=a^5-a \pmod 5.$$

Займемся общим случаем. Формула бинома имеет вид

$$(a+1)^p=a^p+pa^{p-1}+\frac{p(p-1)}{2}a^{p-2}+\frac{p(p-1)(p-2)}{3!}a^{p-3}+\ldots\ldots+\frac{p(p-1)}{2}a^2+pa+1.$$

Биномиальные коэффициенты $C_p^1=p$, $C_p^2=\frac{p(p-1)}{2}$, …, $C_p^k=\frac{p(p-1)\ldots(p-k+1)}{k!}$, …, $C_p^{p-1}=p$ кратны простому числу p. Поэтому $(a+1)^p\equiv a^p+1 \pmod p$, что и требовалось:

$$(a+1)^p-(a+1)\equiv a^p+1-a-1=a^p-a \pmod p.$$ ■

Использованное в доказательстве свойство треугольника Паскаля выполнено только для простых чисел: если p — простой делитель составного числа n, то число $C_n^p/n=(n-1)(n-2)\ldots\ldots(n-p+1)/p!$ не целое, поскольку делимое не кратно p. ■

Таблица 1.

×	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Таблица 2.

×	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	2	4	6	8	10	1	3	5	7	9
3	3	6	9	1	4	7	10	2	5	8
4	4	8	1	5	9	2	6	10	3	7
5	5	10	4	9	3	8	2	7	1	6
6	6	1	7	2	8	3	9	4	10	5
7	7	3	10	6	2	9	5	1	8	4
8	8	5	2	10	7	4	1	9	6	3
9	9	7	5	3	1	10	8	6	4	2
10	10	9	8	7	6	5	4	3	2	1

Что дальше? При $p=13$, если действовать нашим способом, придется возводить в двенадцатую степень числа от 1 до 12 или раскрывать скобки в произведении тринадцати множителей: $a-6, a-5, \ldots, a+5, a+6$. Заниматься этим не хочется, даже если ограничиться возведением в степень чисел 1, 2, 3, 4, 5, 6 или перемножать «всего лишь» шесть скобок: $(a^2-1)(a^2-4)(a^2-9)(a^2-16)(a^2-25)(a^2-36)$. Чем больше p, тем больше вариантов надо перебирать. Поэтому мы прекратим разбор частных случаев и перейдем к доказательству малой теоремы Ферма, которое охватывает сразу все простые числа p. ■

Выпишем в строчку числа 1, 2, 3, …, $p-1$, домножим каждое из них на k, где k не кратно p, и рассмотрим остатки от деления на p. Например, при $p=19$ и $k=4$ получим:

a	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$4a$	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72
$4a \bmod 19$	4	8	12	16	1	5	9	13	17	2	6	10	14	18	3	7	11	15

В нижней строке таблицы — те же самые числа, что и в верхней, только они расположены в другом порядке! Оказывается, это общий закон: не только при $p=19$ и $k=4$, но при любом простом p и не кратном p целом числе k всегда получатся те же самые числа 1, 2, 3, …, $p-1$, возможно, записанные в некотором другом порядке.

Почему? Ну, во-первых, в нижней строке не может появиться 0, ибо произведение не кратных простому числу p чисел a и k не может быть кратно p. Во-вторых, все числа нижней строки разные (это легко доказать «от противного»: если бы числа ak и bk давали при делении на p одинаковые остатки, то разность $ak-bk=(a-b)k$ была бы кратна p, что невозможно, поскольку $a-b$ не кратно p). Этих двух замечаний достаточно: ненулевых остатков от деления на p существует $p-1$ штук, все они вынуждены по одному разу появиться в нижней строке таблицы.

Как вы помните, малая теорема Ферма утверждает, что при любом целом k и простом p число $k^p-k=k(k^{p-1}-1)$ кратно p. Значит, для чисел k, не кратных p, теорему можно формулировать следующим образом:

Теорема 1. ***Если целое число k не кратно простому числу p, то k^{p-1} дает при делении на p остаток 1.***

Доказательство. Поскольку остатки от деления на p чисел k, $2k$, $3k$, … …, $(p-1)k$ — это (с точностью до перестановки) числа 1, 2, 3, …, $p-1$, то

$$k\cdot 2k\cdot 3k\cdot\ldots\cdot(p-1)k\equiv 1\cdot 2\cdot 3\cdot\ldots\cdot(p-1) \pmod p,$$

откуда $k^{p-1}(p-1)!\equiv(p-1)! \pmod p$. Сократив на $(p-1)!$, получим желаемое:

$$k^{p-1}\equiv 1 \pmod p.$$

А тот, кто не знает, почему сравнения можно сокращать (на число, взаимно простое с модулем), может рассуждать следующим образом: поскольку произведение $(k^{p-1}-1)\times(p-1)!$ кратно p, а число $(p-1)!$ не кратно p, то число $k^{p-1}-1$ кратно простому числу p. ■

Таблицы умножения. Рассмотрим все $n-1$ разных ненулевых остатков от деления на n. Составим таблицу умножения, написав на пересечении a-го столбца и b-й строки остаток от деления на n произведения ab. Например, при $n=5$ или 11 получаем таблицы 1 и 2 соответственно.

Поскольку в обоих примерах число n простое, в каждой строке, как и в каждом столбце, возникает некоторая перестановка чисел $1, 2, \ldots, n-1$. Если же рассмотреть составное число, то в таблице обязательно встретится нуль, как, например, при $n=4$ (табл. 3). Не лучше ситуация и при $n=12$ (табл. 4). Опять в некоторых строках есть нули! И вообще, при любом составном числе $n=ab$, где $1<a, b<n$, на пересечении a-й строки и b-го столбца стоит остаток от деления ab на n, то есть 0. ■

Таблица 3.

×	1	**2**	3
1	1	2	3
2	2	**0**	2
3	3	2	1

Таблица 4.

×	1	**2**	**3**	**4**	5	**6**	7	**8**	**9**	**10**	11
1	1	2	3	4	5	6	7	8	9	10	11
2	2	4	6	8	10	**0**	2	4	6	8	10
3	3	6	9	**0**	3	6	9	**0**	3	6	9
4	4	8	**0**	4	8	**0**	4	8	**0**	4	8
5	5	10	3	8	1	6	11	4	9	2	7
6	6	**0**	6	**0**	6	**0**	6	**0**	6	**0**	6
7	7	2	9	4	11	6	1	8	3	10	5
8	8	4	**0**	8	4	**0**	8	4	**0**	8	4
9	9	6	3	**0**	9	6	3	**0**	9	6	3
10	10	8	6	4	2	**0**	10	8	6	4	2
11	11	10	9	8	7	6	5	4	3	2	1

Итак, если n составное,

то имеются делители нуля — ненулевые остатки a и b, произведение ab которых кратно n, иными словами, равно нулю по модулю n. Но даже при составном n в некоторых строках таблицы умножения нет нулей. В таблице 3 таковы первая и третья строки, а в таблице 4 — первая, пятая, седьмая и одиннадцатая. Подумав немного, можно понять, что нули присутствуют в тех и только тех строках, номера которых имеют с числом n общий делитель, отличный от 1 (докажите!). Давайте же вычеркнем из таблицы все такие строки и столбцы. (Если n — простое число, то вычеркивать ничего не придется.) При $n=4$ получим таблицу 5 из двух строк и столбцов. А при $n=12$ останется таблица размером 4×4 (табл. 6). ■

Таблица 5.

×	1	3
1	1	3
3	3	1

Таблица 6.

×	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

Теорема Эйлера.

Чтобы обобщить малую теорему Ферма на случай составного числа n, оставим в таблице умножения только те строки и столбцы, в которых нет нулей, то есть рассмотрим взаимно простые с n остатки от деления на n. В новой таблице строки (и столбцы) отличаются друг от друга лишь порядком, в котором расположены числа. Другими словами, если мы для натурального числа n выпишем все остатки $a_1, a_2, \ldots, a_r$, взаимно простые с n, и домножим каждый из них на взаимно простое с n число k, то получим числа $ka_1, ka_2, \ldots, ka_r$, которые тоже взаимно просты с n и дают разные остатки при делении на n (докажите!).

Итак, строка остатков от деления на n чисел $ka_1, ka_2, \ldots, ka_r$ может отличаться от строки $a_1, a_2, \ldots, a_r$ только порядком расположения чисел. Поэтому точно так же, как для простого p, для составного n имеем:

$$ka_1 ka_2 \ldots ka_r \equiv a_1 a_2 \ldots a_r \pmod{n},$$

откуда

$$(k^r-1)a_1 a_2 \ldots a_r \equiv 0 \pmod{n}.$$

Значит, произведение $(k^r-1)a_1 a_2 \ldots a_r$ кратно n. Поскольку числа $a_1, a_2, \ldots, a_r$ взаимно просты с n, то k^r-1 кратно n. Если n — простое число, то $r=n-1$ и получаем в точности утверждение малой теоремы Ферма. В общем же случае приходим к теореме Эйлера:

Теорема 2. ***Если k — целое число, взаимно простое с натуральным числом n, то k^r-1 кратно n, где r — количество взаимно простых с n натуральных чисел, не превосходящих n, то есть $r=\varphi(n)$.*** ■

Указания к доказательствам утверждений (*см. с. 391*).

2. В произведении четырех последовательных целых чисел обязательно есть множитель, кратный 4. Кроме него, есть еще хотя бы один четный множитель.

4. $a^5-5a^3+4a=a(a^2-1)(a^2-4)=(a-2)(a-1)a(a+1)(a+2)$.

8. Рассмотрите $a=-1$.

9. $56\,786\,730=2\cdot3\cdot5\cdot7\cdot11\cdot13\cdot31\cdot61$.

11. $\underbrace{1\ldots1}_{p-1}=\underbrace{9\ldots9}_{p-1}/9=(10^{p-1}-1)/9$ кратно p по теореме Ферма.

13. Поскольку $\varphi(9)=6$, для любого не кратного 3 числа k, по теореме Эйлера, k^6-1 кратно 9. Далее, $k^6-1=(k^3-1)(k^3+1)$, причем числа k^3-1 и k^3+1 отличаются на 2 и потому не могут одновременно быть кратны 3.

14. Поскольку $7^4\equiv1\pmod{10}$, последняя цифра числа 7^k определяется остатком от деления k на 4. Далее, $7^{2m+1}=(8-1)^{2m+1}\equiv(-1)^{2m+1}=-1 \pmod 4$.

15. Применим теорему Эйлера: $7^{400}\equiv1 \pmod{1000}$. Следовательно, $7^{10\,000}=(7^{400})^{25}\equiv1 \pmod{1000}$. Так как $7^{10\,000}$ оканчивается цифрами 001, последняя цифра числа 7^{9999} равна 3. Значит, из разряда единиц в разряд десятков при умножении 7^{9999} на 7 переносится 2. Поэтому предпоследняя цифра равна 4, из разряда десятков в разряд сотен переносится 3, а в предпредпоследнем разряде находится цифра 1.

17. $n!$ делится на $\varphi(n)$.

18. Если n нечетно, то $k^n+(n-k)^n\equiv k^n+(-k)^n=0 \pmod n$. Сгруппировав первое слагаемое рассматриваемой суммы с последним, второе с предпоследним и так далее, получаем, что она кратна n. Если же n четно, пусть 2^s — наивысшая степень двойки, на которую делится n. Для любого четного k, очевидно, $k^n\equiv0 \pmod{2^s}$; для любого нечетного, по теореме Эйлера, — $k^n=(k^{n/2^{s-1}})^{2^{s-1}}\equiv1 \pmod{2^s}$. Имеем: $1^n+2^n+\ldots+(n-1)^n\equiv n/2\not\equiv0 \pmod{2^s}$, следовательно, сумма не кратна числу n.

19. Пусть $s=2^a\cdot5^b\cdot t$, где a, b — целые неотрицательные числа, t — натуральное число, не кратное ни 2, ни 5. Существует такое натуральное число r, что $10^r\equiv1 \pmod t$. Пусть $n=10^{\max(a,b)}\times(1+10^r+10^{2r}+\ldots+10^{(s-1)r})$. Очевидно, сумма цифр числа n равна s. Поскольку $10^{\max(a,b)}$ делится на $2^a\cdot5^b$ и $10^r+10^{2r}+\ldots+10^{sr}\equiv s\equiv0 \pmod t$, число n кратно s. ■

На вопрос, что он написал в шифровке, Штирлиц ответил: «Не помню. Теперь это знает только Центр». ■

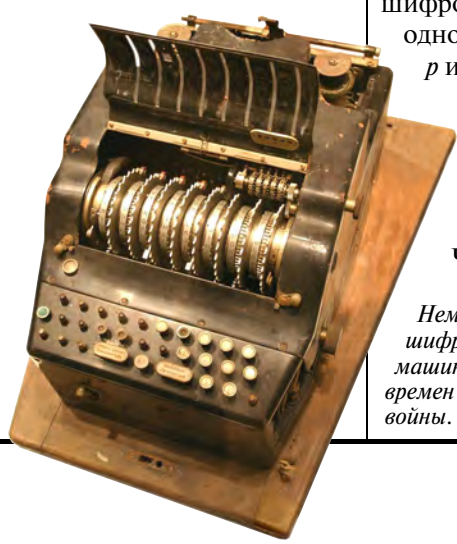
Как возводить в большую степень? Чтобы возвести число в 9007-ю степень, достаточно выполнить 9006 умножений. Но можно обойтись и меньшим числом операций: вычислить x^2, $(x^2)^2=x^4$, $(x^4)^2=x^8$, . . . , $(x^{2048})^2=x^{4096}$, наконец, $(x^{4096})^2=x^{8192}$, и воспользоваться формулой

$$x^{9007}=x\cdot x^2\cdot x^4\cdot x^8\cdot x^{32}\cdot x^{256}\cdot x^{512}\cdot x^{8192},$$

которая основана на том, что в двоичной системе счисления число 9007 имеет вид

$$9007_{10}=10\ 0011\ 0010\ 1111_2.$$

Мы разложили 9007 в сумму $1+2+4+8+32+256+512+8192$ и смогли очень сильно сэкономить: обошлись 13-ю возведениями в квадрат на первом этапе вычислений и 7-ю умножениями на втором. Всего 20 умножений вместо 9006. Огромная экономия! (Для придирчивого читателя отметим, что выше можно говорить не об умножениях, а об умножениях по модулю pq: дабы количество цифр не росло катастрофически, лучше всякий раз не только перемножать, но и брать остаток от деления на pq. Но сейчас разговор не об этом.) Преимущества изложенного метода тем нагляднее, чем больше показатель степени. Например, если показатель степени состоит не из четырех цифр, как 9007, а из нескольких десятков или сотен цифр, то наивный способ не то что утомителен, а неосуществим ни на каких, даже самых мощных компьютерах. А основанный на двоичной системе — работает и в такой ситуации! ■



Немецкая шифровальная машина «Enigma» времен Второй мировой войны.

Шифры с открытым ключом

Вообразите, что вам нужно получить зашифрованное сообщение от вашего друга, но вы с ним не договорились заранее, каким шифром будете пользоваться. Как быть? Существует ли метод шифрования, который можно сообщить всему миру (и друзьям, и врагам), но это не даст врагам возможности расшифровать сообщение? Это был бы замечательный шифр: в отличие от старых шифров, в которых главный секрет — ключ, знание которого позволяет как зашифровывать, так и расшифровывать сообщения, в новом шифре ключи для шифрования и расшифрования разные, так что вы можете рассекретить, например, способ шифрования, оставив сообщения секретными.

Шифр с открытым ключом, скорее всего, уже изобретен! В 1978 г. три математика — Р. Ривест, А. Шамир и Л. Адлеман — зашифровали некоторую английскую фразу и пообещали награду первому, кто расшифрует сообщение

$y=96869613754622061477140922254355882905759991124574319874695120930816298225145708356931476622883989628013391990551829945157815154.$

Способ шифрования они подробно разъяснили. Сначала фразу бесхитростно (a=01, b=02, c=03, . . . , z=26, пробел=00) записали в виде последовательности цифр. Получилось некоторое 78-значное число x. Затем взяли 64-значное простое число p и 65-значное простое число q. Перемножили их (не вручную, разумеется, а на компьютере):

$pq=114381625757888867669325779976146612010218296721242362562561842935706935245733897830597123563958705058989075147599290026879543541.$

Теперь — главное:

$$y\equiv x^{9007}\pmod{pq}.$$

Понимаете? Они опубликовали и произведение pq, и число 9007, и сам метод шифрования (и, разумеется, число y). Было даже сказано, что из чисел p и q одно 64-значное, а другое 65-значное. В секрете остались только сами числа p и q. Требовалось найти x.

В 1994 г. Д. Аткинс, М. Грэфф, А. Ленстра и П. Лейланд расшифровали эту фразу: «The magic words are squeamish ossifrage». (Приведем перевод двух последних слов этой, по всей видимости, бессмысленной фразы: squeamish — «брезгливый, привередливый, обидчивый»; ossifrage — «скопа».)

Числа p и q оказались равны

$p=34905295108476509491478496199038981334177646384933878439908205 77,$

$q=32769132993266709549961988190834461413177642967992942539798288533.$

Это разложение 129-значного числа на множители нашли при помощи так называемого метода квадратичного решета. Выполнение вычислений потребовало колоссальных ресурсов. В работе, возглавлявшейся четырьмя авторами проекта и продолжавшейся после предварительной теоретической подготовки примерно 220 дней, на добровольных началах участвовали около 600 человек и примерно 1600 компьютеров, объединенных сетью Internet.

К сожалению, рассказ о методе квадратичного решета здесь совершенно невозможен: нужно было бы изложить слишком много предварительного материала. Потому мы лишь обсудим основную идею системы RSA (по первым буквам фамилий авторов: Rivest, Shamir, Adleman). ■

Идея алгоритма RSA.

Во-первых, зная p и q, можно найти

$$\varphi(pq)=(p-1)(q-1).$$

Во-вторых (и это главное!), если

$$fg=1+k\varphi(pq),$$

где f, g, k — натуральные числа, то для любого числа x, взаимно простого с pq, по теореме Эйлера имеем

$$x^{fg}=x\cdot(x^k)^{\varphi(pq)}\equiv x\cdot 1=x \pmod{pq}.$$

Вы поняли, что такое f и g? В нашем примере $f=9007$ (единственное обязательное требование к числу f — его взаимная простота с числом $(p-1)(q-1)$; впрочем, брать $f=1$ или $f=(p-1)(q-1)-1$ вряд ли разумно, если хотите сохранить секреты). А число g, как уже было сказано, — решение сравнения

$$fg\equiv 1 \pmod{\varphi(pq)}.$$

(В статье «Алгоритм Евклида» рассказано, как решать такие сравнения.) Сравнения

$$y^g\equiv x^{fg}\equiv x \pmod{pq}$$

показывают, что для нахождения x достаточно найти остаток от деления y^g на pq. (Числа выбраны так, что $x<pq$. При этом x не кратно ни p, ни q. Не думайте, что последнее требование всерьез нас ограничивает: если p и q — большие числа, то вероятность того, что x нацело делится на p или q, пренебрежимо мала. Кроме того, можно предусмотреть в алгоритме, что в случае чего сообщение x будет автоматически как-то чуть-чуть изменено, без изменения его смысла, что x и pq станут взаимно просты.)

Почему многие надеются, что шифр RSA является шифром с открытым ключом? Да потому, что числа pq и f можно сделать общедоступными. Тогда зашифровать сообщение сможет любой, у кого есть компьютер (и какая-нибудь программа, позволяющая выполнять действия с многозначными числами). Расшифровать сообщение легко, если известно число g. Но единственный известный ныне способ нахождения числа g требует нахождения чисел p и q, то есть разложения произведения pq на множители. А эффективных алгоритмов решения этой задачи пока нет (удача 1994 г. не в счет: если бы в числах p и q было не 64 и 65, а хотя бы по 300 цифр, то и ресурсов сети Internet не хватило бы!). Впрочем, нет сейчас и доказательства того, что никто никогда не научится быстро (математик сказал бы: «за время, полиномиальное от количества цифр») разлагать числа на простые множители. ■

Как строят большие простые числа? Для криптографической системы RSA нужны большие (длиной в сотни цифр) простые числа. Эффективный метод построения таких чисел основан на следующей лемме.

Лемма. ***Пусть q — нечетное простое число, r — четное натуральное, $n=qr+1$. Если существует такое целое a, что $a^{n-1}\equiv 1 \pmod n$ и НОД$(a^r-1, n)=1$, то для каждого простого делителя p числа n верно сравнение $p\equiv 1 \pmod{2q}$.***

Доказательство. Обозначим порядок числа a по модулю p буквой k. Поскольку $a^{n-1}\equiv 1 \pmod p$ и $a^{(n-1)/q}=a^r\not\equiv 1 \pmod p$, то k делится на q. Поскольку $p-1$ делится на k, то $p-1$ делится на q. Кроме того, $p-1$ четно.

Следствие. ***Если выполнены условия леммы и $r\leqslant 4q+2$, то n — простое число.***

Доказательство. Пусть n является произведением не менее чем двух простых чисел. Поскольку каждое из них не меньше $2q+1$, получаем противоречие:

$$(2q+1)^2\leqslant n=qr+1\leqslant 4q^2+2q+1.$$

Покажем теперь, как, имея большое простое число q, можно пытаться строить существенно большее простое число n. Выберем случайным образом четное число r, где $q<r\leqslant 4q+2$, и положим $n=qr+1$. Затем проверим n на отсутствие малых простых делителей. (В этом месте мы чуть лукавим: следует не только делить на малые простые числа, но и применять более хитрые методы, основанные на малой теореме Ферма: если $a^{n-1}\not\equiv 1 \pmod n$ для некоторого a, взаимно простого с n, то n составное.) Если выяснится, что n — составное, то следует выбрать новое значение r и повторить вычисления.

Если же есть надежда, что n простое, то можно случайным образом выбрать число a и проверить, выполнены ли соотношения $a^{n-1}\equiv 1 \pmod n$ и НОД$(a^r-1, n)=1$. Если выполнены, то n простое (заметьте: $n>q^2$, так что n записывается примерно вдвое большим количеством цифр, чем q). Если же нет, то можно взять другое значение a, и так далее.

В настоящий момент не доказано, что этот алгоритм работает, и тем более — что он работает быстро. Однако на практике он позволяет строить большие (порядка 10^{300}) простые числа. ■

Числа Фибоначчи

В 1202 г. Леонардо Фибоначчи (Пизанский) в «Книге об абаке» рассмотрел последовательность 1, 1, 2, 3, 5, 8, 13, . . . , каждый следующий член которой равен сумме двух предыдущих. Формулами это можно записать так: $\varphi_1=\varphi_2=1$, $\varphi_{n+2}=\varphi_{n+1}+\varphi_n$ *для любого* n.

«Сколько пар кроликов в один год от одной пары рождается? — спрашивал в 1228 г. Фибоначчи и сам же отвечал: — Некто поместил пару кроликов в некоем месте, огороженном со всех сторон стеной, дабы узнать, сколько пар кроликов родится при этом в течение года, если природа кроликов такова, что через месяц пара кроликов производит на свет другую пару, а рождают кролики со второго месяца после своего рождения. Так как первая пара в первом месяце дает потомство, удвой, и в этом месяце окажутся 2 пары; из них одна пара, а именно, первая, рождает и в следующем месяце, так что во втором месяце оказывается 3 пары; из них в следующем месяце 2 пары дадут потомство, так что в третьем месяце родятся еще 2 пары кроликов, и число пар кроликов в этом месяце достигнет 5; из них в этом же месяце дадут потомство 3 пары, и число пар кроликов в четвертом месяце достигнет 8; из них 5 пар произведут другие 5 пар, которые, сложенные с 8 парами, дадут в пятом месяце 13 пар; из них 5 пар, рожденных в этом месяце, не дают в том же месяце потомства, а остальные 8 пар рождают, так что в шестом месяце оказывается 21 пара; сложенные с 13 парами, которые родятся в седьмом месяце, они дают 34 пары; сложенные с 21 парами, рожденными в восьмом месяце, они дают в этом месяце 55 пар; сложенные с 34 парами, рожденными в девятом месяце, они дают 89 пар; сложенные вновь с 55 парами, рожденными в десятом месяце, они дают 144 пары; снова сложенные с 89 парами, которые рождаются в одиннадцатом месяце, они дают 233 пары; сложенные вновь с 144 парами, рожденными в последнем месяце, они дают 377 пары; столько пар произвела первая пара в данном месте к концу одного года.

Действительно . . . мы складываем первое число со вторым, то есть 1 и 2; и второе с третьим; и третье с четвертым; и четвертое с пятым; и так одно за другим, пока не сложим десятое с одиннадцатым, то есть 144 с 233; и мы получим общее число упомянутых кроликов, то есть 377; и так можно делать по порядку до бесконечного числа месяцев». ■

Вычислим несколько первых членов последовательности:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
φ_n	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610

Рассмотрим цепные дроби

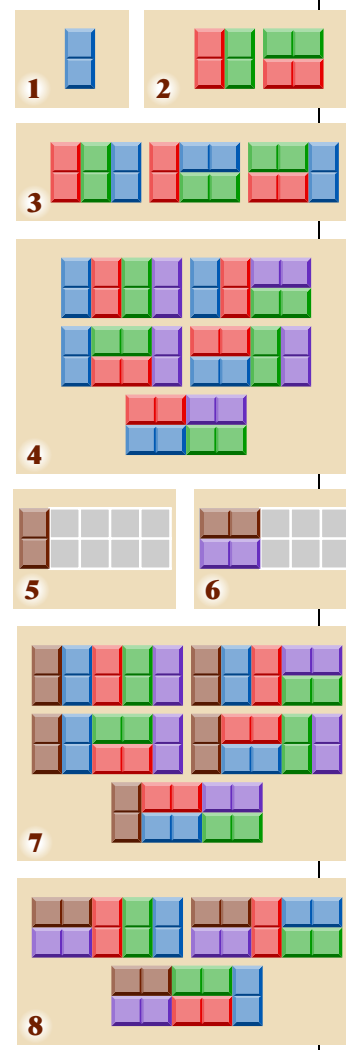
$$1,\qquad 1+\frac{1}{1}=\frac{2}{1},\qquad 1+\frac{1}{1+\frac{1}{1}}=1+\frac{1}{2}=\frac{3}{2},$$

$$1+\frac{1}{1+\frac{1}{1+\frac{1}{1}}}=1+\frac{1}{3/2}=\frac{5}{3},$$

$$1+\frac{1}{1+\frac{1}{1+\frac{1}{1+\frac{1}{1}}}}=1+\frac{1}{5/3}=\frac{8}{5}.$$

Возникают дроби вида φ_{n+1}/φ_n. Да оно и неудивительно:

$$1+\frac{1}{\varphi_{n+1}/\varphi_n}=1+\frac{\varphi_n}{\varphi_{n+1}}=\frac{\varphi_{n+1}+\varphi_n}{\varphi_{n+1}}=\frac{\varphi_{n+2}}{\varphi_{n+1}}.\ \blacksquare$$



Сколькими способами можно разрезать полоску размером $2\times n$ на доминошки, то есть прямоугольники размером 1×2? При маленьких n можно все нарисовать: обозначив искомое число способов через $f(n)$, видим из рисунков 1—4, что $f(1)=1$, $f(2)=2$, $f(3)=3$ и $f(4)=5$. Как найти $f(5)$? Левая верхняя клетка покрыта доминошкой, расположенной либо вертикально (рис. 5), либо горизонтально (рис. 6). Значит, $f(5)$ вариантов распадаются на $f(4)$ вариантов рисунка 7 и $f(3)$ вариантов рисунка 8, то есть

$$f(5)=f(4)+f(3)=5+3=8.$$

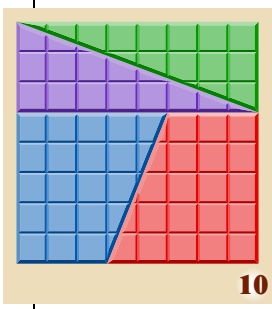
Вообще, при любом $n>1$ верна рекуррентная — выражающая следующее значение через предыдущие — формула $f(n+1)=f(n)+f(n-1)$. Поскольку $f(1)=\varphi_2$ и $f(2)=\varphi_3$, имеем: $f(n)=\varphi_{n+1}$. ■

Сколько существует n-значных чисел, составленных из цифр 2 и 5, в которых никакие две двойки не стоят рядом? Обозначим искомое количество способов через $g(n)$. Очевидно, $g(1)=2$ и $g(2)=3$ (годятся числа 25, 52 и 55). Легко выписать и трехзначные числа: 252, 255, 525, 552 и 555. Значит,

$g(3)=5$. Впрочем, это значение находить даже и необязательно. Важнее то, что любое интересующее нас $(n+1)$-значное число начинается либо с двойки, либо с пятерки. В первом случае после двойки должна идти пятерка, после которой — любое из $g(n-1)$ чисел, во втором случае никаких ограничений пятерка не создает, годится любой из $g(n)$ вариантов. Мы получили рекуррентную формулу $g(n+1)=g(n-1)+g(n)$, совпадающую с формулой Фибоначчи. Поскольку $g(1)=\varphi_3$ и $g(2)=\varphi_4$, имеем: $g(n)=\varphi_{n+2}$. Опять последовательность Фибоначчи! ■

Сколькими способами можно расположить первые n натуральных чисел в строку, чтобы никакое число не отличалось от номера занимаемого номера больше чем на 1? Для $n=1$ и 2 ответы — 1 и 2 способа соответственно. Для $n=3$ — три перестановки 123, 132 и 213, для $n=4$ годятся пять перестановок: 1234, 1324, 2134 и 1234, 2134. Вообще, число n может стоять на n-месте — и таких интересующих нас перестановок столько же, сколько их для $n-1$ чисел; а может на $(n-1)$-м, и тогда на n-месте стоит число $n-1$, а первые $n-2$ числа можно расставлять, не глядя на n и $n-1$.
Итак, количество перестановок, в которых никакое число не сдвигается более чем на 1 — число Фибоначчи! ■

Рассмотрим полоску из k клеток и задумаемся, сколько существует способов пройти из левой клетки полоски в правую, если каждым ходом разрешено переходить в соседнюю справа клетку или перепрыгивать через одну клетку. Очевидно, при $k=1$ идти некуда, да и не нужно; при $k=2$ нужно сделать ровно один шаг. Значит, при $k=1$ или 2 число способов равно 1. Далее, если первый шаг — сдвиг на 1 клетку, то остается полоска из $k-1$ клеток, если же первый шаг — сдвиг на 2 клетки, то остается полоска из $k-2$ клеток. Таким образом, для рассматриваемого числа способов выполнена в точности такая же рекуррентная формула, что и для чисел Фибоначчи. Поэтому количество способов пройти полоску длины k — это число Фибоначчи φ_k. ■


10

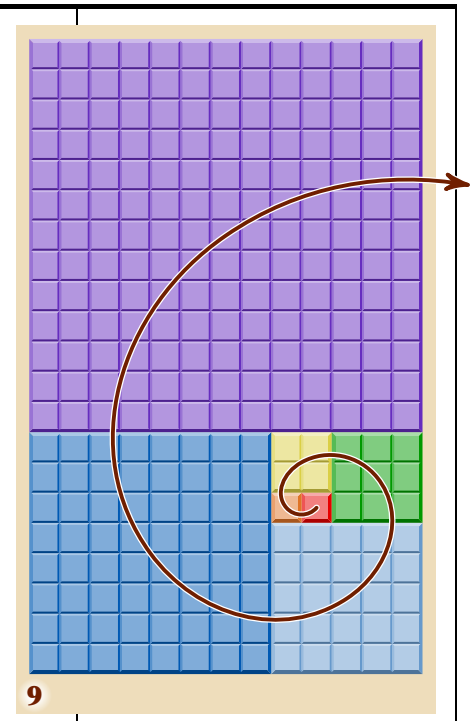
На рисунке 9 числа Фибоначчи выражают длины сторон спиральной последовательности квадратов на клетчатой бумаге. Из этого рисунка очевидна формула

$$\varphi_1^2+\varphi_2^2+\varphi_3^2+\ldots+\varphi_n^2=\varphi_n\varphi_{n+1}.$$

Между числами Фибоначчи есть и другие любопытные соотношения:

$$\varphi_1+\varphi_2+\varphi_3+\ldots+\varphi_n=\varphi_{n+2}-1,$$
$$\varphi_1+\varphi_3+\varphi_5+\ldots+\varphi_{2n-1}=\varphi_{2n},$$
$$\varphi_2+\varphi_4+\varphi_6+\ldots+\varphi_{2n}=\varphi_{2n+1}-1,$$
$$\varphi_{n+1}\varphi_{n-1}-\varphi_n^2=(-1)^n,$$

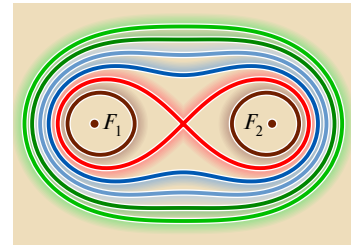
которые легко доказать по индукции. Последнее соотношение открыл в 1680 г. Дж. Д. Кассини. При $n=6$ оно превращается в равенство $13\cdot5-8^2=1$, которое лежит в основе геометрического парадокса: на рисунке 10 шахматная доска разрезана на четыре части, из которых на рисунке 11 сложен прямоугольник размером 5×13. (Аналогичная конструкция при любом n разбивает квадрат со стороной φ_n на четыре части, из которых получается прямоугольник размером $\varphi_{n-1}\times\varphi_{n+1}$. Либо одна клетка теряется, либо возникает лишняя — в зависимости от четности n.) Разгадка парадокса проста: на рисунке 11 линии, соединяющие левый верхний угол с нижним правым углом, образуют не отрезок, а незаметный для глаза параллелограмм.

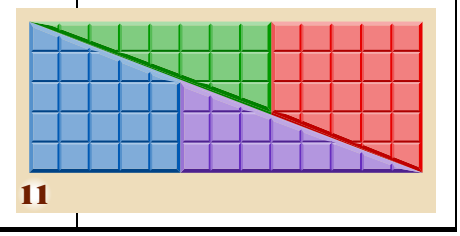

9

Джованни Доменико Кассини (1625—1712) — французский астроном. Открыл вращение Юпитера (1665) и Марса (1666), деление кольца Сатурна на внутреннее и внешнее. Для изучения орбиты Земли пытался использовать овал Кассини — кривую, заданную уравнением

$$(x^2+y^2)^2-2c^2(x^2-y^2)=a^4-c^4.$$

Она состоит из точек M, произведение расстояний от которых до точек $F_1(-c;0)$ и $F_2(c;0)$ равно a^2. При $a\geqslant c\sqrt{2}$ это зеленые выпуклые кривые; при $c<a<c\sqrt{2}$ — синие овалы «с талией»; при $a=c$ — красная лемниската Бернулли; при $a<c$ — два коричневых овала. ■




11

В вершине A правильного восьмиугольника $AB_1C_1D_1ED_2C_2B_2$ находится лягушка. Из любой вершины восьмиугольника, кроме вершины E, она может прыгнуть в любую из двух соседних вершин. Попав в вершину E, лягушка остается там навсегда. Найдем количество способов e_n, которыми лягушка может попасть из вершины A в вершину E ровно за n прыжков.

Раскрасим вершины восьмиугольника через одну в синий и зеленый цвета. Поскольку при каждом прыжке цвет вершины меняется, то $e_{2n-1}=0$. Обозначим через b_n и d_n количества способов, которыми можно за $2n-1$ прыжков попасть из A, соответственно, в B_1 и D_1 (разумеется, столькими же способами можно попасть из A в B_2 и D_2). Очевидно, $e_{2n}=2d_n$. Обозначим через a_n количество способов, которыми за $2n$ прыжков можно, начав движение в вершине A, вернуться в нее, а через c_n — число способов за $2n$ припрыгать в вершину C_1 (столько же способов — в C_2). Поскольку $a_n=2b_n$, $c_n=b_n+d_n$ и $b_{n+1}=a_n+c_n$, $d_{n+1}=c_n$, то $b_{n+1}=a_n+c_n=2b_n+b_n+d_n=3b_n+d_n$ и $d_{n+1}=c_n=b_n+d_n$. Рассмотрим таблицу

n	1	2	3	4	5
b_n	1	3	10	34	116
d_n	0	1	4	14	48

Нетрудно угадать и доказать по индукции закономерности $b_{n+2}=4b_{n+1}-2b_n$ и $d_{n+2}=4d_{n+1}-2d_n$. Найдем все геометрические прогрессии a, aq, aq^2, aq^3, ..., удовлетворяющие этому рекуррентному соотношению:

$$aq^{n+1}=4aq^n-2aq^{n-1},$$

откуда $q^2-4q+2=0$, то есть $q=2\pm\sqrt{2}$. Подберем такие числа α и β, что $d_n=\alpha(2+\sqrt{2})^n+\beta(2-\sqrt{2})^n$. Для этого составим уравнения для $n=1$ и 2:

$$\begin{cases} 0=\alpha(2+\sqrt{2})+\beta(2-\sqrt{2}),\\ 1=\alpha(2+\sqrt{2})^2+\beta(2-\sqrt{2})^2. \end{cases}$$

Решив эту систему, находим $\alpha=1/(4+4\sqrt{2})$ и $\beta=1/(4-4\sqrt{2})$. Следовательно,

$$d_n=\frac{1}{2\sqrt{2}}((2+\sqrt{2})^{n-1}-(2-\sqrt{2})^{n-1})$$

и, таким образом,

$$e_{2n}=2d_n=\frac{1}{\sqrt{2}}((2+\sqrt{2})^{n-1}-(2-\sqrt{2})^{n-1}).\ \blacksquare$$

Если заменить в соотношении Кассини φ_{n-1} на $\varphi_{n+1}-\varphi_n$, то оно примет вид $\varphi_{n+1}^2-\varphi_{n+1}\varphi_n-\varphi_n^2=(-1)^n$. В статье «Уравнения Пелля» доказано, что уравнение $x^2-xy-y^2=\pm1$ других решений в натуральных числах не имеет. ■

Д. Бернулли в 1728 г. опубликовал формулу

$$\varphi_n=\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n-\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n,$$

но о ней позабыли до 1843 г., когда была вновь открыта французом Ж. Бине. Из этой формулы следует, что φ_n растет примерно как геометрическая прогрессия со знаменателем $\alpha=(1+\sqrt{5})/2$, точнее, φ_n равно ближайшему целому числу к $\alpha^n/\sqrt{5}$.

Формулу Бине можно проверить по индукции. Но есть и способ вывести ее, не зная заранее ответ. Идея в том, что мы временно забудем про значения $\varphi_1=\varphi_2=1$, а рассмотрим всевозможные последовательности, сумма каждых двух соседних членов которых равна следующему за ними числу. Например,

$$-1,\ \ 3,\ \ 2,\ \ 5,\ \ 7,\ 12,\ 19,\ 31,\ 50,\ \ldots$$
$$5,\ -1,\ \ 4,\ \ 3,\ \ 7,\ 10,\ 17,\ 27,\ 44,\ \ldots$$

Сумма двух таких последовательностей

$$4,\ \ 2,\ \ 6,\ \ 8,\ 14,\ 22,\ 36,\ 58,\ 94,\ \ldots$$

обладает тем же свойством. В самом деле, если $g_{n+2}=g_{n+1}+g_n$ и $h_{n+2}=h_{n+1}+h_n$, то $g_{n+2}+h_{n+2}=(g_{n+1}+h_{n+1})+(g_n+h_n)$. Среди рассматриваемых последовательностей есть и геометрические прогрессии. Найдем их. Если

$$aq^{n+1}=aq^n+aq^{n-1},$$

то $q^2=q+1$, то есть $q=(1\pm\sqrt{5})/2$. Обозначим $\alpha=(1+\sqrt{5})/2$ и $\beta=(1-\sqrt{5})/2$. Для любых чисел a и b последовательности a, $a\alpha$, $a\alpha^2$, $a\alpha^3$, ... и b, $b\beta$, $b\beta^2$, $b\beta^3$, ... удовлетворяют рекуррентному соотношению Фибоначчи. Значит, удовлетворяет ему и последовательность $a+b$, $a\alpha+b\beta$, $a\alpha^2+b\beta^2$, $a\alpha^3+b\beta^3$, ... Теперь вспомним о первых двух членах: $\varphi_1=\varphi_2=1$. Зная первые два члена последовательности и рекуррентное соотношение, мы можем по очереди найти их все. Поэтому если мы подберем числа a и b так, чтобы были верны равенства $1=a+b$ и $1=a\alpha+b\beta$, то равенство $\varphi_n=a\alpha^{n-1}+b\beta^{n-1}$ будет верно не только для $n=1$ или 2, но и для любого натурального n. Очевидно, $b=1-a$ и $1=a\alpha+(1-a)\beta$, то есть $a=\frac{1-\beta}{\alpha-\beta}=\frac{\alpha}{\sqrt{5}}$ и $b=1-a=\frac{-\beta}{\sqrt{5}}$, а это и есть формула Бине! ■

Производящую функцию для последовательности Фибоначчи

$$f(x)=x+x^2+2x^3+3x^4+5x^5+8x^6+13x^7+21x^8+34x^9+\ldots$$

домножим на x и на x^2:

$$\begin{aligned} xf(x)&=\ \ x^2+\ x^3+2x^4+3x^5+5x^6+\ 8x^7+13x^8+21x^9+34x^{10}+\ldots,\\ x^2f(x)&=\qquad\ x^3+\ x^4+2x^5+3x^6+\ 5x^7+\ 8x^8+13x^9+21x^{10}+34x^{11}+\ldots \end{aligned}$$

Вычитая из $f(x)$ сумму $xf(x)+x^2f(x)$, получаем: $(1-x-x^2)f(x)=x$, то есть

$$\varphi_1x+\varphi_2x^2+\varphi_3x^3+\ldots=\frac{x}{1-x-x^2}.$$

Очевидно, $x^2+x-1=(x+\alpha)(x+\beta)$. Для вывода формулы Бине достаточно разложить производящую функцию на простейшие дроби:

$$\frac{x}{1-x-x^2}=\frac{1}{\sqrt{5}}\left(\frac{\beta}{x+\beta}-\frac{\alpha}{x+\alpha}\right)$$

и воспользоваться формулой суммы бесконечной убывающей геометриче-

ской прогрессии:

$$\frac{\beta}{x+\beta}=\frac{1}{1-x\alpha}=1+x\alpha+x^2\alpha^2+x^3\alpha^3+x^4\alpha^4+\ldots,$$

$$\frac{\alpha}{x+\alpha}=\frac{1}{1-x\beta}=1+x\beta+x^2\beta^2+x^3\beta^3+x^4\beta^4+\ldots\ \blacksquare$$

Арифметика чисел Фибоначчи весьма интересна. Каждое третье число Фибоначчи четно; каждое четвертое кратно 3; каждое двенадцатое оканчивается нулем.

Соседние числа Фибоначчи взаимно просты. Это следует из соотношения Кассини. (Можно обойтись и без него: если бы φ_n и φ_{n+1} имели общий делитель $d>1$, то на d делилось бы и предшествующее число $\varphi_{n-1}=\varphi_{n+1}-\varphi_n$, а вместе с ним и $\varphi_{n-2}=\varphi_n-\varphi_{n-1}$ и так далее; но $\varphi_1=1$ не делится на d.)

Теорема 1. НОД(φ_m, φ_n)=$\varphi_{\text{НОД}(m,n)}$.

Лемма 1. $\varphi_{m+n}=\varphi_m\varphi_{n-1}+\varphi_{m+1}\varphi_n$.

Положив $m=n$, получаем

$$\varphi_{2n}=\varphi_{n+n}=\varphi_n\varphi_{n-1}+\varphi_{n+1}\varphi_n=(\varphi_{n+1}+\varphi_{n-1})\varphi_n=(\varphi_{n+1}+\varphi_{n-1})(\varphi_{n+1}-\varphi_{n-1})=\varphi_{n+1}^2-\varphi_{n-1}^2.$$

Аналогично можно вывести формулу $\varphi_{3n}=\varphi_{n+1}^3+\varphi_n^3-\varphi_{n-1}^3$.

Доказательство леммы. I способ — индукция по n. **База.** При $n=1$ или 2 равенства верны. **Переход.** Если верны равенства

$$\varphi_{m+n}=\varphi_m\varphi_{n-1}+\varphi_{m+1}\varphi_n \quad\text{и}\quad \varphi_{m+n+1}=\varphi_m\varphi_n+\varphi_{m+1}\varphi_{n+1},$$

то $\varphi_{m+n+2}=\varphi_{m+n+1}+\varphi_{m+n}=\varphi_m\varphi_{n-1}+\varphi_{m+1}\varphi_n+\varphi_m\varphi_n+\varphi_{m+1}\varphi_{n+1}=\varphi_m(\varphi_{n-1}+\varphi_n)+$ $+\varphi_{m+1}(\varphi_n+\varphi_{n+1})=\varphi_m\varphi_{n+1}+\varphi_{m+1}\varphi_{n+2}$.

II способ. φ_{m+n} — это количество способов пройти из левой клетки полоски длиной $n+1$ клеток в правую клетку этой же полоски, перепрыгивая не более чем через клетку. Все эти способы можно разбить на два типа: $\varphi_m\varphi_{n-1}$ тех, когда перепрыгиваем через $(m+1)$-ю клетку, и $\varphi_{m+1}\varphi_n$ тех, когда останавливаемся на этой клетке.

Идея доказательства теоремы 1: НОД$(\varphi_{m+n}, \varphi_n)=$НОД$(\varphi_m\varphi_{n-1}+\varphi_{m+1}\varphi_n, \varphi_n)=$ $=$НОД$(\varphi_m\varphi_{n-1}, \varphi_n)=$НОД$(\varphi_m, \varphi_n)$.

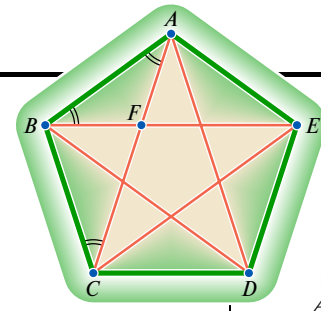
Теорема 2. ***Для любого натурального m среди первых m^2-1 чисел Фибоначчи хотя бы одно число делится на m.***

Доказательство. Рассмотрим пару остатков ($\varphi_k \bmod m$; $\varphi_{k+1} \bmod m$) для $k=1, 2, \ldots, m^2$. Количество остатков от деления на m равно m, поэтому количество пар остатков равно m^2, а мы специально взяли на 1 больше. Поэтому какая-то пара остатков встречается, как минимум, дважды:

$$(\varphi_k \bmod m;\ \varphi_{k+1} \bmod m)=(\varphi_{k+r} \bmod m;\ \varphi_{k+r+1} \bmod m), \qquad (*)$$

где $1\leqslant k<k+r\leqslant m^2+1$. Очевидно, $\varphi_{k-1}=\varphi_{k+1}-\varphi_k\equiv\varphi_{k+r+1}-\varphi_{k+r}=\varphi_{k+r-1}\ (\text{mod}\ m)$, так что $\varphi_{k-1} \bmod m=\varphi_{k+r-1} \bmod m$. Поэтому если $k>1$, то можно вместо k рассмотреть $k-1$, свойство $(*)$ сохранится. Уменьшая и уменьшая таким образом величину k, мы доведем ее до минимально возможного значения $k=$ $=1$ и таким образом получим равенство $(1;\ 1)=(\varphi_{1+r} \bmod m;\ \varphi_{1+r+1} \bmod m)$, откуда $\varphi_r=\varphi_{r+2}-\varphi_{r+1}\equiv 1-1=0\ (\text{mod}\ m)$. ■

Пусть p простое, $p\neq 2$. Если $p=5$, то $\varphi_p\vdots p$. Докажем, что $\varphi_{p\pm1}\vdots p$, где надо брать знак +, если $p\equiv 2$ или 3 (mod 5), и знак −, если $p\equiv 1$ или 4 (mod 5).



Рассмотрим правильный пятиугольник $ABCDE$, длина стороны которого равна 1. Обозначим $AC=d$. Поскольку $CDEF$ — параллелограмм, то $AF=$ $=BF=d-1$. Треугольник ABC подобен треугольнику AFB, следовательно, $\frac{1}{d}=\frac{d-1}{1}$, то есть $d^2-d=1$, откуда $d=\frac{1+\sqrt{5}}{2}$ — так называемое золотое сечение. ■

Рассмотрим прямоугольник размером $1\times d$, где $d>1$. Отрезав от него квадрат 1×1, получим прямоугольник размером $1\times(d-1)$. Он подобен исходному, если $\frac{1}{d}=\frac{d-1}{1}$. Опять золотое сечение! ■

Из равенства $\alpha^2=\alpha+1$ по индукции легко вывести равенство $\alpha^n=\varphi_n\alpha+\varphi_{n-1}$.

Рекуррентную формулу, задающую последовательность Фибоначчи, можно записать в виде $\varphi_n=\varphi_{n+2}-\varphi_{n+1}$. Полагая последовательно $n=0,\ -1,\ -2,\ \ldots$, находим $\varphi_0=0$, $\varphi_{-1}=1$, $\varphi_{-2}=-1$, $\varphi_{-3}=2,\ \ldots,$ и вообще, $\varphi_{-n}=$ $=(-1)^{n+1}\varphi_n$. Формула Бине верна и для чисел Фибоначчи с отрицательными номерами. ■

При $|x|<1/\alpha=(\sqrt{5}-1)/2$ сходится ряд $x+x^2+2x^3+3x^4+$ $+5x^5+8x^6+\ldots$ В частности, при $x=\frac{1}{2}$ имеем $\frac{x}{1-x-x^2}=2$, поэтому $\frac{1}{2}+\frac{1}{4}+\frac{2}{8}+\frac{3}{16}+\frac{5}{32}+\frac{8}{64}+$ $+\frac{13}{128}+\frac{21}{256}+\ldots=\sum\limits_{n=1}^{\infty}\frac{\varphi_n}{2^n}=2$. ■

По индукции легко доказать равенства

$$\varphi_1\varphi_2+\varphi_2\varphi_3+\varphi_3\varphi_4+\ldots$$
$$\ldots+\varphi_{2n-1}\varphi_{2n}=\varphi_{2n}^2,$$
$$\varphi_1\varphi_2+\varphi_2\varphi_3+\varphi_3\varphi_4+\ldots$$
$$\ldots+\varphi_{2n-1}\varphi_{2n}+\varphi_{2n}\varphi_{2n+1}=\varphi_{2n+1}^2-1,$$
$$\varphi_1+2\varphi_2+3\varphi_3+\ldots+n\varphi_n=$$
$$=n\varphi_{n+2}-\varphi_{n+3}+2,$$
$$n\varphi_1+(n-1)\varphi_2+(n-2)\varphi_3+\ldots$$
$$\ldots+2\varphi_{n-1}+\varphi_n=\varphi_{n+4}-n-3.\ \blacksquare$$

Если n четно, то соотношение Кассини можно записать в виде $\varphi_{n-1}\varphi_{n+1}=\varphi_n^2+1$. Поскольку число -1 не является квадратичным вычетом по простому модулю p, где $p\equiv 3 \pmod 4$, то ни φ_{n-1}, ни φ_{n+1} не делится на p. Следовательно, любой нечетный делитель d числа Фибоначчи, номер которого нечетен, удовлетворяет сравнению $d\equiv 1 \pmod 4$. ■

По кругу выложены φ_n карточек оборотной стороной вверх ($n\geqslant 4$). На карточках написаны неизвестные числа. Разрешено переворачивать карточки. Научимся, перевернув не более $n-1$ карточек, находить локальный максимум — такую карточку, что написанное на ней число не меньше чисел обеих ее соседок.

Лемма. ***Если известны числа a, b и c, причем $c\geqslant a$ и $c\geqslant b$, а расположены они вдоль прямой так:***

$$\underbrace{a,\ldots,}_{\varphi_k}\ c,\ \underbrace{\ldots,b}_{\varphi_{k-1}},$$

то есть сначала a, затем φ_k-1 неизвестных чисел, затем c, еще $\varphi_{k-1}-1$ неизвестных чисел и, наконец, b, то можно найти локальный максимум за $k-2$ переворачиваний.

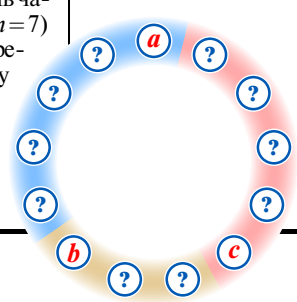
Доказательство — индукция по k. **База.** $k=2$ — ничего переворачивать не надо. **Переход.** Перевернем карточку d, расположенную на расстояниях φ_{k-1} и φ_{k-2} от a и c соответственно:

$$\underbrace{a,\ldots,}_{\varphi_{k-1}}\ d,\ \underbrace{\ldots,c,}_{\varphi_{k-2}}\ \underbrace{\ldots,b.}_{\varphi_{k-1}}$$

Если $d>c$ (или $d\leqslant c$), выбросим все числа, расположенные справа от c (соответственно, слева от d), и воспользуемся предположением индукции.

Вернемся к задаче.

$\varphi_n=\varphi_{n-1}+\varphi_{n-2}=2\varphi_{n-2}+\varphi_{n-3}$. Перевернем две карточки a и b, между которыми $\varphi_{n-2}-1$ других карточек. Пусть для определенности $a\leqslant b$. Разобьем окружность на три дуги, в которых φ_{n-2}, φ_{n-3} и φ_{n-2} карточек соответственно, причем пусть a и b — начальные (для определенности, при движении вдоль окружности против часовой стрелки; на рисунке $n=7$) карточки первых двух дуг. Перевернем начальную карточку третьей дуги и обозначим ее число буквой c. Если $c\leqslant b$, применяем лемму к дуге $a\ldots b\ldots c$; если же $c>b$, то к дуге $a\ldots c\ldots b$. ■



Воспользуемся формулой Бине и сравнением $(x+y)^p\equiv x^p+y^p$, которое выполнено по простому модулю в силу того, что биномиальные коэффициенты C_p^k, где $1\leqslant k<p$, делятся на p, а также тем, что в силу критерия Эйлера и квадратичного закона взаимности $5^{(p-1)/2}\equiv\left(\frac{5}{p}\right)\equiv\mp 1 \pmod p$:

$$\varphi_{p\pm1}=\frac{1}{\sqrt5}\left(\left(\frac{1+\sqrt5}{2}\right)^p\alpha^{\pm1}-\left(\frac{1-\sqrt5}{2}\right)^p\beta^{\pm1}\right)\equiv$$

$$\equiv\frac{1}{2^p\sqrt5}((1+5^{(p-1)/2}\sqrt5)\alpha^{\pm1}-(1-5^{(p-1)/2}\sqrt5)\beta^{\pm1})\equiv$$

$$\equiv\frac{1}{2^p\sqrt5}((1\mp1\sqrt5)\alpha^{\pm1}-(1\pm1\sqrt5)\beta^{\pm1}) \pmod p.$$

Если $p\equiv 1$ или $4 \pmod 5$, то $\varphi_{p-1}\equiv\frac{2}{2^p\sqrt5}((1+\sqrt5)\alpha^{-1}-(1-\sqrt5)\beta^{-1})\equiv 0 \pmod p$.

Аналогично, $\varphi_{p+1}\equiv\frac{2}{2^p\sqrt5}((1-\sqrt5)\alpha-(1+\sqrt5)\beta)\equiv 0 \pmod p$ при $p\equiv 2$ или $3 \pmod 5$. ■

Некоторые числа Фибоначчи кратны своему номеру: например, $\varphi_5\vdots 5$ и $\varphi_{12}=144\vdots 12$. Много ли таких чисел?

Теорема 3. ***Для любого натурального числа n существует такое натуральное число m, что φ_m делится на m, а m, в свою очередь, делится на n.***

Доказательство. Если $\varphi_{n_1}\vdots m_1\vdots n_1$ и $\varphi_{n_2}\vdots m_2\vdots n_2$, то, очевидно,

$$\varphi_{\mathrm{НОК}[n_1,n_2]}\vdots \mathrm{НОК}[m_1,m_2]\vdots \mathrm{НОК}[n_1,n_2].$$

Поэтому теорему 3 достаточно доказать для степеней простых чисел. ■

Начнем со степеней числа 2. Как вы помните,

$$\varphi_{2n}=\varphi_n(\varphi_{n-1}+\varphi_{n+1}).$$

Частное $\frac{\varphi_{2n}}{\varphi_n}=\varphi_{n-1}+\varphi_{n+1}=2\varphi_{n-1}+\varphi_n$ нечетно, если φ_n нечетно. Если же φ_n четно, то φ_{n-1} нечетно (соседние числа Фибоначчи взаимно просты и поэтому не могут быть оба четны). Поэтому если φ_n делится на 4, то число $2\varphi_{n-1}+\varphi_n$ не делится на 4. Если же φ_n не делится на 4, то число $2\varphi_{n-1}+\varphi_n$ — сумма двух четных чисел, не делящихся на 4, — кратно 4.

Поскольку $\varphi_{12}\vdots 2^4$, по индукции получаем: $\varphi_{3\cdot2^k}\vdots 3\cdot2^{k+2}$ при $k\geqslant 2$. ■

Другие степени простых чисел рассмотрим при помощи следующей леммы.

Лемма 2. $\boldsymbol{\varphi_{kn-1}\equiv\varphi_{n-1}^k}$, $\boldsymbol{\varphi_{kn}\equiv k\varphi_n\varphi_{n+1}^{k-1}}$ ***и*** $\boldsymbol{\varphi_{kn+1}\equiv\varphi_{n+1}^k \pmod{\varphi_n^2}}$.

Доказательство — индукция по k. База тривиальна, переход основан на делимости числа φ_{kn} на φ_n и формуле леммы 1 (все сравнения — по модулю φ_n^2):

$$\varphi_{(kn-1)+n}=\varphi_{kn-1}\varphi_{n-1}+\varphi_{kn}\varphi_n\equiv\varphi_{n-1}^k\varphi_{n-1}=\varphi_{n-1}^{k+1},$$

$$\varphi_{n+kn}=\varphi_n\varphi_{kn-1}+\varphi_{n+1}\varphi_{kn}\equiv\varphi_n\varphi_{n-1}^k+\varphi_{n+1}k\varphi_n\varphi_{n+1}^{k-1}\equiv$$

$$\equiv\varphi_n\varphi_{n+1}^k+k\varphi_n\varphi_{n+1}^k=(k+1)\varphi_n\varphi_{n+1}^k,$$

$$\varphi_{n+(kn+1)}=\varphi_n\varphi_{kn}+\varphi_{n+1}\varphi_{kn+1}\equiv\varphi_{n+1}\varphi_{n+1}^k=\varphi_{n+1}^{k+1}.$$

Лемма 3. $\boldsymbol{\varphi_{kn}\equiv\varphi_{n+1}^k-\varphi_{n-1}^k \pmod{\varphi_n^3}}$.

Доказательство — индукция по k, почти как в лемме 2. База тривиальна, да и переход несложен:

$$\varphi_{kn+n}=\varphi_{kn}\varphi_{n-1}+\varphi_{kn+1}\varphi_n\equiv(\varphi_{n+1}^k-\varphi_{n-1}^k)\varphi_{n-1}+\varphi_{n+1}^k\varphi_n=$$
$$=\varphi_{n+1}^k\varphi_{n-1}-\varphi_{n-1}^{k+1}+\varphi_{n+1}^k\varphi_n=\varphi_{n+1}^k(\varphi_{n-1}+\varphi_n)-\varphi_{n-1}^{k+1}=\varphi_{n+1}^{k+1}-\varphi_{n-1}^{k+1}\pmod{\varphi_n^3}.$$

Теорема 4. ***Если $\varphi_n\vdots p$, где p — простое число, $p>2$, то φ_{np} делится на $p\varphi_n$, но не делится на $p^2\varphi_n$.***

Доказательство теоремы 4. В силу леммы 3, по модулю φ_n^3, а значит, и по модулю $p^2\varphi_n$ верно сравнение

$$\varphi_{np}\equiv\varphi_{n+1}^p-\varphi_{n-1}^p=(\varphi_{n+1}-\varphi_{n-1})(\varphi_{n+1}^{p-1}+\varphi_{n+1}^{p-2}\varphi_{n-1}+\ldots+\varphi_{n+1}\varphi_{n-1}^{p-2}+\varphi_{n-1}^{p-1}).$$

Первый множитель равен φ_n. Поскольку $\varphi_{n+1}=\varphi_{n-1}+\varphi_n\equiv\varphi_{n-1}\pmod p$, второй множитель сравним с $p\varphi_{n-1}^{p-1}\equiv0\pmod p$. Докажем, что второй множитель не делится на p^2. Поскольку числа φ_{n+1} и φ_{n-1} не кратны p, то существует такое целое число a, что $\varphi n+1\equiv a\varphi_{n-1}\pmod{p^2}$. При этом второй множитель сравним по модулю p^2 с числом $\varphi_{n-1}^{p-1}(a^{p-1}+a^{p-2}+\ldots+a+1)$, где, очевидно, $a\equiv1\pmod p$. Именно такая ситуация рассмотрена в лемме 2 статьи «Периодические дроби»! ■

Очевидной индукцией из теоремы 4 получаем: $\varphi_{5^n}\vdots5^n$ и $\varphi_{3^n\cdot4}\vdots3^n\cdot4$ для любого натурального n. Таким образом, утверждение теоремы 3 мы уже доказали для степеней простых чисел 2, 3 и 5.

Завершим доказательство теоремы 3. Применим индукцию по наибольшему простому числу, входящему в разложение на простые множители числа n.

Очевидно, как только мы для некоторого простого числа p нашли такое кратное p натуральное число m, что $\varphi_m\vdots m$ и $m\vdots p$, так вследствие теоремы 4 имеем $\varphi_{mp^r}\vdots mp^r$ и $mp^r\vdots p^{r+1}$ для любого натурального r. Как видите, показатель степени, с которым входит число p в разложение числа m на простые множители, можно взять сколь угодно большим, если он хоть раз оказался положительным.

Как же сделать его положительным? При $p>5$, как доказано выше, φ_{p-1} или φ_{p+1} кратно простому числу p. Числа $p+1$ и $p-1$ четные, поэтому в их разложениях на простые множители все сомножители меньше p. Следовательно, можно пользоваться индукционным предположением: можно считать, что существует такое m, что $\varphi_m\vdots m\vdots(p\pm1)$, где знак выбран должным образом. Очевидно, число $\varphi_{p\,\mathrm{НОК}[m,p\pm1]}$ делится на число p НОК$[m,p\pm1]$, которое, в свою очередь, делится на p. ■

Продолжим изучение арифметических свойств чисел Фибоначчи.

Теорема 5. ***Если p и q — простые числа, причем φ_n делится на q, то φ_{np}/φ_n не делится на q.***

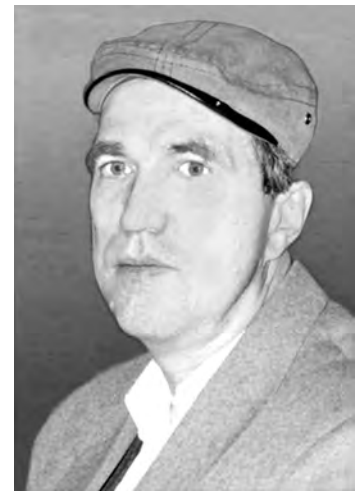
Доказательство. В силу леммы 2 имеем $\varphi_{np}\equiv p\varphi_n\varphi_{n+1}^{p-1}\pmod{\varphi_n^2}$. Ни p, ни φ_{n+1} не кратны q. Теорема 5 доказана.

В книге Н. Н. Воробьева «Числа Фибоначчи» при помощи утверждений типа теорем 4 и 5 доказано, что любое число Фибоначчи, кроме $\varphi_1=\varphi_2=1$, $\varphi_6=8$ и $\varphi_{12}=144$, имеет хотя бы один простой делитель, которым не обладает ни одно из предыдущих чисел Фибоначчи. ■

В доказательстве диофантовости перечислимых подмножеств множества целых чисел Ю. В. Матиясевич использовал следующее утверждение.

Теорема 6. ***φ_{kn} делится на φ_n^2 тогда и только тогда, когда k делится на φ_n.***

Доказательство. В силу леммы 2 имеем $\varphi_{kn}\equiv k\varphi_n\varphi_{n+1}^{k-1}\pmod{\varphi_n^2}$. Поскольку число φ_{n+1} взаимно просто с φ_n, то φ_{kn} делится на φ_n^2 тогда и только тогда, когда k делится на φ_n. ■



Юрий Владимирович Матиясевич родился в 1947 г. В 1964 г. получил золотую медаль на Международной математической олимпиаде. В 1970 г. чрезвычайно сильно продвинулся в решении десятой проблемы Гильберта. Проблема заключалась в том, чтобы научиться по любому многочлену f узнавать, существует ли такой набор рациональных чисел $(x_1,x_2,\ldots,x_k)$, что $f(x_1,x_2,\ldots,x_k)=0$.

Матиясевич заменил слово «рациональные» на «целые» и доказал весьма неожиданный результат: такого алгоритма нет. (В исходной формулировке проблема еще не решена.)

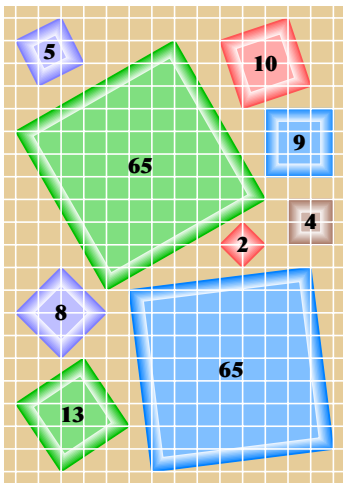
Более того, для любого перечислимого множества $M\subset\mathbb{Z}$ существует такой многочлен $f(m,x_1,x_2,\ldots,x_n)$ с целыми коэффициентами, что $m\in M$ тогда и только тогда, когда для хотя бы одного набора целых чисел x_1, x_2, …, x_n верно равенство $f(m,x_1,x_2,\ldots,x_n)=0$.

Поясним. Множество перечислимо, если существует алгоритм, который выводит на печать его элементы (и только их!). Если можно сделать так, что элементы множества будут появляться по порядку — сначала наименьший элемент, затем второй по величине, третий и так далее, то множество называют разрешимым. Одна из теорем математической логики гласит: существует перечислимое неразрешимое множество. Именно на ней основано доказательство несуществования алгоритма, который по уравнению узнает, разрешимо ли оно в целых числах. ■

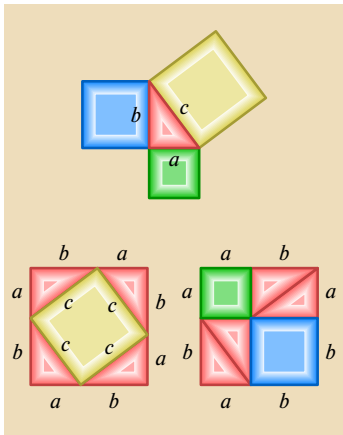
Суммы двух квадратов

«Зачем складывать простые числа? — недоумевал физик Л. Д. Ландау. — Простые числа надо умножать, а не складывать!» «Зачем складывать квадраты целых чисел? — спросите вы. — Почему бы не складывать их кубы или 1000-е степени?» Да, не все задачи достойны пристального внимания. Задача о сумме квадратов — в высшей степени достойна. Чтобы узнать, какие числа представимы, а какие нет, мы используем не только «обычные», но и гауссовы целые числа — прекрасный пример применения абстрактной теории к конкретной арифметической задаче! Хотя эта статья содержит лишь малую часть теории делимости алгебраических чисел, надеемся, ее очарование не оставит вас равнодушным.

Однажды у индийского математика С. Рамануджана (1887—1920) спросили, чем замечательно число 1729. «Так это же наименьшее натуральное число, представимое в виде суммы кубов двух натуральных чисел двумя разными способами!» — воскликнул он. И действительно, $1729 = 1^3 + 12^3 = 9^3 + 10^3$. ■



На клетчатой бумаге изображены квадраты площади $2 = 1^2 + 1^2$, $4 = 2^2 + 0^2$, $5 = 2^2 + 1^2$, $8 = 2^2 + 2^2$, $9 = 3^2 + 0^2$, $10 = 3^2 + 1^2$, $13 = 3^2 + 2^2$ и $65 = 8^2 + 1^2 = 7^2 + 4^2$. Все площади — суммы двух квадратов. Рисунки внизу иллюстрируют теорему Пифагора: площадь квадрата, построенного на гипотенузе прямоугольного треугольника, равна сумме площадей квадратов, построенных на его катетах. ■



Рассмотрим таблицу, в верхней строке и левом столбце которой — квадраты целых чисел, а в других клетках — суммы квадратов.

0	1	4	9	16	25	36	49	64	81	100
1	2	5	10	17	26	37	50	65	82	101
4	5	8	13	20	29	40	53	68	85	104
9	10	13	18	25	34	45	58	73	90	109
16	17	20	25	32	41	52	65	80	97	116
25	26	29	34	41	50	61	74	89	106	125
36	37	40	45	52	61	72	85	100	117	136
49	50	53	58	65	74	85	98	113	130	149
64	65	68	73	80	89	100	113	128	145	164
81	82	85	90	97	106	117	130	145	162	181
100	101	104	109	116	125	136	149	164	181	200

Некоторые числа представимы несколькими способами, например, $25 = 5^2 + 0^2 = 3^2 + 4^2$ и $65 = 8^2 + 1^2 = 7^2 + 4^2$. Не вошедшие в таблицу числа первой сотни (3, 6, 7, 11, 12, 14, 15, ...) в виде суммы двух квадратов не представимы. Наименьшее не представимое число — это 3. Кратные 3 числа 6, 12, 15, 21 тоже не представимы, а вот числа $9 = 3^2 + 0^2$ и $18 = 3^2 + 3^2$ — представимы. Возникает гипотеза: числа, которые кратны 3, но не кратны 9, не представимы в виде суммы двух квадратов. Верно даже более сильное утверждение:

Теорема 1. ***Если сумма квадратов двух целых чисел кратна 3, то слагаемые тоже кратны 3.***

Доказательство. Выпишем остатки от деления квадратов целых чисел на 3:

n^2	0	1	4	9	16	25	36	49	64	81	100
$n^2 \bmod 3$	0	1	1	0	1	1	0	1	1	0	1

Закономерность очевидна: остатки периодически повторяются, и никаких остатков, кроме 0 и 1, не бывает. Точнее говоря, остаток от деления квадрата целого числа x на 3 равен 0, если x кратно 3, то есть представимо в виде $x = 3k$, где k — целое число, и остаток равен 1, если x не кратно 3, то есть представимо в виде $x = 3k \pm 1$. В самом деле, в первом случае $x^2 = 9k^2$ делится на 3 без остатка, а во втором случае $x^2 = 9k^2 \pm 6k + 1$ дает при делении на 3 остаток 1.

Суммы остатков $0 + 1$ и $1 + 1$ не кратны 3. Значит, сумма квадратов $x^2 + y^2$ кратна 3 в том и только том случае, когда x и y кратны 3. ■

Следующее после 3 и 6 не представимое в виде суммы двух квадратов число — это 7. Кратные 7 числа 14, 21, 28, 35, 42, 56, 63 тоже не представимы. Опять возникает гипотеза: если сумма квадратов x^2+y^2 кратна 7, то и сами целые числа x и y кратны 7. Составим таблицу остатков:

n^2	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196
$n^2 \bmod 7$	0	1	4	2	2	4	1	0	1	4	2	2	4	1	0

Сумма никаких двух из чисел 1, 2, 4 не кратна 7. Гипотеза верна! ■

Еще один пример — $p=19$. Составим таблицу остатков:

n^2	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196	225	256	289	324
$n^2 \bmod 19$	0	1	4	9	16	6	17	11	7	5	5	7	11	17	6	16	9	4	1

В верхней строке — квадраты чисел 0, 1, . . . , 18. (Другие можно не рассматривать, поскольку любое целое число x можно представить в виде $x=19q+r$, где q — целое, $0\leqslant r\leqslant 18$, и число $x^2=19^2q^2+38qr+r^2$ дает при делении на 19 такой же остаток, как и r^2.) Поскольку сумма никаких двух из чисел 1, 4, 5, 6, 7, 9, 11, 16 и 17 не кратна 19, приходим к выводу: сумма квадратов двух целых чисел кратна 19 в том и только том случае, когда слагаемые кратны 19. ■

Итак, простые числа $p=3$, 7, 11, 19 обладают тем свойством, что если сумма квадратов кратна p, то каждое из слагаемых кратно p.

Теорема 2. ***Если сумма квадратов a^2+b^2 целых чисел a и b делится на простое число $p=4n+3$, где n — целое неотрицательное число, то числа a и b делятся на p.***

Доказательство. Пусть a не делится на p. Тогда и b не делится на p. Возведем обе части сравнения $a^2\equiv -b^2 \pmod p$ в $(2n+1)$-ю степень:

$$a^{4n+2}\equiv -b^{4n+2} \pmod p.$$

В силу малой теоремы Ферма имеем $a^{4n+2}\equiv 1\equiv b^{4n+2} \pmod p$, поэтому $1\equiv -1 \pmod p$, что невозможно при $p>2$. ■

Произведение сумм квадратов. Если $n=x^2+y^2$, то

$$(x+y)^2+(x-y)^2=x^2+2xy+y^2+x^2-2xy+y^2=2(x^2+y^2)=2n.$$

Значит, если n представимо, то представимо и $2n$. Далее,

$$(2x+y)^2+(x-2y)^2=4x^2+4xy+y^2+x^2-4xy+4y^2=5(x^2+y^2)=5n.$$

Легко проверить и формулы

$$(2x+3y)^2+(3x-2y)^2=13n,$$
$$(4x+y)^2+(x-4y)^2=17n.$$

Все они являются частными случаями общей формулы, которая представляет произведение сумм двух квадратов в виде суммы двух квадратов. Чтобы получить ее, раскроем скобки, прибавим и отнимем $2abxy$ и изменим порядок слагаемых:

$$(a^2+b^2)(x^2+y^2)=a^2x^2+b^2x^2+a^2y^2+b^2y^2=$$
$$=a^2x^2-2axby+b^2y^2+b^2x^2+2bxay+a^2y^2=(ax-by)^2+(bx+ay)^2.$$ ■

Выясним, какие простые числа представимы в виде суммы двух квадратов целых чисел. Очевидно, $5=2^2+1^2$, $13=3^2+2^2$, $17=4^2+1^2$, $29=5^2+2^2$, $37=6^2+1^2$, $41=5^2+4^2$, $53=7^2+2^2$, . . .

Теорема Ферма—Эйлера. ***Любое простое число p, которое при делении на 4 дает остаток 1, представимо в виде суммы квадратов двух натуральных чисел.***

Эту теорему сформулировал П. Ферма, а доказал ее (при помощи любимого Ферма метода бесконечного спуска) в 1747—1749 гг. Л. Эйлер.

В виде суммы двух квадратов не представимы не только простые числа, которые при делении на 4 дают остаток 3, но и вообще никакие такие числа. Точнее говоря, всякое представимое в виде суммы квадратов двух целых чисел нечетное число при делении на 4 дает остаток 1, а не 3. Дело в том, что из двух квадратов, сумма которых нечетна, обязательно один четен, а другой нечетен. Квадрат любого четного числа нацело делится на 4, а квадрат нечетного числа при делении на 4 (а на самом деле даже на 8) дает остаток 1 (проверьте!). ■

Натуральное число представимо в виде суммы квадратов двух целых чисел тогда и только тогда, когда в его разложение на простые множители любой простой множитель вида $4k+3$ входит в четной степени.
Этот критерий впервые сформулировал голландец А. Жирар (1595—1632) в следующем виде: натуральное число представимо в виде суммы двух квадратов тогда и только тогда, когда оно является или квадратом, или числом 2, или простым числом, которое на 1 больше, чем некоторое кратное 4, или произведением нескольких вышеперечисленных чисел. Скорее всего, Жирар опирался лишь на изучение таблиц и не умел доказывать необходимость и достаточность своих условий. ■

Ферма писал: «Если бы выбранное простое число, которое на единицу больше некоторого числа, делящегося на 4, не было суммой квадратов, то существовало бы простое число такой же природы, меньшее заданного, а затем еще и третье, и так далее, бесконечно убывая до тех пор, пока не будет достигнуто простое число 5, которое является наименьшим из всех чисел такой природы; отсюда следовало бы, что 5 не является суммой двух квадратов, что не соответствует действительности. Отсюда сведением к абсурду следует заключить, что все числа такой природы являются суммами двух квадратов». ■

Лемма 1. ***Если сумма квадратов кратна простому числу, являющемуся суммой квадратов, то частное — тоже сумма квадратов.***

Доказательство. Пусть a^2+b^2 делится на простое число $p=r^2+s^2$. Тогда

$$(ar-bs)(ar+bs)=a^2r^2-b^2s^2=(a^2+b^2)r^2-b^2(r^2+s^2)\vdots p$$

Значит, $ar-bs$ или $ar+bs$ кратно p. Пусть, для определенности, $ar-bs\vdots p$. Поскольку $(ar-bs)^2+(as+br)^2=(a^2+b^2)(r^2+s^2)\vdots p$, то $as+br\vdots p$ и, следовательно,

$$\frac{a^2+b^2}{p}=\left(\frac{ar-bs}{p}\right)^2+\left(\frac{as+br}{p}\right)^2$$

— представление числа $(a^2+b^2)/p$ в виде суммы квадратов двух целых чисел. Случай, когда $ar+bs$ делится на p, легко рассмотреть аналогично, воспользовавшись тождеством $(a^2+b^2)(r^2+s^2)=(ar+bs)^2+(as-br)^2$.

Лемма 2. ***Всякое натуральное число, являющееся делителем суммы квадратов двух взаимно простых чисел, является суммой двух квадратов.***

Доказательство. Пусть сумма a^2+b^2 кратна m, причем НОД$(a;\,b)=1$. Представим числа a и b в виде $a=mx+c$ и $b=my+d$, где c и d по абсолютной величине не превосходят $m/2$. Тогда

$$c^2+d^2=(a-mx)^2+(b-my)^2=(a^2+b^2)-2amx+m^2x^2-2bmy+m^2y^2\vdots m.$$

Значит, $c^2+d^2=mn$, где n — целое, причем $n=\dfrac{c^2+d^2}{m}\leqslant\dfrac{(m/2)^2+(m/2)^2}{m}=\dfrac{m}{2}$.

Если $n=0$, то m — общий делитель (взаимно простых!) чисел a и b, так что $m=1=1^2+0^2$.

Пусть $n>0$. Очевидно, m взаимно просто с наибольшим общим делителем чисел c и d. Разделив числа c и d на НОД$(c,\,d)$, видим, что n — делитель суммы квадратов взаимно простых чисел. Если все простые делители числа n — суммы квадратов, то в силу леммы 1 число m — сумма квадратов. Если же хотя бы один из них — не сумма квадратов, аналогично получаем меньшее число, являющееся делителем суммы квадратов взаимно простых чисел и не представимое в виде суммы квадратов, и так далее. Но бесконечной убывающей последовательности натуральных чисел не существует!

Лемма 3. ***Если $p=4n+1$ — простое число, то хотя бы одна из сумм $1^{2n}+2^{2n}$, $2^{2n}+3^{2n}$, . . . , $(4n-1)^{2n}+(4n)^{2n}$ кратна p.***

Доказательство. В силу малой теоремы Ферма каждое из чисел 1^{4n}, 2^{4n}, 3^{4n}, ... , $(4n-1)^{4n}$, $(4n)^{4n}$ дает при делении на p остаток 1. Следовательно, все числа вида $(a+1)^{4n}-a^{4n}=((a+1)^{2n}-a^{2n})((a+1)^{2n}+a^{2n})$, где $1\leqslant a<4n$, кратны p. Если ни одна из сумм $(a+1)^{2n}+a^{2n}$ не кратна p, то все разности $(a+1)^{2n}-a^{2n}$ кратны p и поэтому $a^{2n}\equiv 1\pmod p$ для любого $a=1,\,2,\,\ldots,\,p-1$. Но многочлен степени $2n$ не может иметь $4n$ корней! ■

Другой способ доказательства теоремы Ферма—Эйлера дают теорема Дж. Вильсона (1741—1793), впервые сформулированная англичанином Э. Варингом (1734—1798) и доказанная Ж. Л. Лагранжем, и леммы 4 и 5.

Теорема Вильсона. ***Для любого простого числа p сумма $(p-1)!+1$ кратна p.***

Идею доказательства продемонстрируем на примере числа $p=17$:

$$16!=1\cdot(2\cdot9)\cdot(3\cdot6)\cdot(4\cdot13)\cdot(5\cdot7)\cdot(8\cdot15)\cdot(10\cdot12)\cdot(11\cdot14)\cdot16\equiv16\equiv-1\pmod{17}.$$

Вообще, для любого простого числа $p>3$ числа $2,\,3,\,\ldots,\,p-2$ можно разбить на такие пары $(a;\,b)$, что $ab\equiv1\pmod p$. (Подумайте почему.)

«От противного» легко доказать, что если $(n-1)!+1\vdots n$, где $n>1$, то n — простое. К сожалению, никакой пользы для вычислений это не дает: при сколь-нибудь значительном n число $(n-1)!$ слишком велико. ■

Жозеф Луи Лагранж (1736—1813) — французский математик и механик. Родился в итальянском городе Турине, в семье военного казначея, одиннадцатым ребенком в семье. Учился в артиллерийской школе и в 19 лет стал ее профессором.

Для любой непрерывной в точках a и b и дифференцируемой на интервале $(a;\,b)$ функции f существует такая точка $\xi\in(a;\,b)$, что выполнена формула конечных приращений Лагранжа

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

Эта формула — следствие теоремы Ролля: для любой непрерывной в точках a и b и дифференцируемой на интервале $(a;\,b)$ функции g, принимающей на концах отрезка равные значения, существует такая точка $\xi\in(a;\,b)$, что $g'(\xi)=0$. Чтобы вывести формулу Лагранжа из теоремы Ролля, достаточно рассмотреть

$$g(x)=f(x)-\frac{f(b)-f(a)}{b-a}\cdot x.$$

Для любых точек $(x_0;\,y_0)$, $(x_1;\,y_1)$, . . . , $(x_n;\,y_n)$, все $n+1$ абсцисс которых различны, существует и единственен многочлен f степени не выше n, график которого проходит через данные точки. Единственность очевидна: если бы существовали не равные такие многочлены f_1 и f_2, то их разность была бы многочленом степени не выше n, имеющим более n корней. В силу теоремы Безу многочлен, равный нулю в точках $x_0,\,x_1,\,\ldots,\,x_n$, делится на $x-x_0$, $x-x_1$, . . . , $x-x_n$; поэтому его степень не может быть меньше $n+1$.

А существование проще всего доказать при помощи интерполяционной формулы Лагранжа, которая для $n=1$ имеет вид

$$f(x)=y_0\cdot\frac{x-x_1}{x_0-x_1}+y_1\cdot\frac{x-x_0}{x_1-x_0},$$

при $n=2$ — вид

$$f(x)=y_0\cdot\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}+ +y_1\cdot\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}+ +y_2\cdot\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)},$$

а вообще выглядит так:

$$f(x)=\sum_{0\leqslant k\leqslant n}y_k\prod_{\substack{0\leqslant m\leqslant n,\\ m\neq k}}\frac{x-x_m}{x_k-x_m}.$$

Лагранж заложил основы вариационного исчисления, ввел знак δ вариации. В «Аналитической механике» (1788) свел разнообразие задач механики к одному принципу: чтобы некоторая величина — лагранжиан — принимала минимальное значение. Он писал: «Я имел в виду привести всю теорию механики и искусство решения ее задач к общим формулам, простое развитие которых давало бы все необходимые для решения задачи уравнения». Об особенностях изложения: «В этом сочинении нет чертежей. Методы, в нем излагаемые, не требуют ни геометрических построений, ни механических рассуждений, для них требуются лишь алгебраические операции, подчиненные правильному и однообразному ходу. Любители анализа с удовольствием увидят, что механика становится новою его отраслью, и будут мне признательны за такое расширение его области».

В 1770—1771 гг. Лагранж опубликовал 200-страничный мемуар «Размышления об алгебраическом решении уравнений», где изучал перестановки корней многочлена, заложив основы теории групп и теории Галуа.

Резольвентами Лагранжа называют выражения вида

$$x_1\varepsilon+x_2\varepsilon^2+\ldots+x_n\varepsilon^n,$$

где $x_1, x_2, \ldots, x_n$ — корни уравнения n-й степени, ε — корень n-й степени из 1. Его имя носят метод приведения квадратичной формы к диагональному виду и метод решения задач на условный экстремум. ■

Лемма 4. ***Для любого простого числа $p=4n+1$, где n — натуральное, существует такое целое число m, что m^2+1 кратно p.***

Доказательство. Годится $m=(2n)!$. В самом деле,

$$(p-1)!=1\cdot2\cdot\ldots\cdot(2n-1)\cdot(2n)\cdot(2n+1)\cdot(2n+2)\cdot\ldots\cdot(4n-1)\cdot(4n)= =1\cdot2\cdot\ldots\cdot(2n-1)\cdot(2n)\cdot(p-2n)\cdot(p-(2n-1))\cdot\ldots\cdot(p-2)\cdot(p-1).$$

Это число дает при делении на p такой же остаток, как и число

$$1\cdot2\cdot\ldots\cdot(2n-1)\cdot(2n)\cdot(-1)^{2n}\cdot(2n)\cdot(2n-1)\cdot\ldots\cdot2\cdot1=m^2.$$

Значит, $m^2\equiv(p-1)!\pmod p$. Сумма $(p-1)!+1$ кратна p по теореме Вильсона.

Лемма 5. ***Любой простой делитель p числа m^2+1, где m — целое, представим в виде суммы квадратов двух натуральных чисел.***

Доказательство Ж. Л. Лагранжа. Рассмотрим все такие пары $(r; s)$ целых чисел, что $0\leqslant r$, $s\leqslant\sqrt{p}$, и для каждой пары рассмотрим остаток от деления числа $r+ms$ на p. Поскольку количество таких пар равно $([\sqrt{p}]+1)^2>p$, среди них есть такие две пары $(r_1; s_1)$ и $(r_2; s_2)$, что остатки от деления на p чисел r_1+ms_1 и r_2+ms_2 равны.

При этом число $r+ms$, где $r=r_1-r_2$ и $s=s_1-s_2$, кратно p. Поэтому число

$$r^2+s^2=r^2-m^2s^2+(m^2+1)s^2=(r+ms)(r-ms)+(m^2+1)s^2$$

тоже кратно p. Заметим, что $0<r^2+s^2<p+p=2p$. Единственное кратное p число, которое больше 0, но меньше $2p$, — само число p. Значит, $r^2+s^2=p$. ■

Лемму 5 можно доказать, рассматривая приближения числа m/p рациональными числами. Как доказано в статье «Уравнения Пелля», для любого вещественного числа ξ и любого натурального числа N существуют такие целое число t и натуральное число s, что $s\leqslant N$ и $|s\xi-t|\leqslant\dfrac{1}{N+1}$.

Случай $m=1$ леммы 5 тривиален. Предположим, что $m\geqslant2$. Пусть $\xi=m/p$ и $N=[\sqrt{p}]$. Рассмотрим такие t и s, что

$$|ms-tp|\leqslant\frac{p}{N+1}=\frac{p}{[\sqrt{p}]+1}<\sqrt{p}.$$

Обозначим $ms-tp=r$. Сумма квадратов

$$r^2+s^2=(ms-tp)^2+s^2=s^2(m^2+1)-2mstp+t^2p^2$$

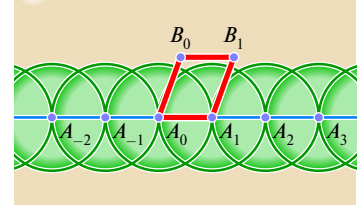
делится на p. Поскольку $s\leqslant N<\sqrt{p}$ и $r<\sqrt{p}$, то $r^2+s^2<2p$. Как и в доказательстве Лагранжа, заключаем: $r^2+s^2=p$. ■

Геометрическое доказательство леммы 5 основано на следующей лемме.

Лемма 6. ***Для любого параллелограмма $OABC$ площади S хотя бы одна из вершин порожденной им решетки удалена от точки O не более чем на $\sqrt{2S/\sqrt{3}}$.***

Доказательство. Выбрав находящиеся на наименьшем расстоянии ρ точки решетки A_0 и A_1 (рис. 1), видим, что решетка содержит все точки A_k, где $k\in\mathbb{Z}$ и $\overrightarrow{A_0A_k}=k\overrightarrow{A_0A_1}$, и не содержит ни одной точки внутри кругов радиуса ρ с центрами в этих точках; поэтому высота основного параллелограмма $A_0A_1B_1B_0$ решетки, опущенная на сторону A_0A_1, не меньше $\rho\sin60^\circ$, откуда $S\geqslant\rho^2\sqrt{3}/2$. ■

1



Лемма Г. Минковского о квадратичной форме. ***Если a, b, c — целые числа, $a>0$ и $ac-b^2=1$, то существуют такие целые числа x и y, что $ax^2+2bxy+cy^2=1$.***

Доказательство. Рассмотрим векторы $\overrightarrow{OA}$ и $\overrightarrow{OC}$ длин $\sqrt{a}$ и $\sqrt{c}$ соответственно, угол φ между которыми выберем так, чтобы скалярное произведение равнялось b, то есть $\cos\varphi=b/\sqrt{ac}$. Площадь S параллелограмма $OABC$ равна

$$S=OA\cdot OC\cdot\sin\varphi=\sqrt{a}\sqrt{c}\sqrt{1-\cos^2\varphi}=\sqrt{a}\sqrt{c}\sqrt{\frac{ac-b^2}{ac}}=1.$$

Существуют такие целые числа x и y, хотя бы одно из которых отлично от нуля, что $|x\overrightarrow{OA}+y\overrightarrow{OC}|\leqslant\sqrt{2/\sqrt{3}}$. Следовательно,

$$ax^2+2bxy+cy^2=(x\overrightarrow{OA}+y\overrightarrow{OC})^2\leqslant 2/\sqrt{3}<2.$$

Поскольку скалярный квадрат любого ненулевого вектора положителен, то $ax^2+2bxy+cy^2>0$. А поскольку между нулем и двойкой есть только одно целое число — единица, то $ax^2+2bxy+cy^2=1$. Лемма Минковского доказана.

Применяя ее к числам $a=p$, $b=m$ и $c=\dfrac{m^2+1}{p}$, получаем для некоторых целых чисел x и y равенство $1=px^2+2mxy+cy^2$. Домножая обе его части на p, получаем

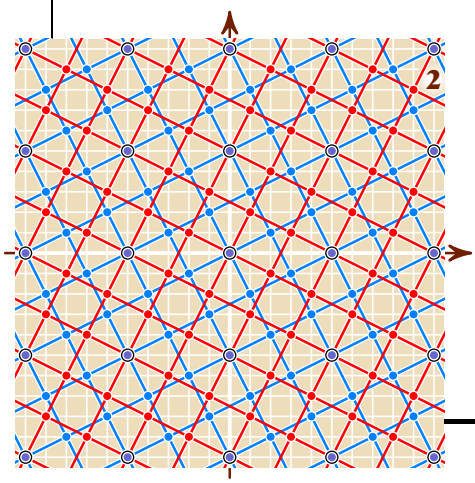
$$p=p^2x^2+2pmxy+(m^2+1)y^2=(px+my)^2+y^2.\ \blacksquare$$

Лемму Минковского о квадратичной форме можно доказать геометрически и другим способом. Он основан на лемме Минковского о выпуклом теле и формуле площади эллипса.

А именно, множество точек $(x;y)$ плоскости, для которых $ax^2+2bxy+cy^2<2$, — внутренность эллипса площади $2\pi/(ac-b^2)=2\pi>4$. В силу леммы Минковского о выпуклом теле, внутри эллипса существует хотя бы одна точка $(x;y)\neq(0;0)$ с целыми координатами. Эта точка искомая, поскольку для нее величина $ax^2+2bxy+cy^2$ меньше 2 и больше 0, то есть равна 1. ■

Изложенные доказательства леммы 5 производят впечатление то ли чудес, то ли фокусов. Проникнуть в суть явления помогает арифметика целых гауссовых чисел. Она не только дает естественное доказательство леммы 5, но и позволяет легко получить формулу для количества представлений числа в виде суммы двух квадратов. ■

Комплексное число $a+bi$ называют целым гауссовым, если a и b — целые числа. Сумма, разность и произведение целых гауссовых чисел — целые гауссовы, так что множество $\mathbb{Z}[i]$ целых гауссовых чисел является, как говорят алгебраисты, кольцом. По определению, целое гауссово число u кратно целому гауссову числу v, если существует такое целое гауссово число w, что $u=vw$.



Отметив на плоскости целые гауссовы числа, получаем решетку. Числа, кратные данному числу z, тоже образуют решетку. Она получается из решетки целых гауссовых чисел растяжением в $|z|$ раз и поворотом на угол $\arg z$. На рисунке 2 синим цветом выделены кратные числа $2+i$, а красным — кратные числа $2-i$. Интересно, какие целые гауссовы числа являются кратными и числа $2+i$,

В седьмом замечании П. Ферма на полях «Арифметики» Диофанта сказано: «Простое число, которое на единицу превосходит кратное четырех, только один раз является гипотенузой прямоугольного треугольника, его квадрат — два раза, куб — три раза, биквадрат (биквадрат — четвертая степень, квадрато-куб — пятая, кубо-куб — шестая. — *Примеч. ред.*) — четыре и так далее до бесконечности.

Это же простое число и его квадрат только одним способом разлагаются на два квадрата (натуральных чисел. — *Примеч. ред.*), его куб и биквадрат — двумя, квадрато-куб и кубо-куб — тремя и так далее до бесконечности.

Если простое число, представимое суммой двух квадратов, умножено на другое простое число, тоже представимое суммой двух квадратов, то произведение дважды представимо суммой двух квадратов; если умножено на квадрат второго простого числа, то произведение трижды представимо суммой двух квадратов; если умножено на куб второго простого числа, то произведение представимо суммой двух квадратов четырьмя способами; и так до бесконечности. . .

Пусть надо найти число, которое было бы гипотенузой семью различными способами. Данное число 7 удваиваем, будет 14. Прибавляем единицу, будет 15. Берем все простые делители числа 15: это 3 и 5. Вычитаем из каждого единицу и берем половины остатков; получаем 1 и 2. Возьмем теперь столько различных простых множителей, сколько здесь чисел, а именно два, и перемножим их между собой с показателями 1 и 2, а именно один на квадрат другого; так получаем число, удовлетворяющее условиям задачи, только бы взятые простые множители превосходили кратные четырех на единицу. . .

А вот метод узнать, сколькими способами данное число может быть составлено из двух квадратов. Пусть данное число 325. Его простые делители, которые превосходят на единицу кратное четырех, это 5 и 13; последний — один раз, а первый — в квадрате. Возьмем показатели 2 и 1. Сложим их произведение и сумму, получится 5; прибавим еди-

ницу, получится число 6, половина которого — 3. Столькими способами данное число составляется из двух квадратов.
Если бы было три показателя, например, 2, 2, 1, то действовать надо было бы так. Произведение двух первых, сложенное с их суммой, даст 8. Умножаем на третий и прибавляем сумму, что дает число 17. Прибавляем к нему единицу, будет 18; половина есть 9. Столькими способами предложенное число составляется из двух квадратов.
Если последнее число, которое нужно разделить пополам, нечетно, то от него следует отнять единицу и взять половину остатка». ■

В «Арифметических исследованиях» К. Ф. Гаусса сказано: «Если $M=2^{\mu}Sa^{\alpha}b^{\beta}c^{\gamma}\ldots$, где a, b, c, ... обозначают различные простые числа вида $4n+1$, и S — произведение всех простых сомножителей числа M вида $4n+3$ (в таком виде можно представить любое положительное целое число, если положить $\mu=0$, когда M — нечетно, и $S=1$, когда M не содержит сомножителей вида $4n+3$), то M не может быть разложено на два квадрата, если S не является квадратом. Если же S есть квадрат, то имеется

$$\frac{1}{2}(\alpha+1)(\beta+1)(\gamma+1)\ldots$$

разложений числа M, когда хотя бы одно из чисел α, β, γ, ... нечетно, и

$$\frac{1}{2}(\alpha+1)(\beta+1)(\gamma+1)\ldots+\frac{1}{2}$$

разложений, когда α, β, γ, ... все четны». Математически точная и почти современная по языку формулировка, не правда ли? ■

Пусть $f(n)$ — количество не превосходящих числа n натуральных чисел, представимых в виде суммы квадратов двух целых чисел. Э. Г. Г. Ландау (1877—1938) доказал в 1908 г., что

$$\lim_{n\to\infty}\frac{f(n)\sqrt{\ln n}}{n}=K,$$

где $K=0{,}764\ldots$, точнее,

$$K=\sqrt{\frac{1}{2}\prod_{\substack{p\ —\ \text{простое},\\ p\equiv 3\ (\mathrm{mod}\ 4)}}\frac{p^2}{p^2-1}}.$$ ■

и числа $2-i$ одновременно? Ответ очевиден: пересечение множеств «синих» и «красных» чисел состоит из чисел, кратных 5. Другими словами, наименьшее общее кратное чисел $2+i$ и $2-i$ равно 5.
Вообще, произведение $(a+bi)(a-bi)=a^2+b^2$ любого комплексного числа $z=a+bi$ и сопряженного к нему числа $\overline{z}=a-bi$ является числом вещественным. Поэтому для любого ненулевого целого гауссова числа z существует кратное ему натуральное число $z\overline{z}=a^2+b^2$. ■

Делители единицы. Очевидно, $1=1\cdot 1=i\cdot(-i)=(-1)\cdot(-1)=(-i)\cdot i$. Значит, числа ± 1 и $\pm i$ — обратимые элементы (делители единицы) кольца $\mathbb{Z}[i]$. Других способов разложить 1 в произведение двух целых гауссовых чисел нет.

Теорема 3. ***В $\mathbb{Z}[i]$ нет делителей единицы, кроме 1, i, -1 и $-i$.***

Доказательство. Если $1=uv$, где u, $v\in\mathbb{Z}[i]$, то $1=|u|\cdot|v|$. Поскольку модули ненулевых целых гауссовых чисел u и v не меньше 1, то $|u|=|v|=1$. ■

Ассоциированными называют числа u и v, если они кратны друг другу: u кратно v и v кратно u. Всякое целое гауссово число z можно представить в виде произведений

$$z=1\cdot z=i(-iz)=(-1)(-z)=(-i)(iz),$$

первый множитель каждого из которых — делитель единицы, а второй ассоциирован с числом z. Далее, если целое гауссово число w кратно числу z, то делителями числа w являются и числа $-z$, iz, $-iz$. Поэтому, рассматривая разложения на множители, можно «не различать» ассоциированные числа. ■

Вернемся к лемме 5, от которой мы отвлеклись, чтобы придать смысл разложению $m^2+1=(m+i)(m-i)$. Числу p не кратен ни один из множителей $m+i$ и $m-i$, но кратно произведение m^2+1. Как может произведение быть кратно p, если ни один из множителей не кратен p? Неужели арифметика гауссовых чисел настолько своеобычна, что в ней не действуют привычные нам законы, например основная теорема арифметики?
Нет, действуют! В статье «Основная теорема арифметики» при помощи алгоритма деления с остатком для кольца $\mathbb{Z}$ целых чисел доказано, что разложение на простые множители единственно. А делить с остатком можно и целые гауссовы числа: для любого целого гауссова числа w и ненулевого целого гауссова числа z расстояние от точки w до ближайшей к ней точке решетки, состоящей из кратных числа z, меньше $|z|$ (и даже не превышает $|z|/\sqrt{2}$).
Поэтому разложение целых гауссовых чисел на простые гауссовы множители единственно в том же смысле, в каком оно единственно для обычных целых чисел — с точностью до перестановки множителей и до ассоциированности. Однако некоторые простые числа p перестают быть простыми при расширении $\mathbb{Z}$ до $\mathbb{Z}[i]$. Например, $2=(1+i)(1-i)=-i(1+i)^2$ и $5=(1+2i)(1-2i)$.

Докажем лемму 5. Делитель p числа $(m+i)(m-i)$ не может быть простым гауссовым числом. Значит, $p=(a+bi)(c+di)$, где целые гауссовы числа $(a+bi)$ и $(c+di)$ — не делители единицы. Поскольку модуль произведения равен произведению модулей, имеем

$$p=\sqrt{a^2+b^2}\sqrt{c^2+d^2},$$

то есть $p^2=(a^2+b^2)(c^2+d^2)$, откуда $p=a^2+b^2=c^2+d^2$. ■

Какие же простые натуральные числа остаются простыми во множестве целых гауссовых чисел, а какие становятся составными? И как устроены разложения «новых составных» чисел?

Гаусс доказал, что в виде суммы квадратов трех целых чисел представимы все натуральные числа, кроме чисел вида $4^m(8n+7)$, где m, n — целые неотрицательные числа.
Современное изложение доказательства этой теоремы есть в «Теории чисел» З. И. Боревича и И. Р. Шафаревича и в «Курсе арифметики» Ж. П. Серра. Оно использует p-адические числа, символ Гильберта и теорему Минковского—Хассе. ■

В силу теоремы Гаусса, для любого натурального n число $8n+3$ представимо в виде суммы квадратов трех целых чисел. Если бы лишь одно из слагаемых было нечетно, то сумма давала бы при делении на 4 остаток 1, а не 3. Поэтому

$$8n+3=$$
$$=(2x+1)^2+(2y+1)^2+(2z+1)^2,$$

что равносильно равенству

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

Значит, всякое натуральное число представимо в виде суммы трех треугольных чисел, то есть чисел вида $(x^2+x)/2$, где x — целое. ■

Методами этой статьи можно доказать следующие утверждения.

1. Число 15 не представимо в виде суммы квадратов двух рациональных чисел. (Этот факт упомянут в «Арифметике» древнегреческого математика Диофанта.)

2. Если $x^2+y^2=1999(z^2+t^2)$, где числа x, y, z и t целые, то $x=y=z=t=0$.

3. Если $a^2+b^2=c^2$, где a, b, c — целые, то abc кратно 60.

4. Никакое число вида n^2+1 не кратно никакому числу вида m^2-1, где m, n — целые числа.

5. Если $x^2y^2=x^2+y^2+z^2$, где x, y и z — целые, то $x=y=z=0$.

6. Если числа x, y, z целые и $4xy-x-y=z^2$, то $x\leqslant 0$ и $y\leqslant 0$.

7. Если $x^3+x^2-2x-1=y^2$, где x и y целые, то $x=-1$ и $y=\pm 1$.

8. Если $y^2=x^3-2$, где x и y целые, то $x=3$ и $y=\pm 5$.

9. Если $y^2=x^3-1$, где x и y целые, то $(x;y)=(1;0)$.

10. Если $y^2=x^3-4$, где x и y целые, то $(x;y)=(2;\pm 2)$ или $(5;\pm 11)$.

11. Уравнение $y^2=x^3+7$ не имеет решений в целых числах. ■

Лемма 7. ***Никакое простое натуральное число p нельзя представить в виде произведения более чем двух целых гауссовых чисел, не являющихся делителями единицы.***

Доказательство. Если $p=(a+bi)(c+di)(e+fi)$, то $|p|=|a+bi|\cdot|c+di|\cdot|e+fi|$, откуда $p^2=(a^2+b^2)(c^2+d^2)(e^2+f^2)$. Квадрат простого числа никак не может быть произведением трех отличных от 1 натуральных чисел.

Следствие. ***Если простое натуральное число p ассоциировано с произведением двух не являющихся делителями единицы целых гауссовых чисел, то эти числа — простые гауссовы.***

Теорема 4. ***Всякое простое натуральное число вида $p=4n+3$ простое и в $\mathbb{Z}[i]$; всякое простое натуральное число вида $p=4n+1$ разлагается на два сопряженных множителя: $p=(a+bi)(a-bi)$, причем множители $a+bi$ и $a-bi$ — простые гауссовы числа; наконец, число 2 ассоциировано с квадратом простого гауссова числа $1+i$.***

Доказательство. Если число $p=4n+3$ представлено в виде произведения двух целых гауссовых чисел $p=(a+bi)(c+di)$, то

$$|p|=|a+bi|\cdot|c+di|,$$

откуда $p^2=(a^2+b^2)(c^2+d^2)$. Значит, либо один из множителей a^2+b^2 и c^2+d^2 равен 1, а другой равен p^2, либо $p=a^2+b^2=c^2+d^2$. В первом случае ясно, что число p было представлено в виде произведения делителя единицы и ассоциированного с p числа. Второй случай невозможен, поскольку p при делении на 4 дает остаток 3, а не 1.
Простое число $p=4n+1$ в силу теоремы Ферма—Эйлера разложимо в сумму квадратов $p=a^2+b^2$, так что $p=(a+bi)(a-bi)$. Множители, в силу следствия леммы 6, являются простыми гауссовыми числами. Число 2 тоже представимо в виде суммы двух квадратов: $2=1^2+1^2=-i(1+i)^2$; число $1+i$ простое в силу этого же следствия. Теорема 4 доказана. ■

Единственность представления простого числа в виде суммы двух квадратов. По теореме Ферма—Эйлера любое простое число p, которое при делении на 4 дает остаток 1, представимо в виде суммы двух квадратов. Оказывается, такое представление единственно с точностью до порядка слагаемых.

Теорема 5. ***Никакое простое число не может быть представлено в виде суммы квадратов двух целых чисел существенно разными (то есть не получающимися один из другого перестановкой слагаемых) способами.***

Доказательство. Если бы простое число p имело два существенно разных представления, $p=a^2+b^2=c^2+d^2$, то разложения $p=(a+bi)(a-bi)=(c+di)\times$ $\times(c-di)$ противоречили бы теореме 4. ■

Сколькими способами число можно представить в виде суммы двух квадратов? В III в. нашей эры Диофант не только знал, что число 65 представимо двумя способами, но и объяснял это тем, что 65 является произведением чисел 13 и 5, каждое из которых — сумма двух квадратов. Комплексных чисел Диофант не знал, иначе он непременно выписал бы разложения $5=(2+i)(2-i)$, $13=(3+2i)(3-2i)$ и продолжил бы свои объяснения следующим образом:

$$65=(2+i)(3+2i)\cdot(2-i)(3-2i)=(4+7i)\cdot(4-7i)=4^2+7^2=$$
$$=(2+i)(3-2i)\cdot(2-i)(3+2i)=(8-i)\cdot(8+i)=8^2+1^2.$$

По-разному группируя множители, получили два разных разложения!

Следующий пример — 25. Это наименьшее число, двумя способами представимое в виде суммы квадратов двух целых неотрицательных чисел:

$$25=(2+i)^2\cdot(2-i)^2=(3+4i)\cdot(3-4i)=3^2+4^2=$$
$$=(2+i)(2-i)\cdot(2+i)(2-i)=5\cdot5=5^2+0^2.$$

Последний пример — число 5746. Как мы хорошо знаем, всякому представлению $5746=a^2+b^2$ соответствует разложение $5746=(a+bi)(a-bi)$ на сопряженные множители. Поэтому разложим рассматриваемое число сначала на простые натуральные, а затем и на простые гауссовы множители:

$$5746=2\cdot13^2\cdot17=(1+i)(1-i)(3+2i)^2(3-2i)^2(4+i)(4-i).$$

Теперь мы должны из нескольких этих множителей составить $a+bi$, да так, чтобы произведение остальных множителей равнялось $a-bi$. Это нетрудно сделать:

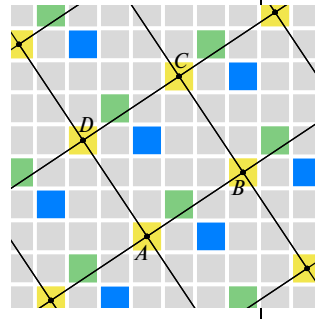
$$a+bi=(1+i)(3+2i)^2(4+i)=-45+61i,$$
$$a-bi=(1-i)(3-2i)^2(4-i)=-45-61i.$$

При этом, разумеется, $45^2+61^2=2025+3721=5746$. Легко найти и еще два варианта:

$$a+bi=(1+i)(3+2i)(3-2i)(4+i)=39+65i,$$
$$a+bi=(1+i)(3-2i)^2(4+i)=75-11i.$$

Они приводят к представлениям $39^2+65^2=1521+4225=5746$ и $75^2+11^2=$ $=5625+121=5746$. Никаких других представлений нет (попытайтесь их придумать — и довольно скоро поймете причину этого).

Аналогично можно найти количество представлений в виде суммы двух квадратов любого натурального числа $n=2^a p_1^{a_1}\ldots p_r^{a_r}Q$, где $p_1,\ldots,p_r$ — попарно различные простые числа, каждое из которых дает остаток 1 при делении на 4, Q — число, не имеющее простых делителей, кроме тех, которые дают остаток 3 при делении на 4. А именно, если Q не является точным квадратом, то n не представимо в виде суммы двух квадратов; если же Q — точный квадрат, то, применив необходимое число раз теорему 2, получаем: количество представлений числа n в виде суммы двух квадратов равно количеству представлений числа $m=$ $=2^a p_1^{a_1}\ldots p_r^{a_r}$ в виде суммы двух квадратов.

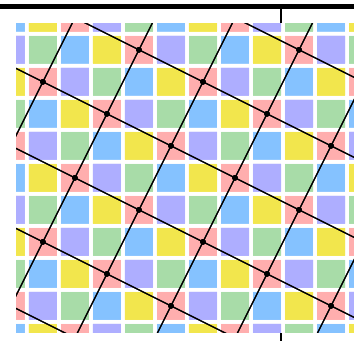


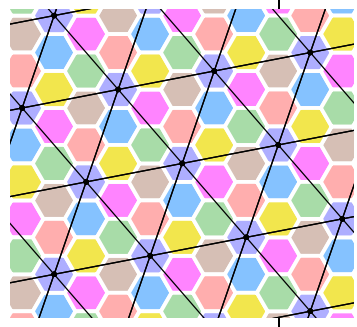
Теорема 6. ***Количество представлений числа m в виде суммы квадратов двух целых неотрицательных чисел равно*** $\left[\dfrac{(a_1+1)\cdot\ldots\cdot(a_r+1)+1}{2}\right]$.

(Если число сомножителей равно 0, то произведение считаем равным 1. Представления, отличающиеся порядком слагаемых, не различаем. Заметьте: Ферма не прибавлял единицу к произведению, а вычитал ее, поскольку он не признавал разложений вида n^2+0^2.)

Следствие. ***Количество точек с целыми координатами на окружности радиуса $\sqrt{n}$ с центром в начале координат (то есть количество решений в целых числах уравнения $x^2+y^2=n$) равно учетверенной разности между количеством натуральных делителей числа n, которые имеют вид $4k+1$, и количеством натуральных делителей вида $4k+3$.***

(Доказательство — индукция по числу простых делителей числа n.) ■





В первом номере журнала «Квант» один из его создателей академик А. Н. Колмогоров опубликовал следующую задачу.

а) На рисунке плоскость покрыта квадратами пяти цветов. Центры квадратов одного и того же цвета расположены в вершинах квадратной сетки. Параллельным переносом решетку квадратов любого данного цвета можно перевести в решетку любого другого цвета. При каком количестве цветов возможно аналогичное заполнение плоскости?

б) Тот же вопрос — о покрытии плоскости шестиугольниками, при котором центры шестиугольников одного и того же цвета образуют решетку из правильных треугольников.

Решение. а) Рассмотрим элементарный квадрат $ABCD$ решетки, образуемой центрами единичных квадратиков одного цвета (на рисунке — желтого). Часть квадрата $ABCD$, закрашенная цветом, участвующим в раскраске, либо представляет собой целый квадратик (например, синий квадратик рисунка), либо состоит из кусочков, из которых можно сложить такой квадратик (например, его можно сложить из двух зеленых или четырех желтых частей). Поэтому площадь квадрата $ABCD$ равна числу цветов раскраски, а по теореме Пифагора она равна сумме квадратов координат вектора $\overrightarrow{AB}$. Значит, количество цветов — сумма квадратов двух целых чисел.

Обратно, для любых целых неотрицательных чисел a и b, где $a>0$, рассмотрим решетку, порожденную векторами $(a;\,b)$ и $(b;\,-a)$. Квадраты, центры которых образуют такую решетку, выкрасим одним цветом; другие решетки получим из этой параллельными переносами.

б) Аналогично, рассмотрев элементарный ромб решетки (состоящий из двух треугольников), вершины которой — центры шестиугольничков одного цвета, получим ответ: a^2+ab+b^2, где a, b — целые числа, хотя бы одно из которых отлично от нуля. ■

Комплексные числа получают, присоединяя к множеству вещественных чисел мнимую единицу i, квадрат которой равен -1. Кватернионы можно получить аналогично, присоединив к множеству $\mathbb{C}$ комплексных чисел мнимую единицу j, обладающую свойствами $j^2=-1$ и $zj=j\overline{z}$ для любого комплексного числа z. Сумму кватернионов z_1+w_1j и z_2+w_2j определяем формулой $(z_1+z_2)+(w_1+w_2)j$, а произведение — формулой $(z_1z_2-w_1\overline{w_2})+(z_1w_2+w_1\overline{z_2})j$.

Нетрудно убедиться, что алгебра кватернионов $\mathbb{H}$ является телом, то есть ассоциативна и не имеет делителей нуля (произведение любых двух ее ненулевых элементов не равно нулю). Обозначив $ij=k$ и представив комплексные числа z и w в виде $z=a+bi$ и $w=c+di$, где a, b, c и $d\in\mathbb{R}$, приходим к формуле $z+wj=a+bi+cj+dk$. Правила умножения запомнить легко: $i^2=j^2=k^2=-1$, $ij=k$, $jk=i$, $ki=j$ и, если идти не по часовой стрелке, а против нее, $ji=-k$, $kj=-i$ и $ik=-j$.

Обозначим векторы $(1;0;0)$, $(0;1;0)$ и $(0;0;1)$ буквами i, j и k соответственно. Тогда кватернион $a+bi+cj+dk$ является суммой числа a и вектора $\vec{v}=bi+cj+dk$. А произведение кватернионов $a_1+\vec{v_1}$ и $a_2+\vec{v_2}$ равно $a_1a_2-\vec{v_1}\vec{v_2}+a_1\vec{v_2}+[\vec{v_1},\vec{v_2}]+a_2\vec{v_1}$, где $\vec{v_1}\vec{v_2}$ — скалярное произведение векторов $\vec{v_1}$ и $\vec{v_2}$, а $[\vec{v_1},\vec{v_2}]$ — их векторное произведение, то есть вектор, перпендикулярный векторам $\vec{v_1}$ и $\vec{v_2}$ и обладающий следующими свойствами: его длина равна площади параллелограмма, натянутого на векторы $\vec{v_1}$ и $\vec{v_2}$, а направлен он так, что тройка векторов $\vec{v_1}$, $\vec{v_2}$ и $[\vec{v_1},\vec{v_2}]$ ориентирована так же, как тройка i, j, k.

Сопряженным кватернионом для кватерниона $s=a+bi+cj+dk$ называют кватернион $\overline{s}=a-bi-cj-dk$. Модуль кватерниона — это число $|s|=\sqrt{s\overline{s}}=\sqrt{a^2+b^2+c^2+d^2}$. Для любых двух кватернионов s_1 и s_2 имеем

$$|s_1s_2|=\sqrt{s_1s_2\overline{s_1s_2}}=\sqrt{s_1s_2\overline{s_2}\,\overline{s_1}}=$$
$$=\sqrt{s_1|s_2|^2\overline{s_1}}=|s_2|\sqrt{s_1\overline{s_1}}=|s_1|\cdot|s_2|.$$

Таким образом, модуль произведения двух кватернионов равен произведению их модулей. Как нетрудно убедиться, это, по сути, и есть формула Эйлера. ■

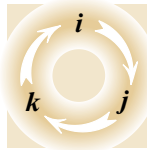
Суммы четырех квадратов

Ж. Л. Лагранж доказал, что любое натуральное число представимо в виде суммы четырех квадратов целых чисел.

Формула Эйлера

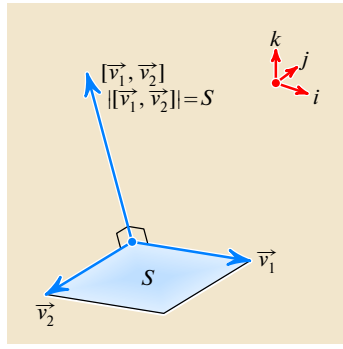
$$(a^2+b^2+c^2+d^2)(x^2+y^2+z^2+t^2)=$$
$$=(ax+by+cz+dt)^2+(ay-bx+ct-dz)^2+(az-bt-cx+dy)^2+(at+bz-cy-dx)^2$$

показывает, что произведение сумм четырех квадратов — тоже сумма четырех квадратов. Поэтому достаточно доказать теорему Лагранжа для простых чисел. Очевидно, $2=1^2+1^2+0^2+0^2$. Пусть p — нечетное простое число.



Лемма. ***Существуют такие целые числа x и y, что x^2+y^2+1 кратно p.***

Доказательство. Рассмотрим числа 0^2, 1^2, 2^2, …, $\left(\frac{p-1}{2}\right)^2$. Если какие-то два из них дают один и тот же остаток при делении на p, то есть если $x^2\equiv y^2 \pmod p$, где $0\leqslant x<y\leqslant(p-1)/2$, то $x^2-y^2=(x-y)(x+y)$ кратно p. Но ни разность $x-y$, ни сумма $x+y$ не кратна p.



Следовательно, рассматриваемые числа дают разные остатки при делении на p. Рассмотрим теперь еще $(p+1)/2$ чисел: $-1-0^2$, $-1-1^2$, $-1-2^2$, …, $-1-\left(\frac{p-1}{2}\right)^2$. Они тоже дают разные остатки. Поскольку всего возможных остатков от деления на p существует p штук, а в каждом из рассматриваемых нами множеств $(p+1)/2$ элементов, то хотя бы одно из чисел вида x^2 дает при делении на p такой же остаток, как и некоторое число вида $-1-y^2$. Значит,

$$x^2\equiv-1-y^2 \pmod p,$$

что и требовалось доказать: $x^2+y^2+1\equiv0 \pmod p$. ■

Числа x и y, как мы помним, не превосходят $(p-1)/2$, поэтому

$$x^2+y^2+1<\left(\frac{p}{2}\right)^2+\left(\frac{p}{2}\right)^2+1<p^2.$$

При этом $x^2+y^2+1^2+0^2=pm$, где $m<p$.

Мы хотим доказать, что число p представимо в виде суммы четырех квадратов целых чисел. Рассмотрим наименьшее натуральное число m, для которого существуют такие целые числа x, y, z, t, что

$$x^2+y^2+z^2+t^2=pm.$$

Как мы уже знаем, $m<p$. Докажем равенство $m=1$ методом бесконечного спуска: предположим, что $m>1$, и докажем, что в таком случае m — не наименьшее. ■

Пусть m четно. Тогда числа x, y, z, t либо все четны, либо все нечетны, либо два из них (для определенности, пусть это x и y) четны, а два (z и t) — нечетны. В любом случае формула

$$\left(\frac{x+y}{2}\right)^2+\left(\frac{x-y}{2}\right)^2+\left(\frac{z+t}{2}\right)^2+\left(\frac{z-t}{2}\right)^2=\frac{x^2+y^2+z^2+t^2}{2}=\frac{pm}{2}$$

показывает, что m — не наименьшее возможное. ■

Пусть m нечетно. Рассмотрим остатки a, b, c, d от деления чисел x, y, z, t на m. Хотя бы один из них отличен от 0: в противном случае сумма квадратов $pm=x^2+y^2+z^2+t^2$ делилась бы на m^2 и (простое!) число p делилось бы на m, хотя $1<m<p$.

Можно считать, что числа a, b, c, d не превосходят $(m-1)/2$. (Если, например, величина a окажется равна $(m+1)/2$ или больше, то можно заменить x на противоположное ему число $-x$. При этом вместо a получим остаток $m-a\leqslant$ $\leqslant m-\dfrac{m+1}{2}=\dfrac{m-1}{2}$.)

Обозначим $n=a^2+b^2+c^2+d^2$. Поскольку

$$a^2+b^2+c^2+d^2\equiv x^2+y^2+z^2+t^2=pm\equiv 0 \pmod m,$$

то $n\equiv 0 \pmod m$, так что $n=mk$, где k — натуральное число. Поскольку все числа a, b, c, d меньше $m/2$, имеем:

$$mk=a^2+b^2+c^2+d^2<4\cdot\left(\frac{m}{2}\right)^2=m^2.$$

Следовательно, $k<m$. Применим формулу Эйлера:

$$(ax+by+cz+dt)^2+(ay-bx+ct-dz)^2+(az-bt-cx+dy)^2+(at+bz-cy-dx)^2=$$
$$=(a^2+b^2+c^2+d^2)(x^2+y^2+z^2+t^2)=npm=m^2pk.$$

Как мы помним, $x\equiv a$, $y\equiv b$, $z\equiv c$ и $t\equiv d \pmod m$. Поэтому по модулю m имеем:

$$\begin{aligned}
&ax+by+cz+dt\equiv x^2+y^2+z^2+t^2=pm\equiv 0,\\
&ay-bx+ct-dz\equiv xy-yx+zt-tz=0,\\
&az-bt-cx+dy\equiv xz-yt-zx+ty=0,\\
&at+bz-cy-dx\equiv xt+yz-zy-tx=0.
\end{aligned}$$

Итак, все числа $ax+by+cz+dt$, $ay-bx+ct-dz$, $az-bt-cx+dy$ и $at+bz-cy-dx$ кратны m; формула

$$pk=\left(\frac{ax+by+cz+dt}{m}\right)^2+\left(\frac{ay-bx+ct-dz}{m}\right)^2+\left(\frac{az-bt-cx+dy}{m}\right)^2+$$
$$+\left(\frac{at+bz-cy-dx}{m}\right)^2$$

представляет число pk в виде суммы четырех квадратов целых чисел. Таким образом, число m не является наименьшим возможным. Теорема Лагранжа доказана. ■

К. Г. Я. Якоби при помощи теории эллиптических функций доказал, что для любого натурального n количество решений уравнения $x^2+y^2+z^2+$ $+t^2=n$ в целых числах равно сумме всех нечетных делителей числа n, умноженной на 24 для четного n и на 8 — для нечетного. ■

Карл Густав Якоб Якоби (1804—1851) — немецкий математик.

Для любых векторов $\vec{u}$, $\vec{v}$ и $\vec{w}$ трехмерного пространства имеет место тождество Якоби

$$[\vec{u},[\vec{v},\vec{w}]]+[\vec{v},[\vec{w},\vec{u}]]+$$
$$+[\vec{w},[\vec{u},\vec{v}]]=\vec{0}.$$

Вместе с антикоммутативностью $[\vec{u},\vec{v}]=-[\vec{v},\vec{u}]$ оно определяет алгебры Ли — в высшей степени интересные объекты изучения современной алгебры.

Дж. Дж. Сильвестер назвал якобианом определитель из частных производных, используемый при замене переменных в многомерных интегралах, чтобы воздать должное трудам Якоби по алгебре. И по сей день не доказана гипотеза якобиана: если для многочленов $f_1, f_2, \ldots, f_n$ от переменных $z_1, z_2, \ldots, z_n$ якобиан — ненулевая постоянная, то отображение

$$(z_1; z_2; \ldots; z_n)\to(f_1; f_2; \ldots; f_n)$$

обратимо, причем обратное отображение тоже задается многочленами.

Подобно тому как функция $y=\sin x$ является обратной к интегралу

$$x=\int_0^y\frac{dt}{\sqrt{1-t^2}},$$

так эллиптическая функция Якоби $y=\mathrm{sn}(x; k)$ обратна к эллиптическому интегралу первого рода

$$x=\int_0^y\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}},$$

где k — постоянная. К функции sn приводит, например, задача о колебаниях тяжелого шарика, подвешенного на нити. ■

Какие числа представимы в виде а) x^2-y^2; б) x^2+2xy; в) x^2+4xy, где x, y — целые числа?

Ответы. а) В виде разности квадратов представимы нечетные числа: $2n+1=(n+1)^2-n^2$. Представимы и числа, делящиеся на четыре: $4n=(n+1)^2-(n-1)^2$. Четные числа, не кратные четырем, непредставимы в виде разности двух квадратов, поскольку произведение двух чисел одной четности либо нечетно, либо делится на 4.

б) Воспользуйтесь равенством $x^2+2xy=x(x+2y)$ или сведите дело к пункту а) при помощи формулы $x^2+2xy=(x+y)^2-y^2$.

в) Представимы числа вида $4n+1=1\cdot(1+4n)$ и числа вида $8n+4=2^2+4\cdot2\cdot n$ и $16n=4^2+4\cdot4\cdot(n-1)$. Непредставимы — вида $4n+2$, $4n+3$ и $16n+8$. ■

Квадрат разрезан на 35 квадратов размером 1×1 и один квадрат большего размера. Какого именно?

Поскольку по крайней мере две стороны основного квадрата граничат только с единичными квадратиками, то длина a стороны основного квадрата — натуральное число. Пусть искомая сторона равна x. Тогда $a^2=35+x^2$, откуда $35=(a-x)(a+x)$. Число 35 разлагается в произведение натуральных множителей лишь двумя способами: $1\cdot35$ и $5\cdot7$. В первом случае $a-x=1$ и $a+x=35$, откуда $x=17$ и $a=18$. Во втором случае $a-x=5$ и $a+x=7$, откуда $x=1$ и $a=6$. По условию задачи, $x>1$. Значит, подходит только $x=17$. ■

Несколько кошек съели 999 919 мышек, причем все кошки съели по одинаковому числу мышек. Сколько было кошек, если каждая кошка съела больше мышек, чем было кошек?

Ответ. $999\,919=1\,000\,000-81=1000^2-9^2=991\cdot1009$. Поскольку числа 991 и 1009 простые, то кошек 991. ■

Решите в целых числах уравнения: а) $x^2+2xy-3y^2=20$; б) $6x^2-xy-12y^2=14$.

Указания.

а) $x^2+2xy-3y^2=(x+y)^2-4y^2=(x-y)(x+3y)$;

б) $6x^2-xy-12y^2=(2x-3y)(3x+4y)$. ■

Уравнения Пелля

П. Л. Чебышёв говорил: «Всякое уравнение, имеющее несколько переменных, подлежит исследованию теории чисел. Но не все они одинаково доступны исследованию и не все имеют одинаковую важность по приложениям своим. Теория чисел до сих пор ограничивается только рассмотрением уравнений, наиболее простых и в то же время имеющих наиболее важные приложения». К числу таких уравнений несомненно относятся уравнения Пелля

$$x^2-dy^2=1,$$

где d — натуральное число, не являющееся точным квадратом.

Левую часть уравнения $x^2-a^2y^2=1$, где a — натуральное число, можно разложить на множители:

$$(x-ay)(x+ay)=1.$$

Число 1 можно представить в виде произведения двух целых чисел двумя способами: $1\cdot1$ и $-1\cdot(-1)$. В первом случае $x-ay=1$ и $x+ay=1$, откуда $x=1$ и $y=0$. Во втором случае $x-ay=-1$ и $x+ay=-1$, откуда $x=-1$ и $y=0$. Ничего особенно интересного в этом нет — мы всего лишь разложили на множители разность квадратов. Действительно поразительные явления происходят, когда d не является точным квадратом.

Уравнениями Пелля можно заниматься по-разному. Что-то может понять даже школьник, только что научившийся раскрывать скобки. С другой стороны, очень важная для математики 10-я проблема Гильберта, поставленная в августе 1900 г. в докладе на Международном математическом конгрессе в Париже, была решена в 1970 г. Ю. В. Матиясевичем при помощи уравнений типа уравнений Пелля. ■

Рассмотрим уравнение $x^2-2y^2=\pm1$ (рис. 1). Не удивляйтесь, что в правой части не 1, а ±1: так легче догадаться до закономерности, о которой вскоре пойдет речь.

Подбором найдем несколько решений: $(x;\,y)=(1;\,0)$, $(1;\,1)$ или $(3;\,2)$. Продолжая вычисления, составим таблицу:

x	1	1	3	7	17	41	99	239
y	0	1	2	5	12	29	70	169
x^2-2y^2	1	−1	1	−1	1	−1	1	−1

Каждый следующий столбец получается из предыдущего по простому правилу: «новое» значение y есть сумма «старых» x и y, а «новое» значение x есть сумма «старого» и «нового» значений y. Точнее, $X=x+2y$ и $Y=x+y$. Конечно, таблицы с несколькими первыми решениями недостаточно для того, чтобы быть уверенным в справедливости этих формул для всего множества решений уравнения; мы должны доказать следующие утверждения.

Теорема 1. ***Если $x^2-2y^2=\pm1$, то пара чисел $(X;\,Y)=(x+2y;\,x+y)$ удовлетворяет равенству $X^2-2Y^2=\mp1$.***

Доказательство. $X^2-2Y^2=(x+2y)^2-2(x+y)^2=x^2+4xy+4y^2-2(x^2+2xy+y^2)=2y^2-x^2=-(x^2-2y^2)$.

Теорема 2. ***Уравнение $x^2-2y^2=\pm1$ не имеет решений в целых неотрицательных числах, кроме тех, что получаются из «тривиального» решения $(1;0)$ при помощи правила $(x;y)\to(x+2y;x+y)$.***

Доказательство теоремы 2 использует метод «бесконечного спуска». Пусть X, Y — натуральные числа, удовлетворяющие равенству $X^2-2Y^2=\pm1$. Рассмотрим систему уравнений

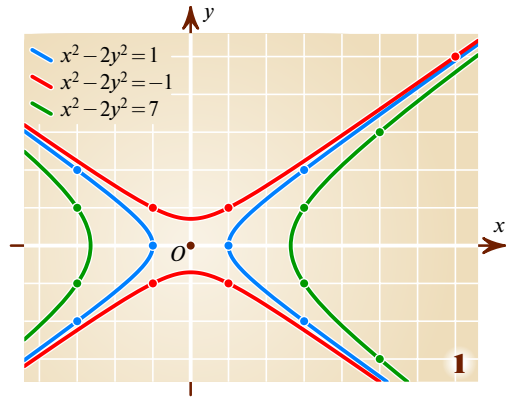
$$\begin{cases} x+2y=X,\\ x+y=Y \end{cases}$$

и решим ее:

$$\begin{cases} x=2Y-X,\\ y=X-Y. \end{cases}$$

Легко проверить, что

$$x^2-2y^2=(2Y-X)^2-2(X-Y)^2=$$
$$=4Y^2-4XY+X^2-2(X^2-2XY+Y^2)=$$
$$=-(X^2-2Y^2).$$



Таким образом каждой паре $(X;Y)$, являющейся решением уравнения $X^2-2Y^2=\pm1$, мы сопоставляем ее «предшественницу» — пару $(x;y)=(2Y-X;X-Y)$, удовлетворяющую равенству $x^2-2y^2=\mp1$.

Лемма 1. ***Если X, Y — натуральные числа и $X^2-2Y^2=\pm1$, то $2Y-X$ и $X-Y$ — неотрицательные числа, причем $X-Y<Y$.***

Доказательство. Рассуждаем «от противного». Если $2Y-X<0$, то $X>2Y$ и $X^2-2Y^2>4Y^2-2Y^2=2Y^2\geqslant2>1$, что противоречит равенству $X^2-2Y^2=\pm1$. Если $X-Y<0$, то $X<Y$ и $X^2-2Y^2<Y^2-2Y^2=-Y^2\leqslant-1$.
Наконец, если $X-Y\geqslant Y$, то $X\geqslant2Y$ и $X^2-2Y^2\geqslant4Y^2-2Y^2=2Y^2\geqslant2>1$, что вновь дает противоречие. Лемма доказана.
Взяв любую пару $(X;Y)$ натуральных чисел, удовлетворяющую равенству $X^2-2Y^2=\pm1$, мы можем рассмотреть ее предшественницу — пару $(x;y)$. При этом $y<Y$. Если x и y — натуральные числа, то у пары $(x;y)$ есть своя предшественница, у которой — своя, и так далее. Бесконечно этот процесс продолжаться не может: неравенство $y<Y$ гарантирует, что начатый с пары $(X;Y)$ процесс образования предшественниц оборвется не более чем через Y шагов. В какой момент обрывается процесс образования пар-предшественниц? Очевидно, когда очередная пара $(x;y)$ состоит не только из натуральных чисел, проще говоря, когда одно из чисел x и y равно нулю. Число x равняться нулю не может, а вот равенство $x^2-2\cdot0^2=\pm1$ возможно. И возможно оно лишь при $x=1$ (напоминаем: $x\geqslant0$).
Итак, для любого решения $(X;Y)$ уравнения $X^2-2Y^2=\pm1$ процесс образования пар-предшественниц остановится, дойдя до пары $(1;0)$. Проследив этот процесс в обратном направлении, то есть не от $(X;Y)$ к $(1;0)$, а от $(1;0)$ к $(X;Y)$, мы видим, что он происходит по формуле $(x;y)\to(x+2y;x+y)$. ■

Рассмотрим последовательности $x_0=1$, $x_1=1$, $x_2=3$, $x_3=7$, $x_4=17$, … и $y_0=0$, $y_1=1$, $y_2=2$, $y_3=5$, $y_4=12$, …, заданные своими начальными членами $x_0=1$, $y_0=0$ и рекуррентными соотношениями $x_{n+1}=x_n+2y_n$ и $y_{n+1}=x_n+y_n$, повнимательнее. По индукции легко доказать равенства $x_n^2-2y_n^2=(-1)^n$, $x_{n+2}=2x_{n+1}+x_n$ и $y_{n+2}=2y_{n+1}+y_n$. (Сделайте это!) ■

Иррациональные числа можно использовать для решения уравнений Пелля. Например,

Сколько решений в целых числах имеет уравнение:
а) $x^2-y^2=2^{100}$; б) $x^2-y^2=p^{100}$, где p — простое число, $p>2$; в) $x^2-9y^2=1\,000\,000$?
Ответы. а) $2\cdot99=198$, поскольку
$$(x;y)=\left(\pm\frac{2^k+2^{100-k}}{2};\pm\frac{2^k-2^{100-k}}{2}\right),$$
где $1\leqslant k\leqslant99$.
б) $2\cdot101=202$.
в) 70. Решения имеют вид
$$\left(\frac{1}{2}(A+B);\frac{1}{6}(A-B)\right),$$
где $AB=10^6$, причем A, B — целые числа. Видим, что A и B должны быть четными и разность $A-B$ должна делиться на 3. Поскольку $AB=10^6\equiv1\pmod3$, последнее выполнено автоматически. Обозначив $A=2a$ и $B=2b$, получаем $ab=2^4\,5^6$. Очевидно, интересующих нас пар $(a;b)$ столько же, сколько чисел вида $\pm2^\alpha5^\beta$, где $0\leqslant\alpha\leqslant4$, $0\leqslant\beta\leqslant6$, то есть $2(4+1)(6+1)=70$ штук. ■

По правилам нового модного танца надо делать либо шаг вперед, либо два шага вперед, либо два шага вперед и сразу же — шаг назад. Сколькими способами танцор может за несколько таких па сдвинуться на 7 шагов от исходного рубежа?
Обозначим через f_n количество способов пройти расстояние в n шагов. Очевидно, $f_0=1$ (никуда не ходить можно единственным способом) и $f_1=2$ (можно сделать либо шаг вперед, либо два шага вперед и шаг назад). Пройти $n+2$ шагов можно трояко: либо пройти сначала $n+1$ шагов и сделать шаг вперед, либо пройти n шагов и сделать два шага, либо пройти $n+1$ шагов, а затем сделать два шага вперед и шаг назад. Значит,
$$f_{n+2}=f_{n+1}+f_n+f_{n+1}=2f_{n+1}+f_n.$$
Ответ: $f_7=408$. ■

Числа $144=12^2$ и $441=21^2$ после зачеркивания двух последних цифр превращаются в $1=1^2$ и $4=2^2$. Найдите наибольший из таких квадратов натуральных чисел, которые не делятся на 10 и остаются квадратами после вычеркивания а) двух; б) четырех; в) $2n$ последних цифр.
Ответы: а) $41^2=1681$; б) $4901^2=24\,019\,801$; в) $\underbrace{49\ldots}_{n-1}\underbrace{90\ldots}_{n-1}01^2$. ■

Существуют ли такие натуральные числа x и y, что x^2+y и y^2+x — квадраты целых чисел? Нет. Пусть для определенности $x \leqslant y$. Тогда

$$y^2 < y^2+x \leqslant y^2+y < (y+1)^2,$$

так что y^2+x заключено между соседними квадратами. ■

Если d — натуральное число, не являющееся квадратом, а z и t — натуральные числа, удовлетворяющие равенству $z^2-dt^2=1$, то натуральные числа a_n и b_n, определенные формулой

$$a_n+b_n\sqrt{d}=(z+t\sqrt{d})^n,$$

обладают тем свойством, что $a_{2n}=2a_n^2-1$ и $b_{2n}=2a_nb_n$. Убедитесь в этом! ■

Существуют ли такие рациональные числа a, b, c, d, что $(a+b\sqrt{2})^2+(c+d\sqrt{2})^2=7+5\sqrt{2}$? Нет. Иначе для сопряженных чисел мы имели бы

$$0 \leqslant (a-b\sqrt{2})^2+(c-d\sqrt{2})^2 = 7-5\sqrt{2} < 0.$$ ■

Если m и n — натуральные числа, то $(5+3\sqrt{2})^m \neq (3+5\sqrt{2})^n$.

Доказательство. I способ. Если

$$(5+3\sqrt{2})^m=(3+5\sqrt{2})^n,$$

то $(5-3\sqrt{2})^m=(3-5\sqrt{2})^n$, что противоречит неравенствам $0<5-3\sqrt{2}<1$ и $5\sqrt{2}-3>1$.

II способ. Если

$$(5+3\sqrt{2})^m=(3+5\sqrt{2})^n,$$

то и $(5-3\sqrt{2})^m=(3-5\sqrt{2})^n$; перемножая эти равенства, получаем $(25-9\cdot 2)^m=(9-25\cdot 2)^n$, то есть $7^m=(-41)^n$. ■

Пусть a, b — целые числа, d — натуральное число, не являющееся квадратом, $x+y\sqrt{d}=\dfrac{1}{a+b\sqrt{d}}$, где x, y — рациональные числа. Убедитесь, что x и y целые тогда и только тогда, когда $a^2-db^2=\pm 1$. ■

Число $[(45+\sqrt{1975})^{30}]$ является нечетным.

Доказательство. Число

$$x=(45+\sqrt{1975})^{30}$$

представимо в виде $a+b\sqrt{1975}$, где a, b — натуральные числа. Рассмотрим сопряженное число:

$$y=(45-\sqrt{1975})^{30}=a-b\sqrt{1975}.$$

Поскольку $x+y=2a$ и $0<y<1$, имеем $[x]=[2a-y]=2a-1$. ■

$$(1+\sqrt{2})^2=1+2\sqrt{2}+2=3+2\sqrt{2};$$
$$(1+\sqrt{2})^3=1+3\sqrt{2}+3\cdot 2+2\sqrt{2}=7+5\sqrt{2};$$
$$(1+\sqrt{2})^4=(1+\sqrt{2})^3(1+\sqrt{2})=(7+5\sqrt{2})(1+\sqrt{2})=17+12\sqrt{2}.$$

Это же решения (3; 2), (7; 5) и (17; 12) уравнения $x^2-2y^2=\pm 1$! Вообще, рассмотрим переход от n-й степени числа $1+\sqrt{2}$ к $(n+1)$-й. Пусть

$$(1+\sqrt{2})^n=x_n+y_n\sqrt{2},$$

где x_n и y_n — натуральные числа. Тогда

$$(1+\sqrt{2})^{n+1}=(1+\sqrt{2})^n(1+\sqrt{2})=(x_n+y_n\sqrt{2})(1+\sqrt{2})=$$
$$=x_n+y_n\sqrt{2}+x_n\sqrt{2}+2y_n=(x_n+2y_n)+(x_n+y_n)\sqrt{2},$$

так что $x_{n+1}=x_n+2y_n$ и $y_{n+1}=x_n+y_n$. Знакомые формулы, не правда ли? ■

Отметим, что число $(1+\sqrt{2})^n$ для любого натурального n представимо в виде $\sqrt{k}+\sqrt{k+1}$, где k — натуральное число:

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

Между прочим,

$$x_{2n}=2x_n^2-(-1)^n \quad \text{и} \quad y_{2n}=2x_ny_n.$$

В самом деле, $x_{2n}+y_{2n}\sqrt{2}=(1+\sqrt{2})^{2n}=((1+\sqrt{2})^n)^2=(x_n+y_n\sqrt{2})^2=x_n^2+2y_n^2+2x_ny_n\sqrt{2}=$ $=(2x_n^2-(-1)^n)+(2x_ny_n)\sqrt{2}$. ■

При возведении числа $1+\sqrt{2}$ в степень мы используем равенство $(\sqrt{2})^2=2$; но число $(-\sqrt{2})^2$ тоже равно 2. Поэтому

$$(1-\sqrt{2})^n=x_n-y_n\sqrt{2}.$$

Такие соображения в алгебре используют часто, есть даже термин: сопряженные числа. В полной общности это важное понятие нам не понадобится. Поэтому скажем только, что для каждого числа вида $a+b\sqrt{2}$, где a, b — рациональные числа, сопряженным числом называют $a-b\sqrt{2}$. Как известно, для любого комплексного числа $a+bi$ сопряженное — это $a-bi$, так что одно слово вроде бы используют в разных целях. На самом деле никакого конфликта обозначений нет: понятие сопряженного зависит не только от самого числа, но и от поля, элементы которого мы рассматриваем. Так что если бы мы подробно поговорили об автоморфизмах полей алгебраических чисел, то все бы встало на свои места. Но это слишком отвлекло бы нас от основной темы.

Тем не менее для нас важно следующее свойство: сопряженное к сумме (разности, произведению, частному) двух чисел равно сумме (разности, произведению, частному) сопряженных к ним. Например, вот как выглядит это для сложения:

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

Чуть бóльших усилий потребует от нас проверка этого свойства для умножения. (Строго говоря, надо бы еще разобраться с разностью и частным, но не будем тратить на это силы: при желании вы легко сделаете это самостоятельно.) Прежде всего вычислим произведение

$$(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2}.$$

Значит, сопряженное к произведению равно $(ac+2bd)-(ad+bc)\sqrt{2}$. Осталось вычислить произведение сопряженных:

$$(a-b\sqrt{2})(c-d\sqrt{2})=(ac+2bd)-(ad+bc)\sqrt{2}.$$

Как видите, результат получился тот же самый.

Отображение $a+b\sqrt{2}\mapsto a-b\sqrt{2}$ называют автоморфизмом поля $\mathbb{Q}[\sqrt{2}]$. А произведение $(a+b\sqrt{2})(a-b\sqrt{2})=a^2-2b^2$ называют нормой числа $a+b\sqrt{2}$. Очень многое из теории уравнений Пелля можно перенести на случай так называемого норменного уравнения в полях алгебраических чисел. ■

Рассмотрим пример: последовательность чисел $x_n=(1+\sqrt{2}+\sqrt{3})^n$. Каждое из них можно привести к виду $x_n=q_n+r_n\sqrt{2}+s_n\sqrt{3}+t_n\sqrt{6}$, где q_n, r_n, s_n, t_n — целые числа. Найдем пределы $\lim\limits_{n\to\infty}\dfrac{r_n}{q_n}$, $\lim\limits_{n\to\infty}\dfrac{s_n}{q_n}$ и $\lim\limits_{n\to\infty}\dfrac{t_n}{q_n}$.

Обозначим: $a=1+\sqrt{2}+\sqrt{3}$, $b=1-\sqrt{2}+\sqrt{3}$, $c=1+\sqrt{2}-\sqrt{3}$ и $d=1-\sqrt{2}-\sqrt{3}$. Наряду с равенством

$$a^n=(1+\sqrt{2}+\sqrt{3})^n=q_n+r_n\sqrt{2}+s_n\sqrt{3}+t_n\sqrt{6}$$

рассмотрим три сопряженных:

$$b^n=q_n-r_n\sqrt{2}+s_n\sqrt{3}-t_n\sqrt{6},$$
$$c^n=q_n+r_n\sqrt{2}-s_n\sqrt{3}-t_n\sqrt{6},$$
$$d^n=q_n-r_n\sqrt{2}-s_n\sqrt{3}+t_n\sqrt{6}.$$

Из этих четырех равенств находим:

$$4q_n=a^n+b^n+c^n+d^n,$$
$$4r_n\sqrt{2}=a^n-b^n+c^n-d^n,$$
$$4s_n\sqrt{3}=a^n+b^n-c^n-d^n,$$
$$4t_n\sqrt{6}=a^n-b^n-c^n+d^n.$$

Следовательно, $\dfrac{r_n}{q_n}=\dfrac{a^n-b^n+c^n-d^n}{(a^n+b^n+c^n+d^n)\sqrt{2}}=\dfrac{1-\left(\frac{b}{a}\right)^n+\left(\frac{c}{a}\right)^n-\left(\frac{d}{a}\right)^n}{\left(1+\left(\frac{b}{a}\right)^n+\left(\frac{c}{a}\right)^n+\left(\frac{d}{a}\right)^n\right)\sqrt{2}}\to\dfrac{1}{\sqrt{2}}$.

(Стремление величин $(b/a)^n$, $(c/a)^n$ и $(d/a)^n$ к нулю следует из того, что все три числа b/a, c/a и d/a по модулю меньше 1.) Аналогично можно вычислить $\lim\limits_{n\to\infty}\dfrac{s_n}{q_n}=\dfrac{1}{\sqrt{3}}$ и $\lim\limits_{n\to\infty}\dfrac{t_n}{q_n}=\dfrac{1}{\sqrt{6}}$. ■

Складывая равенства $(1+\sqrt{2})^n=x_n+y_n\sqrt{2}$, $(1-\sqrt{2})^n=x_n-y_n\sqrt{2}$ и деля на 2, находим

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

А если не сложить, а вычесть, то получим $y_n=\dfrac{(1+\sqrt{2})^n-(1-\sqrt{2})^n}{2\sqrt{2}}$. Это и есть не рекуррентные (когда каждую следующую пару получаем из предыдущей), а явные формулы решений уравнения $x^2-2y^2=\pm1$ в натуральных числах. Заметьте: натуральные x_n и y_n получаются из формул, в которые входит иррациональное число $\sqrt{2}$.

Как известно, $1+2=3$. Легко проверить и равенство

$$1+2+3+\ldots+12+13+14=105=15+16+17+18+19+20.$$

Найдем все такие n, что сумма первых n натуральных чисел равна сумме нескольких последующих — решим в натуральных числах уравнение

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

Решим в целых числах уравнения а) $x(x+1)=4y(y+1)$; б) $x(x+1)(x+7)(x+8)=y^2$; в) $x^2+x=y^4+y^3+y^2+y$; г) $1+x+x^2+x^3+x^4=y^2$; д) $x^2+xy+y^2=x^2y^2$; е) $(x+2)^4-x^4=y^3$.

а) **Ответ:** $(x;y)=(0;0)$, $(0;-1)$, $(-1;0)$ или $(-1;-1)$. Уравнение можно записать в виде $4x^2+4x+1=4(4y^2+4y+1)-3$, откуда $4(2y+1)^2-(2x+1)^2=3$. Число 3 представимо в виде разности квадратов единственным образом: $3=4-1$. Значит, $|2y+1|=1$ и $|2x+1|=1$, то есть $y=0$ или -1 и $x=0$ или -1.

б) Умножив x на $x+8$ и $x+1$ на $x+7$, получаем: $(x^2+8x)\times\times(x^2+8x+7)=y^2$. Решив в целых числах уравнение $z(z+7)=y^2$, получим: $(y;z)=(0;-7)$, $(0;0)$, $(\pm12;-16)$ или $(\pm12;9)$. Осталось решить квадратные уравнения $x^2+8x=-16$, $x^2+8x=-7$, $x^2+8x=0$ и $x^2+8x=9$.

в) **Ответ:** $(x;y)=(0;-1)$, $(-1;-1)$, $(0;0)$, $(-1;0)$, $(5;2)$ или $(-6;2)$. Умножив обе части уравнения на 4 и прибавив к ним по 1, получим $(2x+1)^2=(2y^2+y)^2+3y^2+4y+1=(2y^2+y+1)^2-(y^2-2y)$. Если y — целое и отлично от -1, 0, 1 и 2, то $3y^2+4y+1>0$ и $y^2-2y>0$, так что $(2y^2+y)^2<(2x+1)^2<(2y^2+y+1)^2$. Эти неравенства означают, что $(2x+1)^2$ лежит между двумя последовательными квадратами, а это для целых x невозможно. Подставив в уравнение по очереди $y=-1$, 0, 1 и 2, найдем ответ.

г) Умножим обе части уравнения на 4. Получим: $(2x^2+x)^2=4x^4+4x^3+x^2<(2y)^2\leqslant4x^4+x^2++4+4x^3+8x^2+4x=(2x^2+x+2)^2$. Следовательно, $2y=2x^2+x+1$, $(2x^2+x+1)^2=4(1+x+x^2+x^3+x^4)$, то есть $x^2-2x-3=0$. **Ответ:** $x=3$, $y=11$.

д) Уравнение можно преобразовать к виду $(x+y)^2=xy(xy+1)$.

е) $y^3=8x^3+24x^2+32x+16=8\times\times(x^3+3x^2+4x+2)$. Поэтому $y=2z$, где z — целое и $z^3=x^3+3x^2+4x+2$. Пусть $x\geqslant0$. Тогда $(x+1)^3=x^3++3x^2+3x+1<z^3<x^3+6x^2+12x++8=(x+2)^3$, поэтому $x+1<z<x+2$, что невозможно.

Предположим, что $x\leqslant-2$. Тогда пара чисел $(x_1;y_1)=(-x-2;-y)$ тоже удовлетворяет исходному уравнению, так как $(x_1+2)^4-x_1^4=x^4-(x+2)^4=-y^3=y_1^3$. Но, как было доказано выше, неравенство $x_1\geqslant0$ приводит к противоречию. Таким образом, $-2<x<0$, то есть $x=-1$. При этом, очевидно, $y=0$. ■

Уравнение $x^2+(y^2-1)^2=(y^2)^2$ эквивалентно уравнению $x^2-2y^2=-1$. Поэтому существует бесконечно много различных прямоугольных треугольников, каждый из которых обладает следующими свойствами: длины сторон — целые числа, длина гипотенузы — квадрат целого числа, а один из катетов на единицу короче гипотенузы. ■

Найдем такое наибольшее целое число x, для которого $4^{27}+4^{1000}+4^x$ является квадратом целого числа.

Ответ: $x=1972$. **Решение.** Очевидно, $4^{27}+4^{1000}+4^x=2^{54}\times\times(1+2\cdot2^{1945}+2^{2x-54})$. Если $2x-54=2\cdot1945$, то есть $x=1972$, то выражение в скобках является квадратом суммы.

Если же $x>1972$, то число $1+2\cdot2^{1945}+2^{2(x-27)}$ больше $2^{2(x-27)}$ и меньше $(2^{x-27}+1)^2$, то есть заключено между квадратами двух последовательных натуральных чисел и потому не является точным квадратом. ■

Перед запятой в десятичной записи числа $(\sqrt{2}+\sqrt{3})^{2000}$ стоит цифра 1, а после запятой — не менее 666 девяток.

Доказательство. Для любого целого неотрицательного n обозначим $a_n=(\sqrt{3}+\sqrt{2})^{2n}+(\sqrt{3}-\sqrt{2})^{2n}$, а также $\alpha=5+2\sqrt{6}$ и $\beta=5-2\sqrt{6}$. Тогда

$$\begin{aligned}a_{n+2}&=\alpha^{n+2}+\beta^{n+2}=\\&=(5+2\sqrt{6})^2\alpha^n+(5-2\sqrt{6})^2\beta^n=\\&=(49+20\sqrt{6})\alpha^n+(49-20\sqrt{6})\beta^n=\\&=(50+20\sqrt{6})\alpha^n+(50-20\sqrt{6})\beta^n-\\&\quad-\alpha^n-\beta^n=10(\alpha^{n+1}+\beta^{n+1})-\\&\quad-(\alpha^n+\beta^n)=10a_{n+1}-a_n.\end{aligned}$$

Следовательно, $a_{n+4}=10a_{n+3}-a_{n+2}=10a_{n+3}-10a_{n+1}+a_n$, так что числа a_{n+4} и a_n оканчиваются одной и той же цифрой. Поскольку $a_0=2$, то десятичная запись числа a_{1000} оканчивается цифрой 2. Далее,

$$\begin{aligned}a_{1000}&=(\sqrt{3}+\sqrt{2})^{2000}+\\&+(\sqrt{3}-\sqrt{2})^{2000}>(\sqrt{3}+\sqrt{2})^{2000}=\\&=a_{1000}-(\sqrt{3}-\sqrt{2})^{2000}>\\&>a_{1000}-(1/3)^{2000}>a_{1000}-10^{-666},\end{aligned}$$

откуда и следует, что перед запятой в десятичной записи числа $(\sqrt{3}+\sqrt{2})^{2000}$ стоит цифра 1, а после запятой — не менее 666 девяток. (При помощи компьютера можно проверить, что девяток 995 штук.) ■

то есть $2n(n+1)=k(k+1)$. Умножив обе части на 2, получаем уравнение

$$(2n+1)^2-1=2k^2+2k.$$

Обозначим $x=2n+1$ и еще раз умножим на 2 обе части:

$$2x^2-2=(2k+1)^2-1.$$

Обозначив $2k+1=y$, получаем уравнение $2x^2-1=y^2$, которому, как мы знаем, удовлетворяют числа вида $x=\dfrac{(1+\sqrt{2})^{2m+1}-(1-\sqrt{2})^{2m+1}}{2\sqrt{2}}$. Следовательно, $n=\dfrac{(1+\sqrt{2})^{2m+1}-(1-\sqrt{2})^{2m+1}-2\sqrt{2}}{4\sqrt{2}}$, где m — натуральное число.

Не каждому читателю, по себе знаем, легко привыкнуть пользоваться иррациональными числами для решения уравнений в целых числах. Поэтому мы вернемся к этим методам позже, а пока продолжим рассмотрение примеров. ■

Прямоугольный треугольник со сторонами 3, 4 и 5 обладает тем свойством, что один из его катетов на 1 длиннее другого. Много ли еще таких треугольников, точнее, много ли решений в натуральных числах имеет уравнение $x^2+(x+1)^2=y^2$? Чтобы ответить на этот вопрос, раскроем скобки и приведем подобные: $2x^2+2x+1=y^2$. Домножив обе части на 2, выделим полный квадрат:

$$(2x+1)^2+1=2y^2.$$

Обозначая $z=2x+1$, получаем уравнение $z^2-2y^2=-1$. Любое удовлетворяющее последнему равенству число z нечетно. Поэтому мы свели задачу к уравнению $z^2-2y^2=-1$, где y, z — натуральные числа, причем $z>1$. Отображение $(z;\,y)\to(z+2y;\,z+y)$ меняет 1 в правой части на -1. Чтобы вернуться к 1, отображение надо выполнить дважды:

$$(z;\,y)\to(z+2y;\,z+y)\to((z+2y)+2(z+y);\,(z+2y)+(z+y))=(3z+4y;\,2z+3y).$$

Таким образом, уравнение $x^2-2y^2=-1$ (как и уравнение $x^2-2y^2=1$), имеет бесконечно много решений в натуральных числах. ■

Уравнение $x^2-2y^2=7$. Правило $(x;\,y)\to(3x+4y;\,2x+3y)$ позволяет из одного решения уравнения $x^2-2y^2=7$ получить другое решение. Так из решения $(x;\,y)=(3;\,1)$ получаем $(3\cdot3+4\cdot1;\,2\cdot3+3\cdot1)=(13;\,9)$, из которого получаем $(3\cdot13+4\cdot9;\,2\cdot13+3\cdot9)=(75;\,53)$, из которого можно получить еще одно решение, и так далее. Привычная ситуация, скажете вы? Решения уравнения $x^2-2y^2=1$ получались из «начального» решения $(1;\,0)$ при помощи того же правила $(x;\,y)\to(3x+4y;\,2x+3y)$, так что ничего нового нет? Не торопитесь: $5^2-2\cdot3^2=7$. Решение $(5;\,3)$ не входит в цепочку $(3;\,1)\to(13;\,9)\to\to(75;\,53)\to\ldots$, а порождает свою: $(5;\,3)\to(3\cdot5+4\cdot3;\,2\cdot5+3\cdot3)=(27;\,19)\to\to(3\cdot27+4\cdot19;\,2\cdot27+3\cdot19)=(157;\,111)\to\ldots$

Теорема 3. ***Уравнение $x^2-2y^2=7$ не имеет решений в целых неотрицательных числах, кроме тех, что получаются из одного из двух «начальных» решений* (3; 1) *и* (5; 3) *при помощи правила* $(x;\,y)\to(3x+4y;\,2x+3y)$.**

Доказательство похоже на доказательство теоремы 2. Мы рассматриваем систему

$$\begin{cases}3x+4y=X,\\2x+3y=Y,\end{cases}$$

находим из нее $x=3X-4Y$ и $y=3Y-2X$, замечаем, что

$$x^2-2y^2=(3X-4Y)^2-2(3Y-2X)^2=X^2-2Y^2,$$

затем формулируем и доказываем следующую лемму.

Лемма 2. ***Если X, Y — натуральные числа, удовлетворяющие равенству $X^2 - 2Y^2 = 7$, и выполнено неравенство $Y \geqslant 6$, то $3X - 4Y$ и $3Y - 2X$ — тоже натуральные числа, причем $3X - 4Y < X$.***

Доказательство. Рассуждаем «от противного». Если $3X - 4Y \leqslant 0$, то $X \leqslant \frac{4}{3}Y$ и $7 = X^2 - 2Y^2 \leqslant \frac{16}{9}Y^2 - 2Y^2 < 0$. Если $3Y - 2X \leqslant 0$, то $X \geqslant \frac{3}{2}Y$ и $X^2 - 2Y^2 \geqslant \frac{9}{4}Y^2 - 2Y^2 = \frac{Y^2}{4} > 7$. Наконец, если $3X - 4Y \geqslant X$, то $X \geqslant 2Y$ и $X^2 - 2Y^2 \geqslant 4Y^2 - 2Y^2 = 2Y^2 > 7$, что вновь дает противоречие.
Лемма доказана. Дальнейшее доказательство проводится почти так же, как и доказательство теоремы 2. Чтобы понять, где может остановиться процесс образования пар-предшественниц, достаточно разобрать случаи $Y = 1$, 2, 3, 4, 5. Находим, как и следовало ожидать, два решения: (3; 1) и (5; 3). ■

Уравнение $x^2 - 3y^2 = -1$. Рассмотрим остаток от деления на 3 левой части уравнения $x^2 - 3y^2 = -1$. Поскольку $3y^2$ делится на 3, искомый остаток совпадает с остатком от деления x^2 на 3. Число x можно представить одной из трех формул: $x = 3k$ (если x делится на 3), $x = 3k + 1$ (если x при делении на 3 дает остаток 1) или, наконец, $x = 3k + 2$ (если остаток равен 2). При этом $x^2 = 9k^2$, $9k^2 + 6k + 1$ или $9k^2 + 12k + 4$. Остаток от деления на 3 в первом случае равен 0, а в двух других случаях остаток равен 1.
Итак, левая часть уравнения $x^2 - 3y^2 = -1$ при делении на 3 дает остаток 0 или 1, а правая — остаток 2. Уравнение $x^2 - 3y^2 = -1$ решений в целых числах не имеет. ■

Уравнение $x^2 - 3y^2 = 1$. Равенство $2^2 - 3 = 1$ запишем в виде $(2 + \sqrt{3}) \times (2 - \sqrt{3}) = 1$, а затем возведем обе части в n-ю степень:

$$(2 + \sqrt{3})^n (2 - \sqrt{3})^n = 1.$$

Обозначив через x_n и y_n такие натуральные числа, что $(2 + \sqrt{3})^n = x_n + y_n\sqrt{3}$, получим, заменив знаки перед $\sqrt{3}$, равенство

$$(2 - \sqrt{3})^n = x_n - y_n\sqrt{3}.$$

Следовательно, $1 = (2 + \sqrt{3})^n (2 - \sqrt{3})^n = (x_n + y_n\sqrt{3})(x_n - y_n\sqrt{3}) = x_n^2 - 3y_n^2$. Значит, пара $(x_n; y_n) = \left(\frac{(2 + \sqrt{3})^n + (2 - \sqrt{3})^n}{2}; \frac{(2 + \sqrt{3})^n - (2 - \sqrt{3})^n}{2\sqrt{3}} \right)$ — решение уравнения $x^2 - 3y^2 = 1$. Других решений в натуральных числах нет.

Теорема 4. ***Если $x^2 - 3y^2 = 1$, то пара чисел $(X; Y) = (2x + 3y; x + 2y)$ удовлетворяет равенству $X^2 - 3Y^2 = 1$.***

Доказательство. $(2x + 3y)^2 - 3(x + 2y)^2 = 4x^2 + 12xy + 9y^2 - 3(x^2 + 4xy + 4y^2) = x^2 - 3y^2 = 1$.

Теорема 5. ***Уравнение $x^2 - 3y^2 = 1$ не имеет решений в целых неотрицательных числах, кроме тех, что получаются из «тривиального» решения (1; 0) при помощи правила $(x; y) \to (2x + 3y; x + 2y)$.***

Доказательство теоремы 5 можно провести так же, как и теорем 2 и 3. Поэтому мы не будем это делать, все равно скоро докажем общую теорему! ■

Уравнение $x^2 - xy - y^2 = \pm 1$ не имеет вида $x^2 - dy^2 = 1$. Но умножение на 4 приводит его к виду $4x^2 - 4xy - 4y^2 = \pm 4$, то есть $(2x - y)^2 - 5y^2 = \pm 4$,

---

Сумма квадратов а) трех; б) четырех; в) пяти; г) шести; д) семи; е) восьми; ж) девяти; з) десяти; и) двенадцати последовательных целых чисел не может быть квадратом целого числа.
Доказательство. а) $(x-1)^2 + x^2 + (x+1)^2 = 3x^2 + 2$; но квадрат не может дать остаток 2 при делении на 3.
б) $(x-1)^2 + x^2 + (x+1)^2 + (x+2)^2 = 4x^2 + 4x + 6 \equiv 2 \pmod 4$; но квадрат целого числа не может дать остаток 2 при делении на 4.
в) $5x^2 + 10 = 5(x^2 + 2)$ не может быть квадратом, поскольку $x^2 + 2$ ни при каком целом x не кратно 5.
г) $6x^2 + 6x + 19 \equiv 3 \pmod 4$.
д) $x^2 + 4 = 7z^2$. Значит, $x^2 + 4$ делится на 7, что невозможно.
е) $2x^2 + 2x + 11 = z^2$. Значит, $z^2 \equiv 3 \pmod 4$.
ж) $(x-4)^2 + (x-3)^2 + (x-2)^2 + \ldots + (x+3)^2 + (x+4)^2 = 9x^2 + 60$ делится на 3, но не делится на 9, и поэтому не может быть точным квадратом.
з) $2x(x+1) \not\equiv 3 \pmod 5$, поскольку $(2x+1)^2 \not\equiv 7 \pmod 5$.
и) $y^2 \not\equiv 2 \pmod 4$. ■

Найдем наименьшее натуральное число, квадрат которого представим в виде суммы квадратов 11 последовательных а) целых; б) натуральных чисел.
а) $11^2 = (-4)^2 + (-3)^2 + \ldots + 5^2 + 6^2$. Уравнение

$$(x-5)^2 + (x-4)^2 + \ldots + (x+4)^2 + (x+5)^2 = y^2$$

после раскрытия скобок и приведения подобных принимает вид $11x^2 + 110 = y^2$. Замена $y = 11z$ и сокращение на 11 дают $x^2 + 10 = 11z^2$. Наименьшее по величине натуральное z, удовлетворяющее этому уравнению, равно 1. При этом $y = 11$.
б) Поскольку $x^2 - 1 = 11z^2 - 11$, то $(x-1)(x+1) = x^2 - 1$ делится на 11. Значит, $x = 11t \pm 1$. Значения $x = 1$, 10, 12, 21 не подходят, а при $x = 23$ имеем $z = 7$, то есть $y = 77$. ■

Существует бесконечно много натуральных чисел, квадрат каждого из которых представим в виде суммы квадратов 11 последовательных натуральных чисел.
Доказательство. Поскольку $1^2 - 11 \cdot 1^2 = -10$ и $10^2 - 11 \cdot 3^2 = 1$, то уравнение $x^2 - 11z^2 = -10$ имеет бесконечно много решений в натуральных числах. ■

Если натуральные числа m и n удовлетворяют равенству $2m^2+m=3n^2+n$, то числа а) $m-n$; б) $2m+2n+1$; в) $3m+3n+1$ являются квадратами целых чисел.

Доказательство. а) Обозначим $x=m-n$. Заменив m на $x+n$, раскрыв скобки и упростив, получим равенство

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

Очевидно, числа x и $4n+2x+1$ взаимно просты. Поэтому x — квадрат натурального числа.

б) Заменив $m=(M-1)/2$ и $n=(y-M)/2$, получаем после преобразований равенство

$$y(6M-3y-2)=M^2.$$

Поскольку число y взаимно просто с (нечетным!) числом M, то y — квадрат натурального числа. Осталось вспомнить, что $y=2m+2n+1$. ■

Если число n натуральное число и число

$$m=2+2\sqrt{28n^2+1}$$

целое, то число $\sqrt{m}$ тоже целое. Докажите это самостоятельно. ■

Если a, b — такие натуральные числа, что

$$(\sqrt{3}+\sqrt{2})^{2001}=a\sqrt{3}+b\sqrt{2},$$

то $3a^2-2b^2=1$.

Если a и b — такие натуральные числа, что $3a^2-2b^2=1$, то для некоторого нечетного натурального числа n имеем:

$$a\sqrt{3}+b\sqrt{2}=(\sqrt{3}+\sqrt{2})^n.$$

Доказательство.

$$3a^2-2b^2=$$
$$=(a\sqrt{3}-b\sqrt{2})(a\sqrt{3}+b\sqrt{2})=$$
$$=(\sqrt{3}-\sqrt{2})^{2001}(\sqrt{3}+\sqrt{2})^{2001}=$$
$$=((\sqrt{3}-\sqrt{2})(\sqrt{3}+\sqrt{2}))^{2001}=$$
$$=(3-2)^{2001}=1.$$

Воспользуемся формулой

$$(a\sqrt{3}+b\sqrt{2})(\sqrt{3}\pm\sqrt{2})^2=$$
$$=(5a\pm4b)\sqrt{3}+(5b\pm6a)\sqrt{2}.$$

Пусть a, b — натуральные числа и $3a^2-2b^2=1$. Тогда $3(5a-4b)^2-2(5b-6a)^2=1$. Если $5a-4b\leqslant0$, то

$$3a^2-2b^2\leqslant$$
$$\leqslant3\left(\frac{4}{5}b\right)^2-2b^2=-\frac{2}{25}b^2<0.$$

Значит, $5a-4b>0$. Если $5b-6a\leqslant0$, то

$$3a^2-2b^2\geqslant3\left(\frac{5}{6}b\right)^2-2b^2=\frac{1}{12}b^2.$$

Осталось проверить значения $b=1,2,3$. Подходит только $b=1$, которому соответствует $a=1$. ■

что уже похоже на уравнение Пелля. Впрочем, мы воспользуемся этим преобразованием чуть позже, а здесь решим уравнение в его первоначальном виде. Немного посчитав, можно составить таблицу:

x	0	1	1	2	3	5	8	13	21
y	1	0	1	1	2	3	5	8	13
x^2-xy-y^2	-1	1	-1	1	-1	1	-1	1	-1

Теорема 6. ***Если*** $x^2-xy-y^2=\pm1$, ***то пара чисел*** $(X;\,Y)=(x+y;\,x)$ ***удовлетворяет равенству*** $X^2-XY-Y^2=\mp1$.

Доказательство. $(x+y)^2-(x+y)x-x^2=x^2+2xy+y^2-x^2-xy-x^2=-(x^2-xy-y^2)=\mp1$.

Теорема 7. ***Уравнение*** $x^2-xy-y^2=\pm1$ ***не имеет решений в целых неотрицательных числах, кроме тех, что получаются из «тривиального» решения*** $(0;\,1)$ ***при помощи правила*** $(x;\,y)\to(x+y;\,x)$.

Следствие. ***Все решения уравнения*** $z^2-5y^2=\pm4$ ***в натуральных числах даются формулой*** $(z;\,y)=(\varphi_{n+1}+\varphi_{n-1};\,\varphi_n)$.

Доказательство. Каждой паре целых чисел $(x;\,y)$, удовлетворяющей равенству $x^2-xy-y^2=\pm1$, соответствует пара целых чисел $(z;\,y)=(2x-y;\,y)$, удовлетворяющая равенству $z^2-5y^2=\pm4$, и наоборот (поскольку числа z и y одной четности). Осталось заметить, что если $x=\varphi_{n+1}$ и $y=\varphi_n$, то

$$z=2x-y=2\varphi_{n+1}-\varphi_n=\varphi_{n+1}+\varphi_{n-1}.$$ ■

Теорема 8. ***Целые неотрицательные числа*** x, y ***удовлетворяют уравнению*** $x^2-mxy+y^2=1$ ***(где*** m ***— натуральное число,*** $m>1$***) тогда и только тогда, когда*** x ***и*** y ***— соседние члены последовательности*** $a_0=0$, $a_1=1$, $a_2=m$, $a_3=m^2-1$, $a_4=m^3-2m$, $a_5=m^4-3m^2+1$, ..., ***где*** $a_{k+2}=ma_{k+1}-a_k$ ***при*** $k\geqslant0$.

Доказательство. При помощи индукции проверим, что соседние члены последовательности удовлетворяют уравнению. База. $0^2-m\cdot0\cdot1+1^2=1$. Переход. В последовательности a_0, a_1, a_2, ... за каждой парой $(a_k,\,a_{k+1})=(x,\,y)$, следует пара $(a_{k+1};\,a_{k+2})=(a_{k+1};\,ma_{k+1}-a_k)=(y;\,my-x)$. Очевидно,

$$y^2-my(my-x)+(my-x)^2=y^2-(my-x)(my-(my-x))=y^2-(my-x)x=x^2-mxy+y^2.$$

Докажем, что других решений в целых неотрицательных числах нет. Предположим, что $X^2-mXY+Y^2=1$ и $0\leqslant X\leqslant Y$. Если при этом $X=0$, то, очевидно, $Y=1$. Если же $X>0$, рассмотрим систему уравнений $y=X$ и $my-x=Y$. Очевидно, $x=mX-Y$ и $y=X$. Если $mX-Y<0$, то $mXY<Y^2$ и $1=X^2-mXY+Y^2>X^2\geqslant1$, что невозможно. Значит, $x=mX-Y\geqslant0$. Если $x=0$, то $y=1$. Если же $x>0$, то пара натуральных чисел $(x;\,y)$ удовлетворяет равенству $x^2-mxy+y^2=1$ и условиям $x<y=X\leqslant Y$ (проверьте!). Переходя таким образом от пары $(X;\,Y)$ к предшественнице $(x;\,y)$, затем от $(x;\,y)$ — к ее предшественнице и так далее, мы рано или поздно остановимся — получим решение $(x;\,y)=(0;\,1)$. Идя по цепочке в обратном направлении, то есть начав с $(0;\,1)$ и многократно выполняя преобразование $(x;\,y)\to(X;\,Y)$, мы придем к решению $(X;\,Y)$. ■

При $m=3$ имеем $a_0=0$, $a_1=1$, $a_2=3$, $a_3=3\cdot3-1=8$, $a_4=3\cdot8-3=21$. И вообще, a_n — это $2n$-й член последовательности Фибоначчи. Проще всего это доказать, проверив тождество $\varphi_{n+4}=3\varphi_{n+2}-\varphi_n$.

А можно заменой $y=x-z$ привести уравнение $x^2-3xy+y^2=1$ к виду $z^2+zx-x^2=1$. Решения $(x;\,z)=(\varphi_{2n};\,\varphi_{2n-1})$ соответствуют решениям исходного уравнения в неотрицательных целых числах x и y, удовлетворяющим неравенству $x\geqslant y$. Значит, $(x;\,y)=(x;\,x-z)=(\varphi_{2n};\,\varphi_{2n}-\varphi_{2n-1})=(\varphi_{2n};\,\varphi_{2n-2})$. ■

Гипербола и решетки. Уравнения Пелля встречаются и в геометрии.

Теорема 9. ***а) Существует квадрат, все вершины и все середины сторон которого лежат на гиперболах*** $xy=\pm1$.

б) Существует бесконечно много параллелограммов, одна из вершин каждого из которых — начало координат, две другие лежат на гиперболе $xy=1$, ***а четвертая — на гиперболе*** $xy=-1$.

в) Площадь любого такого параллелограмма равна $\sqrt{5}$.

г) Рассмотрим для некоторого такого параллелограмма $OABC$ ***порожденную им решетку, то есть множество таких точек*** P, ***что*** $\overrightarrow{OP}=m\overrightarrow{OA}+n\overrightarrow{OC}$, ***где*** m, n ***— целые числа. Внутренность «креста», ограниченного гиперболами*** $xy=\pm1$, ***содержит лишь одну точку этой решетки — начало координат; а на самих гиперболах*** $xy=\pm1$ ***лежит бесконечно много точек решетки.***

Доказательство. а) Проанализируем ситуацию. Пусть искомый квадрат существует и выглядит так, как показано на рисунке 2. Обозначим координаты точки A — середины стороны квадрата — через $\left(a;\frac{1}{a}\right)$. Тогда, как легко видеть, $\overrightarrow{AB}=\left(\frac{1}{a};-a\right)$, так что точка B имеет координаты $\left(a+\frac{1}{a};\frac{1}{a}-a\right)$. Условие принадлежности точки B гиперболе $xy=1$ дает уравнение

$$\left(a+\frac{1}{a}\right)\cdot\left(\frac{1}{a}-a\right)=1,$$

откуда $\frac{1}{a^2}-a^2=1$. Этому уравнению удовлетворяет число $a=\sqrt{\frac{\sqrt5-1}{2}}$. Анализ окончен. При найденном a все четыре точки $B\left(a+\frac{1}{a};\frac{1}{a}-a\right)$, $D\left(-a+\frac{1}{a};-\frac{1}{a}-a\right)$, $F\left(-a-\frac{1}{a};-\frac{1}{a}+a\right)$, $H\left(a-\frac{1}{a};\frac{1}{a}+a\right)$ (вершины квадрата) и точки $A\left(a;\frac{1}{a}\right)$, $C\left(\frac{1}{a};-a\right)$, $E\left(-a;-\frac{1}{a}\right)$, $G\left(-\frac{1}{a};a\right)$ (середины сторон) лежат на гиперболах $xy=\pm1$.

б) Рассмотрим точки $A(a;1/a)$ и $C(c;-1/c)$, а также начало координат $O(0;0)$ (рис. 3). Вершина B параллелограмма $OABC$ имеет координаты $\left(a+c;\frac{1}{a}-\frac{1}{c}\right)$. Она лежит на гиперболе $xy=1$ при условии

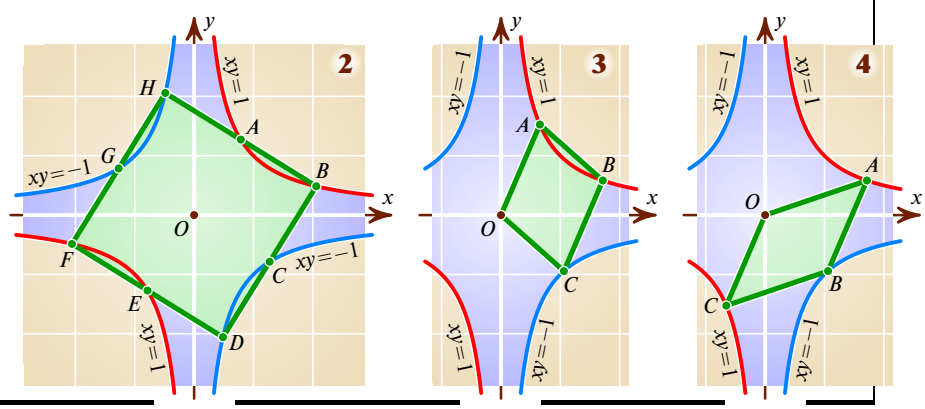
$$(a+c)\cdot\left(\frac{1}{a}-\frac{1}{c}\right)=1,$$

которое можно записать в виде $\frac{c}{a}-\frac{a}{c}=1$, то есть $\frac{c}{a}=\frac{1\pm\sqrt5}{2}$. Этому равенству удовлетворяют бесконечно много пар чисел a и c.

в) Легко доказать, что площадь S параллелограмма $OABC$, где O — начало координат, $\overrightarrow{OA}=(a;b)$ и $\overrightarrow{OC}=(c;d)$, равна $S=|ad-bc|$. Подставляя $b=\frac{1}{a}$ и $d=-\frac{1}{c}$, находим

$$S=\left|\frac{a}{c}+\frac{c}{a}\right|=\left|\frac{2}{1\pm\sqrt5}+\frac{1\pm\sqrt5}{2}\right|=\sqrt5.$$

Но решение еще не закончено: параллелограмм может выглядеть так, как показано на рисунке 4. Его вершины $A(a;1/a)$ и $C(c;1/c)$ лежат



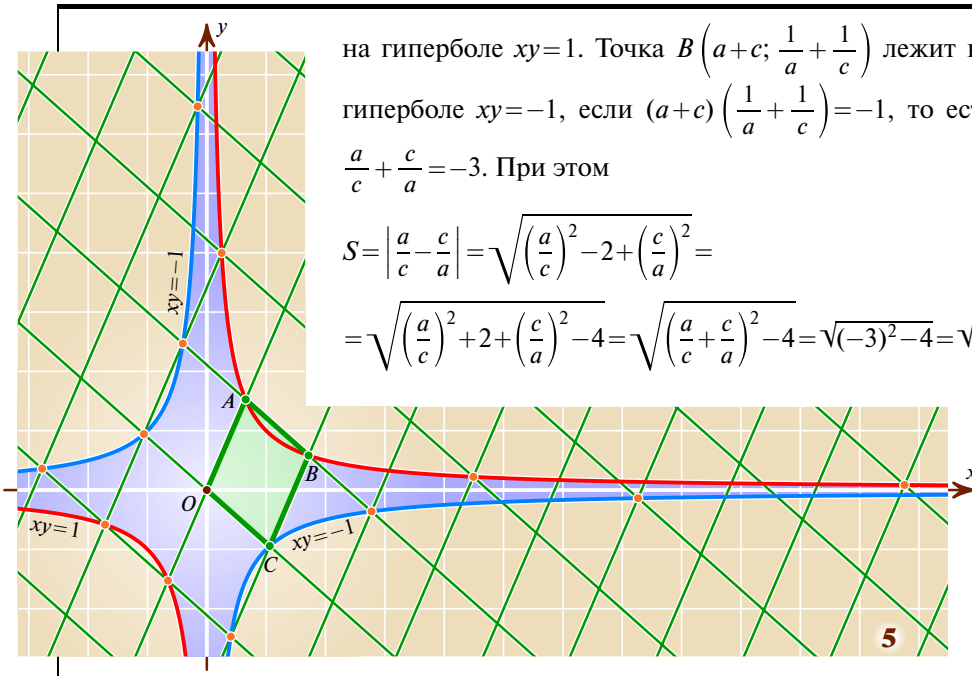
Если все вершины и середины сторон квадрата лежат на гиперболах $xy=\pm1$, то его центр — начало координат.

Доказательство. Рассмотрим 8 точек: вершины и середины сторон некоторого квадрата. Пусть все они лежат на гиперболах $xy=\pm1$. Пусть на некоторой ветви лежат рассматриваемые точки K и L, не лежащие на одной стороне квадрата. С любой стороны от отрезка KL среди рассматриваемых вершин и середин сторон квадрата есть такая точка M, что углы MKL и MLK острые. Получили противоречие: точка M должна лежать в полуполосе, ограниченной отрезком KL и восставленными в точках K и L перпендикулярами, направленной внутрь рассматриваемой ветви гиперболы.

Точки K и L не могут быть и соседними вершинами квадрата (иначе середина отрезка KL не лежит на гиперболах).

Поскольку никакая сторона квадрата не пересекает никакую ветвь гиперболы более чем в двух точках, то каждой ветви принадлежат вершина квадрата и середина одной из выходящих из нее сторон.

Рассмотрим две такие точки A и B. При симметрии относительно начала координат точки A и B переходят в точки A' и B', лежащие на другой ветви той же гиперболы. При этом отрезок $A'B'$ равен и параллелен отрезку BA. Если бы противоположная вершине B вершина квадрата и противоположная середине A середина стороны квадрата не совпадали с точками B' и A' соответственно, то на ветви гиперболы нашлись бы два разных отрезка, равных по длине и параллельных. Получили желанное противоречие: таких отрезков не бывает! ■



на гиперболе $xy=1$. Точка $B\left(a+c;\ \frac{1}{a}+\frac{1}{c}\right)$ лежит на гиперболе $xy=-1$, если $(a+c)\left(\frac{1}{a}+\frac{1}{c}\right)=-1$, то есть $\frac{a}{c}+\frac{c}{a}=-3$. При этом

$$S=\left|\frac{a}{c}-\frac{c}{a}\right|=\sqrt{\left(\frac{a}{c}\right)^2-2+\left(\frac{c}{a}\right)^2}=$$

$$=\sqrt{\left(\frac{a}{c}\right)^2+2+\left(\frac{c}{a}\right)^2-4}=\sqrt{\left(\frac{a}{c}+\frac{c}{a}\right)^2-4}=\sqrt{(-3)^2-4}=\sqrt{5}.$$

г) Рассмотрим порожденную параллелограммом рисунка 3 решетку (рис. 5). Для произвольной точки $P(x;\,y)$ этой решетки $\overrightarrow{OP}=m\overrightarrow{OA}+n\overrightarrow{OC}=\left(ma+nc;\ \frac{m}{a}-\frac{n}{c}\right)$, где m, n — целые числа, имеем

$$|xy|=\left|(ma+nc)\left(\frac{m}{a}-\frac{n}{c}\right)\right|=\left|m^2+mn\left(\frac{c}{a}-\frac{a}{c}\right)-n^2\right|=|m^2+mn-n^2|.$$

Внутренность «креста» из гипербол $xy=\pm1$ задается неравенством $|xy|<1$. Но при целых m и n величина $|m^2+mn-n^2|$ тоже целая. Единственным целым числом, которое по модулю меньше 1, является нуль. Значит, для лежащей внутри креста точки решетки имеем

$$\left|(ma+nc)\left(\frac{m}{a}-\frac{n}{c}\right)\right|=0,$$

откуда $ma+nc=0$ или $mc-na=0$. Ввиду иррациональности отношения a/c это возможно лишь при $m=n=0$.

Значит, внутри «креста» из гипербол расположена единственная точка рассматриваемой решетки — начало координат.

Для решетки, порожденной параллелограммом рисунка 4, решение аналогично, поэтому мы выпишем только формулы

$$\overrightarrow{OP}=m\overrightarrow{OA}+n\overrightarrow{OC}=\left(ma+nc;\ \frac{m}{a}+\frac{n}{c}\right),$$

$$|xy|=\left|(ma+nc)\left(\frac{m}{a}+\frac{n}{c}\right)\right|=\left|m^2+mn\left(\frac{c}{a}+\frac{a}{c}\right)+n^2\right|=$$

$$=|m^2-3mn+n^2|=|(m-n)^2-(m-n)n-n^2|=|k^2-kn-n^2|,$$

где обозначено $k=m-n$.

Итак, внутри «креста гипербол» нет ни одной точки решеток, кроме начала координат. А на самих гиперболах таких точек бесконечно много, ибо уравнения $m^2+mn-n^2=\pm1$ и $k^2-kn-n^2=\pm1$ имеют бесконечно много решений в целых числах. (Первое из них сводится ко второму заменой m на $-k$.) ■

Рассмотрим 14-ю строку треугольника Паскаля: 1, 14, 91, 364, **1001, 2002, 3003,** 3432, 3003, 2002, 1001, 364, 91, 14, 1. (Строки треугольника

Если для натуральных чисел a и b число $\frac{a^2+b^2}{ab-1}$ натуральное, то оно равно 5; а уравнение $x^2-5xy+y^2+5=0$ имеет бесконечно много решений в натуральных числах.

Доказательство. Случай $a=b$ невозможен: число $\frac{2a^2}{a^2-1}=2+\frac{2}{a^2-1}$ нецелое ни при каком натуральном a.

Пусть $t\in\mathbb{N}$ и уравнение

$$x^2-txy+y^2+t=0 \qquad (*)$$

имеет решения в натуральных числах x, y. Рассмотрим наименьшее натуральное $x=a$, для которого существует натуральное $y=b<a$, удовлетворяющее равенству $(*)$. При фиксированных t и b уравнение $x^2-txb+b^2+t=0$ — квадратное относительно x. Если оно имеет натуральный корень a, то по теореме Виета оно имеет и целый корень $A=tb-a$. Если $A\leqslant0$, то $a^2-tab+b^2+t=a(a-tb)+b^2+t>0$, что неверно. Значит, $A\geqslant a$. Если $A=a$, то дискриминант равен нулю:

$$(tb)^2-4(b^2+t)=0,$$

откуда $4t=(t^2-4)b^2\geqslant t^2-4$, так что $t\leqslant4$; но при $t=1,\ 2,\ 3,\ 4$ равенство $4t=(t^2-4)b^2$ не имеет места.

Итак, $A>a$. По теореме Виета, $aA=b^2+t$ и $a+A=tb$. Поэтому

$$b^2+t-tb=aA-a-A=$$
$$=(a-1)(A-1)-1\geqslant b(b+1)-1=$$
$$=b^2+b-1,$$

откуда $t(1-b)\geqslant b-1$. Это возможно лишь при $b=1$, причем все неравенства должны обращаться в равенства, то есть $a=2$, $A=3$, $t=5$.

Два решения уравнения

$$x^2-5xy+y^2+5=0$$

найти легко: $(x;\,y)=(1;\,2)$ и $(1;\,3)$. Из каждого решения $(x;\,y)$, где $x<y$, можно получить новое решение $(y;\,5y-x)$. Действительно,

$$(5y-x)^2-5(5y-x)y+y^2=$$
$$=x^2-5xy+y^2.$$

При этом $5y-x>4y>y$. Таким образом, любые два соседних члена любой из последовательностей

1, 2, 9, 43, 206, 987, . . . ,
1, 3, 14, 67, 321, 1538, . . . ,

где каждый член получается из двух предыдущих чисел x и y по формуле $5y-x$, дают решение интересующего нас уравнения.

На самом деле мы нашли все решения в натуральных числах! Докажем это. Пусть $0<X<Y$ и $X^2-5XY+Y^2+5=0$. Рассмотрим преобразование $(X;\,Y)\to(x;\,y)$, где $x=5X-Y$ и $y=X$. Если $x<X$, то $\min(x,\,y)<\min(X,\,Y)$, так что удалось получить «меньшее» решение в натуральных числах. Если же $5X-Y\geqslant X$, то

$$5=(5X-Y)Y-X^2\geqslant XY-X^2=$$
$$=X(Y-X)\geqslant X.$$

Перебрав значения $X=1,\,2,\,3,\,4,\,5$, находим: $(X;\,Y)=(1;\,2)$ или $(1;\,3)$. ■

Докажем следующие утверждения.

а) Существует бесконечно много таких пар натуральных чисел a и b, что a^2+1 делится на b, а b^2+1 делится на a.

б) Если $x<y$ — натуральные числа и $x^2+y^2+1=3xy$, то $x=\varphi_{2n-1}$ и $y=\varphi_{2n+1}$, где n — некоторое натуральное число.

в) Если a, b и $c=\dfrac{a^2+b^2+1}{ab}$ — натуральные числа, то $c=3$.

г) Если два натуральных числа таковы, что увеличенный на единицу квадрат любого из них делится на другое, то произведение этих чисел на единицу больше квадрата их разности.

д) Уравнение $x^2-(n^2-4)y^2=-4$ не имеет решений в целых числах при натуральном $n\neq 3$.

е) Уравнение $x^2-(n^2-4)y^2=-1$ не имеет решений в целых числах при натуральном $n\neq 3$.

Доказательство. а) Годятся $a=\varphi_{2n-1}$ и $b=\varphi_{2n+1}$, где n — натуральное число.

в) Если $a=b$, то $c=2+\dfrac{1}{a^2}$, так что $a=1$ и $c=3$. Пусть $c\neq 3$ и $a<b$, причем b — наименьшее возможное. Положим $A=ca-b$ и $B=a$. Очевидно,

$$A=ca-b=\frac{a^2+1}{b}<a+1.$$

Значит, $0<A\leqslant a=B<b$ и

$$A^2+B^2+1=A^2+a^2+1=$$
$$=A^2+Ab=A(A+b)=ABc.$$

Получили противоречие: число b не наименьшее из возможных!

д) Пусть $x^2-(n^2-4)y^2=-4$, где $x,\,y$ — натуральные числа. Число x той же четности, что и ny. Значит, число $a=\dfrac{x+ny}{2}$ натуральное. Но $a^2+y^2+1=ayn$. ■

Паскаля с нулевой по пятнадцатую приведены в статье «Числа сочетаний».) Очевидно, $1001+2002=3003$. Значит, $C_{14}^4+C_{14}^5=C_{14}^6$, то есть $C_{15}^5=C_{14}^6$, ибо любое равенство $C_n^{m-2}+C_n^{m-1}=C_n^m$ можно записать в виде $C_{n+1}^{m-1}=C_n^m$.

Теорема 10. ***Равенство $C_x^{y-1}=C_{x-1}^y$ выполнено тогда и только тогда, когда $x=\varphi_{2k}\varphi_{2k+1}$ и $y=\varphi_{2k-1}\varphi_{2k}$, где k — некоторое натуральное число.***

Доказательство. I способ. Выразим числа сочетаний через факториалы:

$$\frac{x!}{(y-1)!\,(x-y+1)!}=\frac{(x-1)!}{y!\,(x-1-y)!},$$

которое очевидными преобразованиями приводим к виду $xy=(x-y+1)\times(x-y)$. Теперь применим некоторый специальный трюк. Обозначим буквой d наибольший общий делитель чисел x и y. Тогда $x=ad$ и $y=bd$, где a и b взаимно просты. Подставив выражения для x и y в уравнение, после сокращения на d получаем равенство

$$abd=(ad-bd+1)(a-b).$$

Поскольку числа $a-b$ и ab взаимно просты и поскольку числа d и $ad-bd+1$ тоже взаимно просты, то $ab=ad-bd+1$ и $d=a-b$, то есть

$$\begin{cases} a=b+d,\\ (b+d)b=(b+d)d-bd+1. \end{cases}$$

Последнее уравнение после упрощений приобретает вид $b^2+bd-d^2=1$. Его решения в натуральных числах нам известны: $b=\varphi_{2k-1}$ и $d=\varphi_{2k}$, где k — натуральное число. Таким образом,

$$\begin{cases} x=ad=(\varphi_{2k-1}+\varphi_{2k})\varphi_{2k}=\varphi_{2k}\varphi_{2k+1},\\ y=bd=\varphi_{2k-1}\varphi_{2k}, \end{cases}$$

как и было обещано. Например, при $k=1,\,2,\,3$ имеем соответственно $(x;\,y)=(2;\,1),\,(15;\,6),\,(104;\,40)$.

Строго говоря, надо бы проверить, что всякая пара чисел $(x;\,y)=(\varphi_{2k}\varphi_{2k+1};\,\varphi_{2k-1}\varphi_{2k})$ удовлетворяет равенству $(x-y+1)(x-y)=xy$. Немного подумав, можно понять, что это очевидно: двигаться «снизу вверх» по только что изложенному решению даже легче, чем «сверху вниз». Впрочем, годится и прямая проверка: $x-y=\varphi_{2k+1}\varphi_{2k}-\varphi_{2k-1}\varphi_{2k}=(\varphi_{2k+1}-\varphi_{2k-1})\varphi_{2k}=\varphi_{2k}^2$ и $x-y+1=\varphi_{2k}^2+1$, так что

$$(x-y+1)(x-y)=(\varphi_{2k}^2+1)\varphi_{2k}^2=\varphi_{2k}\varphi_{2k+1}\varphi_{2k-1}\varphi_{2k}=xy,$$

где мы воспользовались тождеством $\varphi_{2k}^2+1=\varphi_{2k+1}\varphi_{2k-1}$.

II способ. Мы решили уравнение $(x-y+1)(x-y)=xy$, применив довольно неожиданный трюк. Но есть и другой — стандартный — способ. А именно, есть стандартная схема, по которой решают в целых числах уравнения второй степени. Давайте посмотрим, как эта схема работает. Первым делом раскроем скобки и приведем подобные: $x^2-3xy+y^2+x-y=0$.

Теперь освободимся от членов первой степени. Для этого выполним замену $x=X+a$, $y=Y+b$, получив уравнение

$$X^2+2aX+a^2-3XY-3aY-3bX-3ab+Y^2+2bY+b^2+X+a-Y-b=0,$$

и приравняем коэффициенты при X и Y к нулю:

$$\begin{cases} 2a-3b+1=0,\\ -3a+2b-1=0. \end{cases}$$

Решив эту систему, находим $a=-1/5$ и $b=1/5$. При этих значениях a и b уравнение принимает вид $X^2-3XY+Y^2=\dfrac{1}{5}$, где $X=x+\dfrac{1}{5}$ и $Y=y-\dfrac{1}{5}$. Домножив

обе части уравнения на 20, получаем

$$20X^2-60XY+20Y^2=4,$$
$$5(4X^2-12XY+9Y^2)-25Y^2=4,$$
$$(5Y)^2-5(2X-3Y)^2=-4,$$
$$z^2-5t^2=-4,$$

где $z=5Y=5y-1$ и $t=2X-3Y=2x-3y+1$.

Пора воспользоваться следствием из теоремы 7. А именно, все решения уравнения $z^2-5t^2=\pm4$ в натуральных числах даются формулой $(z;\,t)=(\varphi_{n+1}+\varphi_{n-1};\,\varphi_n)$. При этом знаку «+» соответствуют четные n, а знаку «–» — нечетные. Осталось понять, при каких нечетных n число $z=\varphi_{n+1}+\varphi_{n-1}$ дает остаток 4 при делении на 5. Выпишем остатки от деления нескольких первых чисел Фибоначчи на 5:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
φ_n	1	1	2	3	5	8	13	21	34	55	89	144	233	377
$\varphi_n \bmod 5$	1	1	2	3	0	3	3	1	4	0	4	4	3	2
$\varphi_{n-1}+\varphi_{n+1} \bmod 5$	1	3	**4**	2	1	3	**4**	2	1	3	**4**	2	1	

Закономерность очевидна: $n\equiv3 \pmod 4$. Итак,

$$y=\frac{z+1}{5}=\frac{\varphi_{n-1}+\varphi_{n+1}+1}{5},$$

$$x=\frac{t+3y-1}{2}=\frac{\varphi_n+3\dfrac{\varphi_{n-1}+\varphi_{n+1}+1}{5}-1}{2}=\frac{3\varphi_{n-1}+5\varphi_n+3\varphi_{n+1}-2}{10}=$$
$$=\frac{\varphi_n+3\varphi_{n+1}-1}{5}=\frac{\varphi_{n+1}+\varphi_{n+3}-1}{5},$$

где $n\equiv3 \pmod 4$. Обозначив $n=4k-1$, запишем эти формулы в виде

$$x=\frac{\varphi_{4k}+\varphi_{4k+2}-1}{5}=\varphi_{2k}\varphi_{2k+1} \quad \text{и} \quad y=\frac{\varphi_{4k-2}+\varphi_{4k}+1}{5}=\varphi_{2k-1}\varphi_{2k},$$

где мы воспользовались тождеством $\varphi_{2m}+\varphi_{2m+2}-(-1)^m=5\varphi_m\varphi_{m+1}$, которое можно доказать по индукции, завершив тем самым второе доказательство теоремы 10. ■

Следующее вычисление — пожалуй, самое главное в теории уравнений Пелля:

$$(x^2-dy^2)(z^2-dt^2)=x^2z^2-dy^2z^2-dx^2t^2+d^2y^2t^2=$$
$$=x^2z^2+2xzdyt+d^2y^2t^2-dy^2z^2-2dyzxt-dx^2t^2=(xz+dyt)^2-d(xt+yz)^2.$$

А вот как можно получить ту же формулу, если разложить разность квадратов на (иррациональные!) множители и переставить их разумным образом:

$$(x^2-dy^2)(z^2-dt^2)=(x+y\sqrt{d})(x-y\sqrt{d})(z+t\sqrt{d})(z-t\sqrt{d})=$$
$$=(x+y\sqrt{d})(z+t\sqrt{d})\cdot(x-y\sqrt{d})(z-t\sqrt{d})=$$
$$=(xz+dyt+(xt+yz)\sqrt{d})\cdot(xz+dyt-(xt+yz)\sqrt{d})=(xz+dyt)^2-d(xt+yz)^2.$$

Честно говоря, эта выкладка даже длиннее предыдущей. Но она, надеемся, гораздо прозрачнее. Зачем нам нужна доказанная формула? Чтобы строить из одних решений другие! Точнее говоря, формула доказывает следующую важную теорему.

Теорема 11. ***Если $x^2-dy^2=a$ и $z^2-dt^2=b$, то пара чисел $(X;\,Y)=(xz+dyt;\,xt+yz)$ удовлетворяет равенству $X^2-dY^2=ab$.***

Уравнение $(x+1)^3-x^3=y^2$ имеет бесконечно много решений в натуральных числах.

Доказательство. Домножим обе части уравнения $3x^2+3x+1=y^2$ на 4 и выделим полный квадрат: $3(4x^2+4x+1)+1=(2y)^2$, то есть $(2y)^2-3(2x+1)^2=1$. Обозначив $z=2y$ и $t=2x+1$, получаем уравнение Пелля $z^2-3t^2=1$. Нас интересуют не все решения последнего уравнения, а лишь те, где z четно. В любом решении уравнения $z^2-3t^2=1$ одно из чисел z и t четно, а другое нечетно. При переходе $(z;\,t)\to(2z+3t;\,z+2t)$ пара (четное; нечетное) преобразуется в (нечетное; четное), и наоборот. Поэтому нужно рассматривать только «половину» решений, а именно $(z;\,t)=(26;\,15)$, $(362;\,209)$, $(5042;\,2911)$, $(70\,226;\,40\,545)$, $(978\,122;\,564\,719)$, $(13\,623\,482;\,7\,865\,521)$ и так далее. Этим решениям соответствуют пары $(x;\,y)=(7;\,13)$, $(104;\,181)$, $(1455;\,2521)$, $(20\,272;\,35\,113)$, $(282\,359;\,489\,061)$, $(3\,932\,760;\,6\,811\,741)$ и так далее. В частности, $8^3-7^3=13^2$ и $3\,932\,761^3-3\,932\,760^3=6\,811\,741^2$. Согласитесь, последняя формула впечатляет! ■

Уравнение $(x+2)^3-x^3=y^2$ не имеет решений в целых числах, поскольку $6x^2+12x+8\equiv2\not\equiv y^2 \pmod 3$. ■

Если квадрат некоторого натурального числа n представим в виде разности кубов последовательных целых чисел, то число n есть сумма квадратов двух последовательных целых чисел.

Доказательство. $(2n-1)(2n+1)=3(2x+1)^2$. Числа $2n-1$ и $2n+1$ взаимно просты, так что одно из них должно быть квадратом, а другое — утроенным квадратом. Значит, либо $2n-1=3t^2$ и $2n+1=s^2$, либо $2n-1=t^2$ и $2n+1=3s^2$. В первом случае $s^2-3t^2=2$, что невозможно, поскольку квадрат целого числа не может давать остаток 2 при делении на 3. Значит, имеет место второй случай: $2n-1=t^2$. Обозначив $t=2k+1$, из равенства $2n-1=(2k+1)^2$ получаем $n=2k^2+2k+1=k^2+(k+1)^2$. ■

Применяя метод доказательства теорем 2, 3, 5, 8, 12 или напрямую утверждение теоремы 14, нетрудно убедиться, что уравнение $x^2-34y^2=-1$ не имеет решений в целых числах. Верны следующие утверждения.

а) Для любого простого числа p существуют такие целые числа x и y, что $x^2-34y^2\equiv -1 \pmod p$.
б) Если p — нечетное простое число, n — натуральное, x и y — такие целые числа, что x^2-34y^2+1 делится на p^n, то существуют такие целые числа z и t, что $(x+p^nz)^2-34(y+p^nt)^2+1$ делится на p^{n+1}.
в) Если $n>2$ — натуральное число, x и y — такие целые числа, что x^2-34y^2+1 делится на 2^n и не делится на 2^{n+1}, то число $(x+2^{n-1})^2-34y^2+1$ делится на 2^{n+1}!
г) Если m_1 и m_2 — взаимно простые натуральные числа, для которых существуют такие целые числа x_1, y_1, x_2 и y_2, что

$$x_1^2-34y_1^2\equiv -1 \pmod{m_1},$$
$$x_2^2-34y_2^2\equiv -1 \pmod{m_2},$$

то существуют такие целые x и y, что $x^2-34y^2\equiv -1 \pmod{m_1m_2}$.
д) Для любого натурального m сравнение $x^2-34y^2\equiv -1 \pmod m$ имеет решения в целых числах x и y.

Указания. а) Для $p=2$ годятся $x=y=1$. Для $p=17$ годятся $x=4$, $y=0$. Для любого другого простого p рассмотрим числа вида x^2, где $x=0, 1, \ldots, (p-1)/2$, и числа вида $34y^2-1$, где $y=0, 1, \ldots \ldots, (p-1)/2$. Докажите, что как $(p+1)/2$ рассматриваемых чисел вида x^2, так и $(p+1)/2$ рассматриваемых чисел вида $34y^2-1$ дают разные остатки при делении на p. Поскольку $(p+1)/2+(p+1)/2>p$, то хотя бы одно число одного вида сравнимо по модулю p с числом другого вида, то есть найдется такая пара $(x; y)$, что x^2-34y^2+1 делится на p.
г) Воспользуйтесь китайской теоремой об остатках, то есть тем, что существуют такие числа x и y, для которых $x\equiv x_1 \pmod{m_1}$, $x\equiv x_2 \pmod{m_2}$, $y\equiv y_1 \pmod{m_1}$ и $y\equiv y_2 \pmod{m_2}$.

Замечание. Ситуация, когда сравнения имеют решения, а уравнение не имеет, не столь уж редка. Например, для любого натурального числа m сравнение $(3x+1)\times\times(2x+1)\equiv 0 \pmod m$ имеет решения в целых числах, а уравнение $(3x+1)(2x+1)=0$ не имеет целых решений. Тем интереснее знать, что 34 — наименьшее не являющееся точным квадратом натуральное число d, для которого все сравнения вида $x^2-dy^2\equiv\equiv -1 \pmod m$ имеют решения, а уравнение $x^2-dy^2=-1$ целочисленных решений не имеет. ■

Следствие. ***Если d — натуральное число, не являющееся квадратом, $c\neq 0$ и уравнение $x^2-dy^2=c$ имеет хотя бы одно решение в целых числах, то это уравнение имеет бесконечно много решений в натуральных числах.***

Доказательство. Можно считать, что $x\geqslant 0$ и $y\geqslant 0$. Рассмотрим натуральные числа a и b, для которых $a^2-db^2=1$. Тогда числа $x_1=ax+dby$ и $y_1=bx+ay$ натуральные. Формулы $x_{n+1}=ax_n+dby_n$ и $y_{n+1}=bx_n+ay_n$ дают бесконечную последовательность решений.

Теорема 12. ***Если a — наименьшее натуральное число, для которого существует такое натуральное число b, что $a^2-db^2=1$, то уравнение $x^2-dy^2=1$ не имеет решений в целых неотрицательных числах, кроме тех, что получаются из «тривиального» решения $(1; 0)$ при помощи правила $(x; y)\to(ax+dby; bx+ay)$.***

Доказательство. Рассмотрим систему уравнений:

$$\begin{cases} xz+dyt=X,\\ xt+yz=Y. \end{cases}$$

Чтобы найти x, домножим первое уравнение на z, второе — на dt и вычтем затем второе уравнение из первого:

$$zX-dtY=xz^2-dxt^2=x,$$

поскольку $z^2-dt^2=1$. Аналогично, чтобы найти y, домножим первое уравнение на t, второе на z и вычтем второе уравнение из первого:

$$Xt-Yz=dyt^2-yz^2,$$

откуда $y=Yz-Xt$.

Лемма 3. ***Если X, Y — натуральные числа, удовлетворяющие равенству $X^2-dY^2=1$, а z — наименьшее натуральное число, для которого существует такое натуральное число t, что $z^2-dt^2=1$, то $zX-dtY\geqslant 0$ и $Yz-Xt\geqslant 0$, причем $Yz-Xt<Y$.***

Доказательство. Рассуждаем «от противного». Если $zX-dtY<0$, то $X<\dfrac{dtY}{z}$ и, следовательно,

$$1=X^2-dY^2<\left(\frac{dtY}{z}\right)^2-dY^2=dY^2\frac{dt^2-z^2}{z^2}<0.$$

Если $Yz-Xt<0$, то

$$X^2-dY^2>\frac{Y^2z^2}{t^2}-dY^2=\frac{Y^2z^2-dY^2t^2}{t^2}=\frac{Y^2}{t^2}\geqslant 1.$$

(Последнее неравенство следует из того, что наименьшему z отвечает и наименьшее t.) Если же $Yz-Xt\geqslant Y$, то $X\leqslant(Yz-Y)/t$ и

$$X^2-dY^2\leqslant\frac{Y^2(z-1)^2}{t^2}-dY^2=Y^2\frac{z^2-2z+1-dt^2}{t^2}=Y^2\frac{2-2z}{t^2}\leqslant 0.$$

Лемма доказана. Доказательство теоремы проводится в точности так, как доказательство теоремы 2. ■

Использование иррациональностей. Неравенства, неравенства, неравенства. . . Есть ощущение какого-то фокуса, когда все сходится, но причина удачи спрятана и не видна наивному зрителю. Сейчас мы докажем теорему 12 заново. Надеемся, этим мы поможем вам вполне уяснить ее доказательство.

Лемма 4. ***Если $x^2-dy^2>0$ и $x+y\sqrt{d}>0$, то $x>0$.***

Доказательство. $2x=x+y\sqrt{d}+\dfrac{x^2-dy^2}{x+y\sqrt{d}}>0.$

Существует бесконечно много таких натуральных n, что $n!$ делится на n^2+1.

Доказательство. Существует бесконечно много таких пар натуральных чисел m и n, для которых $n^2+1=5m^2$ и $m>5$. При этом $m=\sqrt{\frac{n^2+1}{5}}<\frac{n}{2}$ и $n!$ делится на n^2+1, так как при $m>5$ в произведении $1\cdot 2\cdot 3\cdot 4\cdot \ldots \cdot (n-1)\cdot n$ есть множители 5, m и $2m$. ■

Для любого числа $\alpha>0$ существует бесконечно много таких натуральных n, что $[\alpha n]!$ делится на n^2+1.

Доказательство. В силу равенства $k^2-(k^2+1)\cdot 1=-1$ существуют сколь угодно большие натуральные числа d, для которых уравнение $x^2-dy^2=-1$ имеет хотя бы одно решение в натуральных числах — а следовательно, и бесконечно много.

Есть и другие способы доказательства. Например, можно воспользоваться разложением многочлена $x^{105}+1$ на неприводимые многочлены с целыми коэффициентами или разложением

$$64m^{12}+1=(4m^4+1)\times$$
$$\times(4m^4-4m^3+2m^2-2m+1)\times$$
$$\times(4m^4+4m^3+2m^2+2m+1).$$ ■

Если n — целое неотрицательное число, то $[(1+\sqrt{3})^{2n+1}]$ делится на 2^{n+1} и не делится на 2^{n+2}.

Доказательство. Поскольку

$$-1<(1-\sqrt{3})^{2n+1}<0$$

и число $(1+\sqrt{3})^{2n+1}+(1-\sqrt{3})^{2n+1}$ целое, то

$$[(1+\sqrt{3})^{2n+1}]=$$
$$=(1+\sqrt{3})^{2n+1}+(1-\sqrt{3})^{2n+1}=$$
$$=(1+\sqrt{3})(4+2\sqrt{3})^n+$$
$$+(1-\sqrt{3})(4-2\sqrt{3})^n=$$
$$=2^n\cdot((1+\sqrt{3})(2+\sqrt{3})^n+$$
$$+(1-\sqrt{3})(2-\sqrt{3})^n)=$$
$$=2^n\cdot(2x_n+6y_n)=2^{n+1}(x_n+3y_n),$$

где $x_n=\frac{(2+\sqrt{3})^n+(2-\sqrt{3})^n}{2}$ и $y_n=\frac{(2+\sqrt{3})^n-(2-\sqrt{3})^n}{2\sqrt{3}}$ удовлетворяют равенству $x_n^2-3y_n^2=1$. Осталось заметить, что x_n и y_n — числа разной четности. ■

Есть и другой способ — «от противного». Предположим, что $x\leqslant 0$. Тогда обе части неравенства $y\sqrt{d}>-x$ можно возвести в квадрат: $dy^2>x^2$, что противоречит неравенству $x^2-dy^2>0$.

Лемма 5. ***Если $x^2-dy^2=1$ и $x+y\sqrt{d}>1$, то $y>0$.***

Доказательство. Пусть $y\leqslant 0$. Тогда $x-y\sqrt{d}\geqslant x+y\sqrt{d}>1$. Произведение чисел $x-y\sqrt{d}$ и $x+y\sqrt{d}$, каждое из которых больше 1, не может равняться 1.

Лемма 6. ***Если $a^2-db^2=x^2-dy^2$ и $x+y\sqrt{d}<a+b\sqrt{d}$, причем числа a, b, x и y неотрицательные, то $x<a$ и $y<b$.***

Доказательство. $a-b\sqrt{d}=\frac{a^2-db^2}{a+b\sqrt{d}}<\frac{x^2-dy^2}{x+y\sqrt{d}}=x-y\sqrt{d}$. Сложив неравенства

$$-x+y\sqrt{d}<-a+b\sqrt{d},$$
$$x+y\sqrt{d}<a+b\sqrt{d},$$

получаем: $2y\sqrt{d}<2b\sqrt{d}$. Дальнейшее очевидно.

Лемма 7. ***Пусть a — наименьшее натуральное число, для которого существует такое натуральное число b, что $a^2-db^2=1$. Если x, y — целые числа и $1<x+y\sqrt{d}<a+b\sqrt{d}$, то $x^2-dy^2\neq 1$.***

Доказательство. Предположим противное: $x^2-dy^2=1$. Тогда в силу лемм 4 и 5 числа x и y положительны. В силу леммы 6 имеем $x<a$. Получили противоречие.

Следующая теорема — это другая формулировка теоремы 12.

Теорема 13. ***Пусть a — наименьшее натуральное число, для которого существует такое натуральное число b, что $a^2-db^2=1$. Если x, y — целые числа, $x^2-dy^2=1$ и $x+y\sqrt{d}>0$, то для некоторого целого числа n верно равенство $x+y\sqrt{d}=(a+b\sqrt{d})^n$.***

Доказательство. Обозначим $q=a+b\sqrt{d}$. Поскольку числа a и b натуральные, то $q>1$. Рассмотрим возрастающую геометрическую прогрессию:

$$1<q<q^2<q^3<q^4<q^5<\ldots$$

Она стремится к бесконечности. А убывающая геометрическая прогрессия

$$1<\frac{1}{q}<\frac{1}{q^2}<\frac{1}{q^3}<\frac{1}{q^4}<\frac{1}{q^5}<\ldots$$

стремится к нулю. Поэтому существует такое целое n, что $q^{n-1}<x+y\sqrt{d}\leqslant q^n$. Рассмотрим число $E=(x+y\sqrt{d})/q^{n-1}$. Очевидно, $1<E\leqslant q$. Поскольку

$$\frac{1}{q}=\frac{1}{a+b\sqrt{d}}=\frac{a-b\sqrt{d}}{(a-b\sqrt{d})(a+b\sqrt{d})}=\frac{a-b\sqrt{d}}{a^2-db^2}=a-b\sqrt{d},$$

то $E=(x+y\sqrt{d})(a-b\sqrt{d})^{n-1}$. Воспользовавшись формулой

$$(r+s\sqrt{d})(u+v\sqrt{d})=(ru+dsv)+(rv+su)\sqrt{d},$$

мы заключаем, что число E представимо в виде $E=z+t\sqrt{d}$, где z, t — целые числа. Переходя к сопряженным числам, получаем:

$$z-t\sqrt{d}=(x-y\sqrt{d})(a+b\sqrt{d})^{n-1}.$$

Следовательно,

$$z^2-dt^2=(z+t\sqrt{d})(z-t\sqrt{d})=(x+y\sqrt{d})(x-y\sqrt{d})(a-b\sqrt{d})^{n-1}(a+b\sqrt{d})^{n-1}=$$
$$=(x^2-dy^2)(a^2-db^2)^{n-1}=1.$$

Итак, числа z и t целые, $1<z+t\sqrt{d}\leqslant a+b\sqrt{d}$ и $z^2-dt^2=1$. В силу леммы 7 это возможно лишь в случае равенства $z+t\sqrt{d}=a+b\sqrt{d}$, то есть когда $x+y\sqrt{d}=q^n$. ■

Уравнение $x^2-dy^2=c$.

Доказательства теорем 12 и 13 довольно длинные. Не вполне ясно, что проще: жонглировать неравенствами или иррациональностями. Оказывается, однако, что использованное при доказательстве теоремы 13 рассуждение позволяет весьма ясно показать, как устроены решения в целых числах уравнения $x^2-dy^2=c$.

Напомним обозначения. Как и прежде, d — натуральное число, не являющееся квадратом; a — наименьшее натуральное число, для которого существует такое натуральное число b, что $a^2-db^2=1$; $q=a+b\sqrt{d}$; наконец, c — некоторое целое число, $c\neq 0$.

Пусть x и y — целые числа, $x^2-dy^2=c$ и $x+y\sqrt{d}>0$. Рассмотрим числа вида q^n, где n пробегает множество всех целых чисел. Поскольку $\lim\limits_{n\to-\infty} q^n=0$ и $\lim\limits_{n\to+\infty} q^n=+\infty$, то существует такое целое число n, что

$$q^{n-1}<x+y\sqrt{d}\leqslant q^n.$$

Рассмотрим число $E=(x+y\sqrt{d})/q^{n-1}$. Легко понять, что E представимо в виде $E=z+t\sqrt{d}$, где z и t — целые числа. При этом $z^2-dt^2=c$ и $1<z+t\sqrt{d}\leqslant q$.

Теорема 14. ***Рассмотрим всевозможные пары целых чисел $(z;\,t)$, удовлетворяющие последним двум условиям. Верны следующие утверждения.***

1) ***Если множество M таких пар пусто, то уравнение $x^2-dy^2=c$ не имеет решений в целых числах x и y.***

2) ***Множество M конечно.***

3) ***Все целочисленные решения уравнения $x^2-dy^2=c$ можно получить из формул $x+y\sqrt{d}=\pm(z+t\sqrt{d})q^n$, где $(z;\,t)\in M$, а n — целое число.***

Доказательство. Первое и третье утверждения очевидны. Докажем второе. Пусть $(z;\,t)\in M$. Тогда

$$z-t\sqrt{d}=\frac{c}{z+t\sqrt{d}},$$

так что $|z-t\sqrt{d}|<|c|$ и, следовательно,

$$|z|=\left|\frac{(z+t\sqrt{d})+(z-t\sqrt{d})}{2}\right|<\frac{q+|c|}{2}\quad\text{и}\quad |t|=\left|\frac{(z+t\sqrt{d})-(z-t\sqrt{d})}{2\sqrt{d}}\right|<\frac{q+|c|}{2\sqrt{d}}.\ \blacksquare$$

В 1657 г. — в довольно поздний период своей деятельности — П. Ферма в качестве вызова разослал другим математикам, в частности английским, следующую задачу.

«Сейчас едва ли найдется кто-нибудь, кто предлагает арифметические вопросы, и кто-нибудь, кто их понимает. Не потому ли это происходит, что до сих пор арифметику рассматривали скорее с геометрической, чем с арифметической точки зрения? Так было всегда — и в древних, и в современных работах; примером тому является даже Диофант. Ибо хотя он и более чем другие освободился от геометрии в том отношении, что ограничивает свой анализ рассмотрением рациональных чисел, однако даже у него геометрия не полностью отсутствует. . .

Теперь арифметика имеет, так сказать, собственную область изучения — теорию целых чисел. Евклид лишь слегка затронул ее в своих «Началах», а его последователи недостаточно занимались этой теорией (если только она не содержалась в тех книгах Диофанта, которых мы лишились вследствие разрушительного действия времени); следовательно, арифметикам предстоит развивать или восстанавливать ее.

Поэтому арифметикам, дабы осветить тот путь, по которому надо следовать, предлагаю я эту теорему, чтобы они доказали ее, или эту задачу, чтобы они

Если натуральные числа k, m и n удовлетворяют равенству $m+n\sqrt{3}=(2+\sqrt{3})^k$, где k четно, то число $\sqrt{(m+1)/2}$ целое. А если k нечетно, то целое число $\sqrt{m-1}$.

Указания. Числа m и n удовлетворяют равенству $m^2-3n^2=1$, которое можно записать в виде $(m-1)(m+1)=3n^2$.

Если k четно, то m нечетно; следовательно, НОД$(m-1,\,m+1)=2$. Значит, одно из чисел $m-1$ и $m+1$ имеет вид $2a^2$, а другое — $6b^2$. В случае, когда $m-1=2a^2$ и $m+1=6b^2$, имеем $a^2=3b^2-1$, что невозможно. Следовательно, $m+1=2a^2$.

Если k нечетно, то m четно, так что числа $m-1$ и $m+1$ взаимно просты. Дальнейшее очевидно. ■

Пусть p — простое число и $x^2-py^2=1$, где x, y — натуральные числа. Докажем, что если x (не)четно, то одно из чисел $x-1$ или $x+1$ является (удвоенным) квадратом.

Доказательство. Очевидно, $(x-1)(x+1)=py^2$. Если x четно, то НОД$(x-1,\,x+1)=1$, так что

$$\begin{cases} x-1=a^2,\\ x+1=pb^2 \end{cases}\quad\text{или}\quad \begin{cases} x-1=pa^2,\\ x+1=b^2. \end{cases}$$

Если же x нечетно, то НОД$(x-1,\,x+1)=2$ и имеем системы

$$\begin{cases} x-1=2a^2,\\ x+1=2pb^2 \end{cases}\quad\text{или}\quad \begin{cases} x-1=2pa^2,\\ x+1=2b^2. \end{cases}\ \blacksquare$$

Для любого натурального n между числами n^2 и $n^2+n+3\sqrt{n}$ найдутся три натуральных числа, произведение двух из которых делится на третье.

Доказательство. Пусть $n>2$ и

$$a=(n-x)(n+x+1)=\\ =n^2+n-x^2-x,$$

$$b=(n-x+1)(n+x)=\\ =n^2+n-x^2+x,$$

$$c=(n-x+1)(n+x+1)=\\ =n^2+2n+1-x^2,$$

где x — наибольшее натуральное число, для которого $x^2+x<n$. Очевидно, $ab\,\vdots\,c$ и $n^2<a<b<c$. Докажем неравенство $c<n^2+n+3\sqrt{n}$, то есть $n+1-x^2<3\sqrt{n}$. Пусть $x^2\leqslant n-3\sqrt{n}+1$. Тогда $x<\sqrt{n}-\frac{3}{2}$ и, следовательно,

$$(x+1)^2+(x+1)<\\ <\left(n-\sqrt{n}+\frac{1}{4}\right)+\left(\sqrt{n}-\frac{1}{2}\right)<n,$$

что противоречит выбору x. ■

Для каждого не являющегося квадратом натурального числа $d \leqslant 150$ в таблице указано такое наименьшее натуральное y, что dy^2+1 — квадрат натурального числа. Заметьте: случаи $d=109$ и 149 выделяются на общем фоне.

2	2	54	66	103	22 419
3	1	55	12	104	5
5	4	56	2	105	4
6	2	57	20	106	3 115 890
7	3	58	2574	107	93
8	1	59	69	108	130
10	6	60	4	109	15 140 424 455 100
11	3	61	226 153 980	110	2
12	2	62	8	111	28
13	180	63	1	112	12
14	4	65	16	113	113 296
15	1	66	8	114	96
17	8	67	5967	115	105
18	4	68	4	116	910
19	39	69	936	117	60
20	2	70	30	118	28 254
21	12	71	413	119	11
22	42	72	2	120	1
23	5	73	267 000	122	22
24	1	74	430	123	11
26	10	75	3	124	414 960
27	5	76	6630	125	83 204
28	24	77	40	126	40
29	1820	78	6	127	419 775
30	2	79	9	128	51
31	273	80	1	129	1484
32	3	82	18	130	570
33	4	83	9	131	927
34	6	84	6	132	2
35	1	85	30 996	133	224 460
37	12	86	1122	134	12 606
38	6	87	3	135	21
39	4	88	21	136	3
40	3	89	53 000	137	519 712
41	120	90	2	138	4
42	2	91	165	139	6 578 829
43	531	92	120	140	6
44	30	93	1260	141	8
45	24	94	221 064	142	12
46	3588	95	4	143	1
47	7	96	5	145	24
48	1	97	6 377 352	146	12
50	14	98	10	147	8
51	7	99	1	148	6
52	90	101	20	149	2 113 761 020
53	9100	102	10	150	4

решили ее. Если же преуспеют они в ее доказательстве или решении, то им придется признать, что вопросы такого рода ничем не уступают в отношении красоты, трудности или метода доказательства самым знаменитым вопросам геометрии.

Если дано произвольное число, которое не является квадратом, то найдется бесконечное множество таких квадратов, что если этот квадрат умножить на данное число и к произведению прибавить единицу, то результат будет квадратом.

Пример. Пусть 3, которое не является квадратом, будет данным числом. Если умножить его на квадрат, равный 1, и к произведению добавить 1, то в результате получится 4, что является квадратом. Если то же самое число 3 умножить на квадрат 16, то получится произведение, которое при увеличении на 1 превращается в 49, тоже квадрат. И кроме 1 и 16 можно найти бесконечное множество квадратов с тем же свойством.

Но я спрашиваю об общем правиле решения — когда дано произвольное число, не являющееся квадратом. Например, найдите такой квадрат, что если произведение этого квадрата и числа 149, 109 или 433 увеличить на 1, то получится квадрат». ■

Вступление Ферма к этой задаче ясно показывает, что он желает не традиционного диофантова решения в рациональных числах, а решения в целых числах. (По иронии судьбы ныне слово «диофантово» употребляют, желая получить решения в целых числах, тогда как сам Диофант ни в одной из дошедших до нас работ не занимался решениями в целых числах, а только в рациональных.) Как это ни странно, вступление было опущено одним из посредников в том экземпляре письма, который был передан английским математикам; в результате они сочли задачу совершенно глупой. А именно, можно ввести обозначение $x=1+\frac{m}{n}y$ и подставить в уравнение:

$$\left(1+\frac{m}{n}y\right)^2-dy^2=1,$$

$$\frac{2m}{n}y+\frac{m^2}{n^2}y^2-dy^2=0,$$

$$2mn=(dn^2-m^2)y,$$

откуда

$$y=\frac{2mn}{dn^2-m^2}, \qquad x=\frac{dn^2+m^2}{dn^2-m^2}.$$

Полученные формулы, как легко убедиться, дают бесконечно много решений в рациональных числах.

Когда же дополнительное требование, что x и y должны быть целыми числами, дошло до английских математиков, то они пожаловались, что условие задачи изменили. Конечно, их жалобу можно понять в свете сильной диофантовой традиции, но, как указал Ферма, было наивно надеяться, что он предложил тривиальную задачу. Как видно из таблицы, задача Ферма весьма сложная: для

$d=61$ наименьшее решение — это пара $y=226\,153\,980$ и $x=176\,631\,9049$. (Впрочем, впервые посчитал это не Ферма, а родившийся в 1114 г. индиец Бхаскара Акхария.) А для $d=109$ вообще $y=15\,140\,424\,455\,100$.

Англичанам удалось не только найти частные решения при $d=149$, 109 или 433, но и разработать общую процедуру получения решений для любого значения d. Кто это сделал — неизвестно. Хотя Дж. Валлис первым дал описание процедуры и получил решения в трех частных случаях, он приписывает авторство виконту У. Броункеру. В опубликованной переписке Валлиса нет никаких указаний на то, что Броункер когда-либо сообщал ему что-либо об этом методе, кроме нескольких простых замечаний, которые, быть может, послужили зародышем идеи, развитой впоследствии Валлисом. Возможно, Валлису было важно добиться расположения Броункера и добиться его покровительства, поэтому он и назвал этот метод методом Броункера (ибо Броункер не только принадлежал к знати, но и был первым президентом Королевского общества). Впрочем, некоторые историки считают самого Броункера весьма способным математиком и утверждают, что по своим личным качествам Валлис скорее мог приписать себе чужие заслуги, чем отказаться от своих.

Строго говоря, англичане не решили задачу Ферма, которая заключалась в том, что при данном (не являющемся квадратом) натуральном d существует бесконечно много натуральных x таких, что dx^2+1 является квадратом. Они не доказали, что процедура всегда завершится, и, кажется, даже не понимали, что это нужно доказывать. (Даже Эйлеру не удалось доказать, что английский метод всегда приводит к успеху. Удалось — Лагранжу через 110 лет после того, как Валлис отослал ответ на вызов Ферма.)

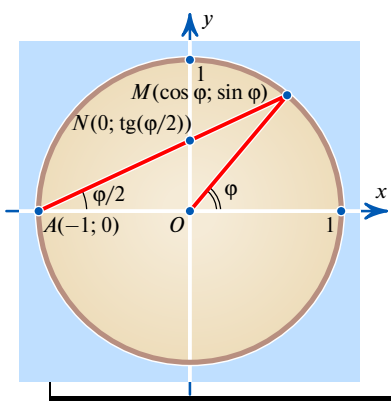
Ферма написал письмо, в котором признал, что англичанам удалось решить его задачу. Однако главным для Ферма в этом письме было убедить англичан, что перед ними была поставлена достойная задача, так что он мог сознательно закрыть глаза на недостатки. Несколько лет спустя он указал, что англичане получили решение только в отдельных частных случаях и не дали общего доказательства. Очевидная интерпретация этого замечания заключается в том, что Ферма заметил отсутствие доказательства того, что предложенный ими процесс всегда приводит к решению; с другой стороны, в нем можно увидеть и менее глубокую критику того, что процесс был описан в недостаточно общих терминах. Ферма утверждает, что он мог бы дать доказательство, «надлежащим образом» применив метод бесконечного спуска. Эти слова, разумеется, нельзя считать достаточным свидетельством в пользу того, что он умел решать свою задачу.

Индийский и английский методы. Легенды гласят, что за несколько веков до нашей эры в Индии было известно равенство $2\cdot 408^2+1=577^2$. Равенство $92\cdot 120^2+1=1151^2$ вместе с изощренной техникой его

Уравнение прямой AM имеет вид $y=k(x+1)$, где $k=\operatorname{tg}\frac{\varphi}{2}$. Подставляя значение y в уравнение окружности $x^2+y^2=1$, получаем $x^2-1+k^2(x+1)^2=0$, то есть $(x+1)(x-1+k^2x+k^2)=0$ и $x=\frac{1-k^2}{k^2+1}$. Значит, $y=k(x+1)=\frac{2k}{k^2+1}$. Мы вывели формулы

$$\cos\varphi=\frac{1-\operatorname{tg}^2\frac{\varphi}{2}}{1+\operatorname{tg}^2\frac{\varphi}{2}}\ \text{и}\ \sin\varphi=\frac{2\operatorname{tg}\frac{\varphi}{2}}{1+\operatorname{tg}^2\frac{\varphi}{2}}.$$

Заметьте: если число k рационально, то и числа x и y рациональны. Верно и обратное: если x и y рациональны, то и k рациональное.



Сумма квадратов 25 последовательных целых чисел может быть квадратом целого числа, а сумма 25 квадратов натуральных чисел — не может.

Доказательство. Обозначим через x среднее из этих чисел. Получаем сумму

$$(x-12)^2+(x-11)^2+(x-10)^2+\ldots$$
$$\ldots+(x+10)^2+(x+11)^2+(x+12)^2.$$

Раскрывая скобки и приводя подобные, получаем уравнение

$$25x^2+1300=y^2,$$

откуда

$$(y-5x)(y+5x)=1300.$$

Поскольку числа $y-5x$ и $y+5x$ отличаются на $10x$, а их произведение оканчивается цифрой 0, то $y-5x$ и $y+5x$ должны оканчиваться нулями. Для натуральных x и y получаем систему

$$\begin{cases} y-5x=10,\\ y+5x=130, \end{cases}$$

которой удовлетворяют числа $x=12$ и $y=70$. ■

Для каких натуральных чисел k множество целочисленных решений уравнения

$$(x+1)^2+(x+2)^2+\ldots+(x+k)^2=y^2$$

непусто и конечно?

Ответ: для $k=n^2$, где n не делится ни на 2, ни на 3.

Для любого натурального числа n существуют такие натуральные x и y, что $x^2-3y^2=1$ и y делится на 3^n, однако степенью тройки y быть не может. ■

Термин «уравнение Пелля» возник в результате ошибки Л. Эйлера. Почему-то — возможно, по причине смутных воспоминаний, оставшихся от чтения «Алгебры» Валлиса, — у Эйлера создалось впечатление, будто Валлис приписывает метод решения этой задачи не Броункеру, а Дж. Пеллю — современнику Валлиса, который много раз упомянут в его работах, но не имел никакого отношения к уравнению $x^2-dy^2=1$. Эйлер впервые сделал эту ошибку в 1730 г., когда ему было 23 года, но она попала и в окончательное издание «Введения в алгебру» (примерно 1770 г.). Эйлер был самым популярным математическим автором своего времени, и поэтому ошибка вошла в историю. . . ■

вывода было получено Брахмагуптой (родился в 598 г.). Общий способ решения уравнения Пелля дал Бхаскара Акхария. Этот метод называют циклическим или индийским.

Познакомимся с ним на примере $d=67$. Наша цель — найти такие натуральные x и y, чтобы разность y^2-67x^2 равнялась 1. В качестве первого приближения рассмотрим равенство

$$8^2-67\cdot 1^2=-3.$$

Вспомнив формулу $(a^2-67b^2)(c^2-67d^2)=(ac+67bd)^2-67(bc+ad)^2$ и, применив ее к равенствам $8^2-67\cdot 1^2=-3$ и $r^2-67\cdot 1^2=s$, где r (а тем самым и s) будут определены позже, получим:

$$(8r+67)^2-67(r+8)^2=-3s.$$

Пытаясь сделать правую часть (по модулю) как можно меньшей только за счет выбора наименьшего по модулю значения s, мы выбрали бы $r=8$, при котором $s=-3$, и получили бы равенство

$$131^2-67\cdot 16^2=9,$$

с которым непонятно что делать дальше.

Идея циклического метода — выбор такого r, чтобы $r+8$ делилось на 3 и s при этом было как можно меньше по модулю. (Когда это сделано, обе части уравнения разделятся нацело на 3^2.) А идея английского метода — выбор такого как можно большего r, что $r^2<d$ и $r+8$ делится на 3. Как видите, методы очень похожи. Оба можно применять для поиска решений при данном d, не будучи заранее уверенным, что это приведет к успеху. (Априори нет никакой уверенности в том, что в общем случае в английском методе после каждого шага r будет существовать. Это надо доказывать!)

Проведем вычисления для циклического метода. Чтобы $r+8$ делилось на 3, число r должно равняться одному из чисел бесконечной в обе стороны арифметической прогрессии $\ldots, -2, 1, 4, 7, 10, 13, 16, \ldots$ Выбор $r=7$ дает наименьшее по модулю значение $s=-18$. Этим r и s соответствует равенство

$$123^2-67\cdot 15^2=54,$$

которое после сокращения на 9 превращается в $41^2-67\cdot 5^2=6$. Теперь — следующий шаг циклического метода:

$$(41r+67\cdot 5)^2-67(5r+41)^2=6s.$$

Число $5r+41$ делится на 6 при $r=5, 11, 17, 23, \ldots$ Выбор $r=5$ дает наименьшее по модулю значение $s=-42$, и мы получаем равенство

$$540^2-67\cdot 66^2=6\cdot(-42),$$

которое после сокращения на 6^2 превращается в $90^2-67\cdot 11^2=-7$. Выполнив еще пять шагов циклического метода, получаем равенство $48\,842^2-67\cdot 5967^2=1$. А именно, взяв $r=9$, имеем $221^2-67\cdot 27^2=-2$; далее опять $r=9$ и $1899^2-67\cdot 232^2=-7$; потом $r=5$ и $3577^2-67\cdot 437^2=6$; на предпоследнем шаге $r=7$ приводит к равенству $9053^2-67\cdot 1106^2=-3$; наконец, $r=8$ дает ответ.

(Оригинальное индийское решение этой задачи использует прием, сокращающий вычисления. Обе части равенства $221^2-67\cdot 27^2=-2$ возводят в квадрат, получая $(221^2+67\cdot 27^2)^2-67\cdot(2\cdot 27\cdot 221)^2=(-2)^2$, после сокращения которого на 4 получаем искомый ответ.)

Легко убедиться, что и индийский, и английский методы позволяют найти решение для $d=67$. Однако ни для английского, ни для индийского метода нет никаких очевидных причин, по которым равенство с правой частью 1 должно обязательно получиться в общем случае. Есть и много других вопросов. Например, если эти методы дадут нам какое-то решение уравнения Пелля, можно ли утверждать, что это решение — наименьшее из возможных?

Существуют такие иррациональные числа $\alpha>1$ и $\beta>1$, что ни при каких натуральных m и n целые части чисел α^m и β^n не совпадают.

Доказательство. I способ. Пусть $\alpha=2+\sqrt{3}$ и $\beta=3\cdot(2+\sqrt{3})$. Тогда числа

$$a=(2+\sqrt{3})^m+(2-\sqrt{3})^m,$$
$$b=3^n(2+\sqrt{3})^n+3^n(2-\sqrt{3})^n$$

целые, причем a не делится на 3, а b — делится. Очевидно, $[(2+\sqrt{3})^m]=a-1$ и, поскольку

$$3(2-\sqrt{3})=\frac{3}{2+\sqrt{3}}<1,$$

то $[3^n(2+\sqrt{3})^n]=b-1$.

II способ. Числа

$$a=((2+\sqrt{3})^m+(2-\sqrt{3})^m)/2,$$
$$b=((1+\sqrt{2})^n+(1-\sqrt{2})^n)/2$$

— натуральные, причем для некоторых натуральных c и d имеем: $a^2-3c^2=1$ и $b^2-2d^2=(-1)^n$. Если $[(2+\sqrt{3})^m]=[(1+\sqrt{2})^n]$ и n нечетно, то $2a-1=2b$; если же n четно, то $2a-1=2b-1$, так что $a=b$ и $3c^2=2d^2$, что тоже невозможно.

III способ. Поскольку континуум несчетен, существуют такие положительные иррациональные α и β, что $\frac{1}{\alpha}+\frac{1}{\beta}=1$ и числа 10^α и 10^β иррациональны. Для любых натуральных чисел m и n при этом $[n\alpha]\neq[m\beta]$, так что числа $[10^{n\alpha}]$ и $[10^{m\beta}]$ имеют разное количество цифр.

IV способ естествен для того, кто знает теорему о стягивающихся отрезках и счетность множества рациональных чисел. Зафиксируем нецелое $\alpha>4$ и докажем существование иррационального β, удовлетворяющее условиям. Обозначим $a_1=[\alpha]+1{,}01$ и $b_1=[\alpha]+1{,}99$. Тогда $b_1>a_1>5$ и $b_1^2-a_1^2=(b_1+a_1)(b_1-a_1)>5$. Перенумеруем рациональные числа отрезка $[a_1; b_1]$, то есть выпишем их всех в виде последовательности $c_1, c_2, c_3, \ldots$ Построим такую последовательность отрезков $[a_k; b_k]$, что

— отрезок $[a_{k+1}; b_{k+1}]$ лежит в отрезке $[a_k; b_k]$;

— $c_k\notin[a_{k+1}; b_{k+1}]$;

— $b_k^{k+1}-a_k^{k+1}>5$;

— для любого числа $\beta\in[a_k; b_k]$ целая часть ни одного из чисел β^n, где $n\leqslant k$, не равна целой части ни одного из чисел α^m ни при каком натуральном m.

Поскольку $4^2-4=12$, то целые части степеней числа α различаются не менее чем на 11. Пусть отрезок $[a_k; b_k]$ построен. Поскольку длина отрезка $[a_k^{k+1}; b_k^{k+1}]$ больше 5, то в нем содержатся хотя бы четыре отрезка с концами в соседних натуральных числах. Хотя бы в одном из них нет (нецелого!) числа c_k^{k+1} и нет ни одной степени числа α; пусть это отрезок $[n; n+1]$. Положим $a_{k+1}=\sqrt[k+1]{n}$ и $b_{k+1}=\sqrt[k+1]{n+1}$. Тогда

$$b_{k+1}^{k+2}-a_{k+1}^{k+2}=b_{k+1}(n+1)-a_{k+1}n>$$
$$=a_{k+1}(n+1)-a_{k+1}n=a_{k+1}>5.$$

Построение закончено. Очевидно, общая точка β всех построенных отрезков удовлетворяет условиям. ■

Существует такая последовательность иррациональных чисел $\alpha_1, \alpha_2, \alpha_3, \ldots$, что равенство $[\alpha_r^m]=[\alpha_s^n]$, где r, s, m и n — натуральные числа, верно лишь при $r=s$ и $m=n$.

Доказательство. Построим такую последовательность положительных иррациональных чисел α_1, α_2, α_3, ... и такую последовательность простых чисел $p_1, p_2, p_3, \ldots$, что для любых натуральных m и n число $[\alpha_m^n]+1$ делится на p_m и не делится на p_k ни при каком $k<m$.

Лемма. ***Для любого натурального числа a уравнение $ax+1=y^2$ имеет бесконечно много решений в натуральных числах.***

Доказательство. Для любого натурального r положим $y=ar+1$ и $x=ar^2+2r$. Лемма доказана.

Начнем построение. Положим $p_1=3$ и $\alpha_1=3(2+\sqrt{3})$. Тогда при любом натуральном n число

$$[\alpha_1^n]+1=3^n(2+\sqrt{3})^n+3^n(2-\sqrt{3})^n$$

целое и даже кратное $p_1=3$.

Предположим, что числа $\alpha_1, \ldots, \alpha_n$ и $p_1, \ldots, p_n$ уже найдены. Рассмотрим произведение $a=p_1p_2\ldots p_n$. Выберем простое число $p_{n+1}>a$ и натуральные числа x и y, для которых $y^2=ax+1$ и $y>p_{n+1}$. Пусть

$$\alpha_{n+1}=p_{n+1}(y+\sqrt{ax}).$$

Тогда

$$[\alpha_{n+1}^m]+1=$$
$$=p_{n+1}^m(y+\sqrt{ax})^m+p_{n+1}^m(y-\sqrt{ax})^m$$

делится на p_{n+1} и не делится ни на одно из чисел $p_1, \ldots, p_n$. ■

Можно! Доказать это проще всего при помощи цепных дробей. Но мы здесь ограничимся неконструктивным доказательством существования решения уравнения Пелля.

Лемма 8. ***Для любого вещественного числа ξ и любого натурального числа N существуют такие целое число a и натуральное b, что $b\leqslant N$ и $|b\xi-a|\leqslant\frac{1}{N+1}$.***

Доказательство. Рассмотрим числа 0 и 1, а также дробные части чисел ξ, $2\xi, \ldots, N\xi$. Если бы все расстояния между этими $N+2$ числами были больше $\frac{1}{N+1}$, получилось бы противоречие. Значит, какое-то из расстояний не превосходит $\frac{1}{N+1}$. Если, например, $|\{b_2\xi\}-\{b_1\xi\}|\leqslant\frac{1}{N+1}$, где $1\leqslant b_1<b_2\leqslant N$, то

$$|(b_2\xi-[b_2\xi])-(b_1\xi-[b_1\xi])|\leqslant\frac{1}{N+1},$$

так что достаточно взять $b=b_2-b_1$ и $a=[b_2\xi]-[b_1\xi]$. Остальные два случая столь же очевидны: если $\{b\xi\}-0\leqslant\frac{1}{N+1}$, то годится $a=[b\xi]$; если же $1-\{b\xi\}\leqslant$ $\leqslant\frac{1}{N+1}$, то можно взять $a=[b\xi]+1$. Лемма доказана. Тем же способом можно доказать, что для любых чисел $\xi_1, \xi_2, \ldots, \xi_k$ и любого натурального числа N существует такое натуральное число b, что $b\leqslant N^k$ и числа $b\xi_1$, $b\xi_2$, ..., $b\xi_k$ отличаются от ближайших к ним целых чисел не более чем на $1/N$. Отсюда следует, что какой бы мы ни нарисовали многоугольник и какое бы малое положительное число ε ни взяли, можно подвергнуть многоугольник такой гомотетии с натуральным коэффициентом, что координаты вершин полученного многоугольника будут отличаться от целых чисел меньше чем на ε. ■

В силу леммы 8, для любого натурального $n>1$ существуют такие натуральные числа a_n и b_n, что $b_n<n$ и $|a_n-b_n\sqrt{d}|\leqslant\frac{1}{n}$. Очевидно,

$$|a_n^2-db_n^2|=|a_n-b_n\sqrt{d}|\cdot|a_n+b_n\sqrt{d}|\leqslant$$
$$\leqslant\frac{1}{n}\cdot|a_n-b_n\sqrt{d}+2b_n\sqrt{d}|\leqslant\frac{1}{n}\cdot\left(\frac{1}{n}+2n\sqrt{d}\right)<1+2\sqrt{d}.$$

Итак, величина $a_n^2-db_n^2$ может принимать лишь конечное число значений. Но n можно брать сколь угодно большим! И при этом в силу неравенства $|a_n-b_n\sqrt{d}|\leqslant\frac{1}{n}$ при $n\to\infty$ имеем $b_n\to\infty$. Значит, хотя бы для одного целого числа c, по модулю меньшего $1+2\sqrt{d}$, существует бесконечно много пар натуральных чисел $(a_n; b_n)$, для которых $a_n^2-db_n^2=c$.

Зафиксируем одно из таких чисел c. Рассмотрим остатки от деления чисел a_n и b_n на $|c|$. Поскольку количество остатков конечно, то существуют такие две разные пары натуральных чисел $(a; b)$ и $(A; B)$, что $a^2-db^2=c=A^2-dB^2$ и

$$a\equiv A,\quad b\equiv B\ (\mathrm{mod}\ |c|).$$

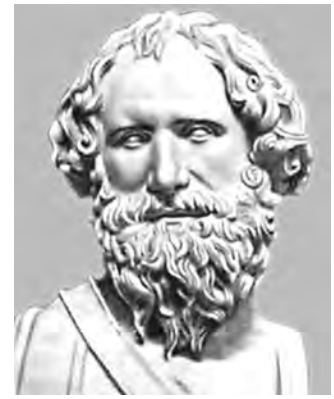
Рассмотрим частное $\frac{A+B\sqrt{d}}{a+b\sqrt{d}}=\frac{(a-b\sqrt{d})(A+B\sqrt{d})}{a^2-db^2}=\frac{aA-bBd+(aB-Ab)\sqrt{d}}{c}$. Поскольку $aA-bBd\equiv a^2-b^2d=c\equiv0\ (\mathrm{mod}\ |c|)$ и $aB-Ab\equiv ab-ab=0\ (\mathrm{mod}\ |c|)$, то числа $x=(aA-bBd)/c$ и $y=(aB-Ab)/c$ целые. Поскольку

$$x^2-dy^2=(x-y\sqrt{d})(x+y\sqrt{d})=\frac{A-B\sqrt{d}}{a-b\sqrt{d}}\cdot\frac{A+B\sqrt{d}}{a+b\sqrt{d}}=\frac{A^2-dB^2}{a^2-db^2}=\frac{c}{c}=1$$

и $y\neq0$, то $(x; y)$ — искомое нетривиальное решение уравнения Пелля! ■

Цепные дроби

Многие полагают: чтобы найти что-нибудь необыкновенное, надо отправиться очень далеко, лучше всего в космос. В обыденной жизни вокруг нас все хорошо известно, и ничего интересного нет. Какое заблуждение! Мы окружены загадочными явлениями, но в упор не замечаем большинство из них. Как известно, Архимед нашел для числа π приближенное значение 22/7. Почему он предпочел седьмые доли, а не восьмые или десятые? И почему високосные годы наступают раз в четыре года? На эти вопросы отвечают цепные дроби.



Один из античных бюстов, считавшихся изображением Архимеда.

Приблизить действительное число α дробью со знаменателем n — это значит из всех дробей со знаменателями n найти ближайшую к числу α. Если на числовой оси нанесены все дроби со знаменателем n, то число α лежит между какими-то двумя дробями, то есть для некоторого целого k имеем

$$\frac{k}{n} \leqslant \alpha < \frac{k+1}{n}.$$

Из этих двух дробей можно выбрать ту дробь $\frac{m}{n}$, которая ближе к α: если точка α ближе к левому концу отрезка $\left[\frac{k}{n}; \frac{k+1}{n}\right]$, разумно взять $m=k$; если ближе к правому концу, то $m=k+1$; если же α — середина отрезка, можно условиться выбирать $m=k$, хотя это несущественно.

Процесс замены числа α его приближенным значением называют *аппроксимацией*. Для аппроксимации можно использовать дроби с любым знаменателем. На практике чаще всего используют десятичные дроби. Однако во времена Архимеда их еще не изобрели, он мог выбрать любые доли. Он выбрал седьмые. Почему? Скоро мы в этом разберемся.

При аппроксимации действительного числа α дробью $\frac{m}{n}$ возникает абсолютная погрешность $\Delta = \left|\alpha - \frac{m}{n}\right|$. Она не превышает $\frac{1}{2n}$. (Если бы мы договорились всегда брать приближение с недостатком или всегда — с избытком, то верхняя граница абсолютной погрешности равнялась бы $1/n$.) Абсолютная погрешность достигает верхней границы $\frac{1}{2n}$, когда α — середина отрезка $[k/n; (k+1)/n]$.

Приближение «выгодное», если при не очень большом знаменателе n оно дает высокую точность. Чтобы охарактеризовать степень выгодности приближения, разумно разделить Δ на $\frac{1}{2n}$, то есть вычислить

$$\left|\alpha - \frac{m}{n}\right| : \frac{1}{2n} = 2|n\alpha - m|.$$

Принято рассматривать половину этой величины $h=|n\alpha - m|$. Назовем h качеством приближения. Очевидно, $h \leqslant \frac{1}{2}$. Если h близко к нулю, то число α близко к одному из концов отрезка $\left[\frac{k}{n}; \frac{k+1}{n}\right]$. Чем ближе h к $\frac{1}{2}$, тем ближе α к середине отрезка. ■

Архимед родился в Сиракузах — богатом торговом городе Сицилии. Отцом его был астроном Фидий, который с детства привил сыну любовь к математике, механике и астрономии. В Александрии Египетской — научном и культурном центре того времени — Архимед познакомился с астрономом Кононом и разносторонним ученым Эратосфеном, с которым переписывался до конца жизни. В то время Александрия славилась своей библиотекой, в которой было собрано более 700 000 рукописей. По-видимому, именно здесь Архимед познакомился с трудами Демокрита, Евдокса и других греческих геометров.

Уже при жизни Архимеда вокруг его имени создавались легенды, поводом для которых служили его поразительные изобретения. Известен рассказ о том, как Архимед сумел определить, сделана ли корона царя Гиерона из чистого золота или ювелир подмешал в нее серебро. Удельные веса золота и серебра были известны, но трудность состояла в том, чтобы найти объем короны: ведь она имела неправильную форму! Архимед все время размышлял над этой задачей. Как-то он принимал ванну, и ему в голову пришла идея: погружая корону в воду, можно определить ее объем, измерив объем вытесненной воды. Согласно легенде, Архимед выскочил на улицу с криком «Эврика!» («Нашел!»). Другая легенда рассказывает, что построенный Гиероном в подарок египетскому царю Птолемею корабль «Сиракосия» не удавалось спустить на воду. Архимед соорудил систему блоков (полиспаст), с помощью которой он проделал работу одним движением руки. Этот случай или размышления Архимеда над принципом

рычага послужили поводом для его слов: «Дайте мне точку опоры, и я переверну Землю».

Архимедов винт для вычерпывания воды до сих пор применяется на практике. Архимед построил планетарий («небесную сферу»), при помощи которого можно было наблюдать движения пяти планет, восходы Солнца и Луны, фазы и затмения Луны, заход обоих тел за линию горизонта.

Инженерный гений Архимеда с особой силой проявился во время осады Сиракуз римлянами в 212 г. А ведь ему было уже 75 лет! Построенные им мощные метательные машины забрасывали римские войска тяжелыми камнями. Думая, что они будут в безопасности у самых стен города, римляне кинулись туда, но в это время легкие метательные машины близкого действия забросали их градом ядер. Мощные краны захватывали железными крюками корабли, приподнимали их и отпускали, отчего они переворачивались и тонули.

Римляне отказались от штурма и перешли к осаде. Историк Полибий писал: «Такова чудесная сила одного человека, одного дарования, умело направленного на какое-либо дело... римляне могли бы быстро овладеть городом, если бы кто-либо изъял из среды сиракузян одного старца». Осенью 212 г. Сиракузы пали вследствие измены. Плутарх рассказывает: «К Архимеду подошел солдат с мечом в руке, чтобы убить его. Но Архимед настойчиво просил подождать одну минуту, чтобы задача, которой он занимался, не осталась нерешенной; солдат, которому не было дела до его доказательства, пронзил его своим мечом».

Архимед рассмотрел последовательность вписанных в окружность и описанных вокруг нее 6-, 12-, 24-, 48- и 96-угольников. Вычислив отношения их периметров к диаметру, он доказал неравенства $3\frac{10}{71}<\pi<3\frac{1}{7}$.

Высшим своим достижением Архимед считал вычисление площади сферы и объема шара. Он завещал выбить на своей могиле шар, вписанный в цилиндр.

Один из создателей дифференциального и интегрального исчислений Г. В. Лейбниц писал: «Внимательно читая Архимеда, перестаешь удивляться всем новейшим открытиям геометров». ■

Эксперимент с числом π. Не следует думать, что чем больше n, тем меньше h. Сделаем опыт с числом π, аппроксимируя его разными дробями:

m/n	3/1	6/2	9/3	13/4	16/5	19/6
Δ	0,1416	0,1416	0,1416	0,1084	0,0584	0,0251
h	0,1416	0,2832	0,4248	0,4336	0,2920	0,1504
m/n	**22/7**	25/8	28/9	31/10	35/11	314/100
Δ	**0,0013**	0,0166	0,0305	0,0416	0,0402	0,0016
h	**0,0089**	0,1327	0,2743	0,4159	0,4424	0,1593

Седьмые доли гораздо выгоднее ближайших соседей. Если бы нам приказали приблизить π, чтобы абсолютная погрешность не превзошла 0,0013, какое n выбрали бы мы? Записав условие $\frac{1}{2n}\leqslant 0{,}0013$, получили бы $n\geqslant 385$, а Архимед достиг той же точности, взяв гораздо меньший знаменатель. Теперь вы убедились, что Архимед выбрал седьмые доли не случайно?

Голландец А. Меций предлагал приближенное значение $\pi\approx 355/113$. Число Меция обладает тем же свойством, что и число Архимеда: знаменатель 113 выгоднее, чем меньшие знаменатели. ■

Преимущества десятичной системы не математические, а зоологические. Если бы у нас на руках было не десять пальцев, а восемь, то человечество пользовалось бы восьмеричной системой. Поэтому откажемся от десятичных дробей и рассмотрим не зависящий от количества пальцев на руках способ приближенного представления чисел — цепные дроби.

Разложить данное число α в цепную дробь — это значит прежде всего выделить его целую часть, то есть представить его в виде $\alpha=[\alpha]+\{\alpha\}$, где $[\alpha]$ — такое целое число, что $[\alpha]\leqslant\alpha<[\alpha]+1$. Обозначаем: $a_0=[\alpha]$. Если α — целое число, то $\{\alpha\}=0$, и процесс разложения в цепную дробь на этом обрывается. Если же $\{\alpha\}>0$, то число α можно представить в виде $\alpha=a_0+\frac{1}{\alpha_1}$, где $\alpha_1>1$. Записав $\alpha_1=[\alpha_1]+\{\alpha_1\}$, находим следующее неполное частное $a_1=[\alpha_1]$. Если $\{\alpha_1\}=0$, то разложение получено. Если же $\{\alpha_1\}>0$, то $\alpha_1=a_1+\frac{1}{\alpha_2}$, где $\alpha_2>1$. И так далее, пока очередное число не окажется целым или — до бесконечности (точнее, пока не наступит конец света).

Если исходное число α иррационально, то и α_1, и α_2, и все возникающие далее такие числа иррациональны, так что процесс разложения в цепную дробь никогда не остановится и даст бесконечную последовательность $a_0, a_1, a_2, \ldots$ элементов цепной дроби — так называемых неполных частных.

Процесс разложения любого рационального числа в цепную дробь заканчивается, поскольку знаменатель разлагаемого числа все время уменьшается. Для любого нецелого рационального числа изложенная конструкция приводит к представлению, последнее неполное частное которого больше 1. ■

Обратите внимание: $\dfrac{1}{6+\frac{1}{4}}=\dfrac{1}{6+\frac{1}{3+\frac{1}{1}}}$ или в сокращенных обозначениях $[0; 6, 4]=[0; 6, 3, 1]$. Такое преобразование (отделение единицы от последнего элемента) можно произвести с любой конечной цепной дробью, последний элемент которой отличен от единицы. Если же последний элемент равен единице, то его, наоборот, можно прибавить к предпоследнему. Например, $[8; 10, 3, 6, 1]=[8; 10, 3, 7]$. Легко доказать, что это — единственная причина неоднозначного представления рационального числа цепной дробью. ■

Подходящие дроби. Цепную дробь можно оборвать, оставив элементы $a_0, a_1, \ldots, a_n$ и отбросив все остальные. Полученное таким образом число $[a_0; a_1, \ldots, a_n]$ называют n-й подходящей дробью. В частности, при $n=0$ имеем нулевую подходящую дробь $a_0/1$.

Пример. $\dfrac{61}{27}=2+\dfrac{7}{27}=2+\dfrac{1}{27/7}=2+\dfrac{1}{3+\frac{6}{7}}=2+\dfrac{1}{3+\frac{1}{7/6}}=2+\dfrac{1}{3+\frac{1}{1+\frac{1}{6}}}$.

Следовательно, подходящие дроби таковы: $\dfrac{p_0}{q_0}=\dfrac{2}{1}$, $\dfrac{p_1}{q_1}=2+\dfrac{1}{3}=\dfrac{7}{3}$, $\dfrac{p_2}{q_2}=[2; 3, 1]=2+\dfrac{1}{[3; 1]}=\dfrac{9}{4}$, наконец, $\dfrac{p_3}{q_3}=[2; 3, 1, 6]=\dfrac{61}{27}$. Обратите внимание: $\dfrac{2}{1}<\dfrac{9}{4}<\dfrac{61}{27}<\dfrac{7}{3}$. ■

Следующий пример — число

$$\pi=3+0{,}14159265\ldots=3+\frac{1}{7+0{,}00885145\ldots}.$$

Очевидно, $\dfrac{p_0}{q_0}=3$ и $\dfrac{p_1}{q_1}=3\dfrac{1}{7}$. Продолжив вычисления, можно найти $\pi=[3; 7, 15, 1, 288, \ldots]$, так что $p_2/q_2=333/106$ и $p_3/q_3=355/1133$.

Можно ли считать, что Архимед и Меций разоблачены: они использовали первую и третью подходящие дроби? Нет, во всяком случае про Архимеда этого утверждать нельзя. Мы решили математическую, но не историческую задачу. Скорее всего, Архимед использовал цепные дроби, но доказать это по дошедшим до нас работам нельзя; историки не пришли к единому мнению. Преимущество дроби 22/7 нетрудно обнаружить и перебором.

Другое дело Меций. Очень трудно (но все равно можно!) дробь 335/113 найти без теории. Вероятно, Меций пользовался цепными дробями. Понятно, почему он остановился на этой дроби: следующие слишком громоздки, чтобы их можно было практически использовать. ■

Цепная дробь числа $\sqrt{2}$. Очевидно, $(\sqrt{2}-1)(\sqrt{2}+1)=2-1=1$ и, следовательно, $\sqrt{2}-1=\dfrac{1}{1+\sqrt{2}}$. Воспользуемся этой формулой многократно:

$$\sqrt{2}=1+(\sqrt{2}-1)=1+\frac{1}{1+\sqrt{2}}=1+\frac{1}{2+(\sqrt{2}-1)}=1+\frac{1}{2+\frac{1}{1+\sqrt{2}}}=[1; 2, 2, 2, \ldots].$$

Подходящие дроби 1/1, $[1; 2]=3/2$, $[1; 2, 2]=7/5$, $[1; 2, 2, 2]=17/12$ знакомы нам по статье «Уравнения Пелля». Может быть, это случайное совпадение? Нет, если n-этажная дробь (в которой n двоек) приводится к несократимому виду x/y, то $(n+1)$-этажная дробь равна

$$1+\frac{1}{1+\frac{x}{y}}=1+\frac{y}{x+y}=\frac{x+2y}{x+y}.$$

Очевидно, $\text{НОД}(x+2y, x+y)=\text{НОД}(y, x+y)=\text{НОД}(x, y)$, так что дробь $\dfrac{x+2y}{x+y}$ тоже несократима. Поэтому увеличение количества дробных черт на единицу — это переход от несократимой дроби $\dfrac{x}{y}$ к несократимой дроби $\dfrac{x+2y}{x+y}$. А это и есть формулы из статьи «Уравнения Пелля»!

Сутки — это период обращения Земли вокруг своей оси. Год — период обращения Земли вокруг Солнца — равен $365^{\text{d}}\ 5^{\text{h}}\ 48^{\text{m}}\ 46^{\text{s}}$ (то есть 365 суток 5 часов 48 минут 46 секунд). Узаконить в обыденной жизни такую длину года невозможно. Если считать, что год — это 365^{d}, за четыре года отставание составит почти сутки. С зимы 1 января постепенно сместится на осень, а потом и на лето. Периодические мероприятия (посев, начало учебного года) нельзя будет связывать с определенными календарными датами.

Выход известен: некоторые годы состоят из 365^{d}, а некоторые — високосные — из 366^{d}. Первым такую систему придумал для Юлия Цезаря александрийский астроном Созиген: каждый четвертый год — високосный. В христианском летоисчислении високосные годы — те, номера которых делятся на 4. Средняя длина юлианского года равна $365^{\text{d}}\ 6^{\text{h}}$, что больше истинной на $11^{\text{m}}\ 14^{\text{s}}$. В 1582 г. папа Григорий XIII дополнил закон чередования обычных и високосных лет правилом: если номер года оканчивается двумя нулями, а число сотен не делится на 4, то год обычный (например, 2000 г. — високосный, а 1900 г. — обычный). Кроме того, считая, что от начала летоисчисления (от «рождества Христова») уже накопилась ошибка в 10^{d}, Григорий XIII сразу прибавил 10^{d}. С тех пор накопились еще 3^{d} (в 1700, 1800 и 1900 гг.). Поэтому сейчас расхождение между юлианским и григорианским календарями составляет 13^{d}.

Средняя длина григорианского года равна $\left(365\dfrac{97}{400}\right)^{\text{d}}=365^{\text{d}}\ 5^{\text{h}}\ 49^{\text{m}}\ 12^{\text{s}}$, что больше истинной на 26^{s}. Весьма простыми средствами достигнута хорошая точность.

В России до 1917 г. пользовались юлианским календарем. Григорианский был введен декретом Совета народных комиссаров в 1918 г.

Длительность года измеряют астрономы и физики. Поэтому говорить о ее рациональности или иррациональности бессмысленно. Для наших целей можно считать, что год длится в точности $365^{\text{d}}\ 5^{\text{h}}\ 48^{\text{m}}\ 46^{\text{s}}$. Рассмотрим цепную дробь: $[365; 4, 7, 1, 3, 5, 20, 6, 12]$. Каждая из подходящих дробей $365, 365\dfrac{1}{4}, 365\dfrac{7}{29}, 365\dfrac{8}{33}, 365\dfrac{31}{128}$

решает проблему календаря. Например, $365\frac{1}{4}$ соответствует юлианскому календарю. Пользоваться приближением $365\frac{7}{29}$ никто не предлагал (следующее приближение $365\frac{8}{33}$ немного сложнее, но значительно точнее). Календарь, по которому високосны восемь лет из каждых тридцати трех, предлагал О. Хайям (1040—1123).

Приближение	Средняя продолжительность года	Погрешность
1/4	$365^d\,6^h\,0^m\,0^s$	$-11^m\,14^s$
7/29	$365^d\,5^h\,47^m\,35^s$	$1^m\,11^s$
8/33	$365^d\,5^h\,49^m\,5^s$	-19^s
31/128	$365^d\,5^h\,48^m\,45^s$	1^s

Четвертый вариант исключительно точен. В 1864 г. астроном Медлер предлагал в юлианском календаре каждые 128 лет пропускать один високосный год (ибо по юлианскому календарю на 128 лет приходится 32, а не 31 високосных).

Средняя длина григорианского года отличается от истинной на 26^s. Выходит, Григорий XIII изобрел календарь более сложный и менее точный, чем хайямовский? Его советники были плохими математиками?

Нет. Тогда продолжительность года была известна не столь точно, как сейчас. Комиссия Григория XIII пользовалась астрономическими таблицами, составленными королем Кастилии Альфонсом X (1221—1284). В них дана следующая продолжительность года: $365^d\,5^h\,49^m\,16^s$.

На основании этих таблиц комиссия считала, что предложенная ей средняя длина года на 4^s отличается от истинной. Если бы она была знакома с предложением Хайяма, то заключила бы, что его календарь дает ошибку в 11^s.

Нет оснований предполагать, что комиссия Григория XIII использовала цепные дроби. Она кропотливо подбирала соотношение обычных и високосных лет. Историки думают, что Хайям владел чем-то вроде цепных дробей: в его эпоху восточная наука во многих отношениях стояла выше европейской. ■

Подходящие дроби 1/1, 3/2, 7/5, 17/12, . . . замечательны тем, что дают (попеременно, слева и справа) весьма точные приближения числа $\sqrt{2}$. А именно,

$$\frac{1}{1}<\frac{7}{5}<\frac{41}{29}<\frac{239}{169}<\ldots<\sqrt{2}<\ldots<\frac{577}{408}<\frac{99}{70}<\frac{17}{12}<\frac{3}{2}.$$

Оценить погрешность приближения несложно:

$$\left|\frac{x}{y}-\sqrt{2}\right|=\left|\frac{(x-y\sqrt{2})(x+y\sqrt{2})}{y(x+y\sqrt{2})}\right|=\left|\frac{x^2-2y^2}{y^2\left(\frac{x}{y}+\sqrt{2}\right)}\right|=\frac{1}{y^2\left(\frac{x}{y}+\sqrt{2}\right)}.$$

Например,

$$0<\frac{17}{12}-\sqrt{2}=\frac{1}{12^2\left(\sqrt{2}+\frac{17}{12}\right)}<\frac{1}{12^2\cdot 2\sqrt{2}}<0{,}0025,$$

$$0<\sqrt{2}-\frac{41}{29}=\frac{1}{29^2\left(\sqrt{2}+\frac{41}{29}\right)}<\frac{1}{12^2\cdot 2\cdot\frac{41}{29}}<0{,}00042.\ \blacksquare$$

Аналогично числу $\sqrt{2}$, разложим в цепную дробь число $\sqrt{3}$. Очевидно,

$$\sqrt{3}=1+(\sqrt{3}-1)=1+\frac{1}{(\sqrt{3}+1)/2}=1+\frac{1}{1+\frac{\sqrt{3}-1}{2}}=1+\frac{1}{1+\frac{1}{\sqrt{3}+1}}=1+\frac{1}{1+\frac{1}{2+(\sqrt{3}-1)}}=$$

$=[1;\,1,\,2,\,1,\,2,\,1,\,2,\,\ldots]$. Подходящие дроби: $[1]=\frac{1}{1}$, $[1;\,1]=\frac{2}{1}$, $[1;\,1,\,2]=\frac{5}{3}$, $[1;\,1,\,2,\,1]=\frac{7}{4}$, $[1;\,1,\,2,\,1,\,2]=\frac{19}{11}$, $[1;\,1,\,2,\,1,\,2,\,1]=\frac{26}{15}$. Заметьте: $1^2-3\cdot 1^2=-2$, $2^2-3\cdot 1^2=1$, $5^2-3\cdot 3^2=-2$, $7^2-3\cdot 4^2=1$, $19^2-3\cdot 11^2=-2$, $26^2-3\cdot 15^2=1$, так что половина дробей «лишние» — они дают решения не уравнения $x^2-3y^2=1$, а уравнения $x^2-3y^2=-2$. (Подходящие дроби числа $\sqrt{2}$ обладают аналогичным свойством.) Мы вскоре докажем, что если x/y — подходящая дробь числа ξ, то $\left|\xi-\frac{x}{y}\right|<\frac{1}{y^2}$, так что цепные дроби дают конструктивное доказательство леммы 8 статьи «Уравнения Пелля». Кроме того, если $x^2-dy^2=1$ и x, y — натуральные числа, то x/y — подходящая дробь числа $\sqrt{d}$ (значит, для поиска решения $(x;\,y)$ уравнения Пелля следует перебирать лишь подходящие дроби числа $\sqrt{d}$). А вот обратное утверждение, как видно на примерах $d=2$ и $d=3$, ложно: не каждая подходящая дробь соответствует решению уравнения Пелля. ■

Примеры убедили вас в полезности цепных дробей. Поэтому мы займемся их систематическим изучением. Рассмотрим иррациональное число α, разложим его в цепную дробь $\alpha=[a_0;\,a_1,\,a_2,\,a_3,\,\ldots]$ и образуем одну за другой подходящие дроби $\frac{p_0}{q_0}=\frac{a_0}{1}$, $\frac{p_1}{q_1}=a_0+\frac{1}{a_1}=\frac{a_0a_1+1}{a_1}$ и $\frac{p_2}{q_2}=a_0+\frac{1}{[a_1;\,a_2]}=a_0+$

$+\frac{a_2}{a_1a_2+1}=\frac{a_0a_1a_2+a_0+a_2}{a_1a_2+1}$. Далее, $\frac{p_3}{q_3}=a_0+\frac{1}{[a_1;\,a_2,\,a_3]}=a_0+\frac{a_2a_3+1}{a_1a_2a_3+a_1+a_3}=$

$=\frac{a_0a_1a_2a_3+a_0a_1+a_0a_3+a_2a_3+1}{a_1a_2a_3+a_1+a_3}$.

Обозначим $p_n=[a_0,\,a_1,\,\ldots,\,a_n]$ и $[\,]=1$. Очевидно, $q_n=[a_1,\,\ldots,\,a_n]$ и

$$[a_0;\,a_1,\,a_2,\,\ldots,\,a_n]=\frac{p_n}{q_n}=\frac{[a_0,\,a_1,\,\ldots,\,a_n]}{[a_1,\,\ldots,\,a_n]}.$$

n	0	1	2	3
p_n	a_0	a_0a_1+1	$a_0a_1a_2+a_0+a_2$	$a_0a_1a_2a_3+a_0a_1+a_0a_3+a_2a_3+1$
q_n	1	a_1	a_1a_2+1	$a_1a_2a_3+a_1+a_3$

Для любого иррационального числа α рассмотрим решетку, порожденную векторами $(1;\alpha)$ и $(0;-1)$, то есть множество точек вида $(n;n\alpha-m)$, где n, m — целые числа.

Теорема. ***Для любых натуральных чисел N и k существует не менее k таких точек $(n;m)$ с целыми координатами, что $1\leqslant n\leqslant Nk$ и $|n\alpha-m|<1/N$.***

Доказательство. Внутри прямоугольника, заданного неравенствами $1\leqslant x\leqslant kN$ и $-\frac{1}{2}\leqslant y\leqslant\frac{1}{2}$, лежат Nk точек решетки, поскольку каждая вертикальная прямая пересекает прямоугольник по отрезку длины 1. Горизонтальными прямыми разрежем прямоугольник на N равных полосок p_1, p_2, …, p_N — прямоугольников высоты $1/N$. Рассмотрим прямоугольник P, заданный неравенствами $|y|<1/N$ и $1\leqslant x\leqslant kN$. Он содержит одну среднюю полоску $p_{(N+1)/2}$, если N нечетно, и состоит из полосок $p_{N/2}$ и $p_{(N+2)/2}$, если N четно.

Если в каждой из полосок p_1, p_2, …, p_N лежит ровно k целочисленных точек, то в $p_{\lfloor N/2\rfloor+1}\subset P$ тоже содержатся k точек, что и требовалось доказать. Если же точки распределены по полоскам не поровну, то хотя бы в одной полоске p_r лежат не менее $k+1$ точек. Обозначим их B_0, B_1, …, …, B_k, … причем пусть B_0 — самая левая из них. Сдвинем точки B_1, B_2, …, B_k на вектор $\overrightarrow{B_0O}$, где O — начало координат. Получим искомые точки A_1, A_2, …, …, A_k, содержащиеся в прямоугольнике P.

Следствие. ***Существует бесконечно много таких дробей m/n, что***

$$\left|\alpha-\frac{m}{n}\right|<\frac{1}{n^2}.$$

Доказательство. Применив теорему при $k=1$, получаем такую дробь m/n, что $|n\alpha-m|<\frac{1}{N}\leqslant\frac{1}{n}$.

При $N\to\infty$ величина $1/N$ стремится к нулю, так что множество таких дробей бесконечно.

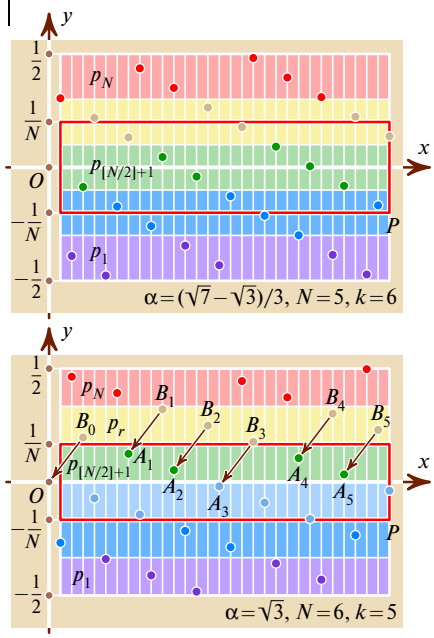
Очевидно, $n\cdot|n\alpha-m|$ — площадь прямоугольника с противоположными вершинами $(n;n\alpha-m)$ и $(0;0)$. Теорема Гурвица—Бореля утверждает, что из трех последовательных подходящих дробей хотя бы одна соответствует точке решетки, лежащей в области $|xy|<1/\sqrt{5}$. ■

Из равенства $[a_0;a_1,a_2,\ldots,a_n]=a_0+\dfrac{1}{[a_1;a_2,\ldots,a_n]}=a_0+\dfrac{[a_2,\ldots,a_n]}{[a_1,a_2,\ldots,a_n]}=$
$=\dfrac{a_0[a_1,a_2,\ldots,a_n]+[a_2,\ldots,a_n]}{[a_1,a_2,\ldots,a_n]}$ получаем формулу

$$[a_0,a_1,a_2,\ldots,a_n]=a_0[a_1,a_2,\ldots,a_n]+[a_2,\ldots,a_n]. \qquad (*)$$

Это рекуррентное соотношение позволяет по очереди вычислять числители и знаменатели подходящих дробей. Если присмотреться, однако, можно обнаружить явную формулу — *правило Эйлера*. Рассмотрим произведение $a_0a_1\ldots a_n$, затем — всевозможные произведения, которые можно получить, вычеркнув пару рядом стоящих букв, затем — произведения, получаемые вычеркиванием двух пар рядом стоящих букв и так далее. Сумма всех таких произведений и равна $[a_0,a_1,\ldots,a_n]$. (Если $n+1$ четно, на последнем шаге отбрасыванием всех элементов получаем произведение нуля множителей; оно по определению равно 1.)

Суть доказательства правила Эйлера в том, что выражение $[a_2,\ldots,a_n]$ состоит из тех слагаемых суммы $[a_0,a_1,a_2,\ldots,a_n]$, в которых вычеркнута пара a_0a_1, а произведение $a_0[a_1,a_2,\ldots,a_n]$ — из тех, где эта пара не вычеркнута. ■



На этих рисунках масштаб по оси Ox в 20 раз меньше масштаба по оси Oy.

Из правила Эйлера следует, что величина $[a_0,a_1,\ldots,a_n]$ не меняется, если записать числа в обратном порядке: $[a_0,a_1,\ldots,a_n]=[a_n,\ldots,a_1,a_0]$.

Теорема 1. ***$p_n=a_np_{n-1}+p_{n-2}$ и $q_n=a_nq_{n-1}+q_{n-2}$. (Чтобы эти равенства были верны и для $n=1$, считаем по определению $p_{-1}=1$ и $q_{-1}=0$.)***

Доказательство. В силу $(*)$ имеем: $p_n=[a_0,\ldots,a_{n-1},a_n]=[a_n,a_{n-1},\ldots,a_0]=$
$=a_n[a_{n-1},\ldots,a_0]+[a_{n-2},\ldots,a_0]=a_n[a_0,\ldots,a_{n-1}]+[a_0,\ldots,a_{n-2}]=a_np_{n-1}+p_{n-2}$.
Доказательство равенства $q_n=a_nq_{n-1}+q_{n-2}$ аналогично.

Теорема 2. ***$p_{n-1}q_n-p_nq_{n-1}=(-1)^n$.***

Доказательство — индукция по n. **База.** $p_0q_1-p_1q_0=a_0a_1-(a_0a_1+1)\cdot1=-1$ (или, если угодно, $p_{-1}q_0-p_0q_{-1}=1\cdot1-a_0\cdot0=1$).

Переход. Пусть для некоторого n равенство верно. Тогда $p_nq_{n+1}-p_{n+1}q_n=$
$=p_n(a_{n+1}q_n+q_{n-1})-(a_{n+1}p_n+p_{n-1})q_n=p_nq_{n-1}-p_{n-1}q_n=-(-1)^n=(-1)^{n+1}$.

Теорема 3. $\dfrac{p_{n-1}}{q_{n-1}}-\dfrac{p_n}{q_n}=\dfrac{(-1)^n}{q_nq_{n-1}}$ ***и*** $\dfrac{p_n}{q_n}-\dfrac{p_{n-2}}{q_{n-2}}=\dfrac{(-1)^na_n}{q_nq_{n-2}}$.

Доказательство. Первое равенство прямо следует из теоремы 2, а второе получаем при помощи теоремы 1 естественным вычислением: $p_nq_{n-2}-p_{n-2}q_n=$
$=(a_np_{n-1}+p_{n-2})q_{n-2}-p_{n-2}(a_nq_{n-1}+q_{n-2})=a_np_{n-1}q_{n-2}-p_{n-2}a_nq_{n-1}=a_n(-1)^n$. ■

Подходящие дроби $p_0/q_0,p_2/q_2,p_4/q_4,\ldots$ иррационального числа α в силу теоремы 3 образуют возрастающую последовательность, а подходящие дроби $p_1/q_1,p_3/q_3,p_5/q_5,\ldots$ — убывающую. Значение цепной дроби заключено между этими двумя последовательностями:

$$\frac{p_0}{q_0}<\frac{p_2}{q_2}<\frac{p_4}{q_4}<\ldots<\alpha<\ldots<\frac{p_5}{q_5}<\frac{p_3}{q_3}<\frac{p_1}{q_1}.$$

Таким образом, α лежит любыми двумя последовательными подходящими дробями и, следовательно,

$$\left|\alpha-\frac{p_n}{q_n}\right|<\left|\frac{p_{n+1}}{q_{n+1}}-\frac{p_n}{q_n}\right|=\frac{1}{q_{n+1}q_n}.$$

Поскольку $q_{n+1}=a_{n+1}q_n+q_{n-1}>a_{n+1}q_n$, то $\left|\alpha-\frac{p_n}{q_n}\right|<\frac{1}{a_{n+1}q_n^2}\leqslant\frac{1}{q_n^2}$.

Далее, для любого n дробь p_{n+2}/q_{n+2} лежит между α и p_n/q_n, следовательно,

$$\left|\alpha-\frac{p_n}{q_n}\right|>\left|\frac{p_{n+2}}{q_{n+2}}-\frac{p_n}{q_n}\right|=\frac{|p_{n+2}q_n-p_nq_{n+2}|}{q_{n+2}q_n}=\frac{a_{n+2}}{q_{n+2}q_n}\geqslant\frac{1}{(q_{n+1}+q_n)q_n},$$

поскольку $a_{n+2}(q_{n+1}+q_n)\geqslant a_{n+2}q_{n+1}+q_n=q_{n+2}$. ■

Сравним полученные неравенства $|q_n\alpha-p_n|>\frac{1}{q_{n+1}+q_n}$ и $|q_{n+1}\alpha-p_{n+1}|<\frac{1}{q_{n+2}}\leqslant\frac{1}{q_{n+1}+q_n}$. Следовательно, с ростом номера подходящей дроби величина h уменьшается. Приближения становятся все качественнее!

Теорема 4. ***Если p, q — целые числа, причем $1\leqslant q<q_n$, то $|q\alpha-p|>|p_n\alpha-q_n|$ и, следовательно,*** $\left|\alpha-\frac{p}{q}\right|=\frac{|q\alpha-p|}{q}>\frac{|q_n\alpha-p_n|}{q_n}=\left|\alpha-\frac{p_n}{q_n}\right|$.

Доказательство. Дробь p/q расположена либо левее числа a_0, и тогда

$$b\left|\alpha-\frac{a}{b}\right|>b|\alpha-a_0|\geqslant|1\cdot\alpha-a_0|=|q_0\alpha-p_0|\geqslant|q_n\alpha-p_n|;$$

либо a/b расположена между двумя промежуточными дробями p_{k-1} и p_{k+1} с номерами одной четности, и тогда имеем

$$\left|\frac{a}{b}-\frac{p_{k-1}}{q_{k-1}}\right|=\frac{|aq_{k-1}-bp_{k-1}|}{bq_{k-1}}\geqslant\frac{1}{bq_{k-1}},$$

$$\left|\frac{a}{b}-\frac{p_{k-1}}{q_{k-1}}\right|<\left|\frac{p_k}{q_k}-\frac{p_{k-1}}{q_{k-1}}\right|=\frac{1}{q_kq_{k-1}},$$

откуда $b>q_k$ и поэтому $n>k$, и

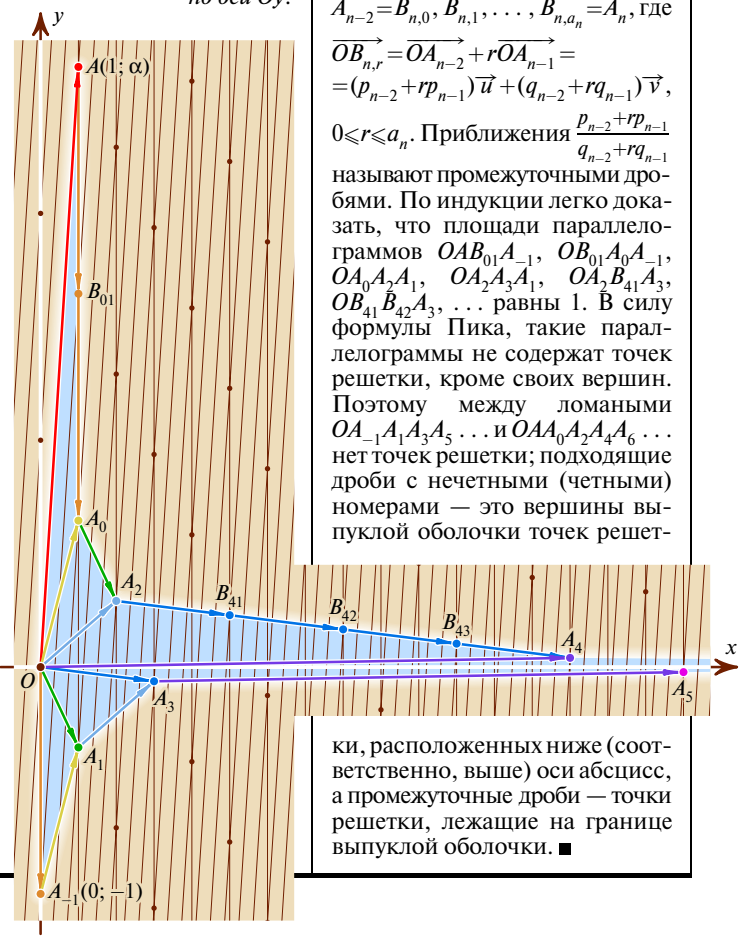
$$b\left|\alpha-\frac{a}{b}\right|\geqslant b\left|\frac{a}{b}-\frac{p_{k+1}}{q_{k+1}}\right|=\frac{|aq_{k+1}-bp_{k+1}|}{q_{k+1}}\geqslant\frac{1}{q_{k+1}}\geqslant$$
$$\geqslant|q_k\alpha-p_k|>|q_n\alpha-p_n|;$$

либо дробь a/b расположена правее числа p_1/q_1, и тогда

$$b\left|\alpha-\frac{a}{b}\right|>b\left|\frac{p_1}{q_1}-\frac{a}{b}\right|=\frac{|p_1b-q_1a|}{q_1}\geqslant\frac{1}{q_1}>$$
$$>|1\cdot\alpha-a_0|=|q_0\alpha-p_0|\geqslant|q_n\alpha-p_n|.$$

Таким образом, получив, например, для числа $\sqrt{2}$ подходящую дробь $\frac{p_3}{q_3}=[1;2,2,2]=\frac{17}{12}$, мы можем быть уверены, что точнее приблизить число $\sqrt{2}$ дробью со знаменателем, меньшим 12, невозможно. ■

На рисунке $\alpha=\sqrt{7}=[2;1,1,1,4,\ldots]$; масштаб по оси Ox в 6 раз меньше масштаба по оси Oy.



Алгоритм «вытягивания носа». Рассмотрим векторы $\overrightarrow{u}=(0;-1)$ и $\overrightarrow{v}=(1;\alpha)$. Пусть для определенности $\alpha>1$ и $\alpha\notin\mathbb{Q}$. Из точки $A(1;\alpha)$ отложим вектор $\overrightarrow{u}$ столь много раз, сколько его можно отложить, не пересекая ось абсцисс, то есть a_0 раз. Пришли в точку A_0, где $\overrightarrow{OA_0}=\overrightarrow{v}+a_0\overrightarrow{u}$. Из точки $A_{-1}(0;-1)$ отложим вектор $\overrightarrow{OA_0}$ столь много раз, сколько его можно отложить, не пересекая ось абсцисс, то есть a_1 раз (фактически мы делим с остатком число 1 на $\{\alpha\}$). Пришли в точку A_1, где $\overrightarrow{OA_1}=\overrightarrow{u}+a_1\overrightarrow{OA_0}=(1+a_1a_0)\overrightarrow{u}+a_1\overrightarrow{v}$.

Вообще, выполняя шаги алгоритма Евклида для чисел α и 1, будем рисовать векторы $\overrightarrow{OA_n}=p_n\overrightarrow{u}+q_n\overrightarrow{v}$. Шаг алгоритма — откладывание от точки A_{n-1} вектора $\overrightarrow{OA_{n-2}}$ столь много раз, сколько можно, не пересекая ось абсцисс, то есть a_n раз. Так получаем вектор

$$\overrightarrow{OA_n}=\overrightarrow{OA_{n-2}}+a_n\overrightarrow{OA_{n-1}}=p_{n-2}\overrightarrow{u}+q_{n-2}\overrightarrow{v}+a_n(p_{n-1}\overrightarrow{u}+q_{n-1}\overrightarrow{v})=$$
$$=(p_{n-2}+a_np_{n-1})\overrightarrow{u}+(q_{n-2}+a_nq_{n-1})\overrightarrow{v}=p_n\overrightarrow{u}+q_n\overrightarrow{v}.$$

На отрезке $A_{n-2}A_n$ лежат точки $A_{n-2}=B_{n,0}, B_{n,1},\ldots,B_{n,a_n}=A_n$, где $\overrightarrow{OB_{n,r}}=\overrightarrow{OA_{n-2}}+r\overrightarrow{OA_{n-1}}=(p_{n-2}+rp_{n-1})\overrightarrow{u}+(q_{n-2}+rq_{n-1})\overrightarrow{v}$, $0\leqslant r\leqslant a_n$. Приближения $\frac{p_{n-2}+rp_{n-1}}{q_{n-2}+rq_{n-1}}$ называют промежуточными дробями. По индукции легко доказать, что площади параллелограммов $OAB_{01}A_{-1}$, $OB_{01}A_0A_{-1}$, $OA_0A_2A_1$, $OA_2A_3A_1$, $OA_2B_{41}A_3$, $OB_{41}B_{42}A_3$, … равны 1. В силу формулы Пика, такие параллелограммы не содержат точек решетки, кроме своих вершин. Поэтому между ломаными $OA_{-1}A_1A_3A_5\ldots$ и $OAA_0A_2A_4A_6\ldots$ нет точек решетки; подходящие дроби с нечетными (четными) номерами — это вершины выпуклой оболочки точек решетки, расположенных ниже (соответственно, выше) оси абсцисс, а промежуточные дроби — точки решетки, лежащие на границе выпуклой оболочки. ■

Адольф Гурвиц (1859—1919) — немецкий математик, ученик Г. Шуберта и Ф. Клейна, профессор Политехнической школы в Цюрихе.
Критерий Рауса—Гурвица по некоторым определителям, составленным из коэффициентов многочлена, определяет, является или нет этот многочлен устойчивым, то есть верно ли, что вещественные части всех его корней отрицательны. (Устойчивые многочлены возникают при изучении устойчивости колебательных систем.) ■

Подходящие дроби, как мы только что доказали, являются наилучшими приближениями. Насколько точны эти приближения?

Теорема 5. ***Для любого натурального n верно хотя бы одно из неравенств***

$$\left|\alpha-\frac{p_{n-1}}{q_{n-1}}\right|<\frac{1}{2q_{n-1}^2} \quad \textit{и} \quad \left|\alpha-\frac{p_n}{q_n}\right|<\frac{1}{2q_n^2}.$$

Доказательство. Поскольку число α расположено между p_{n-1}/q_{n-1} и p_n/q_n, то

$$\left|\frac{p_{n-1}}{q_{n-1}}-\alpha\right|+\left|\alpha-\frac{p_n}{q_n}\right|=\left|\frac{p_{n-1}}{q_{n-1}}-\frac{p_n}{q_n}\right|=\frac{1}{q_nq_{n-1}}<\frac{1}{2q_n^2}+\frac{1}{2q_{n-1}^2}$$

(последнее неравенство выражает собой тот факт, что среднее геометрическое величин $1/q_n^2$ и $1/q_{n-1}^2$ меньше их среднего арифметического; равенство могло бы выполняться лишь при $q_{n-1}=q_n$, что в данном случае невозможно). Отсюда, очевидно, непосредственно следует утверждение теоремы 5.

Теорема 6 (Гурвиц и Борель). ***Для любого натурального n верно хотя бы одно из неравенств***

$$\left|\alpha-\frac{p_{n-1}}{q_{n-1}}\right|<\frac{1}{q_{n-1}^2\sqrt{5}}, \quad \left|\alpha-\frac{p_n}{q_n}\right|<\frac{1}{q_n^2\sqrt{5}} \quad \textit{и} \quad \left|\alpha-\frac{p_{n+1}}{q_{n+1}}\right|<\frac{1}{q_{n+1}^2\sqrt{5}}.$$

Доказательство. Для любого натурального k имеем

$$\alpha=[a_0;a_1,a_2,\ldots,a_k,\alpha_{k+1}]=\frac{\alpha_{k+1}p_k+p_{k-1}}{\alpha_{k+1}q_k+q_{k-1}}.$$

(Эта формула еще не раз пригодится нам.) Значит,

$$\left|\alpha-\frac{p_k}{q_k}\right|=\left|\frac{\alpha_{k+1}p_k+p_{k-1}}{\alpha_{k+1}q_k+q_{k-1}}-\frac{p_k}{q_k}\right|=\frac{1}{(\alpha_{k+1}q_k+q_{k-1})q_k}.$$

Обозначим $\psi_k=\alpha_{k+1}+\dfrac{q_{k-1}}{q_k}$.

Лемма. ***Если $\psi_k\leqslant\sqrt{5}$ и $\psi_{k+1}\leqslant\sqrt{5}$, то $\dfrac{q_k}{q_{k+1}}>\dfrac{\sqrt{5}-1}{2}$.***

Доказательство. Обозначим $x=q_k/q_{k+1}$. Неравенство $\psi_{k+1}\leqslant\sqrt{5}$ запишем в виде $\alpha_{k+2}\leqslant\sqrt{5}-x$, а неравенство $\psi_k\leqslant\sqrt{5}$ преобразуем так:

$$\sqrt{5}\geqslant\alpha_{k+1}+\frac{q_{k-1}}{q_k}=\frac{1}{\alpha_{k+2}}+a_{k+1}+\frac{q_{k-1}}{q_k}=\frac{1}{\alpha_{k+2}}+\frac{q_{k+1}}{q_k},$$

то есть $\dfrac{1}{\alpha_{k+2}}\leqslant\sqrt{5}-\dfrac{1}{x}$. Перемножая, получаем

$$1=\alpha_{k+2}\cdot\frac{1}{\alpha_{k+2}}\leqslant(\sqrt{5}-x)\left(\sqrt{5}-\frac{1}{x}\right),$$

то есть $0\leqslant5-\sqrt{5}\left(x+\dfrac{1}{x}\right)$, $x^2-x\sqrt{5}+1\leqslant0$, $\left(x-\dfrac{\sqrt{5}-1}{2}\right)\times$ $\times\left(x-\dfrac{\sqrt{5}+1}{2}\right)\leqslant0$, наконец, $\dfrac{\sqrt{5}-1}{2}<x<\dfrac{\sqrt{5}+1}{2}$. Первое из этих неравенств и требовалось доказать. (Равенство невозможно, ибо число $x=q_k/q_{k+1}$ рациональное.)
Применяя лемму при $k=n-1$ и $k=n$, получаем противоречащее натуральности числа a_{n+1} неравенство

$$a_{n+1}=\frac{q_{n+1}}{q_n}-\frac{q_{n-1}}{q_n}<\frac{2}{\sqrt{5}-1}-\frac{\sqrt{5}-1}{2}=1.\ \blacksquare$$

Если натуральное число n — не квадрат, то $0<[\sqrt{d}]-\sqrt{d}<-1$ и поэтому $[\sqrt{d}]+\sqrt{d}$ — приведенная квадратичная иррациональность. По теореме Галуа, цепная дробь числа $[\sqrt{d}]+\sqrt{d}$ чисто периодическая. Начальный элемент ее периода равен $2[\sqrt{d}]$, поэтому цепная дробь числа $\sqrt{d}$ имеет предпериод длины 1, а последний элемент ее периода (самое правое красное число) равно $2[\sqrt{d}]$.

d	Цепная дробь числа $\sqrt{d}$	x	y	x^2-dy^2
2	1; 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2	1	1	−1
3	1; 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2	2	1	1
5	2; 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4	2	1	−1
6	2; 2, 4, 2, 4, 2, 4, 2, 4, 2, 4, 2, 4	5	2	1
7	2; 1, 1, 1, 4, 1, 1, 1, 4, 1, 1, 1, 4	8	3	1
8	2; 1, 4, 1, 4, 1, 4, 1, 4, 1, 4, 1, 4	3	1	1
10	3; 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6	3	1	−1
11	3; 3, 6, 3, 6, 3, 6, 3, 6, 3, 6, 3, 6	10	3	1
12	3; 2, 6, 2, 6, 2, 6, 2, 6, 2, 6, 2, 6	7	2	1
13	3; 1, 1, 1, 1, 6, 1, 1, 1, 1, 6, 1, 1	18	5	−1
14	3; 1, 2, 1, 6, 1, 2, 1, 6, 1, 2, 1, 6	15	4	1
15	3; 1, 6, 1, 6, 1, 6, 1, 6, 1, 6, 1, 6	4	1	1
17	4; 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8	4	1	−1
18	4; 4, 8, 4, 8, 4, 8, 4, 8, 4, 8, 4, 8	17	4	1
19	4; 2, 1, 3, 1, 2, 8, 2, 1, 3, 1, 2, 8	170	39	1
20	4; 2, 8, 2, 8, 2, 8, 2, 8, 2, 8, 2, 8	9	2	1
21	4; 1, 1, 2, 1, 1, 8, 1, 1, 2, 1, 1, 8	55	12	1
22	4; 1, 2, 4, 2, 1, 8, 1, 2, 4, 2, 1, 8	197	42	1
23	4; 1, 3, 1, 8, 1, 3, 1, 8, 1, 3, 1, 8	24	5	1
24	4; 1, 8, 1, 8, 1, 8, 1, 8, 1, 8, 1, 8	5	1	1
26	5; 10, 10, 10, 10, 10, 10, 10, 10	5	1	−1
27	5; 5, 10, 5, 10, 5, 10, 5, 10, 5, 10	26	5	1
28	5; 3, 2, 3, 10, 3, 2, 3, 10, 3, 2, 3	127	24	1
29	5; 2, 1, 1, 2, 10, 2, 1, 1, 2, 10, 2	70	13	−1
30	5; 2, 10, 2, 10, 2, 10, 2, 10, 2, 10	11	2	1
31	5; 1, 1, 3, 5, 3, 1, 1, 10, 1, 1, 3, 5	1520	273	1
32	5; 1, 1, 1, 10, 1, 1, 1, 10, 1, 1, 1	17	3	1
33	5; 1, 2, 1, 10, 1, 2, 1, 10, 1, 2, 1	23	4	1
34	5; 1, 4, 1, 10, 1, 4, 1, 10, 1, 4, 1	35	6	1
35	5; 1, 10, 1, 10, 1, 10, 1, 10, 1, 10	6	1	1
37	6; 12, 12, 12, 12, 12, 12, 12, 12	6	1	−1
38	6; 6, 12, 6, 12, 6, 12, 6, 12, 6, 12	37	6	1
39	6; 4, 12, 4, 12, 4, 12, 4, 12, 4, 12	25	4	1
40	6; 3, 12, 3, 12, 3, 12, 3, 12, 3, 12	19	3	1
41	6; 2, 2, 12, 2, 2, 12, 2, 2, 12, 2, 2	32	5	−1
42	6; 2, 12, 2, 12, 2, 12, 2, 12, 2, 12	13	2	1
43	6; 1, 1, 3, 1, 5, 1, 3, 1, 1, 12, 1, 1	3482	531	1
44	6; 1, 1, 1, 2, 1, 1, 1, 12, 1, 1, 1, 2	199	30	1
45	6; 1, 2, 2, 2, 1, 12, 1, 2, 2, 2, 1	161	24	1
46	6; 1, 3, 1, 1, 2, 6, 2, 1, 1, 3, 1, 12	24 335	3588	1
47	6; 1, 5, 1, 12, 1, 5, 1, 12, 1, 5, 1	48	7	1
48	6; 1, 12, 1, 12, 1, 12, 1, 12, 1, 12	7	1	1
50	7; 14, 14, 14, 14, 14, 14, 14, 14	7	1	−1

Теорему 5 можно частично обратить.

Теорема 7. ***Всякая несократимая дробь a/b, удовлетворяющая неравенству $\left|\alpha-\frac{a}{b}\right|<\frac{1}{2b^2}$, является более качественным приближением, чем все дроби с меньшими знаменателями, и поэтому — подходящей дробью числа α.***

Доказательство. Пусть $|d\alpha-c|\leqslant|b\alpha-a|$, где $\frac{a}{b}\neq\frac{c}{d}$. Очевидно,

$$\frac{1}{bd}\leqslant\frac{|bc-ad|}{bd}=\left|\frac{c}{d}-\frac{a}{b}\right|\leqslant\left|\frac{c}{d}-\alpha\right|+\left|\alpha-\frac{a}{b}\right|<\frac{1}{2bd}+\frac{1}{2b^2}=\frac{b+d}{2b^2d},$$

откуда $2b<b+d$, то есть $b<d$, что и требовалось доказать.

Следствие. ***Если $x, y, d\in\mathbb{N}$ и $x^2-dy^2=1$, то x/y — подходящая дробь числа $\sqrt{d}$.***

Доказательство. $\frac{x}{y}-\sqrt{d}=\frac{x^2-dy^2}{y(x+y\sqrt{d})}=\frac{1}{y(x+y\sqrt{d})}<\frac{1}{2y^2}$. ■

Все цепные дроби чисел $\sqrt{d}$, где d — не являющееся квадратом натуральное число, при $d\leqslant 50$, как видно из таблицы, периодические.

Теорема 8 (Ж. Л. Лагранж, 1770 г.). ***Всякая периодическая цепная дробь — квадратичная иррациональность, а всякая квадратичная иррациональность — периодическая цепная дробь.***

Доказательство. Для числа α, изображаемого цепной дробью с периодом длины r и предпериодом длины s, верно равенство $\alpha_{r+s}=\alpha_s$. Выражая из равенств $\alpha=\frac{p_{s-1}\alpha_s+p_{s-2}}{q_{s-1}\alpha_s+q_{s-2}}$ и $\alpha=\frac{p_{r+s-1}\alpha_{r+s}+p_{r+s-2}}{q_{r+s-1}\alpha_{r+s}+q_{r+s-2}}$ числа α_{r+s} и α_s через α, получаем уравнение $\frac{p_{s-2}-q_{s-2}\alpha}{q_{s-1}\alpha-p_{s-1}}=\frac{p_{r+s-2}-q_{r+s-2}\alpha}{q_{r+s-1}\alpha-p_{r+s-1}}$, откуда, освобождаясь от знаменателей, приходим к квадратному уравнению.

Докажем периодичность цепной дроби квадратичной иррациональности α. Подставив $\alpha=\frac{p_{n-1}\alpha_n+p_{n-2}}{q_{n-1}\alpha_n+q_{n-2}}$ в квадратное уравнение $ax^2+bx+c=0$ с целыми коэффициентами a, b, c, которому удовлетворяет число α, получаем уравнение вида $A_n\alpha^2+B_n\alpha+C_n=0$, где

$$A_n=ap_{n-1}^2+bp_{n-1}q_{n-1}+cq_{n-1}^2,\quad C_n=ap_{n-2}^2+bp_{n-2}q_{n-2}+cq_{n-2}^2,$$
$$B_n=2ap_{n-1}p_{n-2}+b(p_{n-1}q_{n-2}+p_{n-2}q_{n-1})+2cq_{n-1}q_{n-2}.$$

Поскольку $\left|\alpha-\frac{p_{n-1}}{q_{n-1}}\right|<\frac{1}{q_{n-1}^2}$, то $p_{n-1}=\alpha q_{n-1}+\varepsilon$, где $|\varepsilon|<\frac{1}{q_{n-1}}$. Следовательно,

$$A_n=a(\alpha q_{n-1}+\varepsilon)^2+b(\alpha q_{n-1}+\varepsilon)q_{n-1}+cq_{n-1}^2=(a\alpha^2+b\alpha+c)q_{n-1}^2+2a\alpha\varepsilon q_{n-1}+a\varepsilon^2+b\varepsilon q_{n-1},$$

откуда $|A_n|=|2a\alpha\varepsilon q_{n-1}+a\varepsilon^2+b\varepsilon q_{n-1}|<2|a\alpha|+|a|+|b|$. Таким образом, A_n (а в силу равенства $C_n=A_{n-1}$ таким свойством обладает и C_n) при изменении n может принимать лишь конечное множество значений.

Как нетрудно доказать (например, непосредственным вычислением),

$$B_n^2-4A_nC_n=(b^2-4ac)(p_{n-1}q_{n-2}-p_{n-2}q_{n-1})^2=b^2-4ac.$$

Поэтому и для B_n имеем лишь конечное множество возможных значений. Таким образом, множество возможных троек $(A_n; B_n; C_n)$ конечно. Поскольку n можно брать сколь угодно большим, то для некоторых r и s имеем $\alpha_s=\alpha_{r+s}$, а это и означает, что цепная дробь числа α периодическая. ■

Феликс Эдуард Жюстен Эмиль Борель (1871—1956) — французский математик, профессор Парижского университета и Нормальной школы (в 1911—1920 гг. был ее директором).

Борелевское множество — это множество, которое можно получить в результате не более чем счетной последовательности операций объединения и пересечения открытых и замкнутых множеств топологического пространства. (Открытые множества — это множества, содержащие вместе с каждой своей точкой некоторую ее окрестность. Замкнутые множества — это дополнения к открытым, то есть множества, содержащие все свои предельные точки.)

Точнее говоря, борелевское множество — это элемент наименьшего семейства множеств, содержащего в себе все открытые множества и замкнутого относительно перехода к дополнению и взятию объединения счетного набора множеств.

Борелевские множества нулевого порядка — это множества открытые или замкнутые. Первого порядка — множества типа F_σ (объединения счетных наборов замкнутых множеств) и G_δ (пересечения счетных наборов открытых множеств). Пример множества типа F_σ, но не G_δ — это множество точек с рациональными координатами. Борелевские множества второго порядка — это множества типа $F_{\sigma\delta}$ (пересечения счетных наборов множеств типа F_σ) и $G_{\delta\sigma}$ (объединения счетных наборов множеств типа G_δ). Так по индукции определяют борелевские множества конечных порядков. Эту классификацию продолжают по трансфинитной индукции на все счетные ординалы; на этом построение борелевских множеств заканчивается. ■

Для золотого сечения $\alpha=(1+\sqrt{5})/2=[1; 1, 1, 1, \ldots]$ имеем:

$$\psi_k=[1;1,1,\ldots]+[0;\underbrace{1,1,\ldots,1}_{k}]\to$$
$$\to\frac{1+\sqrt{5}}{2}+\frac{2}{1+\sqrt{5}}=\sqrt{5}$$

при $k\to\infty$, так что число $\sqrt{5}$ в теореме Гурвица—Бореля нельзя увеличить, сохранив ее утверждение верным для всех без исключений чисел. Чем меньше элементы цепной дроби, тем хуже число приближается дробями! ■

Какие квадратичные иррациональности разлагаются в чисто периодические цепные дроби? На этот вопрос ответил Э. Галуа в 1828 г.

Определение. Иррациональный корень α квадратного уравнения с целыми коэффициентами называют приведенной квадратичной иррациональностью, если $\alpha>1$, а второй корень того же уравнения лежит на интервале $(0;-1)$.

Теорема 9. ***В чисто периодические цепные дроби разлагаются приведенные квадратичные иррациональности и только они.***

Доказательство. Пусть α — чисто периодическая цепная дробь: $\alpha=\alpha_r$, где $r>0$. Тогда $a_0=a_r\geqslant 1$ и поэтому $\alpha>1$. Кроме того, $\alpha=\dfrac{p_{r-1}\alpha_r+p_{r-2}}{q_{r-1}\alpha_r+q_{r-2}}=\dfrac{p_{r-1}\alpha+p_{r-2}}{q_{r-1}\alpha+q_{r-2}}$, откуда

$$q_{r-1}\alpha^2+(q_{r-2}-p_{r-1})\alpha-p_{r-2}=0.$$

Рассмотрим число $\beta=[a_{r-1};a_{r-2},\ldots,a_0,a_{r-1},\ldots,a_1,a_0,\ldots]$; например, если $\alpha=[4;3,2,1,4,3,2,1,\ldots]$, то $\beta=[1;2,3,4,1,2,3,4,\ldots]$. Очевидно, $\beta>1$ и $\beta=$
$=\dfrac{[a_{r-1},a_{r-2},\ldots,a_1,a_0]\beta+[a_{r-1},a_{r-2},\ldots,a_1]}{[a_{r-2},\ldots,a_1,a_0]\beta+[a_{r-2},\ldots,a_1]}=\dfrac{[a_0,a_1,\ldots,a_{r-1}]\beta+[a_1,\ldots,a_{r-2},a_{r-1}]}{[a_0,a_1,\ldots,a_{r-2}]\beta+[a_1,\ldots,a_{r-2}]}=$
$=\dfrac{p_{r-1}\beta+q_{r-1}}{p_{r-2}\beta+q_{r-2}}$, откуда $p_{r-2}\beta^2+(q_{r-2}-p_{r-1})\beta-q_{r-1}=0$. Обозначив $\gamma=-1/\beta$, имеем $-1<\gamma<0$ и

$$p_{r-2}\left(\frac{-1}{\gamma}\right)^2+(q_{r-2}-p_{r-1})\left(\frac{-1}{\gamma}\right)-q_{r-1}=0,$$

то есть $q_{r-1}\gamma^2+(q_{r-2}-p_{r-1})\gamma-p_{r-2}=0$. Значит, γ — корень того же квадратного уравнения, что и α. В одну сторону теорема Галуа доказана. ■

Рассмотрим уравнение $ax^2+bx+c=0$, где a, b, c —целые числа, $a>0$. Пусть α — его иррациональный корень, γ — сопряженное число, $-1<\gamma<0$. Обозначим $d=b^2-4ac$. Поскольку $0>\gamma=\dfrac{b-\sqrt{d}}{2a}$, то $b<\sqrt{d}$. Из неравенства $0<\alpha+\gamma=\dfrac{b}{a}$ имеем $b>0$. Наконец, $a=\dfrac{b+\sqrt{d}}{2\alpha}<\sqrt{d}$.

Докажем, что α_1 — приведенная квадратичная иррациональность. Уравнение $a\left(a_0+\dfrac{1}{\alpha_1}\right)^2+b\left(a_0+\dfrac{1}{\alpha_1}\right)+c=0$ приведем к виду

$$(aa_0^2+ba_0+c)\alpha_1^2+(2aa_0+b)\alpha_1+a=0. \qquad (**)$$

Числу $\alpha_1=\dfrac{1}{\alpha-a_0}$ сопряжено число $\dfrac{1}{\gamma-a_0}$. Очевидны неравенства $\alpha_1>1$, $a_0\geqslant 1$, $\gamma-a_0<-1$, $-1<\dfrac{1}{\gamma-a_0}<0$. Осталось вспомнить, что $\dfrac{1}{\gamma-a_0}=\gamma_1$.

Таким образом, все числа $\alpha,\alpha_1,\alpha_2,\ldots$ — приведенные квадратичные иррациональности. Дискриминант уравнения $(**)$ равен $(2aa_0+b)^2-4(aa_0^2+ba_0+c)a=$
$=b^2-4ac$. Поэтому дискриминанты всех соответствующих квадратных уравнений равны d. Их коэффициенты при x^2 и x — натуральные числа, меньшие $\sqrt{d}$. Поскольку свободный член выражается через коэффициенты при x^2 и x и дискриминант d, то множество интересующих нас квадратных уравнений конечно. Следовательно, $\alpha_s=\alpha_{r+s}$ для некоторых $r>0$ и s.

Если $s>0$, то $-1<a_{s-1}+\dfrac{1}{\gamma_s}<0$. Этими неравенствами число a_{s-1} определено однозначно. Поскольку $\gamma_s=\gamma_{s+r-1}$, то определяемое аналогичными неравенствами число a_{s+r-1} равно a_{s-1}. Значит, $\alpha_{s-1}=a_{s-1}+\dfrac{1}{\alpha_s}=a_{s+r-1}+\dfrac{1}{\alpha_{s+r}}=\alpha_{r+s-1}$.

Таким же образом можно уменьшить s еще на единицу и так далее до тех пор, пока не придем к равенству $s=0$. ■

Александр Яковлевич Хинчин (1894—1959) родился в селе Кондрово Медынского уезда (ныне это районный центр Калужской области). Его отец — главный инженер Кондровской бумажной фабрики. В юности (1908—1914) писал стихи. Увлекался театром: организовал в Кондрове театральную труппу, в которой был режиссером и актером. Брал уроки сценической речи, что очень пригодилось впоследствии при чтении лекций.

Доказал закон повторного логарифма: если $\xi_1,\xi_2,\xi_3,\ldots$ — независимые случайные величины, каждая из которых принимает с вероятностью $1/2$ значение 1 и с вероятностью $1/2$ — значение -1, то для любого положительного числа ε почти всегда для почти всех n верно неравенство $\xi_1+\xi_2+\ldots+\xi_n<(1+\varepsilon)\sqrt{2n\ln\ln n}$ и для бесконечно многих n — неравенство $\xi_1+\xi_2+\ldots+\xi_n>(1-\varepsilon)\sqrt{2n\ln\ln n}$.

Ввел термин «закон больших чисел» и в 1929 г. доказал, что для одинаково распределенных случайных величин $\xi_1,\xi_2,\xi_3,\ldots$, имеющих конечное среднее значение M, для любого положительного числа ε вероятность выполнения неравенства

$$\left|\frac{\xi_1+\xi_2+\ldots+\xi_n}{n}-M\right|>\varepsilon$$

стремится к 0 при $n\to\infty$.

Занимался теорией массового обслуживания, статистической механикой и теорией информации. Доказал, что для почти всех вещественных чисел α (то есть для всех, кроме чисел некоторого множества, которое для любого

положительного числа ε можно покрыть некоторым множеством отрезков, сумма длин которых меньше ε) любое натуральное число k встречается в качестве элемента цепной дроби с плотностью $\log_2\left(1+\frac{1}{k(k+2)}\right)$; среднее геометрическое $\sqrt[n]{a_1a_2\ldots a_n}$ элементов цепной дроби стремится к числу $\prod\limits_{n=1}^{\infty}\left(1+\frac{1}{n(n+2)}\right)^{\log_2 n}$; а величина $\sqrt[n]{q_n}$ стремится к некоторой постоянной величине. (П. Л. Леви доказал, что она равна $\frac{\pi^2}{12\ln 2}$.) В книге «Цепные дроби» изложил эти результаты вместе с доказательством О. Р. Кузьмина (1891—1949) гипотезы Гаусса о том, что если для любого числа $x\in[0;1]$ и для любого натурального числа n обозначить через $M_n(x)$ меру множества таких чисел α отрезка $[0;1]$, для которых $[0;a_n,a_{n+1},a_{n+2},\ldots]<x$, то

$$\lim_{n\to\infty} M_n(x)=\log_2(1+x).$$

Участвовал в редактировании и переработке учебников А. П. Киселева для средней школы. Этими учебниками пользовались почти полвека. Автор книг «Восемь лекций по математическому анализу», «Элементарное введение в теорию вероятностей», «Три жемчужины теории чисел», «Педагогические статьи», статей «О воспитательном эффекте уроков математики», «О формализме в преподавании математики». ■

На компьютере можно вычислить: $\sqrt[3]{2}=[1;3,1,5,1,1,4,1,1,8,1,14,1,10,2,1,4,12,2,3,2,1,3,4,1,1,2,14,3,12,1,15,3,1,4,534,1,1,5,1,1,121,1,2,2,4,10,3,2,2,41,1,1,1,3,7,2,2,9,4,1,3,7,6,1,1,2,9,2,3,3,1,1,69,1,12,\ldots]$. Ограничена или нет последовательность неполных частных цепной дроби числа $\sqrt[3]{2}$ (или любого алгебраического числа степени выше 2), неизвестно. ■

Пусть p — простое число, $m^2\equiv -1 \pmod p$, причем $0<m<p/2$. Ж. А. Серре (1819—1885) доказал, что цепная дробь числа p/m имеет вид $[a_0;a_1,\ldots,a_m,a_m,\ldots,a_1,a_0]$. При этом

$$p=[a_0,a_1,\ldots,a_m]^2+[a_0,a_1,\ldots,a_{m-1}]^2.$$ ■

Теорема Лагранжа — непосредственное следствие теоремы Галуа. В самом деле, для любой квадратичной иррациональности α имеем $\alpha_n>1$ при любом натуральном n. Из равенства $\alpha=\frac{p_{n-1}\alpha_n+p_{n-2}}{q_{n-1}\alpha_n+q_{n-2}}$ находим $\alpha_n=\frac{p_{n-2}-q_{n-2}\alpha}{q_{n-1}\alpha-p_{n-1}}$. Обозначая сопряжение чертой над числом, имеем

$$\overline{\alpha_n}=-\frac{q_{n-2}}{q_{n-1}}\cdot\frac{\overline{\alpha}-\frac{p_{n-2}}{q_{n-2}}}{\overline{\alpha}-\frac{p_{n-1}}{q_{n-1}}}.$$

Разумеется, $q_{n-2}<q_{n-1}$. В случае $\overline{\alpha}<\alpha$ воспользуемся тем, что при любом достаточно большом четном n дробь p_{n-2}/q_{n-2} лежит между $\overline{\alpha}$ и α, а дробь p_{n-1}/q_{n-1} больше числа α; таким образом, неравенства $-1<\alpha_n<0$ очевидны. А если $\overline{\alpha}>\alpha$, достаточно рассмотреть любое достаточно большое нечетное n. ■

Объясним специфический вид цепных дробей чисел вида $\sqrt{d}$, где d — натуральное число, не являющееся квадратом. Обозначим $[\sqrt{d}]=k$. Число $k+\sqrt{d}$ — приведенная квадратичная иррациональность. Пусть $[a_0;a_1,a_2,\ldots,a_{r-2},a_{r-1},\ldots]$ — ее цепная дробь, r — длина периода. По теореме Галуа имеем $[a_{r-1};a_{r-2},a_{r-3},\ldots,a_1,a_0,\ldots]=-\frac{1}{k-\sqrt{d}}=\frac{1}{\sqrt{d}-k}$, так что

$$k+\sqrt{d}=2k+\frac{1}{[a_{r-1};a_{r-2},a_{r-3},\ldots,a_1,a_0,\ldots]}=[2k;a_{r-1},a_{r-2},a_{r-3},\ldots,a_1,a_0,\ldots].$$

Следовательно, $a_r=a_0=2k$, $a_1=a_{r-1}$, $a_2=a_{r-2}$, ... Период цепной дроби числа $\sqrt{d}$ оканчивается на число $2[\sqrt{d}]$, а после отбрасывания этого числа становится палиндромом: читается слева направо так же, как справа налево. ■

Уравнение Пелля легко решить при помощи цепной дроби числа $\sqrt{d}$. (Между прочим, английский метод решения уравнения Пелля — это, по сути, алгоритм разложения числа $\sqrt{d}$ в цепную дробь.)

Пусть r — длина периода (необязательно наименьшая) цепной дроби числа $\alpha=\sqrt{d}$. Как вы помните, $\alpha_r=k+\sqrt{d}$, так что

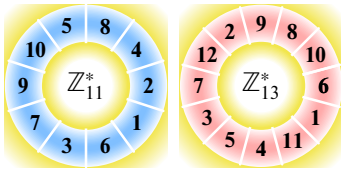
$$\sqrt{d}=\alpha=\frac{\alpha_r p_{r-1}+p_{r-2}}{\alpha_r p_{r-1}+p_{r-2}}=\frac{(k+\sqrt{d})p_{r-1}+p_{r-2}}{(k+\sqrt{d})q_{r-1}+q_{r-2}},$$

откуда $dq_{r-1}+(kq_{r-1}+q_{r-2})\sqrt{d}=kp_{r-1}+p_{r-2}+p_{r-1}\sqrt{d}$. В силу иррациональности числа $\sqrt{d}$ имеем $dq_{r-1}=kp_{r-1}+p_{r-2}$ и $kq_{r-1}+q_{r-2}=p_{r-1}$. Следовательно,

$$(-1)^r=p_{r-1}q_{r-2}-p_{r-2}q_{r-1}=p_{r-1}(p_{r-1}-kq_{r-1})-(dq_{r-1}-kp_{r-1})q_{r-1}=p_{r-1}^2-dq_{r-1}^2.$$

Можно доказать, что $(p_{r-1};q_{r-1})$ — наименьшее решение уравнения $x^2-dy^2=\pm1$ в натуральных числах. Если длина наименьшего периода четна, то уравнение $x^2-dy^2=-1$ решений в целых числах не имеет. ■

Пусть p — простое число, $p\equiv1 \pmod 4$. А. М. Лежандр в 1785 г. доказал, что уравнение $x^2-py^2=-1$ имеет решение в целых числах. Следовательно, длина периода разложения $\sqrt{p}$ нечетна: $\sqrt{p}=[k;a_1,a_2,\ldots,a_m,a_m,\ldots,a_2,a_1,2k,\ldots]$. «Отрезав» первые $m+1$ элементов, получаем чисто периодическую цепную дробь $\alpha=[a_m,\ldots,a_2,a_1,2k,a_1,a_2,\ldots,a_m,\ldots]$. Поскольку ее период симметричен, $\overline{\alpha}=-1/\alpha$. Из доказательства теоремы Галуа имеем $\alpha=\frac{a+\sqrt{p}}{b}$ для некоторых натуральных a и b. Следовательно, $-1=\alpha\overline{\alpha}=\frac{a+\sqrt{p}}{b}\cdot\frac{a-\sqrt{p}}{b}$, то есть $a^2+b^2=p$. Теорема Ферма—Эйлера доказана при помощи цепных дробей! ■



Для любых трех стоящих подряд чисел a, b, c левого рисунка разность b^2-ac кратна 11. И это не курьез, а частный случай общей конструкции: взяв первообразный корень g по простому модулю p, рассмотрим геометрическую прогрессию g, g^2, …, g^{p-2}, g^{p-1} и выпишем вдоль окружности остатки от деления ее членов на p. (Левый рисунок иллюстрирует случай $g=2$ и $p=11$; правый — $g=6$ и $p=13$.)

Для любых трех подряд идущих членов a, b, c геометрической прогрессии выполнено равенство $b^2=ac$. Поскольку мы заменяли числа на их остатки от деления на p, то вместо равенств имеем сравнения по модулю p. ■

Функция Кармайкла

Малая теорема Ферма, цикличность мультипликативной группы вычетов по простому модулю — теорема Гаусса о существовании первообразного корня, вычисление функции Кармайкла $\lambda(n)$ *— нахождение такого наименьшего числа* m*, что* $a^m\equiv 1 \pmod n$ *при* НОД$(a, n)=1$*, а также многие другие классические теоремы и трудные задачи объединены одной важной идеей — периодичностью остатков от деления степеней данного числа* a *на данное число* n*.*

Какие остатки дают степени двойки при делении на 11? Чтобы ответить на этот вопрос, составим таблицу:

n	1	2	3	4	5	6	7	8	9	10	11	12
2^n	2	4	8	16	32	64	128	256	512	1024	2048	4096
$2^n \bmod 11$	2	4	8	5	10	9	7	3	6	1	2	4

Дальше можно не продолжать: $2^{10+n}=2^{10}\cdot 2^n\equiv 1\cdot 2^n=2^n \pmod{11}$, остатки повторяются с периодом 10. Между прочим, средняя строка таблицы излишняя: в нижней строке каждое следующее число — остаток от деления на 11 удвоенного предыдущего числа.

Как бы то ни было, $2^{10}\equiv 1 \pmod{11}$. Ничего удивительного в этом нет, это всего лишь частный случай малой теоремы Ферма. Интереснее другое: в нижней строке таблицы присутствуют все ненулевые остатки от деления на 11. (Например, $3\equiv 2^8$, $5\equiv 2^4$, $7\equiv 2^7$ и $10\equiv 2^5$.) Другими словами, для любого целого числа a, не кратного 11, существует такое s, что $a\equiv 2^s \pmod{11}$. А сейчас — внимание:

$$a^{10}\equiv (2^s)^{10}=(2^{10})^s\equiv 1^s=1 \pmod{11}.$$

Мы проверили малую теорему Ферма не только для $a=2$, но сразу для всех ненулевых остатков! Красиво и неожиданно, не правда ли? ■

Число g называют первообразным корнем по простому модулю p, если числа $g, g^2, \ldots, g^{p-1}$ дают разные (ненулевые) остатки при делении на p. Другими словами, g — первообразный корень, если для любого целого числа a, не кратного числу p, существует такое s, что $a\equiv g^s \pmod p$. ■

Если $a\equiv g^s$ **и** $b\equiv g^t$, то $ab\equiv g^s g^t=g^{s+t} \pmod{11}$. Это сводит умножение по модулю 11 к сложению по модулю 10 (именно по этому модулю рассматриваются числа s и t). Взгляните на таблицу 1 сложения по модулю 10 и таблицу 2 умножения по модулю 11 (строки и столбцы которой переставлены в соответствии с остатками от деления степеней двойки на 11).

Они похожи, как близнецы! Мультипликативная группа вычетов $\mathbb{Z}_{11}^*$ (ее элементы — ненулевые классы вычетов по модулю 11, операция — умножение) изоморфна аддитивной группе $\mathbb{Z}_{10}$ вычетов по модулю 10 (элементы — классы вычетов по модулю 10, операция — сложение). Изоморфизм — это взаимно однозначное отображение, сохраняющее операцию. Например, изоморфизм между $\mathbb{Z}_{10}$ и $\mathbb{Z}_{11}^*$ можно установить, сопоставив каждому из чисел $s=0, 1, \ldots, 9$ число 2^s. При этом сумме $s+t \pmod{10}$ будет сопоставлено произведение $2^s\cdot 2^t \pmod{11}$. ■

Таблица 1. Сложение по модулю 10.

+	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9	0
2	2	3	4	5	6	7	8	9	0	1
3	3	4	5	6	7	8	9	0	1	2
4	4	5	6	7	8	9	0	1	2	3
5	5	6	7	8	9	0	1	2	3	4
6	6	7	8	9	0	1	2	3	4	5
7	7	8	9	0	1	2	3	4	5	6
8	8	9	0	1	2	3	4	5	6	7
9	9	0	1	2	3	4	5	6	7	8

Таблица 2. Умножение по модулю 11.

×	1	2	4	8	5	10	9	7	3	6
1	1	2	4	8	5	10	9	7	3	6
2	2	4	8	5	10	9	7	3	6	1
4	4	8	5	10	9	7	3	6	1	2
8	8	5	10	9	7	3	6	1	2	4
5	5	10	9	7	3	6	1	2	4	8
10	10	9	7	3	6	1	2	4	8	5
9	9	7	3	6	1	2	4	8	5	10
7	7	3	6	1	2	4	8	5	10	9
3	3	6	1	2	4	8	5	10	9	7
6	6	1	2	4	8	5	10	9	7	3

Рассмотрим остатки от деления степеней двойки на 17:

n	1	2	3	4	5	6	7	8
$2^n \bmod 17$	2	4	8	16	15	13	9	1

Зацикливание произошло слишком рано: $2^8 \equiv 1 \pmod{17}$. Поэтому не все ненулевые остатки от деления на 17 — остатки от деления степеней двойки. Например, в нижней строке таблицы нет числа 5, так что разность $2^n - 5$ не кратна 17 ни при каком натуральном n. А теперь начнем с тройки и, не забывая переходить к остатку от деления на 17, будем умножать, умножать и умножать на три:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$3^n \bmod 17$	3	9	10	13	5	15	11	16	14	8	7	4	12	2	6	1

Мы получили все 16 возможных ненулевых остатков от деления на 17. Значит, 3 — первообразный корень по модулю 17. ■

Не для каждого простого числа p в качестве первообразного корня годится 2 или 3. Например, легко проверить, что $2^{11} \equiv 1 \equiv 3^{11} \pmod{23}$, так что ни 2, ни 3 не являются первообразными корнями по модулю 23. (А вот -2 и -3 — являются.) ■

Разбиение на циклы. Пусть целое число a не кратно простому числу p. Рассмотрим множество ненулевых остатков от деления на p. От каждого ненулевого остатка x проведем стрелочку к остатку от деления ax на p. Получим граф, в котором из каждого ненулевого остатка x выходит одна стрелочка и к каждому ненулевому остатку ведет одна стрелочка (если бы к какому-то остатку y вели стрелки от x_1 и от x_2, то выполнялось бы сравнение $ax_1 \equiv y \equiv ax_2$, откуда $x_1 \equiv x_2 \pmod{p}$, так что $x_1 = x_2$).

Поэтому все $p-1$ ненулевых остатков от деления на p разбиваются на циклы вида $\{x, ax, \ldots, a^{k-1}x\}$, каждый из которых состоит из k остатков. Длину k этих циклов называют *порядком* числа a по модулю p. Очевидно, числа $a, a^2, \ldots, a^k$ $(\equiv 1)$ дают при делении на p разные остатки, а дальше последовательность периодична: $a^{k+1} \equiv a$, $a^{k+2} \equiv a^2$, ... При этом

$$a^k \equiv a^{2k} \equiv a^{3k} \equiv \ldots \equiv 1 \pmod{p},$$

а другие степени числа a не сравнимы с 1 по модулю p.

Теорема 1. ***Порядок k не кратного простому числу p целого числа является делителем числа $p-1$. В частности, $a^{p-1} \equiv 1 \pmod{p}$.***

Доказательство. Количество циклов равно $(p-1)/k$ и является целым числом. Возводя обе части сравнения $a^k \equiv 1$ в $(p-1)/k$-ю степень, получаем сравнение $a^{p-1} \equiv 1 \pmod{p}$. ■

Рассмотрев любое натуральное число n вместо простого числа p, аналогичным образом получаем следующие утверждения:

- если целое число a взаимно просто с натуральным числом n, то существует бесконечно много таких натуральных m, что $a^m - 1$ кратно n. Все они кратны наименьшему из них — порядку числа a по модулю n;
- если целое число a взаимно просто с натуральным числом n и если $a^r \equiv a^s \equiv 1 \pmod{n}$, то $a^{\text{НОД}(r,s)} \equiv 1 \pmod{n}$.
- порядок по модулю n взаимно простого с n целого числа a — делитель числа $\varphi(n)$; при этом $a^{\varphi(n)} \equiv 1 \pmod{n}$. ■

Очевидно, -1 является первообразным корнем только по модулю 2 или 3. Далее, из равенства $(a^2)^{(p-1)/2} = a^{p-1}$ следует, что квадрат не может быть первообразным корнем ни по какому нечетному простому модулю p. Э. Артин предположил, что для любого целого числа $g \neq -1$, не являющегося квадратом целого числа, существует бесконечно много таких простых p, что g — первообразный корень по модулю p.

Более того, некоторые вероятностные соображения привели Артина к следующему уточнению гипотезы: если k есть наибольшее такое число, что g является k-й степенью, то отношение количества $\pi_g(n)$ простых чисел, не превосходящих n, по модулю которых g является первообразным корнем, к количеству $\pi(n)$ всех простых чисел, не превосходящих n, стремится при $n \to \infty$ к зависящему только от k пределу

$$\lim_{n\to\infty} \frac{\pi_g(n)}{\pi(n)} = \prod_{k:q} \left(1 - \frac{1}{q-1}\right) \times \prod_{k\nmid q} \left(1 - \frac{1}{q(q-1)}\right),$$

где первое произведение распространено на все простые числа q, являющиеся делителями k, а второе — на все простые числа q, не являющиеся делителями k.

К настоящему времени гипотеза Артина не доказана, хотя некоторый ее аналог, относящийся к полю рациональных функций от одной переменной над конечным полем, доказать удалось. ■

Если a, k, m — натуральные числа, $a > 1$, то $a^m - 1$ делится на $a^k - 1$ тогда и только тогда, когда m делится на k.

Доказательство. Если $m = kn$, то

$$a^m - 1 = (a^k - 1) \times \times (a^{k(n-1)} + a^{k(n-2)} + \ldots + a^k + 1).$$

Обратно, если m не делится на k, то разделим m на k с остатком:

$$m = kn + r,$$

где $0 < r < k$, и рассмотрим формулу

$$a^{kn+r} - 1 = a^{kn+r} - a^r + a^r - 1 = = a^r(a^{kn} - 1) + (a^r - 1).$$

Число $a^r - 1$ не делится на $a^k - 1$, поскольку $0 < a^r - 1 < a^k - 1$. ■

Среди чисел вида 2^n-3 бесконечно много чисел, кратных 5, и бесконечно много кратных 13, но нет ни одного числа, кратного 65 ($=5\cdot 13$).
Дело в том, что $2^n\equiv 3 \pmod 5$ при $n\equiv 3 \pmod 4$, а $2^n\equiv 3 \pmod{13}$ при $n\equiv 4 \pmod{12}$. Число n не может быть и нечетным, и четным одновременно. ■

Ни при каком целом a число a^2+a+1 не кратно ни 5, ни 11, ни 17, и вообще никакому числу вида $6m-1$, где m — натуральное число.
В самом деле, число $6m-1$ имеет хотя бы один простой делитель $p=6k-1$. Если a^2+a+1 кратно p, то $a^3-1=(a-1)(a^2+a+1)$ тоже кратно p. В случае $a\equiv 1 \pmod p$ имеем $a^2+a+1\equiv 1^2+1+1=3$, что невозможно, ибо $p\neq 3$. Значит, порядок числа a по модулю p равен 3, откуда $p-1$ кратно 3. Но $p-1=6k-2$ не кратно 3.
Аналогично, воспользовавшись равенством $a^{12}-1=(a^6-1)\times(a^2+1)(a^4-a^2+1)$, нетрудно доказать, что всякий натуральный делитель числа a^4-a^2+1, где a — целое, дает остаток 1 при делении на 12. ■

Если a^2+1 делится на простое число $p\neq 2$, то $a^2\equiv -1 \pmod p$, откуда

$$a^4=(a^2)^2\equiv(-1)^2=1 \pmod p.$$

Значит, порядок числа a равен 1, 2 или 4. Первые два случая невозможны: сравнение $a^2\equiv 1$ противоречит сравнению $a^2\equiv -1 \pmod p$.
В третьем случае в силу теоремы 1 число $p-1$ делится на 4. Мы доказали часто используемое утверждение: любой нечетный простой делитель числа a^2+1 имеет вид $p=4k+1$ (а не $4k+3$).
Рассуждая аналогично, можно доказать, что если p — нечетный простой делитель числа $a^{2^n}+1$, то $p-1$ делится на 2^{n+1}.
Верно и обратное: для любого простого числа $p=2^{n+1}k+1$ существует кратное ему число вида $a^{2^n}+1$. В самом деле, пусть $a=g^k$, где g — первообразный корень по модулю p. Тогда $a^{2^n}=g^{2^nk}=g^{(p-1)/2}$. Число $g^{(p-1)/2}$ не сравнимо с единицей по модулю p, а его квадрат сравним: $g^{p-1}\equiv 1 \pmod p$, поэтому

$$a^{2^n}=g^{(p-1)/2}\equiv -1 \pmod p.$$ ■

Теорема 1 позволяет решать задачи, которые без нее совершенно неприступны. Рассмотрим два примера.

Теорема 2. ***Если сумма $a^4+a^3+a^2+a+1$, где a — целое число, кратна простому числу p, то $p=5$ или $p\equiv 1 \pmod 5$.***

Доказательство. Если $a\equiv 1 \pmod p$, то $a^4+a^3+a^2+a+1\equiv 1^4+1^3+1^2+1+1=5 \pmod p$, так что p — делитель числа 5; попросту говоря, $p=5$.
Если же $a\not\equiv 1 \pmod p$, из формулы

$$a^5-1=(a-1)(a^4+a^3+a^2+a+1):p$$

заключаем, что порядок числа a по модулю p равен 5. Порядок является делителем числа $p-1$; следовательно, $p-1$ делится на 5.
Верно и обратное утверждение: для $p=5$ годится $a=1$, а для простого числа $p=5k+1$ годится $a=g^k$, где g — первообразный корень по модулю p (существование которого мы докажем ниже). В самом деле, $g^{5k}=g^{p-1}\equiv 1 \pmod p$; произведение $(a-1)(a^4+a^3+a^2+a+1)=a^5-1$ кратно p. Первый множитель не делится на p, поэтому второй делится. ■

Когда 2^n+1 делится на n? При помощи компьютера можно выписать несколько первых таких чисел: $n=1$, 3, 9, 27, 81, 171 (заметьте: предыдущие числа — степени тройки, а $171=19\cdot 9$), 243, 513, 729, 1539, 2187, 3249, 4617, 6561, 9747, 13 203 (впервые возник отличный от 3 и 19 простой множитель: $13\,203=163\cdot 81$), 13 851, 19 683, 29 241, 39 609, 41 553, 59 049, 61 731, 87 723, 97 641, 118 827, 124 659, . . .
Все эти числа (кроме единицы, но она не в счет) делятся на 3. И среди них присутствуют все степени тройки. Как это объяснить?
Со степенями тройки проблем нет: по индукции легко доказать, что $2^{3^k}+1$ делится на 3^{k+1}.
Нетрудно доказать и «лемму о подъеме»: если n кратно 3 и 2^n+1 кратно n, то $2^{3n}+1$ кратно $3n$. В самом деле,

$$2^{3n}+1=(2^n+1)((2^n)^2-2^n+1);$$

первый множитель кратен n, а второй кратен 3. (Почему? Потому что из условия $2^n\equiv -1 \pmod n$ имеем $2^n\equiv -1 \pmod 3$, откуда $(2^n)^2-2^n+1\equiv(-1)^2-(-1)+1\equiv 0 \pmod 3$.)

Теорема 3. ***Если $n>1$ и 2^n+1 кратно n, то n делится на 3.***

Доказательство. Рассмотрим наименьший простой делитель p числа n. Тогда $2^n\equiv -1 \pmod p$. Значит, $2^{2n}\equiv 1 \pmod p$, и поэтому порядок числа 2 по модулю p является делителем числа $2n$. Поскольку порядок числа по модулю p не превосходит $p-1$, а число n не имеет простых делителей, меньших p, есть единственная возможность: порядок числа 2 по модулю p равен 2. Это значит, что $2^2\equiv 1 \pmod p$, то есть $p=3$, что и требовалось доказать. ■

Какие есть первообразные корни по модулю 11, кроме числа 2? Для ответа не нужно перебирать все числа 3, 4, 5, . . . , 9, 10 и составлять для каждого из них таблицу вроде таблицы 1. Некоторые степени двойки можно сразу отбросить: $(2^2)^5=2^{10}\equiv 1$, $(2^4)^5=2^{20}\equiv 1$, $(2^5)^2\equiv 1$, $(2^6)^5\equiv 1$, $(2^8)^5\equiv 1 \pmod{11}$. А вот степени двойки $2^1\equiv 2$, $2^3\equiv 8$, $2^7\equiv 7$ и $2^9\equiv 6$, показатели которых взаимно просты с 10, являются первообразными корнями.
И вообще, если g — первообразный корень по простому модулю p, то g^s является первообразным корнем в том и только том случае, когда s и $p-1$ взаимно просты. ■

Порядки классов вычетов. В следующей таблице для каждого ненулевого остатка $a \bmod 11$ указан его порядок k:

a	1	2	3	4	5	6	7	8	9	10
k	1	10	5	5	5	10	10	10	5	2

Как положено, порядки — делители числа 10. Посчитаем, сколько раз в нижней строке встречаются числа 1, 2, 5 и 10:

Порядок	1	2	5	10
Встречается	1	1	4	4

Видна закономерность? Если нет, посмотрите на таблицу для $p=13$:

a	1	2	3	4	5	6	7	8	9	10	11	12
k	1	12	3	6	4	12	12	4	3	6	12	2

Посчитаем, сколько раз встречаются в нижней строке числа 1, 2, 3, 4, 6 и 12:

Порядок	1	2	3	4	6	12
Встречается	1	1	2	2	2	4

Если вы все еще не догадались, составьте такие таблицы для нескольких других простых чисел p, и рано или поздно увидите, что в нижних строках этих таблиц — значения функции Эйлера: $\varphi(1)=1$, $\varphi(2)=1$, $\varphi(3)=2$, $\varphi(4)=2$, $\varphi(5)=4$, $\varphi(6)=2$, $\varphi(10)=4$ и $\varphi(12)=4$. К. Ф. Гаусс в «Арифметических исследованиях» (1801) доказал, что это не случайность, а общий закон.

Теорема 4. ***Среди $p-1$ ненулевых классов вычетов по простому модулю p порядок k, где k — делитель числа $p-1$, имеют ровно $\varphi(k)$ классов вычетов. В частности, $\lambda(p)=p-1$, а первообразных корней по модулю p ровно $\varphi(p-1)$ штук.***

Доказательство. I способ. Из теоремы Безу следует, что если $a_1, a_2, \ldots, a_m$ — различные корни многочлена $f(x)$, то $f(x)=(x-a_1)(x-a_2)\ldots(x-a_m)g(x)$, где g — некоторый многочлен. Применяя это соображение к многочлену $x^{p-1}-1$, получаем его разложение на линейные множители:

$$x^{p-1}-1\equiv(x-1)(x-2)\ldots(x-p+1),$$

где знак сравнения означает, что если раскрыть все скобки в правой части и вычесть из нее левую, то получим многочлен, коэффициенты которого кратны p. (Подстановка $x=0$, как заметил Лагранж, приводит к теореме Вильсона: $(p-1)!\equiv-1 \pmod p$ для любого простого числа p.)

Если k — делитель числа $p-1$, то многочлен x^k-1 является делителем многочлена $x^{p-1}-1$. Поскольку $x^{p-1}-1$ разлагается в произведение многочленов первой степени, то его делитель x^k-1 является произведением k многочленов первой степени и, следовательно, сравнению $x^k\equiv1\pmod p$ удовлетворяют ровно k классов вычетов по модулю p. Докажем по индукции, что среди ненулевых классов вычетов по простому модулю p существует ровно $\varphi(k)$ классов порядка k.

База. Для $k=1$ утверждение верно.

Переход. Рассмотрим некоторый делитель k числа $p-1$. Предположим, что для любого делителя d числа k, где $d<k$, существует ровно $\varphi(d)$ классов вычетов порядка d. Найдем количество классов вычетов порядка k.

Сравнению $x^k\equiv1\pmod p$ удовлетворяют ровно k классов вычетов. Каждое решение x этого сравнения имеет некоторый порядок по модулю p, причем этот порядок — делитель числа k. В статье «Функция Эйлера» доказано, что сумма чисел вида $\varphi(d)$, где $k:d$, равна k. Следовательно, классов порядка k ровно $\varphi(k)$ штук.

Теорему Э. Безу (1730—1783) при помощи деления многочленов с остатком можно сформулировать и доказать очень коротко. В равенство

$$f(x)=(x-a)g(x)+r,$$

где $g(x)$ — многочлен (неполное частное), а r — число (остаток), подставим $x=a$. Получим:

$$f(a)=(a-a)g(a)+r=r.$$

Значит, остаток r от деления $f(x)$ на $x-a$ равен $f(a)$. Это и есть теорема Безу. Есть у нее и другие, более длинные, но не менее естественные формулировка и доказательство.

Теорема. ***Число a является корнем многочлена $f(x)$ тогда и только тогда, когда $f(x)$ делится на $x-a$, то есть $f(x)=(x-a)g(x)$, где g — некоторый многочлен.***

Доказательство. Если $f(x)=(x-a)g(x)$, то $f(a)=(a-a)\times g(a)=0$. Обратно, пусть $f(a)=0$. Подставим в многочлен

$$f(x)=k_nx^n+k_{n-1}x^{n-1}+\ldots+k_2x^2+k_1x+k_0$$

число a. Получим:

$$0=f(a)=k_na^n+k_{n-1}a^{n-1}+\ldots+k_2a^2+k_1a+k_0.$$

Следовательно,

$$f(x)=f(x)-f(a)=k_n(x^n-a^n)+k_{n-1}(x^{n-1}-a^{n-1})+\ldots+k_2(x^2-a^2)+k_1(x-a).$$

Каждая из разностей

$$x-a,$$
$$x^2-a^2=(x-a)(x+a),$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$x^n-a^n=(x-a)\times(x^{n-1}+x^{n-2}a+\ldots+xa^{n-2}+a^{n-1})$$

кратна $x-a$. ■

Если a^n+1 — простое число, a, n — натуральные числа, $a>1$, то a четно и n — степень числа 2.

Простые числа вида $2^{2^n}+1$ называют числами Ферма. Их известно всего пять: $2^{2^0}+1=3$, $2^{2^1}+1=5$, $2^{2^2}+1=17$, $2^{2^3}+1=257$ и $2^{2^4}+1=65\,537$. Существуют ли другие, неизвестно. ■

Если 2^n-1 делится на 2^m+1, то n делится на $2m$. ■

Для каких натуральных чисел m существует такое n, что 2^n+1 делится на 2^m-1?

Ответ: для $m=1$ или 2. ■

Если a^n-1 простое, $a>1$, $n>1$, то $a=2$ и n — простое.

Не при всяком простом p число 2^p-1 простое: например, $2^{11}-1=2047=23\cdot 89$. Простые числа вида 2^p-1 называют *числами Мерсенна*.

Марен Мерсенн (1588—1648) занимался математикой, теорией музыки, физикой и философией. Он был товарищем Р. Декарта по учебе в иезуитском колледже и членом монашеского ордена минимов. Мерсенн сыграл выдающуюся роль как организатор науки. Он состоял в переписке с Р. Декартом, Ж. Роберваль, Б. Паскалем, Х. Гюйгенсом, Б. Кавальери, Б. Френиклем де Бесси, Дж. Валлисом, П. Ферма и др. Вокруг него образовался кружок ученых, который стал основой для создания Парижской академии наук (1666). В настоящий момент известно 41 число Мерсенна и неизвестно, конечно или бесконечно их множество. В 2004 г. было найдено число Мерсенна $2^{24\,036\,583}-1$; 18 февраля 2005 г. нашли наибольшее из известных на сегодняшний день: $2^{25\,964\,951}-1$ (в его десятичной записи 7 816 230 цифр!). ■

Пусть m — натуральное число, $m\geqslant 3$. Теорема Эйлера утверждает, что

$$a^{2^{m-1}}\equiv 1 \pmod{2^m}$$

для любого нечетного числа a. Докажем по индукции более сильное утверждение:

$$a^{2^{m-2}}\equiv 1 \pmod{2^m}.$$

База — случай $m=3$. Число

$$a^2-1=(a-1)(a+1)$$

кратно 8, поскольку одно из соседних четных чисел $a-1$ и $a+1$ кратно 4.

Переход. Пусть утверждение верно для некоторого $m\geqslant 3$. Рассмотрим разложение на множители:

$$a^{2^{m-1}}-1=(a^{2^{m-2}}-1)(a^{2^{m-2}}+1).$$

По предположению индукции, первый множитель кратен 2^m, а второй четен. Значит, произведение делится на 2^{m+1}, что и требовалось доказать. ■

Всякий простой делитель q числа $a^p\pm 1$, где a — натуральное число, $a>1$, p — простое, $p>2$, является делителем числа $a\pm 1$ или имеет вид $q=2pm+1$, где m — натуральное. ■

II способ не использует равенство $\sum\limits_{k\,\vdots\, d}\varphi(d)=k$ и доказывает даже более общее утверждение: любая конечная подгруппа любого поля циклична.

Лемма. ***Если порядки m и n чисел a и b взаимно просты, то порядок ab равен mn.***

Доказательство. Очевидно, $(ab)^{mn}=(a^m)^n(b^n)^m\equiv 1^n\cdot 1^m=1 \pmod p$. Осталось доказать, что никакой отличный от mn делитель числа mn не годится. Для этого достаточно рассмотреть числа вида mn/q, где q — простой делитель одного из чисел m и n. Пусть для определенности q — делитель числа m. Тогда $(ab)^{mn/q}=a^{mn/q}\cdot(b^n)^{m/q}\equiv a^{mn/q}\not\equiv 1 \pmod p$, ибо mn/q не кратно числу m.

Теперь доказать теорему 4 несложно. Пусть $p-1=q_1^{a_1}q_2^{a_2}\ldots q_s^{a_s}$ — разложение числа $p-1$ в произведение степеней различных простых чисел. Пусть g_1, $g_2,\ldots,g_s$ — такие не кратные p числа, что $g_k^{(p-1)/q_k}\not\equiv 1 \pmod p$ при $k=1,2,\ldots,s$. (Они существуют, поскольку многочлен не может иметь больше корней, чем его степень.) Очевидно, $g=g_1^{(p-1)/q_1^{a_1}}g_2^{(p-1)/q_2^{a_2}}\ldots g_s^{(p-1)/q_s^{a_s}}$ — произведение чисел порядков $q_1^{a_1}$, $q_2^{a_2}$, $\ldots$, $q_s^{a_s}$ — первообразный корень по модулю p. ■

Как известно, $\varphi(360)=\varphi(2^3\cdot 5\cdot 9)=\varphi(2^3)\cdot\varphi(5)\cdot\varphi(9)=4\cdot 4\cdot 6=96$. По теореме Эйлера, для любого целого a, взаимно простого с 360, выполнено сравнение

$$a^{96}\equiv 1 \pmod{360}.$$

А на самом деле верно даже сравнение $a^{12}\equiv 1 \pmod{360}$. Для доказательства достаточно применить теорему Эйлера по каждому из модулей 8, 5 и 9:

$$a^4\equiv 1 \pmod 8,\qquad a^4\equiv 1 \pmod 5,\qquad a^6\equiv 1 \pmod 9;$$

значит, $a^{12}\equiv 1$ по каждому из модулей 8, 5 и 9, а поэтому и по модулю 360. В общем виде это можно сформулировать следующим образом. Рассмотрим разложение

$$n=p_1^{m_1}p_2^{m_2}\ldots p_s^{m_s}$$

числа n в произведение степеней различных простых множителей. Обозначим через $f(n)$ наименьшее общее кратное чисел $\varphi(p_k^{m_k})$, где $k=1,2,\ldots,s$. Например, $f(360)=\text{НОК}[\varphi(2^3),\varphi(3^2),\varphi(5)]=\text{НОК}[4,6,4]=12$. Тогда при любом целом a, взаимно простом с n, справедливы сравнения

$$a^{f(n)}\equiv 1 \pmod{p_k^{m_k}},$$

где $k=1,2,\ldots,s$; следовательно, $a^{f(n)}\equiv 1 \pmod n$. ■

Через $\lambda(n)$ обозначим такое наименьшее натуральное число m, что a^m-1 кратно n для любого числа a, взаимно простого с n. Функцию λ называют *функцией Кармайкла*.

Лемма. ***Для любого натурального числа l, не кратного λ(n), существует такое взаимно простое с n целое число a, что $a^l\not\equiv 1 \pmod n$.***

Доказательство. Разделив с остатком число l на $\lambda(n)$, получаем $l=\lambda(n)q+r$, где q — целое неотрицательное число, $0<r<\lambda(n)$. При этом $a^l=(a^{\lambda(n)})^q\cdot a^r$. Поскольку $r<\lambda(n)$, хотя бы для одного взаимно простого с n числа a сравнение $a^r\equiv 1 \pmod n$ не выполнено. Это и требовалось доказать.

Теорема 5. ***Для любых взаимно простых натуральных чисел m и n***

$$\lambda(mn)=\text{НОК}[\lambda(m),\lambda(n)].$$

Доказательство. Если целое число a взаимно просто с числами m и n, то по определению $a^{\lambda(m)}\equiv 1 \pmod m$ и $a^{\lambda(n)}\equiv 1 \pmod n$, откуда для числа $k=\text{НОК}[\lambda(m),\lambda(n)]$ имеем $a^k\equiv 1 \pmod m$ и $a^k\equiv 1 \pmod n$, следовательно, $a^k\equiv 1 \pmod{mn}$, а потому и $\lambda(mn)\leqslant k$.

Докажем «от противного», что $\lambda(mn)$ делится как на $\lambda(m)$, так и на $\lambda(n)$. Пусть, например, $l=\lambda(mn)$ не делится на $\lambda(m)$. Тогда, в силу леммы, существует такое число b, взаимно простое с m, что $b^l\not\equiv 1 \pmod m$. Рассмотрим число a, для которого $a\equiv b \pmod m$ и a взаимно просто с n. (Почему такое a существует? Например, можно рассмотреть числа вида $b+mx$, где $x=1, 2, \ldots, n$. Они дают разные остатки при делении на n. Поскольку этих чисел n — столько же, сколько классов вычетов по модулю n, — то среди них найдется и нужное нам a.) Очевидно, $a^l\equiv b^l\not\equiv 1 \pmod m$. ■

Функция Кармайкла от степеней простых чисел такова: $\lambda(2)=1$, $\lambda(4)=2$, $\lambda(2^m)=2^{m-2}$ при $m\geqslant 3$, $\lambda(p^n)=p^{n-1}(p-1)$, где p — нечетное простое, n — натуральное. Это легко вывести из теоремы 4 и следующей леммы.

Лемма. ***Порядок числа 5 по модулю 2^m, где $m\geqslant 3$, равен 2^{m-2}. Порядок числа $1+p$ по модулю p^n, где p — простое, причем $p>2$, равен p^{n-1}.***

Идея доказательства. Индукцией по m проверяем, что 5^{2^m} представимо в виде $1+2^{m+2}a$, где a нечетно. Аналогично, при помощи формулы бинома Ньютона индукцией по n убеждаемся, что $(1+p)^{p^n}$ представимо в виде $1+p^{n+1}b$, где b не делится на p.

Следствие. ***Если $n=2, 4, p^m$ или $2p^m$, где p — нечетное простое, m — натуральное, то существует первообразный корень по модулю n. Для всех других натуральных чисел n верно неравенство $\lambda(n)<\varphi(n)$.*** ■

Числа Кармайкла. В силу малой теоремы Ферма, $2^{p-1}\equiv 1 \pmod p$ для любого нечетного простого числа p. Существуют ли составные числа с тем же свойством? Да, существуют:

$$2^{340}\equiv 1 \pmod{341}.$$

В самом деле, $341=11\cdot 31$, причем $2^{10}-1=1023=3\cdot 11\cdot 31$. (Можно проверить, что число 341 — наименьшее составное число n со свойством $2^{n-1}\equiv 1 \pmod n$.) Но почему мы заинтересовались именно случаем $a=2$? Наверное, разумнее спросить: существуют ли такие составные числа n, что для любого a, взаимно простого с n, выполнено сравнение $a^{n-1}\equiv 1 \pmod n$? Такие числа тоже существуют! Их называют *числами Кармайкла*. Наименьшее число Кармайкла — $561=3\cdot 11\cdot 17$, следующие за ним — $1105=5\cdot 13\cdot 17$, $1729=7\cdot 13\cdot 19$, $2465=5\cdot 17\cdot 29$, $2821=7\cdot 13\cdot 31$, $6601=7\cdot 23\cdot 41$, $8911=7\cdot 19\cdot 67$, $10\,585=5\cdot 29\cdot 73$, $15\,841=7\cdot 31\cdot 73$, $29\,341=13\cdot 37\cdot 61$, $41\,041=7\cdot 11\cdot 13\cdot 41$, ... В 1994 г. в журнале Annals of Mathematics (т. 139, с. 703—722) Э. Гренвилль, К. Померанц и У. Элфорд опубликовали (абсолютно недоступное для школьника) доказательство бесконечности множества чисел Кармайкла.

Теорема 6. ***Если составное $n=p_1^{m_1}p_2^{m_2}\cdot\ldots\cdot p_s^{m_s}$, где $p_1, p_2, \ldots, p_s$ — различные простые числа, $m_1, m_2, \ldots, m_s$ — натуральные числа, то n — число Кармайкла тогда и только тогда, когда $m_1=m_2=\ldots=m_s=1$ и $n-1$ кратно каждому из чисел $p_1-1, p_2-1, \ldots, p_s-1$.***

Следствие. ***Если n — число Кармайкла, то для любого целого числа a верно сравнение $a^n\equiv a \pmod n$.***

Доказательство теоремы 6. Пусть n — число Кармайкла. Поскольку при $n>2$ значение функции Кармайкла $\lambda(n)$ четно, то $n-1$ должно быть четным. Следовательно, n нечетно.
Поскольку $\lambda(n)$ делится на $\lambda(p_i^{m_i})=p_i^{m_i-1}(p_i-1)$, а $n-1$ не делится на p_i, то в случае $m_i>1$ получаем противоречие. Следовательно, $m_1=m_2=\ldots=m_s=1$. Завершение доказательства и вывод следствия предоставляем читателю. ■

Для любого натурального n числа 8^n+1 и $5\cdot 4^n+1$ — составные.
Дело в том, что $8^n+1=(2^n)^3+1^3$ кратно числу 2^n+1; а $5\cdot 4^n+1=5\cdot(3+1)^n+1\equiv 5+1\equiv 0 \pmod 3$. ■

Если a, b, c — натуральные числа, $b>1$, то среди чисел вида ab^n+c бесконечно много составных.
Это следует из того, что если p — простой делитель числа $ab+c$, то существует бесконечно много таких натуральных n, что $b^n\equiv b \pmod p$. ■

Для любого целого числа $k\neq -1$ существует бесконечно много натуральных чисел n, для которых число $2^{2^n}+k$ — составное. (Конечно или бесконечно множество составных чисел вида $2^{2^n}+1$, мы не знаем.)
В самом деле, пусть $k-1$ делится на 2^s и не делится на 2^{s+1}. Предположим, что при всех достаточно больших натуральных m число $p=2^{2^m}+k$ простое. Очевидно, если $2^m>s$, то $p-1=2^{2^m}+k-1=2^s h$, где h нечетно.
В силу теоремы Эйлера,

$$2^{\varphi(h)}\equiv 1 \pmod h.$$

Поэтому

$$2^{s+\varphi(h)}\equiv 2^s \pmod{2^s h}.$$

Следовательно, при $m\geqslant s$ имеем

$$2^{m+\varphi(h)}\equiv 2^m \pmod{p-1}.$$

В силу малой теоремы Ферма,

$$2^{2^{m+\varphi(h)}}+k\equiv 2^{2^m}+k\equiv 0 \pmod p.$$

Поскольку

$$2^{2^{m+\varphi(h)}}+k>2^{2^m}+k=p,$$

то число $2^{2^{m+\varphi(h)}}+k$ составное. ■

Если a, b — взаимно простые целые числа, n — натуральное, q — простое, a^n-b^n делится на q, и ни для одного отличного от n делителя m числа n разность a^m-b^m не делится на q, то $q\equiv 1 \pmod n$.
Дж. Д. Биркгоф и Г. Ш. Вандивер, используя свойства многочленов деления круга, в 1902 г. доказали, что для любых, кроме случая $a=2$, $b=1$ и $n=6$, натуральных взаимно простых чисел a и b, где $a>b$, и для любого натурального числа $n>2$ существует простой делитель q разности a^n-b^n, не являющийся делителем ни одной разности a^m-b^m, где $m<n$. ■

Квадратичный закон взаимности

Какие остатки могут давать квадраты целых чисел при делении на данное простое число? Ответить на этот и многие другие вопросы помогает квадратичный закон взаимности. Его пытались доказать Эйлер, Лагранж и Лежандр, а доказал (в 19 лет!) Гаусс. Он неоднократно возвращался к этому закону, придумав несколько разных доказательств. Мы разберем самое элементарное из них (в виде, который ему придал Фробениус).

Всюду в этой статье буква p обозначает простое число, причем $p>2$.

Символ Лежандра $\left(\frac{a}{p}\right)$. По определению, $\left(\frac{a}{p}\right)=1$, если a — квадратичный вычет по модулю p, то есть если существует такое не кратное числу p целое x, что

$$x^2\equiv a \pmod p.$$

Далее, $\left(\frac{a}{p}\right)=0$, если a кратно числу p. Наконец, $\left(\frac{a}{p}\right)=-1$, если a — квадратичный невычет по модулю p, то есть если ни для какого целого x разность x^2-a не делится на p. ■

Вычетов и невычетов — поровну. Рассмотрим числа $1^2, 2^2, \ldots, (p-2)^2, (p-1)^2$. Поскольку остатки от деления чисел x^2 и $(p-x)^2=p(p-2x)+x^2$ на p совпадают, при любом p строка остатков симметрична: читается слева направо так же, как справа налево.

А остатки от деления чисел $1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2$ на p все разные. Ведь если бы какие-то числа r^2 и s^2, где $1\leqslant r<s\leqslant\frac{p-1}{2}$, давали одинаковые остатки, то разность $s^2-r^2=(s-r)(s+r)$ делилась бы на p. Но ни $s-r$, ни $s+r$ не делятся на p. Значит, действительно существует ровно $(p-1)/2$ квадратичных вычетов и $(p-1)/2$ невычетов. ■

Возведем обе части сравнения $a\equiv x^2 \pmod p$, где x не делится на p, в $(p-1)/2$-ю степень:

$$a^{(p-1)/2}\equiv x^{p-1} \pmod p.$$

В силу малой теоремы Ферма $x^{p-1}\equiv 1 \pmod p$. Поскольку в рассматриваемом случае $\left(\frac{a}{p}\right)=1$, то

$$a^{(p-1)/2}\equiv\left(\frac{a}{p}\right) \pmod p.$$

Критерий Эйлера гласит: последнее сравнение верно для *любого* целого числа a, а не только для квадратичных вычетов. ■

Рассмотрим таблицу значений многочлена n^2-2. Среди простых делителей этих значений есть простые числа 7, 17, 23, 31, 47, 71, 79, 97, 167 и 223.

n	n^2-2
3	7
4	$14=2\cdot 7$
5	23
6	$34=2\cdot 17$
7	47
8	$62=2\cdot 31$
9	79
10	$98=2\cdot 7^2$
11	$119=7\cdot 17$
12	$142=2\cdot 71$
13	167
14	$194=2\cdot 97$
15	223

А вот числа 5 среди них нет. Для доказательства достаточно перебрать все остатки (0, 1, 2, 3 и 4), которые может дать целое число при делении на 5, вычислить для каждого из них величину n^2-2 и убедиться, что полученное значение не делится на 5. Можно чуть сократить вычисления, заметив, что по модулю 5 число n сравнимо либо с 0, либо с 1 или -1, либо с 2 или -2, но ни одно из чисел $0^2-2=-2$, $1^2-2=-1$, $2^2-2=2$ не делится на 5.

Какие же простые числа p являются делителями чисел вида n^2-2? Очевидно, $7=8-1$, $17=2\cdot 8+1$, $23=3\cdot 8-1$, $31=4\cdot 8-1$, …, $167=21\cdot 8-1$ и $223=28\cdot 8-1$. Это простые числа вида $p=8n\pm 1$.

А нечетные простые делители $p>2$ чисел вида n^2+2 — это, как можно догадаться при помощи аналогичного эксперимента, числа вида $p=8m+1$ или $p=8m+3$. Казалось бы, при чем здесь 8? . . Объяснение дает квадратичный закон взаимности, который 8 апреля 1796 г. доказал К. Ф. Гаусс. ■

Подставив в критерий Эйлера $b=-1$, имеем:

$$\left(\frac{-1}{p}\right)=(-1)^{(p-1)/2}.$$

Это равенство называется первым дополнением к квадратичному закону взаимности.

Таким образом, существование такого m, что m^2+1 делится на p, равносильно сравнению $p\equiv 1 \pmod 4$. ■

Докажем критерий Эйлера. Все $(p-1)/2$ квадратичных вычетов удовлетворяют сравнению

$$x^{(p-1)/2} \equiv 1 \pmod{p}.$$

Поскольку многочлен не может иметь больше корней, чем его степень, то никакой другой класс вычетов этому сравнению не удовлетворяет. Таким образом, для любого квадратичного невычета b имеем

$$b^{(p-1)/2} - 1 \not\equiv 0 \pmod{p}.$$

Поскольку

$$0 \equiv b^{p-1} - 1 = (b^{(p-1)/2} - 1)(b^{(p-1)/2} + 1) \pmod{p},$$

то $b^{(p-1)/2} + 1 \equiv 0 \pmod{p}$, так что

$$b^{(p-1)/2} \equiv -1 = \left(\frac{b}{p}\right) \pmod{p}.$$

Критерий Эйлера доказан. ■

Мультипликативность символа Лежандра. Для любых двух целых чисел x и y имеем

$$\left(\frac{xy}{p}\right) \equiv (xy)^{(p-1)/2} = x^{(p-1)/2} y^{(p-1)/2} \equiv \left(\frac{x}{p}\right) \cdot \left(\frac{y}{p}\right) \pmod{p}.$$

Следовательно, $\left(\frac{x}{p}\right) \cdot \left(\frac{y}{p}\right) \equiv \left(\frac{xy}{p}\right) \pmod{p}$. Поскольку $p > 2$, а символ Лежандра может равняться лишь 0, 1 или -1, то

$$\left(\frac{x}{p}\right) \cdot \left(\frac{y}{p}\right) = \left(\frac{xy}{p}\right). \blacksquare$$

Критерий Гаусса. Начнем с численного примера:

$$3^8 \cdot 8! = (3 \cdot 1)(3 \cdot 2)(3 \cdot 3)(3 \cdot 4)(3 \cdot 5)(3 \cdot 6)(3 \cdot 7)(3 \cdot 8) =$$
$$= 3 \cdot 6 \cdot 9 \cdot 12 \cdot 15 \cdot 18 \cdot 21 \cdot 24 = 3 \cdot 6 \cdot (-8) \cdot (-5) \cdot (-2) \cdot 1 \cdot 4 \cdot 7 = -8! \pmod{17},$$

откуда $3^8 \equiv -1 \pmod{17}$ и, в силу критерия Эйлера, $\left(\frac{3}{17}\right) = -1$.

Перейдем к общим рассуждениям. Числа $0, 1, 2, \ldots, p-1$ образуют полную систему вычетов по модулю p. Другими словами, любое целое число сравнимо по модулю p с одним и только одним из этих чисел. Полную систему вычетов образуют и числа $0, \pm 1, \pm 2, \ldots, \pm(p-1)/2$. Иначе говоря, любой ненулевой класс вычетов сравним по модулю p с одним из чисел от 1 до $n = (p-1)/2$ или с одним из чисел от -1 до $-n$.

Полусистемой вычетов по модулю p называют любой набор из $n = (p-1)/2$ ненулевых классов вычетов $a_1, a_2, \ldots, a_n$, обладающий тем свойством, что для любого ненулевого класса вычетов $x \bmod p$ выполнено одно из сравнений $x \equiv a_k$ или $x \equiv -a_k \pmod{p}$, где $1 \leqslant k \leqslant n$.

Другими словами, полусистема вычетов — это такое множество из n ненулевых классов вычетов, что ни при каких k и j сумма $a_k + a_j$ не делится на p. Читатель, знакомый с понятиями теории групп, скажет, что для получения полусистемы надо выбрать по одному элементу из каждого класса смежности $\mathbb{Z}_{p-1}^*/\{\pm 1\}$.

Адриен Мари Лежандр (1752—1833) — французский математик и астроном. Развивал теорию геодезических измерений, совместно с астрономами Дж. Д. Кассини и П. Ф. Мешеном определил разницу долгот обсерваторий Парижа и Гринвича, произвел проверку вычислений длины дуги меридиана между Барселоной и Дюнкером, выполненных для определения метра как единицы длины.

В области математического анализа Лежандр доказал приводимость эллиптических интегралов к канонической форме (нормальной форме Лежандра), разлагал их в ряды и составлял таблицы значений.

Своими работами Лежандр внес основополагающий вклад в создание теории чисел. Написал известный учебник геометрии (1794), в некоторых изданиях которого пытался доказать постулат о параллельных. ■

Для любого простого числа p существует такое целое число x, что

$$(x^2 - 2)(x^2 - 3)(x^2 - 6)$$

кратно p. В самом деле, для $p = 2$ или 3 годится $x = 2$ или 3 соответственно. Пусть $p > 3$. Пусть ни для какого целого числа x ни $x^2 - 2$, ни $x^2 - 3$ не делятся на p.

Тогда $\left(\frac{2}{p}\right) = -1$ и $\left(\frac{3}{p}\right) = -1$.

Следовательно,

$$\left(\frac{6}{p}\right) = \left(\frac{2}{p}\right) \cdot \left(\frac{3}{p}\right) = (-1) \cdot (-1) = 1,$$

так что существует целое число x, для которого $x^2 \equiv 6 \pmod{p}$. ■

В силу формулы бинома Ньютона и делимости на p всех чисел p-й строки треугольника Паскаля, кроме двух крайних чисел $C_p^0 = C_p^p = 1$, имеем

$$(1+i)^p = 1 + C_p^1 i + C_p^2 i^2 + \ldots + C_p^{p-2} i^{p-2} + C_p^{p-1} i^{p-1} + i^p \equiv 1 + i^p \pmod p.$$

Пусть $p = 8n+1$. Очевидно,

$$(1+i)^p \equiv 1 + i^{8n+1} = 1 + i \cdot (i^4)^{2n} = 1 + i \pmod p,$$

откуда, сокращая на $1+i$ обе части сравнения, получаем

$$(1+i)^{p-1} \equiv 1 \pmod p.$$

Поскольку $(1+i)^2 = 2i$, имеем $2^{(p-1)/2} = (2i)^{4n} = (1+i)^{4n} \equiv 1 \pmod p$. Вспомнив критерий Эйлера, мы видим, что $\left(\frac{2}{p}\right) = 1$ при $p = 8n+1$.

Аналогично можно разобрать случаи $p = 8n+3$, $8n+5$ или $8n+7$, доказав тем самым второе дополнение к квадратичному закону взаимности $\left(\frac{2}{p}\right) = (-1)^{(p^2-1)/8}$. ■

Символ Якоби $\left(\frac{a}{n}\right)$ определен для любого нечетного натурального числа $n = p_1 p_2 \ldots p_k$, где $p_1, p_2, \ldots, p_k$ — простые числа, и любого целого числа a, взаимно простого с n, формулой $\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\ldots\left(\frac{a}{p_k}\right)$.

Квадратичный закон взаимности верен и для символа Якоби. А именно, для любых нечетных взаимно простых натуральных чисел a и b имеем

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = (-1)^{\frac{a-1}{2}\cdot\frac{b-1}{2}}.$$

Верны и оба дополнения к квадратичному закону взаимности: $\left(\frac{-1}{b}\right) = (-1)^{(b-1)/2}$ и $\left(\frac{2}{b}\right) = (-1)^{(b^2-1)/8}$. (Выведите эти формулы из закона взаимности для символа Лежандра!)

Символ Якоби помогает при вычислении символа Лежандра: например,

$$\left(\frac{103}{1999}\right) = \left(\frac{1999}{103}\right)\cdot(-1)^{\frac{1999-1}{2}\cdot\frac{103-1}{2}} = -\left(\frac{1999}{103}\right) = -\left(\frac{42}{103}\right) = -\left(\frac{2}{103}\right)\times$$

$$\times\left(\frac{21}{103}\right) = -(-1)^{\frac{103^2-1}{8}}\cdot\left(\frac{103}{21}\right)\times$$

$$\times(-1)^{\frac{21-1}{2}\cdot\frac{103-1}{2}} = -\left(\frac{-2}{21}\right) = -\left(\frac{-1}{21}\right)\times$$

$$\times\left(\frac{2}{21}\right) = -(-1)^{\frac{21-1}{4}}\cdot(-1)^{\frac{21^2-1}{8}} = 1.$$ ■

Для любой полусистемы и любого ненулевого класса вычетов x составим таблицу, в верхней строке которой — полусистема вычетов $a_1, a_2, \ldots, a_n$; в средней строке — числа $xa_1, xa_2, \ldots, xa_n$, каждое из которых представлено в виде $xa_k \equiv \varepsilon_k a_{f(k)}$, где $1 \leqslant k \leqslant n$; в нижней строке — числа $\varepsilon_k = \pm 1$. В рассмотренном выше примере $p = 17$ и $x = 3$, полусистема вычетов состоит из первых 8 натуральных чисел, а таблица выглядит так:

$a_k = k$	1	2	3	4	5	6	7	8
xa_k	3	6	$9 \equiv -8$	$12 \equiv -5$	$15 \equiv -2$	$18 \equiv 1$	$21 \equiv 4$	$24 \equiv 7$
ε_k	1	1	-1	-1	-1	1	1	1

Как нетрудно доказать, для любой полусистемы вычетов $a_1, a_2, \ldots, a_n$ и любого x, не делящегося на p, числа $xa_1, xa_2, \ldots, xa_n$ тоже образуют полусистему вычетов, то есть в качестве $a_{f(k)}$, где $1 \leqslant k \leqslant n$, побывают по одному разу все числа $a_1, a_2, \ldots, a_n$. Таким образом,

$$xa_1 \cdot xa_2 \cdot \ldots \cdot xa_n \equiv \varepsilon_1 a_{f(1)} \varepsilon_2 a_{f(2)} \ldots \varepsilon_n a_{f(n)} \pmod p.$$

Сократив обе части на $a_1 a_2 \ldots a_n$, получаем: $x^n \equiv \varepsilon_1 \varepsilon_2 \ldots \varepsilon_n \pmod p$. В силу критерия Эйлера,

$$\left(\frac{x}{p}\right) \equiv x^n \equiv \varepsilon_1 \varepsilon_2 \ldots \varepsilon_n = \pm 1,$$

откуда $\left(\frac{x}{p}\right) = \varepsilon_1 \varepsilon_2 \ldots \varepsilon_n$. Это и есть критерий Гаусса. ■

Пусть p, q — простые нечетные числа, $p \neq q$. Как мы только что доказали, $\left(\frac{q}{p}\right) = (-1)^k$, где k — количество таких целых x, что $1 \leqslant x \leqslant n$ и абсолютно наименьший вычет числа qx отрицателен, то есть

$$-\frac{p}{2} < qx - py < 0$$

для некоторого целого y. Если эти неравенства выполнены для некоторого x, то величина y определена однозначно, причем $py > qx > 0$ и $py < qx + \frac{p}{2} < q \cdot \frac{p}{2} + \frac{p}{2} = p \cdot \frac{q+1}{2}$. Поскольку число y целое, то $1 \leqslant y \leqslant m = \frac{q-1}{2}$.

Поэтому можно сказать, что k есть количество пар натуральных чисел $(x; y)$, удовлетворяющих неравенствам $x \leqslant n$, $y \leqslant m$ и

$$-\frac{p}{2} < qx - py < 0.$$

Аналогично, $\left(\frac{p}{q}\right) = (-1)^K$, где K — количество пар натуральных чисел $(x; y)$, удовлетворяющих неравенствам $x \leqslant n$, $y \leqslant m$ и $-\frac{q}{2} < py - qx < 0$. Последнее неравенство можно записать в виде

$$0 < py - qx < \frac{q}{2}.$$

Поскольку равенство $qx-py=0$ при рассматриваемых значениях x и y невозможно, то

$$\left(\frac{q}{p}\right)\cdot\left(\frac{p}{q}\right)=(-1)^{k+K},$$

где $k+K$ — это количество пар натуральных чисел $(x;y)$, удовлетворяющих неравенствам $x\leqslant n$, $y\leqslant m$ и

$$-\frac{p}{2}<qx-py<\frac{q}{2}.$$

Последнее неравенство задает на координатной плоскости внутреннюю часть полосы, которая ограничена параллельными прямыми $qx-py=-\frac{p}{2}$ и $qx-py=\frac{q}{2}$. А неравенства $0<x<\frac{p+1}{2}$ и $0<y<\frac{q+1}{2}$ задают внутренность прямоугольника с центром

$$S\left(\frac{p+1}{4};\frac{q+1}{4}\right).$$

Таким образом, $k+K$ — это количество целочисленных точек, расположенных в пересечении полосы и прямоугольника. ■

Полоса симметрична относительно точки S! В этом можно убедиться, рассматривая точки пересечения прямых, ограничивающих полосу, со сторонами прямоугольника (рис. 1), или исходя из того, что средняя линия полосы (изображенная красным цветом на рисунке 1) задана уравнением

$$qx-py=\frac{1}{2}\left(-\frac{p}{2}+\frac{q}{2}\right)=\frac{q-p}{4}$$

и, поскольку

$$q\cdot\frac{p+1}{4}-p\cdot\frac{q+1}{4}=\frac{q-p}{4},$$

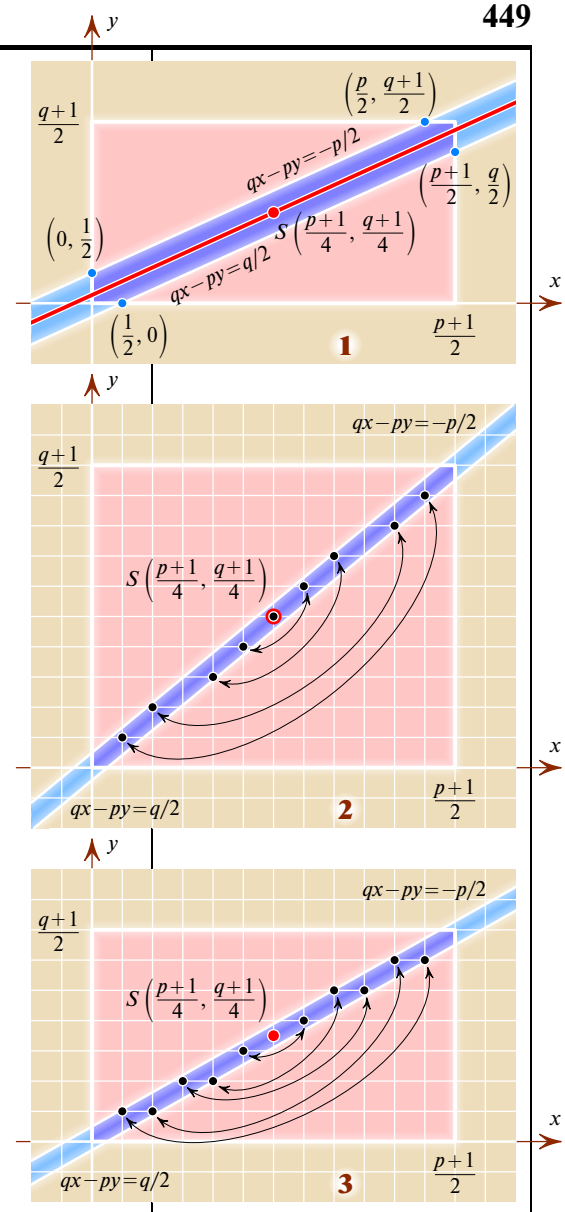
проходит через точку S.

При центральной симметрии относительно точки S любая точка $(x;y)$ с целыми координатами переходит в точку с целыми координатами

$$\left(\frac{p+1}{2}-x;\ \frac{q+1}{2}-y\right).$$

Таким образом точки разбиваются на пары, а точка S симметрична сама себе. Число $k+K$ нечетно, если обе координаты точки S целые, то есть если число $\frac{p-1}{2}\cdot\frac{q-1}{2}$ нечетно, как на рисунке 2, где $p=23$ и $q=19$. Число $k+K$ четно, если хотя бы одна координата точки S не целая, то есть если число $\frac{p-1}{2}\cdot\frac{q-1}{2}$ четно, как на рисунке 3, где $q=13$. Мы доказали квадратичный закон взаимности:

$$\left(\frac{p}{q}\right)\cdot\left(\frac{q}{p}\right)=(-1)^{k+K}=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$ ■



Фердинанд Георг Фробениус (1849—1917) — немецкий математик. В 1877 г. доказал, что всякая конечномерная ассоциативная алгебра без делителей нуля над полем вещественных чисел или одномерна (тем самым совпадая с $\mathbb{R}$), или двумерна (поле $\mathbb{C}$ комплексных чисел), или четырехмерна (тело $\mathbb{H}$ кватернионов). Ввел понятия радикала алгебры, фактор-алгебры, простой и полупростой алгебр.

Именем Фробениуса называют операцию возведения в p-ю степень каждого элемента поля характеристики p. В таком поле верно тождество $(a+b)^p=a^p+b^p$, так что отображение $x\to x^p$ является эндоморфизмом. ■

Геометрия

Ко времени появления письменности люди имели уже некоторый запас геометрических знаний, полученных в результате практической деятельности и первых теоретических обобщений. По дошедшим до нас египетским папирусам и вавилонским текстам видно, что тогда умели вычислять площадь прямоугольного треугольника и трапеции, знали хорошее приближение для длины окружности и площади круга, формулы объема куба, цилиндра, конуса и усеченной пирамиды, а также многие другие формулы, факты и методы.

Но геометрия как наука тогда не существовала, математические знания напоминали собой рецепты из кулинарной книги, их даже и излагали так, как в наши дни советы по домоводству: для решения задачи предлагали рецепт, в правильности которого можно было убедиться на конкретных примерах. Не было главного, что превращает набор фактов в стройную систему, — доказательств.

Примерно такой характер, вероятно, имели и геометрические знания в Греции VII—VI вв. до н. э. Многие сведения греки заимствовали у египтян и вавилонян. Это было время становления демократии в большинстве греческих городов-государств, время бурной общественно-политической жизни, появления научно-философских школ. Ученые впервые в истории человечества попытались понять и объяснить устройство мира не догматически-религиозно, а естественно-научно, логически, философски. Каждая из философских школ Греции старалась доказать правильность своей теории и опровергнуть оппонентов, показав, что их доводы логически противоречивы. В Греции и произошло преобразование способа изучения геометрии.

Большую роль тут сыграла пифагорейская школа. О Пифагоре мы почти ничего не знаем: его имя уже в древности было окружено фантастическими легендами. Известно лишь, что около середины VI в. до н. э. Пифагор переселился с острова Самос в Южную Италию (так называемую Великую Грецию), где находились богатые греческие города-колонии, и основал там союз, имевший и политические, и научные цели. Многие выдающиеся математики V в. до н. э. называли себя пифагорейцами. (Между прочим, теорема Пифагора была известна уже в Древнем Вавилоне.)

Пифагорейцы начали строить геометрию как абстрактную науку, изучающую общие свойства неких идеальных фигур, которые «в чистом виде» в природе не встречаются. Так в геометрии появились линии, имеющие только длину, но не ширину; точки, не имеющие ни длины, ни ширины; фигуры, составленные из таких точек, и так далее. Конечно, представление о прямой линии могло возникнуть как абстракция туго натянутой веревки, струны, луча света. Но идеальный отрезок — это не мел, не струна, даже не луч света.

Создание отвлеченных геометрических понятий было вовсе не легким делом. Далеко не все мыслители древности понимали их пользу. Например, Протагор (481—411 гг. до н. э.) не признавал геометрических абстракций. Он говорил, что никто не видел линий без ширины, не видел, чтобы круг касался линейки только в одной точке (касание всегда происходит по маленькому отрезочку!).

Непосредственным измерением невозможно проверить, что сумма величин углов любого треугольника равна 180°. Ведь любое измерение производят не точно, а с какой-то (пусть даже очень небольшой) погрешностью. Но даже если бы мы научились (с помощью идеальных инструментов) измерять углы идеальных треугольников, то и тогда мы не смогли бы доказать общую теорему: различных треугольников бесконечно много, невозможно перебрать их все.

Доказать все без исключения верные утверждения невозможно: любое доказательство должно опираться на какие-то ранее доказанные утверждения! Поэтому нужно несколько (желательно как можно меньше) утверждений принять без доказательства, а все остальные утверждения выводить из них при помощи логических правил.

Первую такую систему — «Начала» — строил еще в V в. до н. э. Гиппократ Хиосский. Было еще несколько попыток такого рода, но они все были забыты после появления «Начал» Евклида.

Историки считают, что Евклид (или группа математиков, творивших под таким псевдонимом) жил в период примерно от 330 до 275 г. до н. э. Составленные им «Начала» разделены на 13 книг, из которых пятая, седьмая, восьмая, девятая и десятая посвящены теории пропорций и арифметике (изложенным в геометрической форме), остальные являются собственно геометрическими. Многое из того, что уже знал Евклид (например, теория конических сечений), в «Началах» не изложено.

Каждая из книг «Начал» начинается с определения необходимых в ней понятий. Первой книге предпосланы 23 определения. Вот первые восемь из них.

1) Точка есть то, что не имеет частей.
2) Линия есть длина без ширины.
3) Границы линии суть точки.
4) Прямая есть такая линия, которая одинаково расположена по отношению ко всем своим точкам.
5) Поверхность есть то, что имеет только длину и ширину.
6) Границы поверхности суть линии.
7) Плоскость есть поверхность, которая одинаково расположена по отношению ко всем прямым, на ней лежащим.
8) Плоский угол есть взаимное наклонение двух встречающихся линий, расположенных в одной плоскости.

С современной точки зрения эти определения не выдерживают критики: они не используются Евклидом в дальнейших доказательствах и являются лишь описаниями геометрических образов, к тому же довольно наивными. Решил задачу аксиоматизации геометрии в 1899 г. Д. Гильберт, предложив следующую систему аксиом.

I. Аксиомы связи.

I.1. Через любые две точки можно провести хотя бы одну прямую.

I.2. Для любых двух различных точек существует не более чем одна проходящая через них прямая.

I.3. На любой прямой лежат хотя бы две точки. Существуют три точки, не лежащие на одной прямой.

I.4. Через любые три точки, не лежащие на одной прямой, проходит хотя бы одна плоскость. На каждой плоскости лежит хотя бы одна точка.

I.5. Через любые три точки, не лежащие на одной прямой, проходит не более чем одна плоскость.

I.6. Если некоторые две точки некоторой прямой принадлежат данной плоскости, то ей принадлежат все точки этой прямой.

I.7. Если две плоскости имеют общую точку, то они имеют и еще хотя бы одну общую точку.

I.8. Существуют четыре точки, не лежащие на одной плоскости.

II. Аксиомы порядка.

II.1. Если точка B лежит между точками A и C, то A, B, C — различные точки одной прямой, причем точка B лежит между точками C и A.

II.2. Каковы бы ни были точки A и C, существует по крайней мере одна такая точка B на прямой AC, что C лежит между A и B.

II.3. Среди любых трех точек не более чем одна лежит между двумя другими.

II.4 (аксиома Паша). Если точки A, B, C не лежат на одной прямой, а прямая l проходит через некоторую точку, расположенную между точками A и B, то она проходит и через некоторую точку, расположенную между точками A и C, или между точками, расположенными между B и C.

III. Аксиомы конгруэнтности.

Отрезок или угол может находиться в некотором отношении конгруэнтности к другому отрезку или углу, и это отношение удовлетворяет следующим аксиомам.

III.1. Для любых точек A и B данной прямой l и для любой точки A' любой прямой l' можно найти по данную сторону от точки A' на прямой l' одну и только одну такую точку B', что $AB \cong A'B'$. Отношение конгруэнтности симметрично: $AB \cong BA$.

III.2. Если $A'B' \cong AB$ и $A''B'' \cong AB$, то $A'B' \cong A''B''$.

III.3. Если отрезки AB и BC некоторой прямой не имеют ни одной общей внутренней точки и отрезки $A'B'$ и $B'C'$ тоже не имеют ни одной общей внутренней точки, причем $AB \cong A'B'$ и $BC \cong B'C'$, то $AC \cong A'C'$.

III.4. Каждый угол можно однозначно отложить в данной плоскости по данную сторону от данного луча. Точнее говоря, пусть дан угол $\angle(h, k)$ на плоскости α, прямая l на этой же или иной плоскости α' и определенная полуплоскость плоскости α' с граничной прямой l'. Пусть h' — луч прямой l', исходящий из точки O'. Тогда на плоскости α' существует один и только один луч k' такой, что $\angle(h, k) \cong \cong \angle(h', k')$ и все внутренние точки угла $\angle(h', k')$ лежат по заданную сторону от l'. Каждый угол конгруэнтен сам себе: $\angle(h, k) \cong \angle(h, k)$ и $\angle(h, k) \cong \angle(k, h)$.

III.5. Если точки A, B, C не лежат на одной прямой и $AB \cong A'B'$, $AC \cong A'C'$ и $\angle BAC \cong \angle B'A'C'$, то $\angle ABC \cong \angle A'B'C'$ и $\angle ACB \cong \angle A'C'B'$.

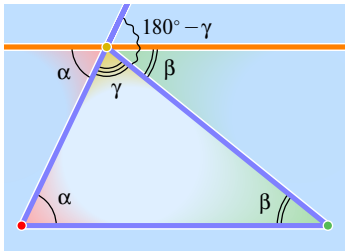
IV. Аксиомы непрерывности.

IV.1 (аксиома Архимеда). Пусть AB и CD — произвольные отрезки. Тогда на прямой AB существуют такие точки $A_1, A_2, \ldots, A_n$, что точка A_1 лежит между A и A_2, точка A_2 лежит между A_1 и A_3 и так далее, причем $AA_1 \cong A_1A_2 \cong \ldots \cong A_{n-1}A_n \cong CD$ и точка B лежит между A и A_n.

IV.2 (аксиома Кантора). Пусть на прямой задана последовательность отрезков $[A_1B_1]$, $[A_2B_2]$, $\ldots$ $\ldots$, $[A_nB_n]$, $\ldots$, где каждый следующий лежит внутри предыдущего, причем для любого заданного отрезка ϵ существует такое n, что ϵ длиннее отрезка $[A_nB_n]$. Тогда существует точка, расположенная внутри всех отрезков.

V. Аксиома параллельности. Через точку, не лежащую на данной прямой l, можно провести не более одной прямой, не пересекающей прямую l.

Вписанные углы

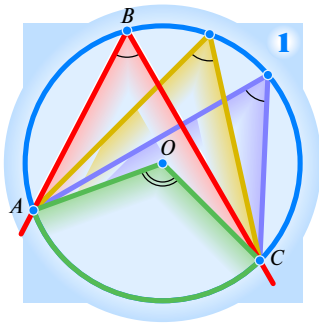


Поскольку $\alpha+\beta+\gamma=180^\circ$, то $\alpha+\beta=180^\circ-\gamma$. Величина внешнего угла $180^\circ-\gamma$ треугольника равна сумме несмежных с ним внутренних углов α и β. ■

Угол ABC называют вписанным в окружность, если его вершина B лежит на окружности, а лучи BA и BC пересекают ее (рис. 1). Теорема о вписанном угле — это равенство

$$\angle ABC=\frac{1}{2}\angle AOC,$$

где O — центр окружности. Угол AOC называют центральным, а его величину называют величиной дуги AC и иногда обозначают $\overset{\smile}{AC}$.



Докажем теорему о вписанном угле. Рассмотрим сначала случай, когда центр окружности лежит на одной из сторон угла, например, $O\in BA$ (рис. 2). Поскольку отрезки OB и OC — радиусы окружности, то $OB=OC$ и поэтому $\angle OBC=\angle OCB$. По теореме о внешнем угле $\angle AOC=$ $=\angle OBC+\angle OCB$. Следовательно, $\angle AOC=2\angle ABC$.

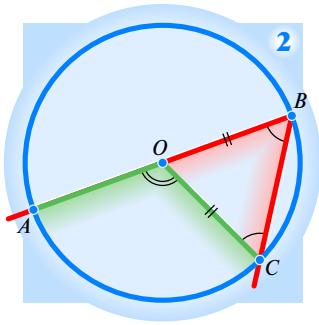
Теперь рассмотрим случай, когда точка O лежит внутри угла ABC (рис. 3). Проведем диаметр BD. Очевидно,

$$\angle AOC=\angle AOD+\angle COD=2\angle ABD+2\angle CBD=2\angle ABC.$$

Для завершения доказательства теоремы осталось рассмотреть рисунок 4, на котором точка O лежит вне угла ABC. Достаточно заменить плюс на минус:

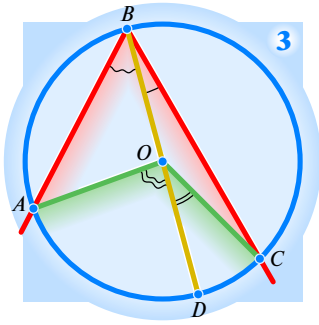
$$\angle AOC=\angle AOD-\angle COD=2\angle ABD-2\angle CBD=2\angle ABC.$$

Теорема о вписанном угле доказана. ■



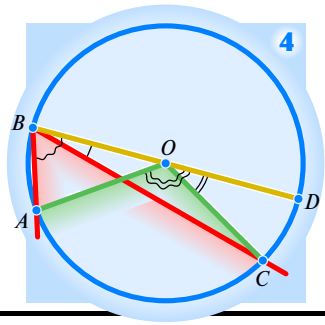
Следствие — теорема об угле между хордой и касательной (рис. 5):

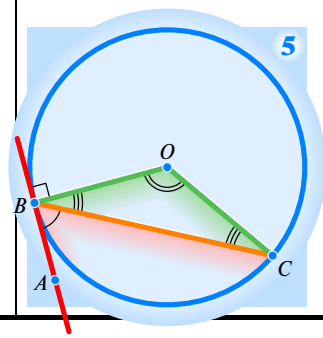
$$\angle ABC=90^\circ-\angle OBC=$$
$$=90^\circ-\frac{1}{2}(\angle OBC+\angle OCB)=90^\circ-\frac{1}{2}(180^\circ-\angle BOC)=\frac{1}{2}\angle BOC.\ \blacksquare$$

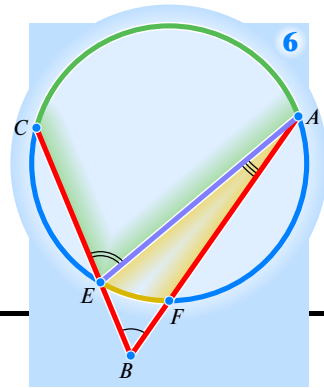


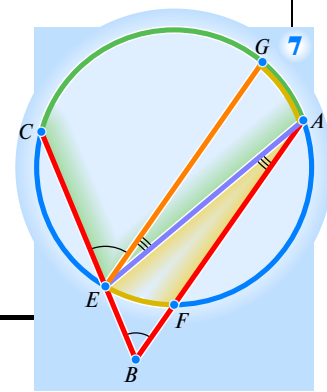
Если вершина угла лежит вне окружности, а его стороны пересекают окружность, то величина угла равна полуразности величин высекаемых на окружности дуг. Доказательство (рис. 6) основано на теореме о внешнем угле:

$$\angle ABC=\angle AEC-\angle EAB=\frac{1}{2}(\overset{\smile}{AC}-\overset{\smile}{EF}).$$





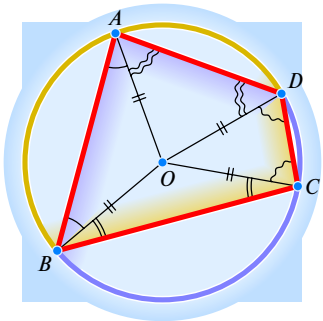




Четырехугольник можно вписать в окружность тогда и только тогда, когда сумма величин его противоположных углов равна 180°. Действительно, если четырехугольник вписан в окружность, то дуги, на которые опираются два его противоположных угла, составляют полную окружность, ее угловая величина — 360°, следовательно, сумма величин противоположных углов равна 180°. Обратно, пусть сумма величин углов A и C четырехугольника $ABCD$ равна 180°. Опишем окружность вокруг треугольника ABD. Угловая величина ее дуги BAD равна

$$\overset{\smile}{BAD} = 360^\circ - 2\angle BAD = \\ = 2(180^\circ - \angle BAD) = 2\angle BCD,$$

а из равенства $\overset{\smile}{BAD} = 2\angle BCD$, как доказано выше, следует, что точка C принадлежит окружности. ■



Соединив вершины вписанного четырехугольника с центром описанной окружности, мы получаем равнобедренные треугольники AOB, BOC, COD и DOA. В случае, когда точка O лежит внутри четырехугольника, имеем:

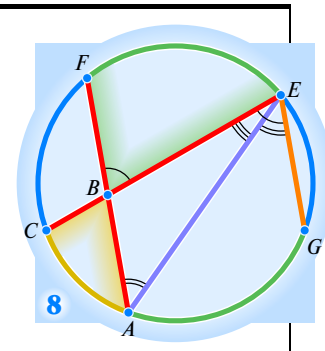
$$\angle DAB + \angle BCD = \angle DAO + \\ + \angle OAB + \angle BCO + \angle OCD = \\ = \angle ADO + \angle OBA + \\ + \angle CBO + \angle ODC = \\ = \angle ADO + \angle ODC + \angle OBA + \\ + \angle CBO = \angle ADC + \angle ABC.$$

Поскольку сумма величин углов четырехугольника равна 360°, а мы доказали, что суммы противоположных углов равны, то каждая из них равна 180°. Мы доказали свойство вписанного четырехугольника без помощи теоремы о вписанном угле! (Случай, когда центр окружности лежит вне четырехугольника, рассматривается аналогично.) ■

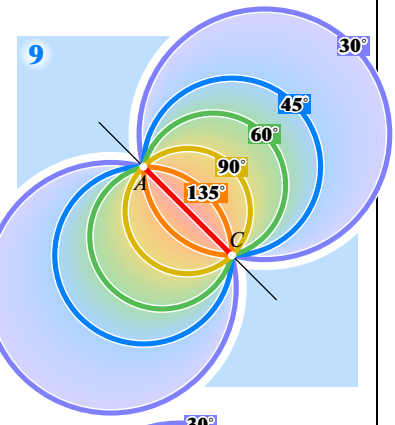
Есть и другой способ (рис. 7): проведя прямую EG параллельно прямой BA, видим, что $\overset{\smile}{AG} = \overset{\smile}{EF}$. Следовательно,

$$\angle ABC = \angle GEC = \frac{1}{2}\overset{\smile}{CG} = \frac{1}{2}(\overset{\smile}{CA} - \overset{\smile}{AG}) = \frac{1}{2}(\overset{\smile}{CA} - \overset{\smile}{EF}),$$

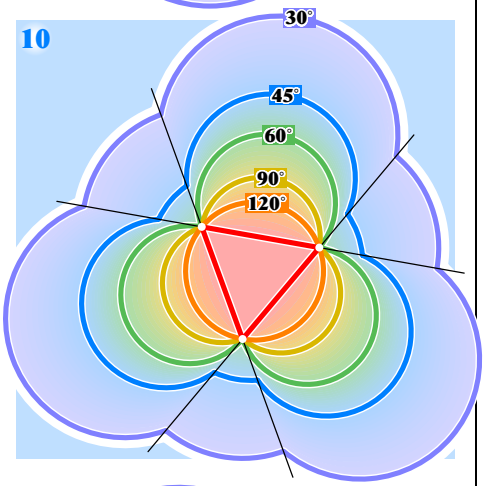
что и требовалось доказать. ■

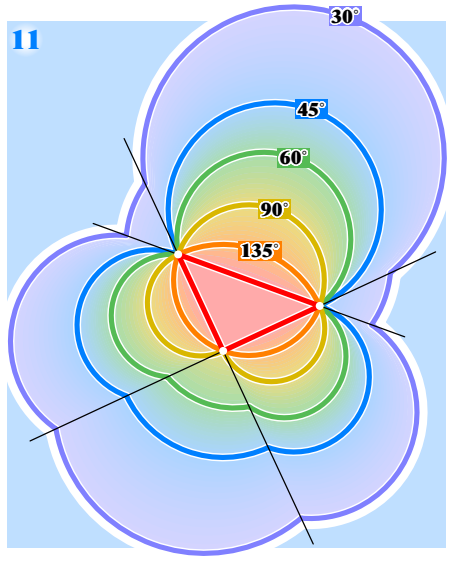


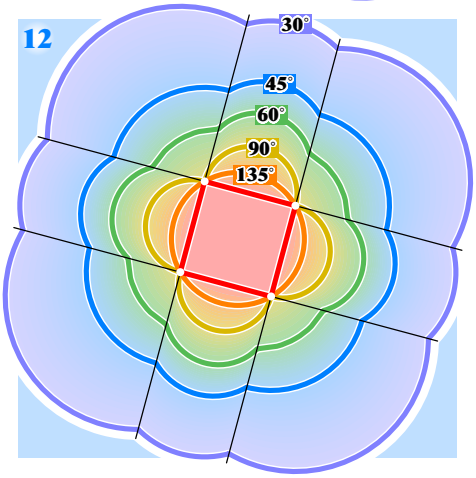
Если вершина внутри окружности, то величина угла равна полусумме величин дуг, высекаемых на окружности им и вертикальным углом. Доказательство можно провести любым из двух рассмотренных способов (рис. 8). ■



«Уши Чебурашки» — это множество точек, из которых данный отрезок AC виден под данным углом φ (рис. 9). Из теоремы о вписанном угле и только что доказанных теорем следует, что это множество — объединение двух дуг окружностей, симметричных относительно прямой AC (сами точки A и C множеству не принадлежат). Если $\varphi = 90^\circ$, то уши Чебурашки — это окружность с диаметром AC, из которой выколоты точки A и C. Множество точек, из которых данный многоугольник виден под данным углом, — объединение нескольких дуг окружностей. На рисунках 10—12 показаны такие множества для правильного треугольника, равнобедренного прямоугольного треугольника и квадрата. ■





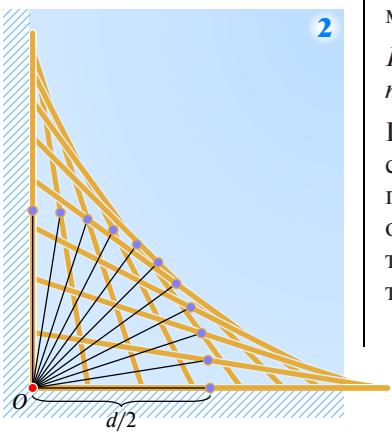


Котенок на лестнице

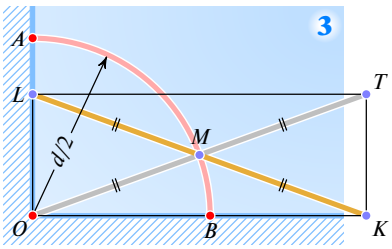
Несколько совершенно разных на первый взгляд сюжетов — задача о траектории середины падающей лестницы, теорема Коперника, разные способы задания астроиды — неожиданно тесно связаны между собой.

На гладком полу у стены стояла лестница и стала скользить вниз (рис. 1). По какой линии движется котенок, сидящий на середине лестницы? Пусть котенок флегматичный и сидит смирно. Тогда получаем такую математическую задачу.

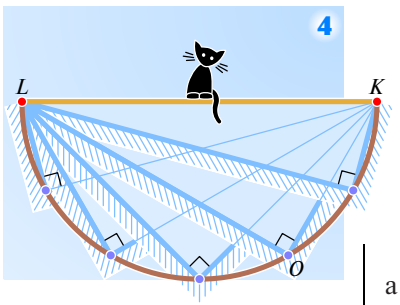
Найти множество середин всевозможных отрезков данной длины d, концы которых лежат на сторонах данного прямого угла.



Попробуем догадаться, что это за множество. Разумеется, когда отрезок скользит концами по сторонам угла, его середина описывает некоторую непрерывную линию. Концы этой линии соответствуют крайним положениям отрезка: вертикальному и горизонтальному. Построим несколько промежуточных положений отрезка (рис. 2). Если сделаем это достаточно аккуратно, то увидим, что все середины находятся на одинаковом расстоянии $d/2$ от вершины O данного угла. Возникает гипотеза: искомая линия — четверть окружности радиуса $d/2$ с центром O.

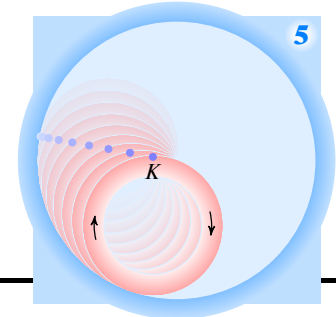


Доказать гипотезу нетрудно, достроив треугольник OKL до прямоугольника $OKTL$ (рис. 3). Диагонали OT и KL равны по длине и точкой пересечения M делятся пополам. Таким образом, середина M отрезка KL лежит на дуге AB окружности с центром O. Эта дуга и есть искомое множество точек. (Заметьте: если лестница OT уперлась одним концом в вершину прямого угла и падает, то траектория сидящего на ее середине котенка — та же самая четверть окружности! Впрочем, при таком падении лестница придавит котенка, в то время как лестница KL оказывалась под ним.)

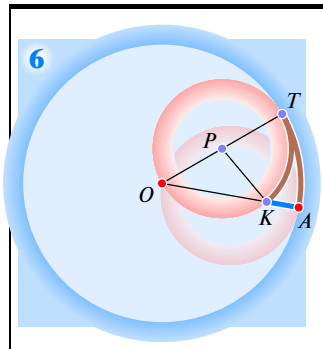


Строго говоря, нужно еще доказать, что любая точка M дуги AB принадлежит искомому множеству. Это нетрудно сделать: через любую точку M дуги AB можно провести луч OM, отложить на нем отрезок $MT = OM$, опустить из точки T перпендикуляры TL и TK на стороны угла — и отрезок KL с серединой M построен. (Это построение выглядит излишним: поскольку точка M описывает «непрерывную линию» с концами A и B, то точка M проходит всю дугу AB, а не какую-то ее часть. По сути это верно, но чтобы так рассуждать, надо владеть понятием «непрерывная функция».) ■

С точки зрения котенка движение лестницы выглядит следующим образом. В его системе координат отрезок KL (лестница) закреплен, а лучи OK и OL вращаются, оставаясь перпендикулярными. Тот факт, что расстояние от середины отрезка длины d до вершины O прямого угла равно $d/2$, превращается в известную теорему: для любых точек K и L плоскости множество точек O, для которых $\angle KOL = 90^\circ$, — окружность с диаметром KL (за вычетом самих точек K и L). ■



Рассмотрим задачу, на первый взгляд никак не связанную с котенком на лестнице. *Пусть по неподвижной окружности, касаясь ее изнутри, катится без скольжения окружность вдвое меньшего радиуса* (рис. 5). *Какова траектория точки K подвижной окружности*?



Ответ на удивление простой: точка K движется по прямой, точнее, по диаметру неподвижной окружности. Прежде чем читать дальше, попробуйте убедиться на опыте в справедливости этой теоремы Коперника. (При этом важно, чтобы внутренний круг катился без скольжения, то есть чтобы длины прокатившихся одна по другой дуг были равны.) ■

Ее нетрудно и доказать.

Пусть подвижная окружность с центром P касается неподвижной в точке T, а точка K подвижной окружности занимала в начальный момент положение A (рис. 6). Длины дуг TA и TK равны. Поскольку радиус подвижной окружности вдвое меньше радиуса неподвижной окружности, то

$$\angle TPK = 2\angle TOA.$$

В силу теоремы о вписанном угле

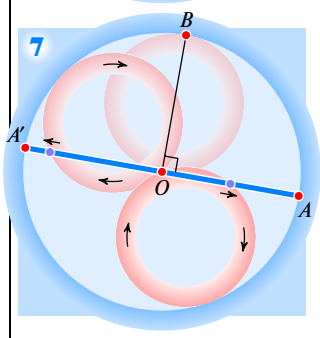
$$\angle TPK = 2\angle TOK.$$

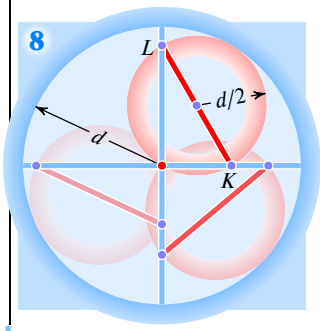
Следовательно,

$$\angle TOA = \angle TOK,$$

то есть точки O, K и A лежат на одной прямой.

Это рассуждение годится вплоть до момента, когда подвижный круг прокатится по четверти большой окружности, точка касания попадет в точку B неподвижной окружности, а точки K и O совпадут (рис. 7). Дальнейшее движение происходит симметрично относительно прямой BO. Когда точка K пройдет весь диаметр AA', подвижный круг начнет движение по нижней половине неподвижной окружности и точка K будет двигаться от A' к A. ■

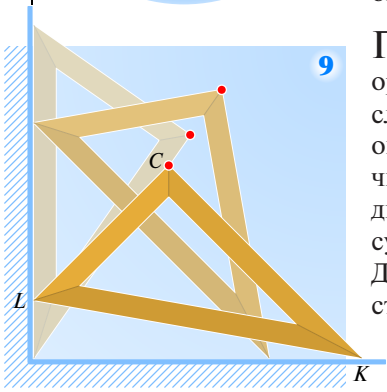




Привлекательность

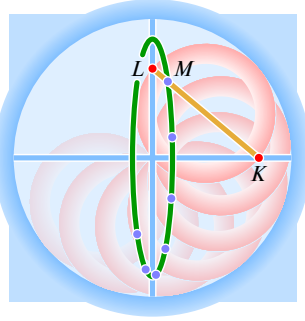
задачи о котенке и теоремы Коперника связана с тем, что в довольно сложных движениях траектории некоторых точек оказываются неожиданно простыми. Но эти задачи связаны не только внешней красотой: движения, рассматриваемые в них, по существу совпадают!

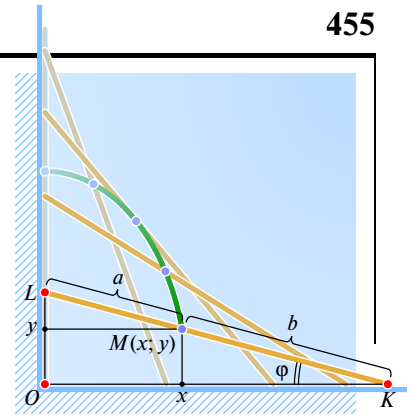
Действительно, пусть по окружности радиуса d катится изнутри окружность радиуса $d/2$, и пусть KL — диаметр этой окружности, жестко связанный с ней. Согласно теореме Коперника точки K и L движутся по неподвижным прямым — диаметрам AA' и BB' большой окружности (рис. 8); диаметр KL при этом скользит своими концами по двум взаимно перпендикулярным прямым. ■



Рассмотрим прямоугольный треугольник KLC,

вершины острых углов которого скользят по сторонам данного прямого угла (рис. 9). Какова траектория точки C?





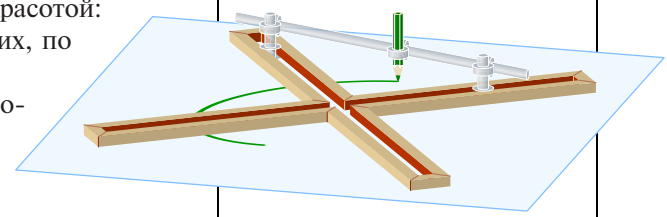
По какой линии будет двигаться котенок, если он сидит не на середине лестницы? На рисунке построено несколько точек одной из таких линий. Видно, что это — не прямая и не окружность. Выясним, что это за кривая. Введем систему координат, взяв в качестве осей стороны прямого угла.

Пусть котенок сидит в точке $M(x; y)$ на расстоянии a от точки L и b — от точки K. Найдем уравнение, связывающее x и y. Если $\angle LKO = \varphi$, то $x = a\cos\varphi$ и $y = b\sin\varphi$, поэтому

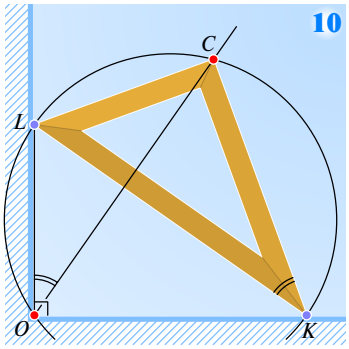
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \cos^2\varphi + \sin^2\varphi = 1.$$

Мы получили каноническое уравнение эллипса. (Между прочим, при $a = b = d/2$ — что соответствует ситуации, когда котенок сидит на середине лестницы, — уравнение превращается в уравнение окружности $x^2 + y^2 = (d/2)^2$.)

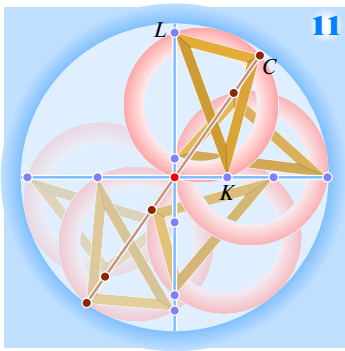
Полученный результат объясняет устройство эллипсографа Леонардо да Винчи — механизма, вычерчивающего эллипсы. ■



В теореме Коперника отмеченная точка лежит на движущейся окружности. А какова траектория точки M, отмеченной не на границе, а внутри движущегося круга? Ответ — эллипс! Для доказательства проведем через точку M диаметр KL движущейся окружности. В силу теоремы Коперника точки K и L движутся по диаметрам неподвижной окружности. А точка M — котенок на лестнице KL. ■



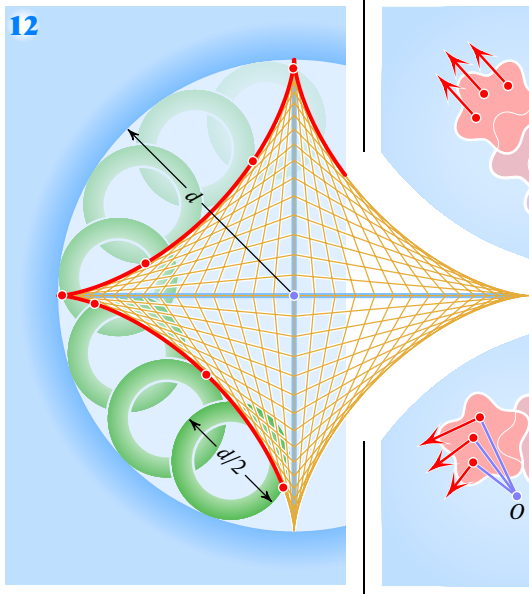
Поскольку $\angle KOL = \angle KCL = 90°$, то четырехугольник $KOLC$ вписанный и, значит, $\angle LOC = \angle LKC$ (рис. 10). Поскольку величина угла LKC треугольника KLC не меняется при его движении, то точка C движется вдоль некоторой прямой, проходящей через начало координат — точку O. Задача решена быстро и естественно? Да. Но посмотрите на рисунок 11: очевидно, при движении треугольника KCL каждая из трех его вершин движется вдоль своего диаметра! ■

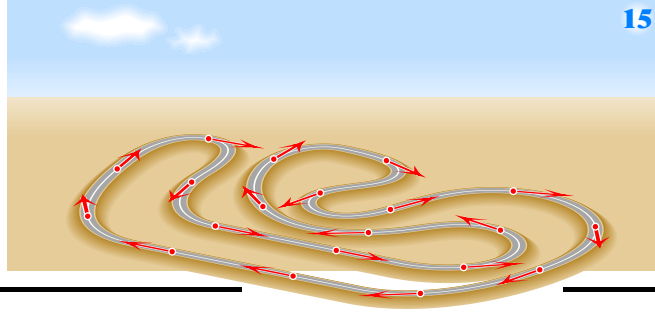


Какое множество точек заметает отрезок KL, то есть каково объединение всевозможных его положений (рис. 12)? Кривую, ограничивающую это множество, называют астроидой. Оказывается, ее можно получить так: заставить круг радиуса $d/4$ катиться изнутри по окружности радиуса d и нарисовать траекторию некоторой точки границы катящегося круга — эта траектория и есть астроида.

Мы докажем это утверждение при помощи важного понятия механики — мгновенного центра вращения. Чтобы познакомиться с этим понятием, сравним два рисунка: рисунок 13, на котором изображено поле скоростей поступательно движущегося тела и говорить о вращении бессмысленно, и рисунок 14, где движение вращательное, а мгновенный центр вращения — точка O. Очевидно, если колесо катится без проскальзывания по (не обязательно ровной!) дороге, то в каждый момент времени мгновенный центр вращения — точка соприкосновения колеса с дорогой.

Далее, при движении точки по любой кривой вектор скорости, очевидно, касается траектории (рис. 15). Рассмотрев круглое колесо, катящееся по дороге (рис. 16), видим, что вектор скорости касается траектории и перпендикулярен хорде окружности, соединяющей рассматриваемую точку с точкой касания колеса с дорогой. Осталось сделать всего лишь два шага. Рассматривая катящиеся без проскальзывания по дороге два круглых колеса, диаметр меньшего из которых равен радиусу большего (рис. 17), приходим к выводу, что желтый диаметр касается синей линии — траектории отмеченной точки. Наконец, заменив произвольную дорогу на окружность радиуса d, внутри которой без проскальзывания катятся окружности радиусов $d/2$ и $d/4$, как показано на рисунке 18, завершаем доказательство. ■





Выведем уравнение астроиды. Обозначив $\angle TQX = \varphi$ (рис. 19), имеем, в силу равенства длин дуг MT и AT, равенство $\angle MQT = 4\varphi$, откуда $\angle MQX = 3\varphi$. Поскольку $\overrightarrow{OM} = \overrightarrow{OQ} + \overrightarrow{QM}$, то

$$\overrightarrow{OM} = \frac{3d}{4}(\cos\varphi;\ \sin\varphi) + \frac{d}{4}(\cos 3\varphi;\ -\sin 3\varphi) =$$

$$= \frac{d}{4}(3\cos\varphi + \cos 3\varphi;\ 3\sin\varphi - \sin 3\varphi).$$

Применив формулы

$$\cos 3\varphi = 4\cos^3\varphi - 3\cos\varphi,$$
$$\sin 3\varphi = 3\sin\varphi - 4\sin^3\varphi,$$

получаем $\overrightarrow{OM} = d(\cos^3\varphi;\ \sin^3\varphi)$. Следовательно, для координат $(x;\ y)$ точки M имеем

$$x^{2/3} + y^{2/3} = d^{2/3}(\cos^2\varphi + \sin^2\varphi) = d^{2/3}.$$ ■

Уравнение астроиды можно получить и другим способом: как уравнение огибающей. Идея простая. Если для некоторой гладкой кривой мы знаем уравнения ее касательных, то можно рассмотреть точку пересечения P двух таких касательных (рис. 20) и устремить касательную m к неподвижной касательной n. При этом точка P стремится к точке N (строго говоря, это верно не для всех кривых: например, нельзя позволить кривой содержать внутри себя отрезок прямой; но мы не будем вдаваться в такие детали).

Итак, для вывода уравнения астроиды надо (рис. 21) решить систему уравнений

$$\begin{cases} \dfrac{x}{d\cos\varphi}+\dfrac{y}{d\sin\varphi}=1, \\ \dfrac{x}{d\cos\psi}+\dfrac{y}{d\sin\psi}=1 \end{cases}$$

и затем устремить ψ к φ. Умножая первое уравнение на $d\sin\varphi$, второе — на $d\sin\psi$, а затем вычитая первое из второго, находим

$$x=\frac{d(\sin\varphi-\sin\psi)\cos\varphi\cos\psi}{\sin\varphi\cos\psi-\cos\varphi\sin\psi}=\frac{d(\sin\varphi-\sin\psi)\cos\varphi\cos\psi}{\sin(\varphi-\psi)}=$$

$$=\frac{d\cdot 2\sin\frac{\varphi-\psi}{2}\cos\frac{\varphi+\psi}{2}\cos\varphi\cos\psi}{2\sin\frac{\varphi-\psi}{2}\cos\frac{\varphi-\psi}{2}}=\frac{d\cos\frac{\varphi+\psi}{2}\cos\varphi\cos\psi}{\cos\frac{\varphi-\psi}{2}}.$$

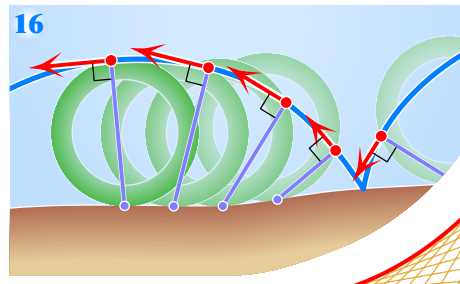
При $\psi\to\varphi$, очевидно, $\cos\frac{\varphi+\psi}{2}\to\cos\varphi$, $\cos\psi\to\cos\varphi$ и $\cos\frac{\varphi-\psi}{2}\to 1$. Поэтому $x\to d\cos^3\varphi$. Аналогично, $y\to d\sin^3\varphi$. (Повторять вычисления не обязательно: можно подставить найденное значение $x=d\cos^3\varphi$ в первое уравнение системы.) Параметрическое уравнение астроиды найдено:

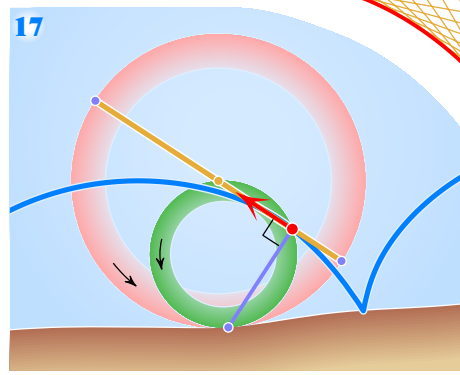
$$(x;\,y)=(d\cos^3\varphi;\,d\sin^3\varphi).$$

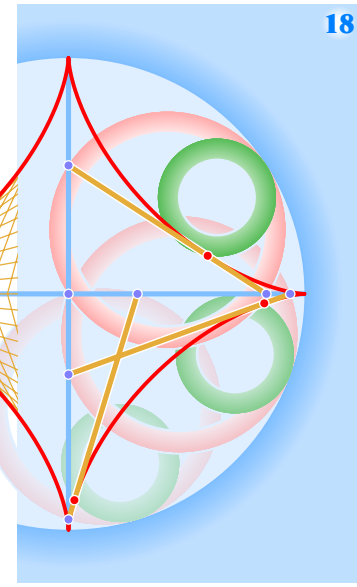
Оно, как нетрудно проверить, равносильно уравнению

$$(d^2-x^2-y^2)^3=27d^2x^2y^2,$$

так что астроида — множество решений уравнения шестой степени. ■

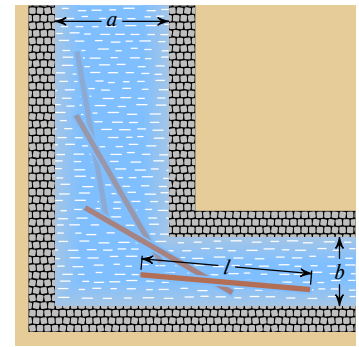


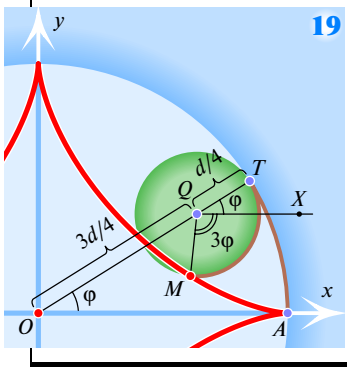


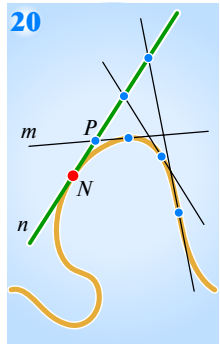


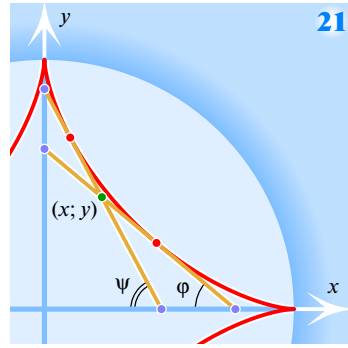
Бревно длины l может проплыть поворот канала, берега которого — параллельные прямые, поворачивающие под прямым углом, причем ширина до поворота равна a, а после поворота — b, если и только если $l^{2/3}\leqslant a^{2/3}+b^{2/3}$.

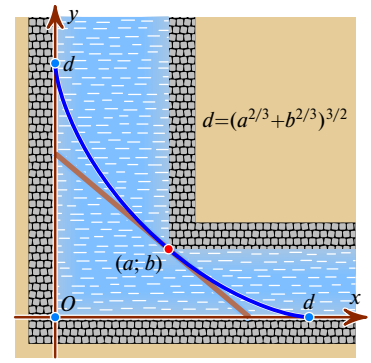
Длина кратчайшего отрезка с концами на сторонах данного прямого угла, проходящего через точку $(a;\,b)$, равна $(a^{2/3}+b^{2/3})^{3/2}$. Этот отрезок касается астроиды, проходящей через точку $(a;\,b)$. ■



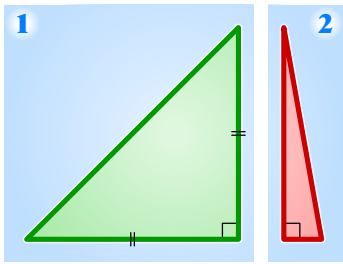








Самый произвольный треугольник

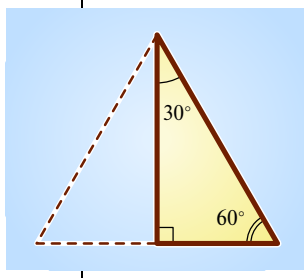


Идею этой статьи можно выразить фразой, помогающей разоблачить агента КГБ: «Он настолько типичен, что этим выделяется». Вряд ли у кого язык повернется назвать произвольным равнобедренный прямоугольный треугольник (рис. 1). Если же один из катетов очень мал (рис. 2), то гипотенуза близка к другому катету, так что треугольник… тоже почти равнобедренный!
Произвольный треугольник должен быть «средним», «типичным», удобным для работы. Как придать этим словам смысл?!

Длины сторон треугольника ограничены размерами листа бумаги. И хотя полезно рисовать большие чертежи, форма треугольника определена величинами его углов. Разобьем задачу поиска самого произвольного равнобедренного треугольника на три в зависимости от того, угол при вершине 1) тупой; 2) меньше угла при основании; 3) больше угла при основании.

1) Пусть α — угол при основании, $\beta = 180^\circ - 2\alpha$ — угол при вершине, причем $\beta > 90^\circ$. Тогда $0^\circ < \alpha < 45^\circ$. Чтобы углы α и β треугольника как можно сильнее отличались друг от друга, нужно выбрать $\alpha \in (0^\circ; 45^\circ)$ так, чтобы была максимальна величина $\beta - \alpha = 180^\circ - 3\alpha$.

Чем меньше α, тем больше $180^\circ - 3\alpha$, но выбрать $\alpha = 0^\circ$ нельзя! Что же получается? Самый произвольный тупоугольный равнобедренный треугольник не существует, и мы можем лишь приблизительно нарисовать его, взяв $\alpha \approx 0^\circ$? Да нет, лучше не позволять углу α быть слишком маленьким. (Можно еще потребовать, чтобы β не слишком приближалось к 180°, но это получится само собой, поскольку $180^\circ - \beta = 2\alpha$.) Итак, ищем $\alpha \in (0^\circ; 45^\circ)$, чтобы наименьшая из величин α и $180^\circ - 3\alpha$ была как можно больше.



Поскольку при $0^\circ < \alpha < 45^\circ$ верно неравенство $\alpha < 180^\circ - 3\alpha$, ответа опять нет: можно брать α сколь угодно близким к 45°, но в точности 45° — нельзя, треугольник получается прямоугольным, а не тупоугольным.

Изменим требования: пусть наименьшая из величин α и $\beta - 90^\circ = 90^\circ - 2\alpha$ будет как можно больше. График функции $y = \min(x, 90 - 2x)$ показан на рисунке 3. Максимальное значение достигается при $x = 30$.

Итак, задаче о равнобедренном тупоугольном треугольнике мы дали три разные формулировки. В двух из них определенного ответа нет, а в третьей $\alpha = 30^\circ$ и $\beta = 120^\circ$ (рис. 4). Является ли этот ответ единственно правильным? Нет, возможна ситуация, когда нас будут интересовать другие величины, и тогда задачу придется ставить и решать заново.

2) Пусть $\alpha > \beta$, где α — угол при основании, а β — угол при вершине равнобедренного треугольника. Тогда $\beta < 60^\circ$. Если интересоваться величиной

$$\max_{\substack{0^\circ<\beta<\alpha,\\ \alpha=(180^\circ-\beta)/2}} \min(\beta, \alpha-\beta) = \max_{0^\circ<\beta<60^\circ} \min\left(\beta, 90^\circ - \frac{3}{2}\beta\right),$$

Рассмотрим прямоугольный треугольник с острыми углами α и β, где $\alpha \leqslant \beta$. Пусть α как можно сильнее отличается от β, а β — от 90°. Иначе говоря, пусть разности $\beta - \alpha$ и $90^\circ - \beta$ будут побольше. (Может быть, пусть еще и угол α не будет слишком мал? Нет, $\alpha = 90^\circ - \beta$, ничего нового это не даст.)

Как следить за несколькими величинами сразу? Что лучше: $\alpha = 11^\circ$ и $\beta = 79^\circ$, когда $\beta - \alpha = 68^\circ$ и $90^\circ - \alpha = 11^\circ$, или $\alpha = 28^\circ$ и $\beta = 62^\circ$, когда разности равны 34° и 28°? Будем следить за минимальной из величин, то есть будем искать α и β, для которых максимально велика величина

$$\min_{0<\alpha\leqslant\beta<90^\circ}(\beta-\alpha, 90^\circ-\beta) = \min_{\alpha\in(0^\circ;45^\circ]}(90^\circ-2\alpha, \alpha).$$

При $0^\circ < \alpha \leqslant 30^\circ$ имеем

$$\min(\alpha, 90^\circ - 2\alpha) \leqslant \alpha \leqslant 30^\circ,$$

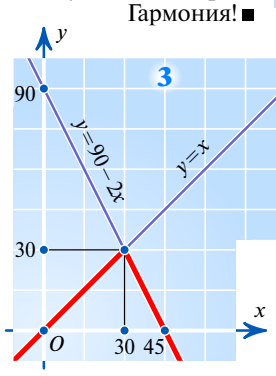
а при $30^\circ < \alpha \leqslant 45^\circ$ имеем

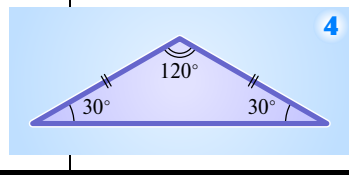
$$\min(\alpha, 90^\circ - 2\alpha) \leqslant 90^\circ - 2\alpha < 30^\circ.$$

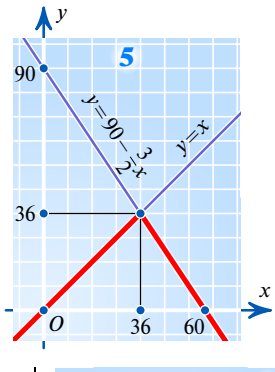
Следовательно,

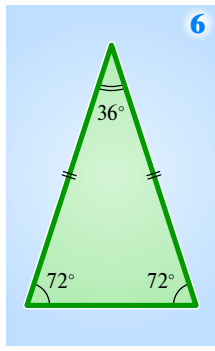
$$\max_{0^\circ<\alpha\leqslant45^\circ} \min(\alpha, 90^\circ - 2\alpha) = 30^\circ,$$

причем максимум достигается при $\alpha = 30^\circ$. Треугольник с углами 30°, 60° и 90° — половина равностороннего треугольника. Он не низок, не высок, он не узок, не широк. Гармония! ■





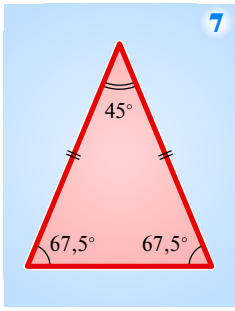




то ответ легко найти, построив график $y=\min\left(x, 90-\frac{3}{2}x\right)$ (рис. 5). Максимум достигается при $x=36$ (рис. 6). Если же мы рассмотрим

$$\max_{\substack{0^\circ<\beta<60^\circ\\ \alpha=(180^\circ-\beta)/2}} \min(\beta, \alpha-\beta, 90^\circ-\alpha, 90^\circ-\beta),$$

то, построив график, видим, что максимум достигается при $\beta=45^\circ$ (рис. 7). Какой из треугольников, изображенных на рисунках 6 и 7, милее вашему сердцу? Нам больше нравится второй, поскольку он сильнее отличается от прямоугольного. Впрочем, возможны другие постановки задачи, так что все зависит от выбранного критерия!



3) Мы ищем «самый произвольный» остроугольный равнобедренный треугольник, угол при вершине которого больше угла при основании: $90^\circ>180^\circ-2\alpha>\alpha$, то есть $45^\circ<\alpha<60^\circ$. Если рассмотреть величину

$$\max_{\substack{45^\circ<\alpha<60^\circ\\ \beta=180^\circ-2\alpha}} (\beta-\alpha),$$

то ничего интересного не увидим: максимум не достигается, поскольку точка $\alpha=45^\circ$ не входит в рассматриваемый интервал. А вот если рассмотреть

$$\max_{\substack{45^\circ<\beta<60^\circ,\\ \alpha=(180^\circ-\beta)/2}} \min(\beta, \beta-\alpha, 90^\circ-\alpha, 90^\circ-\beta),$$

то максимум достигается при $\alpha=54^\circ$ и $\beta=72^\circ$ (рис. 8). ■

Неравнобедренный треугольник.

Обозначим величины углов буквами α, β и γ. Пусть для определенности $0^\circ<\alpha\leqslant\beta\leqslant\gamma<180^\circ$. Постараемся максимизировать величину

$$\Delta=\min(\alpha, \beta-\alpha, \gamma-\beta, 180^\circ-\gamma).$$

Заменим γ на $180^\circ-\alpha-\beta$. Очевидно,

$$\Delta=\min(\alpha, \beta-\alpha, 180^\circ-\alpha-2\beta, \alpha+\beta),$$

а условие $\beta\leqslant\gamma$ принимает вид $\alpha+2\beta\leqslant180^\circ$. На координатной плоскости множество точек $(\alpha;\beta)$, удовлетворяющих последнему неравенству и неравенствам $0<\alpha\leqslant\beta$, — треугольник OKL (рис. 9).

Очевидно, $\beta+\alpha>\alpha$. Выясним, для каких точек треугольника OKL какое из чисел α, $\beta-\alpha$, $180^\circ-\alpha-2\beta$ минимально. Для этого воспользуемся чем-то вроде «метода интервалов» — посмотрим, где эти величины равны, то есть нарисуем прямые, заданные уравнениями $\alpha=\beta-\alpha$, $\beta-\alpha=180^\circ-\alpha-2\beta$ и $\alpha=180^\circ-\alpha-2\beta$. Они делят треугольник OKL на треугольники OLM, OMK и KLM. Как нетрудно сообразить (проанализировав неравенства или произвольно выбрав в каждом из этих треугольников по точке), в треугольниках OLM, OMK и KML из величин α, $\beta-\alpha$, $180^\circ-\alpha-2\beta$ минимальной является соответственно α, $\beta-\alpha$ и $180^\circ-\alpha-2\beta$ (рис. 10).

Максимальное значение функция $\Delta(\alpha, \beta)$ принимает в точке $M(30^\circ; 60^\circ)$.

На роль «самого неравнобедренного» треугольника претендует прямоугольный треугольник! Хорошо ли это? В некоторых случаях, наверное, хорошо. Но чаще всего — плохо! Постараемся, чтобы треугольник не был прямоугольным.

При поиске самого произвольного прямоугольного треугольника мы рассматривали величину $\min(\alpha, 90^\circ-2\alpha)$. Рассмотрим сумму квадратов

$$\alpha^2+(90-2\alpha)^2.$$

Трехчлен

$$5\alpha^2-360\alpha+8100$$

принимает свое наименьшее значение при $\alpha=\frac{360^\circ}{2\cdot5}=36^\circ$. Вот еще один претендент на роль самого произвольного прямоугольного треугольника! ■

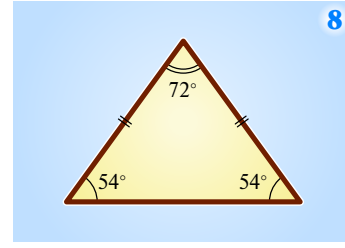
При поиске наипроизвольнейшего равнобедренного треугольника можно рассмотреть сумму квадратов

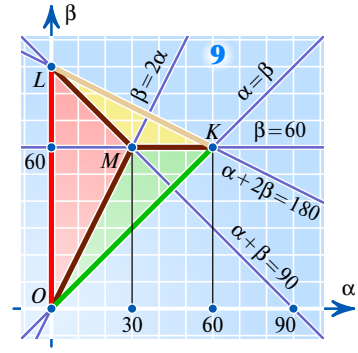
$$(\beta-\alpha)^2+\alpha^2=(180-3\alpha)^2+\alpha^2.$$

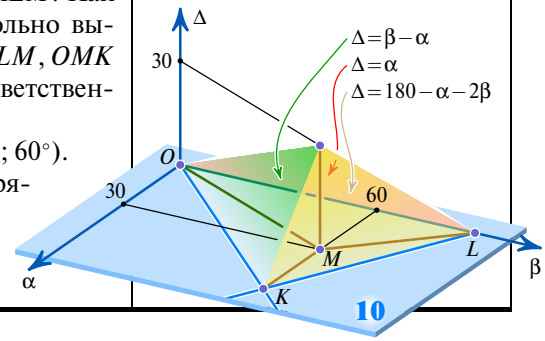
Квадратный трехчлен

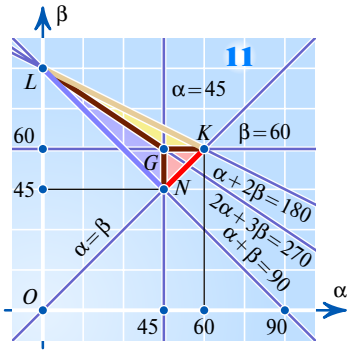
$$10\alpha^2-6\cdot180\alpha+32400$$

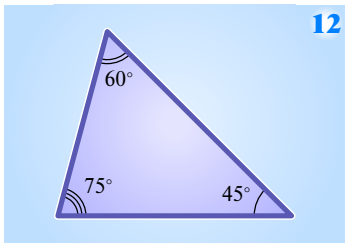
принимает наименьшее значение при $\alpha=\frac{6\cdot180^\circ}{2\cdot10}=54^\circ$. ■

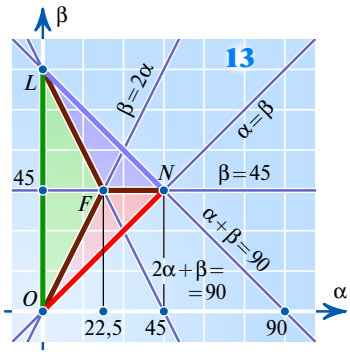


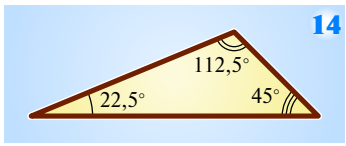


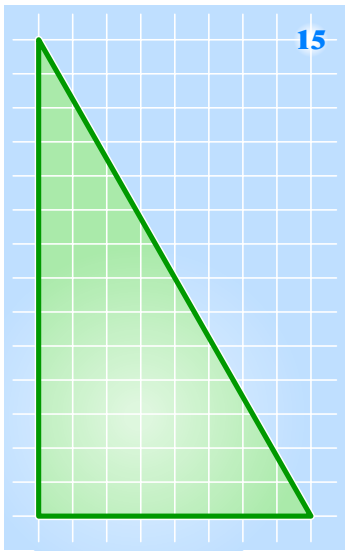












Очевидно, $\alpha \leqslant 60^\circ$. Поэтому за «непрямоугольность» отвечают величины $90^\circ - \beta$ и $|90^\circ - \gamma| = |\alpha + \beta - 90^\circ|$. Будем максимизировать величину

$$\Delta = \min(\alpha,\ \beta - \alpha,\ 180^\circ - \alpha - 2\beta,\ 90^\circ - \beta,\ |\alpha + \beta - 90^\circ|).$$

Отдельно исследуем случаи 1) остроугольного и 2) тупоугольного треугольника.

1) Поскольку $\alpha > \alpha + \beta - 90^\circ$ и $90^\circ - \beta > \alpha + \beta - 90^\circ$, вместо пяти величин можно рассмотреть всего лишь три:

$$\Delta = \min(\alpha,\ \beta - \alpha,\ 180^\circ - \alpha - 2\beta,\ 90^\circ - \beta,\ \alpha + \beta - 90^\circ) = \\ = \min(\beta - \alpha,\ 180^\circ - \alpha - 2\beta,\ \alpha + \beta - 90^\circ).$$

Прямые, заданные уравнениями $\beta - \alpha = 180^\circ - \alpha - 2\beta$, $180^\circ - \alpha - 2\beta = \alpha + \beta - 90^\circ$ и $\beta - \alpha = \alpha + \beta - 90^\circ$, разбивают треугольник KLN (рис. 11) на треугольники KLG, KNG и LNG. Максимального значения величина Δ достигает в точке $G(45^\circ; 60^\circ)$ (рис. 12).

2) Поскольку $180^\circ - \alpha - 2\beta > 90^\circ - \alpha - \beta$ и $90^\circ - \beta > \alpha$, то

$$\Delta = \min(\alpha,\ \beta - \alpha,\ 180^\circ - \alpha - 2\beta,\ 90^\circ - \beta,\ 90^\circ - \alpha - \beta) = \min(\alpha,\ \beta - \alpha,\ 90^\circ - \alpha - \beta).$$

Прямые $\alpha = \beta - \alpha$, $\beta - \alpha = 90^\circ - \alpha - \beta$ и $\alpha = 90^\circ - \alpha - \beta$ разбивают треугольник OLN (рис. 13) на треугольники OLF, ONF и LNF. Максимального значения величина Δ достигает в точке $F(22{,}5^\circ; 45^\circ)$ (рис. 14). ■

Когда на уроке надо будет изобразить произвольный треугольник, все быстро и бездумно нарисуют его. А мы с вами будем транспортиром углы вымерять? Или схватимся за циркуль? Засмеют!

Возражение резонное — найденные нами треугольники должны быть удобными для рисования в школьной тетради (желательно без транспортира и циркуля!).

Далеко не каждый угол можно построить, соединив узлы сетки. Но... нужна ли нам абсолютная точность? Нет!

Задача, таким образом, сменилась на поиск наилучших приближений некоторых иррациональных чисел рациональными. На помощь приходят цепные дроби.

Начнем с наипроизвольнейшего прямоугольного треугольника. Отношение длин катетов равно

$$\operatorname{tg} 60^\circ = \sqrt{3} = [1;\ 1,\ 2,\ 1,\ 2,\ 1,\ 2,\ 1,\ 2,\ \ldots].$$

«Обрывая» эту дробь в разных местах, получаем все более точные приближения числа $\sqrt{3}$ обыкновенными дробями:

$$1 + \frac{1}{1} = 2, \quad 1 + \frac{1}{1 + \frac{1}{2}} = \frac{5}{3}, \quad 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1}}} = \frac{7}{4}, \quad 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{2}}}} = \frac{19}{11},$$

дальше не выписываем, поскольку числитель и знаменатель получаются слишком большие для тетрадной страницы. Поэтому имеет смысл строить прямоугольный треугольник с катетами 19 и 11 клеток. Неплох и треугольник с катетами 7 и 4, а также подобные ему — 14 и 8 или 21 и 12, отличие углов от 30° и 60° всего лишь около 1°. Очень удобен треугольник с катетами 14 и 8 (рис. 15) — он не слишком велик, и его стороны легко поделить пополам (например, если нужно провести медиану).

Следующий — треугольник с углами $67{,}5^\circ$, $67{,}5^\circ$ и 45°. Поскольку

$$1 = \operatorname{tg} 45^\circ = \frac{2 \operatorname{tg} 22{,}5^\circ}{1 - \operatorname{tg}^2 2{,}5^\circ},$$

то, решив квадратное уравнение $\operatorname{tg}^2 22{,}5^\circ + 2\operatorname{tg} 22{,}5^\circ - 1 = 0$, находим $\operatorname{tg} 22{,}5^\circ = \sqrt{2} - 1$. Поскольку $\sqrt{2} = [1; 2, 2, 2, \ldots]$, подходящие дроби числа

$$\operatorname{tg} 67{,}5^\circ = \operatorname{ctg} 22{,}5^\circ = \frac{1}{\sqrt{2}-1} = \sqrt{2}+1$$

таковы:

$$2 + \frac{1}{2} = \frac{5}{2},$$

$$2 + \cfrac{1}{2 + \cfrac{1}{2}} = \frac{12}{5},$$

$$2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2}}} = \frac{29}{12},$$

. . . . . .

Приближение 29/12 не уместится на тетрадную страницу, а предпоследнее дает треугольник с высотой 12 и основанием 10 (рис. 16). При необходимости (например, если нужно провести среднюю линию или медиану) размеры можно увеличить вдвое.

Следующий — равнобедренный треугольник с углом при основании 54°. Тангенс этого угла в явном виде вычислить сложнее, чем $\operatorname{tg} 60^\circ$. Но можно бесхитростно посчитать на калькуляторе:

$$\operatorname{tg} 54^\circ = [1; 2, 1, 1, 1, 10, \ldots].$$

Эта цепная дробь как будто сама указывает, где ее лучше оборвать:

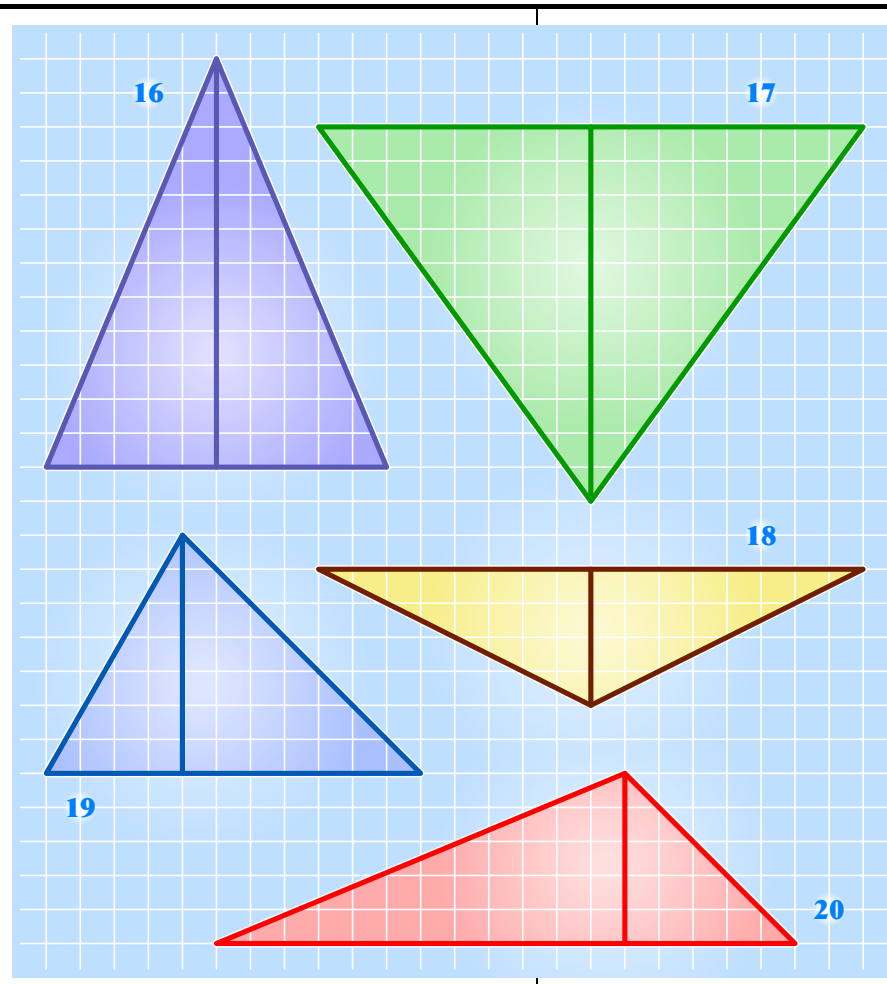
$$\operatorname{tg} 54^\circ \approx 1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1}}}} = \frac{11}{8}.$$

Треугольник имеет высоту 11 и основание $2 \cdot 8 = 16$ (рис. 17).

Равнобедренный треугольник с углом при основании 30° состоит из двух прямоугольных треугольников с углами 30° и 60°, сложенных меньшими катетами (рис. 18). Можно и удвоить, а то уж очень низенький получился.

Остроугольный неравнобедренный треугольник с углами 45°, 60° и 75° состоит из двух прямоугольных треугольников, величины углов одного из которых равны 45° и 45°, а второго — 60° и 30°. С первым проблем нет, а второй мы уже строили (рис. 19).

Треугольник с углами 22,5°, 45° и 112,5° состоит из двух прямоугольных треугольников, один из которых — равнобедренный прямоугольный, а величины углов второго — 22,5° и 67,5°. Приближение $\operatorname{tg} 67{,}5^\circ \approx 12/5$ дает треугольник с высотой 5 и основанием $12 + 5 = 17$ (рис. 20). ■



Можно придумать еще много разных критериев наипроизвольности — и разумных, и вздорных. Подготовиться на все возможные случаи все равно не удастся. (И заметьте: кроме треугольников надо уметь рисовать произвольный четырехугольник, трапецию, параллелограмм и даже семиугольную пирамиду.)

Не всегда нужно рисовать самую типичную фигуру, о которой идет речь в задаче. Иной раз эскиз полезнее, чем точный чертеж. Иногда надо анализировать частные случаи или предельные ситуации. Рисование чертежа — дело творческое!

Итак, когда вам скажут «нарисуйте произвольный треугольник», быстренько составьте целевую функцию, найдите, где она принимает экстремальное значение, разложите ответ в цепную дробь, найдите ее подходящие дроби — и рисуйте на здоровье! ■

Точка Торричелли

Эванджелиста Торричелли (1608—1647) — итальянский физик и математик, ученик и помощник Г. Галилея. После смерти Г. Галилея в 1642 г. был назначен «философом и первым математиком великого герцога Тосканского».

За тридцать лет до Ньютона Торричелли подошел к формулировке теоремы об обратности операций интегрирования и дифференцирования, хотя доказал ее лишь в терминах механики и для случая равномерно ускоренного движения.

Около 1641 г. Э. Торричелли написал книгу «О движении свободно падающих и брошенных тяжелых тел», в которой развил положения динамики Галилея и решил ряд задач баллистики. В 1644 г. он изобрел ртутный барометр и объяснил факт подъема ртути в трубке наличием воздушного давления. ■

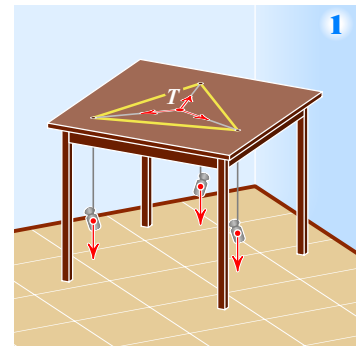
Точку T, сумма расстояний от которой до вершин данного треугольника минимальна, называют точкой Торричелли. Если величины углов треугольника меньше 120°, *то все его стороны видны из нее под углом* 120°, *а если величина одного из углов треугольника не меньше* 120°, *то точкой Торричелли является вершина этого угла.*

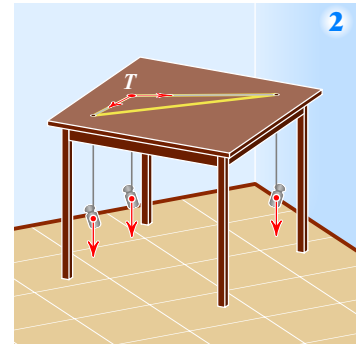
На деревянном столе нарисуем треугольник, просверлим в вершинах дырочки, проденем веревочки, свяжем их в одной точке над столом, привяжем к свободным концам грузики одинаковой массы и отпустим. Грузики будут стремиться опуститься как можно ниже (точнее, минимизировать потенциальную энергию системы трех грузиков). В наинизшем положении силы уравновесятся во внутренней точке T треугольника (рис. 1) или притянут точку T к одной из вершин и будут тянуть ее «под стол» (рис. 2). Сумма векторов одинаковой длины равна нулю (рис. 3) тогда и только тогда, когда они составляют друг с другом углы величиной 120°.

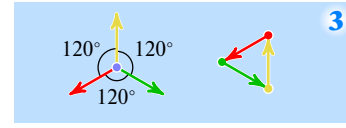
Так решают задачу физики. Решение верное, нужно только придать математический смысл словам «потенциальная энергия», «сила» и прочим использованным физическим идеям и понятиям. ■

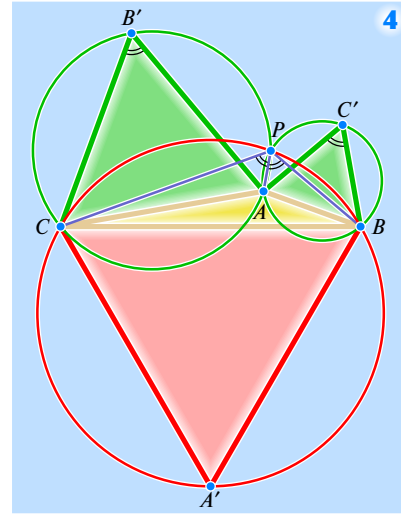
Опишем окружности вокруг равносторонних треугольников ABC' и CAB', построенных во внешнюю сторону треугольника ABC. Если величина угла A больше 120° (рис. 4), то окружности пересекаются вне треугольника ABC; в силу теоремы о вписанном угле точка пересечения P принадлежит описанной окружности треугольника CBA'.

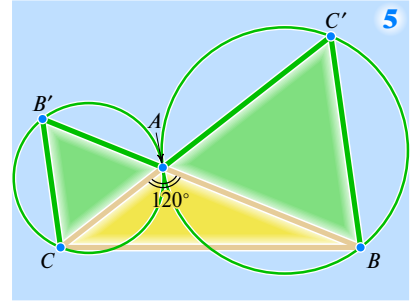
Если $\angle BAC = 120°$ (рис. 5), то зеленые окружности касаются.



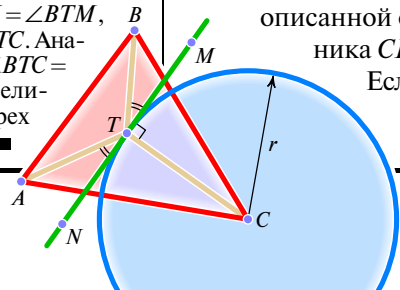


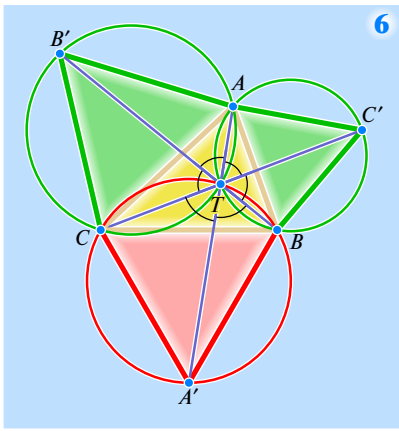






Пусть точка Торричелли T не совпадает ни с одной из вершин треугольника ABC. Докажем равенство углов ATC и BTC. Зафиксируем расстояние $CT = r$ и будем минимизировать сумму $AT + BT$. Поскольку окружность вблизи точки T от касательной MN «мало отличима» (смысл этим словам могут придать те читатели, которые изучали математический анализ), а на прямой MN точка, сумма расстояний от которой до точек A и B минимальна, выделена условием равенства углов падения и отражения, то $\angle ATN = \angle BTM$, откуда $\angle ATC = \angle BTC$. Аналогично имеем $\angle BTC = \angle BTA$, так что величины всех этих трех углов равны 120°. ■

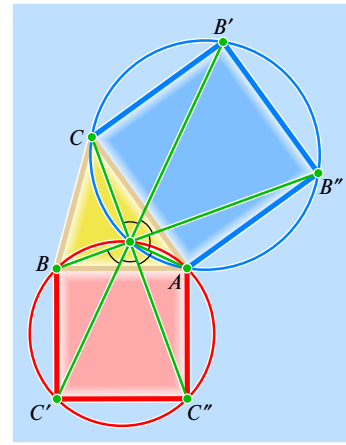




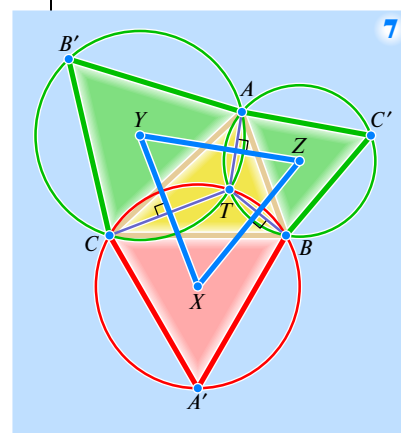
Если величины всех углов треугольника ABC меньше 120° (рис. 6), то зеленые окружности пересекаются внутри треугольника ABC в некоторой точке T, причем по теореме о вписанном угле величины всех помеченных на рисунке углов равны 60° и, значит, точка T принадлежит и описанной окружности треугольника CAB'; таким образом, отрезки AA', BB' и CC' пересекаются в точке T под углом 60° друг к другу. ■

Отрезок YZ, соединяющий центры зеленых окружностей, перпендикулярен общей хорде TA этих окружностей (рис. 7). Аналогично $XY \perp TC$ и $ZX \perp TB$. Поскольку величина угла между прямыми TB и TC равна 120°, то $\angle YXZ = 360° - 90° - 90° - 120° = 60°$. Аналогично доказываем равенства 60° величин углов Y и Z треугольника XYZ. Таким образом, треугольник XYZ равносторонний. Его называют внешним треугольником Наполеона Бонапарта (1769—1821).

Построив на сторонах произвольного треугольника ABC равносторонние треугольники не во внешнюю, а во внутреннюю сторону, получаем внутренний треугольник Наполеона $X'Y'Z'$ (рис. 8). Нетрудно убедиться, что описанные окружности треугольников ABC'', BCA'' и CAB'' пересекаются в одной точке, поэтому внутренний треугольник Наполеона тоже равносторонний. ■



Построим на сторонах треугольника ABC во внешнюю сторону квадраты $ABC'C''$ и $ACB'B''$. Опишем вокруг квадратов окружности. По теореме о вписанном угле величины всех помеченных на рисунке углов равны 45°. Поскольку $4 \cdot 45° = 180°$, прямые BB'', $B'C'$ и CC'' пересекаются в одной точке. ■



Равенство сторон треугольника Наполеона можно доказать и без использования точки Торричелли. По теореме косинусов, $XY^2 = XC^2 + YC^2 - 2XC \cdot YC \cos(30° + \gamma + 30°)$. Поскольку $XC = \frac{a}{\sqrt{3}}$, $YC = \frac{b}{\sqrt{3}}$ и $\cos(60° + \gamma) = \frac{1}{2}\cos\gamma - \frac{\sqrt{3}}{2}\sin\gamma$, имеем

$$XY^2 = \frac{a^2}{3} + \frac{b^2}{3} - \frac{ab}{3}(\cos\gamma - \sqrt{3}\sin\gamma) =$$

$$= \frac{a^2}{6} + \frac{b^2}{6} + \left(\frac{a^2}{6} + \frac{b^2}{6} - \frac{2ab\cos\gamma}{6}\right) + \frac{ab}{\sqrt{3}}\sin\gamma =$$

$$= \frac{a^2 + b^2 + c^2}{6} + \frac{2S_{ABC}}{\sqrt{3}}.$$

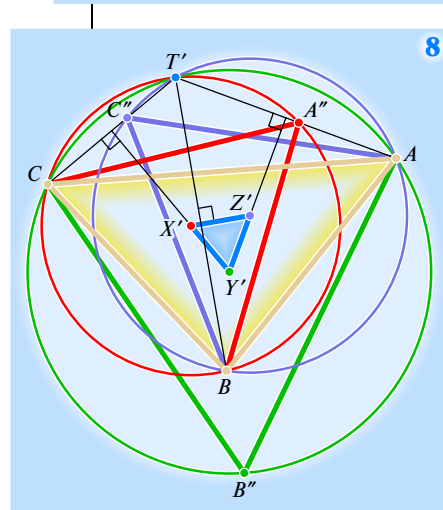
Формула симметричная, треугольник Наполеона равносторонний! Аналогично для стороны внутреннего треугольника Наполеона

$$X'Y'^2 = XC^2 + YC^2 - 2XC \cdot YC \cos(\gamma - 60°) = \frac{a^2 + b^2 + c^2}{6} - \frac{2S_{ABC}}{\sqrt{3}}.$$

Следовательно, $XY^2 - X'Y'^2 = 4S_{ABC}/\sqrt{3}$ и, значит,

$$S_{XYZ} - S_{X'Y'Z'} = \frac{\sqrt{3}}{4}(XY^2 - X'Y'^2) = S_{ABC}.$$

Разность площадей внешнего и внутреннего треугольников Наполеона равна площади исходного треугольника! ■

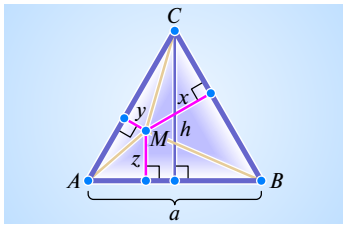


Центры внутреннего и внешнего треугольников Наполеона совпадают с точкой M пересечения медиан треугольника ABC.

Площадь равностороннего треугольника ABC со стороной a и высотой h можно вычислить как по формуле $S=ah/2$, так и по формуле

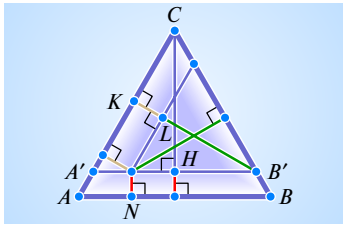
$$S=S_{BCM}+S_{CAM}+S_{ABM}=(ax+ay+az)/2.$$

Следовательно, $x+y+z=h$.



Равенство $x+y+z=h$ можно доказать и без помощи площадей, проведя через точку M отрезок $A'B'\|AB$ и заметив, что $z=MN$ и

$$h-z=CH=B'K=B'L+LK=x+y.$$ ■



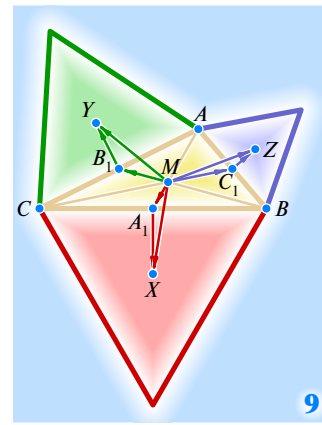
Повернем треугольник YAZ на 120° по часовой стрелке вокруг точки Y. Аналогично, повернем вокруг точки X треугольник XBZ на 120° против часовой стрелки. При этих поворотах точки A и B перейдут в точку C. Как легко проверить, сумма величин углов YAZ, ZBX и XCY равна 360°. Поскольку к тому же $AZ=ZB$, то при рассматриваемых поворотах отрезки AZ и BZ перейдут в один и тот же отрезок CD.

Дельтоид $XDYZ$ состоит из конгруэнтных треугольников XYZ и XYD. Поскольку $\angle ZXD=120^\circ=\angle ZYD$, то треугольники XYZ и XYD равносторонние! ■

Докажем это. Обозначим через A_1, B_1 и C_1 середины сторон треугольника ABC (рис. 9). Очевидно,

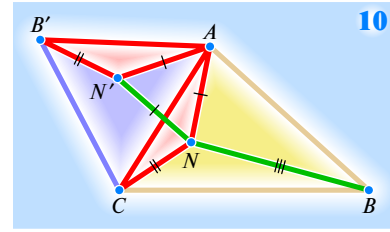
$$\overrightarrow{MX}+\overrightarrow{MY}+\overrightarrow{MZ}=$$
$$=(\overrightarrow{MA_1}+\overrightarrow{A_1X})+(\overrightarrow{MB_1}+\overrightarrow{B_1X})+(\overrightarrow{MC_1}+\overrightarrow{C_1X})=$$
$$=(\overrightarrow{MA_1}+\overrightarrow{MB_1}+\overrightarrow{MC_1})+(\overrightarrow{A_1X}+\overrightarrow{B_1X}+\overrightarrow{C_1X})=\overrightarrow{0}.$$

Первая скобка равна $\overrightarrow{0}$, поскольку точка M является центром тяжести треугольника $A_1B_1C_1$, а вторая равна $\overrightarrow{0}$, так как векторы $\overrightarrow{A_1X}$, $\overrightarrow{B_1Y}$ и $\overrightarrow{C_1Z}$ можно получить из векторов $\overrightarrow{BC}$, $\overrightarrow{CA}$ и $\overrightarrow{AB}$ соответственно поворотом на 90° и уменьшением длин в $2\sqrt{3}$ раз.



9

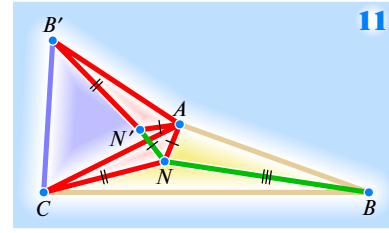
Равенство $\overrightarrow{MX}+\overrightarrow{MY}+\overrightarrow{MZ}=\overrightarrow{0}$ означает, что M — центр тяжести треугольника XYZ. Доказательство для внутреннего треугольника Наполеона аналогично, только поворот на 90° будет в другую сторону. ■



10

При повороте на 60° вокруг точки A точка B' переходит в точку C, а B — в C'. Поэтому отрезок BB' переходит в $C'C$ и, следовательно, $BB'=CC'$. На самом деле

$$AT+BT+CT=AA'=BB'=CC'.$$



11

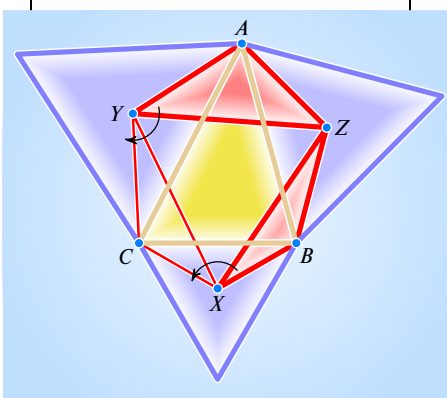
Доказать это (а заодно и выяснить, где расположена точка Торричелли) поможет тот же самый поворот (рис. 10). Поскольку треугольник ACN при повороте на 60° вокруг точки A по часовой стрелке переходит в треугольник $AB'N'$, то $CN=B'N'$ и, таким образом,

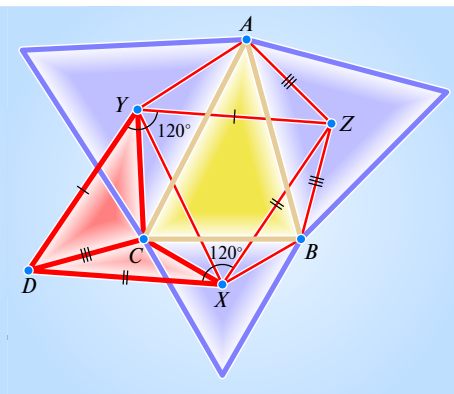
$$BB'\leqslant BN+NN'+N'B'=BN+NA+NC.$$

Значит, $BB'\leqslant AN+BN+CN$, а равенство достигается лишь в случае, когда точки B, N, N' и B' лежат на одной прямой (в указанном порядке). Это значит, что величины углов треугольника ABC не превышают 120°, а сумма расстояний от точки N до вершин треугольника минимальна для точки T, из которой все стороны треугольника видны под углом 120°.

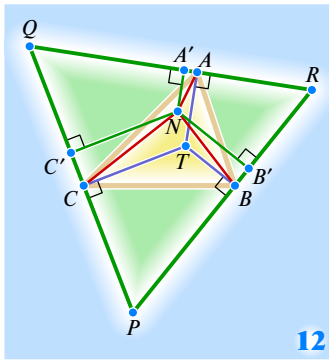
При помощи поворота нетрудно исследовать и случай $\angle BAC\geqslant 120^\circ$ (рис. 11): поскольку длина объемлющей ломаной $BNN'B'$ не меньше длины объемлемой выпуклой ломаной BAB', то

$$BA+AC=BA+AB'\leqslant BN+NN'+N'B'=BN+AN+CN.$$ ■

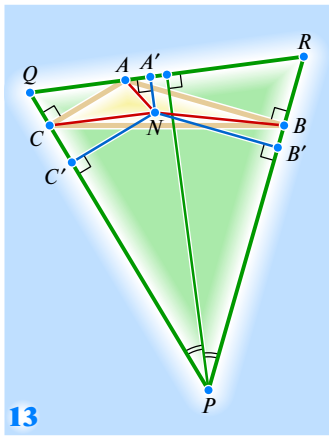




Рассмотрим треугольник ABC, величины всех углов которого меньше 120°. Внутри него существует точка T, из которой все стороны видны под углом 120°. Восставим в точках A, B и C перпендикуляры к отрезкам TA, TB и TC (рис. 12). Получаем равносторонний треугольник PQR. Поскольку сумма расстояний



12



13

от любой точки N равностороннего треугольника до его сторон равна высоте треугольника и тем самым не зависит от того, какую именно точку мы рассматриваем, имеем

$$AT+BT+CT=NA'+NB'+NC'\leqslant NA+NB+NC.$$

Сумма $AT+BT+CT$ вдвое больше высоты треугольника XYZ (см. рис. 7) и в $\sqrt{3}$ раз меньше длины стороны треугольника Наполеона. ■

В случае $\angle BAC\geqslant 120^\circ$ точку Торричелли можно найти тем же способом. Восставим перпендикуляры в точках B и C к отрезкам AB и AC соответственно (рис. 13) и обозначим точку их пересечения буквой P. Через точку A проведем перпендикуляр к биссектрисе угла BPC. Очевидно, $PR=PQ\geqslant QR$. Для любой точки N треугольника PQR имеем

$$NA'\cdot QR+NB'\cdot PR+NC'\cdot PQ=2S_{PQR}=$$
$$=2(S_{PAR}+S_{PAQ})=AB\cdot PR+AC\cdot PQ.$$

Разделив на $PR=PQ$, получаем:

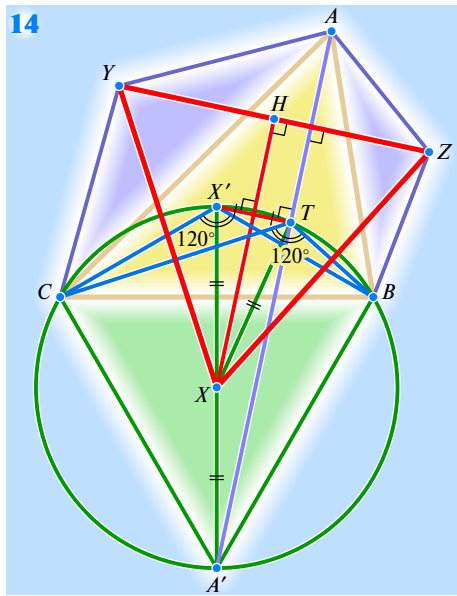
$$NA'\cdot\frac{QR}{PR}+NB'+NC'=AB+AC,$$

откуда

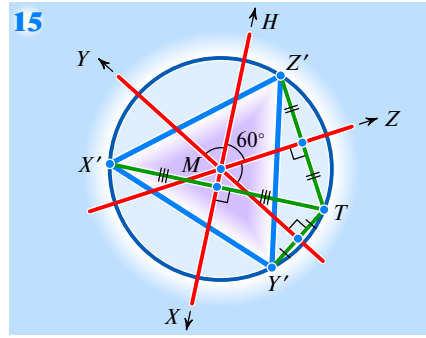
$$NA+NB+NC\geqslant NA'+NB'+NC'\geqslant AB+AC.\ ■$$

Радиус описанной окружности внутреннего треугольника Наполеона равен TM (рис. 14). В самом деле, поскольку $\angle BTC=\angle BX'C=120^\circ$, то точки T и X' лежат на описанной окружности треугольника $A'BC$. Поскольку X — центр этой окружности, то $TX=XX'=XA'$. Медиана TX треугольника $TA'X'$ равна половине стороны, к которой она проведена; значит, $TX'\perp TA'$.

Как вы помните, $TA\perp YZ$. Высота XH внешнего треугольника Наполеона проходит через его центр M и тоже перпендикулярна стороне YZ. Поэтому точки T и X' симметричны относительно прямой MX. Аналогично, Y' и Z' — образы точки T при симметрии относительно прямых MY и MZ соответственно. Следовательно, треугольник $X'Y'Z'$ равносторонний (рис. 15), причем точка T лежит на его описанной окружности. ■

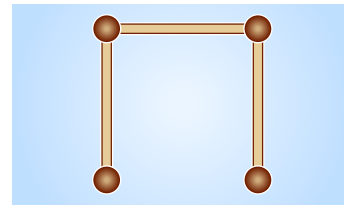


14

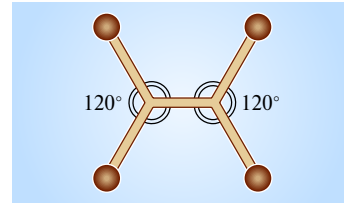


15

Солдаты, разместившиеся в нескольких окопах, хотят прокопать систему траншей наименьшей возможной суммарной длины, чтобы из любого окопа можно было добраться до любого другого. Что мы им посоветуем? Например, верно ли, что если окопы находятся в вершинах квадрата со стороной длины 1, то надо копать траншеи вдоль трех сторон квадрата?



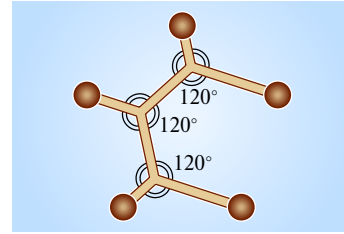
Нет, для квадрата сеть Штейнера (именно так называют искомую сеть траншей) содержит две развилки! Сумма длин траншей равна, как легко подсчитать, $1+\sqrt{3}<3$.



В общем случае сеть Штейнера получается добавлением к системе k окопов некоторого числа r развилок, причем в каждой развилке сходятся три траншеи под 120° друг к другу. Сеть является деревом (из любого цикла мы удалили бы любое ребро без потери связности графа) и поэтому состоит из $k+r-1$ отрезков. Поскольку из любой развилки выходят три траншеи, а из окопа — не менее чем одна, то

$$k+r-1\geqslant\frac{3r+k}{2},$$

откуда $r\leqslant k-2$. (Пополам поделили, поскольку у отрезка два конца.) Вот какова сеть Штейнера для вершин правильного пятиугольника:



Тригонометрические тождества

Вся тригонометрия основана на двух формулах:

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta, \qquad (1)$$

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta. \qquad (2)$$

В разных доказательствах этих формул вам встретятся теорема синусов, повороты, скалярное произведение векторов, теорема Птолемея и комплексные числа.

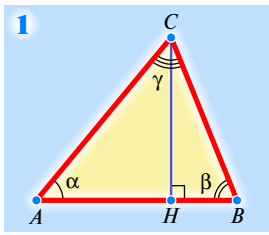
Опустим высоту CH на сторону AB треугольника ABC (рис. 1):

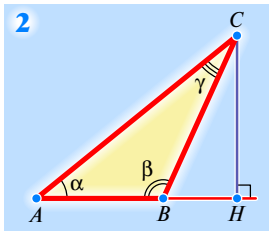
$$AB=AH+HB=AC\cos\alpha+BC\cos\beta.$$

По теореме синусов, $AB=2R\sin\gamma$, $AC=2R\sin\beta$ и $BC=2R\sin\alpha$, где R — радиус описанной окружности треугольника ABC. Подставив и разделив на $2R$, получаем

$$\sin\gamma=\sin\beta\cos\alpha+\sin\alpha\cos\beta.$$

Осталось заметить, что $\sin\gamma=\sin(180^\circ-\alpha-\beta)=\sin(\alpha+\beta)$. Формула (1) доказана. Существенный недостаток этого доказательства в том, что на α и β оно накладывает условия $0<\alpha$, $\beta<90^\circ$ — иначе треугольник ABC с нужными нам углами либо вообще не существует, либо же высота CH падает не на сторону, а на ее продолжение. Впрочем, эти ограничения можно ослабить, заменив их на неравенства $0<\alpha$, $0<\beta$ и $\alpha+\beta<180^\circ$: если один из углов A и B тупой (рис. 2), то $AB=AH-BH=AC\cos\alpha-BC\cos(180^\circ-\beta)=AC\cos\alpha+BC\cos\beta$. Однако в формуле (1) величины α и β могут быть любыми, не обязательно положительными; и даже если они положительны, их сумма $\alpha+\beta$ не обязательно меньше 180°. Как быть? Приходится использовать формулы приведения: $\cos\alpha=\sin(90^\circ-\alpha)$, $\sin(\alpha+180^\circ)=-\sin\alpha$, ... При помощи них можно свести общий случай к ситуации, для которой формула доказана. ■

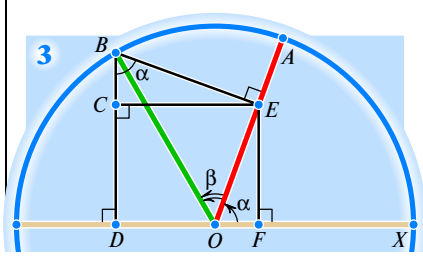


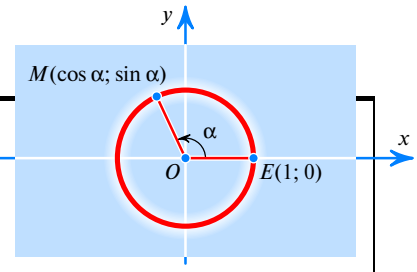


Рассмотрим окружность единичного радиуса с центром в точке O и отложим от некоторого луча OX против часовой стрелки угол величиной α, а от полученного луча OA — угол величиной β (рис. 3). Опустим перпендикуляры BD и BE на прямые OX и OA. Затем из точки E опустим перпендикуляры EF и EC на прямые OX и BD. Очевидно, $\angle CBE=\angle EOF=\alpha$. Имеем:

$$\sin(\alpha+\beta)=BD=BC+CD=BC+EF=$$
$$=BE\cos\alpha+OE\sin\alpha=\sin\beta\cos\alpha+\cos\beta\sin\alpha,$$
$$\cos(\alpha+\beta)=-OD=OF-DF=OF-CE=$$
$$=OE\cos\alpha-BE\sin\alpha=\cos\beta\cos\alpha-\sin\beta\sin\alpha.$$

Это доказательство верно для любых величин α и β: положительных и отрицательных, больших и маленьких. Использованные в доказательстве величины «сами собой» в нужных случаях положительны, в нужных — отрицательны. ■



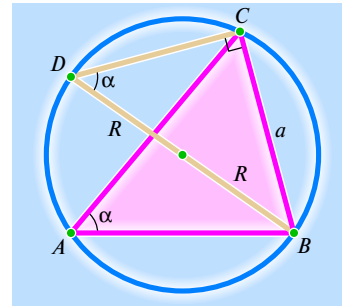


Повернув вокруг начала координат точку $E(1;0)$ на угол величиной φ против часовой стрелки, получаем точку M, абсциссу и ординату которой называют, соответственно, косинусом и синусом угла φ.

Проведя диаметр BD описанной окружности треугольника ABC, получаем прямоугольный треугольник BDC. В силу теоремы о вписанном угле

$$\sin\alpha=\sin BDC=\frac{a}{2R},$$

откуда $2R=a/\sin\alpha$. Это и есть знаменитая теорема синусов. ■



В какую точку переходит точка $(x;y)$ при отражении относительно прямой, проходящей через начало координат под углом φ к положительному направлению оси абсцисс?

Сначала выполним поворот на угол φ по часовой стрелке. Прямая при этом перейдет в ось абсцисс, а точка $(x;y)$ — в точку $(x\cos\varphi+y\sin\varphi;\ y\cos\varphi-x\sin\varphi)$. При симметрии относительно оси абсцисс меняется знак ординаты, так что получаем точку $(x\cos\varphi+y\sin\varphi;\ -y\cos\varphi+x\sin\varphi)$. При повороте на угол φ эта точка переходит в точку, абсциссу которой вычисляем по формуле

$$(x\cos\varphi+y\sin\varphi)\cos\varphi-$$
$$-(-y\cos\varphi+x\sin\varphi)\sin\varphi=$$
$$=x(\cos^2\varphi-\sin^2\varphi)+2y\sin\varphi\cos\varphi=$$
$$=x\cos2\varphi+y\sin2\varphi.$$

Аналогично вычисляем ординату:

$$(-y\cos\varphi+x\sin\varphi)\cos\varphi+$$
$$+(x\cos\varphi+y\sin\varphi)\sin\varphi=$$
$$=y(-\cos^2\varphi+\sin^2\varphi)+2x\sin\varphi\cos\varphi=$$
$$=-y\cos2\varphi+x\sin2\varphi.$$ ■

Скалярное произведение векторов $\vec{a}=(\cos\alpha;\sin\alpha)$ и $\vec{b}=(\cos\beta;-\sin\beta)$ по определению равно произведению длин этих векторов на косинус угла между ними:

$$\vec{a}\cdot\vec{b}=1\cdot1\cdot\cos(\alpha+\beta).$$

С другой стороны, скалярное произведение равно сумме произведений соответствующих координат:

$$\vec{a}\cdot\vec{b}=\cos\alpha\cos\beta-\sin\alpha\sin\beta.$$

Сравнивая последние две формулы, получаем формулу (2)! ■

Поворот вокруг точки O на угол φ — это отображение плоскости, при котором всякая точка M переходит в такую точку N, что $OM=ON$ и $\angle MON=\varphi$. Угол φ откладывают против часовой стрелки, если φ положителен; если же $\varphi<0$, то откладывают по часовой стрелке угол величиной $-\varphi$. Поворот обычно обозначают так: R_O^φ, по первой букве слова rotation (англ.) — вращение (сравните: ротор, ротация).

Рассмотрим повороты R^α и R^β вокруг начала координат на углы α и β. Последовательное выполнение этих двух поворотов дает поворот на угол $\alpha+\beta$:

$$R^\beta\circ R^\alpha=R^{\alpha+\beta}.$$

При повороте на угол α вокруг начала координат точка $(1;0)$ переходит, как следует из определений косинуса и синуса, в точку с координатами $(\cos\alpha;\sin\alpha)$. Точка $(0;1)$ при этом же повороте переходит в точку $(-\sin\alpha;\cos\alpha)$.

Вектор $(x;y)$ можно разложить в сумму векторов, параллельных осям координат, по формуле $(x;y)=x\cdot(1;0)+y\cdot(0;1)$. Поскольку при повороте любой параллелограмм переходит в параллелограмм, то для любых чисел x, y и любых векторов $\vec{a}$ и $\vec{b}$ имеем

$$R(x\vec{a}+y\vec{b})=xR(\vec{a})+yR(\vec{b}).$$

Значит, вектор $(x;y)$ при повороте R^α переходит в вектор

$$xR^\alpha(1;0)+yR^\alpha(0;1)=x(\cos\alpha;\sin\alpha)+y(-\sin\alpha;\cos\alpha)=$$
$$=(x\cos\alpha-y\sin\alpha;\,x\sin\alpha+y\cos\alpha).$$

Последняя формула позволяет написать выражения для координат $(x_1;y_1)$ точки, в которую переходит точка $(x;y)$ при повороте на угол α вокруг начала координат:

$$x_1=x\cos\alpha-y\sin\alpha,$$
$$y_1=x\sin\alpha+y\cos\alpha.$$

Для координат $(x_2;y_2)$ точки, в которую переходит точка $(x_1;y_1)$ при повороте R^β, верны аналогичные формулы

$$x_2=x_1\cos\beta-y_1\sin\beta,$$
$$y_2=x_1\sin\beta+y_1\cos\beta.$$

Подставив в формулу для x_2 выражения для x_1, y_1, получим

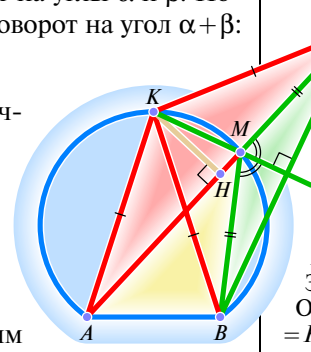
$$x_2=(x\cos\alpha-y\sin\alpha)\cos\beta-(x\sin\alpha+y\cos\alpha)\sin\beta.$$

Раскроем скобки и сгруппируем:

$$x_2=x\cos\alpha\cos\beta-y\sin\alpha\cos\beta-x\sin\alpha\sin\beta-y\cos\alpha\sin\beta=$$
$$=x(\cos\alpha\cos\beta-\sin\alpha\sin\beta)-y(\sin\alpha\cos\beta+\cos\alpha\sin\beta).$$

Когда в 1986 г. «Задачник «Кванта» приближался к М1000, был объявлен конкурс на лучшую задачу года. Поступили разные задачи, но ни одна из них не понравилась настолько, чтобы присвоить ей столь «круглый» номер. И тогда было принято решение: объявить победителем конкурса Архимеда из Сиракуз. В дугу AB вписана ломаная AMB из двух отрезков ($AM>MB$). Докажите, что основание перпендикуляра KH, опущенного из середины K дуги AB на отрезок AM, делит ломаную пополам: $AH=HM+MB$.

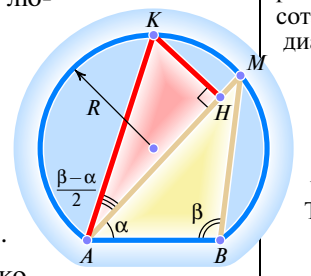
Геометрическое решение использует неожиданное дополнительное построение: на продолжении отрезка AM отложим точку C такую, что $MC=MB$.



Угол BMC как внешний угол треугольника AMB равен сумме углов BAM и ABM. В то же время угол KMA как вписанный измеряется половиной дуги AK и, следовательно, равен полусумме углов BAM и ABM. Поэтому MK — биссектриса угла BMC. Значит, $MK\perp BC$ и $KC=KB$. Отсюда следует, что $KA=KB=KC$. Поскольку в равнобедренном треугольнике AKC высота KH является заодно и медианой, имеем:

$$AH=HC=HM+MB,$$

что и требовалось.



Тригонометрическое решение не требует фантазии: обозначим $\angle BAM=\alpha$ и $\angle ABM=\beta$. Тогда

$$AM=2R\sin\beta,$$
$$BM=2R\sin\alpha,$$

где R — радиус описанной окружности треугольника ABM. Поэтому

$$AM+MB=2R(\sin\alpha+\sin\beta)=$$
$$=4R\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}.$$

С другой стороны,

$$AK=2R\sin\frac{\alpha+\beta}{2}$$

и $\angle KAM=\dfrac{\beta-\alpha}{2}$. Следовательно,

$$AH=AK\sin\angle KAM=$$
$$=2R\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}=$$
$$=\frac{1}{2}(AM+MB).\ ■$$

Подставим $-\beta$ вместо β в формулы (1) и (2):

$$\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta,$$
$$\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta.$$

Из этих формул и формул (1), (2) получаем

$$\sin(\alpha+\beta)+\sin(\alpha-\beta)=2\sin\alpha\cos\beta,$$
$$\cos(\alpha+\beta)+\cos(\alpha-\beta)=2\cos\alpha\cos\beta,$$
$$\cos(\alpha-\beta)-\cos(\alpha+\beta)=2\sin\alpha\sin\beta.$$

Эти формулы можно использовать для преобразования произведений тригонометрических функций в суммы:

$$\sin\alpha\cos\beta=(\sin(\alpha+\beta)+\sin(\alpha-\beta))/2,$$
$$\cos\alpha\cos\beta=(\cos(\alpha+\beta)+\cos(\alpha-\beta))/2,$$
$$\sin\alpha\sin\beta=(\cos(\alpha-\beta)-\cos(\alpha+\beta))/2.$$

Положим теперь $x=\alpha+\beta$, $y=\alpha-\beta$. Тогда $\alpha=\frac{x+y}{2}$, $\beta=\frac{x-y}{2}$, и мы получаем формулы для преобразования сумм и разностей тригонометрических функций в произведения:

$$\sin x+\sin y=2\sin\frac{x+y}{2}\cos\frac{x-y}{2},$$
$$\cos x+\cos y=2\cos\frac{x+y}{2}\cos\frac{x-y}{2},$$
$$\cos y-\cos x=2\sin\frac{x+y}{2}\sin\frac{x-y}{2}.$$

Из этих формул можно вывести и другие тригонометрические формулы. Например, подставив $\beta=\alpha$, получим:

$$\sin 2\alpha=2\sin\alpha\cos\alpha,$$
$$\cos 2\alpha=\cos^2\alpha-\sin^2\alpha.$$

Легко получить и формулу тангенса суммы:

$$\operatorname{tg}(\alpha+\beta)=\frac{\sin(\alpha+\beta)}{\cos(\alpha+\beta)}=$$
$$=\frac{\sin\alpha\cos\beta+\cos\alpha\sin\beta}{\cos\alpha\cos\beta-\sin\alpha\sin\beta}=$$
$$=\frac{\frac{\sin\alpha}{\cos\alpha}+\frac{\sin\beta}{\cos\beta}}{1-\frac{\sin\alpha\sin\beta}{\cos\alpha\cos\beta}}=\frac{\operatorname{tg}\alpha+\operatorname{tg}\beta}{1-\operatorname{tg}\alpha\operatorname{tg}\beta}.$$

Заменив β на $-\beta$, получим формулу тангенса разности:

$$\operatorname{tg}(\alpha-\beta)=\frac{\operatorname{tg}\alpha-\operatorname{tg}\beta}{1+\operatorname{tg}\alpha\operatorname{tg}\beta}.$$

Формулы для тангенса разности и тангенса суммы верны не при любых значениях переменных α и β, а только при тех, для которых $\operatorname{tg}\alpha$ и $\operatorname{tg}\beta$ существуют (то есть $\alpha,\ \beta\neq\pi/2+\pi n$ ни при каком целом n) и знаменатель не обращается в ноль (в случае тангенса суммы $\operatorname{tg}\alpha\operatorname{tg}\beta\neq1$, а в случае тангенса разности $\operatorname{tg}\alpha\operatorname{tg}\beta\neq-1$). ■

Теперь заметим, что

$$x_2=x\cos(\alpha+\beta)-y\sin(\alpha+\beta),$$

и сравним две формулы для x_2. В одной коэффициент при x равен $\cos\alpha\cos\beta-\sin\alpha\sin\beta$, в другой — коэффициент при x равен $\cos(\alpha+\beta)$. Получаем формулу (1). Аналогично, сравнивая коэффициенты при y, получаем формулу (2). ■

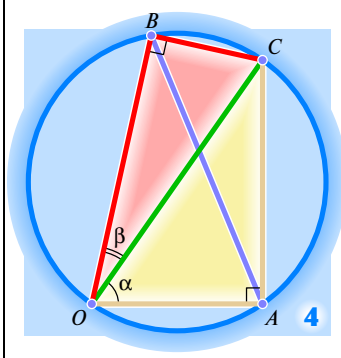
Теорема Птолемея. Рассмотрим прямоугольные треугольники OAC и OBC с общей гипотенузой OC длиной 1. Оба они вписаны в окружность с диаметром OC (рис. 4). В силу теоремы Птолемея,

$$OC\cdot AB=OA\cdot BC+OB\cdot AC.$$

Поскольку $OC=1$, $OA=\cos\alpha$, $OB=\cos\beta$, $CB=\sin\beta$ и $CA=\sin\alpha$, то

$$AB=\cos\alpha\cdot\sin\beta+\cos\beta\cdot\sin\alpha.$$

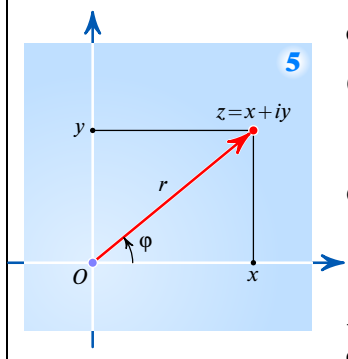
По теореме синусов, $AB=1\cdot\sin(\alpha+\beta)$ (в самом деле, 1 — это диаметр описанной окружности треугольника OAB, $\alpha+\beta$ — величина угла AOB). Формула (1) оказалась частным случаем теоремы Птолемея! (Разумеется, наше доказательство имеет смысл только для острых углов α и β.) Между прочим, Птолемей использовал свою теорему именно для вычисления тригонометрических функций одних углов через функции других углов. Поэтому можно сказать, что исторически теорема Птолемея (II в. нашей эры) предшествовала формулам (1), (2). В современную форму тригонометрию привел Л. Эйлер (1707—1783). ■



Произведение комплексных чисел

$\cos\alpha+i\sin\alpha$ и $\cos\beta+i\sin\beta$ равно

$$(\cos\alpha+i\sin\alpha)\cdot(\cos\beta+i\sin\beta)=$$
$$=\cos\alpha\cos\beta+i\cos\alpha\sin\beta+$$
$$+i\sin\alpha\cos\beta-\sin\alpha\sin\beta.$$



С другой стороны, чтобы перемножить комплексные числа, заданные в тригонометрической форме, достаточно перемножить их модули и сложить аргументы. Модули чисел $\cos\alpha+i\sin\alpha$ и $\cos\beta+i\sin\beta$ равны 1; аргументы равны α и β соответственно. Значит,

$$(\cos\alpha+i\sin\alpha)\cdot(\cos\beta+i\sin\beta)=\cos(\alpha+\beta)+i\sin(\alpha+\beta).$$

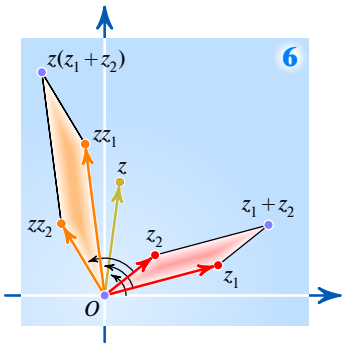
Мы в очередной раз доказали формулы (1) и (2)! ■

Во многих учебниках при рассказе о тригонометрической форме комплексных чисел использованы тригонометрические формулы. Поэтому, чтобы избежать порочного круга, изложим (интересный и сам по себе) геометрический способ определения комплексных чисел и операций над ними. Рассмотрим плоскость, в которой задана система координат (рис. 5). Любую точку z плоскости можно задавать не только декартовыми координатами $(x;y)$, но и полярными координатами $(r;\varphi)$, где $r=\sqrt{x^2+y^2}$ — расстояние

от точки z до начала координат; φ — угол, на который можно повернуть против часовой стрелки положительную полуось числовой прямой до того положения, при котором она пройдет через точку z.

Как видим, всегда $r \geqslant 0$, причем $r = 0$ только при $z = 0$. Кроме того, для всех $z \neq 0$ величина φ определена с точностью до кратных 360°, так что в интервале $0 \leqslant \varphi < 360°$ величина φ определена однозначно. Величину r называют *модулем* (или *абсолютной величиной*) числа z, а φ — *аргументом* числа z. Обозначения: $r = |z|$, $\varphi = \arg z$.

Сумму $z_1 + z_2$ любых двух точек z_1 и z_2 определим как сумму векторов, то есть по правилу параллелограмма (рис. 6). Произведение $z_1 z_2$ чисел, полярные координаты которых суть $(r_1; \varphi_1)$ и $(r_2; \varphi_2)$, определим как точку с полярными координатами $(r_1 r_2; \varphi_1 + \varphi_2)$ (модули перемножаем, аргументы складываем).

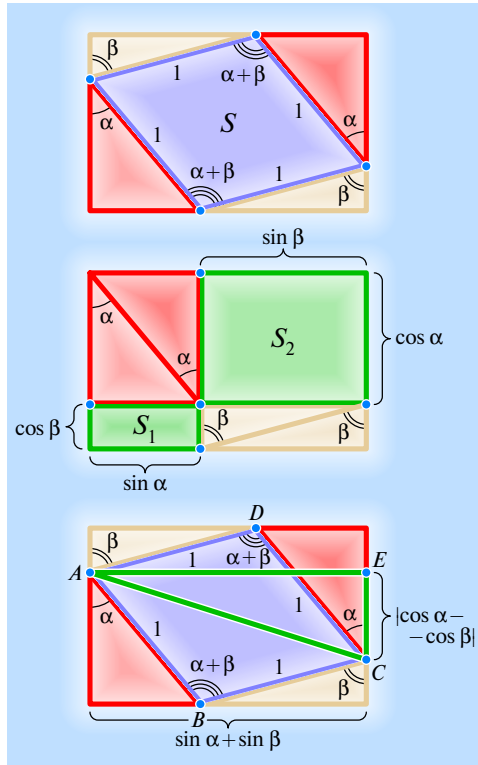


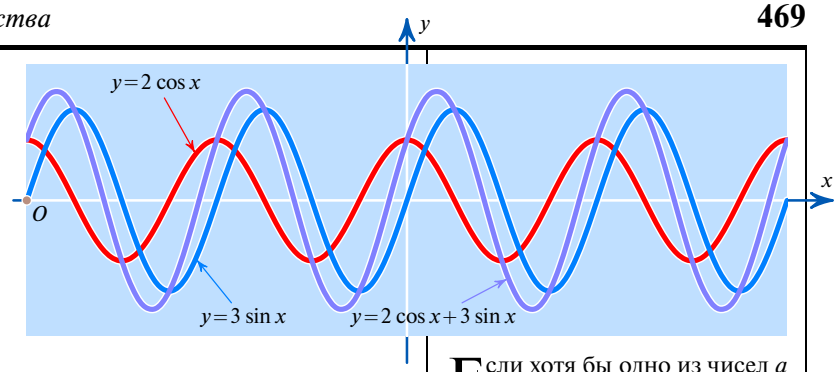
Очевидно, для точек вещественной числовой прямой вышеопределенные операции сложения и умножения не выводят за пределы этой прямой и соответствуют обычным операциям сложения и умножения вещественных чисел. (Проверьте свойства умножения, особенно правило «минус на минус дает плюс».)

Из основных свойств операций умножения и сложения для комплексных чисел неочевиден только распределительный закон (другими словами, дистрибутивность):

$$z(z_1 + z_2) = zz_1 + zz_2.$$

Для нас этот закон очень важен: именно он позволяет раскрывать скобки, при умножении комплексных чисел. Давайте его докажем. Геометрически умножение на z означает последовательное (в любом порядке) применение гомотетии с коэффициентом $|z|$ и поворота вокруг начала координат на угол φ. При гомотетии параллелограмм переходит в параллелограмм. При повороте — то же самое, параллелограмм переходит в параллелограмм (рис. 6). А это и есть дистрибутивность! Иначе говоря, сложить сначала векторы z_1 и z_2 и увеличить затем длину полученной суммы в $|z|$ раз, осуществив к тому же поворот на угол φ, — это все равно что сначала каждый из векторов z_1 и z_2 увеличить в $|z|$ раз и повернуть на угол φ, а уже затем сложить полученные векторы. ■





Если хотя бы одно из чисел a и b отлично от нуля, то

$$a\cos x + b\sin x = \sqrt{a^2+b^2} \times \left(\frac{a}{\sqrt{a^2+b^2}}\cos x + \frac{b}{\sqrt{a^2+b^2}}\sin x\right) =$$
$$= \sqrt{a^2+b^2}(\sin\varphi\cos x + \cos\varphi\sin x) = \sqrt{a^2+b^2}\sin(x+\varphi),$$

где φ — такой угол, что $\sin\varphi = a/\sqrt{a^2+b^2}$ и $\cos\varphi = b/\sqrt{a^2+b^2}$. Мы получили так называемую формулу введения дополнительного угла. Если $x = \omega t$, где ω — ненулевая постоянная величина, t — время, то получаем, что сумма двух гармонических колебаний $y = a\cos\omega t$ и $y = b\sin\omega t$ — гармоническое колебание. Величину φ называют фазой, а $\sqrt{a^2+b^2}$ — амплитудой гармонического колебания $y = a\cos x + b\sin x$. ■

Рассмотрим ромб со стороной 1 и пристроим к его сторонам прямоугольные треугольники, как показано на рисунке. Получим прямоугольник со сторонами $\sin\alpha + \sin\beta$ и $\cos\alpha + \cos\beta$. Этот же прямоугольник можно составить из тех же четырех треугольников и двух прямоугольников. Очевидно, что площадь ромба $S = \sin(\alpha+\beta)$ равна сумме $S_1 + S_2 = \sin\alpha\cos\beta + \sin\beta\cos\alpha$. Формула синуса суммы доказана (к сожалению, только для острых углов α, β). ■

Вычислим двумя способами длину диагонали ромба. По теореме косинусов

$$AC^2 = 1^2 + 1^2 - 2\cdot 1\cdot 1\cdot\cos(\alpha+\beta) = 2 - 2\cos(\alpha+\beta).$$

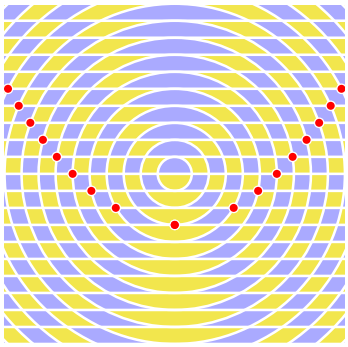
А по теореме Пифагора

$$AC^2 = AE^2 + EC^2 = (\sin\beta + \sin\alpha)^2 + (\cos\alpha - \cos\beta)^2 =$$
$$= 2 + 2\sin\alpha\sin\beta - 2\cos\alpha\cos\beta.$$

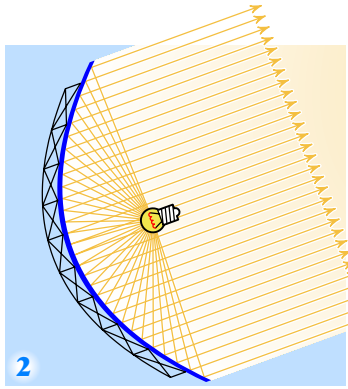
Сравнивая полученные выражения, получаем (для острых углов α и β) формулу косинуса суммы (2). ■

Парабола

Для алгебраиста парабола — это график квадратного трехчлена $y=ax^2+bx+c$, где $a\neq 0$. Для геометра — множество точек, равноудаленных от данной точки (фокуса параболы) и не проходящей через эту точку данной прямой (директрисы).



Сетка линий состоит из концентрических окружностей с радиусами 1, 2, 3, 4, ... и центром в точке O, прямой l, проходящей через точку O, и всевозможных параллельных прямых l касательных к окружностям. Полученные клетки раскрашены в шахматном порядке. Каждые две соседние красные точки — противоположные вершины фиолетовой клетки. Объясните, почему точки такой бесконечной цепочки лежат на одной параболе, так что рисунок словно соткан из парабол. ■



Параболу, заданную уравнением

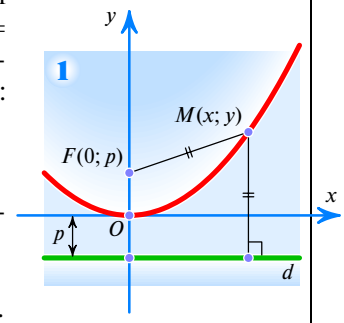
$$y=ax^2+bx+c=$$
$$=a\left(x+\frac{b}{2a}\right)^2-\frac{b^2}{4a}+c,$$

параллельный перенос на вектор $\left(\frac{b}{2a};\ \frac{b^2}{4a}-c\right)$ переводит в параболу, заданную уравнением $y=ax^2$. Гомотетия $(x;\ y)\to(ax;\ ay)$ переводит полученную параболу в параболу $\frac{y}{a}=a\left(\frac{x}{a}\right)^2$, то есть $y=x^2$. Таким образом, все параболы подобны. (Нетрудно доказать, что в случае $a\neq 1$ параболы $y=ax^2+bx+c$ и $y=x^2$ не только подобны, но и гомотетичны.) ■

Равносильность алгебраического и геометрического определений доказать легко. Обозначив расстояние между фокусом и директрисой через $2p$ (рис. 1), выразим расстояние от произвольной точки $M(x;\ y)$ плоскости до прямой d по формуле $\rho(M,\ d)=|y+p|$, а расстояние между точками M и F — по формуле $MF=\sqrt{x^2+(y-p)^2}$. Приравняем эти расстояния:

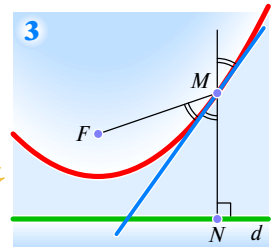
$$|y+p|=\sqrt{x^2+(y-p)^2},$$
$$y^2+2yp+p^2=x^2+y^2-2yp+p^2,$$

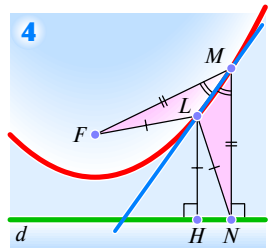
то есть $4py=x^2$, а это и есть алгебраическое определение параболы. ■

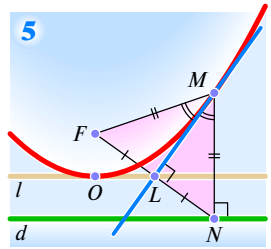


Поместим лампочку в фокус параболы. Отразившись, лучи света становятся параллельными (рис. 2)! Другими словами, прямая MF и перпендикуляр MN, опущенный на директрису, составляют равные углы с касательной к параболе (рис. 3). Доказать это свойство параболы проще всего от противного. Предположим, что биссектриса угла FMN пересекает параболу, кроме точки M, в точке L (рис. 4). Тогда треугольники NML и FML равны по двум сторонам и углу между ними. Следовательно, $LN=LF$. Опустим из точки L на директрису перпендикуляр LH. Получаем противоречие:

$$LF=LN>LH=LF.\ ■$$

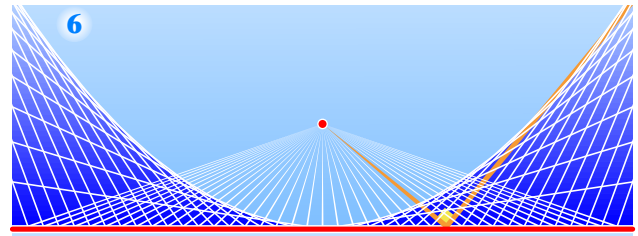




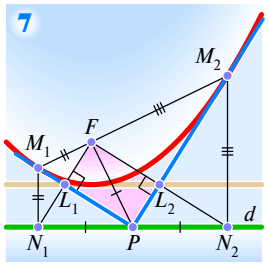


Касательная ML к параболе — биссектриса угла FMN равнобедренного треугольника — является заодно его высотой и медианой (рис. 5). Таким образом, когда точка M пробегает параболу, точка L движется по прямой l — касательной к параболе в вершине O.

На это можно посмотреть с другой точки зрения: перемещать не точку M по параболе, а точку L — по прямой l. А именно, для каждой точки L прямой l восставим к прямой FL перпендикуляр в точке L (рис. 6). Возникает огибающая перпендикуляров — парабола. ■

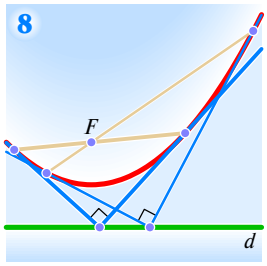


На директрисе d отметим некоторую точку P и проведем из нее касательные PM_1 и PM_2 (рис. 7). Очевидно, $PN_1 = PF = PN_2$. Поэтому угол N_1FN_2 прямой и PL_1FL_2 — прямоугольник. Значит, из всех точек своей директрисы парабола видна под прямым углом (рис. 8). При этом



$$\angle PFM_1 = \angle PN_1M_1 = 90^\circ = \angle PN_2M_2 = \angle PFM_2,$$

так что точки M_1, F и M_2 лежат на одной прямой. ■



Найдем расстояние от данной точки $A(0;\,b)$ до параболы, заданной уравнением $y=x^2$, то есть наименьшее из расстояний AM, где $M(x;\,x^2)$ — точка параболы. Для этого рассмотрим функцию

$$f(x) = AM^2 = x^2 + (b-x^2)^2$$

и найдем ее наименьшее значение. Обозначив $t=x^2$, имеем: $f(x)=t^2+(1-2b)t+b^2$. Квадратичная функция принимает свое минимальное значение в точке $t_0=(2b-1)/2$. При этом, как легко посчитать,

$$f\left(\sqrt{\frac{2b-1}{2}}\right) = b - \frac{1}{4}.$$

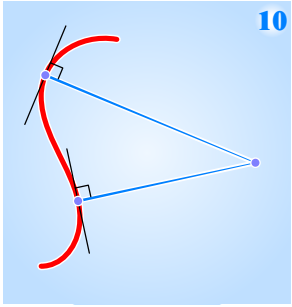
Впрочем, надо помнить о том, что $t=x^2\geqslant 0$: если $2b-1<0$, то минимальное значение функция $f(x)$ принимает при $x=0$. Итак,

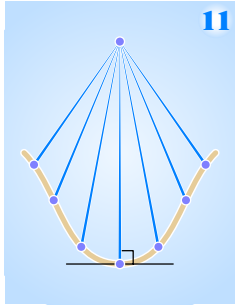
$$\min\sqrt{f(x)} = \begin{cases} \sqrt{b-\dfrac{1}{4}}, & \text{если } b>1/2, \\ |b|, & \text{если } b\leqslant 1/2. \end{cases}$$

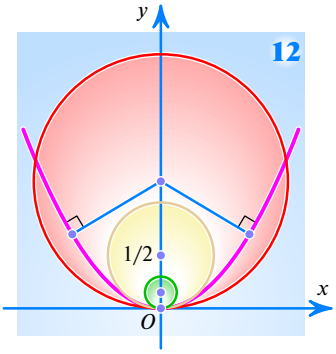
График этой функции изображен на рисунке 9. ■

Перпендикуляр из точки на кривую не всегда единственен (рис. 10) и не всегда является кратчайшим из отрезков, соединяющих данную точку с точками кривой: бывает даже, что перпендикуляр является наидлиннейшим из таких отрезков (рис. 11)!

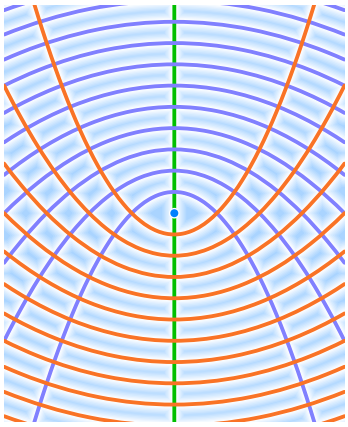
Если $b\leqslant 1/2$, то ближайшая к A точка параболы — начало координат. Если же $b>1/2$, то окружность радиуса b с центром $A(0;\,b)$ пересекает параболу более чем в одной точке (рис. 12); из точки $(0;\,b)$ можно опустить на параболу не один, а три перпендикуляра. ■

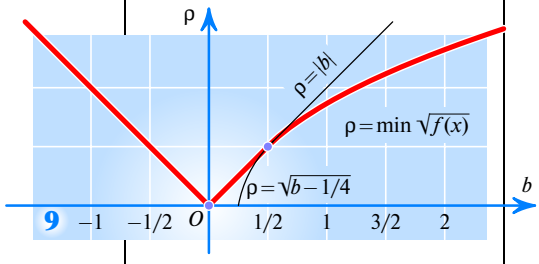






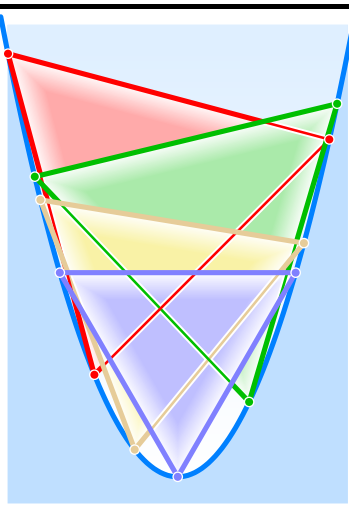
Рассмотрим все параболы с данным фокусом и данной вертикальной осью. Они естественно разбиваются на два семейства: у парабол одного семейства ветви идут вверх, у другого — вниз. Докажите, что любая из парабол одного семейства перпендикулярна любой параболе второго семейства. (По определению, угол между кривыми в точке их пересечения — это угол между касательными к этим кривым.) ■



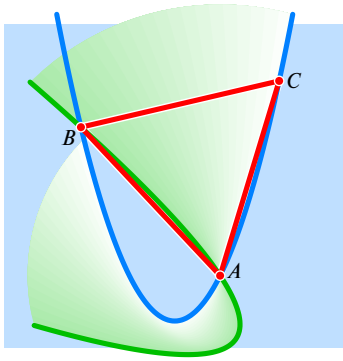


Параболические антенны так названы, поскольку их отражатель — параболоид — поверхность, получаемая при вращении параболы вокруг оси симметрии. Радиоволны, отразившись от параболоида, собираются в фокусе, где расположен приемник. ■

Одна из антенн системы сверхдлиннобазового интерферометра (VLBA), протянувшегося на 8000 км от Гавай до Антильских островов. Диаметр антенны 25 м. Фото NRAO/AUI.



В параболу можно вписать бесконечно много равносторонних треугольников. Чтобы построить такой треугольник ABC, достаточно взять любую точку A параболы и, повернув параболу на 60°, обозначить буквой B отличную от A точку пересечения параболы с ее образом, а буквой C — прообраз точки B при рассматриваемом повороте. ■



Сколько же перпендикуляров можно провести из данной точки $(a;\,b)$ к параболе $y=x^2$? Рассмотрим квадрат расстояния от точки $(a;\,b)$ до точки $(x;\,x^2)$ — функцию

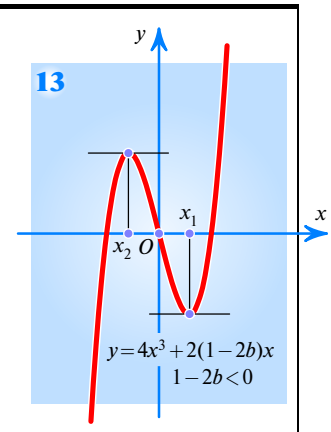
$$f(x)=(x-a)^2+(x^2-b)^2=x^4+(1-2b)x^2-2ax+a^2+b^2$$

и продифференцируем ее:

$$(f(x))'=4x^3+2(1-2b)x-2a.$$

Минимальное значение функция $f(x)$ достигает в одной из точек, где производная обращается в нуль. Если уравнение

$$4x^3+2(1-2b)x-2a=0$$



имеет единственное решение, то это и есть точка минимума, а перпендикуляр единственен. Если же решений больше одного, то перпендикуляров из точки $(a;\,b)$ на параболу можно опустить не один, а два или три.

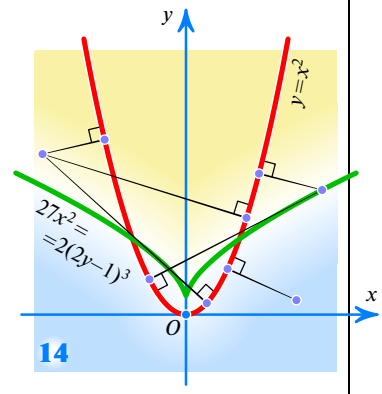
Интересно, когда перпендикуляр один, когда их два, а когда три? Если $1-2b\geqslant 0$, то функция $4x^3+2(1-2b)x$ является возрастающей и каждое свое значение принимает один раз. Так что при $b\leqslant 1/2$ на параболу можно опустить только один перпендикуляр. Если же $1-2b<0$, то график функции $y=4x^3+2(1-2b)x$ выглядит примерно так, как показано на рисунке 13. Найти точки локального максимума и минимума этой функции легко: ее производная равна

$$(4x^3+2(1-2b)x)'=12x^2+2(1-2b)$$

и обращается в нуль в точках $x_{1,2}=\pm\sqrt{(2b-1)/6}$. Вычислив

$$4x_{1,2}^3+2(1-2b)x_{1,2}=\mp\frac{2\sqrt{2}}{3\sqrt{3}}(2b-1)^{3/2},$$

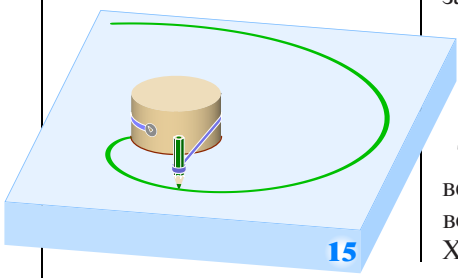
и обозначив для краткости $A=\dfrac{\sqrt{2}}{3\sqrt{3}}(2b-1)^{3/2}$,



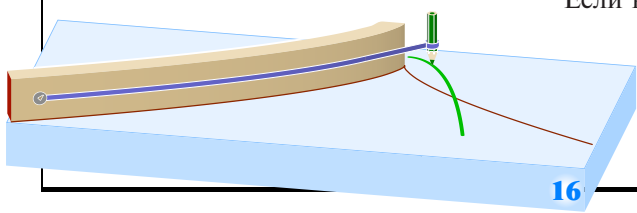
мы получаем ответ: из точки $(a;\,b)$, где $b>1/2$, на параболу $y=x^2$ можно опустить один перпендикуляр, если $|a|>A$; два перпендикуляра, если $a=\pm A$; три перпендикуляра, если $|a|<A$.

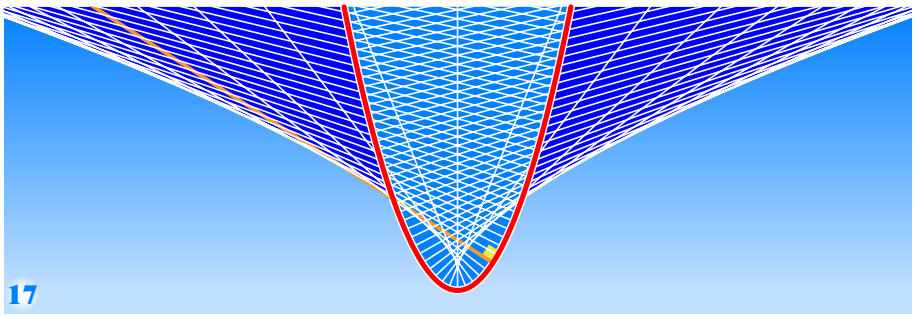
Чтобы сделать этот ответ наглядным, рассмотрим на плоскости Oxy кривую, заданную уравнением $27x^2=2(2y-1)^3$ — так называемую полукубическую параболу (рис. 14). Тогда из точек, расположенных выше этой кривой, на параболу $y=x^2$ можно опустить три перпендикуляра, из точек самой этой кривой, кроме ее «клюва» $(0;\,1/2)$, — два, а из клюва и из точек, лежащих ниже полукубической параболы — один перпендикуляр.

Такую полукубическую параболу называют *эволютой* (от лат. *evolutus* — «развернутый») параболы, а саму параболу — *эвольвентой* (от лат. *evolvens* — «разворачивающий») этой полукубической параболы. Эти термины ввел в 1673 г. Х. Гюйгенс (1629—1695). Эвольвента окружности изображена на рисунке 15.



Если ветвь эволюты параболы изготовить из жесткого материала и в одной из ее точек закрепить кнопкой нить такой длины, что конец нити при обтекании полукубической параболы попадет в вершину исходной параболы, то карандаш, привязанный к этому концу нити, при обратном движении опишет дугу параболы (рис. 16). ■





Отложив от каждой точки параболы $y=x^2$ во внешнюю и внутреннюю стороны по перпендикуляру отрезки длины 1/3, 1/2, 1, 2 и 3, получаем, соответственно, синие, сиреневые, зеленые, желтые и красные кривые. Заметьте: желтая и красная внутренние кривые имеют петли, причем красная петля вылезает из исходной параболы наружу.

Восставим перпендикуляр к параболе $y=x^2$

в каждой ее точке (рис. 17). Оказывается, все эти перпендикуляры касаются полукубической параболы, только что найденной нами. Поскольку $(x^2)'=2x$ и поскольку произведение угловых коэффициентов двух взаимно перпендикулярных прямых равно -1, то угловой коэффициент перпендикуляра, восставленного к параболе в точке $(c;\,c^2)$, равен $-\dfrac{1}{2c}$. Значит, этот перпендикуляр задан уравнением

$$y=-\frac{1}{2}(x-c)+c^2,$$

которое можно записать в виде

$$2cy=-x+c+2c^3.$$

Чтобы найти огибающую, рассмотрим аналогичное уравнение, в котором параметр c заменен на $c+\varepsilon$, причем в дальнейшем устремим ε к нулю. (Таким образом, чтобы найти огибающую, мы рассматриваем «близкие» прямые, находим точку их пересечения и переходим к пределу, превращая «близкие» прямые в, если позволено так выразиться, «бесконечно близкие».) Итак, рассмотрим систему уравнений

$$\begin{cases} 2cy=-x+c+2c^3, \\ 2(c+\varepsilon)y=-x+c+\varepsilon+2(c+\varepsilon)^3. \end{cases}$$

Вычтем первое уравнение из второго:

$$2\varepsilon y=\varepsilon+6c^2\varepsilon+6c\varepsilon^2+2\varepsilon^3,$$

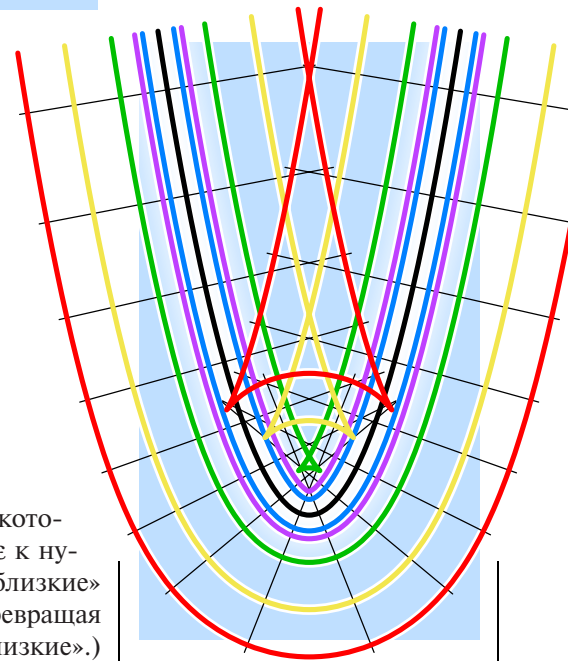
откуда $y=\dfrac{1}{2}+3c^2+3c\varepsilon+\varepsilon^2$. Устремив ε к нулю, находим $y=\dfrac{1}{2}+3c^2$. Подставив это значение в первое уравнение системы, имеем:

$$c+6c^3=-x+c+2c^3,$$

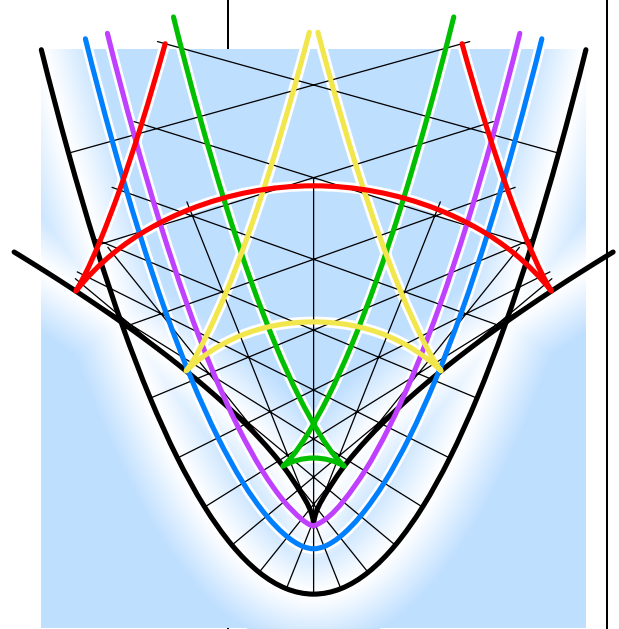
откуда $x=-4c^3$. Итак,

$$(x;\,y)=\left(-4c^3;\,\frac{1}{2}+3c^2\right).$$

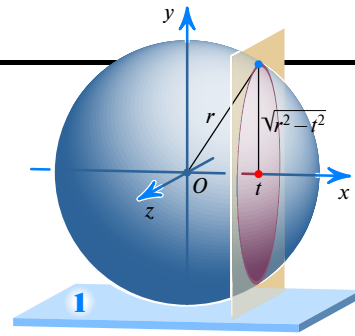
Мы нашли огибающую для семейства нормалей (то есть перпендикуляров) к параболе. Нетрудно убедиться, что найденные значения x и y удовлетворяют равенству $27x^2=2(2y-1)^3$, так что огибающая — обещанная полукубическая парабола! ■



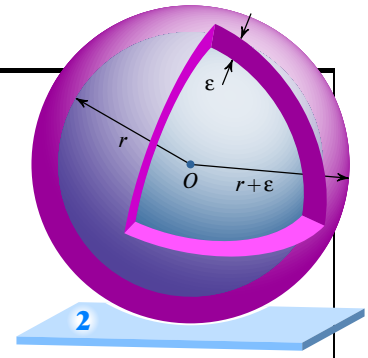
Точки возврата всех петель лежат на полукубической параболе (подумайте почему). ■



Шар и сфера



1



2

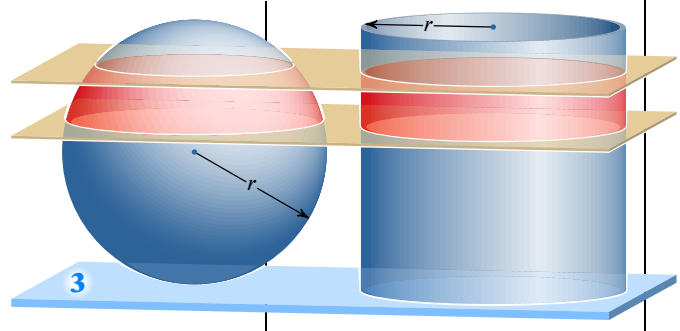
Шар состоит из точек, удаленных от данной точки (центра) не более чем на данное расстояние (радиус). Сфера — это граница шара. Сфера радиуса r с центром в начале координат задана уравнением $x^2+y^2+z^2=r^2$, а шар — неравенством $x^2+y^2+z^2\leqslant r^2$.

Объем шара радиуса r равен

$$V=\int_{-r}^{r} S(t)\,dt,$$

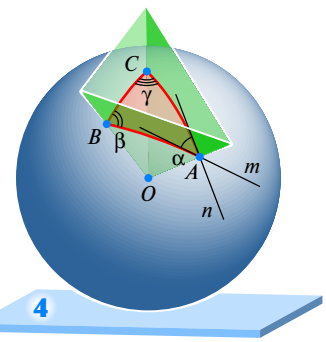
где $S(t)=\pi(r^2-t^2)$ — площадь сечения шара (рис. 1) плоскостью, заданной уравнением $x=t$. Поэтому

$$V=\int_{-r}^{r}\pi(r^2-t^2)\,dt=\pi\left(r^2t-\frac{t^3}{3}\right)\Bigg|_{-r}^{r}=\frac{4}{3}\pi r^3.\ \blacksquare$$



3

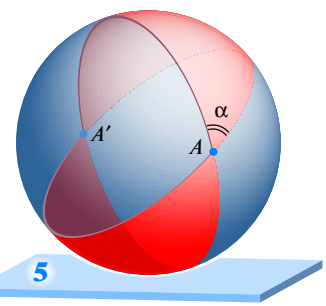
Площадь S сферы легко найти, рассмотрев шар радиуса $r+\varepsilon$ с тем же центром, где $\varepsilon>0$ (рис. 2). Разность объемов — объем шарового слоя толщины ε — равна $\frac{4}{3}\pi((r+\varepsilon)^3-r^3)=\frac{4}{3}\pi(3r^2\varepsilon+3r\varepsilon^2+\varepsilon^3)$. Тот же самый объем при маленьких ε с довольно высокой точностью равен εS. Значит, $S=\frac{4}{3}\pi\cdot 3r^2=4\pi r^2$. ■



4

На сфере и на цилиндре одинакового радиуса любые две плоскости, пересекающие сферу и параллельные основаниям цилиндра (рис. 3), высекают «пояски» одинаковой площади. В частности, площадь всей сферы равна площади боковой поверхности цилиндра: $4\pi r^2=2r\cdot 2\pi r$. При помощи современных обозначений доказать эту теорему Архимеда нетрудно. Сферический поясок является поверхностью вращения графика функции $f(x)=\sqrt{r^2-x^2}$, рассматриваемой на отрезке $a\leqslant x\leqslant b$, где $-r\leqslant a<b\leqslant r$, вокруг оси абсцисс. Поэтому его площадь равна

$$\int_a^b 2\pi f(x)\sqrt{1+(f'(x))^2}\,dx=\int_a^b 2\pi\sqrt{r^2-x^2}\sqrt{1+\left(\frac{-2x}{2\sqrt{r^2-x^2}}\right)^2}\,dx=\int_a^b 2\pi r\,dx=2\pi r(b-a).$$ ■



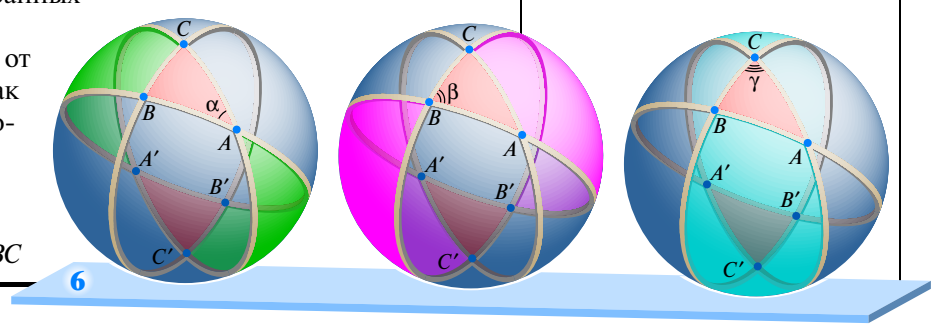
5

Трехгранный угол с вершиной в центре сферы высекает на ней сферический треугольник (рис. 4). Стороны сферического треугольника — дуги больших кругов. Поскольку касательные m и n к сфере перпендикулярны радиусу OA, то величины α, β и γ углов сферического треугольника равны величинам соответствующих двугранных углов трехгранного угла.

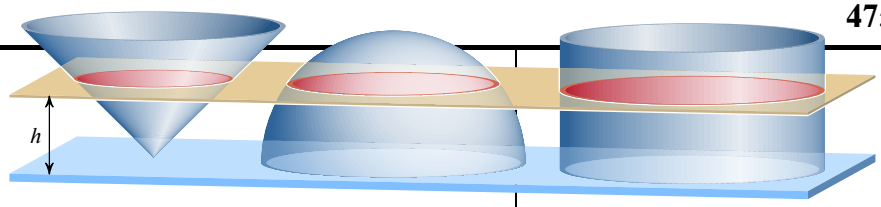
«Двуугольник» (рис. 5) составляет от площади сферы такую же часть, как угол 2α от угла 2π. Поэтому его площадь равна

$$\frac{2\alpha}{2\pi}\cdot 4\pi r^2=4\alpha r^2.$$

Теперь продолжим стороны AB, BC



6



и CA сферического треугольника до больших кругов. Три двуугольника покрывают сферический треугольник ABC и симметричный ему треугольник $A'B'C'$ в три слоя, а остальную часть сферы — в один слой (рис. 6). Поэтому

$$4\alpha r^2+4\beta r^2+4\gamma r^2=4\pi r^2+2S_{ABC}+2S_{A'B'C'},$$

где $S_{ABC}=S_{A'B'C'}$ — площадь сферического треугольника ABC, откуда

$$S_{ABC}=(\alpha+\beta+\gamma-\pi)\,r^2.$$

Следовательно, сумма величин углов сферического треугольника больше 180°. ■

Сколько существует сфер, касающихся всех четырех плоскостей граней тетраэдра? Всегда существуют одна вписанная (зеленая на рисунке 7) сфера и 4 вневписанные (сиреневые) сферы, причем

$$V=\frac{1}{3}(S_{ABC}+S_{ABD}+S_{ACD}+S_{BCD})\,r=\frac{1}{3}(S_{ABC}+S_{ABD}+S_{ACD}-S_{BCD})\,r_a,$$

где r — радиус вписанной сферы, r_a — радиус сферы, касающейся грани BCD и продолжений других трех граней. Если существует (коричневая) сфера, касающаяся «продолжений за ребро» граней ABC и ABD и «продолжений за вершины» A и B граней ACD и BCD, то

$$V=\frac{1}{3}(S_{ACD}+S_{BCD}-S_{ABC}-S_{ABD})\,R,$$

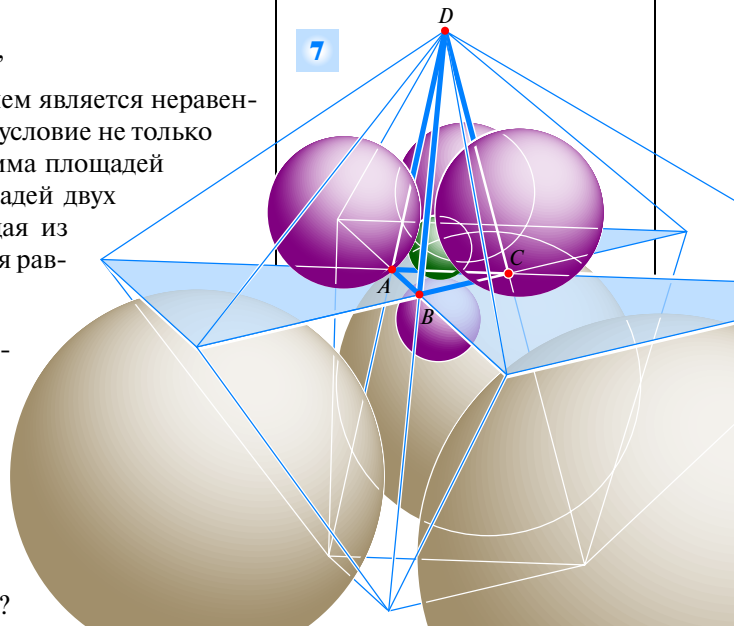
где R — радиус сферы. Необходимым для этого условием является неравенство $S_{ACD}+S_{BCD}>S_{ABC}+S_{ABD}$. Можно доказать, что это условие не только необходимое, но и достаточное. В частности, если сумма площадей никаких двух граней тетраэдра не равна сумме площадей двух других граней, то существуют $1+4+3=8$ сфер, каждая из которых касается всех четырех плоскостей граней. А для равногранного тетраэдра таких сфер всего $1+4=5$. ■

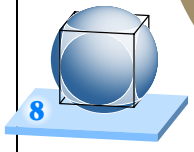
Рассмотрим куб размером $3\times3\times3$. Отметим центры 12 единичных кубиков, которые выходят на поверхность куба двумя своими гранями и не являются ни угловыми, ни центральными на гранях. Поскольку эти 12 точек удалены от центра куба на расстояние $\sqrt{2}$, то сферы радиуса $1/\sqrt{2}$ с центрами в них (рис. 8) касаются центрального шара (рис. 9). Можно ли расположить 13 одинаковых шаров, чтобы все они касались одного шара того же радиуса? Дж. Грегори (1638—1675) надеялся, что можно. И. Ньютон (1643—1727) утверждал, что нельзя. Точку в их споре поставила в 1953 г. статья К. Шютте и Б. Л. Ван дер Вардена. Прав оказался Ньютон. ■

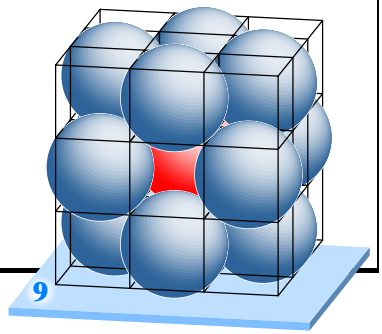
Шары рисунка 9 являются частью важной конструкции. Рассмотрим все точки пространства с целыми координатами и отметим те из них, сумма координат которых четна. Рассмотрим шары радиуса $1/\sqrt{2}$ с центрами в отмеченных точках. Доля объема, которую занимают эти шары, равна $\pi/\sqrt{18}\approx0{,}7405$. И. Кеплер в 1611 г. выдвинул гипотезу, что это наиболее плотная упаковка шаров. Эта задача вошла (под номером 18) и в список наиболее важных задач математики XX в., составленный в 1900 г. Д. Гильбертом. В 1998 г. Хайес и Фергюсон при помощи весьма сложных рассуждений свели задачу к рассмотрению примерно 5000 конфигураций, которое и выполнили на компьютере. ■

Полушар радиуса r, а также два тела — цилиндр и конус, радиусы оснований и высота каждого из которых равны r, Архимед расположил, как показано на рисунке, и заметил, что площадь πr^2 любого горизонтального сечения цилиндра равна сумме $\pi h^2+\pi(r^2-h^2)$ площадей сечений конуса и полушара. Таким образом, сумма объемов полушара и конуса равна объему цилиндра. Зная объем конуса $\pi r^3/3$ и цилиндра πr^3, Архимед нашел объем шара

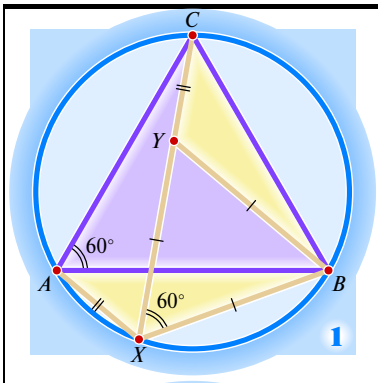
$$2\left(\pi r^3-\frac{1}{3}\pi r^3\right)=\frac{4}{3}\pi r^3.$$ ■

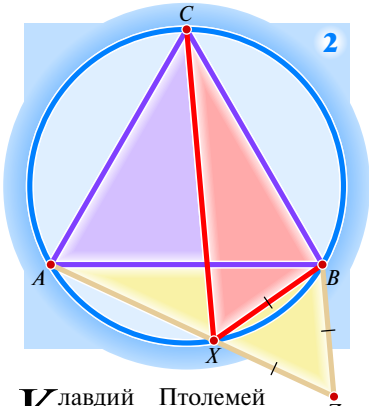






ВПИСАННЫЕ МНОГОУГОЛЬНИКИ





Клавдий Птолемей (ок. 100—178) — астроном, математик, географ эпохи позднего эллинизма, развивавший геоцентрическую систему. Уроженец Египта, жил и работал в Александрии. О жизни его почти ничего не известно.

Главное произведение — «Великое собрание» («Алмагест»). В нем содержатся сведения по тригонометрии, в том числе сферической. Ввел деление градуса на минуты и секунды. В «Географии» заложил основы картографии. Исходя из шарообразности Земли, строил разные проекции, но на практике пользовался стереографической. Положение определял по долготе и широте. Следуя Гиппарху из Никеи, применял эксцентрические круги и эпициклы для объяснений видимых движений Солнца, Луны и планет.

Исходя из равенства произведения диагоналей вписанного четырехугольника сумме произведений его противоположных сторон, определил длины хорд дуг величиной 1,5° и 0,75° и вычислил по ним приближенно хорду дуги величиной 1°. Составил таблицу хорд, соответствующих дугам от 0° до 180°. По Птолемею, $\sin 1^\circ = 0{,}017268$ и $\pi = 3{,}14166$ (на самом деле 0,0127453... и 3,14159...). ■

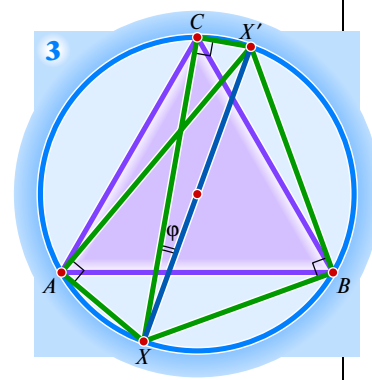
Чему равна сумма расстояний от точки X дуги AB описанной окружности равностороннего треугольника ABC (рис. 1) до вершин A и B? Что такое теорема Птолемея? Для каких n существует вписанный в окружность выпуклый n-угольник $A_1A_2\ldots A_n$ и расположенная внутри многоугольника отличная от центра окружности точка P, из которой все стороны видны под равными углами, а длины всех отрезков A_1P, A_2P, ..., A_nP — целые числа? Эти вопросы неожиданным образом очень тесно взаимосвязаны.

Докажем равенство $AX + BX = CX$. **I способ.** Отложим на XC отрезок $XY = XB$. По теореме о вписанном угле, $\angle CXB = \angle CAB = 60^\circ$. Поэтому $\triangle XBY$ равносторонний. При повороте вокруг B на 60° точка C переходит в A, а Y — в X. Поэтому треугольники CBY и ABX конгруэнтны, $CY = AX$. Следовательно, $CX = CY + YX = AX + BX$.

II способ. Построим на отрезке XB вовне правильный треугольник XZB (рис. 2). При повороте вокруг B на 60° точка C переходит в точку A, X — в Z, а $\triangle CXB$ — в $\triangle AZB$. Значит, $CX = AZ$. Поскольку сумма величин противоположных углов вписанного четырехугольника $AXBC$ равна 180°, имеем $\angle AXB = 120^\circ$, так что точки A, X, Z лежат на одной прямой. Следовательно, $AX + BX = AX + XZ = AZ = CX$.

III способ. Обозначим $AB = BC = CA = l$, $AX = a$, $BX = b$ и $CX = c$. По теореме косинусов из $\triangle AXC$ и $\triangle CXB$ находим

$$l^2 = a^2 + c^2 - 2ac\cos 60^\circ,$$
$$l^2 = b^2 + c^2 - 2bc\cos 60^\circ,$$



Вычитая почленно, получаем $a^2 - b^2 - ac + bc = 0$, откуда

$$(a-b)(a+b-c) = 0.$$

Осталось разделить на $a - b$. (Случай $a = b$ разберите самостоятельно.)

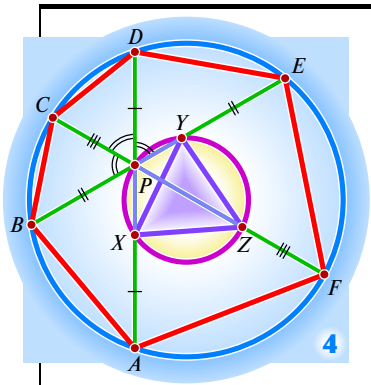
IV способ. Проведем диаметр XX' (рис. 3). Обозначим $\angle CXX' = \varphi$. Тогда треугольники XAX', XBX', XCX' — прямоугольные с гипотенузой XX'. Следовательно, $AX = XX'\cos(60^\circ - \varphi)$ и $BX = XX'\cos(60^\circ + \varphi)$, откуда

$$AX + BX = XX'(\cos(60^\circ - \varphi) + \cos(60^\circ + \varphi)) = XX' \cdot 2\cos 60^\circ \cos\varphi = XX'\cos\varphi = CX.$$

V способ. Для вписанного четырехугольника $AXBC$ запишем теорему Птолемея: $AX \cdot BC + BX \cdot AC = CX \cdot AB$. Осталось разделить на длину стороны треугольника ABC. ■

Мы пятью способами доказали равенство $AX + BX = CX$. Выведем из него следующую теорему.

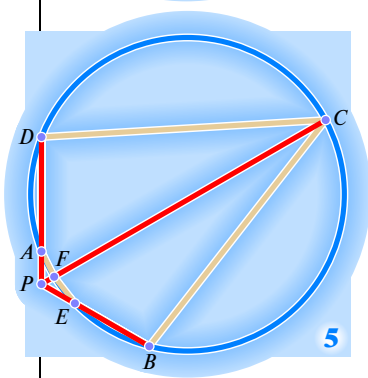
Теорема 1. ***Если диагонали AD, BE и CF вписанного шестиугольника $ABCDEF$ пересекаются в точке P под углом 60° друг к другу (рис. 4), то $AP + CP + EP = BP + DP + FP$.***



Доказательство. Проведем через точку P окружность, концентрическую описанной окружности шестиугольника. Тогда

$$AP+CP+EP=AX+XP+CP+EY+YP,$$
$$BP+DP+FP=BP+DP+FZ+ZP.$$

Поскольку $AX=DP$, $BP=EY$, $CP=FZ$, осталось проверить равенство $XP+YP=ZP$. В силу теоремы о вписанном угле, $\angle YXZ=\angle YPZ=60^\circ$ и $\angle XYZ=\angle XPZ=60^\circ$. Треугольник XYZ равносторонний! ■



Аналогичное теореме 1 утверждение можно сформулировать и в случае, когда точка P лежит вне шестиугольника. А именно, если под углом 60° друг к другу провели три прямые, которые пересекли некоторую окружность, как показано на рисунке 5, то легко доказать равенство

$$-AP+CP-EP=BP+DP-FP.$$

Некоторые отрезки «взяты со знаком минус», только и всего. ■

Для вписанного шестиугольника $ABCDEF$, диагонали AD, BE, CF которого под равными углами пересекаются в точке P, как мы доказали, выполнены равенства

$$\begin{cases} a+c+e=b+d+f, \\ ad=be=cf, \end{cases} \qquad (*)$$

где $a=PA$, $b=PB$, $c=PC$, $d=PD$, $e=PE$ и $f=PF$. Выясним, могут ли все длины a, b, . . . , f выражаться целыми числами, если точка P не является центром окружности.

Теорема 2. ***Если на пересекающихся под углом 60° друг к другу прямых от точки их пересечения P отложить отрезки PA=a, PB=b, . . . , PF=f, для которых выполнены равенства (*), то получим вписанный шестиугольник ADCDEF.***

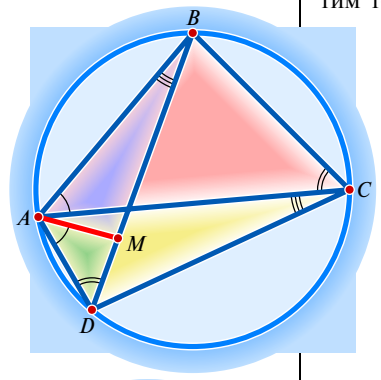
Доказательство. Отложим сначала на двух прямых отрезки $PA=a$, $PB=b$, $PD=d$ и $PE=e$. Точки A, B, D и E лежат на одной окружности. Рассмотрим точки C' и F' пересечения этой окружности с третьей прямой и обозначим $c'=PC'$, $f'=PF'$. Тогда

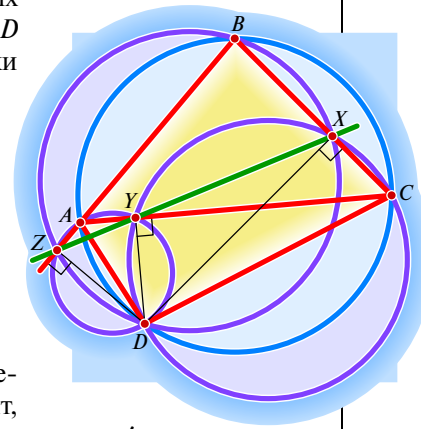
$$\begin{cases} a+c'+e=b+d+f', \\ ad=be=c'f'. \end{cases} \qquad (**)$$

Из систем $(*)$ и $(**)$ имеем:

$$\begin{cases} c-c'=f-f', \\ cf=c'f'. \end{cases}$$

Запишем первое уравнение в виде $c-f=c'-f'$, возведем в квадрат и прибавим к результату учетверенное второе уравнение: $(c+f)^2=(c'+f')^2$. Значит, $c+f=c'+f'$. Вспомнив уравнение $c-f=c'-f'$, получаем: $c=c'$, $f=f'$. (Можно доказать эти равенства и иначе: c и $-f$ — корни квадратного уравнения $x^2-(c-f)x=cf$, «другими» корнями которого являются c' и $-f'$.)





Теорема Птолемея. ***Если четырехугольник ABCD вписан в окружность, то AC·BD=AB·CD+AD·BC.***

Доказательство. I способ. Домножим равенство Птолемея на $\frac{1}{2}\sin\varphi$, где φ — угол между диагоналями четырехугольника:

$$\frac{1}{2}AC\cdot BD\sin\varphi=$$
$$=\frac{1}{2}BC\cdot DA\sin\varphi+\frac{1}{2}AB\cdot CD\sin\varphi,$$

Левая часть полученного равенства — площадь четырехугольника, а в правой части находятся площади $\frac{1}{2}BC\cdot DA\sin\varphi$ и $\frac{1}{2}AB\times$ $\times CD\sin\varphi$ треугольников DAB' и DCB', где B' — образ точки B при симметрии относительно серединного перпендикуляра отрезка AC. (Дело в том, что

$$\angle B'AD=\frac{1}{2}(\overset{\smile}{B'C}+\overset{\smile}{CD})=$$
$$=\frac{1}{2}(\overset{\smile}{AB}+\overset{\smile}{CD})=\varphi.$$

Аналогично, $\angle B'CD=180^\circ-\varphi$.)

II способ. На диагонали BD отметим точку M так, что $\angle MAD=\angle BAC$. Тогда $\triangle MAD\sim\triangle BAC$, $MD/BC=AD/AC$. Одновременно $\triangle BAM\sim\triangle CAD$, $BM/CD=AB/AC$. Следовательно,

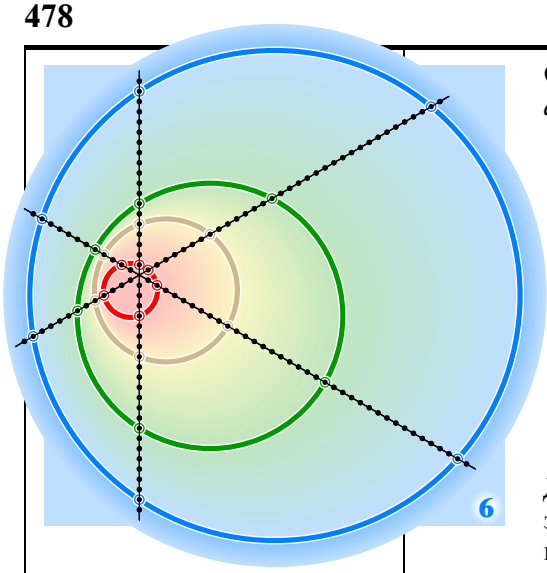
$$BD=BM+MD=$$
$$=\frac{AB\cdot CD}{AC}+\frac{AD\cdot BC}{AC},$$

откуда $AC\cdot BD=AB\cdot CD+AD\cdot BC$.

III способ. Пусть X, Y и Z — основания перпендикуляров, опущенных из точки D на прямые BC, CA и AB соответственно. Четырехугольник $DCXY$ вписан в окружность с диаметром CD. Следовательно, по теореме синусов, $XY=2DC\sin\angle ACB$. Аналогично, $YZ=2DA\times$ $\times\sin\angle CAB$ и $ZX=2DB\times$ $\times\sin\angle ABC$. По теореме Симсона, точки X, Y и Z лежат на одной прямой, так что $XZ=XY+YZ$. Следовательно,

$$DC\sin\angle ACB=$$
$$=DA\sin\angle BAC+$$
$$+DB\sin\angle ABC.$$

Осталось домножить это равенство на диаметр описанной окружности треугольника ABC. ■


6

Осталось предъявить решение системы (∗) в натуральных числах a, b, c, d, e, f, не все из которых равны друг другу. На рисунке 6 приведены даже четыре примера.

Как придумать бесконечно много таких примеров? Положим $a=pqx$, $d=rsy$, $b=pry$, $e=qsx$, $c=psx$ и $f=qry$, где $q\neq s$. Равенства $ad=be=cf$ верны. Равенство $a+c+e=b+d+f$ принимает вид

$$x(pq+ps+qs)=y(pr+rs+qr).$$

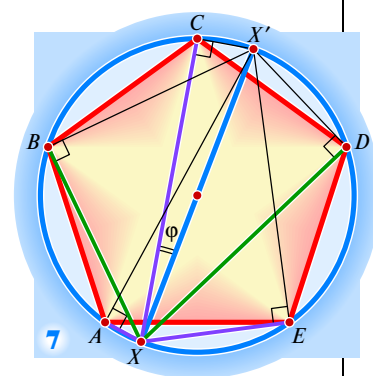
Значит, достаточно взять $x=pr+rs+qr$ и $y=pq+ps+qs$. ■

Продолжим изучение вписанных многоугольников.

Теорема 3. ***Если точка X лежит на дуге AE описанной окружности правильного пятиугольника ABCDE, то AX+CX+EX=BX+DX.***

Доказательство. I способ. Проведем диаметр XX' (рис. 7). Обозначим $\angle CXX'=\varphi$ и $XX'=d$. Тогда треугольник AXX' прямоугольный, откуда $AX=XX'\cos\angle AXX'=d\cos(72^\circ+\varphi)$. Аналогично, $BX=d\cos(36^\circ+\varphi)$, $CX=d\cos\varphi$, $DX=d\cos(36^\circ-\varphi)$ и $EX=d\cos(72^\circ-\varphi)$. Значит, осталось проверить тождество

$$\cos(72^\circ+\varphi)+\cos\varphi+\cos(72^\circ-\varphi)=\cos(36^\circ+\varphi)+\cos(36^\circ-\varphi).$$

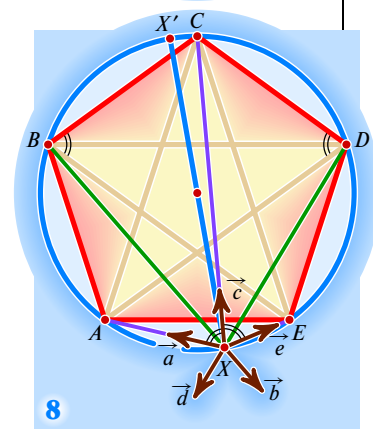

7

Поскольку $\cos(72^\circ+\varphi)+\cos(72^\circ-\varphi)=2\cos 72^\circ\times\cos\varphi$ и $\cos(36^\circ+\varphi)+\cos(36^\circ-\varphi)=2\cos 36^\circ\cos\varphi$, достаточно доказать равенство $\cos 36^\circ-\cos 72^\circ=1/2$. Домножим его левую часть на $\sin 36^\circ$:

$$\cos 36^\circ\sin 36^\circ-\cos 72^\circ\sin 36^\circ=\frac{1}{2}\sin 72^\circ-\frac{1}{2}(\sin 108^\circ-\sin 36^\circ)=\frac{1}{2}\sin 36^\circ.$$

II способ. Отложим векторы $\vec a$, $\vec c$, $\vec e$ единичной длины вдоль лучей XA, XC, XE, а векторы $\vec b$, $\vec d$ — вдоль лучей BX и DX (рис. 8). Тогда $AX=\vec a\cdot\overrightarrow{XX'}$, $CX=\vec c\cdot\overrightarrow{XX'}$, $EX=\vec e\cdot\overrightarrow{XX'}$, $BX=-\vec b\cdot\overrightarrow{XX'}$, $DX=-\vec d\cdot\overrightarrow{XX'}$. Значит,

$$AX+CX+EX-BX-DX=(\vec a+\vec c+\vec e+\vec b+\vec d)\overrightarrow{XX'}.$$


8

Проверим равенство $\vec a+\vec c+\vec e+\vec b+\vec d=\vec 0$. Для этого заметим, что по теореме о вписанном угле прямые AX, BX, CX, DX и EX пересекаются под равными углами. Если бы сумма $\vec s=\vec a+\vec c+\vec e+\vec b+\vec d$ была отлична от $\vec 0$, то вектор $\vec s$ изменялся бы при повороте на 72°. Но при этом повороте слагаемые $\vec a$, $\vec d$, $\vec b$, $\vec e$, $\vec c$ всего лишь переставляются местами, переходя в $\vec d$, $\vec b$, $\vec e$, $\vec c$, $\vec a$ соответственно.

III способ. Применим теорему Птолемея к четырехугольникам $ABCX$, $BCDX$, $CDEX$, $DEXA$ и $EXAB$:

$$AX\cdot BC+CX\cdot AB=BX\cdot AC,$$
$$CX\cdot BD=BX\cdot CD+DX\cdot BC,$$
$$CX\cdot DE+EX\cdot CD=DX\cdot CE,$$
$$AX\cdot DE+EX\cdot AD=DX\cdot AC,$$
$$AX\cdot BE+EX\cdot AB=BX\cdot AE.$$

Рассмотрим любой треугольник ABC и любую точку X. Всегда ли $AX+BX\geqslant CX$? Конечно, нет: точки A, B, X могут быть расположены близко друг к другу, а точка C — далеко от них.

А для равностороннего треугольника неравенство $AX+BX\geqslant CX$ выполнено. Выясним, как должны быть расположены точки A, B, C, чтобы это неравенство выполнялось для любой точки X.

Если оно выполнено для любой точки X, то оно верно, в частности, для $X=A$. Поэтому $AA+BA\geqslant CA$, то есть $BA\geqslant CA$.

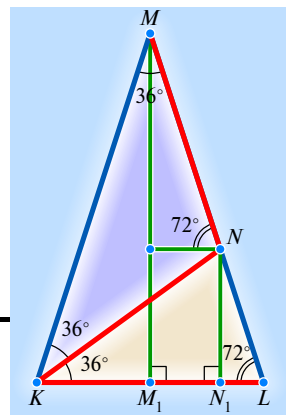
Аналогично, подставив $X=B$, приходим к неравенству $AB\geqslant CB$. Значит, AB — наибольшая сторона треугольника ABC. (Впрочем, ABC — не обязательно треугольник. Рассуждения годятся и для случая, когда точки лежат на прямой (C — между A и B), то есть треугольник «вырождается» в отрезок.)

Из неравенства Птолемея $AX\cdot BC+BX\cdot AC\geqslant CX\cdot AB$ следует, что верно и обратное: если $AB\geqslant AC$ и $AB\geqslant BC$, то $AX+BX\geqslant CX$ для любой точки X. ■

Проведем в треугольнике KLM с углами $\angle K=\angle L=72^\circ$ и $\angle M=36^\circ$ биссектрису KN. Треугольники LKN и KNM равнобедренные: $MN=NK=KL$. Пусть для определенности длины этих трех отрезков равны 1.

Опустим перпендикуляры MM_1 и NN_1 на основание треугольника:

$$M_1N_1=MN\cos 72^\circ=\cos 72^\circ,$$
$$\cos 36^\circ-\cos 72^\circ=KN_1-M_1N_1=KM_1=1/2.$$ ■



Складывая, получаем

$$AX(2a+d)+CX(2a+d)+EX(2a+d)=BX(2a+d)+DX(2a+d),$$

где a — длина стороны, d — длина диагонали. Осталось разделить обе части этого равенства на $2a+d$.

Аналогичным теореме 3 и точно так же доказываемым свойством обладает любой правильный многоугольник с нечетным числом сторон. ■

Если хорды KM и LN окружности пересекаются в точке P, то $KP\cdot MP=LP\cdot NP$. Верно и обратное: если отрезки KM и LN пересекаются в точке P и $KP\cdot MP=LP\cdot NP$, то точки K, L, M и N лежат на одной окружности. Поэтому, взяв $KP=4$, $MP=1$ и $LP=NP=2$ (или $KP=6$, $MP=2$, $LP=3$ и $NP=4$), получаем вписанный четырехугольник (рис. 9), стороны которого видны под равными углами из точки P, не являющейся центром окружности. ■

Пусть $n=2k$ и $k>3$. Предположим, что все стороны вписанного n-угольника видны из точки P под равными углами, а центр O окружности лежит внутри угла A_2PA_3 (рис. 10). Обозначим $\angle OPA_2=\theta$. Опустим перпендикуляры PB_1, PB_2 и PB_3 на прямые PA_1, PA_2 и PA_3. Тогда $PB_1=OP\cos\left(\frac{\pi}{k}+\theta\right)$, $PB_2=OP\cos\theta$ и $PB_3=OP\cos\left(\frac{\pi}{k}-\theta\right)$, откуда

$$PB_1+PB_3=2OP\cos\theta\cos\frac{\pi}{k}=2PB_2\cos\frac{\pi}{n},$$

то есть $\cos\frac{\pi}{k}=\frac{PB_1+PB_3}{2PB_2}$.

Если бы длины всех отрезков PA_i были целыми, то длины

$$PB_1=\frac{PA_1-PA_{k+1}}{2},\quad PB_2=\frac{PA_2-PA_{k+2}}{2},\quad PB_3=\frac{PA_3-PA_{k+3}}{2}$$

были бы рациональными. Но при $k>3$ число $\cos\frac{\pi}{k}$ иррационально. ■

Пусть n нечетно и $n>3$. Предположим, что все стороны вписанного n-угольника $A_1A_2\ldots A_n$ видны из точки P под равными углами, а центр O окружности не совпадает с точкой P. Продолжим каждый из лучей A_kP, где $k=1,\ldots,n$, до пересечения с окружностью в точке B_k. По свойству хорд,

$$PA_1\cdot PB_1=PA_2\cdot PB_2=\ldots=PA_n\cdot PB_n=c$$

для некоторого числа c. Предположим, что все длины PA_1, PA_2, $\ldots$, PA_n целые. Подставив выражения $PB_k=\frac{c}{PA_k}$ в равенство

$$PA_1+PA_2+\ldots+PA_n=PB_1+PB_2+\ldots+PB_n,$$

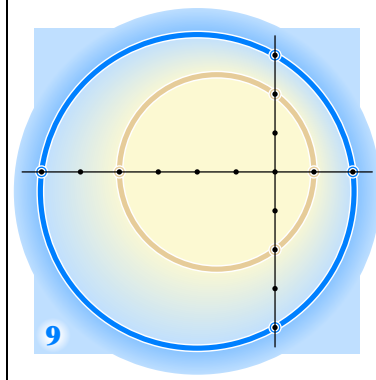
получаем

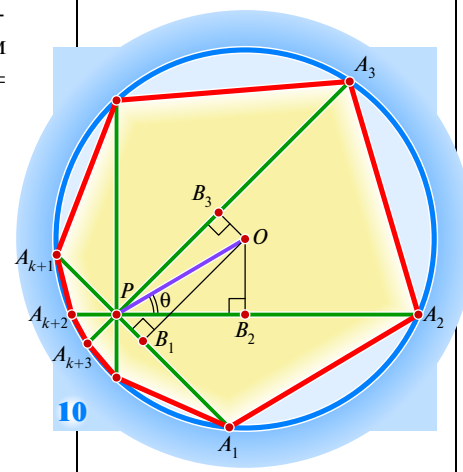
$$PA_1+PA_2+\ldots+PA_n=c\left(\frac{1}{PA_1}+\frac{1}{PA_2}+\ldots+\frac{1}{PA_n}\right),$$

откуда c — рациональное число. Значит, рациональны и все $PB_k=\frac{c}{PA_k}$. Но мы уже доказали, что такого не бывает. ■

Расстояния от точки D до вершин A и B равностороннего треугольника ABC равны 2 и 3. Чему может равняться DC?

Ответ: $1\leqslant DC\leqslant 5$. ■





Правильный n-угольник $A_1A_2\ldots A_n$ вписан в окружность радиуса R с центром O. Если P — произвольная точка этой окружности, то сумма проекций на прямую OP всех n векторов, соединяющих точку P с вершинами n-угольника, равна nR.

Доказательство. Сумма векторов с началом в точке P и концами в вершинах n-угольника равна

$$\overrightarrow{PA_1}+\overrightarrow{PA_2}+\ldots+\overrightarrow{PA_n}=\overrightarrow{PO}+\overrightarrow{OA_1}+\overrightarrow{PO}+\overrightarrow{OA_2}+\ldots+\overrightarrow{PO}+\overrightarrow{OA_n}=n\overrightarrow{PO}+(\overrightarrow{OA_1}+\overrightarrow{OA_2}+\ldots+\overrightarrow{OA_n})=n\overrightarrow{PO}.$$ ■

Если точка X лежит на дуге CD описанной окружности квадрата $ABCD$, то $AX+CX=BX\sqrt{2}$.

Указание. Чтобы доказать это равенство, примените теорему Птолемея к вписанному четырехугольнику $ABCX$. ■

Радикальная ось

Радикальная ось двух неконцентрических окружностей — это множество точек M, степени которых относительно этих двух окружностей одинаковы: $O_1M^2-r_1^2=O_2M^2-r_2^2$, где O_1 и O_2 — центры окружностей, r_1 и r_2 — их радиусы.

Рассмотрим окружность с центром O и радиусом r. Проведем через некоторую точку P, расположенную внутри окружности, две прямые, одна из которых пересекает окружность в точках A_1 и B_1, а другая — в точках A_2 и B_2 (рис. 1). По теореме о вписанном угле, $\angle A_1A_2P=\angle PB_1B_2$ и $\angle A_2A_1P=\angle PB_2B_1$. Значит, треугольники A_1A_2P и B_2B_1P подобны, так что $\dfrac{A_1P}{B_2P}=\dfrac{A_2P}{B_1P}$, откуда $A_1P\cdot B_1P=A_2P\cdot B_2P$.

Проведя диаметр A_3B_3 через точку P и обозначив $OP=p$, имеем $A_3P=r-p$ и $B_3P=r+p$, откуда

$$A_1P\cdot B_1P=A_2P\cdot B_2P=A_3P\cdot B_3P=(r-p)(r+p)=r^2-p^2.$$

Аналогично, взяв точку Q вне окружности, получаем из подобия треугольников A_1A_2Q и B_2B_1Q равенство $A_1Q\cdot B_1Q=A_2Q\cdot B_2Q$ (рис. 2). Проведя касательную QK, из подобия треугольников QKA_1 и QB_1K находим $\dfrac{QK}{QA_1}=\dfrac{QB_1}{QK}$, то есть $QA_1\cdot QB_1=QK^2$. Очевидно, $QK^2=QA_3\cdot QB_3=(q-r)(q+r)=q^2-r^2$, где $q=OQ$. ■

Степень точки относительно окружности — это разность квадрата расстояния от этой точки до центра окружности и квадрата радиуса этой окружности. Таким образом, степени точек, расположенных вне окружности, положительны, степени внутренних точек отрицательны, а степени точек самой окружности равны нулю.

Для касающихся окружностей радикальная ось — это общая касательная (рис. 3). ■

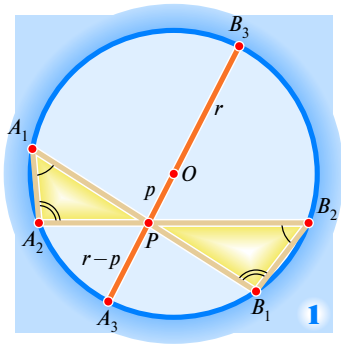
Введем координаты на плоскости так, что $O_1=(0;0)$ и $O_2=(d;0)$. Для точки $M(x;y)$ имеем $O_1M^2=x^2+y^2$ и $O_2M^2=(x-d)^2+y^2$. Поэтому равенство степеней точек можно записать в виде

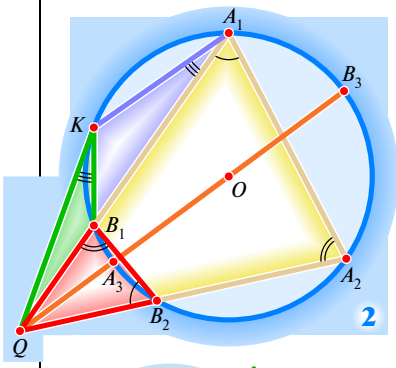
$$x^2+y^2-r_1^2=(x-d)^2+y^2-r_2^2,$$

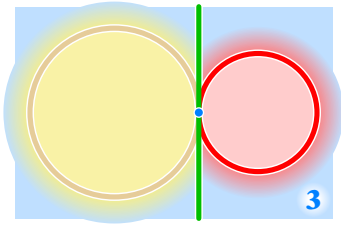
то есть $2xd=d^2+r_1^2-r_2^2$, так что радикальная ось двух окружностей — прямая, перпендикулярная прямой O_1O_2.

Таким образом, для окружностей, пересекающихся в двух точках A и B, радикальная ось — это прямая AB, поскольку степени обеих точек A и B относительно обеих окружностей равны нулю (рис. 4). ■

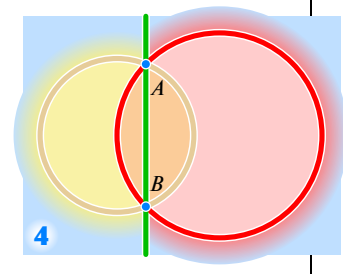
Рассмотрим три окружности с разными центрами. Для каждых двух из них есть радикальная ось. Если центры окружностей лежат на одной прямой, то радикальные оси перпендикулярны этой прямой и поэтому параллельны между собой (рис. 5). Если же центры не лежат на одной прямой, то параллельности нет и поэтому

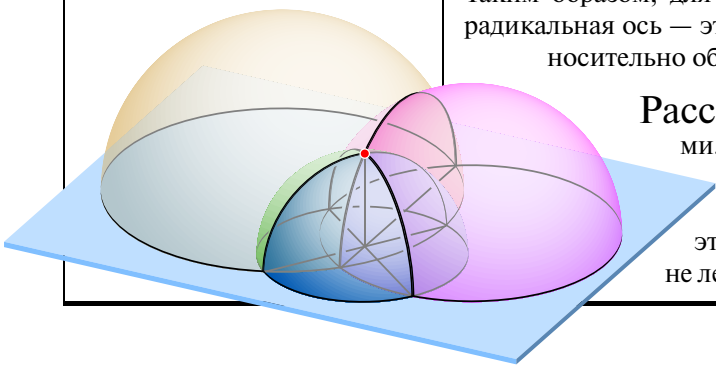


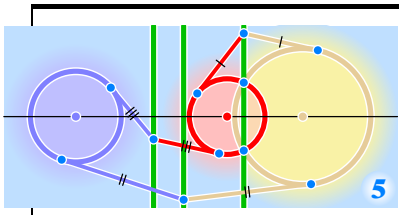




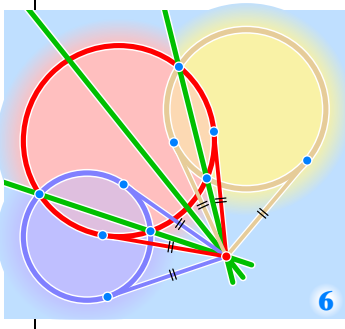
Рассмотрим три пересекающиеся полусферы. Пересечение двух полусфер — полуокружность, перпендикулярная плоскости. Один конец этой полуокружности лежит внутри, а другой — вне третьей окружности. Поэтому полуокружность пересекает третью полусферу в некоторой точке. В ней пересекаются все три полусферы. Поскольку при ортогональном проецировании полуокружности переходят в общие хорды окружностей, то эти три хорды проходят через проекцию точки пересечения полуокружностей. ■



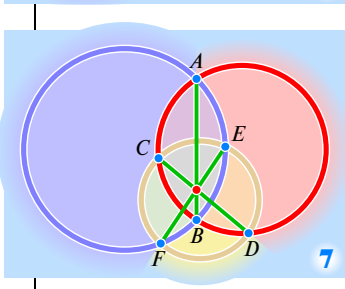




можно рассмотреть точку пересечения радикальной оси первой и второй окружностей с радикальной осью второй и третьей окружности. Эта точка обладает двумя свойствами: во-первых, равны ее степени относительно первой и второй окружностей, во-вторых, равны ее степени относительно второй и третьей окружностей. Поскольку из равенств $a=b$ и $b=c$ следует равенство $a=c$, то степени рассматриваемой точки относительно первой и третьей окружностей тоже совпадают и все три радикальные оси пересекаются в одной точке — радикальном центре трех окружностей.



Если точка пересечения лежит вне окружностей, как на рисунке 6, то длины касательных, проведенных из нее к окружностям, равны. Если же точка лежит внутри окружностей (рис. 7), то она является точкой пересечения хорд AB, CD и EF. ■



Наличие общей точки пересечения хорд AB, CD и EF можно доказать и без помощи радикальных осей. В обозначениях рисунка 8 имеем

$$EP\cdot PF_1 = AP\cdot PB = CP\cdot DB = EP\cdot PF_2,$$

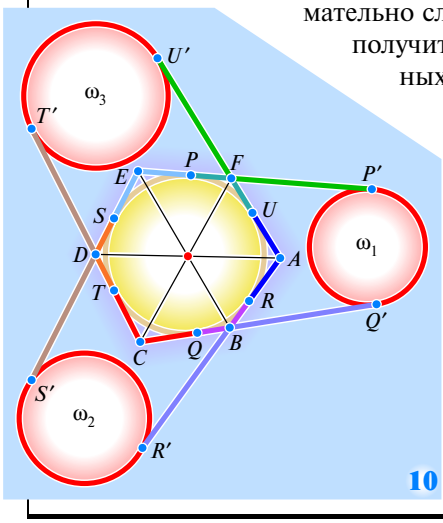
откуда $EP\cdot PF_1 = EP\cdot PF_2$, то есть $F_1 = F_2$. ■

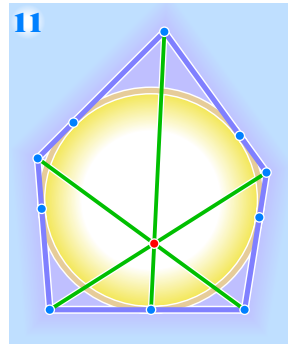
Ш. Ж. Брианшон (1783—1864) доказал, что если все стороны шестиугольника касаются окружности, то его диагонали пересекаются в одной точке (рис. 9). Для доказательства на лучах EF, CB, AB, ED, CD и AF возьмем точки P', Q', R', S', T' и U' так, что

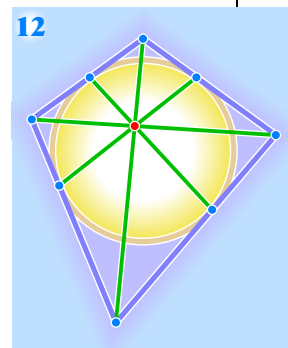
$$PP' = QQ' = RR' = SS' = TT' = UU'$$

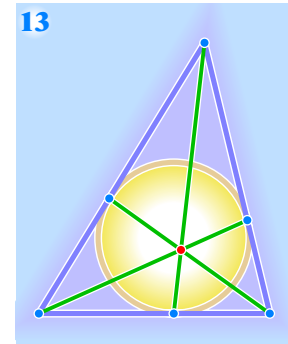
(длина этих отрезков может иметь любое удобное значение) и построим окружности ω_1, ω_2 и ω_3 (рис. 10). Очевидно, прямая AD — радикальная ось окружностей ω_2 и ω_3, BE — радикальная ось ω_1 и ω_2, а CF — ω_1 и ω_3. А радикальные оси пересекаются в одной точке!

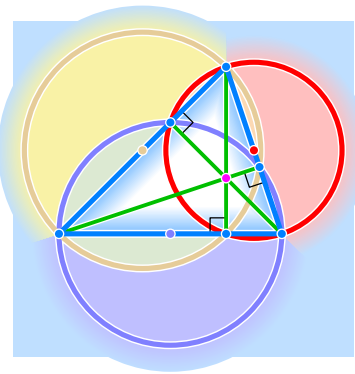
Если позволить вершинам шестиугольника сливаться, внимательно следя при этом за обозначениями, то можно получить несколько интересных теорем: об описанных пятиугольнике (рис. 11), четырехугольнике (рис. 12) и даже треугольнике (рис. 13). ■



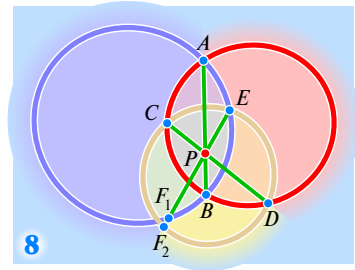


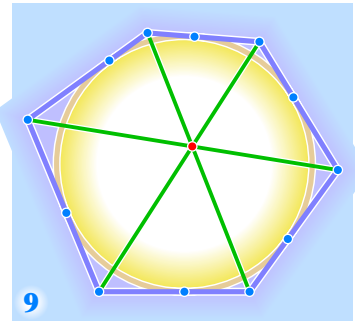






Высоты треугольника являются радикальными осями окружностей, построенных на его сторонах как на диаметрах. Поскольку радикальные оси пересекаются в одной точке, то высоты любого треугольника пересекаются в одной точке — ортоцентре треугольника! ■





Площадь суммы фигур

Герман Минковский (1864—1909) — немецкий математик, друг Д. Гильберта. Продолжил исследования Гаусса и Дирихе квадратичных форм, использовал при этом геометрические методы и тем самым основал «геометрию чисел». Для любой симметричной относительно точки O ограниченной выпуклой фигуры F длина по Минковскому отрезка AB — это AB/OG, где $AB \| OG$ и точка G лежит на границе фигуры F. Для любых точек A, B и C верно неравенство треугольника $|AB|_M + |BC|_M \geqslant |AC|_M$. Таким образом плоскость превращается в метрическое пространство, круг единичного радиуса которого — фигура F.

Для любого числа $p \geqslant 1$ и любых чисел $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ он доказал получившее его имя неравенство

$$\left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{1/p} \leqslant \leqslant \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p}.$$

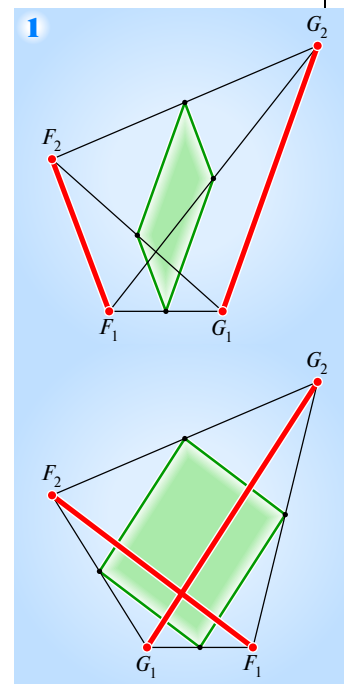
Минковский доказал, что для любой системы векторов пространства, не все из которых лежат в одной плоскости и сумма которых равна $\vec{0}$, существует и с точностью до параллельного переноса единственен такой многогранник, что эти векторы перпендикулярны его граням, а их длины равны площадям этих граней.

Рассматривая квадратичную форму $x^2 + y^2 + z^2 - t^2$, Минковский дал математическое обоснование специальной теории относительности А. Эйнштейна. ■

Для любых двух выпуклых ограниченных фигур F и G рассмотрим множество середин отрезков, у которых один конец принадлежит фигуре F, а другой — G. Брунн и Минковский доказали, что площадь этого множества не меньше числа $(\sqrt{S_F} + \sqrt{S_G})^2/4$.

Найдем множество середин M всевозможных отрезков AB, где $A \in F_1F_2$ и $B \in G_1G_2$ (рис. 1). Сначала пусть $B = G_1$. При движении точки A по отрезку F_1F_2 точка M пробегает среднюю линию треугольника $F_1F_2G_1$. Сдвинем точку B по направлению к G_2 — средняя линия сдвинется в том же направлении на вдвое меньшее расстояние. Следовательно, искомое множество — параллелограмм с вершинами в серединах отрезков F_1G_1, G_1F_2, F_2G_2 и G_2F_1 (параллелограмм вырождается в отрезок, если $F_1F_2 \| G_1G_2$).

Если отрезки F_1F_2 и G_1G_2 пересекаются, то искомое множество — параллелограмм с вершинами в серединах сторон выпуклого четырехугольника $F_1G_1F_2G_2$. Если отрезки F_1F_2 и G_1G_2 не лежат в одной плоскости, то параллелограмм является сечением тетраэдра $F_1F_2G_1G_2$ плоскостью, параллельной ребрам F_1F_2 и G_1G_2 и проходящей посередине между ними (рис. 2). ■

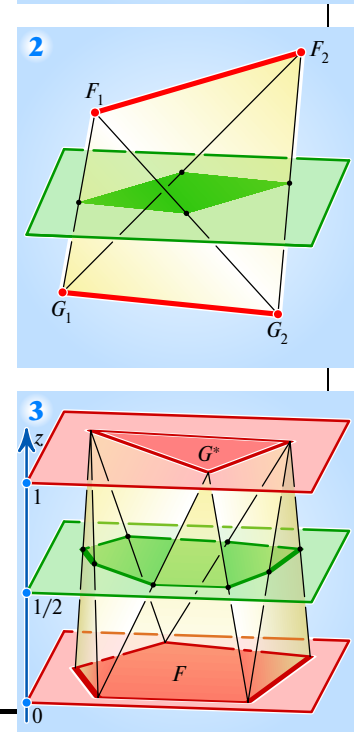


Вместо отрезков рассмотрим на плоскости Oxy выпуклые ограниченные фигуры F и G и поднимем G на единицу вдоль оси аппликат, получив фигуру G^*. Рассечем выпуклую оболочку фигур F и G^* плоскостью, заданной уравнением $z = 1/2$, то есть проходящей посередине между плоскостями фигур F и G^*. Каждый отрезок AB будет пересечен в его середине. Поскольку середина отрезка, соединяющего точки $A(x_f; y_f; 0)$ и $B^*(x_g; y_g; 1)$, — это точка

$$M^*\left(\frac{x_f + x_g}{2}; \frac{y_f + y_g}{2}; \frac{1}{2}\right),$$

то сечение, если его спроецировать на плоскость Oxy, дает множество середин M отрезков AB. Это множество будем называть полусуммой фигур F и G (рис. 3).

Полусумма любых двух выпуклых многоугольников — выпуклый многоугольник, периметр которого — полусумма их периметров. Например, полусумма прямоугольников размеров $a \times b$ и $c \times d$ с соответственно параллельными сторонами — прямоугольник размера $\frac{a+c}{2} \times \frac{b+d}{2}$. ■



Рассмотрим плоскость, заданную уравнением $z=\mu$, где $0<\mu<1$. Обозначим $\lambda=1-\mu$. Плоскость делит отрезок PQ в отношении $PM^*/M^*Q=\mu/\lambda$. Поэтому $M^*=(\lambda x_f+\mu x_g;\ \lambda y_f+\mu y_g;\ \mu)$.

Определение. Фигура $\lambda F+\mu G$ — это множество точек $M(\lambda x_f+\mu x_g;\ \lambda y_f+\mu y_g)$, где $(x_f;\ y_f)\in F$ и $(x_g;\ y_g)\in G$. ■

Откажемся от равенства $\lambda+\mu=1$. Взяв $\lambda=\mu=1$, получаем сумму Минковского фигур F и G. Конечно, если сдвинуть начало координат, то фигура $F+G$ сдвинется на тот же вектор, в отличие от полусуммы и вообще от комбинации $\lambda F+\mu G$, где $\lambda+\mu=1$, которые не шелохнутся. Но если условиться не различать (считать эквивалентными) фигуры, получающиеся одна из другой параллельным переносом, то можно не указывать, где выбрано начало координат, — выражение $\lambda F+\mu G$ определено с точностью до параллельного переноса. Например, если G — круг радиуса r, то $F+G$ — (с точностью до параллельного переноса) объединение всех кругов радиуса r с центрами в точках фигуры F.

Неравенство Брунна—Минковского. ***Для любых неотрицательных чисел λ и μ верно неравенство***

$$S_{\lambda F+\mu G}\geqslant(\lambda\sqrt{S_F}+\mu\sqrt{S_G})^2=\lambda^2S_F+2\lambda\mu\sqrt{S_FS_G}+\mu^2S_G.$$

Доказательство. Проверим сначала неравенство для прямоугольников F и G размеров $a\times b$ и $c\times d$ с соответственно параллельными сторонами. Очевидно, $S_{\lambda F+\mu G}=(\lambda a+\mu c)(\lambda b+\mu d)$, так что неравенство приобретает вид

$$(\lambda a+\mu c)(\lambda b+\mu d)\geqslant\lambda^2ab+2\lambda\mu\sqrt{abcd}+\mu^2cd.$$

Раскрывая скобки и приводя подобные слагаемые, приходим к неравенству

$$\lambda\mu(ad+bc)\geqslant2\lambda\mu\sqrt{abcd},$$

то есть $ad+bc\geqslant2\sqrt{abcd}$, что очевидно: $(\sqrt{ad}-\sqrt{bc})^2\geqslant0$.

Сведем общий случай к только что разобранному. Если фигура F разрезана вертикальной прямой l на части F_- и F_+, а фигура G разрезана вертикальной прямой m на части G_- и G_+, то фигура $\lambda F_-+\mu G_-$ лежит слева от прямой $\lambda l+\mu m$, а $\lambda F_++\mu G_+$ — справа.

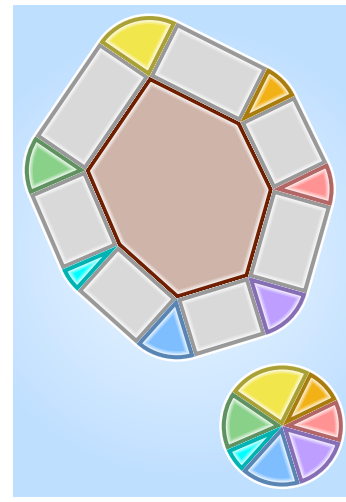
Для произвольного натурального числа n разобьем фигуру F прямыми, перпендикулярными оси абсцисс, на n частей $F_1, F_2, \ldots, F_n$ равной площади. Фигуру G тоже разобьем на части $G_1, G_2, \ldots, G_n$ равной площади.

Для любого $k=1, 2, \ldots, n$ фигура $\lambda F+\mu G$ содержит в себе фигуру $\lambda F_k+\mu G_k$, причем эти n фигур могут пересекаться только по границе, но никак не по внутренним точкам (рис. 4):

$$\lambda F+\mu G\supset\bigcup_{k=1}^{n}\lambda F_k+\mu G_k.$$

(Заметьте: мы складываем только части с одинаковыми номерами!) Каждую из n фигур F_k и G_k приблизим прямоугольником, стороны которого параллельны осям абсцисс и ординат. Площади таких прямоугольников для фигуры F равны S_F/n (с точностью до маленьких погрешностей, влияние которых можно устранить, устремив n к бесконечности), а для фигуры G — (с той же оговоркой) S_G/n. Для прямоугольников неравенство Брунна—Минковского уже доказано, поэтому имеем

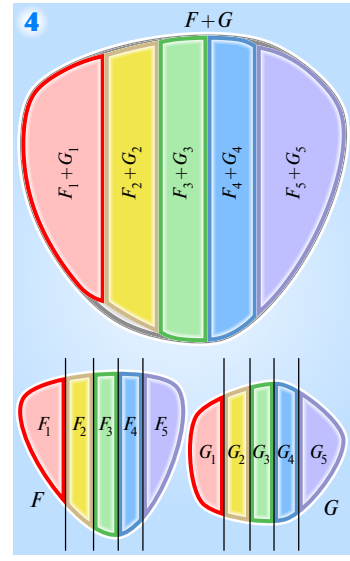
$$S_{\lambda F+\mu G}\geqslant n\cdot(\lambda\sqrt{S_F/n}+\mu\sqrt{S_G/n})^2=(\lambda\sqrt{S_F}+\mu\sqrt{S_G})^2.\ ■$$



Сумма многоугольника F периметра P и площади S и круга G единичного радиуса состоит из самого многоугольника F, прямоугольников, у каждого из которых одна сторона совпадает со стороной F, а длина другой равна 1, и секторов, из которых можно при помощи параллельных переносов собрать круг радиуса 1. Таким образом, неравенство Брунна—Минковского приобретает вид

$$S+P+\pi\geqslant S+2\sqrt{S\pi}+\pi,$$

то есть $S\leqslant\dfrac{P^2}{4\pi}=\pi\cdot\left(\dfrac{P}{2\pi}\right)^2$. Как известно, $P/(2\pi)$ — это радиус окружности периметра P. Значит, площадь любого выпуклого многоугольника не превышает площади круга такого же периметра. ■



Длины биссектрис треугольника

Можно доказать, что циркулем и линейкой в общем случае невозможно построить треугольник по его биссектрисам. Тем не менее для любых трех положительных чисел существует и единственен треугольник, длины биссектрис которого — данные числа.

Квадрат длины биссектрисы можно выразить через длины сторон треугольника:

$$AL^2 = AB\cdot AC\cdot\left(1-\frac{BC^2}{(AB+AC)^2}\right),$$

где AL — биссектриса угла A треугольника ABC.

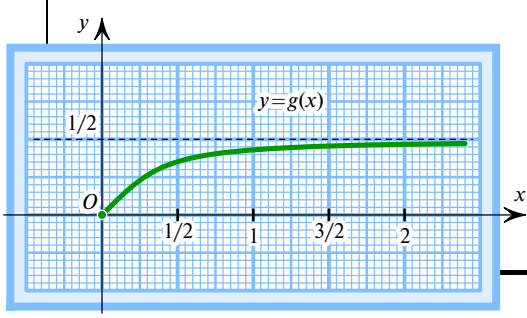
Пусть даны положительные числа l_a, l_b и l_c. Задача о существовании и единственности треугольника с предписанными длинами биссектрис сводится к доказательству утверждения: система уравнений и неравенств

$$\begin{cases} l_a^2 = bc\left(1-\dfrac{a^2}{(b+c)^2}\right), \\ l_b^2 = ca\left(1-\dfrac{b^2}{(c+a)^2}\right), \\ l_c^2 = ab\left(1-\dfrac{c^2}{(a+b)^2}\right), \\ a+b>c,\ \ b+c>a,\ \ c+a>b \end{cases}$$

имеет единственное решение в положительных числах $a=BC$, $b=AC$ и $c=AB$. ■

Обозначим $P=a+b+c$, $x=a/P$, $y=b/P$ и $z=c/P$. Тогда $x+y+z=\dfrac{a+b+c}{P}=1$ и система приобретает вид

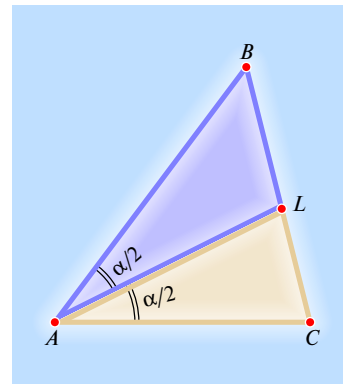
$$\begin{cases} l_a^2 = P^2yz\left(1-\dfrac{x^2}{(y+z)^2}\right), \\ l_b^2 = P^2zx\left(1-\dfrac{y^2}{(z+x)^2}\right), \\ l_c^2 = P^2xy\left(1-\dfrac{z^2}{(x+y)^2}\right), \\ x+y+z=1, \\ x+y>z,\ \ y+z>x,\ \ z+x>y. \end{cases}$$



Заменив в первом уравнении системы $y+z$ на $1-x$, получаем

$$l_a^2 = P^2yz\left(1-\frac{x^2}{(1-x)^2}\right) = P^2yz\frac{1-2x}{(1-x)^2}.$$

А неравенство $y+z>x$ эта замена переводит в неравенство $1-x>x$, то есть $x<1/2$. ■



Биссектриса AL треугольника ABC делит его на треугольники ABL и ACL. Площадь треугольника ABC можно вычислить как по известной формуле

$$S_{ABC} = \frac{AB\cdot AC\cdot\sin\alpha}{2},$$

так и по формуле

$$S_{ABC} = S_{ABL} + S_{ACL} = \frac{AB\cdot AL\cdot\sin\frac{\alpha}{2}}{2} + \frac{AC\cdot AL\cdot\sin\frac{\alpha}{2}}{2}.$$

Следовательно,

$$AB\cdot AC\cdot\sin\alpha = AB\cdot AL\cdot\sin\frac{\alpha}{2} + AC\cdot AL\cdot\sin\frac{\alpha}{2},$$

откуда

$$AL = \frac{AB\cdot AC\cdot\sin\alpha}{(AB+AC)\sin\frac{\alpha}{2}}.$$

Воспользовавшись формулой $\sin\alpha = 2\sin\dfrac{\alpha}{2}\cos\dfrac{\alpha}{2}$, получаем равенство

$$l_a = \frac{2bc\cos\frac{\alpha}{2}}{b+c}.$$

Возводя его в квадрат, при помощи формулы

$$\cos^2\frac{\alpha}{2} = \frac{\cos\alpha+1}{2}$$

и теоремы косинусов, которую можно записать в виде

$$\cos\alpha = \frac{b^2+c^2-a^2}{2bc},$$

где $a=BC$, $b=AC$ и $c=AB$, получаем после несложных преобразований формулу, выражающую квадрат длины биссектрисы треугольника через длины его сторон. ■

Обозначив $f(x)=\dfrac{x(1-x)^2}{1-2x}$ и $t=P^2xyz$, получаем

$$\begin{cases} f(x)=t/l_a^2,\\ f(y)=t/l_b^2,\\ f(z)=t/l_c^2,\\ x+y+z=1,\\ 0<x,\,y,\,z<1/2. \end{cases}$$

Взглянув на график функции f, приходим к выводу, что f возрастает на промежутке $[0;\,1/2)$, причем $f(0)=0$ и $\lim\limits_{x\to 1/2} f(x)=\infty$. Доказать возрастание функции f на $(0;\,1/2)$ проще всего при помощи производной: легко проверить, что

$$f'(x)=\frac{(1-x)(4x^2-3x+1)}{(1-2x)^2}>0$$

при $x\in(0;\,1/2)$.

Таким образом, существует определенная на $[0;+\infty)$ непрерывная функция g, обратная к функции f (точнее, к ограничению этой функции на промежуток $[0;\,1/2)$, только и интересующий нас). Система приобретает вид

$$\begin{cases} x=g(t/l_a^2),\\ y=g(t/l_b^2),\\ z=g(t/l_c^2),\\ g(t/l_a^2)+g(t/l_b^2)+g(t/l_c^2)=1. \end{cases}$$ ■

Сумма возрастающих функций $h(t)=g(t/l_a^2)+g(t/l_b^2)+g(t/l_c^2)$ является возрастающей функцией, причем $h(0)=0$ и $\lim\limits_{t\to+\infty} h(t)=3/2$. Поэтому в силу теоремы о промежуточном значении непрерывной функции существует единственное положительное t, для которого $h(t)=1$. Зная t, находим x, y и z, затем $P=\sqrt{\dfrac{t}{xyz}}$ и, наконец, $a=Px$, $b=Py$ и $c=Pz$. ■

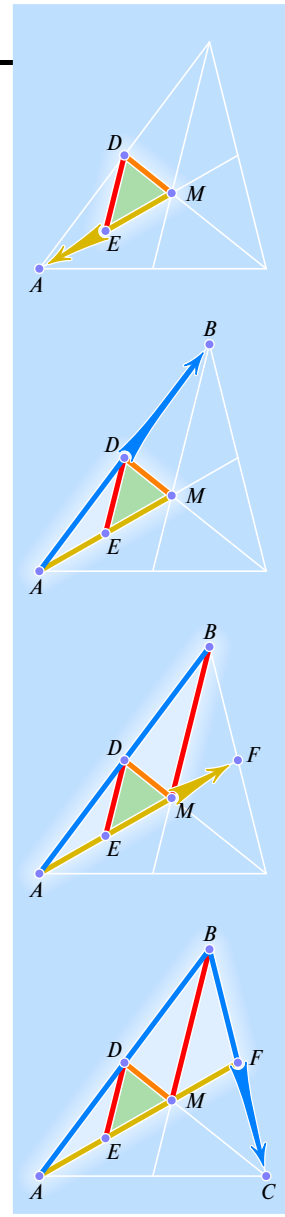
Рассмотрим случай $l_a=l_b=1$ и $l_c=3$. Уравнение $h(t)=1$ принимает вид $2g(t)+g(t/9)=1$. Обозначим $g(t)=u$. Тогда $g(t/9)=1-2u$. Вспоминая определение функции g, имеем $t=f(u)$ и $t/9=f(1-2u)$, то есть

$$\frac{u(1-u)^2}{1-2u}=9\cdot\frac{(1-2u)(2u)^2}{4u-1},$$

откуда получаем уравнение третьей степени

$$(1-u)^2(4u-1)=36u(1-2u)^2.$$

Легко доказать, что корень этого уравнения иррационален. В «Алгебре» Б. Л. Ван дер Вардена (и многих других курсах высшей алгебры) приведено доказательство того, что такой треугольник циркулем и линейкой построить нельзя. ■

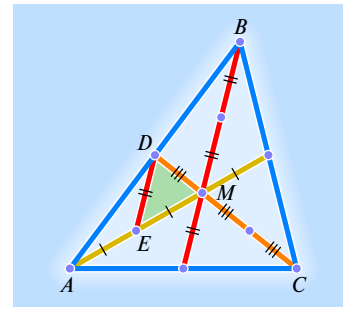


Длины m_a, m_b и m_c медиан треугольника ABC удовлетворяют неравенствам

$$\begin{cases} m_a<m_b+m_c,\\ m_b<m_c+m_a,\\ m_c<m_a+m_b. \end{cases}$$

В самом деле, длины сторон треугольника MDE, где M — центр тяжести $\triangle ABC$, а D и E — середины отрезков AB и AM, — это $m_a/3$, $m_b/3$ и $m_c/3$. Записанные выше неравенства — это, по сути, неравенства треугольника для $\triangle MDE$.

Треугольник MDE помогает доказать, что для любых трех удовлетворяющих этим неравенствам положительных чисел m_a, m_b и m_c существует и единственен треугольник с такими длинами медиан. А именно, построив $\triangle MDE$, мы можем найти вершину A (отразив точку M относительно E), затем — B (отразив A относительно D) и, наконец, C (отразив E относительно M и затем B относительно F). ■



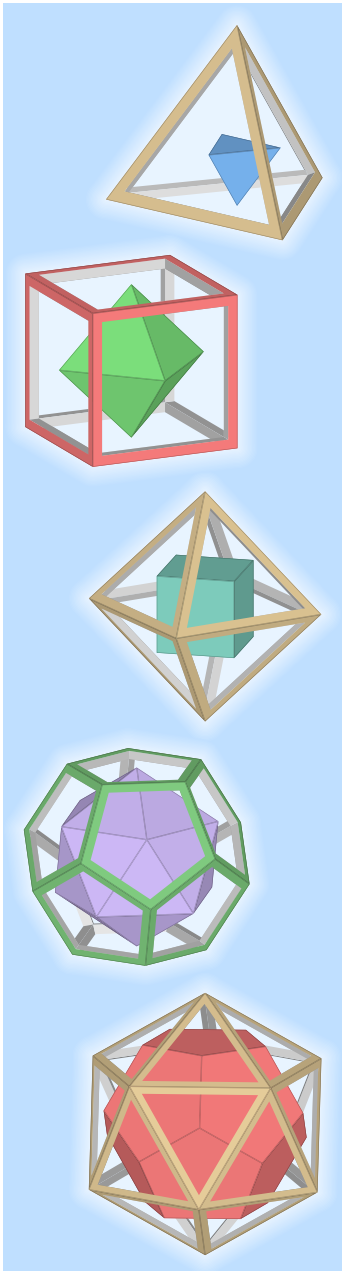
Если h_a, h_b и h_c — длины высот треугольника ABC, то для его площади S имеем

$$S=\frac{ah_a}{2}=\frac{bh_b}{2}=\frac{ch_c}{2},$$

откуда $a=2S/h_a$, $b=2S/h_b$ и $c=2S/h_c$. Неравенства $a<b+c$, $b<c+a$ и $c<a+b$ приобретают вид

$$\begin{cases} \dfrac{1}{h_a}<\dfrac{1}{h_b}+\dfrac{1}{h_c},\\ \dfrac{1}{h_b}<\dfrac{1}{h_c}+\dfrac{1}{h_a},\\ \dfrac{1}{h_c}<\dfrac{1}{h_a}+\dfrac{1}{h_b}. \end{cases}$$

Построив треугольник с длинами сторон $1/h_a$, $1/h_b$ и $1/h_c$, при помощи гомотетии легко проверить, что эта система неравенств не только необходима, но и достаточна для существования треугольника, длины высот которого суть h_a, h_b и h_c. ■



Центры граней правильного тетраэдра являются вершинами меньшего правильного тетраэдра. А если мы отметим центры граней куба, то получим не 8 точек, а всего лишь 6. Очевидно, это вершины октаэдра. Менее очевидно — но верно! — что центры граней додекаэдра являются вершинами икосаэдра, а центры граней икосаэдра, в свою очередь, — вершинами додекаэдра. Говорят, что тетраэдр двойственен сам себе; двойственны также куб и октаэдр, додекаэдр и икосаэдр. ■

Правильные и полуправильные многогранники

Правильный многогранник — это выпуклый многогранник, все грани которого — равные правильные многоугольники, а все многогранные углы равны. Полуправильный выпуклый многогранник — это многогранник, все грани которого — правильные многоугольники (не обязательно равные между собой), а все вершины «одинаково устроены» в том смысле, что можно любую из них при помощи самосовмещения многогранника совместить с любой другой вершиной.

Правильных многогранников всего 5 (с точностью до перемещений и гомотетий): тетраэдр, куб, октаэдр, додекаэдр и икосаэдр (рис. 1—5). Если подойти близко-близко к одной из граней выпуклого многогранника, то она покажется огромной, а все остальные грани расположатся внутри нее (рис. 6). На рисунках 7—11 в таком ракурсе изображены все правильные многогранники. Эти рисунки позволяют легко посчитать, сколько у какого из них вершин (B), ребер (P) и граней (Γ). В следующей таблице теперь все ясно, кроме двух столбцов: n и k. Немного подумав, вы сообразите, что n — количество вершин (или сторон) грани, а k — количество граней (если угодно, ребер), сходящихся в одной вершине.

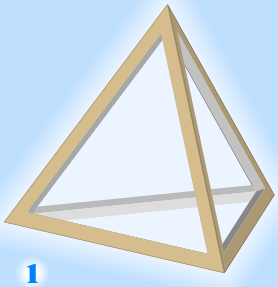
Многогранник	n	B	P	Γ	k
Тетраэдр	3	4	6	4	3
Куб	4	8	12	6	3
Октаэдр	3	6	12	8	4
Додекаэдр	5	20	30	12	3
Икосаэдр	3	12	30	20	5

Числа n, B, P, Γ и k взаимосвязаны: $kB=2P=n\Gamma$. Объяснить это легко: разрезав каждое ребро пополам, мы получаем $2P$ отрезков, объединенных в B «кустов» по k отрезков в каждом (на рисунке 12 это показано на примере тетраэдра). Второе равенство аналогично: разрезав многогранник вдоль ребер на отдельные грани, видим, что количество ребер удвоилось, причем всего имеем Γ отдельных один от другого n-угольников.

Рассмотрим еще несколько примеров многогранников:

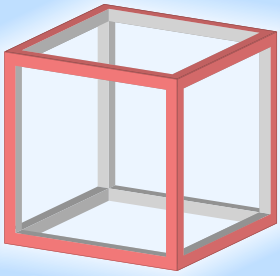
Многогранник	B	P	Γ
n-угольная пирамида (рис. 13)	$n+1$	$2n$	$n+1$
n-угольная призма (рис. 14)	$2n$	$3n$	$n+2$
n-угольная антипризма (рис. 15)	$2n$	$4n$	$2n+2$
Усеченный тетраэдр	12	18	8
Кубооктаэдр	12	24	14
Усеченный октаэдр	24	32	14

Закономерность очевидна: $B-P+\Gamma=2$. Это — знаменитая формула Эйлера. Она верна для любого графа, нарисованного на сфере или на плоскости так,



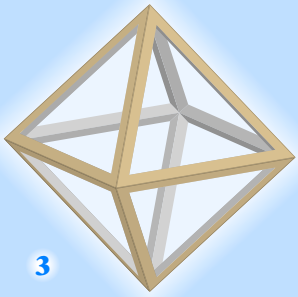
1

Тетраэдр



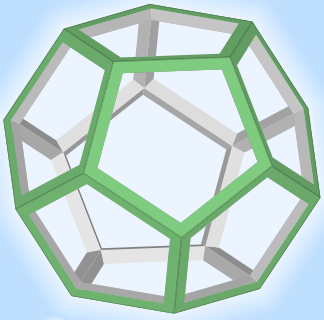
2

Куб



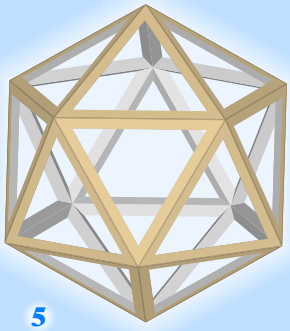
3

Октаэдр



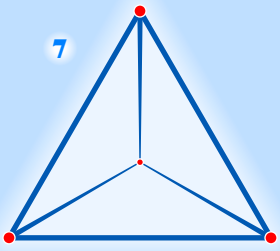
4

Додекаэдр

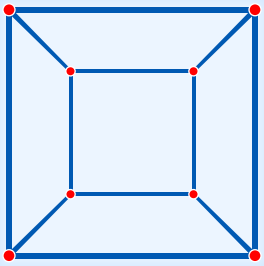


5

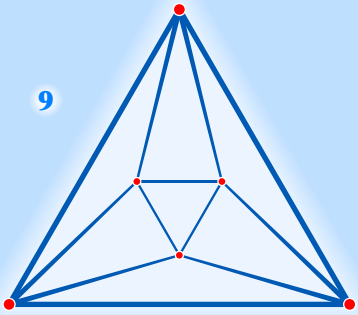
Икосаэдр



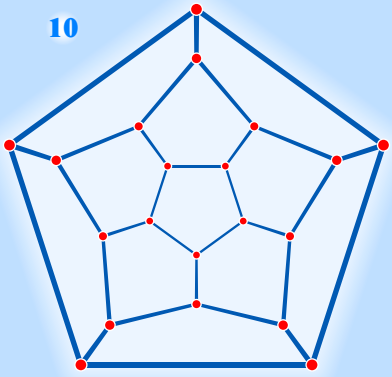
7



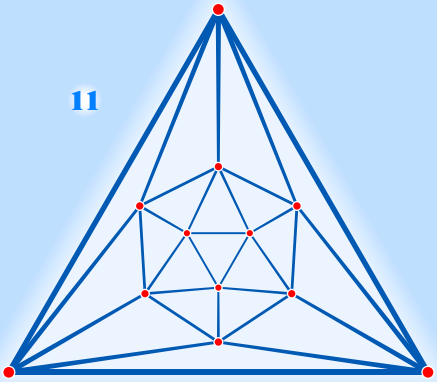
8



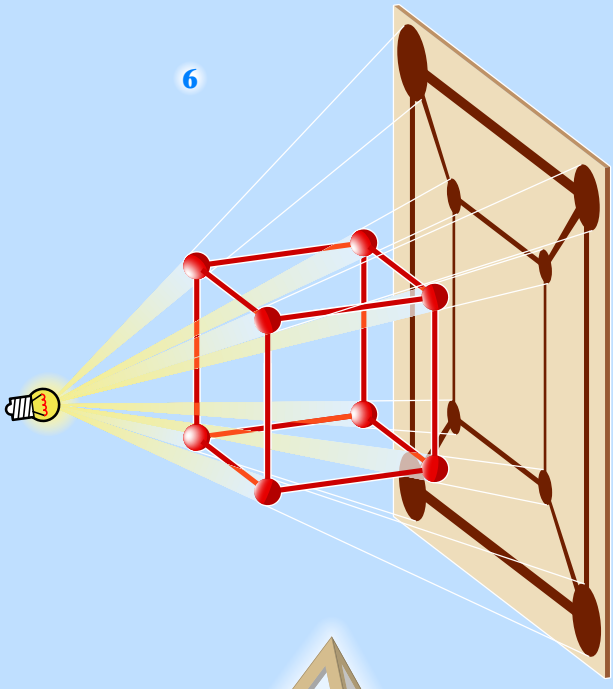
9



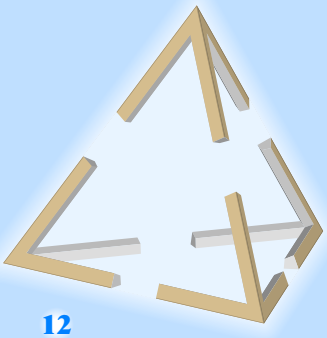
10



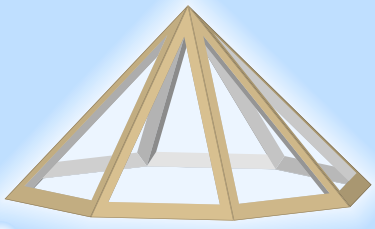
11



6

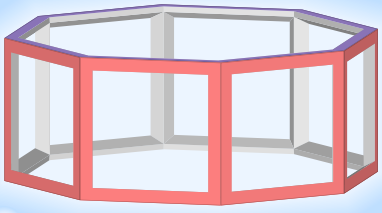


12



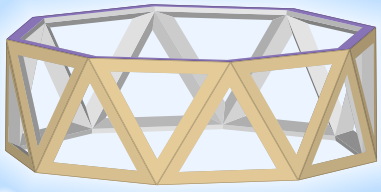
13

Восьмиугольная пирамида



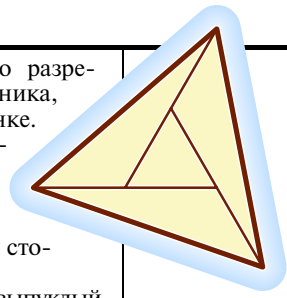
14

Восьмиугольная призма



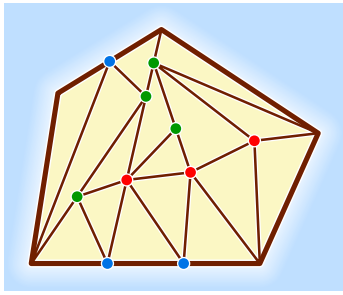
15

Восьмиугольная антипризма



Треугольник можно разрезать на 4 треугольника, как показано на рисунке. Но если $n>3$, то в любом разрезании выпуклого n-угольника можно найти два треугольника с общей стороной!

Доказательство. Пусть выпуклый n-угольник разрезан на t треугольников, ни у каких двух из которых нет общей стороны. Обозначим буквой a количество (синих на рисунке) вершин треугольников, лежащих внутри сторон исходного n-угольника; b — количество (зеленых) вершин, каждая из которых лежит внутри исходного многоугольника и при этом на стороне некоторого из t треугольников, на которые разрезан исходный многоугольник; c — количество



остальных (красных) вершин. (На рисунке $n=5$, $t=16$, $a=3$, $b=4$ и $c=3$.) Подсчитаем сумму величин углов всех треугольников двумя способами:

$$180^\circ t=180^\circ(n-2)+180^\circ a+\\+180^\circ b+360^\circ c,$$

то есть $t=n-2+a+b+2c$.

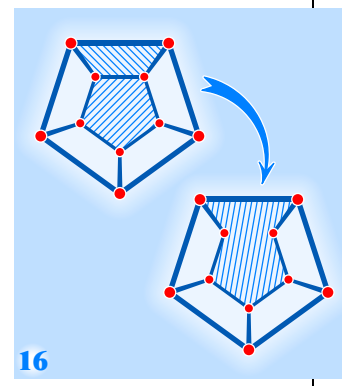
Для каждой синей точки и для каждой вершины n-угольника отметим сторону треугольника разбиения, правым концом которой она является (точнее говоря, при обходе против часовой стрелки этот конец стороны должен встретиться раньше другого). Для каждой зеленой точки отметим три стороны: ту, внутренней точкой которой она является, и две лежащие на той же прямой стороны, являющиеся продолжением одна другой и для которых рассматриваемая точка является вершиной. Очевидно, каждая сторона каждого треугольника будет отмечена — следовательно, $n+a+3b\geqslant 3t$ и, значит,

$$n+a+3b\geqslant 3n-6+3a+3b+6c,$$

откуда $0\geqslant(2n-6)+2a+6c\geqslant 2$. Противоречие! ■

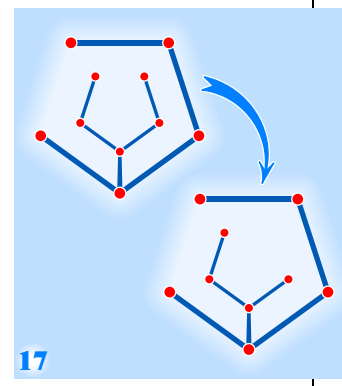
что ребра не пересекаются между собой (при этом, разумеется, в вершине графа может сходиться сколь угодно много ребер; запрещены лишь пересечения во внутренних точках ребра). Доказательство формулы Эйлера проведем в два этапа.

Если в графе есть хотя бы один цикл, рассмотрим любое ребро этого цикла и сотрем его (рис. 16). Количество ребер уменьшится на 1, а две соседние грани сольются в одну; тем самым количество граней тоже уменьшится на 1, а эйлерова характеристика не изменится: $B-(P-1)+(\Gamma-1)=B-P+\Gamma$. Таким образом, стирая ребра и уничтожая тем самым цикл за циклом, мы рано или поздно получим дерево — граф без циклов (рис. 17). Выбрав любой его лист — вершину степени 1, — мы можем стереть его вместе с выходящим из него ребром. Количества вершин и ребер уменьшаются на 1, а эйлерова характеристика не меняется: $(B-1)-(P-1)+\Gamma=B-P+\Gamma$. Стирая вершины, мы в конце концов получим граф, состоящий из одной вершины. Для него $B-P+\Gamma=1-0+1=2$, что и требовалось доказать. ■



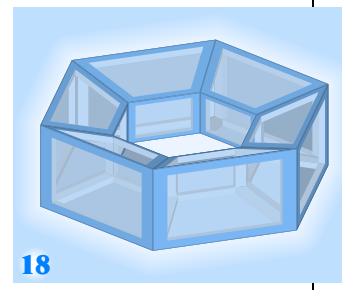
16

Для правильного многогранника, как вы помните, $B=2P/k$ и $\Gamma=2P/n$. Поэтому $\frac{2P}{k}-P+\frac{2P}{n}=2$, то есть $\frac{1}{k}+\frac{1}{n}=\frac{1}{2}+\frac{1}{P}>\frac{1}{2}$. Поскольку в эту формулу n и k входят симметрично, достаточно разобрать случай $k\leqslant n$, когда $\frac{1}{k}\geqslant\frac{1}{n}$ и, следовательно, $\frac{2}{k}\geqslant\frac{1}{k}+\frac{1}{n}>\frac{1}{2}$, откуда $k<4$, то есть на самом деле $k=3$. При $k=3$ получаем $\frac{1}{n}>\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$, откуда $n<6$. Значит, $(k;n)=(3;3)$, $(4;3)$ или $(5;3)$. Вспомнив, что числа k и n можно поменять местами, находим еще две пары: $(3;4)$ и $(3;5)$. Формула $\frac{1}{k}+\frac{1}{n}=\frac{1}{2}+\frac{1}{P}$ позволяет в каждом из этих случаев посчитать величину P. Зная P, n и k, легко найти B и Γ. ■

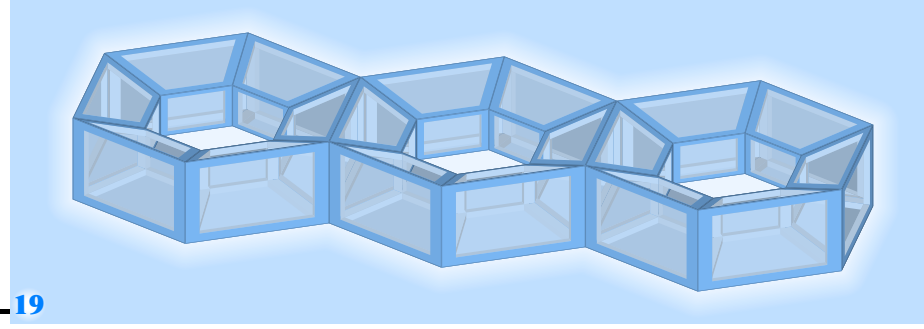


17

На рисунке 18 изображена «гайка» с 24 вершинами, 48 ребрами и 24 гранями. Очевидно, эйлерова характеристика равна 0. А для рисунка 19 имеем $B-P+\Gamma=64-136+68=-4$. Вообще, можно доказать, что в случае g «дыр» эйлерова характеристика равна $2-2g$. ■



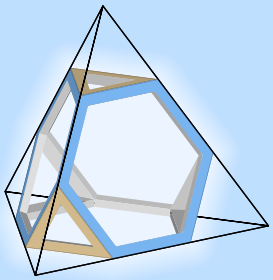
18



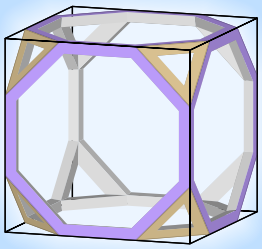
19

ПОЛУПРАВИЛЬНЫЕ МНОГОГРАННИКИ

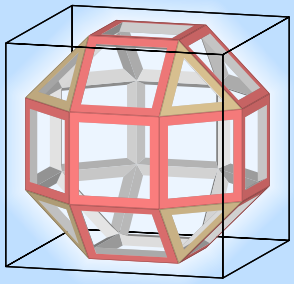
На всех чертежах показаны также правильные многогранники, из которых усечением получаются полуправильные.



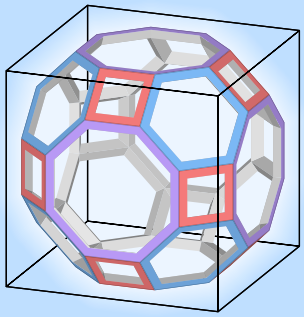
Усеченный тетраэдр



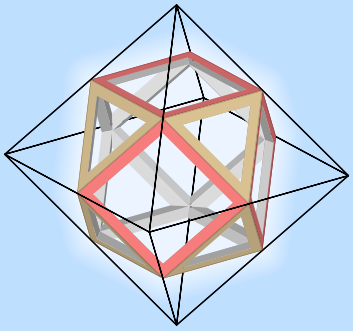
Усеченный куб



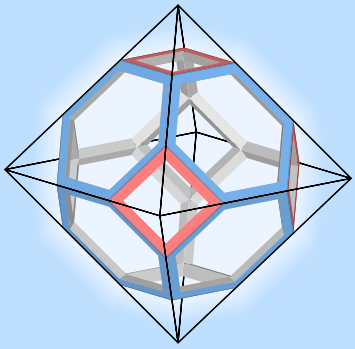
Ромбокубооктаэдр



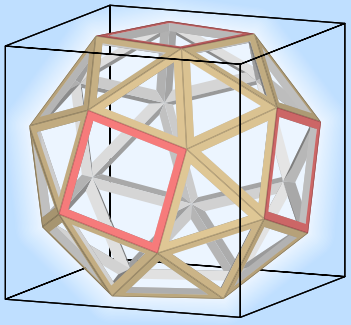
Усеченный кубооктаэдр



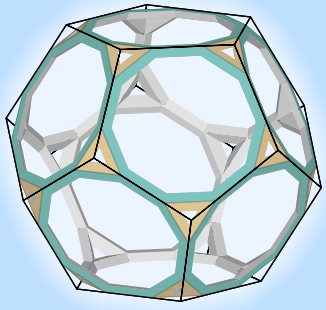
Кубооктаэдр



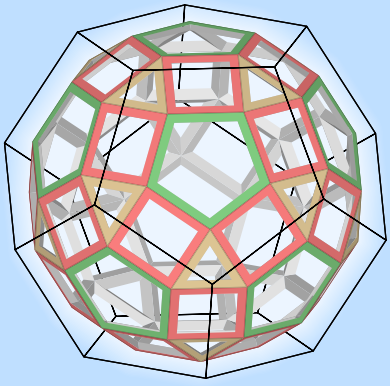
Усеченный октаэдр



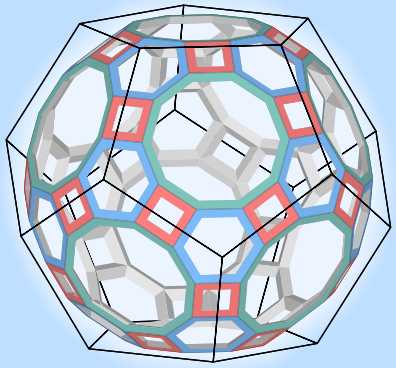
Курносый куб



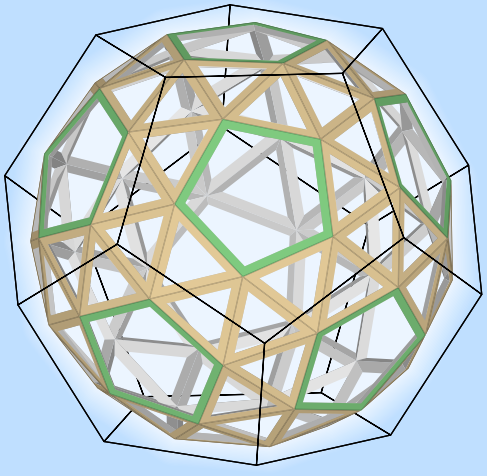
Усеченный додекаэдр



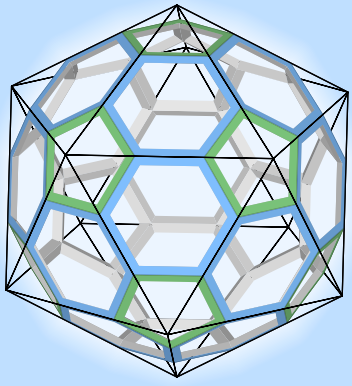
Ромбоикосододекаэдр



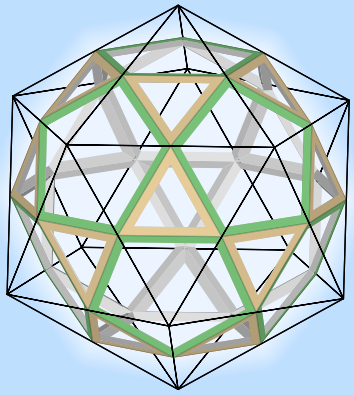
Усеченный икосододекаэдр



Курносый додекаэдр



Усеченный икосаэдр



Икосододекаэдр

АЛГЕБРА

Уравнение $x^2+px+q=0$ замена $x=y-\frac{p}{2}$ приводит к виду $y^2=\frac{p^2}{4}-q$, откуда $y=\pm\frac{\sqrt{p^2-4q}}{2}$, то есть

$$x=y-\frac{p}{2}=\frac{-p\pm\sqrt{p^2-4q}}{2}.$$

Таким образом решение квадратного уравнения сведено к арифметическим операциям и извлечению квадратного корня.

Уравнение третьей степени $x^3+ax^2+bx+c=0$ аналогичная замена $x=y-\frac{a}{3}$ приводит к виду

$$y^3+\left(b-\frac{a^2}{3}\right)y+\frac{2a^3}{27}-\frac{ab}{3}+c=0,$$

так что достаточно решить уравнение вида $y^3+py+q=0$, в котором коэффициент при y^2 равен 0.

Сделаем замену $y=u+v$. Уравнение приобретет вид

$$u^3+v^3+(3uv+p)(u+v)+q=0.$$

Чтобы его упростить, потребуем от чисел u и v выполнения равенства $uv=-p/3$, откуда $u^3v^3=-p^3/27$, а уравнение принимает вид $u^3+v^3=-q$. Таким образом, числа u^3 и v^3 — корни квадратного уравнения

$$z^2+qz-\frac{p^3}{27}=0,$$

откуда $u^3=\frac{-q+\sqrt{q^2+\frac{4p^3}{27}}}{2}$ и $v^3=\frac{-q-\sqrt{q^2+\frac{4p^3}{27}}}{2}$.

Извлекая кубические корни и складывая, получаем ответ — формулу, которую придумал Шипионе дель Ферро (1465—1526) в 1506 г., а опубликовал в книге «Великое искусство» Дж. Кардано (1501—1576), которому эту формулу рассказал переоткрывший ее в 1535 г. Н. Тарталья (ок. 1499—1557).

Может возникнуть вопрос: почему получился один ответ, хотя уравнение третьей степени в общем случае имеет три корня? Дело в том, что $1^3=1$ и $\left(\frac{1}{2}\pm\frac{i\sqrt{3}}{2}\right)^3=1$ — корень третьей степени из 1 (а поэтому и корень третьей степени из любого ненулевого числа) можно извлечь тремя способами. Но теперь возникает другой вопрос: почему решений не 9 — ведь можно тремя способами извлечь корень и из числа u, и тремя из v? Ответ очевиден: $uv=-p/3$, так что выбор числа u однозначно определяет величину v.

Ученик Кардано Л. Феррари (1522—1565) решил уравнение четвертой степени $x^4+ax^3+bx^2+cx+d=0$. Начало рассуждения такое же, как и при решении уравнений второй и третьей степеней: замена $y=x-\frac{a}{4}$ аннулирует коэффициент при третьей степени неизвестной, так что достаточно решить уравнение

$$y^3+py^2+qy+r=0.$$

Перепишем его в виде

$$(y^2+z)^2=(2z-p)y^2-qy+(z^2-r),$$

где z — некоторый параметр, значение которого мы вскоре определим. Если бы правая часть была полным квадратом, то мы извлекли бы квадратный корень (не забыв поставить $\pm$) и, решив полученные квадратные уравнения, решили бы тем самым исходное уравнение четвертой степени. Как известно, корни квадратного трехчлена кратные тогда и только тогда, когда его дискриминант равен нулю:

$$D=q^2-4(2z-p)(z^2-r).$$

Таким образом, уравнение $D=0$ — это уравнение третьей степени относительно z. Его называют резольвентой уравнения четвертой степени. Мы умеем решать уравнение третьей степени — значит, можем найти такое z (годится любое из трех возможных значений), что дискриминант окажется равен 0 и тем самым мы сможем решить и уравнение четвертой степени.

Можно ли решить уравнение пятой степени? Нет, в 1824 г. норвежский математик Н. Х. Абель (1802—1829) доказал следующую теорему.

Общее алгебраическое уравнение с одной неизвестной степени выше четвертой неразрешимо в радикалах, то есть не существует формулы, выражающей корни уравнения степени выше четвертой через коэффициенты с помощью операций сложения, вычитания, умножения, деления, возведения в степень и извлечения корней.

Доступное для школьника доказательство изложено в «Алгебре» Б. Л. Ван дер Вардена и в «Теореме Абеля в задачах и решениях» В. Б. Алексеева. Первая из этих книг — классическая монография, а вторая написана на основе лекций, прочитанных ее автором и В. И. Арнольдом в школе-интернате № 18 им. А. Н. Колмогорова.

Доказательство основано на изучении группы перестановок корней уравнения и выяснении того, какие составленные из этих корней выражения и как меняются под действием этих перестановок. Поясним, о чем речь. Выражение $x+2y$ при замене x на y, а y на x превращается в $2x+y$. А выражения $\sigma_1=x+y$, $\sigma_2=xy$, x^2+y^2 и $x^{35}+x^9y^4+x^4y^9+y^{35}$ переходят сами в себя.

Заметьте: $x^2+y^2=\sigma^2-2\sigma_2$. Оказывается, и любая симметрическая функция от x и y выражается в виде многочлена от xy и $x+y$. Далее, любая симметрическая функция трех переменных x, y и z выражается в виде многочлена от $\sigma_1=x+y+z$, $\sigma_2=xy+yz+zx$ и $\sigma_3=xyz$. И вообще, рассмотрев для любых n переменных $x_1, x_2, \ldots, x_n$ многочлены $\sigma_1, \sigma_2, \ldots, \sigma_n$, определяемые из формулы

$$(y-x_1)(y-x_2)\ldots(y-x_n)=y^n-\sigma_1y^{n-1}+\sigma_2y^{n-2}-\ldots$$
$$\ldots+(-1)^{n-1}\sigma_{n-1}y+(-1)^n\sigma_n,$$

мы утверждаем: любая симметрическая функция от $x_1, x_2, \ldots, x_n$ является многочленом от $\sigma_1, \sigma_2, \ldots, \sigma_n$. Это утверждение называют основной теоремой о симметрических функциях. Докажем ее в случае $n=2$ индукцией по степени многочлена. Пусть $f(x; y)$ — симметрический многочлен от двух переменных, то есть $f(x; y)=f(y; x)$. Рассмотрим многочлен $g(x; y)=f(x; y)-f(0; x+y)$. Очевидно, g — тоже симметрическая функция, причем $g(0; y)=0$. Поэтому многочлен $g(x; y)$ делится на y — а в силу своей симметричности и на xy, то есть

$$g(x; y)=xyh(x; y),$$

где h — некоторый симметрический многочлен. Степень многочлена g меньше степени многочлена f, поэтому в силу предположения индукции h выражается в виде многочлена от $x+y$ и xy, то есть $h(x; y)=k(x+y; xy)$, где k — некоторый многочлен. Осталось вспомнить, что

$$f(x; y)=g(x; y)+f(0; x+y)=\sigma_2k(\sigma_1, \sigma_2)+f(0; \sigma_1),$$

и основная теорема о симметрических многочленах для $n=2$ доказана.

Не желая далее углубляться в эту тему, заметим главное: очень часто помогает рассмотрение группы симметрий, то есть множества преобразований, которые переводят некоторый объект в себя. Например, чем параллелограмм лучше произвольного четырехугольника? Центральной симметричностью. Ромб отличается от произвольного параллелограмма тем, что у него есть ось две симметрии, а квадрат от произвольного ромба — тем, что осей симметрии не две, а четыре.

Группа симметрий правильного треугольника изоморфна группе перестановок S_3 всех трех его вершин. Группа симметрий квадрата состоит из 8 элементов — 4 осевых симметрий и 4 поворотов (на 0°, 90°, 180° и 270°). Еще интереснее группа симметрий куба или додекаэдра. И один из способов доказать теорему Абеля о неразрешимости уравнения пятой степени в радикалах состоит именно в исследовании свойств группы симметрий додекаэдра! Не зря в шутку говорят, что математика — искусство называть совершенно разные вещи одним словом!

Алгебраические идеи применяют в самых разных разделах математики. В статье «Группа кос» рассказано, как можно умножать довольно причудливые геометрические объекты — косы. А в следующей за ней статье «Теория узлов» рассказано о том, что произвольному узлу можно сопоставить некоторый многочлен. Если один узел можно перевести в другой, не разрывая веревку, а только деформируя ее в пространстве, то этим узлам оказывается сопоставлен один и тот же многочлен; это позволяет, если многочлены некоторых двух узлов разные, утверждать, что узлы не могут быть деформированы один в другой. Так алгебру применяют для решения топологических задач.

В статье «Многочлены деления круга» рассказано о, казалось бы, отдельной задаче — разложении многочлена x^n-1 на множители. Но она оказывается тесно связана со свойствами правильного n-угольника и даже с построениями циркулем и линейкой!

И это не исключение, а общее для современной математики явление: многообразие задач, в которых применяются алгебраические конструкции, поистине огромно. Не зря говорят об алгебраической топологии, алгебраической геометрии, алгебраической теории чисел и многих других алгебраических науках.

Линейная алгебра изучает n-мерные векторные пространства и их отображения друг в друга, алгебра логики — логические операции над высказываниями, теория групп и алгебр Ли чрезвычайно важна для геометрии и физики.

Таким образом, возникнув как наука о решении уравнений и как правило переноса слагаемых из одной части равенства в другую, алгебра чрезвычайно расширила область своей применимости.

И если в 1799 г. Гаусс, доказав теорему о существовании для каждого отличного от константы многочлена f такого комплексного числа z, что $f(z)=0$, с полным правом утверждал, что он доказал основную теорему алгебры, то сейчас у каждого из десятков разделов, на которые подразделяется современная алгебра, свои основные теоремы — ничуть не менее важные и ничуть не легче доказываемые, чем основная теорема алгебры (название «приклеилось» навечно, как и название «основная теорема арифметики»).

Комплексные числа

Первопричиной появления комплексных чисел послужило то обстоятельство, что некоторые квадратные уравнения с вещественными коэффициентами имеют вещественные решения, а некоторые (дискриминанты которых отрицательны) не имеют. Математику трудно смириться с тем, что какая-то задача не имеет решения. Поэтому в таких случаях стараются так расширить основные понятия, чтобы эту невозможность устранить. Так приходят к расширению поля $\mathbb{R}$ вещественных чисел (числовой прямой) до поля $\mathbb{C}$ комплексных чисел («числовой плоскости»).

Одной из привлекательных черт теории комплексных чисел является ее подлинная комплексность: в ней сочетаются алгебраические, аналитические, геометрические и топологические методы. Наряду с конкретными прикладными задачами решают и весьма общие абстрактные задачи. Понятия и методы комплексного анализа используют во всех разделах математики.

Что такое комплексное число? Новые числа в математике вводят, когда старых оказывается недостаточно. Изобретение целых чисел, то есть расширение множества $\mathbb{N}=\{1, 2, 3, \ldots\}$ натуральных чисел до множества $\mathbb{Z}=\{\ldots, -2, -1, 0, 1, 2, \ldots\}$, дает возможность решить, например, уравнение $x+7=5$. Построив еще более широкое множество $\mathbb{Q}=\left\{\frac{m}{n} \mid m\in\mathbb{Z},\ n\in\mathbb{N}\right\}$ рациональных чисел, получаем возможность решать уравнения вроде $3x=8$. Желание измерить диагональ единичного квадрата (или, что то же, решить уравнение $x^2=2$) приводит к очередному расширению множества чисел до множества $\mathbb{Q}[\sqrt{2}]$ чисел вида $a+b\sqrt{2}$, где $a, b\in\mathbb{Q}$. Очевидно, сумма, разность и произведение чисел вида $a+b\sqrt{2}$ — число такого же вида. С делением тоже все в порядке, например:

$$\frac{1+\sqrt{2}}{3-2\sqrt{2}}=\frac{(1+\sqrt{2})(3+2\sqrt{2})}{(3-2\sqrt{2})(3+2\sqrt{2})}=7+5\sqrt{2},\quad \frac{2-5\sqrt{2}}{3+\sqrt{2}}=\frac{(2-5\sqrt{2})(3-\sqrt{2})}{(3+\sqrt{2})(3-\sqrt{2})}=\frac{16}{7}-\frac{17}{7}\sqrt{2}.$$

Видите, как просто? В общем виде это выглядит так:

$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{(a+b\sqrt{2})(c-d\sqrt{2})}{(c+d\sqrt{2})(c-d\sqrt{2})}=\frac{ac-2bd+(bc-ad)\sqrt{2}}{c^2-2d^2}.$$

В этих вычислениях использовано то, что квадрат числа $\sqrt{2}$ равен 2. Комплексные числа получим, введя в рассмотрение число i, квадрат которого равен -1. Может показаться, что «такого не бывает», ведь уравнение $x^2+1=0$ не имеет решений не только в рациональных, но и в вещественных числах. Однако число $\sqrt{2}$, заметьте, тоже «не существовало» до тех пор, пока мы рассматривали только рациональные числа. ■

Рассмотрим выражения вида $a+bi$, где a, b — вещественные числа. Эти выражения мы и будем называть комплексными числами. Сумму и произведение определим естественными формулами

$$(a+bi)+(c+di)=(a+c)+(b+d)i,$$
$$(a+bi)\cdot(c+di)=(ac-bd)+(ad+bc)i.$$

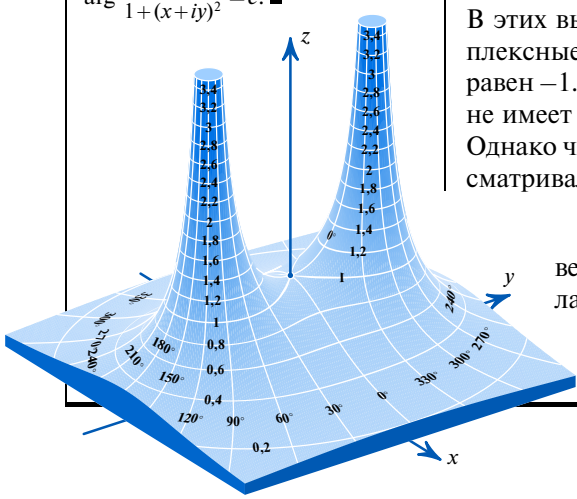
Впервые геометрически проиллюстрировал расширение множества $\mathbb{R}$ вещественных чисел до множества $\mathbb{C}$ комплексных чисел в 1799 г. датчанин К. Вессель (1745—1818), но его сочинение «Об аналитическом представлении направлений» долгое время оставалось неизвестным. В 1806 г. геометрическую интерпретацию независимо от Весселя открыл швейцарец Ж. Р. Арган (1768—1822). Впрочем, К. Ф. Гаусс, скорее всего, пользовался этими наглядными представлениями раньше Весселя и Аргана. А еще раньше комплексные числа возникали при исследовании уравнений третьей и более высоких степеней. ■

Обстоятельный анализ свойств функций немыслим без выхода в комплексную область. Например, функция $f(x)=\frac{1}{1+x^2}$ одинаково прекрасна (бесконечно дифференцируема) во всех точках числовой оси, а ее ряд Тейлора

$$\frac{1}{1+x^2}=1-x^2+x^4-x^6+x^8-\ldots$$

не сходится при $x\geqslant 1$. Причина этого неясна, пока мы остаемся в действительной области: ведь точки $x=\pm 1$, разделяющие множества сходимости и расходимости, ничем не примечательны. А выход в комплексную область сразу разъясняет явление: на окружности $|z|=1$ лежат точки $z=\pm i$, при приближении к которым функция стремится к бесконечности.

На рисунке изображен график (рельеф) функции $z=\frac{1}{|1+(x+iy)^2|}$. На нем проведены линии уровня $\left|\frac{1}{1+(x+iy)^2}\right|=c$ и перпендикулярные им линии $\arg\frac{1}{1+(x+iy)^2}=c$. ■



Последняя формула, быть может, нуждается в комментарии: $(a+bi)\cdot(c+di)=ac+adi+bci+bdi^2=ac+adi+bci-bd$. Это именно комментарий, а не доказательство, поскольку пользоваться обычными правилами раскрытия скобок можно только после того, как даны определения сложения и умножения комплексных чисел и проверены эти «обычные правила», то есть формулы

$z_1+z_2=z_2+z_1$ (переместительный закон, или коммутативность сложения),

$z_1z_2=z_2z_1$ (коммутативность умножения),

$(z_1+z_2)+z_3=z_1+(z_2+z_3)$ (сочетательный закон, или ассоциативность сложения),

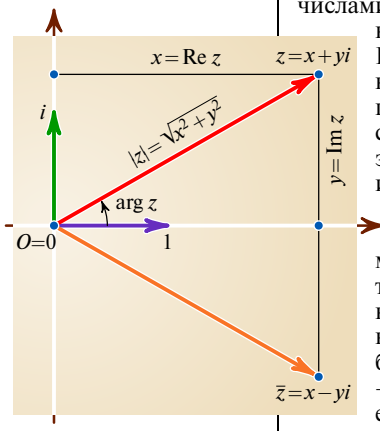
$(z_1z_2)z_3=z_1(z_2z_3)$ (ассоциативность умножения),

$(z_1+z_2)z_3=z_1z_3+z_2z_3$ (распределительный закон, или дистрибутивность). ■

Геометрическая интерпретация. Формулы сложения и умножения комплексных чисел позволяют отождествить комплексное число $a+0i$ с вещественным числом a. Поэтому в дальнейшем мы будем писать не $a+0i$, а попросту a.

Отождествим ось абсцисс координатной плоскости с вещественной осью (то есть множеством всех вещественных чисел); единичный вектор (1; 0) оси абсцисс обозначим просто 1, а единичный вектор (0; 1) оси ординат обозначим через i. Произвольный вектор $z=(x;y)$ плоскости можно теперь записать в виде $z=x(1;0)+y(0;1)=x+yi$. Принято вещественные числа x и y называть вещественной и мнимой частями комплексного числа z. Обозначения: $x=\operatorname{Re} z$, $y=\operatorname{Im} z$. Сложение комплексных чисел — это обычное сложение векторов. А умножение определяется, как мы уже видели, более «хитрой» формулой. ■

Модулем (абсолютной величиной) числа $z=x+yi$ называют расстояние $|z|=\sqrt{x^2+y^2}$ от начала координат до точки $(x;y)$.



Теорема. ***Модуль произведения комплексных чисел равен произведению их модулей*:**

$$|(a+bi)(x+yi)|=|a+bi|\cdot|x+yi|.$$

Доказательство.

$$|(a+bi)(x+yi)|=|(ax-by)+(ay+bx)i|=$$
$$=\sqrt{(ax-by)^2+(ay+bx)^2}=\sqrt{(a^2+b^2)(x^2+y^2)}=$$
$$=|a+bi|\cdot|x+iy|.\ \blacksquare$$

Сопряженные числа. Уравнение $z^2=-1$ имеет два корня: i и $-i$. Поскольку при вычислениях используется именно равенство $i^2=-1$, возникает идея заменить i на $-i$. Верное равенство при одновременной замене всех входящих в него символов i на $-i$ останется верным!

Точная реализация этой идеи такова: два комплексных числа, действительные части которых равны, а мнимые части равны по абсолютной величине и противоположны по знаку, называют сопряженными. Число, сопряженное с $z=x+yi$, обозначают $\overline{z}=x-yi$. Геометрический смысл перехода от числа к сопряженному — симметрия относительно оси абсцисс. Легко проверить тождества

$$\overline{u+v}=\overline{u}+\overline{v} \quad \text{и} \quad \overline{u\cdot v}=\overline{u}\cdot\overline{v},$$

которые и позволяют заменять в формулах все числа на сопряженные.

Между прочим, $|z|^2=x^2+y^2=(x+iy)(x-iy)=z\overline{z}$. Это позволяет изящно доказать теорему:

$$|uv|^2=(uv)(\overline{uv})=uv\,\overline{u}\,\overline{v}=(u\overline{u})(v\overline{v})=|u|^2\cdot|v|^2.\ \blacksquare$$

Переход к комплексным числам является очередным шагом в последовательности: натуральные числа — целые числа — рациональные числа — действительные числа — комплексные числа. Может сложиться впечатление, что до действительных чисел это на самом деле числа, а комплексные числа — это уже не числа, а объекты более сложной природы. Конечно, терминология может быть принята любая, однако в действительности комплексные числа вполне заслуживают, чтобы их называли числами.

Первое возражение против этого может состоять в том, что это не числа, а пары чисел. Вспомним, однако, что подобным же образом вводятся рациональные числа. Рациональное число — это класс эквивалентных дробей, где дроби — это пары целых чисел, записываемые в виде $\frac{m}{n}$ (где $n\neq0$); дроби $\frac{m_1}{n_1}$ и $\frac{m_2}{n_2}$ эквивалентны, если $m_1n_2=m_2n_1$. Действия над рациональными числами — это просто действия над парами целых чисел. Поэтому первое возражение несостоятельно. Другое возражение может состоять в том, что числа — это то, чем можно что-то измерять. Если понимать под этим, что числа — это то, чем можно измерять все, что угодно, то тогда надо запретить, например, отрицательные числа, так как не бывает отрезков длиной -3 см, а поезд не может ехать -4 дня. Придется запретить и слишком большие: температура манной каши не бывает больше 1000 °C. Если же считать, что числа — это то, чем можно (или удобно) измерять хоть что-нибудь, то тогда комплексные числа оказываются ничем не хуже других чисел — ими очень удобно описывать, например, ток, напряжение и сопротивление в электрических цепях переменного тока, и это широко используют в электротехнике.

Таким образом, переход от действительных чисел к комплексным является таким же естественным, как, например, переход от целых чисел к рациональным. ■

Многочлены деления круга

Разложения на множители многочленов вида x^n-1 тесно связаны с задачей о делении окружности на n равных частей. Их изучение позволило К. Ф. Гауссу в 1796 г. доказать, что правильный n-угольник может быть построен циркулем и линейкой, если $\varphi(n)$ — степень двойки. (В частности, можно построить правильный $(2^m\cdot 3\cdot 5\cdot 17\cdot 257\cdot 65\,537)$-угольник.) Не обойтись без многочленов деления круга и в теории Галуа, позволяющей по алгебраическому уравнению сказать, разрешимо оно в радикалах или нет. Важнейшие объекты алгебры и арифметики — корни из единицы, функция Эйлера $\varphi(n)$ и функция $\tau(n)$ (количество натуральных делителей числа n) — встречаются на первых же шагах изучения многочленов деления круга.

Известны формулы сокращенного умножения

$$x^2-1=(x-1)(x+1),$$
$$x^3-1=(x-1)(x^2+x+1),$$
$$x^4-1=(x^2-1)(x^2+1)=(x-1)(x+1)(x^2+1),$$
$$x^5-1=(x-1)(x^4+x^3+x^2+x+1).$$

Раскрыв скобки, легко проверить общую формулу

$$x^n-1=(x-1)(x^{n-1}+x^{n-2}+\ldots+x+1),$$

известную как формула суммы геометрической прогрессии. ■

Можно ли разложить многочлен

$$f_n(x)=x^{n-1}+x^{n-2}+\ldots+x+1$$

на множители с целыми коэффициентами? При некоторых n — можно! Прежде чем формулировать ответ, рассмотрим «экспериментальный материал».

$x^2-1=(x-1)(x+1)$. Обозначим $\Phi_1(x)=x-1$ и $\Phi_2(x)=x+1$.

$x^3-1=(x-1)(x^2+x+1)$. Обозначим $\Phi_3(x)=x^2+x+1$. Многочлен Φ_3 нельзя разложить на множители с целыми коэффициентами. В самом деле, если $x^2+x+1=(ax+b)(cx+d)$, то $ac=1$ и $bd=1$, откуда следует, что числа a, b, c и d могут быть равны только ±1. Без ограничения общности можно считать $a=1$ и $c=1$. Дальнейшее очевидно. (Впрочем, можно рассуждать и проще: квадратный трехчлен x^2+x+1 не имеет корней и поэтому не может быть разложен в произведение линейных множителей.)

$x^4-1=(x-1)(x+1)(x^2+1)$. Обозначим $\Phi_4(x)=x^2+1$.

$x^5-1=(x-1)(x^4+x^3+x^2+x+1)$. Обозначим $\Phi_5(x)=x^4+x^3+x^2+x+1$. Неразложимость многочлена Φ_5 на множители с целыми коэффициентами не вполне очевидна. Можно рассуждать так. Делителей первой степени нет, поскольку в противном случае многочлен Φ_5 имел бы рациональный корень, который заодно был бы корнем многочлена x^5-1, то есть равнялся бы числу 1. Значит, надо привести к противоречию разложение

$$x^4+x^3+x^2+x+1=(ax^2+bx+c)(dx^2+ex+f).$$

Фердинанд Готхольд Макс Эйзенштейн (1823—1852) — ученик К. Ф. Гаусса. Вместе с К. Г. Я. Якоби сформулировал и доказал кубический закон взаимности, рассмотрев для этого поле разложения многочлена x^3-1. В 1840-х гг. определил целое алгебраическое число как корень многочлена с целыми коэффициентами, старший коэффициент которого равен 1. Например, если целое число d не делится на квадрат никакого простого числа, причем $d\neq0$ и $d\neq1$, то целые числа поля $\mathbb{Q}(\sqrt{d})$ при $d\not\equiv3 \pmod 4$ — это числа вида $a+b\sqrt{d}$, где a, b — целые; а при $d\equiv3 \pmod 4$ — это числа вида $a+b(1+\sqrt{d})/2$, где a, b — целые. Сумма и произведение целых алгебраических чисел — тоже целые алгебраические. Например, чтобы найти многочлен с целыми коэффициентами, одним из корней которого является $\sqrt{2}+\sqrt[3]{3}$, достаточно рассмотреть произведение шести скобок

$$(x-\sqrt{2}-\sqrt[3]{3})(x+\sqrt{2}-\sqrt[3]{3})\times$$
$$\times(x-\sqrt{2}-\zeta\sqrt[3]{3})(x+\sqrt{2}-\zeta\sqrt[3]{3})\times$$
$$\times(x-\sqrt{2}-\zeta^2\sqrt[3]{3})(x+\sqrt{2}-\zeta^2\sqrt[3]{3}),$$

где $\zeta=(-1+i\sqrt{3})/2$.

Признак Эйзенштейна. Если все коэффициенты многочлена $f(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0$, кроме старшего коэффициента a_n, делятся на простое число p, а свободный член a_0 не делится на p^2 (но делится, как уже было сказано, на p), то многочлен f не разложим на множители с целыми коэффициентами.

Доказательство. Если

$$f(x)=$$
$$=(b_kx^k+b_{k-1}x^{k-1}+\ldots+b_1x+b_0)\times$$
$$\times(c_mx^m+c_{m-1}x^{m-1}+\ldots+c_1x+c_0),$$

то $b_0c_0=a_0$. Значит, одно из чисел b_0, c_0 делится на p, а другое не делится. Пусть для определенности b_0 делится на p, а c_0 — не делится. Рассматривая по очереди выражения для коэффициентов a_1, a_2, …, a_k, убеждаемся, что все числа b_1, b_2, …, b_k делятся на p. А в таком случае число $a_n=b_kc_m$ делится на p, что противоречит условию. ■

В 1845 г. Эйзенштейн дал следующее доказательство квадратичного закона взаимности.
Лемма. ***Если n — нечетное натуральное число, то***

$$\sin nx=(-4)^{(n-1)/2}\sin x\times$$
$$\times\prod_{1\leqslant k\leqslant(n-1)/2}\left(\sin^2 x-\sin^2\frac{2\pi k}{n}\right).$$

Идея доказательства. $\sin nx$ — многочлен степени n от $\sin x$, корни которого — 0 и $\pm\sin 2\pi k/n$, где $1\leqslant k\leqslant(n-1)/2$. Множитель $(-4)^{(n-1)/2}$ получаем по индукции или применяя формулу Эйлера $\sin x=(e^{ix}-e^{-ix})/2$ и сравнивая коэффициенты при e^{inx} в левой и правой частях.
Доказательство квадратичного закона взаимности. Пусть p и q — нечетные простые числа, $p\neq q$. В силу критерия Гаусса

$$\left(\frac{p}{q}\right)=\prod_{1\leqslant k\leqslant(q-1)/2}\frac{\sin(2\pi kp/q)}{\sin(2\pi k/q)}.$$

В силу леммы

$$\left(\frac{p}{q}\right)=\prod_{1\leqslant k\leqslant(q-1)/2}(-4)^{(p-1)/2}\times$$
$$\times\prod_{1\leqslant m\leqslant(p-1)/2}\left(\sin^2\frac{2\pi k}{q}-\sin^2\frac{2\pi m}{q}\right)=$$
$$=(-4)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\times$$
$$\times\prod_{\substack{1\leqslant k\leqslant(q-1)/2,\\ 1\leqslant m\leqslant(p-1)/2}}\left(\sin^2\frac{2\pi k}{q}-\sin^2\frac{2\pi m}{q}\right).$$

Меняя роли p и q, имеем

$$\left(\frac{q}{p}\right)=(-4)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\times$$
$$\times\prod_{\substack{1\leqslant k\leqslant(q-1)/2,\\ 1\leqslant m\leqslant(p-1)/2}}\left(\sin^2\frac{2\pi m}{q}-\sin^2\frac{2\pi k}{q}\right).$$

Множители в формулах для $\left(\frac{p}{q}\right)$ и $\left(\frac{q}{p}\right)$ отличаются лишь знаками. Количество же скобок равно $\frac{p-1}{2}\cdot\frac{q-1}{2}$, поэтому

$$\left(\frac{p}{q}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\left(\frac{q}{p}\right).$$ ■

Разумеется, $ad=1$ и $cf=1$. Следовательно, коэффициенты a, c, d и f могут равняться лишь ±1. Дальнейшее очевидно, хотя и требует перебора вариантов.
$\boldsymbol{x^6-1=(x^3-1)(x^3+1)=(x-1)(x^2+x+1)(x+1)(x^2-x+1).}$ Как и раньше, возник один новый неразложимый делитель — многочлен $\Phi_6(x)=x^2-x+1$.
$\boldsymbol{x^7-1=(x-1)(x^6+x^5+x^4+x^3+x^2+x+1).}$ Второй множитель, как обычно, обозначим Φ_7. Неразложимость многочлена Φ_7, как и любого многочлена f_p, где p — простое число, следует из признака Ф. Г. Эйзенштейна:

$$f_p(y+1)=\frac{(y+1)^p-1}{(y+1)-1}=y^{p-1}+C_p^1y^{p-2}+\ldots+C_p^{p-2}y+C_p^{p-1},$$

где мы воспользовались биномом Ньютона. (Советуем проделать эти выкладки самостоятельно при $p=3$, 5 или 7.) Как известно, $C_p^k=\frac{p!}{k!(p-k)!}$. При $0<k<p$ числитель делится на p, а знаменатель не делится. Поэтому все коэффициенты многочлена $f_p(y+1)$, кроме старшего, делятся на p. Осталось заметить, что свободный член $C_p^1=p$ не делится на p^2. По признаку Эйзенштейна многочлен $f_p(y+1)$, а вместе с ним и $f_p(x)$, неприводим.
$\boldsymbol{x^8-1=(x^4-1)(x^4+1)=(x-1)(x+1)(x^2+1)(x^4+1)}$ **и** $\boldsymbol{x^9-1=(x^3-1)(x^6+x^3+1)=}$ $\boldsymbol{=(x-1)(x^2+x+1)(x^6+x^3+1).}$ Неразложимость многочленов $\Phi_8(x)=x^4+1$ и $\Phi_9(x)=x^6+x^3+1$ следует из признака Эйзенштейна: $(x+1)^4+1=x^4+4x^3+$ $+6x^2+4x+2$ и $(x+1)^6+(x+1)^3+1=x^6+6x^5+15x^4+21x^3+18x^2+9x+3$.
Продолжая в том же духе, легко построить таблицу, в которой под каждым из значений n выписано, на сколько неразложимых множителей можно разложить многочлен x^n-1. Множителей в точности столько, сколько делителей у числа n. Вскоре мы это докажем.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\tau(n)$	1	2	2	3	2	4	2	4	3	4	2	6	2	4

Многочлен $x^{15}-1$ можно разложить и как разность кубов, и как разность пятых степеней:

$$(x^5)^3-1^3=(x^5-1)(x^{10}+x^5+1),\quad (x^3)^5-1^5=(x^3-1)(x^{12}+x^9+x^6+x^3+1).$$

Как «объединить» эти два разложения в одно? Оказывается, $x^{10}+x^5+1$ делится на x^2+x+1. Поделив «в столбик», получаем

$$x^{15}-1=(x-1)(x^2+x+1)(x^4+x^3+x^2+x+1)(x^8-x^7+x^5-x^4+x^3-x+1).$$

Неразложимость многочлена $\Phi_{15}=x^8-x^7+x^5-x^4+x^3-x+1$ не очевидна; тем не менее, все многочлены Φ_n, как мы вскоре докажем, неразложимы. ■

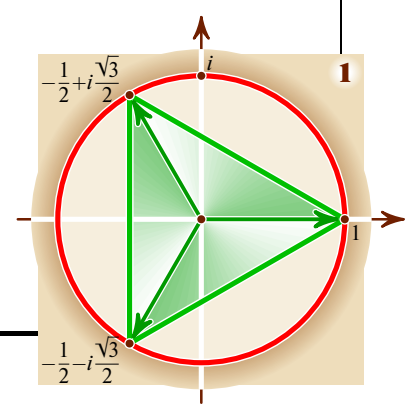
Разложения с комплексными коэффициентами.

Чтобы понять, как устроены многочлены Φ_n и почему их степень есть функция Эйлера, мы будем, как это ни странно на первый взгляд, разлагать x^n-1 на множители с комплексными коэффициентами. Опять начнем с примеров.
Уравнение $\boldsymbol{x^3-1=0}$ имеет корень $x=1$ и еще два комплексных корня, которые легко найти, решив квадратное уравнение x^2+ $+x+1=0$ по обычной формуле:

$$x=\frac{-1\pm\sqrt{1^2-4\cdot1}}{2}=\frac{-1\pm\sqrt{-3}}{2}=\frac{-1\pm i\sqrt3}{2}.$$

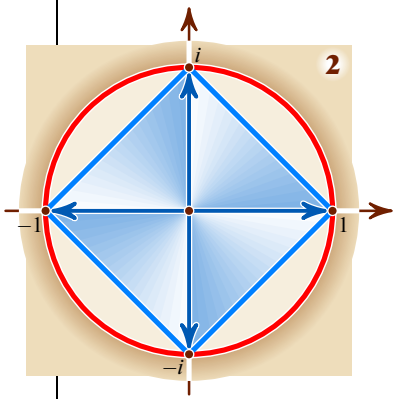
Итак, $x^3-1=(x-1)\left(x+\frac12+\frac{i\sqrt3}{2}\right)\left(x+\frac12-\frac{i\sqrt3}{2}\right)$.

Числа 1, $-\frac12+\frac{i\sqrt3}{2}$, $-\frac12-\frac{i\sqrt3}{2}$ — вершины правильного треугольника (рис. 1).



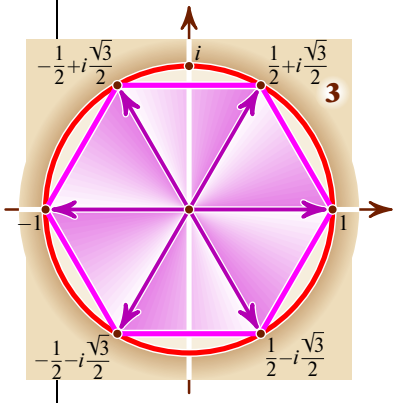
$x^4-1=(x-1)(x+1)(x+i)(x-i)$. Числа $1, i, -1, -i$ — вершины квадрата (рис. 2). Отложим на время случай $n=5$ и разберем более простой (ибо само число составное) случай.

$x^6-1=(x^3-1)(x^3+1)=(x-1)(x^2+x+1)(x+1)(x^2-x+1)$, так что



$$x^6-1=(x-1)\left(x+\frac{1}{2}+\frac{i\sqrt{3}}{2}\right)\left(x+\frac{1}{2}-\frac{i\sqrt{3}}{2}\right)\times$$
$$\times(x+1)\left(x-\frac{1}{2}+\frac{i\sqrt{3}}{2}\right)\left(x-\frac{1}{2}-\frac{i\sqrt{3}}{2}\right).$$

Числа 1, $\frac{1}{2}+\frac{i\sqrt{3}}{2}$, $-\frac{1}{2}+\frac{i\sqrt{3}}{2}$, -1, $-\frac{1}{2}-\frac{i\sqrt{3}}{2}$, $\frac{1}{2}-\frac{i\sqrt{3}}{2}$ — вершины правильного шестиугольника (рис. 3).



Теперь рассмотрим случай $n=5$. Чтобы решить уравнение

$$x^4+x^3+x^2+x+1=0,$$

разделим на x^2 и сгруппируем:

$$x^2+x+1+\frac{1}{x}+\frac{1}{x^2}=0,$$
$$x^2+\frac{1}{x^2}+x+\frac{1}{x}+1=0,$$
$$\left(x+\frac{1}{x}\right)^2-2+\left(x+\frac{1}{x}\right)+1=0.$$

Сделав замену $x+\frac{1}{x}=y$, получим квадратное уравнение $y^2+y-1=0$, откуда $y=\frac{-1\pm\sqrt{5}}{2}$. Осталось решить уравнения $x+\frac{1}{x}=\frac{-1\pm\sqrt{5}}{2}$. Это легко сделать, но получаются громоздкие ответы. И по ним не очевидно, что полученные корни (вместе с числом 1) делят единичную окружность на 5 равных частей. ■

Пользуясь тригонометрической формой комплексных чисел и тем, что при умножении аргументы складываются, получаем формулу Муавра:

$$(r\cdot(\cos\alpha+i\sin\alpha))^n=r^n\cdot(\cos n\alpha+i\sin n\alpha).$$

Следовательно, модули всех решений уравнения $x^n=1$ равны 1, а аргументы удовлетворяют условию $n\alpha=360^\circ k$, где k — целое число. Следовательно, корни степени n из 1 — это числа вида $\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}$, где $k=1,\ldots,n$, то есть вершины правильного n-угольника, вписанного в окружность радиуса 1.

Корни n-й степени из единицы, то есть решения уравнения $x^n=1$, заодно являются и корнями mn-й степени: $x^{nm}=1$. Например, всякий корень 3-й степени является заодно корнем 12-й степени.

Определение. Корень называют первообразным степени n, если он не удовлетворяет никакому уравнению $x^m=1$ при натуральном $m<n$.

Например, 1 — единственный первообразный корень степени 1; -1 — первообразный корень степени 2; $-\frac{1}{2}\pm\frac{i\sqrt{3}}{2}$ — первообразные корни степени 3.

Теперь мы готовы дать определение многочлена деления круга.

Эварист Галуа (1811—1832) родился в небольшом городе Бур-ля-Рен близ Парижа. Его отец руководил учебным заведением для юношей, а с 1813 г. был мэром города. Эварист рос болезненным и впечатлительным мальчиком. Первоначальное образование получил под руководством матери, большой поклонницы античной культуры. Двенадцати лет поступил в Королевский коллеж в Париже. Первые три года его считали хорошим учеником, хотя и отмечали несколько необычные манеры. Один из учителей писал: «Страсть к математике владеет им; я думаю, для него было бы лучше, если бы родители согласились, чтобы он занимался только этой наукой: здесь он теряет время и навлекает на себя наказания».

Галуа изучал сочинения Ж. Л. Лагранжа, Л. Эйлера, К. Ф. Гаусса и Н. Х. Абеля. Несчастья обрушились на него в 1827—1829 гг.: отец покончил с собой вследствие политической интриги, сам он, еще не кончив курса лицея, провалил экзамен по математике в Политехническую школу. Ко второму экзамену готовился под руководством Л.-П. Ришара (1795—1849), у которого в это же время занимались будущий математик Ш. Эрмит и будущий астроном У. Ж. Ж. Леверье (открывший в 1841 г. планету Нептун). Самым способным учеником Ришар считал Галуа. В 1829 г. Галуа опубликовал «Доказательство одной теоремы о периодических цепных дробях». Тем не менее он не поступил в Политехническую школу и во второй раз:

ему показалось, что экзаменаторы смеются над ним, задавая слишком простые вопросы; придя в ярость, он отказался отвечать.

По совету Ришара поступил в Нормальную школу, которая давала стипендию и готовила преподавателей для учебных заведений. За год пребывания в этой школе написал несколько научных работ, которые представил в Парижскую академию наук. Статьи попали к непременному секретарю академии Ж. Б. Фурье, который вскоре умер. В его бумагах нашли только часть работ Галуа.

В январе 1831 г. Галуа вновь передал в академию рукопись исследования о решении уравнений и просил президента академии «по крайней мере прочесть со вниманием» свой труд. Рукопись передали двум академикам, которые не смогли разобраться в ее идеях — слишком новы они были. Академия отвергла работу.

В это время Галуа уже был исключен из школы за республиканские политические взгляды (он опубликовал статью о директоре Нормальной школы) и сидел в тюрьме за тост, который был воспринят как оскорбление Луи Филиппа. Суд оправдал Галуа, приняв во внимание его юный возраст, но через месяц как один из вожаков манифестации молодежи он был арестован и после долгого следствия в конце 1831 г. приговорен к 6 месяцам тюрьмы. Вскоре после освобождения, в мае 1832 г., был убит на дуэли вследствие какой-то темной любовной истории. В ночь перед дуэлью пересмотрел и дополнил свою рукопись и послал ее своему другу О. Шевалье со следующей просьбой: «Ты публично попросишь Якоби или Гаусса дать заключение не о справедливости, а о важности этих теорем. После этого, я надеюсь, найдутся люди, которые найдут свою выгоду в расшифровке всей этой путаницы».

Работы Галуа (60 страниц небольшого формата) были опубликованы Ж. Лиувиллем в 1846 г. Усилиями А. Кэли, Ж.-А. Серре и других открытия Галуа превращены в теорию Галуа. «Трактат о подстановках и алгебраических уравнениях» М. Э. К. Жордана (1870) представил ее в систематическом общепонятном изложении. ■

Определение. $\Phi_n(x)=\prod\limits_{\substack{1\leqslant k\leqslant n,\\ \text{НОД}(k,n)=1}}(x-\zeta^k)$.

Таким образом, корни многочлена Φ_n — это в точности первообразные корни степени n из единицы, то есть числа вида $\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}$, где НОД$(k, n)=1$ и $1\leqslant k\leqslant n$. Степень многочлена Φ_n равна $\varphi(n)$. Поскольку каждый корень n-й степени из единицы является первообразным корнем d-й степени ровно для одного делителя d числа n, то

$$x^n-1=\prod_{n\vdots k}\Phi_k(x),$$

где произведение взято по всем делителям k числа n, знак $\vdots$ читается «делится». Покажем, как можно использовать эту формулу для нахождения Φ_n. Чтобы посчитать Φ_{81}, выпишем два разложения:

$$x^{81}-1=\Phi_1(x)\Phi_3(x)\Phi_9(x)\Phi_{27}(x)\Phi_{81}(x),$$
$$x^{27}-1=\Phi_1(x)\Phi_3(x)\Phi_9(x)\Phi_{27}(x).$$

Поделив одно на другое, получим: $\Phi_{81}(x)=(x^{81}-1)/(x^{27}-1)=x^{54}+x^{27}+1$. Таким же образом доказывается общая формула $\Phi_{p^m}=(x^{p^m}-1)/(x^{p^{m-1}}-1)$, где p — простое число, m — натуральное, а также равенства

$$\Phi_{pq}=\frac{(x^{pq}-1)(x-1)}{(x^p-1)(x^q-1)}\quad\text{и}\quad\Phi_{pqr}=\frac{(x^{pqr}-1)(x^p-1)(x^q-1)(x^r-1)}{(x^{pq}-1)(x^{qr}-1)(x^{rp}-1)(x-1)},$$

где p, q, r — различные простые числа.

Тот, кто знаком с функцией Мёбиуса, легко докажет формулу

$$\Phi_n(x)=\prod_{n\vdots d}(x^k-1)^{\mu(n/k)}.$$

Поскольку в этой формуле все коэффициенты — целые числа, а старшие коэффициенты равны 1, то Φ_n — многочлен с целыми коэффициентами. (В определении, как вы помните, участвовали комплексные числа!)

Теорема 1. ***Φ_n неразложим на множители с целыми коэффициентами.***

Доказательство. Рассмотрим простое число p, на которое не делится число n. Число ζ^p, как и ζ, — первообразный корень степени n. Пусть $f(x)=0$ и $g(x)=0$ — неразложимые уравнения с целыми коэффициентами, которым удовлетворяют числа ζ и ζ^p соответственно.

Лемма. ***$f(x)=\pm g(x)$.***

Доказательство леммы. Многочлен x^n-1 имеет общий с многочленом f корень ζ, а с g — корень ζ^p. Следовательно, x^n-1 делится и на $f(x)$, и на $g(x)$. Предположим, что многочлены f и g существенно различны (то есть отличаются не только множителем ±1). Тогда

$$x^n-1=f(x)g(x)h(x)$$

для некоторого многочлена h с целыми коэффициентами.

Многочлен $g(x^p)$ имеет число ζ своим корнем, поэтому $g(x^p)=f(x)k(x)$ для некоторого многочлена k с целыми коэффициентами. Из статьи «Малая теорема Ферма» известно сравнение $g(x^p)\equiv(g(x))^p\pmod p$. Следовательно,

$$(g(x))^p\equiv f(x)\,k(x)\pmod p.$$

Рассмотрим неразложимый в кольце $\mathbb{Z}_p[x]$ делитель ψ многочлена f. Он является делителем многочлена g^p, а поэтому и делителем g. Следовательно, x^n-1 делится на $\psi^2(x)$; поэтому его производная nx^{n-1} делится на $\psi(x)$. В силу условия $n\not\equiv0\pmod p$ единственный (с точностью до умножения на не кратную p

Все коэффициенты первых 104 многочленов деления круга равны 0 или ± 1. Но

$$\Phi_{105}=x^{48}+x^{47}+x^{46}-x^{43}-x^{42}-\mathbf{2}x^{41}-x^{40}-x^{39}+x^{36}+x^{35}+x^{34}+x^{33}+x^{32}+x^{31}-x^{28}-x^{26}-x^{24}-x^{22}-x^{20}+x^{17}+x^{16}+x^{15}+x^{14}+x^{13}+x^{12}-x^{9}-x^{8}-\mathbf{2}x^{7}-x^{6}-x^{5}+x^{2}+x+1. \blacksquare$$

Построение правильного 17-угольника циркулем и линейкой. Число $\zeta=\cos\frac{2\pi}{17}+i\sin\frac{2\pi}{17}$ — корень многочлена $\Phi_{17}=x^{16}+x^{15}+\ldots+x^2+x+1$. Геометрически ζ — вектор, соединяющий начало координат с вершиной правильного 17-угольника, вписанного в окружность радиуса 1. Число 3 является первообразным корнем по модулю 17. Поэтому ненулевые остатки от деления на 17 можно выписать по окружности, как на рисунке. Зеленые восьмиугольники подсказывают выписать следующие числа:

$$\alpha_1=\zeta^3+\zeta^{10}+\zeta^5+\zeta^{11}+\zeta^{14}+\zeta^7+\zeta^{12}+\zeta^6,$$
$$\alpha_2=\zeta+\zeta^9+\zeta^{13}+\zeta^{15}+\zeta^{16}+\zeta^8+\zeta^4+\zeta^2,$$

Легко проверить, что $\alpha_1+\alpha_2=-1$ и $\alpha_1\alpha_2=-4$. Следовательно, числа α_1 и α_2 — корни квадратного уравнения $\alpha^2+\alpha-4=0$. Красные квадраты подсказывают рассмотреть

$$\beta_1=\zeta^3+\zeta^5+\zeta^{14}+\zeta^{12},$$
$$\beta_2=\zeta^9+\zeta^{15}+\zeta^8+\zeta^2,$$
$$\beta_3=\zeta^{10}+\zeta^{11}+\zeta^7+\zeta^6,$$
$$\beta_4=\zeta+\zeta^{13}+\zeta^{16}+\zeta^4.$$

Как легко проверить, $\beta_2+\beta_4=\alpha_2$, $\beta_2\beta_4=-1$, $\beta_1+\beta_3=\alpha_1$, $\beta_1\beta_3=-1$. Следовательно, числа β_2 и β_4 удовлетворяют уравнению $\beta^2-\alpha_2\beta-1=0$. А числа β_1 и β_3 — корни уравнения $\beta^2-\alpha_1\beta-1=0$. Желтые диаметры подсказывают рассмотреть числа

$$\gamma_4=\zeta^{13}+\zeta^4 \text{ и } \gamma_8=\zeta+\zeta^{16}.$$

Сложение и умножение дают $\gamma_4+\gamma_8=\beta_4$ и

$$\gamma_4\gamma_8=\zeta^{14}+\zeta^{29}+\zeta^5+\zeta^{20}=\zeta^{14}+\zeta^{12}+\zeta^5+\zeta^3=\beta_1,$$

так что γ_4 и γ_8 — корни уравнения $\gamma^2-\beta_4\gamma-\beta_1=0$.

Наконец, число ζ удовлетворяет уравнению $\zeta+\zeta^{-1}=\gamma_8$, то есть $\zeta^2-\gamma_8\zeta+1=0$.

Значит, число ζ может быть найдено последовательным решением квадратных уравнений. А это означает, что правильный 17-угольник можно построить циркулем и линейкой. $\blacksquare$

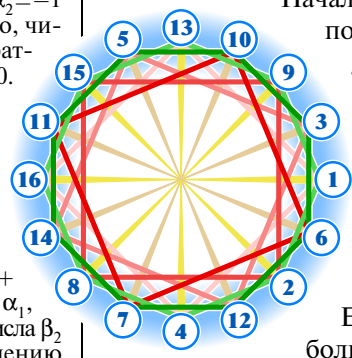
$$\cos\frac{2\pi}{17}=\frac{1}{16}\left(-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}+\sqrt{68+12\sqrt{17}-2(3+\sqrt{17})\sqrt{34-2\sqrt{17}}}\right)$$

постоянную) неразложимый делитель многочлена nx^{n-1} — это x. Но x^n-1 не делится на x. Противоречие. Лемма доказана.

Теперь доказать теорему несложно. Пусть ζ^m — первообразный корень степени n из единицы, то есть ни один из простых делителей $p_1, p_2, \ldots, p_r$ числа $m=p_1p_2\ldots p_r$ не является делителем числа n. (Числа $p_1, p_2, \ldots, p_r$ не обязательно все разные.) Так как ζ является корнем многочлена f, в силу леммы таково же и число ζ^{p_1}. Повторное применение леммы показывает, что корнем многочлена f является и число $\zeta^{p_1p_2}$. Применив лемму r раз, доказываем равенство $f(\zeta^m)=0$. Следовательно, все корни многочлена Φ_n являются корнями многочлена f. Поскольку многочлен f неразложим, а Φ_n не имеет кратных корней, то $f(x)=\pm\Phi_n(x)$. $\blacksquare$

В 1937 г. в парижском журнале «Comptes Rendus» была высказана гипотеза: ни при каком простом p многочлен $f_p(x)=x^{p-1}+x^{p-2}+\ldots+1$ не представим в виде произведения отличных от константы многочленов с вещественными неотрицательными коэффициентами. Мы дадим три доказательства, первое из которых использует комплексные числа и неразложимость f_p на множители с целыми коэффициентами, второе опирается на разложение с вещественными коэффициентами, а третье «ничего не использует», но зато требует от читателя высокой степени сосредоточенности и аккуратности.

Теорема 2. ***Если p — простое число, то в любом разложении многочлена f_p в произведение отличных от константы множителей с вещественными коэффициентами встретится хотя бы один отрицательный коэффициент.***



Начало всех трех способов доказательства одинаково. Предположим, что $f_p(x)$ разложен в произведение многочленов $g(x)=a_mx^m+a_{m-1}x^{m-1}+\ldots+a_1x+a_0$ и $h(x)=b_nx^n+b_{n-1}x^{n-1}+\ldots+b_1x+b_0$, все коэффициенты которых неотрицательны. Поскольку любой из многочленов g и h можно разделить на положительное число, умножив одновременно другой на это же число, мы будем считать, что $a_m=1$. Тогда, поскольку старший коэффициент произведения есть произведение старших коэффициентов, обязательно $b_n=1$.

Если теперь какой-нибудь коэффициент a_k окажется больше 1, сразу возникнет противоречие: при перемножении g и h члены a_kx^k и x^n дадут a_kx^{n+k}, коэффициент при котором будет больше 1, и этот коэффициент уже никак не уменьшится, ибо все другие неотрицательны. Поэтому все $a_k\leqslant 1$. Разумеется, и все $b_l\leqslant 1$.

I способ. Разложение $f_p(x)=g(x)h(x)$ имеет вид

$$x^{p-1}+x^{p-2}+\ldots+x+1=(x^m+a_{m-1}x^{m-1}+\ldots+a_1x+1)(x^n+b_{n-1}x^{n-1}+\ldots+b_1x+1).$$

Коэффициент при x^1 равен $1=a_1+b_1$. Значит, хотя бы одно из чисел a_1, b_1 отлично от 0. Пусть $a_1\neq 0$. Обратим внимание на произведения $1\cdot x^n$ и $a_1x\cdot b_{n-1}x^{n-1}$. Если $b_{n-1}>0$, то коэффициент при x^n в правой части оказывается больше 1.

Если же $b_{n-1}=0$, то противоречие получается по-другому. Вспомним, что многочлен $h(x)$ разлагается на множители вида

$$h(x)=(x-\varepsilon^{k_1})(x-\varepsilon^{k_2})\ldots(x-\varepsilon^{k_n}),$$

где $k_1, \ldots, k_n$ — натуральные числа, $\varepsilon=\cos(2\pi/p)+i\sin(2\pi/p)$. Вычислив коэффициент при x^{n-1}, получим:

$$b_{n-1}=-\varepsilon^{k_1}-\varepsilon^{k_2}-\ldots-\varepsilon^{k_n}.$$

Поскольку $b_{n-1}=0$, многочлен $x^{k_1}+x^{k_2}+\ldots+x^{k_n}$ (заметьте — многочлен с целыми коэффициентами!) имеет общий корень ε с многочленом $f_p(x)$. Но последний, как мы знаем, неприводим. Поэтому он не может иметь общий корень с многочленом, который на него не делится.

II способ. Начнем с определения. Многочлен $P(x)=\sum\limits_{i=0}^{n} a_i x^i$ называют возвратным, если для любого $0\leqslant i\leqslant n$ справедливо равенство $a_{n-i}=a_i$.
Очевидно, многочлен $P(x)$ степени n возвратный тогда и только тогда, когда $P(x)=x^n\cdot P\left(\frac{1}{x}\right)$. Произведение возвратных многочленов — возвратно.

Лемма 2. ***Многочлены $g(x)$ и $h(x)$ — возвратные.***

Лемма 3. ***Все коэффициенты многочленов g и h равны 0 или 1.***

Очевидно, теорема из них следует: подставляя $x=1$ в разложение $f_p(x)=g(x)\,h(x)$, получим противоречие с простотой числа $p=f_p(1)$.
Доказательство леммы 2. При $p=2$ утверждение очевидно. При нечетных p все множители правой части полученного из $(*)$ разложения

$$f_p(x)=\prod_{k=1}^{(p-1)/2}\left(x^2-2x\cos\frac{2\pi k}{p}+1\right)$$

являются возвратными. Значит, $g(x)$ и $h(x)$ — возвратные.
Доказательство леммы 3. Предположив противное, обозначим буквой k наименьшее из таких чисел, что хотя бы одно из чисел a_k, b_k отлично от 0 и 1. Пусть, для определенности, a_k не равно ни 0, ни 1. По лемме 1 имеем $a_{m-k}=a_k$. Значит, $a_{m-k}>0$.
Если $b_k>0$, то коэффициент при x^m после раскрытия скобок произведения $g(x)\,h(x)$ получается больше 1 (сообразите, почему!). Если же $b_k=0$, то значение выражения $a_kb_0+a_{k-1}b_1+\ldots+a_0b_k$ не равно ни 0, ни 1 (а оно должно равняться 1).
III способ не использует комплексных чисел.

Лемма 2′. ***Если возвратный многочлен $f(x)$ с неотрицательными коэффициентами разложен в произведение многочленов $g(x)$ и $h(x)$ с неотрицательными коэффициентами, причем все коэффициенты многочлена f не превосходят величины его свободного члена, то g и h — тоже возвратные.***

Лемма 3′. ***Если в условиях леммы 1′ все коэффициенты многочлена f равны 0 или 1, то и все коэффициенты многочленов g, h равны 0 или 1.***

Они сильнее, чем леммы 1 и 2, поэтому вывод теоремы 2 из лемм остается прежним. Доказательство леммы 3′, по сути, не отличается от доказательства леммы 3, поэтому нам осталось только доказать лемму 2′.
Можно считать, что $a_0=b_0=b_n=a_m=1$. Пусть хотя бы один из полиномов g, h не является возвратным. Рассмотрим наименьшее k, для которого не выполнено хотя бы одно из равенств $a_k=a_{m-k}$, $b_k=b_{n-k}$.
Коэффициент при x^k произведения вычисляется по формуле $a_0b_k+a_1b_{k-1}+\ldots+a_kb_0$. Он должен равняться коэффициенту при x^{m+n-k}, то есть сумме $a_{m-k}b_n+\ldots+a_mb_{n-k}$. Поскольку при всех $j<k$ выполнены равенства $a_j=a_{m-j}$, $b_j=b_{n-j}$, в рассматриваемых двух суммах все слагаемые, кроме крайних, совпадают. Значит,

$$a_0b_k+a_kb_0=a_{m-k}b_n+a_mb_{n-k},$$

то есть $b_k+a_k=a_{m-k}+b_{n-k}$.
Осталось доказать равенства $b_ka_{m-k}=0$ и $a_kb_{n-k}=0$, рассмотрев, соответственно, коэффициенты при x^m и x^n, и завершить доказательство леммы 2′. ■

Пусть $1+x+x^2+\ldots+x^{n-1}=F(x)\times G(x)$, $n>1$, $F(x)$ и $G(x)$ — многочлены с неотрицательными коэффициентами. Докажем, что а) все коэффициенты этих многочленов — нули и единицы;
б) один из многочленов $F(x)$, $G(x)$ можно представить в виде $(1+x+\ldots+x^{k-1})\,T(x)$, где $k>1$, а коэффициенты многочлена $T(x)$ — нули и единицы.
Доказательство. Будем изображать многочлен $x^a+x^b+\ldots$ системой отрезков $[a,a+1]\cup[b,b+1]\cup\ldots$ Умножение этого многочлена на x^t будем представлять как перенос на t единиц. Тогда, например, разложению

$$1+x+x^2+\ldots+x^{11}=$$
$$=(1+x+x^6+x^7)(1+x^2+x^4)$$

будет соответствовать система отрезков $[0,2]\cup[6,8]$, сдвиги которой на 0, 2 и 4 единицы покрывают в точности отрезок $[0,12]$. (Один множитель задает систему отрезков, другой — величины сдвигов.) Теперь задачу можно сформулировать так.
На отрезке задана некоторая система S содержащихся в нем не пересекающихся друг с другом отрезочков. Доказать, что если отрезок можно покрыть параллельными сдвигами системы S (отрезочки не должны накладываться внутренними точками, а могут только «стыковаться» в концах), то длины всех отрезочков системы S одинаковы.
Можно считать, что все сдвиги выполняются только направо. (Если есть и сдвиги налево, то возьмем наибольший из них. Вместо системы S можно рассмотреть этот ее сдвиг.)
Рассмотрим самый левый отрезочек. Очевидно, среди сдвигов должен быть сдвиг на длину k этого отрезочка. Из этого следует, что длины всех отрезочков системы не превосходят k (иначе длинный отрезочек накладывался бы на себя при сдвиге).
Если среди отрезочков системы есть отрезочек длины меньше k, рассмотрим самый левый такой отрезочек. Противоречие очевидно: точки, расположенные у его правого конца, не могут быть покрыты ни сдвигами самого этого отрезочка, ибо величины сдвигов не могут быть меньше k, ни сдвигами никакого более левого отрезочка, иначе он наложился бы на k-сдвиг самого этого отрезочка. ■

Гауссовы суммы

Сумма векторов, соединяющих центр правильного многоугольника с его вершинами, равна $\vec{0}$. *К. Ф. Гаусс в «Арифметических исследованиях» рассмотрел более сложно устроенные суммы таких векторов. Они оказались очень важны для теории чисел и получили название «гауссовы суммы».*

Проведем векторы из центра O правильного n-угольника во все его вершины $A_1, A_2, \ldots, A_n$.

Теорема. $\overrightarrow{OA_1}+\overrightarrow{OA_2}+\ldots+\overrightarrow{OA_n}=\vec{0}$.

Доказательство. Если бы сумма не равнялась нулю, то при повороте всех векторов на угол $360^\circ/n$ она должна была бы одновременно и повернуться на угол $360^\circ/n$, и остаться неизменной, поскольку при повороте векторы переходят «по циклу» друг в друга. ■

Разложим левую часть уравнения $z^n-1=0$ на множители:

$$(z-1)(z^{n-1}+z^{n-2}+\ldots+z+1)=0.$$

Значит, если $z^n=1$ и $z\neq 1$, то

$$z^{n-1}+z^{n-2}+\ldots+z+1=0.$$

Как известно, уравнение $z^n=1$ имеет n решений — «корней из единицы». Они являются вершинами правильного n-угольника, вписанного в единичную окружность, и имеют вид

$$\zeta^k=\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n},$$

где $\zeta=\cos\frac{2\pi}{n}+i\sin\frac{2\pi}{n}$, $k=1,\ldots,n$.

Зная все n корней $\zeta, \zeta^2, \ldots, \zeta^n\ (=1)$ многочлена z^n-1, мы можем разложить его на множители:

$$z^n-1=(z-\zeta)(z-\zeta^2)\ldots(z-\zeta^{n-1})(z-1).$$

Сократив обе части на $z-1$, получаем

$$z^{n-1}+z^{n-2}+\ldots+z+1=(z-\zeta)(z-\zeta^2)\ldots(z-\zeta^{n-1}).$$

Подставим в последнее равенство вместо z число 1:

$$n=(1-\zeta)(1-\zeta^2)\ldots(1-\zeta^{n-1}).$$

(Заметьте: мы подставили $z=1$ в равенство, которое получили при помощи деления на $z-1$. Хотя «на ноль делить нельзя», никакой ошибки мы не допустили: равенство многочленов не может нарушаться только в одной точке.)

Переходя к модулям, получаем: произведение $A_1A_n\cdot A_2A_n\cdot\ldots\cdot A_{n-1}A_n$ длин сторон и диагоналей, выходящих из вершины правильного n-угольника $A_1A_2\ldots A_n$, вписанного в окружность единичного радиуса, равно n. (Зная это, легко посчитать, что произведение длин всех сторон и диагоналей правильного n-угольника, вписанного в окружность радиуса R, равно $n^{n/2}\cdot R^{n(n-1)/2}$.)

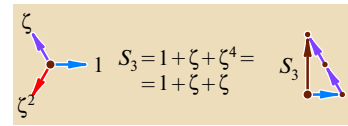
Пусть n нечетно. Тогда все множители правой части можно разбить на сопряженные (то есть симметричные относительно оси абсцисс) пары чисел $1-\zeta^k=1-\cos\frac{2\pi k}{n}-i\sin\frac{2\pi k}{n}$ и $1-\zeta^{n-k}=1-\cos\frac{2\pi k}{n}+i\sin\frac{2\pi k}{n}$. Взяв из каждой

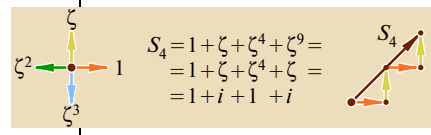
На рисунках изображены гауссовы суммы S_n для $n=3$, 4, 5, 6, и 7. Их значения при небольших n вычислить нетрудно:

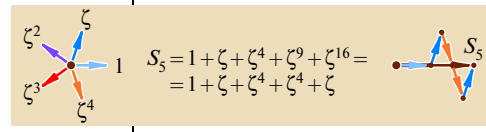
$$S_3=1+\zeta+\zeta^4=1+2\zeta=$$
$$=1+2\cdot\left(\cos\frac{2\pi}{3}+i\sin\frac{2\pi}{3}\right)=-i\sqrt{3};$$

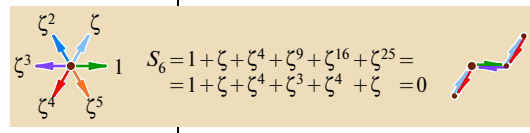
$$S_4=1+i+i^4+i^9=2(1+i);$$

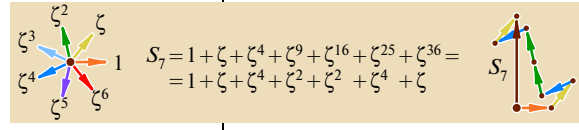
$$S_5=1+\zeta+\zeta^4+\zeta^9+\zeta^{16}=1+2\zeta+2\zeta^4=$$
$$=1+4\cos\frac{2\pi}{5}=1+4\cdot\frac{-1+\sqrt{5}}{4}=\sqrt{5}.\ ■$$











Пусть число p простое, $p>2$. Поскольку $\overline{\zeta}=\zeta^{-1}$, сопряженное к S_p число есть сумма $\overline{S_p}=1+\zeta^{-1}+\zeta^{-4}+\ldots+\zeta^{-(p-1)^2}$. Если $p\equiv 1\pmod 4$, то -1 является квадратичным вычетом по модулю p и поэтому последняя сумма отличается от $S_p=1+\zeta+\zeta^4+\ldots+\zeta^{(p-1)^2}$ только порядком слагаемых, так что $\overline{S_p}=S_p$ и $S_p^2=S_p\overline{S_p}=p$. Значит, $S_p=\sqrt{p}$ или $-\sqrt{p}$. Если же $p\equiv 3\pmod 4$, то рассматриваемые суммы имеют только одно общее слагаемое — число 1. При этом

$$S_p+\overline{S_p}=2(1+\zeta+\zeta^2+\ldots+\zeta^{p-1})=0,$$

так что $S_p^2=S_p\cdot(-\overline{S_p})=-p$. Значит, $S_p=i\sqrt{p}$ или $-i\sqrt{p}$.

Как видите, значение гауссовой суммы при простом p мы «почти нашли» — осталось выбрать лишь одно из двух значений. Но именно это самое трудное! ■

Петер Густав Лежён Дирихле (1805—1859) — немецкий математик. Доказал теорему о существовании бесконечного множества простых чисел во всякой арифметической прогрессии, первый член и разность которой — взаимно простые натуральные числа. В 1829 г. доказал, что если функция непрерывна всюду, кроме конечного множества точек, то ее ряд Фурье сходится во всех точках непрерывности к ее значению, а в точках разрыва — к среднему арифметическому пределов функции слева и справа.

При исследовании группы обратимых элементов кольца целых алгебраических чисел конечного расширения поля $\mathbb{Q}$ (частные случаи этой задачи — уравнения Пелля) применил «принцип ящиков»: если жильцов больше, чем квартир, то хотя бы двое живут в одной квартире. Это простое замечание помогает при решении многих трудных задач теории чисел и названо принципом Дирихле.

Функция Дирихле, равная 0 в иррациональных точках и 1 в рациональных, задается формулой

$$\lim_{m\to\infty}\lim_{n\to\infty}(\cos m!\pi x)^{2n}.$$

Она не интегрируема по Риману ни на каком отрезке, но, будучи почти всюду равна 0, интегрируема по Лебегу.

Принцип Дирихле в теории гармонических функций гласит, что среди принимающих заданные значения на границе области G функций f та функция, для которой минимален интеграл

$$\iiint_G \left(\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2\right)\times \times dx\,dy\,dz,$$

является гармонической. ■

пары сопряженных множителей только один множитель, мы получим число, модуль которого — квадратный корень из модуля произведения:

$$\sqrt{n}=|(1-\zeta)(1-\zeta^2)\ldots(1-\zeta^{(n-1)/2})|.$$

Без комплексных чисел это равенство можно записать так:

$$\sqrt{n}=2^{(n-1)/2}\sin\frac{\pi}{n}\sin\frac{2\pi}{n}\cdot\ldots\cdot\sin\frac{(n-1)\pi}{2n}.$$

Для четного n, как нетрудно убедиться,

$$\sqrt{n/2}=|(1-\zeta)(1-\zeta^2)\cdot\ldots\cdot(1-\zeta^{(n-2)/2})|=2^{(n-2)/2}\sin\frac{\pi}{n}\sin\frac{2\pi}{n}\cdot\ldots\cdot\sin\frac{(n-2)\pi}{2n}.\ \blacksquare$$

Гауссова сумма — это

$$S_n=1+\zeta+\zeta^4+\zeta^9+\ldots+\zeta^{(n-1)^2}.$$

После ряда безуспешных попыток Гаусс в 1811 г. доказал, что

$$S_n=\frac{1+i^{-n}}{1-i}\sqrt{n}=\begin{cases}\sqrt{n}, & \text{если } n\equiv 1\ (\mathrm{mod}\ 4),\\ 0, & \text{если } n\equiv 2\ (\mathrm{mod}\ 4),\\ i\sqrt{n}, & \text{если } n\equiv 3\ (\mathrm{mod}\ 4),\\ (1+i)\sqrt{n}, & \text{если } n\equiv 0\ (\mathrm{mod}\ 4).\end{cases}$$

В 1835 г. Дирихле при помощи рядов Фурье получил другое доказательство этого равенства. ■

Модуль гауссовой суммы, в отличие от ее точного значения, нетрудно найти. Пусть сначала n нечетно. Поскольку модуль числа равен корню из произведения числа и его сопряженного, достаточно доказать формулу

$$S_n\overline{S_n}=n,$$

то есть

$$(1+\zeta+\zeta^4+\zeta^9+\ldots+\zeta^{(n-1)^2})(1+\overline{\zeta}+\overline{\zeta^4}+\overline{\zeta^9}+\ldots+\overline{\zeta^{(n-1)^2}})=n.$$

Как известно, $\overline{\zeta}=\zeta^{-1}$. Раскроем скобки. При умножении взятого из первой скобки числа ζ^{k^2}, где $k=0,\ldots,n-1$, на взятое из второй скобки слагаемое ζ^{-m^2}, где $m=0,\ldots,n-1$, получаем $\zeta^{k^2-m^2}$. Обозначим через a и b остатки от деления на n чисел $k-m$ и $k+m$. Очевидно, $\zeta^{k^2-m^2}=\zeta^{ab}$. Любой паре остатков $(a;b)$ соответствует единственная пара $(k;m)$. Поэтому при суммировании встретятся по одному разу все n^2 разных пар $(a;b)$ и, следовательно,

$$S_n\overline{S_n}=\sum_{b=0}^{n-1}\sum_{a=0}^{n-1}\zeta^{ab}.$$

При $b=0$ все n слагаемых вида ζ^{ab} равны 1. При $1\leqslant b<n$ сумма $\sum\limits_{a=0}^{n-1}\zeta^{ab}$ равна 0. Равенство доказано.

Если $n=2m$, где m нечетно, то $\zeta^{(m+t)^2}=\zeta^{m^2+2mt+t^2}=(\zeta^m)^m\cdot(\zeta^n)^t\cdot\zeta^{t^2}=(-1)^m\cdot\zeta^{t^2}=-\zeta^{t^2}$ при $t=1,\ldots,m$, так что $S_n=0$.

Если же $n=4k$, то, обозначив $a=k-m$, имеем

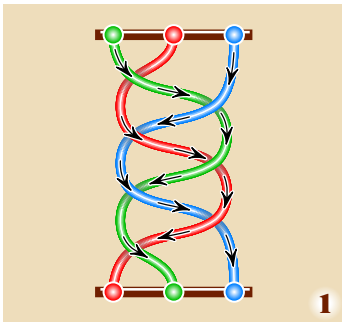
$$S_n\overline{S_n}=\sum_{k=0}^{n-1}\sum_{m=0}^{n-1}\zeta^{k^2-m^2}=\sum_{a=0}^{n-1}\sum_{m=0}^{n-1}\zeta^{(a+m)^2-m^2}=\sum_{a=0}^{n-1}\sum_{m=0}^{n-1}\zeta^{a^2+2am}=\sum_{a=0}^{n-1}\zeta^{a^2}\sum_{m=0}^{n-1}\zeta^{2am}.$$

При $a=0$ или $a=n/2$ все n слагаемых суммы $\sum\limits_{m=0}^{n-1}\zeta^{2am}$ равны 1. При всех остальных значениях a сумма $\sum\limits_{m=0}^{n-1}\zeta^{2am}$ равна 0. Следовательно,

$$S_n\overline{S_n}=(1+\zeta^{(n/2)^2})n=2n.\ \blacksquare$$

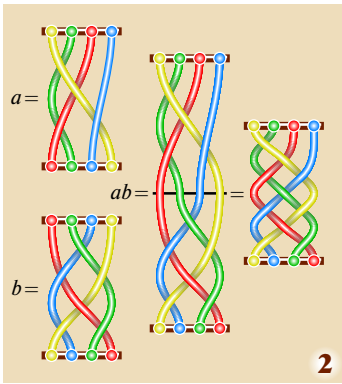
Группа кос

Теория кос — раздел математики, в равной степени относящийся к топологии и алгебре, и имеющий важные приложения в классической механике, теории функций комплексного переменного, квантовой теории поля и др.


1

Математическая коса состоит из n нитей (непрерывных кривых в пространстве), которые начинаются в n точках горизонтальной прямой и заканчиваются в n точках другой горизонтальной прямой, расположенной ниже. При этом нити должны быть нисходящими, то есть при движении вдоль нити точка должна непрерывно двигаться вниз (рис. 1).

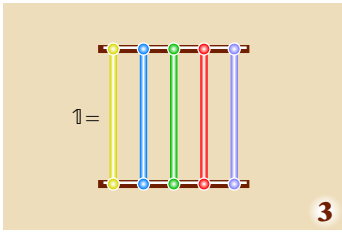
Две косы *эквивалентны*, если одну из них можно перевести в другую, растягивая и сжимая, но не разрывая нити; в этом процессе нити всегда должны оставаться нисходящими. ■



2

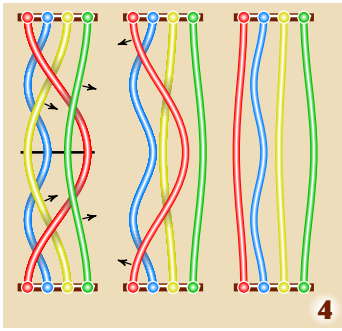
Введем операцию умножения кос a и b с одинаковыми количествами нитей: соединим нижние концы нитей косы b с верхними концами нитей косы a (рис. 2); полученную косу, сжатую в два раза в вертикальном направлении, называют произведением этих двух кос и обозначают ab.

Очевидно, умножение кос ассоциативно, то есть для любых трех кос a, b и c косы $(a\cdot b)\cdot c$ и $a\cdot(b\cdot c)$ эквивалентны. Тривиальная коса (рис. 3) играет роль единицы (и поэтому обозначается $\mathbb{1}$): для любой косы a выполнены равенства $a\cdot\mathbb{1}=a=\mathbb{1}\cdot a$. В самом деле, «подклеив» к любой косе тривиальную, мы можем так «пошевелить» новую косу, что снова получится данная коса.

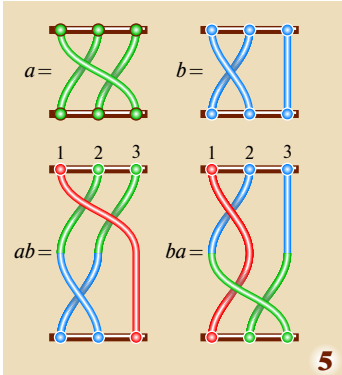


3

Пусть имеется коса b. Как построить обратную косу, то есть такую косу b^{-1}, при умножении которой на b получается $\mathbb{1}$? Очень просто: нужно зеркально отразить косу b относительно горизонтальной плоскости. Легко понять, что косу, «склеенную» из двух симметричных, можно «расплести» (рис. 4).


4

Мы привыкли, что произведение двух чисел не зависит от порядка сомножителей. Увы, для кос это не так: умножение кос некоммутативно. Придумать две косы a и b, произведение которых зависит от порядка, несложно (рис. 5). Косы ab и ba действительно различны (неэквивалентны): первая нить косы ab приходит в крайнее правое положение, а первая нить косы ba — в крайнее левое. Подведем итог в виде теоремы.



5

Теорема о косах. ***Умножение кос обладает следующими свойствами*:**

1) ***для любых кос a, b и c***

$$(a\cdot b)\cdot c=a\cdot(b\cdot c)$$ **(*ассоциативность*);**

2) ***существует такая коса*** $\mathbb{1}$**, *что для любой косы a***

$$a\cdot\mathbb{1}=a=\mathbb{1}\cdot a$$ **(*наличие единицы*);**

3) ***для любой косы b найдется такая коса*** b^{-1}**, *что***

$$bb^{-1}=\mathbb{1}=b^{-1}b$$ **(*наличие обратного элемента*).**

Иными словами, косы образуют группу, которую обычно обозначают B_n (англ. braid — «коса»). Каждой косе соответствует перестановка — элемент группы S_n. Таким образом, имеем естественный гомоморфизм $B_n\to S_n$. Поэтому умножение кос, как и умножение перестановок, при $n>2$ некоммутативно. Группа B_2 изоморфна аддитивной группе целых чисел.

Обозначим через b_k косу из n нитей, k-я нить которой проходит «под» $(k+1)$-й, а остальные нити вертикальны (рис. 6). Символами $b_1, b_2, \ldots, b_{n-1}$ можно закодировать любую косу из n нитей: «разрежем» косу так, чтобы между двумя соседними линиями разреза оказался ровно один «перекресток», то есть коса b_k или b_k^{-1}, и выпишем подряд все полученные буквы (рис. 7).

Умножение кос в этих обозначениях записывается очень просто, например,

$$b_1 b_3^{-1} b_2 \cdot b_3^{-1} b_2^2 b_1^{-1} = b_1 b_3^{-1} b_2 b_3^{-1} b_2^2 b_1^{-1}$$

(b_2^2 означает $b_2 b_2$), то есть достаточно стереть знак умножения. Некоторые записи можно сократить: например, $b_1 b_3 b_3^{-1} b_2^{-1} b_2^2$ и $b_1 b_2$ — одна и та же коса, поскольку

$$b_k b_k^{-1} = \mathbb{1} = b_k^{-1} b_k. \qquad (1)$$

Между символами $b_1, b_2, \ldots, b_{n-1}$ есть и другие соотношения. Например,

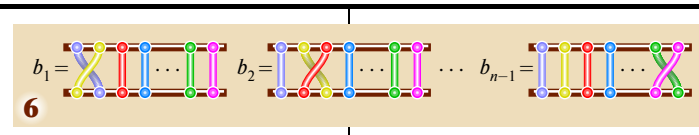
$$b_k b_j = b_j b_k. \qquad (2)$$

Обратите внимание, что это соотношение выполнено не для всех k и j, иначе умножение было бы коммутативным. Но если $k+1<j$ (случай $j+1<k$ полностью аналогичен), косу $b_j b_k$ можно получить из косы $b_k b_j$, «сдвинув» точку пересечения k-й и $(k+1)$-й нитей вниз, а j-й и $(j+1)$-й — вверх (рис. 8). Таким образом, эти косы эквивалентны. Кроме того, имеет место формула

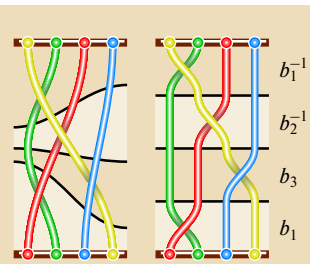
$$b_k b_{k+1} b_k = b_{k+1} b_k b_{k+1} \qquad (3)$$

(косу $b_k b_{k+1} b_k$ можно превратить в $b_{k+1} b_k b_{k+1}$ последовательными преобразованиями, рис. 9).

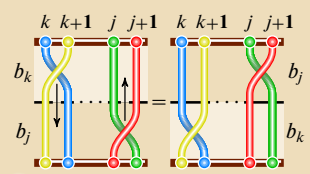
Теорема Артина. ***В группе кос все равенства вытекают из соотношений*** **(1), (2), (3).**



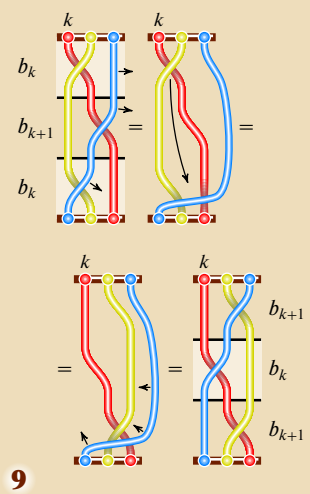
6



7



8



9

Иными словами, любое верное равенство можно получить, комбинируя эти соотношения. Например, равенство

$$b_1 b_2 b_4 = b_2 b_1 b_2 b_4 b_1^{-1}$$

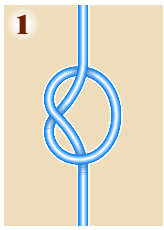
можно получить, перемножив равенства $b_1 b_2 b_1 = b_2 b_1 b_2$ и $b_1^{-1} b_4 = b_4 b_1^{-1}$ (при этом из-за некоммутативности умножать нужно обязательно с одной и той же стороны). При помощи нескольких таких операций можно получить любое равенство в группе кос. ■

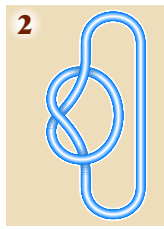
Существует алгоритм, придуманный французским математиком П. Деорнуа, который позволяет выяснить, эквивалентны ли две данные косы. Идея в том, что косы нужно закодировать и к полученным словам, состоящим из букв $b_1, b_2, \ldots, b_{n-1}$ (возможно, в отрицательных степенях), применять некоторые алгебраические преобразования. Таким образом в решении топологической задачи помогает алгебра. ■

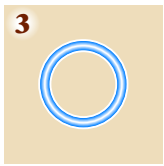
Эмиль Артин (1898—1962) — немецкий математик, один из создателей аксиоматической алгебры. Занимался также топологией и теорией чисел. В топологии построил удивительные примеры «диких» узлов. В теории чисел доказал знаменитый закон взаимности. В алгебре, кроме группы кос, придумал класс колец (множеств с двумя операциями), которые теперь носят его имя: артиновы кольца.

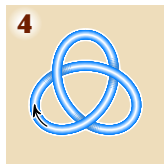
Соотношение (3) он обнаружил при следующих обстоятельствах. Администрация некоторой ткацкой фабрики предложила ему как математику изучить, как можно переплетать нити, чтобы получить новые ткани. Он этим занялся и даже кое-что полезное для ткацкого дела придумал, но не это оказалось главным. А главное — он очень увлекся математической стороной дела и придумал теорию кос. ■

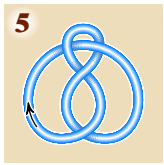
Соотношение (3) называют равенством Артина или уравнением кос, а в последнее время еще и уравнением Янга—Бакстера. Лауреат Нобелевской премии физик Янг вывел это уравнение, исследуя теорию элементарных частиц на прямой, которые меняются местами. Математик и специалист по статистической физике Бакстер, рассматривая превращения воды в лед, пришел к тому же уравнению. Теория кос тесно связана с физикой. Недаром одну из филдсовских медалей (присуждаемых только за математические достижения) получил физик Виттен. ■

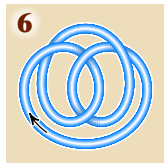


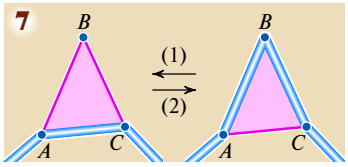


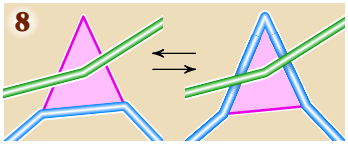


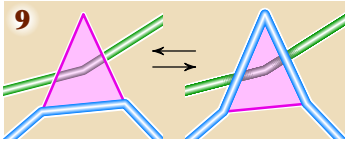


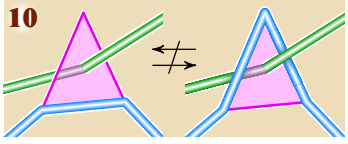


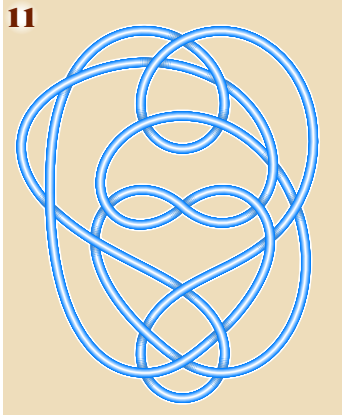












Теория узлов

Теория узлов — ветвь топологии, изучающая математические модели реальных узлов (из веревок, шнуров, ниток). Разумеется, в математике узел — это некая абстракция: веревку моделирует бесконечно тонкая, гибкая и растяжимая нить, то есть, с математической точки зрения, гладкая замкнутая кривая в пространстве.

Рассматривая математический узел, нужно либо как-то зафиксировать его концы (обычно говорят, что один конец уходит в бесконечность «вверх», а другой — в бесконечность «вниз», рис. 1), либо просто соединить их (рис. 2). В последнем случае модель узла — замкнутая несамопересекающаяся кривая в пространстве. Будем считать, что эта кривая является ломаной, то есть состоит из отрезков (впрочем, на рисунках мы почти всегда будем изображать узлы в виде гладких кривых, считая отдельные звенья ломаной очень маленькими). Самый простой узел — тривиальный (рис. 3). Узел называется нетривиальным, если он не эквивалентен тривиальному, то есть его нельзя распутать — «пошевелить» (возможно растягивая, но не разрывая нить) так, чтобы он превратился в тривиальный.

Вот несколько примеров нетривиальных узлов: узел на рисунке 4 называется трилистником, узел на рисунке 5 — восьмеркой. (Обычно узлы рассматривают с ориентацией, то есть считают, что задано направление обхода кривой, это направление изображается стрелкой.) Подумайте, как назвать узел, изображенный на рисунке 6. ■

Дадим определение эквивалентности узлов математически строго. Напомним, что узел — это ломаная. С этой ломаной можно производить следующие *элементарные операции* (рис. 7):

(1) два последовательных звена AB и BC ломаной ABC заменить звеном AC;

(2) звено AC заменить двузвенной ломаной ABC;

обе операции разрешены, только если треугольник ABC не пересекается (в пространстве) ни с какими другими кусками нашего узла. Например, в ситуациях, показанных на рисунках 8, 9, эти операции производить можно, а в ситуации, показанной на рисунке 10, — нельзя.

Теперь назовем два узла *эквивалентными*, если их можно элементарными операциями превратить в совершенно одинаковые (совмещаемые сдвигом) узлы. Под словом «узел» мы часто будем понимать не только конкретную кривую, но и все множество кривых, эквивалентных данной.

Зная, какие узлы считаются эквивалентными, естественно поставить задачу о классификации узлов: перечислить все узлы, изобразив по одному из каждого класса эквивалентности, и указав способ (алгоритм) показывающий к какому классу относится данный узел. Частным случаем этой задачи является задача распутывания: по данному изображению узла узнать, является ли он тривиальным или нет. Например, тривиален ли узел, показанный на рисунке 11?

В настоящее время, задача классификации не имеет удовлетворительного решения (хотя и решена в принципе). Для узлов с малым числом перекрестий (до 16) она решается с помощью таблиц узлов, в которых изображаются узлы с данным числом перекрестий. На рисунке 12 изображены все простые (определение см. ниже) узлы с 3, 4, 5, 6, 7 и 8 перекрестиями.

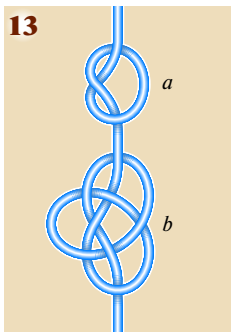
А как доказать, что два узла не эквививаленты (в частности, как доказать, что какой-нибудь узел не тривиален)? Для этого используются инварианты узлов (см. ниже), например, полином Конвея. Если сосчитать полином Конвея трилистника 3_1, тривиального узла и узла 6_3, получим, соответственно, x^2+1, 1 и $-x^4+x-1$, откуда следует, что узлы 3_1 и 6_3 не тривиальны и не эквивалентны между собой. ■

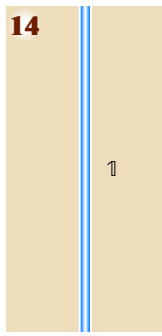
Умножение узлов. Оказывается, в множестве узлов существует своеобразная арифметика, напоминающая арифметику натуральных чисел относительно умножения. Но как умножать узлы? Если считать узлы кривыми, концы которых уходят в бесконечность, то умножение узлов определяется естественным образом: произведение узлов a и b — это просто нить, на которой завязан сначала узел a, затем узел b (рис. 13). Это умножение ассоциативно: для любых узлов a, b и c

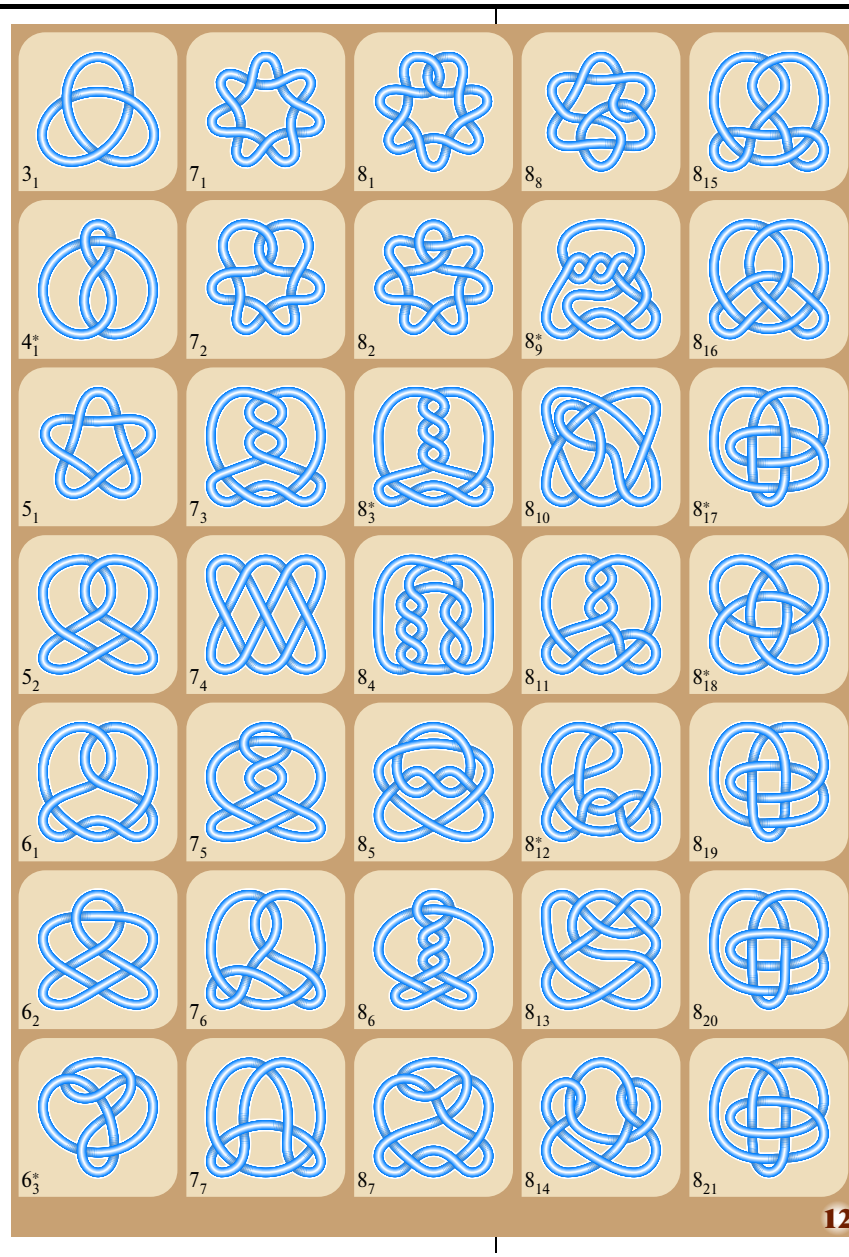
$$(a\cdot b)\cdot c=a\cdot(b\cdot c).$$

Ясно, что тривиальный узел $\mathbb{1}$ (то есть просто вертикальная прямая, рис. 14) является единичным элементом: $a\cdot\mathbb{1}=a=\mathbb{1}\cdot a$ для любого узла a. Но, в отличие от кос, ни один нетривиальный узел не имеет обратного: если мы завяжем на нити два узла, а затем потянем за концы, то узлы не развяжутся.

Покажем, что два узла, завязанных на одной нити, можно переставить. Действительно, пусть на нити завязан сначала узел a, затем узел b

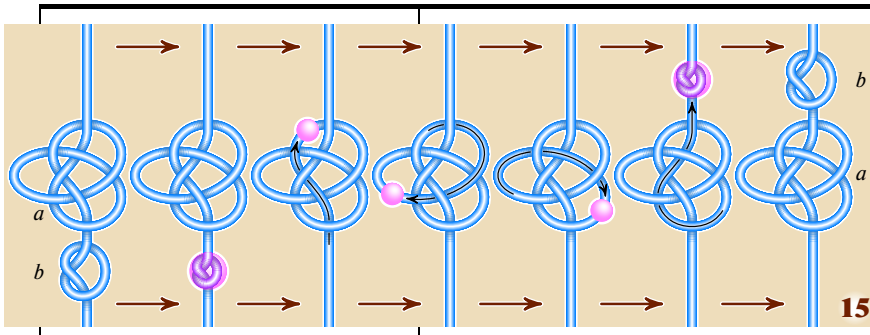






n	Количество узлов с n перекрестиями	n	Количество узлов с n перекрестиями
0	1	9	49
1	0	10	165
2	0	11	552
3	1	12	2176
4	1	13	9988
5	2	14	46 972
6	3	15	253 293
7	7	16	1 388 705
8	21		

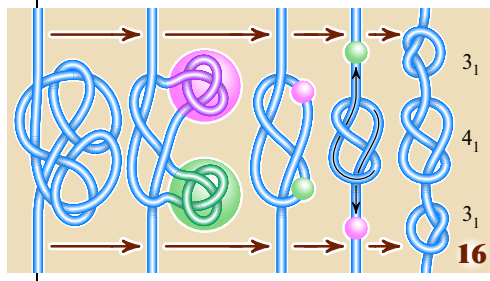
Каждый узел в таблице занумерован числом перекрестий и индексом — его условным номером среди узлов с данным числом перекрестий. Звездочкой отмечены узлы, эквивалентные своему зеркальному отражению. В 1877 г. английский математик и физик П. Дж. Тейт классифицировал узлы не более чем с 7 перекрестиями; на сегодняшний день при помощи компьютеров классифицированы узлы, в которых не более 16 перекрестий. Было бы интересно найти формулу для количества простых узлов с n перекрестиями, но пока в этом поиске успехов нет. ■



(рис. 15). Сперва, не трогая узел a, «затянем» узел b в маленький узелок. Затем заключим этот узелок в стеклянный шарик и будем двигать его вдоль нити. В итоге шарик окажется с другой стороны от узла a и его можно превратить опять в узел b. Таким образом, умножение узлов коммутативно:

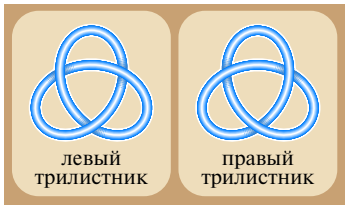
$$a \cdot b = b \cdot a.$$

Теорема об узлах. ***Узлы образуют ассоциативную и коммутативную систему относительно умножения. В этой системе есть единичный элемент, но нет обратных элементов.***

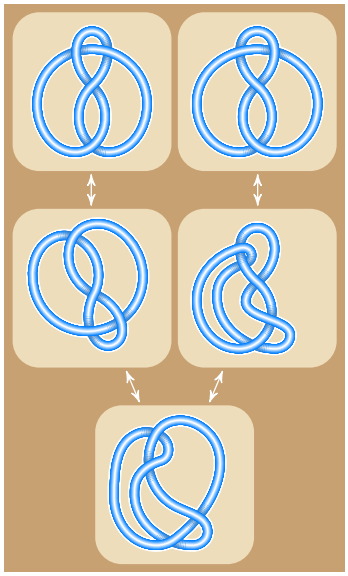


У узлов есть еще одно арифметическое свойство — среди них можно выделить *простые узлы*, то есть такие узлы, которые нельзя представить в виде произведения двух (нетривиальных) узлов. Например, все узлы, приведенные в таблице узлов (рис. 12), — простые. На рисунке 16 показано, как некоторый составной узел превращается в произведение трех простых. Как и для натуральных чисел, можно показать, что разложение на простые узлы единственно с точностью до порядка. ■

Построив зеркальное отражение трилистника, получаем узел — правый трилистник. Никакой последовательностью элементарных операций его нельзя превратить в исходный (левый) трилистник: они не эквивалентны. Это — теорема, которая доказывается очень трудно. Требуется очень тонкий инвариант, который чувствует, что эти узлы хотя и симметричны, но не эквивалентны.



Посмотрите на восьмерку и ее зеркальное отражение. Казалось бы, есть левая восьмерка, а есть правая. Но на самом деле эти две восьмерки эквивалентны! На рисунке показано, как преобразовать одну из них в другую. ■



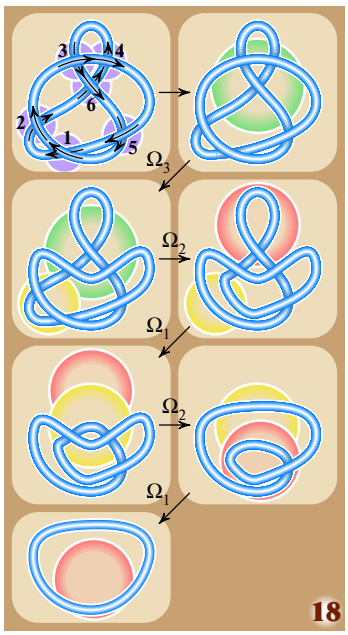
Вернемся к задаче распутывания узлов.

Не может ли компьютер помочь в ее решении? Оказывается, хорошо запрограммированный компьютер развязывает (тривиальные) узлы гораздо быстрее человека, даже наделенного незаурядным пространственным воображением. Чтобы понять, как компьютер это делает, нам потребуется немного теории (придуманной, кстати, в докомпьютерную эру живыми людьми). Первый шаг в этой теории состоит в сведении (сложной) пространственной задачи развязывания узла к (более простой) задаче применения некоторых операций к диаграмме узла на плоскости. Эти операции придумал в 1920-е гг. немецкий математик К. В. Ф. Рейдемейстер (1893—1971), они изображены на рисунке 17.

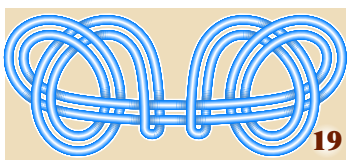
Эти операции выполняются так. Пусть дана плоская диаграмма (проекция) узла, который мы намереваемся распутать, скажем, диаграмма узла рисунка 18. Проведем на диаграмме окружность так, чтобы внутри ее оказалась одна из конфигураций, обведенных на рисунке 17, и заменим ее внутри окружности на парную конфигурацию. Так, на рисунке 18 мы провели окружность, охватывающую конфигурацию из трех дуг, и, применив операцию Ω_3, получили вторую диаграмму рисунка 18. Затем мы провели окружность, охватывающую два перекрестия и, применив операцию Ω_2, получили следующую диаграмму. Читателю предлагается самостоятельно отследить процесс распутывания до конца, поочередно обращаясь к рисункам 17 и 18.

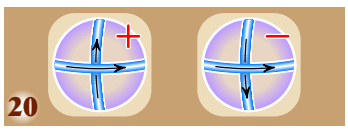
Оказывается, что любой (тривиальный!) узел можно развязать таким образом. А именно, имеет место следующее утверждение.

Лемма Рейдемейстера. ***Два узла в пространстве эквивалентны тогда и только тогда, когда плоскую диаграмму одного из них можно превратить в диаграмму другого с помощью операций Рейдемейстера. В частности, узел тривиален тогда и только тогда, когда его плоскую диаграмму можно распутать с помощью операций Рейдемейстера.*** ■



У этой истории есть мораль, выходящая за рамки теории узлов. А именно, пример развязывания узлов показывает, что жадность (здесь — желание сиюминутно упростить или улучшить ситуацию) не всегда приводит к цели. В деле распутывания узлов, как и в жизни, следует проявлять дальновидность: прежде чем упрощать, иногда сначала целесообразно еще более усложнить ситуацию, и только после этого удастся упростить ее окончательно. ■





На первый взгляд кажется, что лемма Рейдемейстера не только сводит задачу распутывания узла к выполнению простеньких операций над кривой на плоскости, но и дает быстрый и эффективный алгоритм для распутывания. Действительно, заметив, что среди операций Рейдемейстера две уменьшают число перекрестков диаграммы узла (а именно, убирание петель Ω_1 и уничтожение парных перекрестков Ω_2), при распутывании будем применять только эти «упрощающие» операции. Такой алгоритм (стремящийся на каждом шагу упростить узел, то есть уменьшить число перекрестков) называется *жадным*: он «жаден» до упрощения, он не думает вперед и на каждом шагу делает ход, добиваясь сиюминутного упрощения.

Вернемся к рисунку 18 и посмотрим, как бы повел себя жадный алгоритм в этой ситуации. Ясно, что он не стал бы применять операцию Ω_3 (как это сделали мы), а применил бы операцию Ω_2, убрав одним махом два перекрестка! Читателю мы предлагаем нарисовать окружность на рисунке 18, охватывающую парный перекресток, применить операцию Ω_2, а затем, выступая в роли жадного алгоритма, довести распутывание до конца. Ясно, что для выполнения этих указаний человека вполне может заменить компьютер.

И что же, мы нашли простой способ распутывания узлов? Увы. Жадный алгоритм не всегда умеет распутывать тривиальные узлы. В самом деле, посмотрите на узел рисунка 19. Ясно, что перед ним жадный алгоритм бессилен: нельзя произвести ни одной упрощающей операции Рейдемейстера. А между тем, поглядев немного на рисунок 19, легко убедиться, что узел-то — тривиальный! В полном соответствии с леммой Рейдемейстера, узел рисунка 19 можно развязать и с помощью операций Ω_1, Ω_2, Ω_3, но при этом начинать придется с операции Ω_2, которая увеличивает число перекрестков на 2.

Раз жадный алгоритм не годится, по какому же алгоритму компьютеру распутывать узлы? Очень просто — по алгоритму полного перебора с запоминанием. Действует он так. Диаграмма узла рисуется на экране мышкой, а затем кодируется самим компьютером в виде строки из чисел (номеров перекрестков), букв В, Н (обозначающих проход сверху или снизу) и знаков +, − (обозначающих знак ориентации перекрестков, рис. 20). Например, первый узел рисунка 18 кодируется так:

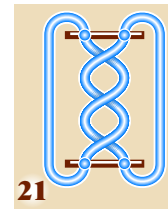
$$1\text{В}+\ 2\text{В}-\ 3\text{В}-\ 4\text{В}+\ 5\text{В}+\ 1\text{Н}+\ 2\text{Н}-\ 6\text{Н}+\ 4\text{Н}+\ 3\text{Н}-\ 6\text{В}+\ 5\text{Н}+. \qquad (*)$$

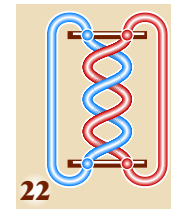
Глядя на диаграмму узла (вернее, на ее код), компьютер находит все возможные конфигурации, к которым применимы операции $\Omega_1, \Omega_2, \Omega_3$, и выполняет их. Например, он находит парные перекрестия 1 и 2 и применяет операцию Ω_2, то есть из кода (*) выбрасывает куски 1В+ 2В− и 1Н+ 2Н−. Все полученные коды он запоминает, каждый вновь преобразует всеми возможными способами и так далее. Если исходный узел был тривиальным, переборный алгоритм рано или поздно (в силу леммы Рейдемейстера) его распутает, то есть вычеркнет все символы кода, получив «пустое слово». Если же в компьютер ввести диаграмму нетривиального узла, алгоритм будет «работать вечно».

Алгоритм полного перебора страдает очевидным недостатком: он требует огромных ресурсов машинной памяти, поскольку сильно ветвится и требует запоминания всех промежуточных кодов. На практике он эффективно работает только когда не нужно выполнять операцию Ω_2 в сторону увеличения числа перекрестков. Например, узел рисунка 19 (имеющий всего 32 перекрестия) ему не по зубам.

Для более эффективного машинного распутывания узлов требуются принципиально другие способы их кодирования. Один такой способ был недавно найден российским математиком И. Дынниковым, но его описание выходит за рамки этой статьи. Отметим лишь, что компьютер, оснащенный программой Дынникова, без труда распутывает узлы с 500 перекрестками! ■

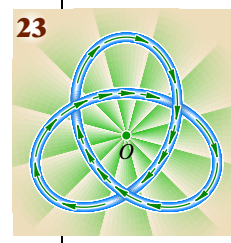
Отметим важную связь между узлами и косами. Имеется очевидный способ превратить косу в узел: надо замкнуть ее, то есть соединить верхние концы нитей с нижними (не запутывая между собой соединяющие нити). Например, в результате замыкания косы из двух нитей, изображенной на рисунке 21, получается узел, эквивалентный трилистнику. Но не всегда при замыкании косы образуется узел: после замыкания косы рисунка 22 получаются две замкнутые кривые, которые между собой зацеплены (одна как бы обматывается вокруг другой).


21

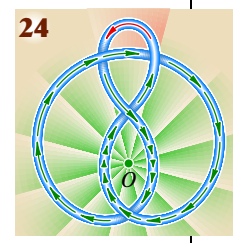

22

Возникает естественный вопрос: всякий ли узел можно получить, замкнув некоторую косу? Ответ дает теорема Дж. У. Александера (1888—1971).

Теорема Александера. ***Любой узел — это замкнутая коса.***



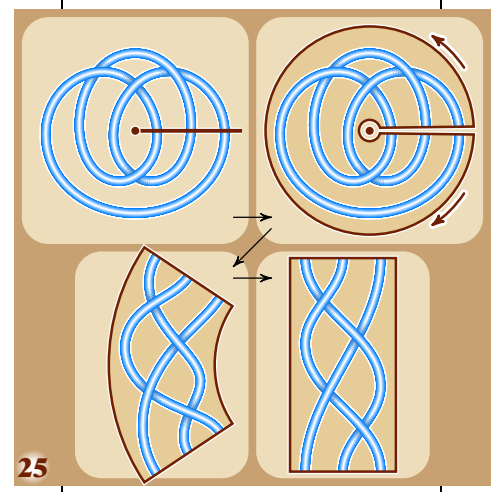
23



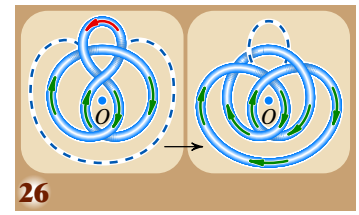
24

Узел трилистник обладает замечательным свойством: если смотреть из точки O на любой его участок, направление движения по этому участку всегда видится слева направо (рис. 23), то есть узел как бы обматывается вокруг этой точки. Тем же свойством обладает узел, изображенный на рисунке 6. А вот восьмерка этим свойством не обладает: можно проверить, что какую бы точку O мы ни взяли, всегда найдутся и участки, которые обходятся слева направо, и участки, которые обходятся справа налево (рис. 24). Значит, этот узел не обматывается.

Если узел обматывается вокруг некоторой точки, то очень легко построить косу, замыканием которой он является: разрежем наш узел по лучу, проведенному из этой точки, и развернем его (можно представить, что узел заключен в подкову, а мы выпрямляем эту подкову, превращая ее в прямоугольник, рис. 25). Легко видеть, что замыканием полученной косы является исходный узел.


25

Как же быть, если узел не обматывается? Оказывается, довольно несложно превратить его в узел, который обматывается. Мы не будем этого подробно доказывать, но продемонстрируем основную идею. Возьмем участок, который обходится «не в ту сторону», и перекинем его через точку O так, чтобы он обходился в нужную сторону (рис. 26). (Кстати, это показывает, что узел рисунка 6 — тоже восьмерка.) Так можно поочередно избавиться от всех «неправильных» участков (на самом деле это не совсем просто, и в строгом доказательстве есть некоторые тонкие моменты). ■



26

Инварианты узлов. Связь между узлами и косами, развитая в 1940-е гг. в работах нашего соотечественника А. А. Маркова, привела новозеландского математика В. Джонса к открытию знаменитого полинома (многочлена) Джонса, за изобретение которого он удостоился медали Филдса. Построение полинома Джонса слишком сложно для этой статьи, однако мы здесь опишем другой инвариант узлов — полином Дж. Х. Конвея, — который весьма похож на полином Джонса, но определяется проще.

Полином Конвея $\nabla_L(x)$ узла (или зацепления) L задается тремя аксиомами:

I. Эквивалентные узлы (зацепления) имеют один и тот же полином Конвея:

$$L \sim L' \quad \Rightarrow \quad \nabla_L(x) = \nabla_{L'}(x).$$

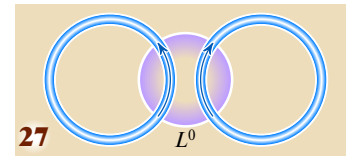
II. Полином Конвея тривиального узла равен 1 (то есть полиному нулевой степени со свободным членом 1): $\nabla_{\bigcirc}(x) = 1$.

III. Пусть L^+, L^-, L^0 — три узла (или зацепления), диаграммы которых идентичны вне некоторой окружности, а внутри ее имеют вид

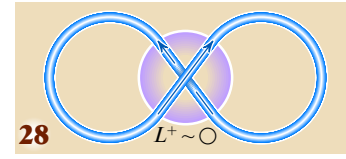
L^+:  , L^-:  , L^0:  .

Тогда между их полиномами выполняется соотношение

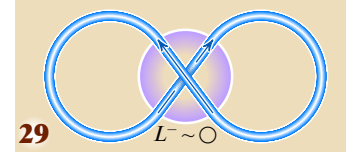
$$\nabla_{L^+}(x) - \nabla_{L^-}(x) = x\,\nabla_{L^0}(x).$$



27



28



29

Можно доказать, что для любого узла (зацепления) L, полином $\nabla_L(x)$ однозначно определяется аксиомами I, II, III. Доказательство (использующее, в частности, лемму Рейдемейстера) мы опускаем, но покажем, как эти аксиомы работают на практике.

Начнем с диаграммы «зацепления», состоящего из двух незацепленных окружностей. Проведем окружность так, как показано на рисунке 27, и назовем это зацепление L^0. Соответствующие зацепления L^+ и L^- (см. аксиому III) показаны на рисунках 28 и 29. Оба этих зацепления эквивалентны тривиальному узлу, поэтому по аксиомам I и II имеем: $\nabla_{L^+}(x)=\nabla_{L^-}(x)=1$ и, значит, $1-1=x\nabla_{\bigcirc\bigcirc}(x)$, то есть

$$\nabla_{\bigcirc\bigcirc}(x)=0.$$

Рассмотрим теперь так называемое *зацепление Хопфа* H^+. Проведем окружность, как показано на рисунке 30, и назовем это зацепление L^+; соответствующие зацепления L^- и L^0 показаны на рисунках 31 и 32. Вновь применяя аксиомы I—III и воспользовавшись равенством $\nabla_{\bigcirc\bigcirc}(x)=0$, получаем:

$$\nabla_{H^+}(x)=\nabla_{L^+}(x)=x\nabla_{L^0}(x)+\nabla_{L^-}(x)=x\nabla_{\bigcirc}(x)+\nabla_{\bigcirc\bigcirc}(x)=x.$$

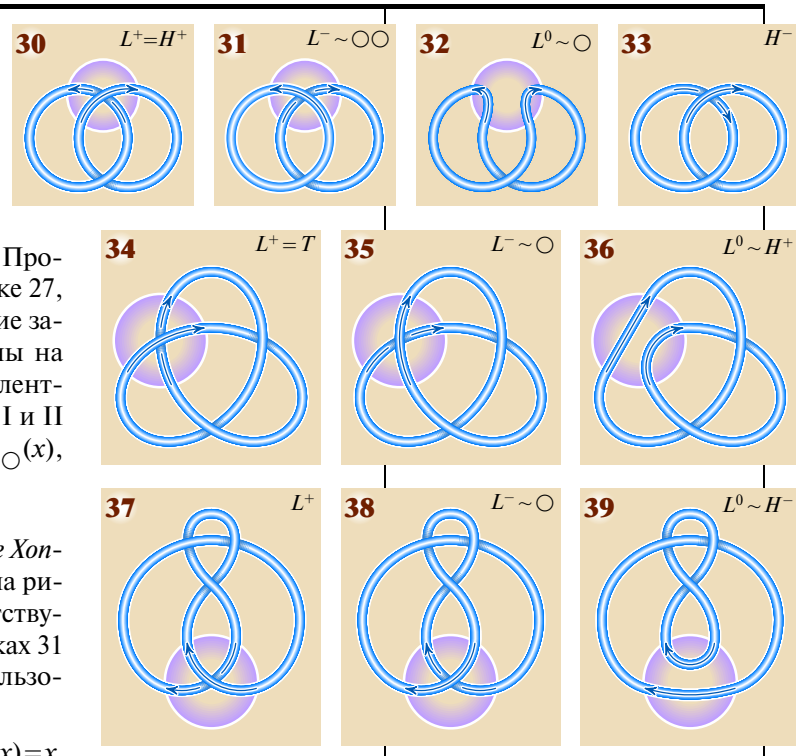
Если изменить ориентацию на одной из окружностей зацепления Хопфа H^+, получим другое зацепление Хопфа H^- (рис. 33). Аналогичный подсчет (который мы советуем читателю провести самостоятельно) показывает, что $\nabla_{H^-}(x)=-x$. Полученные результаты показывают, что зацепление Хопфа не тривиально, то есть его окружности нельзя расцепить, поскольку

$$\nabla_{H^+}=x\neq 0=\nabla_{\bigcirc\bigcirc}(x).$$

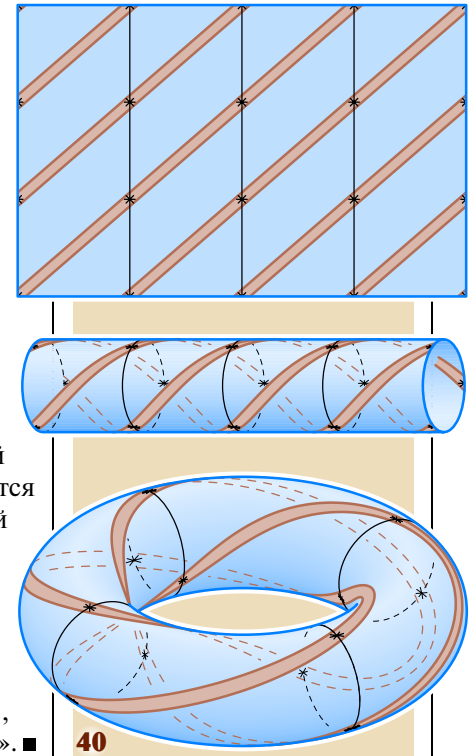
Вычислим теперь полином Конвея правого трилистника T. Для этого проведем окружность так, как показано на рисунке 34, и назовем этот узел L^+; зацепления L^- и L^0 показаны на рисунках 35 и 36. Очевидно, $L^-\sim\bigcirc$, поэтому $\nabla_{L^-}(x)=1$. Далее, $L^0\sim H^+$, поэтому $\nabla_{L^0}(x)=x$, откуда получаем (по аксиоме III)

$$\nabla_T(x)=\nabla_{L^+}(x)=\nabla_{L^-}(x)+x\nabla_{L^0}(x)=1+x^2.$$

Этот подсчет показывает, что трилистник нельзя развязать. Читателю мы предлагаем вычислить полином Конвея для другого (левого) трилистника и убедиться, что он тоже равен $1+x^2$: полином Конвея не различает левый трилистник от правого. Также предлагаем при помощи рисунков 37—39 убедиться, что полином Конвея восьмерки равен $1-x^2$, так что восьмерка не эквивалентна трилистнику. ■



Зависит ли результат вычисления полинома Конвея от порядка выбора перекрестий? Оказывается, нет. Но это непростая теорема. ■



В настоящее время теория узлов, которая достигла наибольшей интенсивности своего развития в конце 1980-х — начале 1990-х гг., остается живой наукой. Она продолжает развиваться в работах математиков, ей интересуются физики-теоретики, биологи и биохимики (в особенности те, которые изучают ДНК). В ней еще есть много нерешенных вопросов и таинственных, до конца не понятных закономерностей.

Есть, впрочем, и вполне законченные разделы. Например, про торические узлы (узлы, расположенные на поверхности тора) известно практически все. Они классифицируются парой взаимно простых чисел (p, q), $p, q\geqslant 2$. На рисунке 40 показан торический узел $(3, 4)\sim 8_{19}$ и его «развертки». ■

Математический анализ

Одно из важных открытий греческой математики — несоизмеримость длины диагонали квадрата и его стороны. Доказать это проще всего, разложив $\sqrt{2}$ в цепную дробь:

$$\sqrt{2} = [1; 2, 2, 2, 2, 2, 2, \ldots].$$

Дробь бесконечна, а рациональные числа, как известно, разлагаются в конечные цепные дроби.

Можно доказать иррациональность числа $\sqrt{2}$ и без цепных дробей, методом «от противного». В самом деле, пусть $\sqrt{2} = m/n$ — несократимая дробь, то есть m и n — взаимно простые натуральные числа. Возводя равенство в квадрат и освобождаясь от знаменателя, получаем:

$$2n^2 = m^2. \qquad (*)$$

Очевидно, m четно, то есть $m = 2a$ для некоторого натурального a. Следовательно, $2n^2 = 4a^2$, откуда

$$n^2 = 2a^2$$

и число n должно быть четным, что противоречит несократимости дроби m/n.

Так возникает необходимость изучать не только рациональные, но и иррациональные числа. Правда, число $\sqrt{2}$ не совсем «дикое»: оно является корнем многочлена $x^2 - 2$, то есть алгебраическим числом. Существуют ли числа неалгебраические, то есть числа, не являющиеся корнями ни одного многочлена с целыми коэффициентами? Да, такие числа — их называют трансцендентными — существуют. Чтобы доказать это, нам нужно прежде всего выяснить, что же это такое — вещественное число.

Определить вещественное число можно многими разными способами. Бесконечная десятичная дробь — самое простое на первый взгляд определение — требует уточнений. Дело в том, что число 0,999... равно числу 1. Эта неоднозначность записи некоторых чисел в виде бесконечных десятичных дробей приводит к тому, что если строго доказывать все теоремы, то приходится довольно часто, особенно в начале развития теории, особо разбирать такого рода исключения.

Другое определение — дедекиндовы сечения. Идея в следующем. Для любой точки x вещественной прямой можно рассмотреть все рациональные числа, расположенные слева от x, и рациональные числа, расположенные справа от x. Если число x иррациональное, то таким образом мы разбили на два класса все рациональные числа; в противном случае число x можно отнести к любому из двух классов. По сути, левый класс дедекиндова сечения — это приближения числа x рациональными числами снизу, а правый — приближения сверху. Дальше надо научиться складывать, умножать, делить дедекиндовы сечения и превратить множество дедекиндовых сечений в поле вещественных чисел.

Точную реализацию этой идеи и другие построения теории вещественных чисел вы найдете в многочисленных учебниках математического анализа, а здесь мы докажем две важные теоремы: счетность множества алгебраических чисел и несчетность множества вещественных чисел.

Множество называют счетным, если оно равномощно натуральному ряду, то есть если существует взаимно однозначное соответствие между этим множеством и натуральным рядом. Например, элементы множества рациональных чисел можно пересчитать следующим образом. На n-м этапе, где $n = 1, 2, 3, \ldots$, выписываем в порядке возрастания дроби, знаменатели которых не превосходят n, числители не превосходят по абсолютной величине числа n^2, а сама эта дробь еще не была выписана на предыдущих этапах. Получаем такую последовательность: $\frac{-1}{1}, \frac{0}{1}, \frac{1}{1}, \frac{-4}{2}, \frac{-3}{2}, \frac{-1}{2}, \frac{1}{2}, \frac{3}{2}, \frac{4}{2}, \frac{-9}{3}, \frac{-8}{3}, \frac{-7}{3}, \frac{-5}{3}, \frac{-4}{3}, \frac{-2}{3}, \frac{-1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{4}{3}, \frac{5}{3}, \frac{7}{3}, \frac{8}{3}, \frac{9}{3}, \frac{-16}{4}, \ldots$

Очевидно, каждое рациональное число рано или поздно будет выписано — а это и значит, что множество рациональных чисел счетно.

Аналогично нетрудно доказать счетность множества всех вещественных алгебраических чисел. Сопоставим каждому многочлену сумму s его степени и сумму абсолютных величин всех его коэффициентов. Поскольку множество корней у каждого отличного от тождественного нуля многочлена конечно и поскольку для любого натурального s существует лишь конечное множество многочленов с такой величиной s, то можно выписать все алгебраические числа в последовательность, как это выше сделано для рациональных чисел.

Докажем теперь несчетность множества всех вещественных чисел. Мы докажем даже несчетность отрезка [0; 1]. Рассмотрим любую последовательность чисел $x_1, x_2, x_3, \ldots$ Разделим отрезок [0; 1] на три равные части. Точка x_1 не может принадлежать

всем трем. Пусть I_1 — одна из частей, которой не принадлежит число x_1. Разделим отрезок I_1 на три равные части. Точка x_2 не может принадлежать всем трем. Пусть I_2 — одна из этих частей, не содержащая числа x_2. Далее будем поступать совершенно аналогично и построим для любого натурального n отрезок I_n, которому не принадлежит число x_n. Получили последовательность отрезков, вложенных друг в друга:

$$[0;\,1]\supset I_1\supset I_2\supset I_3\supset\ldots$$

Очевидно, существует общее для всех этих отрезков число ξ. Для любого натурального n имеем $\xi\in I_n$ и $x_n\notin I_n$, следовательно, $\xi\neq x_n$. Таким образом, число ξ не совпадает ни с одним из чисел $x_1, x_2, x_3, \ldots$ Теорема Кантора о несчетности континуума (1874) доказана.

Обратите внимание на выделенное слово «очевидно». Это — ссылка на очень важное свойство вещественных чисел — теорему о вложенных отрезках. Эта теорема равносильна тому, что всякая возрастающая ограниченная последовательность стремится к некоторому пределу.

Обратимся теперь к доказательству теоремы Коши о промежуточном значении: ***непрерывная на отрезке функция принимает все промежуточные значения.*** (Другими словами, если непрерывная функция принимает два разных значения, то она принимает и все значения, заключенные между ними.)

График непрерывной функции, наивно говоря, можно нарисовать, не отрывая карандаша от бумаги. Точное определение таково. Функция f непрерывна в точке a, если для любого $\varepsilon>0$ существует такое число $\delta>0$, что если $|x-a|<\delta$, то $|f(x)-f(a)|<\varepsilon$. Функция непрерывна на отрезке, если она непрерывна в каждой точке отрезка. Из этого определения следует, что если непрерывная функция в какой-то точке не равна нулю, то она сохраняет знак на некотором интервале, содержащем эту точку.

Теорема Коши равносильна тому, что непрерывная на отрезке функция, принимающая на концах значения разных знаков, принимает на этом отрезке нулевое значение.

Докажем ее методом дихотомии (деления пополам). Поделим отрезок пополам. Если нуль в середине отрезка, все доказано. Если же в середине не нуль, то на концах одного из отрезков функция принимает значения разных знаков. Делим его пополам и так далее. Если мы не наткнемся по ходу дела на нуль, рассмотрим общую точку всей построенной цепочки отрезков. Если в этой точке значение функции не равно нулю, то получаем противоречие с непрерывностью функции!

Теорема Коши доказана. Ее непосредственное следствие — теорема о существовании вещественного корня у многочлена любой нечетной степени (который, как и любой многочлен, — непрерывная функция). В самом деле, пусть $f(x)=x^{2n+1}+a_1x^{2n}+\ldots$ $\ldots+a_{2n}x+a_{2n+1}$ — многочлен нечетной степени. Тогда при $x>0$ имеем

$$f(x)=x^{2n+1}\left(1+\frac{a_1}{x}+\ldots+\frac{a_{2n+1}}{x^{2n+1}}\right),$$

поэтому при всех достаточно больших x имеем $f(x)>0$. Аналогично доказывается, что для всех достаточно больших по модулю отрицательных x число $f(x)$ отрицательное. Следовательно, многочлен f имеет вещественный корень.

Следующая важная теорема — теорема Вейерштрасса: ***непрерывная на отрезке функция достигает на нем своего минимального и максимального значения.*** Докажем ее. Пусть f непрерывна на некотором отрезке I. Прежде всего докажем, что f ограничена на I. Допустим, что это не так, и f принимает на I сколь угодно большие положительные значения. Тогда для любого натурального n на отрезке I есть такая точка x_n, что $f(x_n)>n$. Мы построили бесконечное множество точек $\{x_n\,|\,n\in\mathbb{N}\}$. Разделим отрезок пополам. На одном из двух отрезков останется бесконечное множество точек. Его разделим пополам и так далее. По лемме о вложенных отрезках существует точка, принадлежащая всем отрезкам. (Заметьте, мы доказали важную теорему: любое бесконечное подмножество отрезка имеет хотя бы одну предельную точку. Иногда ее формулируют так: из любой ограниченной последовательности можно выделить сходящуюся подпоследовательность.) Из определения непрерывности следует, что на некотором интервале, содержащем эту точку, функция f ограничена. Противоречие!

Итак, f ограничена сверху. Пусть f не достигает своего максимума. Это означает, что существует такое число M, что $f(x)<M$ для всех x из I, и вместе с тем функция принимает значения, сколь угодно близкие к M. Найдем теперь для каждого натурального числа m такую точку y_m, что $f(y_m)>M-\frac{1}{m}$.

Мы опять построили бесконечное множество точек. Снова делим отрезок пополам и поступаем в том же духе, как при доказательстве ограниченности. Найдем точку η, принадлежащую всем отрезкам. По построению и из определения непрерывности следует, что $f(\eta)$ не может быть не чем иным, как M. Теорема Вейерштрасса доказана. Из нее, рассмотрев вместо отрезка круг, нетрудно вывести основную теорему алгебры (попробуйте!).

Множества и операции над ними

Понятие множества — одно из самых фундаментальных в математике. Множество определяется некоторым свойством S, которым могут обладать или не обладать рассматриваемые объекты; те объекты, которые обладают свойством S, образуют множество. Например, если мы рассматриваем натуральные числа и свойство S — «быть простым», то соответствующее множество состоит из всех простых чисел $2, 3, 5, 7, 11, 13, 17, 19, \ldots$
Возникшая в конце XIX в. теория множеств выросла в весьма обширную и содержательную научную теорию. Но для большинства математиков важны не ее последние достижения, а в первую очередь то, что язык теории множеств стал общепринятым языком математики.

Что такое множество? В русском языке для обозначения тех или иных множеств используют разные слова. Например, говорят: школьный класс, букет цветов, куча мусора, спортивная команда, театральная труппа, коллекция марок, стая птиц, набор инструментов, коллектив авторов, стадо коров, рой пчел, шайка бандитов, колония микробов. . .

В математике для обозначения совокупностей употребляют, как правило, единый термин — *множество*. Можно говорить о множестве граней куба, множестве решений уравнения, множестве букв русского алфавита, множестве всех побывавших в Эфиопии ирландцев, и так далее.

Множество состоит из элементов: например, «Властелин колец» — один из элементов множества всех книг, а Стерлитамак — один из элементов множества всех городов Башкирии. Утверждение «элемент x принадлежит множеству A» символически записывают так: $x \in A$; запись $x \notin A$ означает, что элемент x не принадлежит множеству A. Если каждый элемент множества A принадлежит множеству B, то A называют *подмножеством* множества B и пишут $A \subseteq B$ или $B \supseteq A$. Например, множество корней многочлена $x^2 - x$ (состоящее из двух элементов — чисел 0 и 1) является подмножеством множества корней многочлена $x^3 - x$ (состоящего из трех элементов — чисел ± 1 и 0).

Если $A \subseteq B$ и $A \neq B$, то пишут $A \subset B$. Впрочем, в некоторых книгах знак $\subseteq$ не используют — считают, что он означает то же самое, что и $\subset$. Если множества A и B состоят из одних и тех же элементов, то есть $(x \in A) \Rightarrow (x \in B)$ и $(x \in B) \Rightarrow (x \in A)$, то эти множества называют *равными* и пишут $A = B$. (Знак $\Rightarrow$ означает «следовательно».) ■

Иногда мы не знаем заранее, содержит ли некоторое множество (например, множество корней данного уравнения) хотя бы один элемент. Поэтому целесообразно ввести понятие *пустого* множества, то есть множества, не содержащего ни одного элемента. Его обозначают символом $\varnothing$. Любое множество содержит $\varnothing$ в качестве подмножества. Подмножества некоторого множества, отличные от него самого и от $\varnothing$, называют *собственными*.

Множества бывают *конечные* и *бесконечные*. Так, множество двузначных чисел — конечное (оно содержит 90 элементов), а множество $\mathbb{N}$ всех натуральных чисел бесконечно.

Наивная теория множеств довольно быстро приводит к противоречиям. Сначала напомним известные парадоксы.

- Еще в Древней Греции знали «парадокс лжеца». Представьте себе, что некто говорит: «Я лгу». Лжет он при этом или нет?

Или представьте, что вы читаете в книге: «То, что здесь написано, — неправда». Так что же тут написано? Если правда, то тогда — это неправда; а если неправда — то это правда!

- В одном полку брадобрею приказали брить всех тех, кто не бреется сам. Должен ли брадобрей брить сам себя?
- Может ли Всемогущий Бог создать камень, который он сам поднять не сможет?
- Каждое натуральное число можно назвать, произнеся несколько слов. Например, число 2 задается одним словом, а число 22 — двумя. Давайте рассмотрим *наименьшее число, которое нельзя задать меньше чем десятью словами*. Его описание состоит всего из 9 слов, что противоречит его основному свойству.
- Будем называть прилагательное самоприменимым, если оно обладает своим свойством. Например, прилагательные «русский» и «трехсложный» самоприменимы, а «глиняный» и «двусложный» — нет. Является ли прилагательное «несамоприменимый» самоприменимым?
- (Парадокс Рассела.) Многие множества не являются своими элементами. Например, множество $\mathbb{N}$ натуральных чисел само не является натуральным числом и поэтому не является своим элементом. А множество всех множеств — тоже множество и поэтому является своим элементом. Рассмотрим множество $P = \{M \mid M \notin M\}$ всех множеств M, не являющихся своими собственными элементами. Интересно, $P \in P$ или нет? Если $P \in P$, то по определению $P \notin P$. А если $P \notin P$, то мы должны добавить элемент P в множество P!

Для борьбы с парадоксом Рассела математики запрещают (при помощи одной из аксиом теории множеств) рассматривать множества, являющиеся своими собственными элементами. В частности, нет понятия «множество всех множеств» — запрещено для борьбы с парадоксом Рассела! ■

Для записи конечных множеств используют фигурные скобки, в которых записывают элементы этого множества, причем каждый элемент записывают только один раз, а порядок, в котором элементы следуют друг за другом, не имеет значения. Например, множество цифр можно записать в виде

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

Оно состоит из 10 элементов. Множество $\varnothing$ состоит из 0 элементов; $\{\varnothing\}$ — из 1 элемента (пустого множества), $\{\varnothing, \{\varnothing\}\}$ — из 2 элементов; $\{\varnothing, \{\{\varnothing, 5\}, \{\varnothing, \{\varnothing, 23\}\}\}\}$ — тоже из 2 элементов (один из элементов — пустое множество, другой — множество, состоящее из двух элементов — $\{\varnothing, 5\}$ и $\{\varnothing, \{\varnothing, 23\}\}$).
Бесконечное множество невозможно задать, явно выписав все его элементы. Поэтому часто пишут формулы наподобие следующей:

$$D = \{x \in \mathbb{N} \mid x \text{ имеет нечетное число делителей}\}.$$

Вертикальную черту | можно читать «таких, что» или «обладающих следующим свойством»; таким образом, D — множество натуральных чисел, количество натуральных делителей которых нечетно. Легко доказать, что $D = \{k^2 \mid k \in \mathbb{N}\}$, так что одно и то же множество D можно задать разными способами. ■

Операции над множествами. Пусть A и B — произвольные множества. Их объединением $A \cup B$ называют множество, состоящее из элементов, принадлежащих хотя бы одному из множеств A и B (рис. 1):

$$A \cup B = \{x \mid x \in A \text{ или } x \in B\}.$$

Объединением произвольного конечного или бесконечного семейства множеств A_α называют множество всех элементов, принадлежащих хотя бы одному из этих множеств:

$$\bigcup A_\alpha = \{x \mid \exists \alpha\, x \in A_\alpha\}.$$

(Знак $\exists$ читают «существует». Первую букву слова Exists развернули, чтобы ни с чем ее не спутать.)
Аналогично определяют объединение любого (конечного или бесконечного) набора множеств. Объединение множеств $A_1, A_2, \ldots, A_n$ обозначают $\bigcup\limits_{k=1}^{n} A_k$, объединение бесконечной последовательности множеств $A_1, A_2, A_3, \ldots$ обозначают $\bigcup\limits_{k=1}^{\infty} A_k$.

Пересечением множеств A и B называют множество, состоящее из элементов, принадлежащих обоим этим множествам (рис. 2):

$$A \cap B = \{x \mid x \in A \text{ и } x \in B\} = \{x \in A \mid x \in B\}.$$

Например, пересечение множества всех четных чисел и множества всех чисел, делящихся на 3, состоит из всех чисел, делящихся без остатка на 6. Пересечением любого (конечного или бесконечного) набора множеств A_α называют множество, состоящее из всех тех элементов, которые принадлежат всем без исключения множествам A_α. Формулой это записывают так:

$$\bigcap A_\alpha = \{x \mid \forall \alpha\, x \in A_\alpha\}.$$

(Знак $\forall$ читают «для любого». Первую букву слова All букву поставили с ног на голову, чтобы ни с чем ее не спутать. Знаки $\exists$ и $\forall$ — так называемые кванторы — укорачивают запись многих математических утверждений. Чем дальше вы будете углубляться в математику, тем полезнее они окажутся.)
Операции пересечения и объединения по самому своему определению коммутативны:

$$A \cap B = B \cap A, \qquad A \cup B = B \cup A,$$

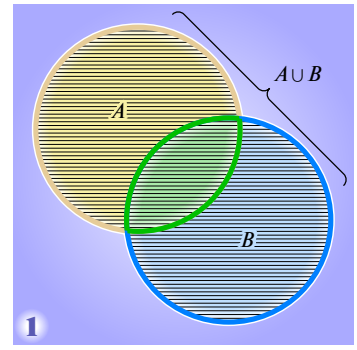
и ассоциативны:

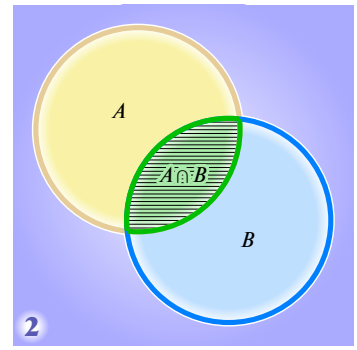
$$(A \cap B) \cap C = A \cap (B \cap C), \qquad (A \cup B) \cup C = A \cup (B \cup C).$$

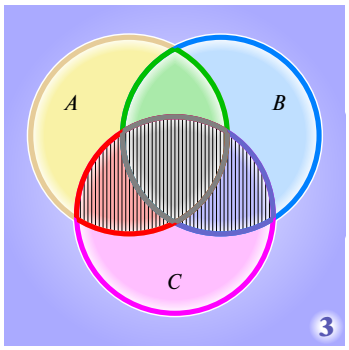
Множество $\mathbb{N}$ натуральных чисел является подмножеством множества $\mathbb{Z}$ целых чисел, которое является подмножеством множества $\mathbb{Q}$ рациональных чисел, которое, в свою очередь, является подмножеством множества $\mathbb{R}$ вещественных (или, что то же самое, действительных; впрочем, для понимания этой статьи не обязательно знать, что это такое) чисел, которое является подмножеством множества $\mathbb{C}$ комплексных (это знать еще менее обязательно!) чисел, которое является подмножеством $\mathbb{H}$ кватернионов (а что это такое, вы узнаете — если узнаете — совсем не скоро, во всяком случае существенно позже, чем о комплексных числах):

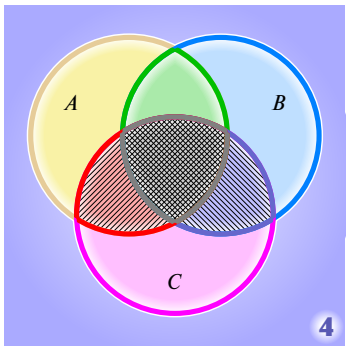
$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} \subset \mathbb{H}.$$

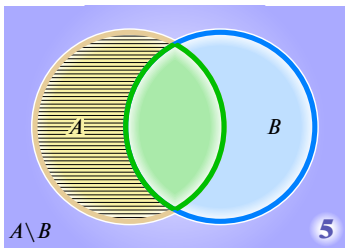
При запоминании поможет этимология: обозначения $\mathbb{N}$, $\mathbb{R}$, $\mathbb{C}$ — первые буквы слов natural, real и complex; $\mathbb{Z}$ — первая буква немецкого слова Zahl (число); $\mathbb{Q}$ — от слова quotient (отношение). ■

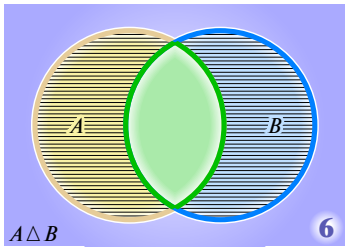












Кроме того, они взаимно дистрибутивны:

$$(A\cup B)\cap C=(A\cap C)\cup(B\cap C),$$
$$(A\cap B)\cup C=(A\cup C)\cap(B\cup C).$$

Проверим, например, первое из этих двух равенств. Проще всего это сделать при помощи так называемых кругов Эйлера—Венна: на рисунке 3 заштриховано множество $(A\cup B)\cap C$, а на рисунке 4 показаны $A\cap C$ и $A\cap B$.

Можно обойтись и без рисунков. Пусть элемент x принадлежит множеству $(A\cup B)\cap C$, то есть $x\in(A\cup B)\cap C$. По определению, это означает, что $x\in A\cup B$ и $x\in C$. Значит, x принадлежит множеству C и хотя бы одному из множеств A и B. Стало быть, x принадлежит хотя бы одному из множеств $A\cap C$ и $B\cap C$, то есть x принадлежит их объединению.

Обратно, пусть $x\in(A\cap C)\cup(B\cap C)$. Тогда $x\in\in A\cap C$ или $x\in B\cap C$. Поскольку множества $A\cap C$ и $B\cap C$ являются подмножествами множества $(A\cup B)\cap C$, то в любом случае $x\in\in(A\cup B)\cap C$, что и требовалось доказать.

Разностью множеств A и B называют множество всех тех элементов множества A, которые не принадлежат множеству B (рис. 5):

$$A\setminus B=\{x\in A\mid x\notin B\}.$$

Симметрической разностью множеств A и B называют множество всех элементов, которые принадлежат только одному из множеств A и B, но не обоим вместе (рис. 6):

$$A\triangle B=\{x\mid(x\in A\text{ и }x\notin B)\vee(x\in B\text{ и }x\notin A)\}.$$

(Знак $\vee$ читают «или».)

Название «симметрическая разность» для операции $A\triangle B$ не вполне отражает ее сущность; эта операция во многом аналогична операции взятия суммы множеств A и B. Действительно, $A\cup B$ означает, что мы связываем *неисключающим* «или» два утверждения: «элемент принадлежит A» и «элемент принадлежит B»; $A\triangle B$ означает, что те же два утверждения связаны *исключающим* «или»: элемент принадлежит множеству $A\triangle B$ тогда и только тогда, когда он принадлежит либо только A, либо только B. Множество $A\triangle B$ можно назвать «суммой по модулю два» множеств A и B (берем объединение этих двух множеств, но те элементы, которые встретились дважды, выбрасываем). ■

Дополнение и принцип двойственности.

Если в некоторой задаче все элементы рассматриваемых множеств принадлежат некоторому известному множеству S, то для каждого из рассматриваемых в этой задаче множеств (являющихся подмножествами множества S) вводят дополнение: $\overline{A}=S\setminus A$.

В теории множеств и ее приложениях весьма важную роль играет принцип двойственности, основанный на следующих двух соотношениях:

Георг Фердинанд Людвиг Филипп Кантор (1845—1918) родился в Санкт-Петербурге в семье коммерсанта. В 1856 г. из-за болезни главы семьи Канторы покинули Россию и после нескольких переездов обосновались во Франкфурте-на-Майне. Начальную школу окончил в Петербурге. В Германии сначала учился в гимназии в Висбадене, затем во Франкфурте-на-Майне, в Дармштадте, наконец, в 1860 г. перешел на общий курс Высшей ремесленной школы в Цюрихе. У него рано проявился интерес к математике, однако отец противился, поскольку занятие математикой не сулило материального благополучия. В 1862 г. отец уступил просьбам сына. Математическое образование Кантор начал в университете в Цюрихе, но в следующем же семестре вынужден был покинуть его из-за смерти отца и продолжил обучение в 1863 г. в Берлинском университете, который окончил в 1867 г., представив диссертацию «О неопределенных уравнениях второй степени». Его сочинение на право чтения лекций (хабилитация) также было посвящено теории чисел (1869).

Кантор был одним из основателей немецкого математического общества, с 1890 по 1893 г. — его первым председателем. Внес большой вклад в подготовку первого Международного конгресса математиков, состоявшегося в 1897 г. в Цюрихе.

Публикациям о бесконечных множествах предшествовал цикл статей о тригонометрических рядах, условиях их сходимости и свойствах сумм таких рядов. В завершающей цикл статье (1872) Кантор рассматривает особенности, которые может иметь ряд на различных точечных множествах. Поэтому он исследует струк-

туру точечных множеств на прямой и вводит многие понятия, используемые до сих пор, например, понятие предельной точки множеств. Это была первая его публикация, непосредственно касающаяся свойств множеств. Кантор установил счетность множеств рациональных, алгебраических чисел. Он убедился, что множество действительных чисел несчетно. Развил теорию ординальных (порядковых) чисел. Теория ординалов — один из наиболее завершенных разделов теории множеств Кантора. Вместе с тем это раздел теории множеств, в котором наиболее отчетливо виден принципиально новый взгляд на бесконечное, принятый Кантором. В традиционном понимании бесконечное представлялось как нечто «неограниченно увеличивающееся», например, как множество натуральных чисел в «естественном порядке». В теории Кантора бесконечное математически зафиксировано в определенной форме завершенного бесконечного. Кантор к такому пониманию пришел под давлением логики, почти против своей воли, ибо оно противоречило традициям, которые он ценил.

В период с 1872 по 1884 г. Кантор подготовил основные труды, в которых излагается его учение о множествах. В 1895—1897 гг. опубликовано его фундаментальное сочинение «К обоснованию учения о трансфинитных множествах». В процессе развертывания своей теории Кантор столкнулся со многими принципиальными трудностями. Одна из них — так называемая континуум-гипотеза, то есть вопрос, имеются ли мощности большие, чем мощность множества натуральных чисел, но меньшие, чем мощность континуума. Надо заметить, что эта проблема была решена только в середине XX в.: выяснилось, что ни доказать, ни опровергнуть ее нельзя!

Выявились и другие трудности: были обнаружены парадоксы, затрагивающие суть самого понятия множества. Кантор испытывал в связи со всеми фактами, касающимися основ его учения, огромное напряжение, что пагубно сказалось на здоровье. Примерно с начала XX в. оно видимо ухудшилось. В 1913 г. он вышел в отставку. ■

- дополнение объединения любого (даже бесконечного) семейства множеств равно пересечению дополнений: $S\setminus\bigcup\limits_{\alpha} A_\alpha=\bigcap\limits_{\alpha}(S\setminus A_\alpha)$;
- дополнение пересечения любого семейства множеств равно объединению дополнений: $S\setminus\bigcap\limits_{\alpha} A_\alpha=\bigcup\limits_{\alpha}(S\setminus A_\alpha)$.

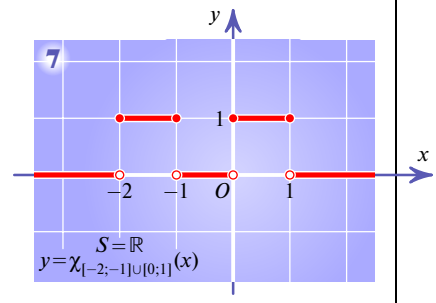
Докажем, например, равенство $S\setminus\bigcup\limits_{\alpha} A_\alpha=\bigcap\limits_{\alpha}(S\setminus A_\alpha)$. Условие $x\in S\setminus\bigcup\limits_{\alpha} A_\alpha$ означает, что $x\notin\bigcup\limits_{\alpha} A_\alpha$, то есть что x не принадлежит ни одному из множеств A_α. Значит, x принадлежит каждому из множеств $S\setminus A_\alpha$, то есть $x\in\bigcap\limits_{\alpha}(S\setminus A_\alpha)$.

Принцип двойственности состоит в том, что из любого равенства, относящегося к системе подмножеств фиксированного множества S, совершенно автоматически может быть получено другое (двойственное) равенство путем замены всех рассматриваемых множеств их дополнениями, объединений — пересечениями, а пересечений — объединениями. Например, зная принцип двойственности, мы можем не заботиться о доказательстве равенства

$$(A\cap B)\cup C=(A\cup C)\cap(B\cup C),$$

раз уж мы уже доказали равенство $(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$. ■

Характеристические функции. Для подмножеств множества S вводится *характеристическая функция* χ_A, определяемая так: характеристическая функция множества A равна в точке x единице, если $x\in A$, и нулю — если $x\notin A$. Например, если S — это множество всех вещественных чисел, а множество A является объединением отрезка $[-2;-1]$ и отрезка $[0;1]$, то график характеристической функции множества A можно изобразить так, как это сделано на рисунке 7. А если в роли A взять множество $\mathbb{Q}$ рациональных чисел, то график характеристической функции χ_A нарисовать весьма затруднительно. Тем не менее, эту функцию настолько часто используют в математическом анализе, что она даже получила специальное название — *функция Дирихле*.



Зная характеристические функции множеств A и B, легко найти характеристические функции пересечения, разности, симметрической разности: для любого m имеем

$$\chi_{A\cap B}(m)=\chi_A(m)\cdot\chi_B(m),$$
$$\chi_{A\setminus B}(m)=\chi_A(m)-\chi_A(m)\cdot\chi_B(m),$$
$$\chi_{A\triangle B}(m)=\chi_A(m)+\chi_B(m)-2\chi_A(m)\cdot\chi_B(m).$$

Чуть сложнее догадаться до формулы для характеристической функции объединения двух множеств:

$$\chi_{A\cup B}(m)=\chi_A(m)+\chi_B(m)-\chi_A(m)\cdot\chi_B(m)$$

Есть аналогичная формула и для трех множеств:

$$\chi_{A\cup B\cup C}(m)=\chi_A(m)+\chi_B(m)+\chi_C(m)-$$
$$-\chi_A(m)\cdot\chi_B(m)-\chi_A(m)\cdot\chi_C(m)-\chi_B(m)\cdot\chi_C(m)+\chi_A(m)\cdot\chi_B(m)\cdot\chi_C(m).$$

Вывести ее и обобщить на любое число объединяемых множеств — то есть доказать формулу включений-исключений — помогает переход к дополнению. Рассмотрим пример — случай трех множеств A, B, C, содержащихся в некотором множестве U. Очевидно, $\chi_{A\cup B\cup C}(m)=1-\chi_{\overline{A\cup B\cup C}}(m)=1-\chi_{\overline{A}\cap\overline{B}\cap\overline{C}}(m)=$ $=1-\chi_{\overline{A}}(m)\chi_{\overline{B}}(m)\chi_{\overline{C}}(m)=1-(1-\chi_A(m))(1-\chi_B(m))(1-\chi_C(m))$, осталось только раскрыть скобки и привести подобные! ■

Показательная функция и логарифм

Можно ли возводить числа в нецелые степени? Или даже в степень с иррациональным показателем? Что такое логарифм? На эти вопросы отвечают эта и следующая статьи энциклопедии.

Пусть $a>1$. Показательную функцию $x\to a^x$ мы определим сначала для натуральных x, затем для целых, для рациональных и, наконец, для иррациональных x.

Первый этап самый бесхитростный: $a^1=a$, $a^2=a\cdot a$, $a^3=a^2\cdot a$, вообще, для любого натурального n полагаем по определению $a^{n+1}=a^n\cdot a$. ■

Для любых натуральных m и n докажем по индукции равенства

$$a^m\cdot a^n=a^{m+n}, \qquad (*)$$
$$(a^m)^n=a^{mn}. \qquad (**)$$

При $n=1$ доказываемые равенства верны. Очевидны и переходы от n к $n+1$:

$$a^m\cdot a^{n+1}=a^m\cdot(a^n\cdot a)=(a^m\cdot a^n)\cdot a=a^{m+n}\cdot a=a^{(m+n)+1}=a^{m+(n+1)},$$
$$(a^m)^{n+1}=(a^m)^n\cdot a^m=a^{mn}\cdot a^m=a^{mn+m}=a^{m(n+1)}.\ ■$$

Распространить определение с натуральных на целые показатели, сохранив в силе формулы $(*)$ и $(**)$, можно единственным способом: из равенств

$$a=a^{1+0}=a^1\cdot a^0=a\cdot a^0$$

получаем $a^0=1$; затем из равенств

$$1=a^0=a^{n-n}=a^n\cdot a^{-n}$$

видим, что a^{-n} надо определять по формуле $a^{-n}=1/a^n$.

Прежде чем идти дальше, хорошо бы убедиться, что формулы $(*)$ и $(**)$ верны не только для любых натуральных, но и для любых целых чисел m и n. Сделать это нетрудно, хотя и хлопотно: каждое из чисел m и n может быть положительным, отрицательным или нулем. Девять случаев! А если числа m и n разных знаков, то их сумма $m+n$ то ли положительна, то ли отрицательна, то ли равна нулю. . .

Нет смысла тратить на это бумагу. Рассмотрим для примера лишь один случай формулы $(*)$. Пусть $m>0$, $n<0$ и $m+n<0$. Обозначим $m+n=-k$. Имеем:

$$a^m\cdot a^n=a^m\cdot a^{-m-k}=\frac{a^m}{a^{m+k}}=\frac{1}{a^k}=a^{-k}.\ ■$$

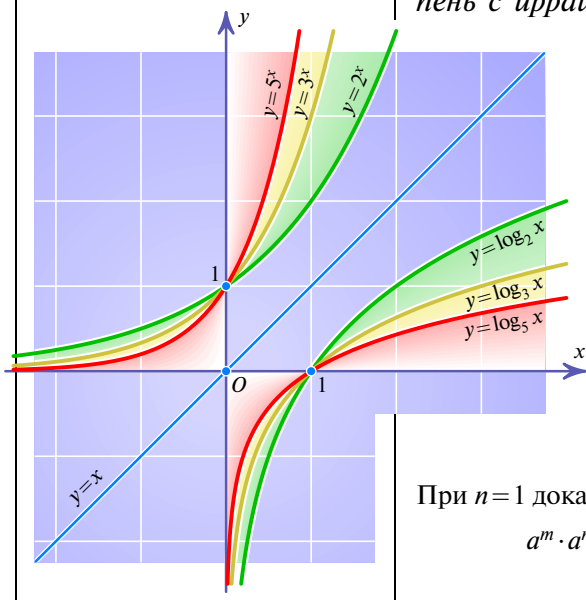
Третий этап — возведение числа a в степень с любым рациональным показателем m/n, где число m целое, а n — натуральное.

Рассмотрим сначала случай $m=1$. Частный случай формулы $(**)$ — равенство $a=a^1=(a^{1/n})^n$. Значит, величину $a^{1/n}$ разумно определять как корень n-й степени из числа a, то есть такое положительное число x, что $x^n=a$. Сразу же возникают вопросы: почему такое число x существует и почему оно единственно?

Логарифмическая функция — это функция, обратная показательной: равенство $a^x=y$, где $a>0$ и $a\ne1$, при помощи логарифмов записывается в виде

$$x=\log_a y.$$

График функции $y=\log_a x$ получается из графика функции $y=a^x$ симметрией относительно биссектрисы первого и третьего квадрантов. ■



Логарифм произведения равен сумме логарифмов:

$$\log_a xy=\log_a x+\log_a y.$$

В самом деле, обозначим $\log_a x=u$, $\log_a y=v$ и $\log_a xy=w$. По определению, $a^u=x$, $a^v=y$ и $a^w=xy$. Значит,

$$a^w=xy=a^u\cdot a^v=a^{u+v},$$

откуда $w=u+v$. ■

От одного основания логарифмической функции нетрудно перейти к другому:

$$\log_a x=\frac{\log_b x}{\log_b a}.$$

Для доказательства опять введем обозначения: $\log_a x=u$, $\log_b x=p$ и $\log_b a=q$. По определению, $a^u=x$, $b^p=x$ и $b^q=a$. Очевидно,

$$b^p=x=a^u=(b^q)^u=b^{qu},$$

откуда $p=qu$, то есть $u=p/q$. Таким образом, логарифм числа x по основанию a — это дробь, где в числителе стоит логарифм по основанию b числа x, а в знаменателе — логарифм числа a по тому же основанию b. ■

Существует — по теореме о промежуточном значении непрерывной функции. А именно, функция $f(x)=x^n$ непрерывна и обладает следующими свойствами: $f(1)=1<a<a^n=f(a)$. Таким образом, значение функции в левом конце отрезка $[1;\,a]$ меньше числа a, а в правом конце — больше. При движении вдоль оси абсцисс от точки $x=1$ к точке $x=a$ непрерывно зависящая от x величина $f(x)$ не может «перепрыгнуть» через величину $y=a$.
Единственно — поскольку функция f строго возрастает и поэтому каждое свое значение принимает лишь единожды. ■

Для любого целого m и натурального n определим $a^{m/n}$ как m-ю степень числа $\sqrt[n]{a}$. Поскольку представление рационального числа в виде дроби не единственно (например, $2/3=4/6=6/9=\ldots$), необходимо проверить корректность определения: для любых двух представлений рационального числа $m/n=r/s$ в виде дроби доказать равенство $(\sqrt[n]{a})^m=(\sqrt[s]{a})^r$.
Обозначим $\sqrt[n]{a}=x$ и $\sqrt[s]{a}=y$. По определению, $x^n=a=y^s$. В силу равенства $\dfrac{m}{n}=\dfrac{r}{s}$ имеем $ms=rn$. Значит,

$$x^{nms}=(x^n)^{ms}=a^{ms}=a^{rn}=(y^s)^{rn}=y^{srn}.$$

Из полученного равенства $(x^m)^{ns}=(y^r)^{ns}$ имеем $x^m=y^r$, поэтому

$$(\sqrt[n]{a})^m=x^m=y^r=(\sqrt[s]{a})^r.$$

Корректность определения степени с рациональным показателем доказана. Аналогично для любых дробей m/n и p/q можно подтвердить формулы $(*)$ и $(**)$, то есть доказать равенства $a^{m/n}\cdot a^{p/q}=a^{(mq+pn)/(nq)}$ и $(a^{m/n})^{p/q}=a^{mp/(nq)}$.

Построенная функция $\dfrac{m}{n}\to a^{m/n}$ строго возрастающая: если x и y — рациональные числа, причем $x<y$, то $a^y=a^{y-x+x}=a^{y-x}\cdot a^x>a^x$, поскольку $a^{y-x}>1$. ■

Для любого вещественного числа α рассмотрим множество всех рациональных чисел x, расположенных слева от α, и множество всех расположенных справа от α рациональных чисел y. Поскольку $x<\alpha<y$, то $a^x<a^y$. Определим a^α как разделяющее число этих множеств — такое число $c=a^\alpha$, что для любых рациональных x и y, между которыми расположено число α, верны неравенства $a^x<c<a^y$.
Число c существует в силу теоремы о разделяющем числе. Докажем его единственность — убедимся в том, что разность a^y-a^x при подходящем выборе чисел x и y может быть сделана сколь угодно малой. Поскольку

$$a^y-a^x=a^x(a^{y-x}-1)<c(a^{y-x}-1),$$

достаточно доказать, что величина a^t стремится к единице, когда (положительная!) величина t стремится к нулю. Предположим противное: пусть для некоторого положительного числа ε для любого положительного t верно неравенство $a^t>1+\varepsilon$. Взяв $t=1/n$, где n — натуральное число, получаем неравенство $\sqrt[n]{a}>1+\varepsilon$. Возведя обе части в n-ю степень и применив неравенство Бернулли, получаем неравенства

$$a>(1+\varepsilon)^n>1+n\varepsilon,$$

откуда $n<(a-1)/\varepsilon$, что противоречит неограниченности натурального ряда. ■

Мы убедились в корректности данного нами определения показательной функции. Поскольку любое вещественное число сколь угодно точно можно приблизить рациональными числами, то легко убедиться, что формулы $(*)$ и $(**)$ справедливы для любых вещественных чисел. ■

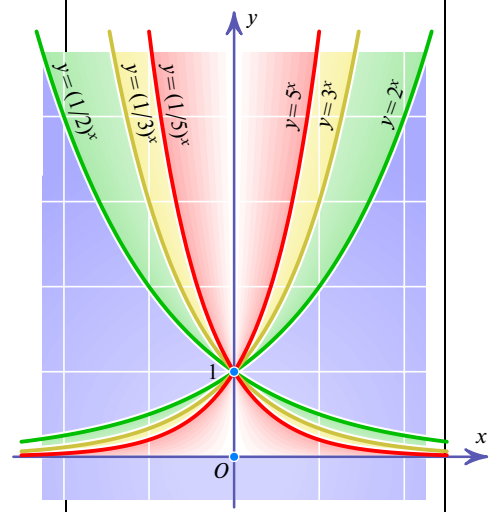
Смотрите, как тремя разными способами можно высказать одну в сущности мысль:

$$2^3=8,$$
$$2=\sqrt[3]{8},$$
$$3=\log_2 8.$$ ■

Величину b^x, где $0<b<1$, определяем формулой

$$b^x=\frac{1}{(1/b)^x}.$$

Функция $x\to b^x$ убывающая, ее график симметричен графику функции $x\to(1/b)^x$ относительно оси ординат. ■



Теорема о разделяющем числе гласит: для любых двух непустых множеств A и B вещественных чисел, обладающих тем свойством, что всякий элемент a множества A расположен левее всякого элемента b множества B, существует разделяющее множества A и B число c, то есть такое число, что $a\leqslant c\leqslant b$ для любых $a\in A$ и $b\in B$. ■

Для любого натурального числа n и для любого числа a, удовлетворяющего неравенству $a\geqslant -1$, верно следующее неравенство Бернулли:

$$(1+a)^n\geqslant 1+na.$$

Его легко доказать по индукции. **База** очевидна: $(1+a)^1\geqslant 1+a$. Нетруден и индукционный **переход**: предположив, что

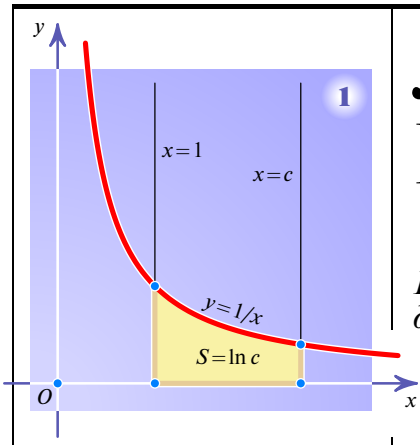
$$(1+a)^n\geqslant 1+na,$$

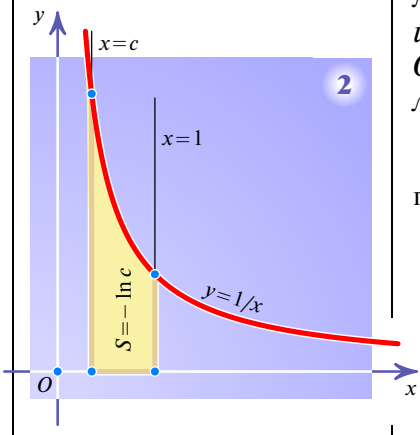
имеем

$$(1+a)^{n+1}=(1+a)^n\cdot(1+a)\geqslant$$
$$\geqslant(1+na)(1+a)=1+na+a+na^2\geqslant$$
$$\geqslant 1+(n+1)a.$$ ■

Логарифм и экспонента

В предыдущей статье показательную функцию определили сначала для натуральных значений показателя, затем для целых, рациональных и, наконец, для иррациональных. Логарифм определили как функцию, обратную показательной. Ни о производных и дифференциальных уравнениях, ни о числе e, ни о связи логарифмов с гармоническим рядом не было сказано ни слова: на том пути изложения добраться до этих понятий не так-то просто.
Однако есть и «царский путь»: сначала определить натуральный логарифм, а затем обратную к нему функцию — экспоненту.





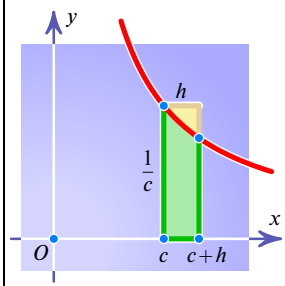
Рассмотрим гиперболу — график функции $y=1/x$. Пусть c — положительное число (рис. 1, 2). Обозначим

$$\ln c=\int_1^c \frac{dx}{x}.$$

Для тех, кто не знаком с определением интеграла, поясним: $\ln c$ — это число, абсолютная величина которого равна площади фигуры (криволинейной трапеции), ограниченной графиком функции $y=1/x$, осью абсцисс и прямыми $x=1$ и $x=c$. По определению, натуральный логарифм $\ln c$ положителен при $c>1$, равен нулю при $c=1$ и отрицателен при $0<c<1$.
Таким образом, мы определили функцию $y=\ln x$ для положительных x. Ее значения положительны при $x>1$ и отрицательны при $0<x<1$ (рис. 3). ■

Основное свойство логарифма выражено формулой

$$\ln ab=\ln a+\ln b. \qquad (*)$$



Докажем его:

$$\ln(ab)=\int_1^{ab}\frac{dx}{x}=\int_1^{a}\frac{dx}{x}+\int_a^{ab}\frac{dx}{x}=\ln a+\int_a^{ab}\frac{dx}{x}.$$

Замена $x=at$ завершает доказательство:

$$\int_a^{ab}\frac{dx}{x}=\int_1^{b}\frac{d(at)}{at}=\int_1^{b}\frac{dt}{t}=\ln b.\ ■$$

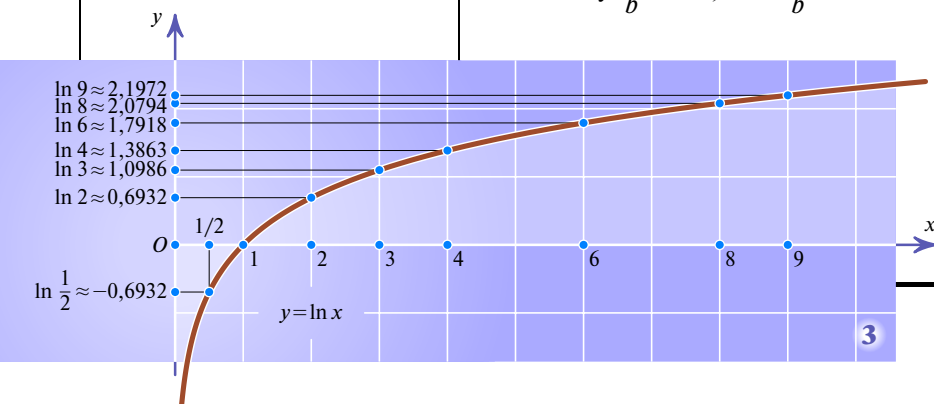
Научимся логарифмировать частное. Пусть $a>0$ и $b>0$. Поскольку $\frac{a}{b}\cdot b=a$, то $\ln\frac{a}{b}+\ln b=\ln a$, так что $\ln\frac{a}{b}=\ln a-\ln b$, то есть логарифм частного равен разности логарифмов числителя и знаменателя. ■

Формула (*) позволяет уточнить поведение функции $y=\ln x$. Прежде всего убедимся, что $\ln x$ неограниченно возрастает при возрастании x. Действительно, поскольку

Для любого положительного числа c и любого числа $h>0$ разность $\ln(c+h)-\ln c$ равна площади криволинейной трапеции, закрашенной на рисунке зеленым цветом. Чем меньше h, тем больше эта трапеция походит на прямоугольник ширины h и высоты $1/c$. Поэтому

$$\lim_{h\to 0}\frac{\ln(c+h)-\ln c}{h}=\frac{1}{c}.$$

Таким образом, производная натурального логарифма $(\ln x)'$ равна $1/x$. ■



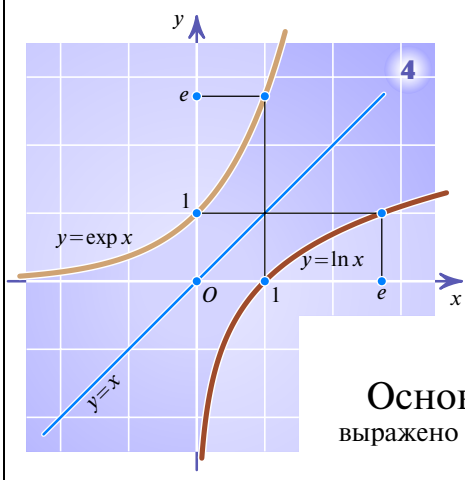
$\ln 2 > 0$, то $\ln 2^n = n \ln 2$ стремится к бесконечности при $n \to \infty$. Так что между осями координат и гиперболой $y = 1/x$ заключена бесконечно большая площадь! Исследуем теперь поведение логарифма при x, близких к нулю. Так как

$$\ln \frac{1}{2^n} = \ln 1 - \ln 2^n = -n \ln 2,$$

то при $0 < x < \frac{1}{2^n}$ имеем $\ln x < -n \ln 2$.

Итак, функция $y = \ln x$ принимает и положительные, и отрицательные сколь угодно большие по абсолютной величине значения. Воспользовавшись теоремой о промежуточном значении непрерывной функции, приходим к важному выводу: множество значений логарифмической функции — вся вещественная ось. При этом, в силу монотонности, каждое свое значение функция принимает лишь единожды. Таким образом, для любого вещественного числа y уравнение $\ln x = y$ имеет единственное решение. Его обозначают $x = \exp(y)$. ■

Функцию $y \to \exp(y)$, обратную к функции $x \to \ln x$, называют экспонентой. Для любого вещественного числа x и для любого положительного y верны равенства $\ln(\exp x) = x$ и $\exp(\ln y) = y$. График экспоненты симметричен графику натурального логарифма относительно биссектрисы первого и третьего квадрантов: равенство $y = \exp(x)$ равносильно равенству $\ln y = x$, так что графики получаются друг из друга преобразованием плоскости, при котором точка $(x; y)$ переходит в точку $(y; x)$, то есть симметрией относительно прямой $y = x$ (рис. 4).



Таким образом, экспонента — возрастающая функция, определенная на всей числовой прямой и принимающая только положительные значения. Очевидно, экспонента стремится к $+\infty$ при $x \to +\infty$ и стремится к нулю при $x \to -\infty$. ■

Основное свойство экспоненты

выражено следующей формулой:

$$\exp(a) \cdot \exp(b) = \exp(a + b). \qquad (**)$$

Для доказательства обозначим $\exp(a) = x$, $\exp(b) = y$ и $\exp(a + b) = z$. В силу основного свойства логарифма, $\ln(xy) = \ln x + \ln y$. Следовательно,

$$\ln(xy) = \ln x + \ln y = a + b = \ln z,$$

откуда $xy = z$, что и требовалось доказать. Теперь легко получить формулы

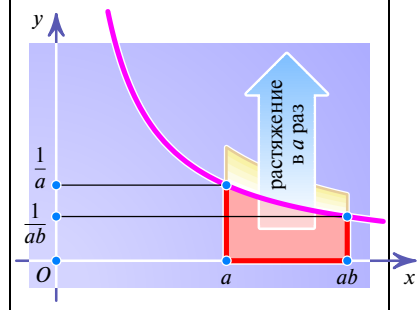
$$\exp(-x) = 1/\exp(x),$$
$$\exp(x_1 + x_2 + \ldots + x_n) = \exp(x_1) \cdot \exp(x_2) \cdot \ldots \cdot \exp(x_n). \blacksquare$$

Обозначим $\exp(1) = e$. Тогда $1 = \ln e$, так что e — это единственное число, натуральный логарифм которого равен единице (см. рис. 4). Пользуясь основным свойством экспоненты, мы докажем сначала для любого натурального, а затем для любого целого и для любого рационального числа x равенство $\exp(x) = e^x$. Прежде всего при натуральном $x = m$ имеем

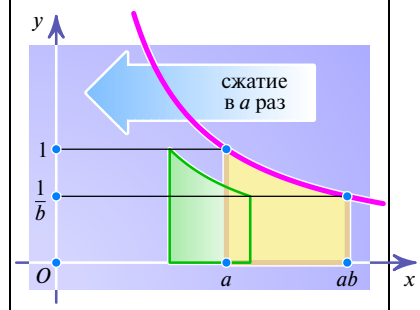
$$\exp m = \exp(1 + 1 + \ldots + 1) = \exp 1 \cdot \exp 1 \cdot \ldots \cdot \exp 1 = e^m,$$
$$\exp(-m) = 1/\exp(m) = 1/e^m = e^{-m}.$$

Поскольку $e^0 = 1 = \exp 0$, мы проверили равенство $\exp x = e^x$ для любого целого

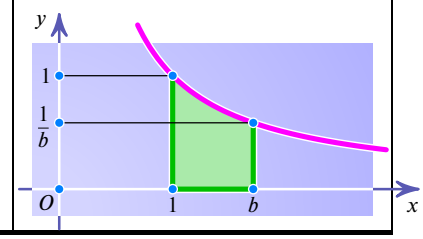
Равенство $\int_a^{ab} \frac{dx}{x} = \int_1^b \frac{dx}{x}$ геометрически очевидно. Растянем плоскость в a раз вдоль оси ординат, то есть каждую точку $(x; y)$ переведем в точку $(x; ay)$. Площади всех фигур увеличились в a раз.



Сожмем плоскость в a раз вдоль оси абсцисс, то есть каждую точку $(x; y)$ переведем в точку $(x/a; y)$. Площади всех фигур уменьшились в a раз.



Композиция $(x; y) \to (x/a; ay)$ растяжения вдоль оси ординат и сжатия вдоль оси абсцисс ширину любого прямоугольника, стороны которого параллельны осям координат, уменьшает в a раз, а высоту во столько же раз увеличивает. Значит, площади прямоугольников со сторонами, параллельными осям координат, не меняются. Вместе с ними не меняются и площади ступенчатых фигур, вписанных в криволинейные трапеции, а значит, не меняются и площади этих трапеций. В частности, площади красной и зеленой криволинейных трапеций равны. ■



числа x. Далее, для любого натурального n имеем

$$(\exp(1/n))^n = \exp\left(\underbrace{\frac{1}{n}+\frac{1}{n}+\ldots+\frac{1}{n}}_{n\text{ слагаемых}}\right) = \exp 1 = e,$$

так что $\exp(1/n) = \sqrt[n]{e} = e^{1/n}$. Для произвольного рационального числа $x = m/n$ получаем

$$\exp(m/n) = (\exp(1/n))^m = (\sqrt[n]{e})^m = e^{m/n},$$

следовательно, $\exp x = e^x$ для любого рационального числа x. Если же число x иррациональное, то формулу $e^x = \exp x$ удобно считать определением степени числа e с показателем x. (Впрочем, если показательную функцию уже определили, продолжив ее по непрерывности с рациональных значений показателя на иррациональные, то формула $e^x = \exp x$ останется в силе: ведь экспонента — непрерывная функция!)

Показательную функцию $y = a^x$ и логарифм $\log_a x$, где $a > 0$ и $a \neq 1$, теперь легко определить через экспоненту и натуральный логарифм:

$$a^x = \exp(x \ln a) = e^{x \ln a}, \qquad \log_a x = \frac{\ln x}{\ln a}.$$

Легко убедиться, что эти функции совпадают с теми, что были определены в статье «Показательная функция и логарифм». Конечно же, для любых $x > 0$ и $y > 0$ имеем

$$a^{x+y} = a^x \cdot a^y, \qquad \log_a(xy) = \log_a x + \log_a y, \qquad a^{\log_a x} = x.$$

Рассмотрим теперь свойства логарифма и экспоненты, где существенно, что в роли основания взято именно число e, а не какое-то другое.

На рисунке 5 криволинейная трапеция $ABCD$ содержится в прямоугольнике $ABC'D$ и содержит внутри себя прямоугольник $ABCD'$. Поскольку ширина обоих прямоугольников равна x, а высоты равны соответственно $BC' = AD = 1$ и $BC = AD' = 1/(1+x)$, то для любого $x > 0$ имеем

$$\frac{x}{1+x} = S_{ABCD'} < S_{ABCD} = \ln(1+x) < S_{ABC'D} = x.$$

Неравенства $\dfrac{x}{1+x} < \ln(1+x) < x$ аналогично можно доказать и при $x \in (-1; 0)$. Впрочем, сейчас нас интересуют положительные значения $x = 1/n$, где n — натуральное число. Из неравенств

$$\frac{1/n}{1+\frac{1}{n}} < \ln\left(1+\frac{1}{n}\right) < \frac{1}{n}$$

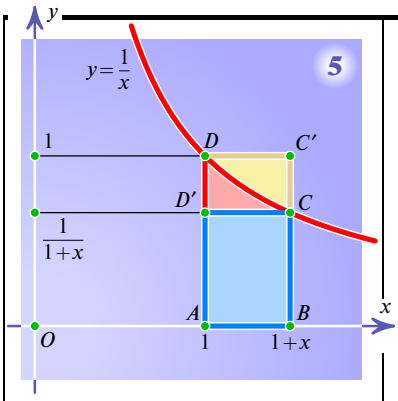
получаем: $1 < (n+1)\ln\left(1+\dfrac{1}{n}\right)$ и $n \ln\left(1+\dfrac{1}{n}\right) < 1$, от-

Логарифмическая линейка состоит из корпуса с передвигающимся в нем движком и бегунка — рамки со стеклом, на котором проведена визирная линия (выделена красным цветом). На движке и корпусе нанесены шкалы A, B, C и D, где A размечена так же, как B, а C — как D. Шкалы логарифмические в том смысле, что расстояние от начального штриха, обозначающего 1, до штриха, обозначающего число n, равно $\mu \lg n$, где μ для обычной линейки равно 12,5 см для шкал B и C, $\mu = 25$ см для шкал E и F, $\mu = 8{,}33$ см для шкалы A.

Устанавливая начало или конец шкалы E движка напротив штриха m шкалы F самой линейки, против штриха n шкалы E, на шкале F читаем произведение mn. Частное m/n находим, совмещая штрихи m и n шкал F и E и читая ответ напротив начала или конца шкалы E. Шкалы B и C дают квадраты чисел шкал F и E соответственно. Шкала A дает кубы чисел шкалы F. Равномерная шкала G дает мантиссы (дробные части) десятичных логарифмов чисел шкалы F. Шкала D помогает находить частное, представляя собой нанесенную в обратном порядке шкалу E.

На обратной стороне движка логарифмической линейки помещены шкалы тригонометрических величин: синусов S, тангенсов T, синусов (они же тангенсы) малых углов ST. Если вставить движок в корпус линейки обратной стороной, то на шкале F можно прочитать умноженные на 10 значения синусов и тангенсов углов шкал S и T соответственно и умноженные на 100 значения синусов углов шкалы ST. ■

До изобретения компьютеров ученые и инженеры для вычислений могли использовать таблицы логарифмов, синусов, тангенсов и т. п., а также логарифмическую линейку, дававшую две-три значащие цифры результата. (На практике более высокая точность нужна нечасто!) На фотографии показано умножение числа $\sqrt{2} \approx 1{,}41$ (шкала F) на $\sqrt{3} \approx 1{,}73$ (шкала E). Произведение равно 2,44 (шкала F).



Никола Шюке (ок. 1445 — ок. 1550) — французский математик и врач. Родился в Париже, работал в Лионе, крупном торговом городе. Шюке использовал термин «миллион», который ввел, видимо, путешественник Марко Поло (1254—1323) для описания богатств Востока. Шюке ввел термины «биллион», «триллион», «квадриллион» и так далее до «нониллиона». Как и Архимед в «Псаммите», Шюке сопоставил арифметическую и геометрическую прогрессии, записав одну под другой:

$$\begin{array}{cccccc} 1, & 2, & 3, & \ldots, & n \\ a, & a^2, & a^3, & \ldots, & a^n. \end{array}$$

Сумме двух членов верхней прогрессии соответствует произведение двух стоящих под ними членов нижней прогрессии, что позволяет заменять операцию умножения на сложение. В XVII в. Дж. Непер и Й. Бюрги на основе сопоставления этих прогрессий изобрели логарифмы.

В трактате «Наука о числах» (1484) Шюке впервые применил отрицательные и нулевые показатели степеней. Неизвестную величину x он обозначал как premier (первое число) или nombre linear (линейное число), что указывает на связь с геометрией. Для обозначения степеней Шюке справа от коэффициента ставил показатель степени. У него $12x$ записывается как 12^1, а $12x^2$ как 12^2, $12/x$ как 12^{1m} (m — минус). Само число 12 он записывал как 12^0. Шюке пишет: «8^3, умноженное на 7^{1m}, дает 56^2», что означает $8x^3 \cdot 7x^{-1} = 56x^2$. Математики арабского Востока не знали алгебраических символов. У них был только термин «шай» («вещь») для обозначения неизвестной величины. Создание алгебраической символики — длительный процесс, протекавший в Европе в XV—XVII вв. Шюке внес в него значительный вклад. ■

куда $\ln e < \ln\left(1+\frac{1}{n}\right)^{n+1}$, $\ln\left(1+\frac{1}{n}\right)^n < \ln e$, то есть $\left(1+\frac{1}{n}\right)^n < e < \left(1+\frac{1}{n}\right)^{n+1}$. Следовательно,

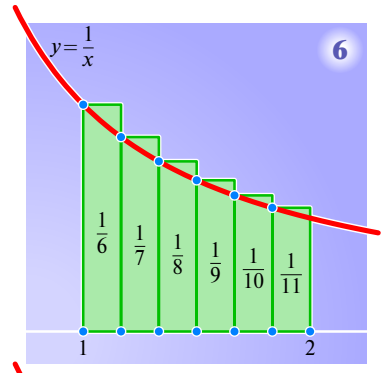
$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

(второй замечательный предел). Аналогичными рассуждениями можно получить и общую формулу для экспоненты:

$$\exp x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n. \blacksquare$$

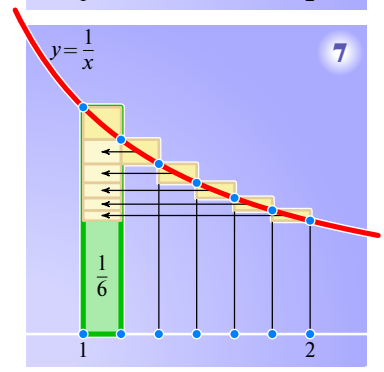
Сумма площадей зеленых прямоугольников рисунка 6, на котором отрезок [0; 1] разбит на 6 равных частей, равна

$$\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}.$$



А сумма площадей желтых прямоугольничков меньше 1/6 (рис. 7), поскольку при помощи параллельных переносов все эти прямоугольнички можно поместить в левый прямоугольник площади 1/6. Следовательно,

$$0 < \frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}-\ln 2 < \frac{1}{6}.$$



Разумеется, отрезок можно было разбить не на 6, а на n равных частей:

$$0 < \frac{1}{n}+\frac{1}{n+1}+\ldots+\frac{1}{2n-1}-\ln 2 < \frac{1}{n},$$

так что

$$\ln 2 = \lim_{n\to\infty}\sum_{k=n}^{2n-1}\frac{1}{k}. \qquad (***)$$

Вместо отрезка [1; 2] мы могли рассмотреть отрезок $[1; 1+x]$, получив тем самым формулу

$$\lim_{n\to\infty}\left(\frac{1}{n}+\frac{1}{n+x}+\frac{1}{n+2x}+\ldots+\frac{1}{n+(n-1)x}\right) = \ln(1+x).$$

Равенство (***) можно записать в довольно неожиданной форме, применив следующее алгебраическое преобразование (показанное ниже для $n=6$, хотя его можно провести для любого n):

$$\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11} =$$
$$= \left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right)+\left(\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}\right)-\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}\right) =$$
$$= \left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}\right)-2\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\frac{1}{8}+\frac{1}{10}\right) =$$
$$= 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\frac{1}{7}-\frac{1}{8}+\frac{1}{9}-\frac{1}{10}+\frac{1}{11}.$$

Таким образом,

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\frac{1}{7}-\frac{1}{8}+\frac{1}{9}-\frac{1}{10}+\frac{1}{11}-\frac{1}{12}+\ldots = \ln 2. \blacksquare$$

Степенные ряды помогают вычислять и логарифм, и экспоненту. Начнем с логарифма. Замена $t=1+y$ дает нам равенство

$$\int_1^{1+x}\frac{dt}{t}=\int_0^x\frac{dy}{1+y}.$$

Пусть $|x|<1$. Тогда величину $1/(1+y)$ можно разложить в сумму бесконечной убывающей геометрической прогресии:

$$\frac{1}{1+y}=1-y+y^2-y^3+y^4-y^5+\ldots$$

Проинтегрировав почленно (законность этой операции доказывают в курсах математического анализа при изучении равномерно сходящихся функциональных рядов), получаем формулу

$$\ln(1+x)=$$
$$=\int_0^x dy-\int_0^x y\,dy+\int_0^x y^2\,dy-\int_0^x y^3\,dy+\int_0^x y^4\,dy-\int_0^x y^5\,dy+\ldots=$$
$$=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\frac{x^5}{5}-\frac{x^6}{6}+\ldots$$

Ряд для экспоненты не менее красив. Хотя окончательная формула верна для всех x (не только вещественных, а даже и комплексных!), мы ограничимся случаем $x>0$. Обозначим $T_n(x)=\left(1+\frac{x}{n}\right)^n$. Как мы уже знаем, $\lim\limits_{n\to\infty}T_n(x)=\exp x$. Применим бином Ньютона:

$$\left(1+\frac{x}{n}\right)^n=1+n\cdot\frac{x}{n}+\frac{n(n-1)}{1\cdot2}\cdot\frac{x^2}{n^2}+\frac{n(n-1)(n-2)}{1\cdot2\cdot3}\cdot\frac{x^3}{n^3}+\ldots$$
$$\ldots+\frac{n(n-1)\ldots(n-k+1)}{k!}\cdot\frac{x^k}{n^k}+\ldots+\frac{x^n}{n^n}=$$
$$=1+x+\frac{1-\frac{1}{n}}{1\cdot2}x^2+\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)}{1\cdot2\cdot3}x^3+\ldots$$
$$\ldots+\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)}{k!}x^k+\ldots+\frac{x^n}{n^n}.$$

Числа, стоящие в правой части этого равенства в круглых скобках, все меньше единицы. Поэтому $T_n(x)<S_n(x)$, где $S_n(x)=1+x+\frac{x^2}{1\cdot2}+\ldots+\frac{x^n}{n!}$.

С другой стороны, отбрасывая в правой части разложения $T_n(x)$ по биному все слагаемые, кроме $k+1$ первых, получаем при $n>k$ неравенство

$$\left(1+\frac{x}{n}\right)^n>1+x+\left(1-\frac{1}{n}\right)\frac{x^2}{2!}+$$
$$+\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\frac{x^3}{3!}+\ldots+\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\frac{x^k}{k!}.$$

При постоянном k и при $n\to\infty$ правая часть стремится к $S_k(x)$, поскольку каждый из сомножителей в скобках стремится к единице. А поскольку $T_n(x)$ стремится к $\exp x$, то для любого натураль-

Николаус Меркатор (Mercator — латинизированная форма фамилии Kaufmann) (ок. 1620—1687) — немецкий математик, астроном и инженер. Учился и работал в Копенгагене, затем в Лондоне и Париже. Основное математическое сочинение — «Логарифмотехника» (1668) — содержит разложение функции $\ln(1+x)$ при $|x|<1$ в степенной ряд. Это второй в истории (первый — формула суммы бесконечно убывающей геометрической прогрессии) пример разложения функции в степенной ряд. ■

$e=[2;\,1,\,2,\,1,\,1,\,4,\,1,\,1,\,6,\,1,\,1,\,8,\,1,\,1,\ldots]=$

$$=2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{4+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{8+\cfrac{1}{1+\cfrac{1}{1+\ddots}}}}}}}}}}}}}}$$

$\frac{e-1}{e+1}=[0;\,2,\,6,\,10,\,14,\,18,\ldots]$

$\frac{e-1}{2}=[0;\,1,\,6,\,10,\,14,\,18,\ldots]$

$\sqrt{e}=[1;\,1,\,1,\,1,\,5,\,1,\,1,\,1,\,9,\ldots]$

Гиперболические синус и косинус — это функции $\operatorname{sh}x=(e^x-e^{-x})/2$ и $\operatorname{ch}x=(e^x+e^{-x})/2$. Поскольку

$$\frac{e^{x+y}+e^{-x-y}}{2}=\frac{e^x+e^{-x}}{2}\cdot\frac{e^y+e^{-y}}{2}+$$
$$+\frac{e^x-e^{-x}}{2}\cdot\frac{e^y-e^{-y}}{2},$$

то $\operatorname{sh}(x+y)=\operatorname{sh}x\operatorname{ch}y+\operatorname{ch}x\operatorname{sh}y$. Аналогично проверяются и формулы $\operatorname{ch}(x+y)=\operatorname{ch}x\operatorname{ch}y+\operatorname{sh}x\operatorname{sh}y$, $\operatorname{ch}^2x-\operatorname{sh}^2x=1$, $\operatorname{ch}2x=\operatorname{ch}^2x+\operatorname{sh}^2x$, $\operatorname{sh}2x=2\operatorname{sh}x\operatorname{ch}x$.

График $y=\operatorname{ch}x$ — цепная линия. Именно по линии $y=a\operatorname{ch}(x/a)$ устанавливается в равновесии гибкая и нерастяжимая тяжелая нить (цепь, провод и т. п.), подвешенная за оба конца. ■

ного k верно неравенство

$$\exp x > S_k(x).$$

Значит, последовательность $S_1(x), S_2(x), S_3(x), \ldots$ возрастающая и ограничена сверху числом $\exp x$. Следовательно, эта последовательность имеет предел, причем

$$\lim_{k\to\infty} S_k(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^k}{k!} + \ldots \leqslant \exp x.$$

Кроме того, из неравенства $T_n(x) < S_n(x)$ получаем, переходя к пределу при $n \to \infty$, что

$$\lim_{n\to\infty} S_n(x) \geqslant \exp x.$$

Следовательно,

$$\exp x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \blacksquare$$

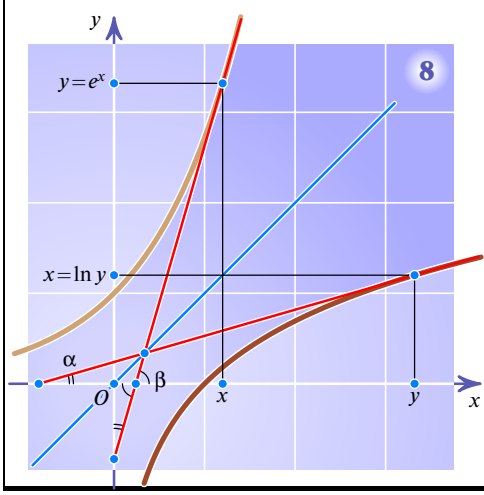
Продифференцируем экспоненту. Поскольку $(x^n)' = nx^{n-1}$ и $n! = n\cdot(n-1)!$, то $\left(\dfrac{x^n}{n!}\right)' = \dfrac{x^{n-1}}{(n-1)!}$ и, дифференцируя ряд почленно, получаем равенство $(e^x)' = e^x$. Впрочем, обоснование законности почленного дифференцирования в школьную программу не входит. Поэтому лучше воспользуемся тем, что экспонента — это функция, обратная к натуральному логарифму (рис. 8): очевидно, $\alpha + \beta = 90°$, так что $\operatorname{tg}\alpha \operatorname{tg}\beta = 1$, где $\operatorname{tg}\alpha$ — производная натурального логарифма в точке y, так что $\operatorname{tg}\alpha = 1/y$ и, значит,

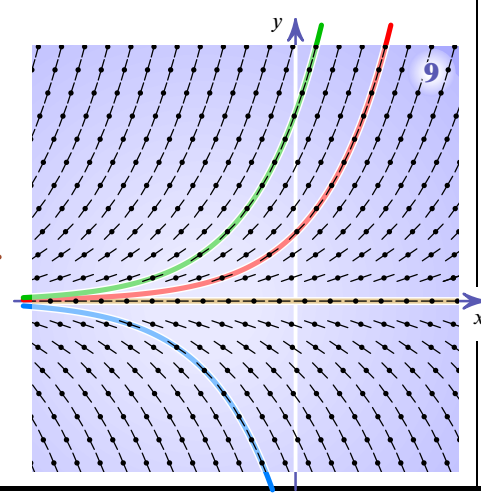
$$(e^x)' = \operatorname{tg}\beta = \frac{1}{1/y} = y = e^x.$$

Таким образом, функция $y = e^x$ — а вместе с ней и все функции вида $y = c\cdot e^x$ — удовлетворяет дифференциальному уравнению $y' = y$.

На рисунке 9 в каждой точке $(x; y)$ координатной плоскости нарисована прямая (точнее, ее маленький отрезок) с угловым коэффициентом y. Эти отрезочки являются касательными к графикам функций $y = c\cdot e^x$.

Случай $c = 0$ приводит к тождественно равной нулю функции $y(x) = 0$, являющейся одним из решений дифференциального уравнения $y' = y$. Ее график — ось абсцисс — отделяет получающиеся друг из друга параллельными переносами вдоль оси абсцисс графики вида $y = c\cdot e^x$, где $c > 0$, от графиков, для которых $c < 0$. ■





Джон Непер (1550—1617) — шотландский математик — изобретатель логарифмов. Его труды «Описание удивительной таблицы логарифмов» (1614) и «Построение удивительной таблицы логарифмов» (1619) содержат определение и свойства логарифмов, таблицы логарифмов синусов, косинусов, тангенсов, принципы вычисления таблиц, приложения логарифмов к сферической тригонометрии. Его кинематическое определение в современных терминах является определением логарифма как решения дифференциального уравнения $y' = 1/x$ с начальным условием $y(1) = 0$. Число e иногда называют неперовым числом.

Неперовы аналогии — это формулы для решения сферических треугольников а) по двум данным сторонам a, b и углу C между ними и б) по двум данным углам A, B и прилежащей к ним стороне c:

$$\operatorname{tg}\frac{A-B}{2} = \frac{\sin\dfrac{a-b}{2}}{\sin\dfrac{a+b}{2}}\operatorname{ctg}\frac{C}{2}, \qquad (1)$$

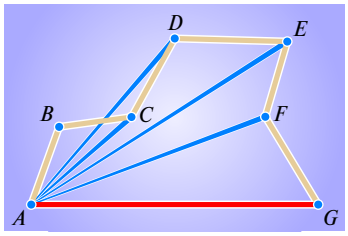
$$\operatorname{tg}\frac{A+B}{2} = \frac{\cos\dfrac{a-b}{2}}{\cos\dfrac{a+b}{2}}\operatorname{ctg}\frac{C}{2}, \qquad (2)$$

$$\operatorname{tg}\frac{a-b}{2} = \frac{\sin\dfrac{A-B}{2}}{\sin\dfrac{A+B}{2}}\operatorname{ctg}\frac{c}{2}, \qquad (3)$$

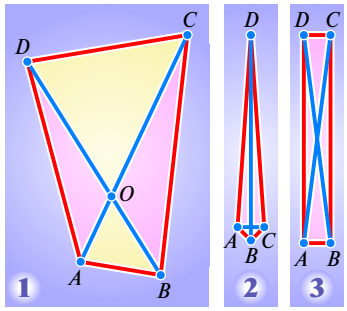
$$\operatorname{tg}\frac{a+b}{2} = \frac{\cos\dfrac{A-B}{2}}{\cos\dfrac{A+B}{2}}\operatorname{ctg}\frac{c}{2}. \qquad (4)$$

Для решения задачи а) сначала применяют формулы (1) и (2), затем (3) или (4); для задачи б) — сначала формулы (3) и (4), а затем одну из (1), (2). ■

Неравенство треугольника

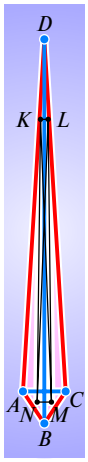


Длина любой ломаной не меньше расстояния между ее концами. Доказать это нетрудно: $AG \leqslant AF + FG \leqslant (AE + EF) + FG \leqslant (AD + DE) + EF + FG \leqslant \ldots \leqslant AB + BC + CD + DE + EF + FG$. ■



Какие значения может принимать отношение

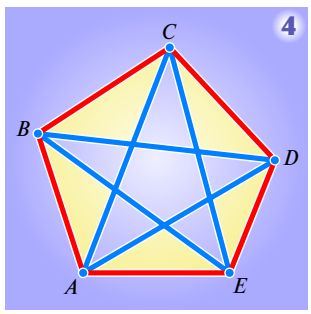
$$\frac{KM + LN}{AC + BD},$$

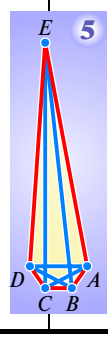


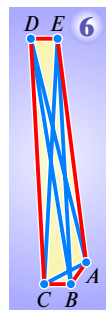
где $KLMN$ — выпуклый четырехугольник, расположенный внутри выпуклого четырехугольника $ABCD$? Очевидно,

$$KM + LN < KL + LM + MN + NK < AB + BC + CD + DA < 2(AC + BD).$$

Чтобы понять, что полученная оценка неулучшаема, достаточно рассмотреть четырехугольник $ABCD$ рисунка 2 и расположить точки K и L вблизи точки D, как показано на рисунке, а точки M и N — вблизи точек A, B, C. ■







Длина любой стороны треугольника меньше суммы длин двух других его сторон. Это неравенство позволяет решить много интересных геометрических задач.

Какие значения может принимать отношение $\dfrac{AC + BD}{AB + BC + CD + DA}$ суммы длин диагоналей выпуклого четырехугольника к его периметру? Обозначив буквой O точку пересечения диагоналей рассматриваемого четырехугольника, имеем (рис. 1):

$$AB + CD < (AO + OB) + (CO + OD) = (AO + OC) + (BO + OD) = AC + BD.$$

Аналогично, $BC + AD < AC + BD$. Следовательно, $AB + CD + BC + AD < 2(AC + BD)$. С другой стороны, сложив неравенства $AC < AB + BC$, $AC < AD + DC$, $BD < BA + AD$ и $BD < BC + CD$, получаем $2(AC + BD) < 2(AB + BC + CD + DA)$. Таким образом,

$$\frac{1}{2} < \frac{AC + BD}{AB + BC + CD + DA} < 1.$$

Можно ли заменить здесь 1/2 на какое-то большее число? Нет, нельзя. Чтобы доказать это, рассмотрим четырехугольник, вершины A, B и C которого расположены очень близко друг к другу, а вершина D — довольно далеко от них (рис. 2). Очевидно, сумма длин диагоналей этого четырехугольника близка к длине отрезка AD, а периметр мало отличается от $2AD$. (Слова «мало», «очень близко», «довольно далеко», строго говоря, бессмысленны. Следовало бы рассмотреть не один четырехугольник, а целое семейство. Предел интересующего нас отношения оказался бы равен 1/2. Но это настолько естественно, что на такую «борьбу за строгость» мы не будем тратить ваше внимание.)
Теперь рассмотрим прямоугольник $ABCD$, сторона AB которого очень короткая (рис. 3). Очевидно, сумма диагоналей такого прямоугольника почти равна его периметру. Таким образом, нельзя ни заменить 1/2 на большее число, ни 1 — на меньшее. ■

Какие значения может принимать отношение суммы длин диагоналей выпуклого пятиугольника к сумме длин его сторон? Рассмотрим выпуклый пятиугольник $ABCDE$ (рис. 4). Применив неравенство треугольника к желтым треугольникам, видим, что сумма длин диагоналей больше периметра. Далее, сложив неравенства $AC < AB + BC$, $BD < BC + CD$, $CE < CD + DE$, $DA < DE + EA$ и $EB < EA + AB$, имеем $AC + BD + CE + DA + EB < 2(AB + BC + CD + DE + EA)$. Таким образом,

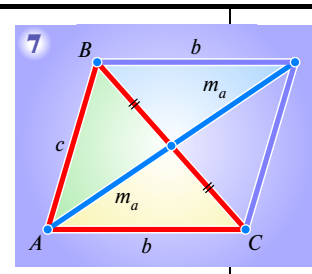
$$1 < \frac{AC + BD + CE + DA + EB}{AB + BC + CD + DE + EA} < 2.$$

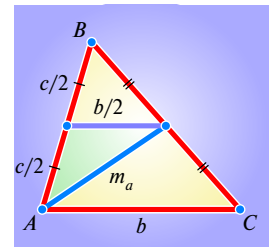
Число 1 нельзя заменить на большее, а 2 — на меньшее: для доказательства первого из этих двух утверждений достаточно расположить точки A, B, C и D вблизи друг друга, а точку E — далеко от них (рис. 5); а для неулучшаемости оценки сверху достаточно передвинуть точку D к точке E (рис. 6). ■

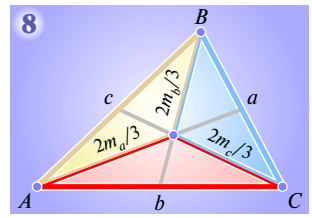
Какие значения может принимать отношение $\frac{m_a+m_b+m_c}{a+b+c}$, где m_a, m_b и m_c — длины медиан треугольника, a, b, c — длины его сторон? Продлив медиану m_a на ее длину (рис. 7), получаем неравенство $2m_a<b+c$. Аналогично, $2m_b<a+c$ и $2m_c<a+b$. Сложив эти три неравенства, получаем $2(m_a+m_b+m_c)<2(a+b+c)$. С другой стороны, поскольку медианы в точке пересечения делятся в отношении 2:1, считая от вершины (рис. 8), то $a<\frac{2}{3}m_b+\frac{2}{3}m_c$, $b<\frac{2}{3}m_a+\frac{2}{3}m_c$ и $c<\frac{2}{3}m_a+\frac{2}{3}m_b$. Сложив эти неравенства, получаем $a+b+c<\frac{4}{3}(m_a+m_b+m_c)$. Итак,

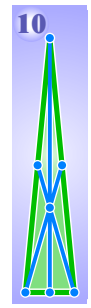
$$\frac{3}{4}<\frac{m_a+m_b+m_c}{a+b+c}<1.$$

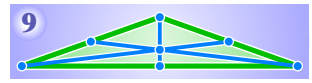
Для доказательства неулучшаемости оценок снизу и сверху достаточно рассмотреть равнобедренные треугольники, величина угла при вершине у первого из которых близка к 180° (рис. 9), а у второго — к 0° (рис. 10). ■





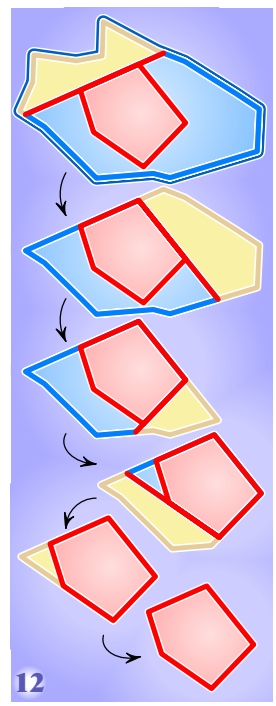






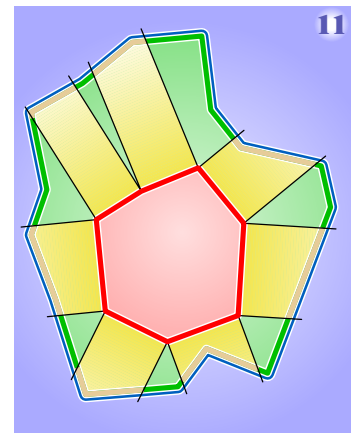
Пусть выпуклый многоугольник лежит внутри некоторого многоугольника. Обязательно ли периметр внутреннего меньше периметра внешнего? Да! Доказать это можно, восставив перпендикуляры к каждой из сторон внутреннего многоугольника (рис. 11): расстояние между параллельными прямыми не превосходит длины любой ломаной, концы которой лежат соответственно на этих двух прямых (а показанные зеленым цветом части сторон внешнего многоугольника можно даже не учитывать).

Есть и другой способ (рис. 12): «совершенствовать» внешний многоугольник, постепенно превращая его во внутренний. А именно, продолжив одну из сторон внутреннего многоугольника до пересечения со сторонами внешнего, видим, что при замене ломаной на отрезок периметр внешнего многоугольника уменьшается. Выполнив еще несколько аналогичных операций, мы в конце концов уничтожим все отличия внешнего многоугольника от внутреннего. При каждой операции периметр уменьшался и в конце, поскольку многоугольники совпали, стал равен периметру внутреннего. Следовательно, изначально периметр внешнего многоугольника был больше периметра внутреннего. ■



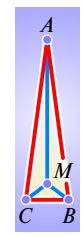
В «Началах» Евклида доказанное нами утверждение кратко сформулировано так: «Объемлющая больше объемлемой». Оба изложенных доказательства работают и в стереометрии: площадь поверхности любого выпуклого многогранника меньше площади поверхности любого многогранника, внутри которого первый расположен. Интересно, останется ли это в силе, если рассматривать не площадь поверхности, а сумму длин ребер многогранника? Если не ограничивать количество вершин внутреннего многогранника, то отрицательный

Неравенство $m_a<(b+c)/2$ можно доказать не только при помощи достраивания треугольника до параллелограмма, но и проведя среднюю линию. ■



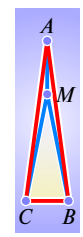
Внутри треугольника ABC отмечена точка M. Чему может быть равно отношение

$$\frac{MA+MB+MC}{AB+BC+CA}?$$



Сложив неравенства $AM+MB>AB$, $BM+MC>BC$, $CM+MA>CA$, имеем: $2(AM+BM+CM)>AB+BC+CA$. С другой стороны, так как объемлющая больше объемлемой, верны неравенства $AM+MB<AC+CB$, $BM+MC<BA+AC$, $CM+MA<AB+BA$. Сложим их: $2(AM+BM+CM)<2(AB+BC+CA)$. Таким образом,

$$\frac{1}{2}<\frac{AM+BM+CM}{AB+BC+CA}<1.$$



Для доказательства неулучшаемости полученных оценок достаточно рассмотреть треугольник ABC, вершины B и C которого расположены очень близко, а вершина A — далеко от них. Осталось расположить точку M сначала вблизи от B и C, а затем — около точки A. ■

Парадокс? Точкой, сумма расстояний от которой до вершин выпуклого четырехугольника $ABCD$ минимальна, является точка O пересечения его диагоналей. Доказать это очень легко: для любой точки M плоскости по неравенству треугольника имеем $AM+CM\geqslant AC$ и $BM+DM\geqslant BD$, откуда

$$AM+CM+BM+DM\geqslant AC+BD=\\=AO+OC+BO+OD.$$

Пусть точки A и B будут неподвижны, а точки C и D устремятся к вершине E равностороннего треугольника ABE. Точка O пересечения диагоналей тоже устремится к точке E. Мы доказали, что сумма расстояний от точки O до вершин четырехугольника минимальна. В пределе четырехугольник превращается в треугольник ABE. Значит, сумма расстояний от предельного положения точки O — минимальная из всевозможных сумм расстояний от точек плоскости до вершин треугольника ABE.

Однако сумма расстояний от центра описанной окружности до вершин треугольника *меньше* суммы $AE+BE$. Парадокс! ■

Рассмотрим на плоскости точки $O(0;0)$, $A(a;b)$ и $B(a+x;b+y)$. Теорема Пифагора позволяет записать неравенство $OB\leqslant OA+AB$ в виде

$$\sqrt{(a+x)^2+(b+y)^2}\leqslant\\\leqslant\sqrt{a^2+b^2}+\sqrt{x^2+y^2}.$$

Возведя обе части этого неравенства в квадрат, раскрыв в левой части скобки и упростив, получаем неравенство Коши—Буняковского

$$ax+by\leqslant\sqrt{(a^2+b^2)(x^2+y^2)}.$$

Если бы точки брали не на плоскости, а в пространстве, аналогично получили бы неравенство

$$ax+by+cz\leqslant\\\leqslant\sqrt{(a^2+b^2+c^2)(x^2+y^2+z^2)}.$$

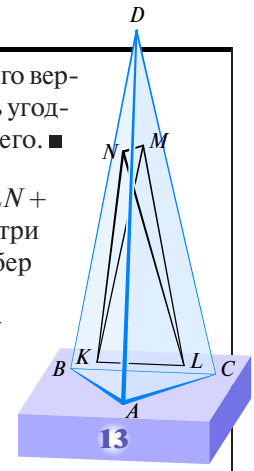
И вообще, для любого натурального n и для любых чисел $a_1, a_2, \ldots, a_n$ и $x_1, x_2, \ldots, x_n$ верно неравенство

$$\sum_{k=1}^{n}a_kx_k\leqslant\sqrt{\left(\sum_{k=1}^{n}a_k^2\right)\left(\sum_{k=1}^{n}x_k^2\right)}.$$

Оно обращается в равенство только в случае, когда вектор $(a_1;a_2;\ldots;a_n)$ сонаправлен вектору $(x_1;x_2;\ldots;x_n)$. ■

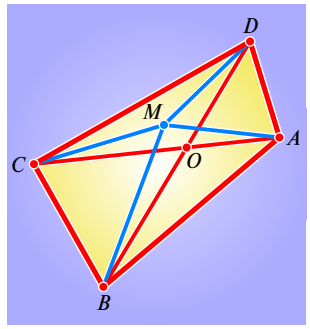
ответ очевиден: внутренний многогранник может иметь много вершин, расположенных так, что сумма длин его ребер в сколь угодно большое число раз превзойдет сумму длин ребер внешнего. ■

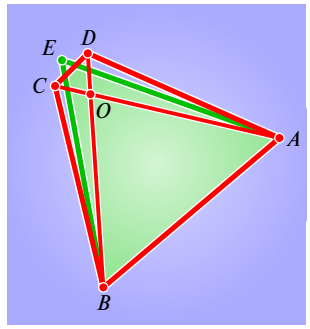
Может ли сумма $P_{KLMN}=KL+KM+KN+LM+LN+MN$ длин ребер тетраэдра $KLMN$, расположенного внутри тетраэдра $ABCD$, оказаться больше суммы P_{ABCD} длин ребер тетраэдра $ABCD$?



13

Расположим точки A, B, C, K и L близко друг к другу (рис. 13), а точки D, M и N — далеко от них и тоже близко одну к другой. Отношение P_{KLMN}/P_{ABCD} может оказаться не только больше 1, но даже сколь угодно близко к 4/3.



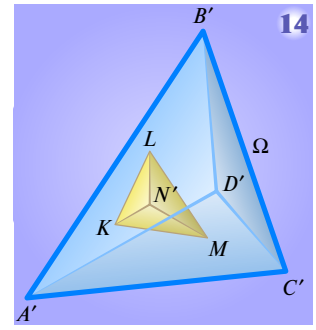


Может ли отношение P_{KLMN}/P_{ABCD} быть больше или равно 4/3? Ответить на этот вопрос смогли лишь четверо из 53 школьников, решавших эту задачу в 1982 г. в Одессе на Всесоюзной математической олимпиаде. Все решения оказались очень красивыми. ■

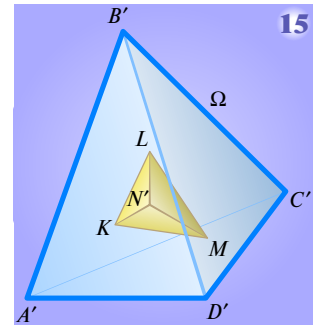
И. Зиганшин и Г. Перельман предложили геометрическое решение. Пусть для определенности KLM — грань тетраэдра $KLMN$, имеющая наибольший периметр среди всех граней этого тетраэдра. Спроецируем тетраэдр $ABCD$ на плоскость KLM. Проекции вершин тетраэдра обозначим, соответственно, через A', B', C' и D'. Проекция — это либо треугольник (рис. 14), либо выпуклый четырехугольник (рис. 15). Ломаную, ограничивающую проекцию тетраэдра $ABCD$, обозначим буквой Ω (на рисунках 14, 15 она показана синим цветом). Треугольник KLM лежит внутри Ω. Периметр ломаной Ω, треугольника KLM и сумму длин проекций ребер тетраэдра $ABCD$ обозначим, соответственно, через P_Ω, P_{KLM} и $P_{A'B'C'D'}$. Коротко решение можно записать следующим образом:

$$P_{KLMN}\leqslant 2P_{KLM}\leqslant 2P_\Omega<2\cdot\frac{2}{3}P_{A'B'C'D'}\leqslant\frac{4}{3}P_{ABCD}.$$

(Докажите каждое из этих неравенств!) ■



14



15

Другие два решения выходят за рамки школьной программы. Обычно более других ценят олимпиадные задачи, для решения которых знание высших разделов математики не особенно помогает и математик-профессионал не имеет преимущества перед школьниками. Однако не все задачи таковы. И вот здесь как раз тот случай, когда знание некоторых фактов «нешкольной» математики позволяет получить очень простое и естественное доказательство. ■

Решение А. Савкина основано на следующей лемме.

Лемма. ***Пусть в пространстве заданы две системы отрезков, причем если рассмотреть любую прямую и спроецировать все отрезки на нее, то сумма длин проекций отрезков первой системы будет не больше суммы длин проекций отрезков второй системы. Тогда сумма длин отрезков первой системы не превосходит суммы длин отрезков второй системы.***

Мы не умеем доказывать эту лемму, не используя интегрирование. А при помощи интегралов — легко: средняя длина проекции отрезка на всевозможные прямые пространства пропорциональна длине отрезка и, разумеется, не зависит от его направления.

На какую бы прямую l мы ни спроецировали ребра тетраэдра $ABCD$, получим некоторый отрезок Δ_1, никакая внутренняя точка которого не покрыта менее чем трижды, то есть проекциями менее чем трех рассматриваемых отрезков. (Доказать это проще всего, заметив, что любая перпендикулярная отрезку Δ_1 в любой его внутренней точке плоскость пересекает тетраэдр по треугольнику или четырехугольнику, а поэтому пересекает три или четыре ребра.)

Спроецировав на ту же прямую l ребра тетраэдра $KLMN$, получим некоторый отрезок Δ_2, причем $\Delta_2 \subseteq \Delta_1$ (поскольку тетраэдр $KLMN$ лежит внутри тетраэдра $ABCD$) и каждая точка отрезка Δ_2 покрыта проекциями не более чем четырежды (поскольку сечение тетраэдра плоскостью — это точка, отрезок, треугольник или четырехугольник). ■

А. Спивак зафиксировал тетраэдр $ABCD$ и поставил следующую задачу на максимум: *расположить точки K, L, M и N так, чтобы ни одна из них не вышла за пределы тетраэдра $ABCD$ и периметр P_{KLMN} был бы наибольшим возможным*.

Прежде всего возникает вопрос: имеет ли задача на максимум решение? (Например, невозможно найти максимальное значение функции $\dfrac{-1}{1+x^2}$. Сколь угодно близким к нулю отрицательным числом значение может быть, а нулем — не может.) Оказывается, оно существует в силу теоремы Вейерштрасса о наибольшем значении непрерывной функции, определенной на компакте (в данном случае роль этого компакта играет двенадцатимерное множество — произведение четырех тетраэдров).

Рассмотрим ту четверку точек K, L, M, N, для которой величина P_{KLMN} достигает наибольшего значения. Докажем, что любая из этих четырех точек обязана совпадать с одной из вершин тетраэдра $ABCD$. Действительно, пусть точка N не совпадает ни с A, ни с B, ни с C, ни с D. Рассмотрим отрезок N_1N_2 с серединой в точке N, все точки которого являются точками тетраэдра $ABCD$. Поскольку $KN \leqslant (KN_1 + KN_2)/2$, $LN \leqslant (LN_1 + LN_2)/2$, $MN \leqslant (MN_1 + MN_2)/2$, причем хотя бы одно из этих неравенств строгое, то

$$KN + LN + MN < \frac{1}{2}(KN_1 + KN_2 + LN_1 + LN_2 + MN_1 + MN_2),$$

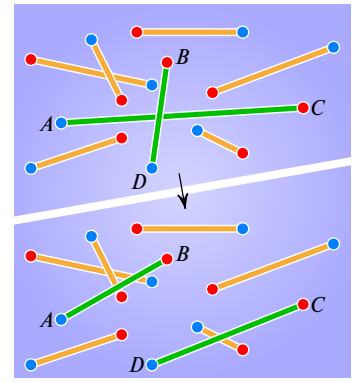
откуда следует, что $P_{KLMN} < (P_{KLMN_1} + P_{KLMN_2})/2$, а это неравенство противоречит неравенствам $P_{KLMN} \geqslant P_{KLMN_1}$ и $P_{KLMN} \geqslant P_{KLMN_2}$.

Осталось рассмотреть случаи, когда каждая из точек K, L, M, N совпадает с одной из точек A, B, C, D. Сделайте это самостоятельно!

Решение закончено, но скажем еще несколько слов. В последнем решении нашел отражение важный принцип: максимальное значение выпуклой вниз функции на многограннике достигается в вершине многогранника. Особенно важен этот принцип в линейном программировании — разделе математики, имеющем многочисленные приложения. ■

Докажем, что n красных точек и n синих точек плоскости, никакие три из которых не лежат на одной прямой, можно соединить n непересекающимися отрезками, концы каждого из которых разного цвета (красный и синий).

Доказательство. Разобьем точки на n пар, каждую из которых образуют две точки разного цвета. Проведем соответствующие n отрезков. Если никакие два из них не пересекутся, то мы получили искомую систему отрезков. В противном случае пусть точки A и D — синие, а B и C — красные, причем отрезки AC и BD пересекаются. Заменим эти два пересекающихся отрезка на (непересекающиеся!) отрезки AB и CD.

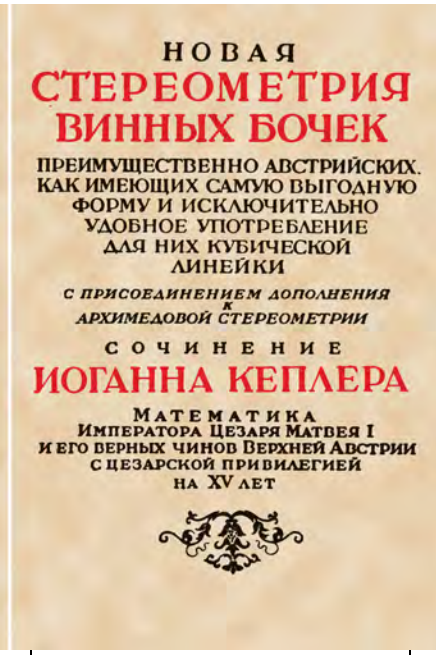


Так будем делать и дальше, пока можно: берем любую пару пересекающихся отрезков, у каждого из которых концы разного цвета, и заменяем на пару непересекающихся отрезков, концы которых тоже разного цвета. Остановится ли этот процесс? (На рисунке мы избавились от точки пересечения отрезков AC и BD, но приобрели вместо нее 3 новые точки пересечения!)

Да, остановится! Чтобы это доказать, рассмотрим не количество точек пересечения (как мы видели, оно может не только уменьшаться, но и увеличиваться), а сумму длин всех n отрезков. Поскольку сумма $AC + BD$ длин диагоналей любого выпуклого четырехугольника $ABCD$ больше суммы $AB + CD$ длин двух его противоположных сторон, каждая замена уменьшает рассматриваемую сумму длин. Бесконечно много таких уменьшений быть не может: $2n$ точек можно соединить между собой n отрезками лишь конечным числом способов. ■

Кеплер и винные бочки


NOVA
STEREOMETRIA
DOLIORVM VINARIORVM, IN PRImis Auſtriaci, figuræ omnium aptiſsimæ;
ET
USUS IN EO VIRGÆ CUBIcæ compendioſiſſimus & plane ſingularis.
Acceſſit
STEREOMETRIÆ ARCHIMEdeæ Supplementum.
Authore
Ioanne Kepplero, Imp. Cæſ Matthiæ I
cjuſq, ſidd. Ordd. Auſtriæ ſupia Anaſum Mathematico.
ANNO M. DC. XV.
LINCII
Excudebat JOANNES PLANCVS, ſumptibus Authoris.


НОВАЯ
СТЕРЕОМЕТРИЯ ВИННЫХ БОЧЕК
ПРЕИМУЩЕСТВЕННО АВСТРИЙСКИХ, КАК ИМЕЮЩИХ САМУЮ ВЫГОДНУЮ ФОРМУ И ИСКЛЮЧИТЕЛЬНО УДОБНОЕ УПОТРЕБЛЕНИЕ ДЛЯ НИХ КУБИЧЕСКОЙ ЛИНЕЙКИ
С ПРИСОЕДИНЕНИЕМ ДОПОЛНЕНИЯ К АРХИМЕДОВОЙ СТЕРЕОМЕТРИИ
СОЧИНЕНИЕ
ИОГАННА КЕПЛЕРА
МАТЕМАТИКА ИМПЕРАТОРА ЦЕЗАРЯ МАТВЕЯ I И ЕГО ВЕРНЫХ ЧИНОВ ВЕРХНЕЙ АВСТРИИ С ЦЕЗАРСКОЙ ПРИВИЛЕГИЕЙ НА XV ЛЕТ

Одна из предтеч математического анализа — книга «Новая стереометрия винных бочек. . .» И. Кеплера (1571—1630). Мы расскажем, как он разгадал секрет формы винных бочек, решая задачи на максимум и минимум.

«Я небеса измерял; ныне тени земли измеряю. Дух на небе мой жил; здесь же тень тела лежит». Мы не знаем, была ли написана на могильной плите Кеплера эта сочиненная им самим эпитафия. Неизвестно и то, так ли все было, как он рассказал в предисловии к «Новой стереометрии. . .». Но история столь хороша, что ее с удовольствием пересказывают, приукрашивая или сокращая, многие историки: «В ноябре прошлого года ко мне на дом было принесено и поставлено несколько бочек вина, а через четыре дня пришел продавец с мерной линейкой».

Кеплер удивился, как с помощью одного измерения (рис. 1) можно узнать объем, — ведь бочки бывают разной формы! — и «счел для себя подходящим взять новый предмет математических занятий и исследовать геометрические законы такого удобного измерения и выяснить его основания, если таковые имеются».

Напомним сказанные по другому поводу слова А. Н. Колмогорова: «В каждый данный момент существует лишь тонкий слой между «тривиальным» и недоступным. В этом слое и делаются математические открытия. Заказная прикладная задача поэтому в большинстве случаев или решается тривиально, или вообще не решается. . .» Поставленная Кеплером задача попала в этот «тонкий слой»: через три дня он «очинил перо для отделки и записи доказательства, готового в уме».

Начал он с того, что «бочка имеет форму пузатого цилиндра, или, говоря точнее, бочка представляется как бы разделенной на два усеченных конуса, вершины которых, направленные в противоположные стороны, отсечены деревянными днищами бочки, а основание общее, разделяющее конусы и образующие наибольший круг, опоясывающий бочки». Другими словами, Кеплер предложил такую математическую модель винной бочки: две одинаковые половины, полученные вращением равнобоких трапеций $ABEF$ и $BCDE$ вокруг их общей оси симметрии (рис. 2). Продавец измеряет расстояние $d = AE$. Сначала рассмотрим случай $r = R$ (австрийская бочка). В конце статьи мы рассмотрим случай $r < R$ (рейнская бочка).■

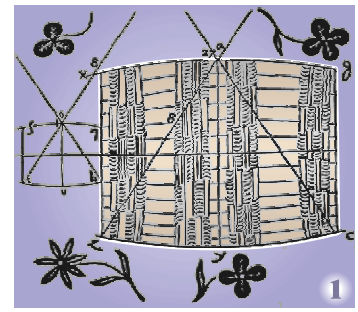
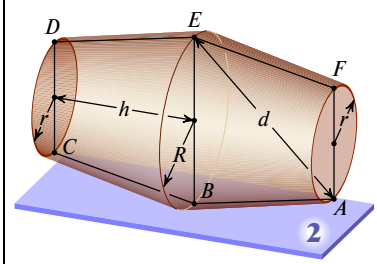

1

Рисунок из 1-го издания «Новой стереометрии. . .» (1615).



Как австрийские бочары умудрялись делать бочки абсолютно одинаковой формы (но разного размера)? Можно вообразить, что существовал какой-то австрийский стандарт. Однако кто, как и зачем его изобрел?

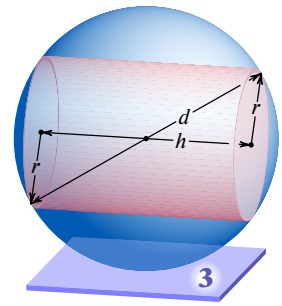
Линейка размечена по «кубическому закону»: при увеличении линейных размеров в k раз объем увеличивается в k^3 раз, ибо в k раз увеличиваются и длина, и ширина, и высота. (Однажды в редакцию «Докладов Академии наук» один биолог принес статью, где опытным путем доказывал обнаруженный им закон: объем инфузории пропорционален третьей степени ее длины. Больших трудов стоило убедить его, что закон верен не только для инфузорий, а для проверки не нужны эксперименты на мухах, воробьях, крысах, обезьянах, медведях, слонах и китах.)■

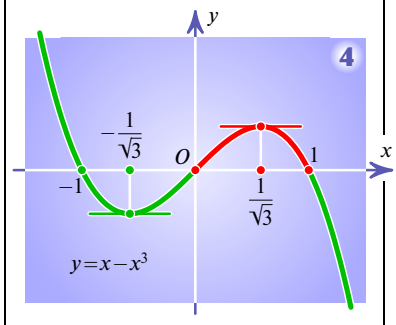
Кеплер нашел ответ: «Непузатые цилиндрические бочки более удлиненной или более укороченной формы, чем австрийские, вместительны менее последних». Точнее, из всех вписанных в шар с диаметром $EA=d$ цилиндров наибольший объем имеет тот, который описан вокруг куба, то есть у которого $h=r\sqrt{2}$.

Докажем это утверждение. По теореме Пифагора $d^2=h^2+(2r)^2$ (рис. 3). Объем цилиндра равен произведению площади основания на высоту:

$$V=\pi r^2h=\pi\frac{d^2-h^2}{4}h=\frac{\pi}{4}d^3(1-x^2)x,$$

где использовано обозначение $x=h/d$. Таким образом, мы должны найти максимальное значение функции $y=x-x^3$ на отрезке $[0;\,1]$. Ее график (даже на большей области определения) изображен на рисунке 4. ■





Проще всего найти точку максимума тому, кто знает, что для ее нахождения достаточно выбрать наибольшее из значений в тех точках, где производная равна нулю или не существует, а также в концах отрезка:

$$y'=1-3x^2,$$

производная на отрезке $[0;\,1]$ равна нулю лишь при $x=1/\sqrt{3}$; на концах функция равна нулю. Поэтому наибольшее значение функция принимает при $x=1/\sqrt{3}$ (при этом, заметьте, $h=d/\sqrt{3}$ и $r=\sqrt{(d^2-h^2)/4}=d/\sqrt{6}$, так что $h/r=\sqrt{2}$), максимальное значение объема цилиндрической бочки равно $2V=\dfrac{\pi}{3\sqrt{3}}d^3$. ■

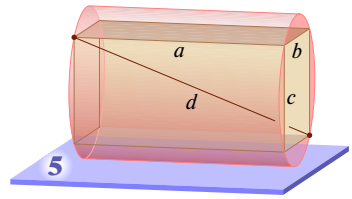
Можно обойтись и без производных. В любой цилиндр можно вписать прямоугольный параллелепипед (рис. 5), основание которого — квадрат со стороной длины $r\sqrt{2}$. Объем такого параллелепипеда равен $2r^2h$, а объем цилиндра — πr^2h. Таким образом, отношение объемов цилиндра и параллелепипеда равно $\pi/2$ и не зависит ни от r, ни от h. Значит, вместо цилиндра можно искать (вписанный в сферу диаметра d) прямоугольный параллелепипед максимального объема!

Кеплер писал: «Из всех прямоугольных параллелепипедов с квадратными основаниями, вписанных в одну и ту же сферу, куб имеет наибольший объем». Кеплер не пользовался алгеброй, поэтому его рассуждение растянулось на несколько страниц. А мы применим неравенство о среднем арифметическом и среднем геометрическом

$$abc\leqslant\left(\frac{a+b+c}{3}\right)^3$$

и неравенство о среднем арифметическом и среднем квадратичном

$$\frac{a+b+c}{3}\leqslant\sqrt{\frac{a^2+b^2+c^2}{3}}.$$



(Последнее легко доказать, возведя обе части в квадрат и сведя дело к неравенству $(a-b)^2+(b-c)^2+(c-a)^2\geqslant 0$.)

Если a, b, c — длины ребер прямоугольного параллелепипеда (заметьте — мы даже не требуем, чтобы основание было квадратом), вписанного в шар диаметра d, то $d^2=a^2+b^2+c^2$ и

$$abc\leqslant\left(\frac{a+b+c}{3}\right)^3\leqslant\left(\sqrt{\frac{a^2+b^2+c^2}{3}}\right)^3=\left(\frac{d}{\sqrt{3}}\right)^3.$$

Величина $(d/\sqrt{3})^3$ — объем куба, вписанного в сферу диаметра d. ■

Найдем $\max\limits_{0\leqslant x\leqslant 1}(x-x^3)$, не используя производных. Пусть $x=\dfrac{1}{\sqrt{3}}+t$. Тогда $t\in\left[-\dfrac{1}{\sqrt{3}};\,\dfrac{2}{\sqrt{3}}\right]$ и $y=\dfrac{1}{\sqrt{3}}+t-\left(\dfrac{1}{\sqrt{3}}+t\right)^3=\dfrac{2}{3\sqrt{3}}-t^2\sqrt{3}-t^3$. При $t>0$ имеем $y<\dfrac{2}{3\sqrt{3}}$. При $t\in\left[-\dfrac{1}{\sqrt{3}};\,0\right)$, очевидно,

$$-t^2\sqrt{3}-t^3=-t^2(\sqrt{3}+t)<0,$$

так что неравенство $y<\dfrac{2}{3\sqrt{3}}$ выполнено и в этом случае. Как видите, доказательство не очень сложное, но весьма искусственное — непонятно, как мы догадались до замены переменной. (А догадались — приравняв нулю производную!) ■

Не менее искусственный способ — воспользоваться неравенством о среднем арифметическом и среднем геометрическом. Как известно, если a, b, c — неотрицательные числа, то $\sqrt[3]{abc}\leqslant(a+b+c)/3$. Запишем это неравенство в виде

$$abc\leqslant(a+b+c)^3/27$$

и положим $a=2x$, $b=(1-x)\times(\sqrt{3}+1)$ и $c=(1+x)(\sqrt{3}-1)$. Тогда $a+b+c=2\sqrt{3}$ и неравенство принимает вид

$$2x(1-x)(\sqrt{3}+1)(1+x)\times(\sqrt{3}-1)\leqslant 8/(3\sqrt{3}),$$

откуда $x-x^3\leqslant 2/(3\sqrt{3})$. ■

В книге «Вы, конечно, шутите, мистер Фейнман!» американский физик Ричард Фейнман (1918—1988) рассказывает такую историю: «Когда я был в Массачусетском технологическом институте, я любил подшучивать над людьми. Однажды в кабинете черчения кто-то поднял лекало (кусок пластмассы для рисования гладких кривых — забавно выглядящая штука в завитушках) и спросил: «Имеют ли кривые на этих штуках какую-либо формулу?» Я немного подумал и ответил: «Несомненно. Это такие специальные кривые. Дай-ка я покажу тебе. — Я взял свое лекало и начал его медленно поворачивать. — Лекало сделано так, что, независимо от того, как ты его повернешь, в наинизшей точке каждой кривой касательная горизонтальна».

Все парни в кабинете начали крутить свои лекала под различными углами, подставляя карандаш к нижней точке и по-всякому прилаживая его. Они обнаружили, что касательная горизонтальна. Все были крайне возбуждены от этого открытия, хотя уже много прошли по математике и даже «выучили», что производная (касательная) в минимуме (нижней точке) для любой кривой равна нулю (горизонтальна). Они не совмещали эти факты. Они не знали даже того, что они уже «знали».

Я плохо представляю, что происходит с людьми: они не учатся путем понимания. Они учатся каким-то другим способом — путем механического запоминания или как-то иначе. Их знания так хрупки!» ■

Реактивные сопла первых реактивных самолетов, установленные на крыльях вблизи фюзеляжа, представляли опасность для хвостового оперения. Конструкторы, знавшие обсуждаемую формулу, повернули двигатели на небольшой угол α. Хвостовое оперение было спасено (отклонение реактивной струи пропорционально α), а результирующая сила тяги практически не изменилась (потеря $\approx \alpha^2/2$, где α — угол в радианах; для угла в 3° теряется всего порядка 1/800 мощности). ■

Слава австрийским бочарам!

Ответ у Кеплера получился тот же, что и у нас. И он не смог сдержать своего восхищения австрийскими бочарами, которые «как бы по здравому и геометрическому смыслу при построении бочки соблюдают правило, чтобы за радиус днища брать треть длины клепок. Именно, при таком устройстве цилиндр, мысленно построенный между двумя днищами, будет самым вместительным, хотя бы при постройке бочки от точных правил несколько и отступили, потому что по обе стороны от места наибольшего значения убывание вначале нечувствительно. . . Бочары за длину клепки берут. . . полуторную величину диаметра основания, что дает приближение к наивместительнейшей фигуре, потому что клепки изгибаются и с обеих сторон выходят за обручи, которые охватывают и сжимают днища, так что излишек в длине против полуторного диаметра основания и приходится на эти выступающие оконечности. . .»

Скептик скажет, что для обоснования приближения $\sqrt{2} \approx 3/2$ столь глубокомысленные рассуждения излишни: довольно того, что запомнить число 3/2 легче, чем 1,41421356237. . . Но поймите и Кеплера: он желал вызвать у читателя чувство уважения к науке (для этого рассматривал даже бочки, полученные при вращении дуг эллипсов, гипербол, парабол, . . .). И на жизнь он зарабатывал не математическими и даже не астрономическими занятиями, а выпуском календарей, содержащих предсказания всякого рода, и составлением гороскопов, то есть предсказаний судьбы на основании расположения светил на небе. «Лучше издавать альманахи с предсказаниями, — писал Кеплер, — чем просить милостыню». «Астрология — дочь астрономии, хотя и незаконная, и разве не естественно, чтобы дочь кормила свою мать, которая иначе могла бы умереть с голоду?»

Проницательный читатель уже обратил внимание не на восторги, а на суть дела: «по обе стороны от места наибольшего значения убывание вначале нечувствительно». Что хотел сказать этим Кеплер? ■

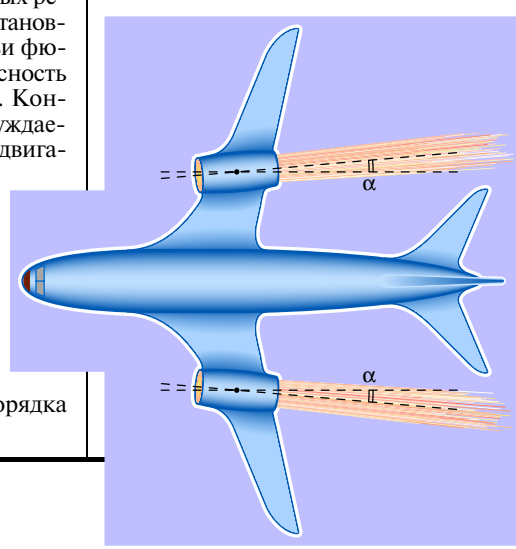
«Убывание нечувствительно».

Тому, кто не знаком с математическим анализом, легко объяснить эту мысль, если вспомнить формулу

$$y = \frac{2}{3\sqrt{3}} - t^2\sqrt{3} - t^3.$$

При малых t величины $t^2\sqrt{3}$ и t^3 очень малы. Если, например, $t = 0{,}001$, то $t^2 = 0{,}000001$ и $t^3 = 0{,}000000001$. Эти две величины малы даже по сравнению с (тоже маленькой) величиной t.

А тот, кто изучал анализ, может обойтись и без замены переменной. По определению, для функции $f(x)$, дифференцируемой в точке x_0, имеет место равенство



$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

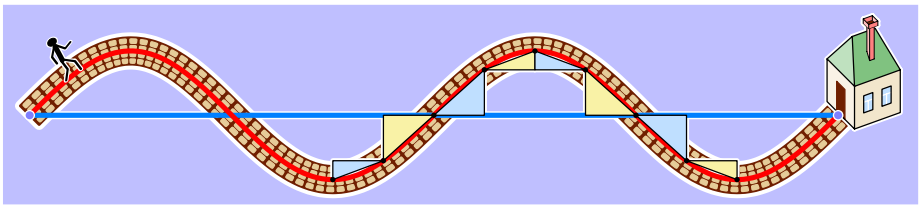
По определению предела имеем

$$\frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) + \alpha(h),$$

где $\alpha(h)$ — бесконечно малая величина, то есть $\alpha(h) \to 0$ при $h \to 0$. Значит,

$$f(x_0 + h) = f(x_0) + f'(x_0)\,h + \alpha(h) \cdot h. \quad (*)$$

(Последняя формула настолько важна, что в университетских курсах анализа именно ее берут за определение производной.)



Так что же мы видим? Если $f'(x_0)\neq 0$, то отклонение $f(x_0+h)$ от $f(x_0)$ при малых h почти пропорционально h (величина $\alpha(h)\cdot h$ при $h\to 0$ стремится к нулю быстрее, чем $f'(x_0)h$). Если же $f'(x_0)=0$, то отклонение равно $\alpha(h)\cdot h$ и стремится к нулю быстрее, чем h. Это значит, что вблизи точки, где производная равна 0, малое изменение аргумента функции сказывается на изменении функции существенно слабее (Кеплер пишет: «нечувствительно»), чем вблизи точки, где производная отлична от 0. Именно по этой причине небольшие отклонения австрийских бочек от стандарта практически не влияли на точность измерения их объема кубической линейкой.

«Нечувствительность изменения» функции вблизи точки, где производная обращается в ноль, является чрезвычайно важным математическим и даже общенаучным фактом. Например, кто из нас не клял нескончаемо длинные зимние ночи? И ведь как они начинаются, так несколько месяцев темень и темень. Ноябрь, декабрь, январь, февраль — короткий день и длинная ночь! Не то обидно, что 22 декабря день очень короток. Обидно, что так обстоит дело не только 22 декабря. Казалось бы, продолжительность дня должна меняться, а она почти не меняется! На рисунке 6 на графике продолжительности дня красным (синим) цветом выделены участки, где она отличается от максимальной (минимальной) менее чем на 30 минут.

Но в свой срок приходит весна (производная становится положительной, сказал бы прозаик), продолжительность дня довольно быстро возрастает, и, к нашему удовольствию, долго, весь май и все лето, дни длинные, а ночи короткие. «Изменения нечувствительны» не только в точке минимума, но и в точке максимума!

При малых x (где величина x измеряется в радианах) $\sin x\approx x$, точнее говоря,

$$\lim_{x\to 0}\frac{\sin x}{x}=1 \qquad \text{(первый замечательный предел).}$$

Производная функции $\cos x$ в точке $x=0$ равна нулю, поэтому $\cos x$ при малых отклонениях x от нуля «меняется нечувствительно». Точнее,

$$\cos x=\cos(2\cdot x/2)=1-2\sin^2(x/2)\approx 1-x^2/2.$$

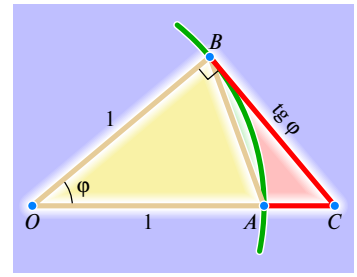
Рассмотрев прямоугольный треугольник с катетами a, b и гипотенузой c, где b мало по сравнению с c, и обозначив буквой x величину угла, противоположного катету b, получим:

$$x\approx\sin x=\frac{b}{c}, \qquad a=c\cos x\approx c\left(1-\frac{b^2}{2c^2}\right),$$

откуда $c-a\approx\dfrac{b^2}{2c}$. Как вы помните, катет b мал по сравнению с гипотенузой. Поэтому последняя формула означает, что катет a отличается от гипотенузы c на величину гораздо меньшую, чем b («нечувствительно», как сказал бы Кеплер).

Таким образом, больший катет вытянутого прямоугольного треугольника практически столь же длинен, как и гипотенуза, а разность длин примерно равна $\dfrac{b^2}{2c}$ и тем самым гораздо меньше длины меньшего катета b. ■

Предположим, вы возвращаетесь домой по синусоиде. Насколько ваш путь длиннее, чем если бы вы шли прямо? Первое впечатление (что вдвое), конечно, преувеличивает длину. Все же кажется, что путь по синусоиде длиннее раза в полтора. На самом деле примерно на 20%. Причина в том, что бóльшая часть синусоиды слабо наклонена к оси, поэтому соответствующие гипотенузы практически не длиннее катетов. ■



Площадь круга радиуса 1 равна π. Поэтому площадь сектора OAB равна $\pi\cdot\dfrac{\varphi}{2\pi}=\dfrac{\varphi}{2}$. Площади треугольников OAB и OBC равны, соответственно,

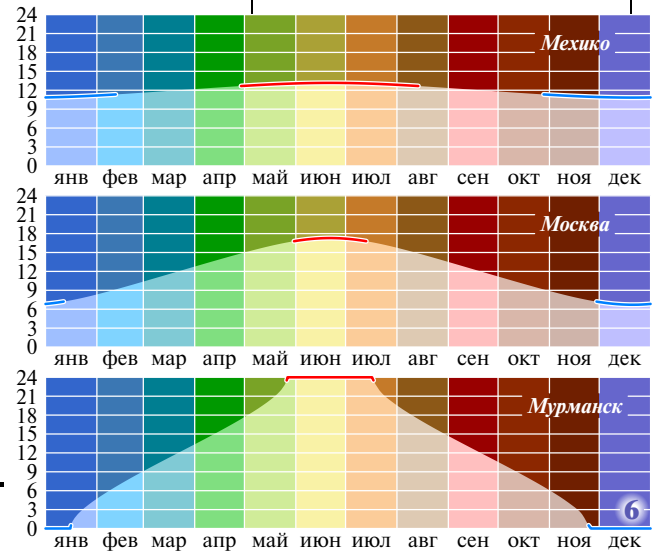
$$\frac{OA\cdot OB\sin\varphi}{2}=\frac{\sin\varphi}{2}$$

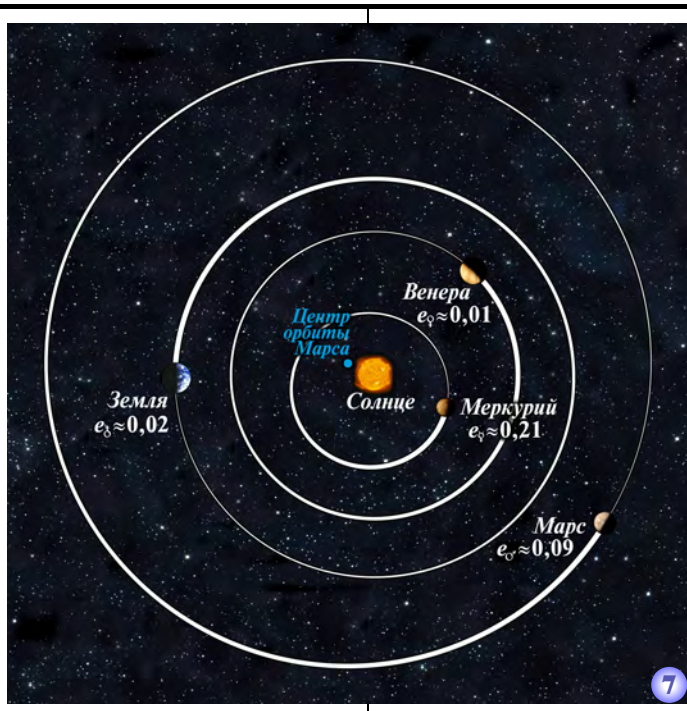
и $\dfrac{OB\cdot BC}{2}=\dfrac{1\cdot\operatorname{tg}\varphi}{2}=\dfrac{\sin\varphi}{2\cos\varphi}$.

Следовательно,

$$\frac{\sin\varphi}{2}<\frac{\varphi}{2}<\frac{\sin\varphi}{2\cos\varphi},$$

откуда $1>\dfrac{\sin\varphi}{\varphi}>\cos\varphi$. Поскольку $\cos\varphi$ стремится к 1 при $\varphi\to 0$, то и $\dfrac{\sin\varphi}{\varphi}\to 1$ при $\varphi\to 0$. ■





Учитель Кеплера Тихо Браге в обсерватории «Ураниборг» в течение 20 лет скрупулезно измерял положения планет Солнечной системы. После смерти учителя Кеплер взялся за математическую обработку результатов этих наблюдений и обнаружил, что орбита Марса — эллипс. . .
Чтобы понять рассуждения Кеплера, нам потребуется формула

$$b/a = \sqrt{1-e^2} \approx 1 - e^2/2,$$

где a и b — большая и малая полуоси эллипса, а e — эксцентриситет.
Поясним эту формулу. Если вы знакомы с производными, то знаете, что производная функции $f(x) = \sqrt{x}$ равна $\dfrac{1}{2\sqrt{x}}$, и можете подставить $x_0 = 1$ и $h = -e^2$ в формулу $(*)$. А если не знакомы, то попросту возведите в квадрат:

$$\left(1 - \frac{e^2}{2}\right)^2 = 1 - e^2 + \frac{e^4}{4} \approx 1 - e^2.$$

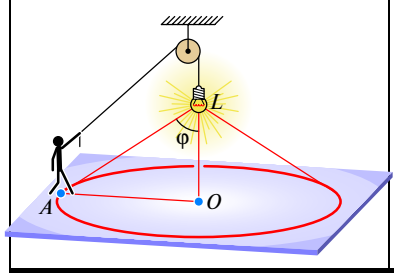
Таким образом, эллипс с малым эксцентриситетом практически неотличим от окружности. Например, если $e = 0{,}1$, то малая ось короче большой всего на 1/200. Для эллипса с длиной большой оси 1 м малая ось короче большой всего на полсантиметра, так что на глаз отличие такого эллипса от окружности вообще не заметно. Фокусы же смещены от центра на 5 см, что очень заметно.
Сначала Кеплер думал, что орбита Марса — окружность. Однако Солнце оказалось не в центре, а сдвинутым примерно на 1/10 часть радиуса (рис. 7). Но Кеплер не остановился на этом (уже замечательном) результате — потому что он знал теорию конических сечений. Кеплер знал, что эллипс с маленьким эксцентриситетом очень похож на окружность, и проверил, как ведет себя то небольшое отклонение орбиты от окружности, которое еще оставалось.
Орбита Марса оказалась слегка сплюснутой в направлении, перпендикулярном диаметру, на котором лежит Солнце — примерно на полпроцента, то есть на $e^2/2$. Так Кеплер пришел к мысли об эллиптических орбитах планет.
Как видите, разница между «чувствительным» и «нечувствительным» изменением важна не только для австрийской бочкотары, но и для астрономии! ■

Ошибка, которую не сделал Кеплер. «Рукопись этой книги, — сказано в «Новой стереометрии. . .», — пролежала шестнадцать месяцев у аугсбургского книготорговца, и вопреки данному мне обещанию не была напечатана.
. . . С этого времени у меня явилось намерение напечатать эту книжку самому, несмотря на большой недостаток средств. При этом мне представилась возможность не только исправить ее, но и продвинуть в отношении объема сравнительно с первоначально написанной. Не скрою, что на эти размышления было затрачено некоторое время, уделенное от прочих занятий, но я не жалею об этой потере, так как никоим образом невозможно, чтобы пожал плод бессмертия труд, не посеявший некоторого времени».
Кеплер не скрывает от читателя свои методы и даже заблуждения. Например, доказав, что из всех вписанных в данный круг прямоугольников наибольшую площадь имеет квадрат (словами Кеплера: «Осевые сечения прямых цилиндров, имеющих равные диагонали, имеют неравные площади, за исключением

Пусть электрическая лампа может передвигаться (например, на блоке) по вертикальной прямой OL. На каком расстоянии от горизонтальной плоскости ее надо поместить, чтобы в точке A плоскости получить наибольшую освещенность?
Освещенность E пропорциональна косинусу угла падения лучей ($\angle ALO = \varphi$) и обратно пропорциональна квадрату расстояния AL, то есть $E = \dfrac{C\cos\varphi}{AL^2}$, где величина C зависит лишь от силы света лампы. Учитывая, что $AL = \dfrac{AO}{\sin\varphi}$, имеем:

$$E = \frac{C}{AO^2}\sin^2\varphi\cos\varphi.$$

Найдите максимум этой функции самостоятельно или прочитайте рассказ о конусе максимального объема на следующей странице. **Ответ:** на расстоянии $LO = \dfrac{AO}{\sqrt{2}}$. ■



того случая, когда у них одинаковые или обратные отношения диаметра основания к высоте; наибольшая площадь среди них у того, который получается от сечения цилиндра с высотой, равной диаметру основания»), он признается: «Не хочу скрыть ошибки, в которые меня первоначально ввергло поверхностное рассмотрение этой теоремы, ибо это напоминание предупредит читателя, чтобы он остерегался подобных же (заблуждений. — *Прим. ред.*)». И дальше подробно объясняет, в чем дело, доказывая странную для нынешнего читателя теорему III: «Отношения объемов прямых цилиндров, осевые сечения которых имеют одну и ту же диагональ, не аналогичны (не равны. — *Прим. ред.*) отношениям площадей осевых сечений, и при наибольшей площади сечения объем не наибольший».

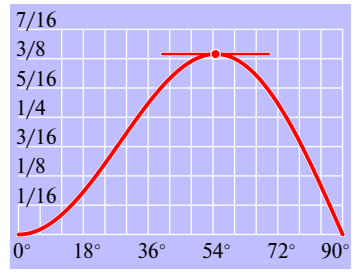
В другом месте читаем: «Кто, избавившись от заблуждения приписывать наибольший объем тому из цилиндров с данной диагональю, у которого площадь осевого сечения наибольшая, и узнав, что самым вместительным будет цилиндр, в котором отношение диаметра основания к высоте равно $\sqrt{2}$, оказался бы столь проницательным и осторожным, чтобы тотчас не предположить того же самого и об объеме усеченного конуса? Я же это подумал и держался такого мнения последние полтора года и даже дошел до того, что, опираясь на это основание, считал все рейнские бочки, без различия их пузатости, в отношении емкости ниже австрийских. . . Поэтому я отношу к пользе, полученной от настоящего печатания, что при подготовке издания геометрия потеребила меня за ухо. . .» ■

Рейнская бочка. Объем усеченного конуса, радиусы оснований которого r и R, а высота (то есть расстояние между плоскостями оснований) h, равен

$$\frac{\pi h}{3}(r^2+rR+R^2).$$

По теореме Пифагора $d^2=(r+R)^2+h^2$. Поэтому объем усеченного конуса равен

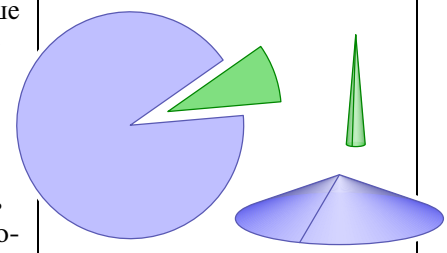
$$\frac{\pi h}{3}(d^2-h^2-Rr).$$



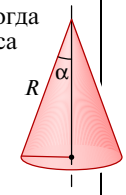
Если d и h зафиксировать (и если, разумеется, $0<h<d$), то будет фиксирована сумма $r+R$. Но произведение rR при этом фиксировано не будет! Объем будет наибольшим, когда наименьшим будет произведение rR. Меньше нуля произведение стать не может. А нулем — может. Наибольший объем равен $\frac{\pi h}{3}(d^2-h^2)$ и достигается для конуса (при этом $r=0$ и $R=\sqrt{d^2-h^2}$).

Осталось выяснить, при каком h (разумеется, $0<h<d$) величина $h(d^2-h^2)$ максимальна. Почти такую же задачу мы уже решали, изучая австрийскую бочку. Ответ — при $h=\frac{d}{\sqrt{3}}$. Объем бочки при этом равен $\frac{4\pi}{9\sqrt{3}}d^3$. Это и есть наибольший возможный объем бочки — очень похожей, по словам Кеплера, на те, что приходили с Рейна! (Напомним, что объем австрийской бочки равен $\frac{\pi}{3\sqrt{3}}d^3$, что составляет лишь 75% объема рейнской бочки. Кеплер комментировал это так: «Предполагая, что бочки представляют собой просто удвоенные усеченные конусы, заключаем, что продолговатые умеренно пузатые вместительнее цилиндрических того же поперечного размера, и никогда не делают бочек столь чудовищно пузатых, чтобы они оказались снова менее вместительны, чем цилиндрические того же продольного размера».) ■

Конус максимального объема. Рассмотрим круг радиуса R. Вырежем из него сектор и свернем из оставшейся части «фантик» — конус. Ясно, что если вырезанный сектор очень мал, то



высота конуса и его объем тоже малы. Но и вырезать почти все, оставив лишь маленький сектор, тоже неразумно, если мы хотим получить конус сколь-нибудь значительного объема: узкий конус хотя и будет иметь высоту, мало отличающуюся от R, площадь его основания будет мала. Чтобы найти конус максимального объема, обозначим через α половину величины угла осевого сечения конуса. Тогда радиус основания конуса равен $R\sin\alpha$, высота равна $R\cos\alpha$, а объем равен



$$\frac{1}{3}\cdot\pi(R\sin\alpha)^2\cdot R\cos\alpha=$$

$$=\frac{\pi}{3}R^3\sin^2\alpha\cos\alpha.$$

График функции $\sin^2\alpha\cos\alpha$ изображен на рисунке. (Заметьте: замена $x=\cos\alpha$ приводит к функции $(1-x^2)x$, где $0<x<1$.) Производная равна

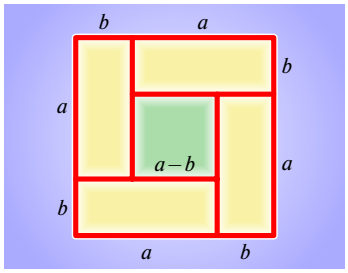
$$2\sin\alpha\cos^2\alpha-\sin^3\alpha$$

и обращается в нуль (при $0<\alpha<\pi$), когда $\alpha=\operatorname{arctg}\sqrt{2}$. Объем конуса при этом равен $\frac{2\pi R^3}{9\sqrt{3}}$. ■

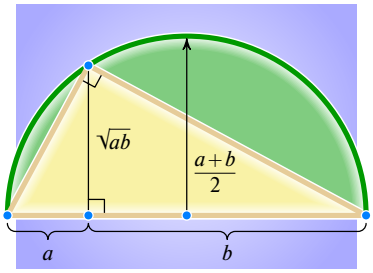
О молекуле метана (CH_4) в учебнике химии сказано: «Валентные связи атома углерода направлены к вершинам тетраэдра, угол между ними составляет 109°28′». Это и есть угол $2\operatorname{arctg}\sqrt{2}$. Совпадение не случайно: такой угол образуется, если из центра правильного тетраэдра провести лучи в две его вершины. Другими словами, ребро правильного тетраэдра в $2\sqrt{2}$ раз больше расстояния от центра этого тетраэдра до середины ребра. ■

Квадрат со стороной $a+b$, где $a>b>0$, можно разрезать на четыре прямоугольника размером $a\times b$ и квадрат со стороной $a-b$. Ничего удивительного в этом нет:

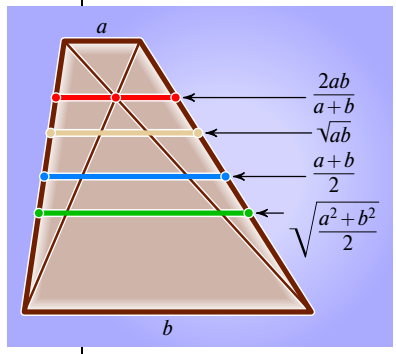
$$(a+b)^2=4ab+(a-b)^2.\ \blacksquare$$



Высота $\sqrt{ab}$ прямоугольного треугольника, разбивающая гипотенузу на отрезки длин a и b, не превосходит радиуса описанной окружности $\dfrac{a+b}{2}$. $\blacksquare$

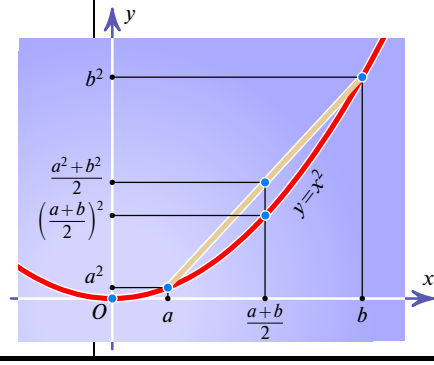


Если параллельно основаниям a и b трапеции проведены отрезки длин $\dfrac{2ab}{a+b}$, $\sqrt{ab}$, $\dfrac{a+b}{2}$ и $\sqrt{\dfrac{a^2+b^2}{2}}$, то первый из них проходит через точку пересечения диагоналей, второй разбивает трапецию на две подобные между собой трапеции, третий — средняя линия, а четвертый разбивает трапецию на две трапеции равной площади. $\blacksquare$



Середина отрезка, соединяющего точки $(a;a^2)$ и $(b;b^2)$, — точка $\left(\dfrac{a+b}{2};\dfrac{a^2+b^2}{2}\right)$ — расположена выше параболы:

$$\frac{a^2+b^2}{2}\geqslant\left(\frac{a+b}{2}\right)^2.\ \blacksquare$$



Неравенства о средних

Классические неравенства: о средних арифметическом и геометрическом, арифметическом и квадратичном, Коши—Буняковского, Минковского, Гёльдера, Йенсена и Мюрхеда — доказаны в этой статье, некоторые — несколькими способами.

Квадрат любого числа неотрицателен: $(a-b)^2\geqslant 0$. Следовательно, $a^2+b^2\geqslant 2ab$. Прибавив к обеим частям $2ab$, получаем $a^2+2ab+b^2\geqslant 4ab$, то есть $(a+b)^2\geqslant 4ab$. Если числа a и b неотрицательны, неравенства можно записать в виде

$$\sqrt{\frac{a^2+b^2}{2}}\geqslant\sqrt{ab}\qquad\text{и}\qquad\frac{a+b}{2}\geqslant\sqrt{ab}.$$

Величины $\dfrac{a+b}{2}$, $\sqrt{\dfrac{a^2+b^2}{2}}$ и $\sqrt{ab}$ называют средним арифметическим, средним квадратичным и средним геометрическим чисел a и b. $\blacksquare$

Выясним, что больше — среднее квадратичное или среднее арифметическое. Возведем $\sqrt{(a^2+b^2)/2}$ и $(a+b)/2$ в квадрат:

$$\frac{a^2+b^2}{2}\vee\frac{a^2+2ab+b^2}{4},$$

то есть $2a^2+2b^2\vee a^2+2ab+b^2$. (Мы использовали знак $\vee$, поскольку пока не знаем, какой должен быть знак неравенства — $\leqslant$ или $\geqslant$.) Перенесем все из правой части в левую:

$$a^2-2ab+b^2\vee 0.$$

Дело свелось к неравенству $(a-b)^2\geqslant 0$. Значит, $\vee$ — это знак $\geqslant$; среднее квадратичное не меньше среднего арифметического. $\blacksquare$

Подставив в неравенство о среднем арифметическом и среднем геометрическом $1/a$ и $1/b$ вместо a и b, получаем неравенство

$$\frac{\frac{1}{a}+\frac{1}{b}}{2}\geqslant\sqrt{\frac{1}{a}\cdot\frac{1}{b}},$$

то есть $\sqrt{ab}\geqslant\dfrac{2ab}{a+b}$, где $\dfrac{1}{\frac{\frac{1}{a}+\frac{1}{b}}{2}}=\dfrac{2ab}{a+b}$ — среднее гармоническое чисел a и b. Итак, для любых положительных чисел a и b верны неравенства

$$\frac{2ab}{a+b}\leqslant\sqrt{ab}\leqslant\frac{a+b}{2}\leqslant\sqrt{\frac{a^2+b^2}{2}}.\ \blacksquare$$

Сравним среднее арифметическое $\frac{a+b+c}{3}$ и среднее квадратичное $\sqrt{\frac{a^2+b^2+c^2}{3}}$ трех чисел:

$$\left(\frac{a+b+c}{3}\right)^2 \vee \frac{a^2+b^2+c^2}{3},$$
$$(a+b+c)^2 \vee 3(a^2+b^2+c^2),$$
$$a^2+b^2+c^2+2ab+2ac+2bc \vee 3a^2+3b^2+3c^2,$$
$$0 \vee 2a^2+2b^2+2c^2-2ab-2bc-2ac,$$

то есть $0 \vee (a-b)^2+(b-c)^2+(c-a)^2$. Очевидно, $\vee$ — это $\leqslant$. Аналогично для четырех чисел, раскрыв скобки в неравенствах $(a-b)^2 \geqslant 0$, $(a-c)^2 \geqslant 0$, $(a-d)^2 \geqslant 0$, $(b-c)^2 \geqslant 0$, $(b-d)^2 \geqslant 0$, $(c-d)^2 \geqslant 0$ и сложив их все, получаем

$$a^2-2ab+b^2+a^2-2ac+c^2+a^2-2ad+d^2+b^2-2bc+c^2+b^2-2bd+d^2+c^2-2cd+d^2 \geqslant 0,$$

откуда $4a^2+4b^2+4c^2+4d^2 \geqslant a^2+b^2+c^2+d^2+2ab+2ac+2ad+2bc+2bd+2cd$, то есть

$$4(a^2+b^2+c^2+d^2) \geqslant (a+b+c+d)^2,$$

что и требовалось доказать: $\sqrt{(a^2+b^2+c^2+d^2)/4} \geqslant (a+b+c+d)/4$.

Теорема 1. ***Среднее квадратичное любых n чисел $x_1, x_2, \ldots, x_n$ не меньше их среднего арифметического, то есть $\sqrt{\frac{x_1^2+x_2^2+\ldots+x_n^2}{n}} \geqslant \frac{x_1+x_2+\ldots+x_n}{n}$. Равенство достигается тогда и только тогда, когда $x_1=x_2=\ldots=x_n$.***

Доказательство. Выпишем все неравенства вида $(x_m-x_k)^2 \geqslant 0$, где $1 \leqslant k < m \leqslant n$, и сложим их. Получим неравенство

$$(n-1)(x_1^2+x_2^2+\ldots+x_n^2)-2\sum_{1\leqslant k<m\leqslant n} x_k x_m \geqslant 0,$$

которое можно привести к виду

$$n(x_1^2+x_2^2+\ldots+x_n^2) \geqslant x_1^2+x_2^2+\ldots+x_n^2+2\sum_{1\leqslant k<m\leqslant n} x_k x_m,$$

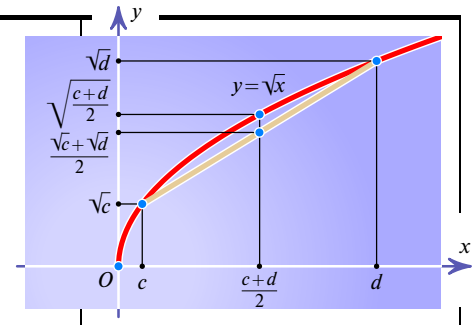
то есть $n(x_1^2+x_2^2+\ldots+x_n^2) \geqslant (x_1+x_2+\ldots+x_n)^2$, что и требовалось доказать. ■

Неравенство Коши—Буняковского — обобщение неравенства о среднем арифметическом и среднем квадратичном.

Теорема 2. ***Для любых вещественных чисел $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ верно неравенство $a_1b_1+a_2b_2+\ldots+a_nb_n \leqslant \sqrt{a_1^2+a_2^2+\ldots+a_n^2}\sqrt{b_1^2+b_2^2+\ldots+b_n^2}$.***

Доказательство. I способ. Пусть $A=\sqrt{a_1^2+a_2^2+\ldots+a_n^2}$ и $B=\sqrt{b_1^2+b_2^2+\ldots+b_n^2}$. Если $A=0$ (или $B=0$), то $a_1=a_2=\ldots=a_n=0$ (соответственно, $b_1=b_2=\ldots=b_n=0$), так что неравенство выполнено. Если же $A>0$ и $B>0$, то

$$\frac{a_1b_1+a_2b_2+\ldots+a_nb_n}{AB} = \frac{a_1}{A}\cdot\frac{b_1}{B}+\frac{a_2}{A}\cdot\frac{b_2}{B}+\ldots+\frac{a_n}{A}\cdot\frac{b_n}{B} \leqslant$$
$$\leqslant \frac{1}{2}\left(\frac{a_1^2}{A^2}+\frac{b_1^2}{B^2}+\frac{a_2^2}{A^2}+\frac{b_2^2}{B^2}+\ldots+\frac{a_n^2}{A^2}+\frac{b_n^2}{B^2}\right) =$$
$$= \frac{1}{2}\left(\frac{a_1^2+a_2^2+\ldots+a_n^2}{A^2}+\frac{b_1^2+b_2^2+\ldots+b_n^2}{B^2}\right) = \frac{1+1}{2} = 1.$$



Середина отрезка, соединяющего точки $(c; \sqrt{c})$ и $(d; \sqrt{d})$ графика функции $y=\sqrt{x}$, лежит под графиком:

$$\frac{\sqrt{c}+\sqrt{d}}{2} \leqslant \sqrt{\frac{c+d}{2}}.$$

Получили новое неравенство? Нет, подстановка $c=a^2$ и $d=b^2$ сводит его к неравенству о среднем квадратичном и среднем арифметическом. ■

Неравенство Чебышёва для монотонных последовательностей. Если $0 \leqslant a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n$ и $0 \leqslant b_1 \leqslant b_2 \leqslant \ldots \leqslant b_n$, то

$$(a_1+a_2+\ldots+a_n)(b_1+b_2+\ldots+b_n) \leqslant$$
$$\leqslant n(a_1b_1+a_2b_2+\ldots+a_nb_n).$$

Доказать его проще всего, проверив тождество

$$n\sum_{1\leqslant k\leqslant n} a_kb_k - \sum_{1\leqslant k\leqslant n} a_k \sum_{1\leqslant k\leqslant n} b_k =$$
$$= \sum_{1\leqslant k<m\leqslant n} (a_m-a_k)(b_m-b_k). \blacksquare$$

Виктор Яковлевич Буняковский (1804—1889) математическое образование получил в Париже. Преподавал математику и механику в Первом кадетском корпусе, Морском корпусе, Институте путей сообщения и Петербургском университете. С 1864 г. — вице-президент Петербургской академии наук. Большое значение для установления русской терминологии имел его «Лексикон чистой и прикладной математики» (1839). Буняковский писал и учебники для средней школы (1844, 1849). Его «Основания математической теории вероятностей» (1846) содержат не только теорию, но и ее приложения к страхованию и демографии. Остроумно и систематически критиковал попытки доказать пятый постулат Евклида, но труды Н. И. Лобачевского по достоинству не оценил. ■

Среднее арифметическое любых трех неотрицательных чисел a, b, c не меньше их среднего геометрического:

$$\frac{a+b+c}{3} \geqslant \sqrt[3]{abc}.$$

Равенство достигается тогда и только тогда, когда $a=b=c$.

Доказательство. Обозначим $a=x^3$, $b=y^3$ и $c=z^3$. Неравенство приобретает вид

$$\frac{x^3+y^3+z^3}{3} \geqslant xyz,$$

то есть

$$x^3+y^3+z^3-3xyz \geqslant 0.$$

Воспользуемся разложением на множители

$$x^3+y^3+z^3-3xyz=(x+y+z)\times \\ \times(x^2+y^2+z^2-xy-xz-yz).$$

Первый множитель — сумма неотрицательных чисел, а неотрицательность второго множителя следует из неравенства Коши—Буняковского при $a_1=x$, $a_2=y$, $a_3=z$, $b_1=y$, $b_2=z$ и $b_3=x$. ■

Пусть числа $y_1, y_2, \ldots, y_n$ положительные и $y_1y_2\ldots y_n=1$. Докажем по индукции неравенство $y_1+y_2+\ldots+y_n \geqslant n$. (Вывести из него неравенство

$$(x_1+x_2+\ldots+x_n) \geqslant n\sqrt[n]{x_1x_2\ldots x_n}$$

легко, взяв $y_k=\dfrac{x_k}{\sqrt[n]{x_1x_2\ldots x_n}}$ при $k=1, 2, \ldots, n$.)

База — $n=1, y_1=1$ — тривиальна. Предположим, что утверждение верно для некоторого натурального числа n, и докажем его для $n+1$. Среди положительных чисел, произведение которых равно 1, обязательно есть хотя бы одно число, не превосходящее 1, и хотя бы одно число, не меньшее 1. Пусть, для определенности, $y_n \leqslant 1$ и $y_{n+1} \geqslant 1$. Тогда

$$(1-y_n)(y_{n+1}-1) \geqslant 0,$$

то есть $y_{n+1}+y_n \geqslant y_ny_{n+1}+1$. Следовательно,

$$y_1+y_2+\ldots+y_{n-1}+y_n+y_{n+1} \geqslant \\ \geqslant y_1+y_2+\ldots+y_{n-1}+y_ny_{n+1}+1,$$

что не меньше $n+1$ в силу предположения индукции, примененного к числам $y_1, y_2, \ldots, y_{n-1}$, y_ny_{n+1}. ■

II способ. Для любого числа t сумма $\sum\limits_{k=1}^{n}(a_kt-b_k)^2$ неотрицательна. Следовательно, дискриминант квадратичного трехчлена $t^2\sum\limits_{k=1}^{n}a_k^2-2t\sum\limits_{k=1}^{n}a_kb_k+\sum\limits_{k=1}^{n}b_k^2$ неположителен:

$$4\left(\sum_{k=1}^{n}a_kb_k\right)^2-4\sum_{k=1}^{n}a_k^2\sum_{k=1}^{n}b_k^2 \leqslant 0,$$

что и требовалось доказать.

III способ — тождество Лагранжа:

$$\left(\sum_{k=1}^{n}a_k^2\right)\left(\sum_{k=1}^{n}b_k^2\right)=\left(\sum_{k=1}^{n}a_kb_k\right)^2+\sum_{1\leqslant k<m\leqslant n}(a_kb_m-a_mb_k)^2. \blacksquare$$

О. Л. Коши при помощи индукции из неравенства $\dfrac{a+b}{2} \geqslant \sqrt{ab}$ вывел неравенство о среднем арифметическом и среднем геометрическом для n чисел.

Теорема 3. $\dfrac{a_1+a_2+\ldots+a_n}{n} \geqslant \sqrt[n]{a_1a_2\ldots a_n}$ *для любых неотрицательных чисел* $a_1, a_2, \ldots, a_n$.

Доказательство состоит из двух этапов. Сначала из неравенства n чисел выведем неравенство для $2n$ чисел; затем (уже без индукции, при помощи алгебраической подстановки) из неравенства для n чисел выведем неравенство для m чисел, где $m<n$. Таким образом, сначала поднимаясь по степеням двойки, а затем спускаясь к интересующему нас числу, получаем требуемое. Первый этап продемонстрируем для $n=2$:

$$\frac{a+b+c+d}{4}=\frac{\dfrac{a+b}{2}+\dfrac{c+d}{2}}{2} \geqslant \sqrt{\frac{a+b}{2}\cdot\frac{c+d}{2}} \geqslant \sqrt{\sqrt{ab}\sqrt{cd}}=\sqrt[4]{abcd},$$

а также для $n=4$:

$$\frac{a+b+c+d+e+f+g+h}{8}=\frac{\dfrac{a+b+c+d}{4}+\dfrac{e+f+g+h}{4}}{2} \geqslant \\ \geqslant \sqrt{\frac{a+b+c+d}{4}\cdot\frac{e+f+g+h}{4}} \geqslant \sqrt{\sqrt[4]{abcd}\sqrt[4]{efgh}}=\sqrt[8]{abcdefgh}.$$

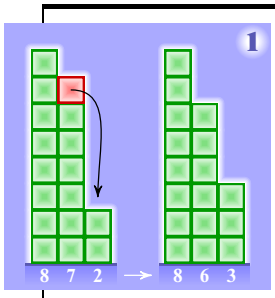
Теперь продемонстрируем второй этап на примере спуска от 4 к 3. А именно, рассмотрим любые неотрицательные числа a, b, c и неравенство о средних для чисел a, b, c и $(a+b+c)/3$. Имеем:

$$\frac{a+b+c+\dfrac{a+b+c}{3}}{4} \geqslant \sqrt[4]{abc\frac{a+b+c}{3}}.$$

Левая часть, как легко проверить, равна $(a+b+c)/3$. Возведем обе части неравенства в четвертую степень:

$$\left(\frac{a+b+c}{3}\right)^4 \geqslant abc\frac{a+b+c}{3}.$$

Осталось разделить на $(a+b+c)/3$, и неравенство для трех чисел выведено! (В случае $a=b=c=0$, отметим для придирчивого читателя, среднее арифметическое и среднее геометрическое равны 0.) В общем случае дело обстоит таким же образом (убедитесь, спустившись, например, от 8 к 5). ■



Неравенства Мюрхеда. На рисунке 1 изображены две диаграммы Юнга, которым мы сопоставим следующие симметрические функции от трех переменных:

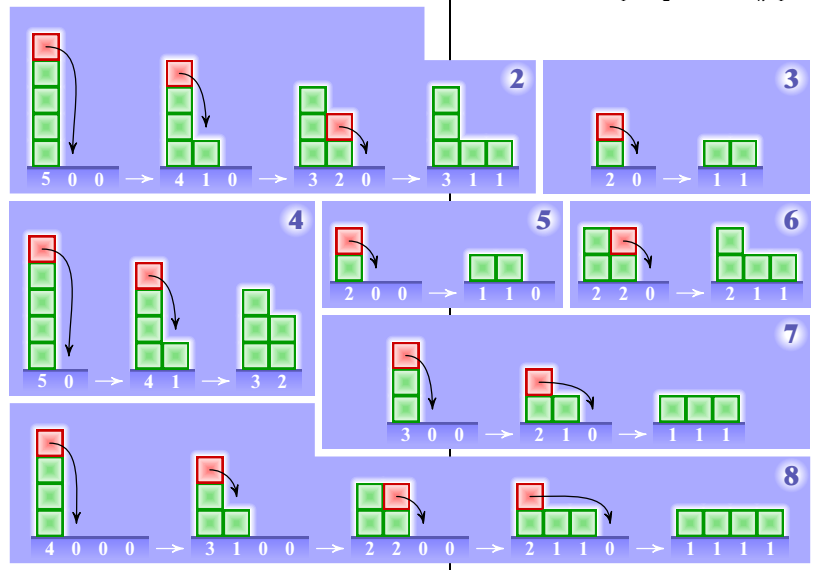
$$f(x;y;z)=x^8y^7z^2+x^8z^7y^2+y^8x^7z^2+y^8z^7x^2+z^8x^7y^2+z^8y^7x^2,$$
$$g(x;y;z)=x^8y^6z^3+x^8z^6y^3+y^8x^6z^3+y^8z^6x^3+z^8x^6y^3+z^8y^6x^3.$$

Очевидно, разность $f(x;y;z)-g(x;y;z)$ равна

$$x^8y^2z^2(y^5+z^5-y^4z-z^4y)+y^8x^2z^2(x^5+z^5-x^4z-z^4x)+$$
$$+z^8x^2y^2(x^5+y^5-x^4y-y^4x)=$$
$$=x^8y^2z^2(y^4-z^4)(y-z)+y^8x^2z^2(x^4-z^4)(x-z)+z^8x^2y^2(x^4-y^4)(x-y)\geqslant 0.$$

Так обстоит дело и в общем случае: если при перемещении клетки диаграммы Юнга мы вновь получаем диаграмму Юнга (то есть если эта клетка становится ближе к горизонтальной прямой), то значение симметрической функции не увеличивается. Не увеличивается оно и при нескольких таких преобразованиях. На рисунке 2 показаны преобразования диаграмм Юнга, соответствующие неравенствам

$$2(x^5+y^5+z^5)\geqslant$$
$$\geqslant x^4y+x^4z+y^4x+y^4z+z^4x+z^4y\geqslant$$
$$\geqslant x^3y^2+x^3z^2+y^3x^2+y^3z^2+z^3x^2+z^3y^2\geqslant$$
$$\geqslant 2(x^3yz+y^3xz+z^3xy).$$



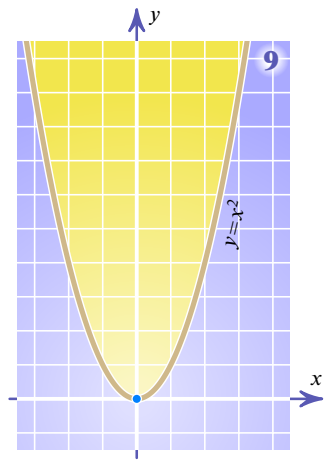
На рисунках 3—8 показаны преобразования диаграмм Юнга, приводящие к неравенствам

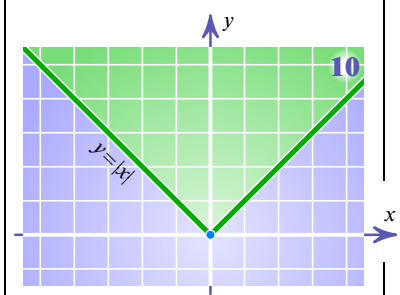
$$x^2+y^2\geqslant 2xy,$$
$$x^5+y^5\geqslant x^3y^2+x^2y^3,$$
$$x^2+y^2+z^2\geqslant xy+xz+yz,$$
$$x^2y^2+x^2z^2+y^2z^2\geqslant x^2yz+y^2xz+z^2xy,$$
$$x^3+y^3+z^3\geqslant 3xyz,$$
$$x^4+y^4+z^4+t^4\geqslant 4xyzt.$$

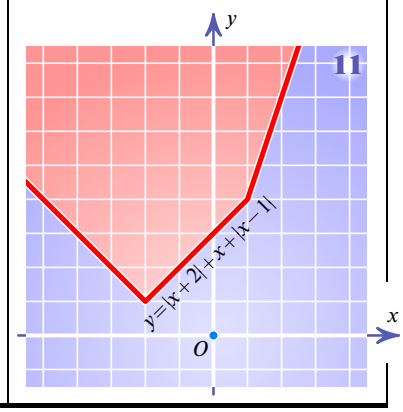
Обратите особое внимание на последние два неравенства. Диаграмме, состоящей из одного столбца высотой в n клеток, соответствует функция $(n-1)!\,(x_1^n+x_2^n+\ldots+x_n^n)$, а строке из n клеток — $n!\,x_1x_2\ldots x_n$. Следовательно,

$$x_1^n+x_2^n+\ldots+x_n^n\geqslant nx_1x_2\ldots x_n.$$

Замена $x_k^n=y_k$, где $k=1,2,\ldots,n$, показывает, что это — одна из форм записи неравенства о среднем арифметическом и среднем геометрическом. ■







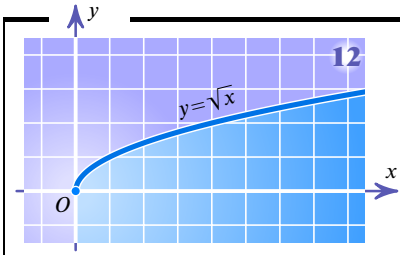
Выпуклые функции. На рисунках 9—11 изображены надграфики функций x^2, $|x|$ и $|x+2|+x+|x-1|$. Все они — выпуклые множества: каждая их хорда лежит не ниже графика функции. (Множество называют выпуклым, если вместе с любыми двумя своими точками оно содержит все точки соединяющего эти точки отрезка.) Функцию f называют выпуклой вниз, если ее надграфик — выпуклое множество.

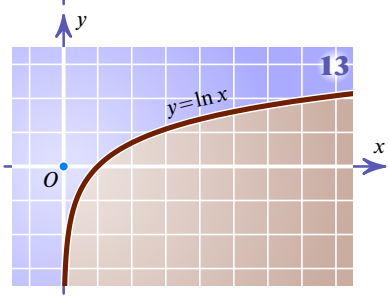
Из диаграммы Юнга, состоящей из столбцов высотой $a_1\geqslant a_2\geqslant\ldots\geqslant a_n$, можно получить элементарными преобразованиями Мюрхеда («падениями клеток») диаграмму, состоящей из столбцов высотой $b_1\geqslant b_2\geqslant\ldots\geqslant b_n$, тогда и только тогда, когда

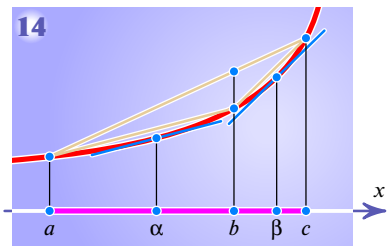
$$a_1+a_2+\ldots+a_n=b_1+b_2+\ldots b_n$$

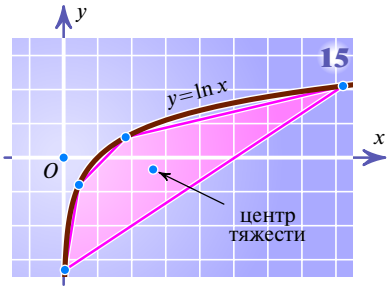
и выполнена система неравенств

$$a_1\geqslant b_1,$$
$$a_1+a_2\geqslant b_1+b_2,$$
$$\ldots\ldots\ldots\ldots$$
$$a_1+a_2+\ldots+a_{n-1}\geqslant$$
$$\geqslant b_1+b_2+\ldots+b_{n-1}.$$ ■









На рисунках 12 и 13 изображены подграфики функций $\sqrt{x}$ и $\ln x$ — тоже выпуклые множества. Функция f выпукла вверх, если ее подграфик — выпуклое множество.

Теорема 4. ***Если функция f дифференцируема на интервале и ее производная является неубывающей функцией, то f выпукла вниз.***

Доказательство. Рассмотрим любой лежащий внутри данного интервала отрезок $[a; c]$ и его внутреннюю точку b (рис. 14). По формуле конечных приращений Лагранжа существуют такие точки $\alpha\in(a; b)$ и $\beta\in(b; c)$, что $f'(\alpha)=\frac{f(b)-f(a)}{b-a}$ и $f'(\beta)=\frac{f(c)-f(b)}{c-b}$. Поскольку $f'(\alpha)>f'(\beta)$, то $\frac{f(b)-f(a)}{b-a}\geqslant\frac{f(c)-f(b)}{c-b}$, то есть

$$f(b)\leqslant\frac{c-b}{c-a}f(a)+\frac{b-a}{c-a}f(c).$$

Значит, точка $(b; f(b))$ лежит не выше точки $\left(b;\ \frac{c-b}{c-a}f(a)+\frac{b-a}{c-a}f(c)\right)$ отрезка, соединяющего точки $(a; f(a))$ и $(c; f(c))$ графика функции f. ■

На графике функции $y=\ln x$ отметим точки $(x_1; \ln x_1)$, $(x_2; \ln x_2)$, . . . , $(x_n; \ln x_n)$ и поместим в каждую из них грузик единичной массы (рис. 15). Центр тяжести — точка $\left(\frac{x_1+x_2+\ldots+x_n}{n};\ \frac{\ln x_1+\ln x_2+\ldots+\ln x_n}{n}\right)$ — лежит на графике, если $x_1=x_2=\ldots=x_n$; в остальных случаях центр тяжести лежит ниже графика:

$$\ln\frac{x_1+x_2+\ldots+x_n}{n}\geqslant\frac{\ln x_1+\ln x_2+\ldots+\ln x_n}{n}=\ln\sqrt[n]{\ln x_1+\ln x_2+\ldots+\ln x_n}.$$

Получили весьма наглядное доказательство теоремы 3. ■

И. Л. Йенсен (1859—1925) в 1906 г. сформулировал это в общем виде так: если функция f выпукла вниз на некотором интервале, $x_1, x_2, \ldots, x_n$ — точки этого интервала, а $m_1, m_2, \ldots, m_n$ — положительные числа, то

$$f\left(\frac{m_1x_1+m_2x_2+\ldots+m_nx_n}{m_1+m_2+\ldots+m_n}\right)\leqslant\frac{m_1f(x_1)+m_2f(x_2)+\ldots+m_nf(x_n)}{m_1+m_2+\ldots+m_n}.$$ ■

Функция $y=x^p$, где $p>1$, выпукла вниз, поэтому

$$\left(\frac{m_1x_1+\ldots+m_nx_n}{m_1+\ldots+m_n}\right)^p\leqslant\frac{m_1x_1^p+\ldots+m_nx_n^p}{m_1+\ldots+m_n},$$

то есть $(m_1x_1+\ldots+m_nx_n)^p\leqslant(m_1+\ldots+m_n)^{p-1}(m_1x_1^p+\ldots+m_nx_n^p)$. Обозначим $q=p/(p-1)$ и возведем обе части неравенства в степень $1/p$:

$$m_1x_1+\ldots+m_nx_n\leqslant(m_1+\ldots+m_n)^{1/q}(m_1x_1^p+\ldots+m_nx_n^p)^{1/p}.$$

Огюстен Луи Коши (1789—1857) — французский математик. Окончил Политехническую школу (1807) и школу мостов и дорог (1810). В 1810—1813 гг. работал инженером в Шербурге. В 1816—1830 гг. преподавал в Политехнической школе и Коллеж де Франс, с 1848 г. — в Коллеж де Франс и Парижском университете. После революции 1830 г. как сторонник Бурбонов был в эмиграции до 1838 г.

В его курсе анализа (1821—1828) даны определения предела, непрерывной функции, определенного интеграла как предела интегральных сумм, сходящегося ряда. Систематически развивал основы теории функций комплексной переменной (уравнения Коши—Римана). В 1831 г. понял, что радиус сходимости ряда Тейлора аналитической функции равен расстоянию до ближайшей особой точки.

Заметил, что в точке $x=0$ все производные функции вещественной переменной, равной 0 при $x=0$ и e^{-1/x^2} при $x\neq0$, равны 0; поэтому ее ряд Тейлора не сходится к самой функции.

В 1821 г. утверждал, что сумма сходящегося ряда непрерывных функций непрерывна. Причина ошибки — отсутствие понятия равномерной сходимости.

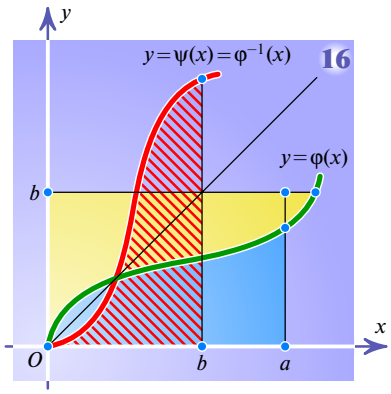
Дал первые доказательства существования решений дифференциальных уравнений. Применял две идеи: 1) получал искомую кривую как предел ломаных; 2) искал формальный степенной ряд, удовлетворяющий уравнению, а затем доказывал сходимость, подбирая мажоранту. Второй метод распространил и на комплексную область. ■

Положив теперь $m_k = b_k^q$ и $x_k = a_k b_k^{1-q}$, где $k = 1, \ldots, n$, получаем обобщение неравенства Коши—Буняковского — неравенство Х. Л. Гёльдера (1859—1937)

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} b_i^q\right)^{1/q},$$

где $p > 1$, $q > 1$ и $\frac{1}{p} + \frac{1}{q} = 1$. ■

Пусть $y = \varphi(x)$ — непрерывная возрастающая функция, причем $\varphi(0) = 0$. Пусть $a > 0$ и функция ψ обратна функции φ. Рассматривая рисунок 16, приходим к неравенству Юнга



$$\int_0^a \varphi(x)\,dx + \int_0^b \psi(y)\,dy \geqslant ab,$$

причем равенство имеет место лишь при $b = \varphi(a)$. Например, для $\varphi(x) = x^{p-1}$ получаем неравенство

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}. \blacksquare$$

Обозначим $A = \left(\sum_{1 \leqslant k \leqslant n} a_k^p\right)^{1/p}$ и $B = \left(\sum_{1 \leqslant k \leqslant n} b_k^q\right)^{1/q}$. Для $k = 1, 2, \ldots, n$ напишем неравенство Юнга $\frac{a_k b_k}{AB} \leqslant \frac{a_k^p}{pA^p} + \frac{b_k^q}{qB^q}$ и сложим эти n неравенств:

$$\frac{\sum_{1 \leqslant k \leqslant n} a_k b_k}{AB} \leqslant \frac{\sum_{1 \leqslant k \leqslant n} a_k^p}{pA^p} + \frac{\sum_{1 \leqslant k \leqslant n} b_k^q}{qB^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Следовательно, $\sum_{1 \leqslant k \leqslant n} a_k b_k \leqslant AB$ — неравенство Гёльдера доказано второй раз! ■

В пространстве $\mathbb{R}^n$ длину вектора $(x_1; x_2; \ldots; x_n)$ определим по формуле $\left(\sum_{1 \leqslant k \leqslant n} |x_k|^p\right)^{1/p}$. Оказывается, при $p \geqslant 1$ верно неравенство треугольника.

Теорема 5 (неравенство Минковского). ***Если $p \geqslant 1$, то для любых неотрицательных чисел $x_1, y_1, x_2, y_2, \ldots, x_n, y_n$ верно неравенство***

$$\left(\sum_{1 \leqslant k \leqslant n} (x_k + y_k)^p\right)^{1/p} \leqslant \left(\sum_{1 \leqslant k \leqslant n} x_k^p\right)^{1/p} + \left(\sum_{1 \leqslant k \leqslant n} y_k^p\right)^{1/p}.$$

Доказательство. Случай $p = 1$ и случай $x_1 = y_1 = x_2 = y_2 = \ldots = x_n = y_n = 0$ тривиальны. В остальных случаях применим неравенство Гёльдера:

$$\sum_{1 \leqslant k \leqslant n} (x_k + y_k)^p = \sum_{1 \leqslant k \leqslant n} x_k (x_k + y_k)^{p-1} + \sum_{1 \leqslant k \leqslant n} y_k (x_k + y_k)^{p-1} \leqslant$$
$$\leqslant \left(\sum_{1 \leqslant k \leqslant n} x_k^p\right)^{1/p} \left(\sum_{1 \leqslant k \leqslant n} (x_k + y_k)^p\right)^{1/q} + \left(\sum_{1 \leqslant k \leqslant n} y_k^p\right)^{1/p} \left(\sum_{1 \leqslant k \leqslant n} (x_k + y_k)^p\right)^{1/q}.$$

Осталось разделить обе части на $\left(\sum_{1 \leqslant k \leqslant n} (x_k + y_k)^p\right)^{1/q}$. ■

Докажем неравенство Минковского

$$\sqrt[n]{a_1 \cdot \ldots \cdot a_n} + \sqrt[n]{b_1 \cdot \ldots \cdot b_n} \leqslant \sqrt[n]{(a_1 + b_1) \cdot \ldots \cdot (a_n + b_n)}.$$

Делим обе части на $\sqrt[n]{a_1 \cdot \ldots \cdot a_n}$:

$$1 + \left(\frac{b_1}{a_1}\right)^{1/n} \cdot \left(\frac{b_2}{a_2}\right)^{1/n} \cdot \ldots \cdot \left(\frac{b_n}{a_n}\right)^{1/n} \leqslant \left(1 + \frac{b_1}{a_1}\right)^{1/n} \cdot \left(1 + \frac{b_2}{a_2}\right)^{1/n} \cdot \ldots \cdot \left(1 + \frac{b_n}{a_n}\right)^{1/n}.$$

Обозначив $x_k = \ln \frac{a_k}{b_k}$, где $k = 1, \ldots, n$, получаем

$$1 + e^{(x_1 + \ldots + x_n)/n} \leqslant (1 + e^{x_1})^{1/n} \cdot \ldots \cdot (1 + e^{x_n})^{1/n}.$$

Прологарифмируем обе части этого неравенства:

$$\ln(1 + e^{(x_1 + \ldots + x_n)/n}) \leqslant \frac{1}{n}(\ln(1 + e^{x_1}) + \ldots + \ln(1 + e^{x_n})).$$

Это в точности неравенство Йенсена для выпуклой вниз функции $\ln(1 + e^x)$. ■

Зафиксируем сумму $s = x_1 + x_2 + \ldots + x_n$ и в множестве неотрицательных значений переменных найдем максимальное значение произведения

$$p = x_1 x_2 \ldots x_n.$$

Выписываем функцию Лагранжа

$$f(x_1, x_2, \ldots, x_n) = x_1 x_2 \ldots x_n - \lambda(x_1 + x_2 + \ldots x_n)$$

и дифференцируем ее:

$$\frac{\partial f}{\partial x_k} = \frac{p}{x_k} - \lambda.$$

Значит, все частные производные равны 0 лишь в точке $x_1 = x_2 = \ldots = x_n = s/n$. Поскольку на границе рассматриваемой области максимум произведения не достигается, то $p \leqslant \left(\frac{s}{n}\right)^n$, а это и есть неравенство о среднем арифметическом и среднем геометрическом.

Методы математического анализа, одним из которых является метод множителей Лагранжа, оказываются весьма полезны и результативны при доказательстве и многих других неравенств. ■

Для нахождения простых чисел можно использовать решето Эратосфена: выписываем подряд натуральные числа 1, 2, 3, 4, ... (чем больше, тем лучше), а затем зачеркиваем сначала четные числа 4, 6, 8, 10, ..., затем делящиеся на 3 числа 9, 15, 21, ..., на следующем шаге — кратные 5 числа 25, 35, 55, ... Незачеркнутыми оказываются только простые числа и число 1, причем к моменту, когда зачеркиваем кратные числа n, все незачеркнутые числа, меньшие n^2, — простые.

Решета Эратосфена и Флавия отличаются тем, что Эратосфен зачеркивает числа, а Флавий — вычеркивает. ■

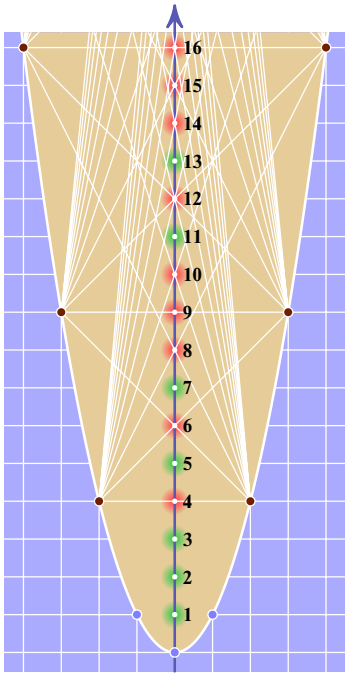
Отметим на параболе $y=x^2$ все точки с целыми кординатами, кроме $(0;0)$ и $(\pm1;1)$. Проведем всевозможные хорды с концами в отмеченных точках, пересекающие ось ординат. Прямая, проходящая через точки $(-a;a^2)$ и $(b;b^2)$, задана уравнением

$$\frac{x+a}{b+a}=\frac{y-a^2}{b^2-a^2}.$$

Подставив $x=0$ и освободившись от знаменателя, получаем

$$a(b-a)=y-a^2,$$

то есть $y=ab$. Значит, множество точек пересечения — это множество составных чисел! ■



Решето Иосифа Флавия

Вычеркнем из натурального ряда каждое второе число (2, 4, 6, ...), затем — каждое третье из оставшихся чисел, потом — каждое четвертое из оставшихся и так далее. Останется последовательность Иосифа Флавия: 1, 3, 7, 13, 19, 27, 39, 49, 63, 79, 91, 109, 133, 147, 181, 207, 223, 253, 289, 307, 349, ... Ее 3826-й член равен 11 499 769. При этом $4\cdot\dfrac{11\,499\,769}{3826^2}\approx3{,}1423\approx\pi$.

Мы докажем, что частное от деления n-го члена этой последовательности на n^2 стремится к $\pi/4$.

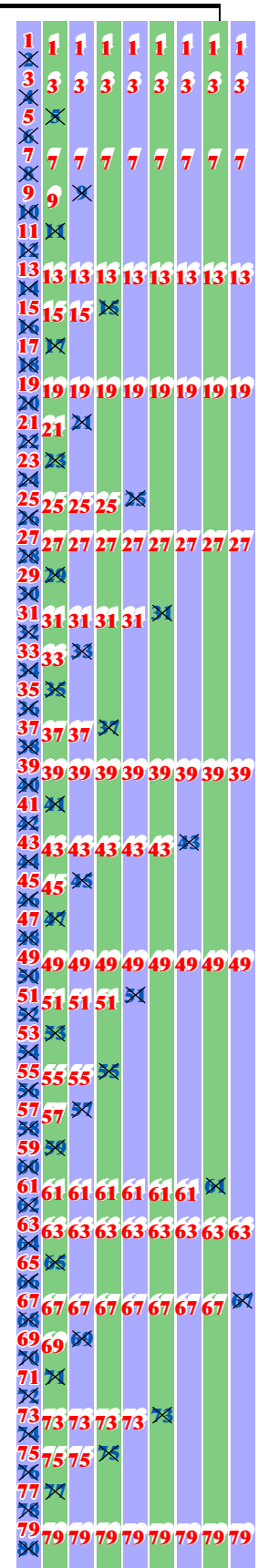
Обозначим n-е число Флавия через N.

Для вычислений возьмем не весь натуральный ряд, а лишь первые N натуральных чисел. Сначала вычеркиваем каждое второе число, оставляя невычеркнутыми $N-[N/2]$ чисел, то есть приблизительно $N/2$. Затем вычеркиваем каждое третье число из оставшихся, так что останется примерно 2/3 от 1/2, то есть приблизительно $N/3$. (Точную формулу все еще можно написать: $N-[N/2]-\left[\dfrac{N-[N/2]}{3}\right]$. Но чем дальше, тем больше понадобится знаков целой части, так что лучше уж писать приближенные формулы, чем нагромождать квадратные скобки.) Когда вычеркнем каждое четвертое из оставшихся чисел, останется примерно 3/4 предыдущего количества, то есть $\frac{3}{4}N/3=N/4$. И вообще, после вычеркивания каждого k-го числа останется примерно N/k чисел.

Казалось бы, все сделано: вычеркивания заканчиваются в момент, когда впервые $k>N/k$, то есть $k^2>N$. Значит, n-й член последовательности должен быть примерно равен n^2. Только вот компьютерные вычисления убедительно свидетельствуют: правильный ответ — скорее всего $\pi n^2/4$, но никак не n^2.

Причина большой ошибки — в накоплении маленьких. Заменяя целую часть числа им самим, мы делаем не очень большую — не больше 1 — ошибку. Но за n шагов ошибки накапливаются и существенно искажают результат.

Заметьте: в начале вычислений k невелико, а N/k огромно, так что относительная ошибка невелика и возрастает с уменьшением величины N/k. Значит, последние стадии вычеркиваний надо исследовать какими-то иными средствами, чем первые. ■



Вычеркивания среди первых N натуральных чисел прекращаются, когда k становится больше n. К этому моменту уцелело ровно n чисел. Перед этим довольно продолжительное время k приблизительно равно n; на каждом таком шаге вычеркиваем по одному числу. Перед этим несколько более короткое, но тоже продолжительное время вычеркивали по два числа, перед этим — по три и так далее.

Придадим этому более точную форму. Пусть последний шаг вычислений — вычеркивание n-го числа из $n+1$ уцелевших к этому моменту. Тогда перед этим вычеркнули $(n-1)$-е число из $n+2$, раньше — $(n-2)$-е из $n+3$, и так далее, $(n-a)$-е — из $n+1+a$. Переход на другой режим происходит в момент, когда выполняется неравенство $2(n-a)\geqslant n+1+a$, то есть когда $a\approx n/3$. При этом $k\approx 2n/3$, а уцелели примерно $2k$ чисел.

Формулы $k\approx\frac{2}{3}n-b$ и $r\approx\frac{4}{3}n+2b$, где r — количество еще не вычеркнутых чисел, описывают процесс вычеркивания по два числа. Переход на режим вычеркивания по три числа соответствует равенству $3\left(\frac{2}{3}n-b\right)\approx\frac{4}{3}n+2b$, из которого находим $b\approx\frac{1}{5}\cdot\frac{2}{3}n$, откуда $k\approx\frac{2}{3}\cdot\frac{4}{5}n$ и $r\approx 3k$.

Продолжая в том же духе и обозначив для краткости $c_m=\frac{2}{3}\cdot\frac{4}{5}\cdot\ldots\cdot\frac{2m}{2m+1}$, находим, что переход между режимами вычеркивания по m и по $m+1$ чисел происходит при $k\approx c_m n$ и $r\approx(m+1)k$. Эти формулы довольно точны при маленьких m и все менее и менее точны при возрастании m. Эффект такой же, какой мы уже наблюдали, когда рассматривали процесс от начала к концу: ошибка накапливается и оказывается при больших m сопоставимой с самой исследуемой величиной. ■

Стыковка. Так что же делать, если и при рассмотрении от начала к концу, и при рассмотрении от конца к началу успевает накопиться ошибка, хотя каждый из этих процессов сначала ведет себя хорошо и лишь потом безобразничает? Надо рассмотреть момент стыковки:

$$k\approx c_m n \qquad \text{и} \qquad \frac{N}{k}\approx(m+1)c_m n,$$

откуда $N\approx(m+1)c_m^2 n^2$. В силу формулы Валлиса, $(m+1)c_m^2\to\pi/4$, что и требовалось!

Чтобы превратить это рассуждение в доказательство, нужно оценить ошибки, то есть превратить все приближенные равенства в безупречные неравенства. Это сделать можно, но мы этим трудоемким и скучным делом здесь заниматься не будем. ■

Имя Иосифа Флавия носит и следующая простая задача. Выпишем первые n натуральных чисел по кругу и будем вычеркивать каждое второе, пока не останется лишь одно число $f(n)$. Например, при $n=5$ вычеркиваем 2, 4, 3 и 1, поэтому $f(5)=3$. А при $n=12$ вычеркиваем 2, 4, 6, 8, 10, 12, 3, 7, 11, 5 и 1, поэтому $f(12)=9$.

В общем случае, если $n=2^k+m$, где $m<2^k$, останется число $f(n)=2m+1$. Ключевой момент доказательства состоит в том, что $f(2^k)=1$. (Проверьте!) Если же $n=2^k+m$, количество чисел после первых m вычеркиваний становится степенью двойки. Очередным числом в этот момент является $2m+1$, оно и уцелеет! ■

Английский математик Джон Валлис (1616—1703) в 1665 г. опубликовал формулу

$$\lim_{m\to\infty} m\left(\frac{2}{3}\cdot\frac{4}{5}\cdot\frac{6}{7}\cdot\ldots\cdot\frac{2m}{2m+1}\right)^2=\frac{\pi}{4}.$$

Доказательство начнем издалека. Пусть $a_m=\int\limits_0^\pi \sin^m x\,dx$.

Тогда $a_0=\pi$ и $a_1=2$. Далее — интегрирование по частям в стиле вытягивающего себя за волосы из болота Мюнхгаузена:

$$a_{m+1}=\int\limits_0^\pi \sin^{m+1}x\,dx=$$
$$=-\int\limits_0^\pi \sin^m x\,d\cos x=$$
$$=-\sin^m x\cos x\Big|_0^\pi+$$
$$+m\int\limits_0^\pi \cos^2 x\sin^{m-1}x\,dx=$$
$$=m\int\limits_0^\pi(1-\sin^2 x)\sin^{m-1}x\,dx=$$
$$=ma_{m-1}-ma_{m+1},$$

откуда $a_{m+1}=\frac{m}{m+1}a_{m-1}$.

Теперь легко находим $a_2=\frac{1}{2}a_0=\frac{\pi}{2}$, $a_3=\frac{2}{3}a_1=2\cdot\frac{2}{3}$, $a_4=\frac{3}{4}a_2=\frac{1}{2}\cdot\frac{3}{4}\pi$, и вообще

$$a_{2m}=\frac{1}{2}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{2m-1}{2m}\pi,$$
$$a_{2m+1}=2\cdot\frac{2}{3}\cdot\frac{4}{5}\cdot\ldots\cdot\frac{2m}{2m+1}.$$

Из неравенств $a_{2m-1}>a_{2m}>a_{2m+1}$ следуют неравенства

$$1>\frac{a_{2m}}{a_{2m-1}}>\frac{2m}{2m+1},$$

из которых легко получить формулу Валлиса. Обычно ее приводят в равносильном виде

$$\frac{\pi}{2}=\prod_{m=1}^{\infty}\frac{2m}{2m-1}\cdot\frac{2m}{2m+1}.$$ ■

Иосиф Флавий — античный историк — в составе отряда из 41 иудейского воина был, как гласит легенда, загнан римлянами в пещеру. Предпочитая самоубийство плену, воины выстроились в круг и последовательно убивали каждого третьего из живых до тех пор, пока все не погибнут. Однако Иосиф (даже не один, а вместе со своим другом) спасся — вычислил два спасительных места в страшном круге, на которые и встали Иосиф и его друг. ■

Комбинаторика

Желая дать представление о предмете одной фразой, говорят, что комбинаторика — это раздел математики, в котором изучают, сколько подчиненных тем или иным условиям комбинаций можно составить из данных объектов. Или говорят, что комбинаторика изучает конечные множества и разные структуры на них.

Зададим себе вопрос: *сколько существует трехзначных чисел*? Самое большое трехзначное число — это 999. Самое большое двузначное — 99. Поэтому существует $999-99=900$ трехзначных чисел.

Тот же ответ можно получить и другим способом. Вообразите, что мы пишем, цифра за цифрой, трехзначное число. Сначала напишем любую из девяти цифр 1, 2, . . . , 9 в разряд сотен. Затем любую из десяти цифр 0, 1, . . . , 9 — в разряд десятков; наконец, какую-нибудь (любую) цифру — в разряд единиц. Ответ: $9\cdot 10\cdot 10=900$.

Сколькими способами можно расположить на шахматной доске белую и черную ладьи, чтобы они не били одна другую? Сначала поставим на любую из 64 клеток доски белую ладью. Для черной ладьи останется 49 полей. Ответ: $64\cdot 49=3136$.

Заметьте: мы перемножаем числа 64 и 49, а не складываем их. Одна из стандартных ошибок, которую делают начинающие, состоит в том, что они путают, когда надо складывать, а когда умножать. Между тем все просто: сумма $m+n$ есть количество элементов объединения двух непересекающихся множеств, одно из которых состоит из m, а другое из n элементов; а произведение mn — это количество пар вида $(x; y)$, где x может быть любым из m элементов некоторого множества (например, это может быть множество 64 возможных полей для белой ладьи), а y, при каждом фиксированном x, может быть любым из некоторых n элементов (например, y — одно из 49 возможных полей для черной ладьи).

Сколько сторон и диагоналей у выпуклого 15-угольника? Можно, конечно, нарисовать 15-угольник на большом листе бумаги и посчитать. Но лучше найти общую формулу.

Очевидно, у треугольника 3 стороны и ни одной диагонали; у четырехугольника 4 стороны и 2 диагонали; у пятиугольника 5 сторон и столько же диагоналей, а у шестиугольника 6 сторон и 9 диагоналей. Получили ряд чисел: 3, 6, 10, 15. Их называют *треугольными числами* и вычисляют по формуле $n(n-1)/2$. В частности, у 15-угольника $15\cdot 14/2=105$ диагоналей и сторон.

Докажем эту формулу по индукции. Очевидно, $(n+1)$-угольник можно получить из n-угольника добавлением одной новой вершины. Из этой вершины выходят отрезки — стороны и диагонали — во все остальные n вершин. Значит, $(n+1)$-угольник имеет ровно на n больше сторон и диагоналей, чем n-угольник. Поскольку

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

мы видим, как преобразуется формула при переходе от n к $n+1$. Точнее говоря, если формула давала верное значение для n-угольника, то и для $(n+1)$-угольника все будет правильно. Мы помним, что при $n=3$ она верна. Значит, по индукции, она верна и при любом n.

Впрочем, формулу $n(n-1)/2$ можно доказать и проще. Из каждой вершины n-угольника выходят $n-1$ отрезков (2 стороны и $n-3$ диагоналей). Умножая n на $n-1$ и деля на 2 (ибо у каждого отрезка два конца), получаем ответ!

От простых задач перейдем к довольно трудной. Комбинаторика занимается не только конечными множествами, но последовательностями чисел, иной раз весьма экзотическими. Дж. Лагариас, И. Рейнс и Н. Слоан придумали следующую последовательность. Первые два ее члена $a_1=1$ и $a_2=2$, а рекуррентное правило таково: a_{n+1} — наименьшее натуральное число, которое еще не встретилось в последовательности и не взаимно просто с a_n.

Первые 20 членов последовательности таковы: 1, 2, 4, 6, 3, 9, 12, 8, 10, 5, 15, 18, 14, 7, 21, 24, 16, 20, 25 и 30. За очевидную нерегулярность поведения авторы назвали ее ЭКГ-последовательностью (буквы ЭКГ означают «электрокардиограмма»). Докажем, что каждое натуральное число входит в нее.

Лемма 1. ***ЭКГ-последовательность содержит бесконечно много четных чисел.***

Доказательство. Предположим противное: лишь несколько членов ЭКГ-последовательности четны, то есть существует такое натуральное m, что все числа a_m, a_{m+1}, a_{m+2}, . . . нечетны. Поскольку все члены ЭКГ-последовательности различны, существует такое натуральное $k\geqslant m$, что a_{k+1} больше каждого из чисел $a_1, a_2, \ldots, a_k$. Обозначим буквой p наименьший простой делитель числа a_k. Очевидно, $a_{k+1}\geqslant a_k+p$ (иначе числа a_{k+1} и a_k были бы взаимно простыми). Поскольку сумма a_k+p двух нечетных

чисел четна, то $a_{k+1} > a_k + p$. Следовательно, a_{k+1} не является наименьшим натуральным числом, отличным от $a_1, a_2, \ldots, a_k$ и взаимно простым с a_k. Противоречие.

Лемма 2. ***Пусть p — простое число. Если ЭКГ-последовательность содержит бесконечно много чисел, кратных p, то она содержит все кратные числа p.***

Доказательство. Пусть число pn не содержится в ЭКГ-последовательности, а все предыдущие кратные числа p — содержатся. Почти все (то есть все, начиная с некоторого номера m) члены последовательности больше pn. Рассмотрим такое кратное p число a_k, что $k > m$. Очевидно, pn претендует на роль a_{k+1}. Противоречие.

Теперь — собственно доказательство того, что ЭКГ-последовательность содержит все натуральные числа. По лемме 1, последовательность содержит бесконечно много четных чисел. Следовательно, по лемме 2, она содержит все четные числа и поэтому для любого простого числа p содержит бесконечно много чисел, кратных p. Еще раз применив лемму 2, заключаем, что для любого простого p ЭКГ-последовательность содержит все натуральные числа, кратные p. Утверждение доказано.

Вычислив на компьютере первые несколько тысяч членов ЭКГ-последовательности и нарисовав ее график, можно прийти к такой гипотезе: если a_n — простое число, то $a_n \approx n/2$; если a_n — утроенное простое число, то $a_n \approx 3n/2$; в остальных случаях $a_n \approx n$. Однако пока доказаны лишь неравенства

$$\frac{n}{260} \leqslant a_n \leqslant 14n.$$

Расскажем еще об одном достижении комбинаторики. В июле 2005 г., когда работа над этой книгой уже завершалась, Б. Ласс вывел из тождества Гаусса—Якоби, о котором рассказано в статье «Тождества и биекции», доказательство гипотезы Д. Дюмона: для любых натуральных чисел n и k количество представлений числа n в виде суммы k квадратов нечетных натуральных чисел (порядок слагаемых учитываем, то есть $1^2 + 5^2$ не отождествляем с $5^2 + 1^2$) равно $A - (-1)^k B$, где A — количество решений уравнения

$$x_1x_2 + x_2x_3 + \ldots + x_{k-1}x_k + x_kx_1 = n$$

в натуральных числах $x_1, x_2, \ldots, x_k$, дающих остаток 1 при делении на 4, а B — количество решений того же уравнения в натуральных числах, дающих при делении на 4 остаток 3.

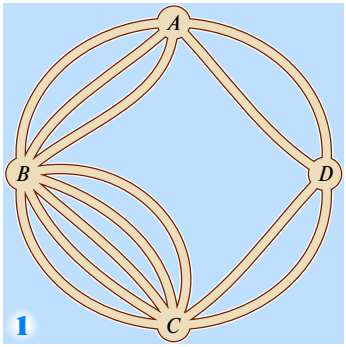
Из этого результата можно вывести классические теоремы Лагранжа, Гаусса, Якоби и Кронекера о количествах представлений натуральных чисел в виде суммы двух, трех или четырех квадратов. Таким образом, статьи «Суммы двух квадратов» и «Суммы четырех квадратов» раздела «Арифметика» связаны со статьей «Тождества и биекции» раздела «Комбинаторика». Раньше об этом и не подозревали!

Тут уместно процитировать Д. Гильберта. Более 100 лет назад он завершил свой знаменитый доклад «Математические проблемы» так:

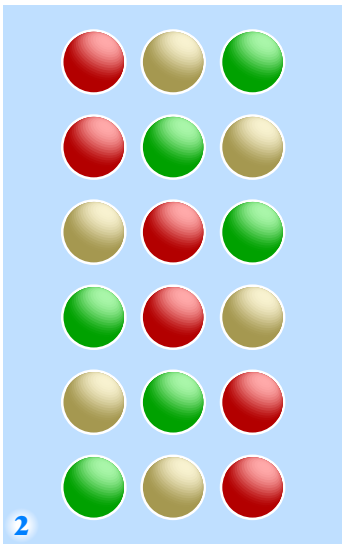
«Названные проблемы — это только образцы проблем; но их достаточно, чтобы показать, как богата, многообразна и широка математическая наука уже сейчас; перед нами встает вопрос, предстоит ли математике когда-нибудь то, что с другими науками происходит с давних пор: не распадется ли она на отдельные частные науки, представители которых будут едва понимать друг друга и связь между которыми будет становиться все меньше. Я не верю в это и не хочу этого. Математическая наука, на мой взгляд, представляет неделимое целое, организм, жизнеспособность которого обусловливается связностью его частей. Ведь при всем различии математического материала в частностях мы все же очень ясно видим тождественность логических вспомогательных средств, родство образования идей в математике в целом и многочисленные аналогии в ее различных областях. Мы также замечаем, что чем дальше развивается математическая теория, тем гармоничнее и более едино оформляется ее сооружение и между до сих пор разделенными областями открываются неожиданные связи. Так получается, что при расширении математики ее единый характер не теряется, а становится все более отчетливым.

Но — спросим мы — при расширении математического знания не станет ли в конце концов невозможным для отдельного исследователя охватить все его части? В качестве ответа я хочу сослаться на то, что существо математической науки таково, что каждый действительный успех в ней идет рука об руку с нахождением более сильных вспомогательных средств и более простых методов, которые одновременно облегчают понимание более ранних теорий и устраняют затруднительные старые рассуждения; поэтому отдельному исследователю, благодаря тому что он усвоит эти более сильные вспомогательные средства и более простые методы, удастся легче ориентироваться в различных областях математики, чем это имеет место для какой-нибудь другой науки.

Единый характер математики обусловлен внутренним существом этой науки; ведь математика — основа всего точного естествознания. А для того чтобы в совершенстве выполнить это высокое назначение, пусть в грядущем столетии она обретет гениальных мастеров и многочисленных, пылающих благородным рвением приверженцев».



1


2

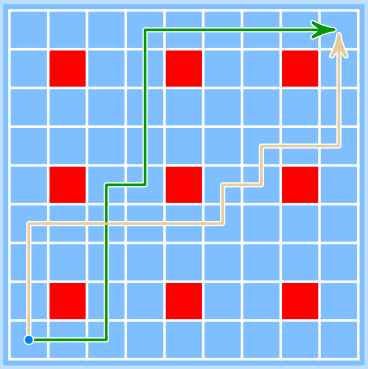
Перестановки

Комбинаторика — это раздел математики, в котором изучают сколько комбинаций, подчиненных тем или иным условиям, можно составить из данных объектов. Прежде чем переходить к общим принципам, рассмотрим несколько простых примеров.

1) **Сколькими способами** можно проехать из A в C, если система дорог такова, как показано на рисунке 1? Прежде чем попасть из A в C, надо или любым из трех возможных способов попасть из A в B, а затем любым из пяти способов — из B в C; или же любым из двух способов попасть из A в D, а затем любым из двух способов — из D в C. **Ответ:** $3\cdot 5+2\cdot 2=19$.

2) Сколькими способами можно выложить в ряд красный, желтый и зеленый шарики? На рисунке 2 изображены все 6 способов.

3) Сколькими способами можно пройти из левой нижней клетки квадрата 9×9 в правую верхнюю, двигаясь на каждом шаге на единицу вправо или на единицу вверх и ни разу не побывав ни на одной закрашенной красным цветом клетке (рис. 3)? На рисунке 3 показаны два разных маршрута, но нет терпения рисовать все варианты! Что же делать? Поставим перед собой более скромную цель: найдем количество путей из левой нижней клетки не в правую верхнюю, а например, в клетку A (рис. 4). Очевидно, таких путей два. Теперь легко понять, что в клетку B можно пройти четырьмя способами, а в клетку C — восемью.


3

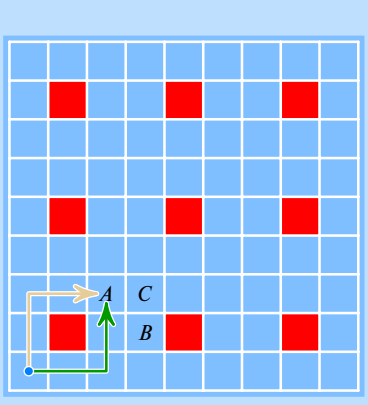
Впрочем, даже сейчас мы ставим перед собой слишком трудные вопросы. Лучше заполнять таблицу последовательно, клетку за клеткой, многократно используя правило суммы: если в некоторую клетку Z можно попасть из клеток X и Y, то число способов попасть в клетку Z есть сумма чисел способов попасть в X и Y. На рисунке 5, например, можно заполнить клетку Z, написав туда $24=7+17$. Так потихонечку и заполним всю доску (рис. 6).

4) Сколько слов (не обязательно осмысленных) можно получить, переставляя буквы слова МАМА? **Ответ:**

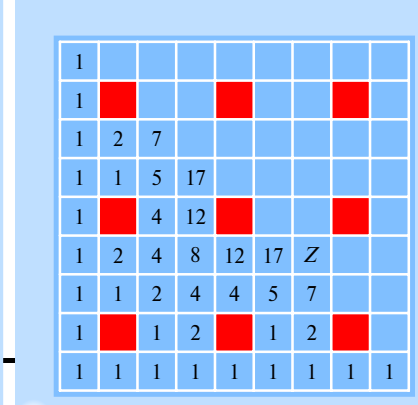
МАМА, МААМ, ММАА,
АМАМ, АММА, ААММ

— всего 6 способов.

Мы разобрали уже 4 разные комбинаторные задачи. Перейдем к более общим понятиям. ■



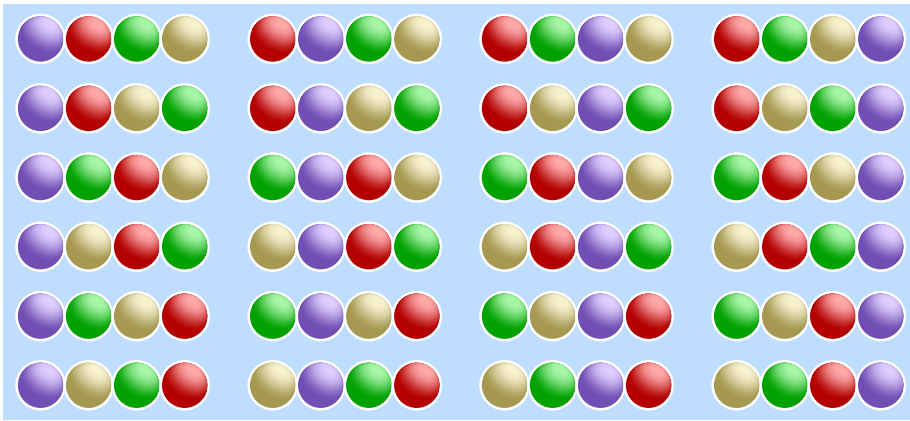
4



5

6

Перестановки. Факториал. Два элемента a и b могут быть выписаны в строчку всего двумя способами: ab и ba. Для трех элементов, как мы знаем из четвертого примера, существует 6 вариантов. Нетрудно посчитать и число перестановок множества из 4 элементов:



Всего 24 перестановки, расположенные в 4 столбца по 6 перестановок в каждом. Очевидно, перестановки на 5 элементах можно расположить в 5 столбцов, по 24 в каждом. Значит, всего существует $5\cdot 24=120$ таких перестановок. Для числа перестановок n элементов есть обозначение: $n!$ (читается «эн факториал»). Факториал равен произведению всех натуральных чисел от 1 до n. Например, $4!=1\cdot 2\cdot 3\cdot 4$. Функция $n!$ возрастает очень быстро. Так, $1!=1$, $2!=2$, $3!=6$, . . . , $10!=3\,628\,800$. Главное свойство факториала очевидно из определения:

$$(n+1)!=(n+1)\cdot n!.$$

Подставим в эту формулу $n=0$. Имеем: $1!=1\cdot 0!$, откуда $0!=1$. И действительно, во многих формулах для единообразной записи очень удобно пользоваться обозначением $0!=1$. (А вот определить $(-1)!$ невозможно: равенство $0!=0\cdot(-1)!$ невозможно ни при каком значении $(-1)!$.) ■

Размещения. Следующее важное понятие комбинаторики — размещение. Рассмотрим такую ситуацию: в классе, в котором 25 учеников, нужно выбрать старосту, его заместителя и помощника заместителя. Сколькими способами это можно сделать?

Очевидно, сначала 25 способами можно выбрать любого ученика в старосты. Затем из 24 оставшихся — заместителя старосты, а после этого любой из 23 оставшихся может оказаться помощником заместителя. По правилу произведения, всего имеем $A_{25}^3=25\cdot 24\cdot 23$ вариантов.

Вообще, через A_n^k (читается «а из эн по ка») обозначают число способов выбрать из данных n элементов сначала первый элемент, потом второй, третий, . . . , k-й. Вычисляют его по формуле

$$A_n^k=n(n-1)\ldots(n-k+1).$$

Заметьте: в правой части ровно k множителей, и последний из них равен $n-k+1$, а вовсе не $n-k$, как могло показаться на первый взгляд. Формулу можно записать и через факториалы: $A_n^k=\dfrac{n!}{(n-k)!}$. Для числа размещений есть и другое обозначение: $A_n^k=n^{\underline{k}}$. ■

Факториалы возникают в комбинаторике очень часто. Поэтому принято считать, что если ответ выражен через факториал, то все сделано. Этому в немалой степени способствует открытая в 1730 г. Дж. Стирлингом (1692—1770) формула

$$n!\approx\sqrt{2\pi n}\,n^n/e^n.$$

Относительная ошибка в этом приближении очень мала и стремится к нулю при увеличении числа n.

Для доказательства формулы Стирлинга воспользуемся логарифмами, интегралами и формулой Валлиса (см. статью «Решето Иосифа Флавия»). $\ln n!=$ $=\ln 2+\ln 3+\ldots+\ln n$ — это сумма площадей прямоугольников (догадайтесь, как расположенных относительно графика $y=\ln x$). Ширина каждого из них равна 1, а высота n-го равна $\ln n$. Значит, $\ln n!-\int\limits_1^n \ln t\,dt$ — это сумма площадей криволинейных треугольников (каждый из которых получается выбрасыванием из прямолинейного треугольника маленького кусочка, заключенного между графиком и его хордой. Интегрируя по частям, получаем

$$\int\limits_1^n \ln t\,dt=t\ln t\Big|_1^n-\int\limits_1^n t\,d\ln t=$$

$$=n\ln n-\int\limits_1^n dt=n\ln n-n.$$

Мы уже выявили самые большие сомножители формулы Стирлинга: n^ne^{-n}. Осталось понять, откуда берутся $\sqrt{n}$ и $\sqrt{2\pi}$. Поскольку $\ln\sqrt{n}=\dfrac{\ln n}{2}$, для объяснения смысла множителя $\sqrt{n}$ достаточно заметить, что $\dfrac{\ln n}{2}$ — это сумма площадей не криволинейных, а обычных треугольников: ведь их площади равны $\dfrac{\ln 2}{2}$, $\dfrac{\ln 3-\ln 2}{2}$, $\dfrac{\ln 4-\ln 4}{2}$, . . . , $\dfrac{\ln n-\ln(n-1)}{2}$.

А множитель $\sqrt{2\pi}$ можно найти при помощи формулы Валлиса. Для это нужно доказать, что сумма площадей зеленых «добавок» стремится к некоторой постоянной величине и тем самым доказать формулу

$$n!\approx k\sqrt{n}\,n^n/e^n$$

с некоторым неизвестным пока числом k. Подставив это выражение для факториала в формулу Валлиса, находим $k=\sqrt{2\pi}$. ■

Числа сочетаний

Важные для комбинаторики и теории вероятностей числа сочетаний C_n^k можно определить многими разными способами, например формулой

$$C_n^k = A_n^k/k! = n^{\underline{k}}/k!. \qquad (*)$$

Но гораздо интереснее выяснить их комбинаторный смысл: числа сочетаний возникают в самых разных задачах, на первый взгляд не имеющих ничего общего.

Число сочетаний из n по k — это количество k-элементных подмножеств в множестве, состоящем из n элементов. Другими словами, это количество слов длины n, составленных из k букв А и $n-k$ букв Б. Или число способов пройти из начала координат в точку $(k; n-k)$, если каждый шаг — сдвиг на единицу вверх или вправо. ■

На рисунке 1 изображен треугольник Б. Паскаля. Его число, стоящее на k-м месте в n-м ряду, обозначают C_n^k. В частности, $C_0^0=1$, $C_1^0=C_1^1=1$, $C_2^0=C_2^2=1$, $C_2^1=2$, . . . , $C_6^3=20$, . . . Очевидно, $C_n^0=1=C_n^n$. При помощи рекуррентной формулы

$$C_{n+1}^k = C_n^{k+1} + C_n^k$$

можно вычислить любое число треугольника Паскаля как сумму стоящих над ним слева и справа. Очевидно, треугольник симметричен относительно вертикальной прямой, что выражается формулой

$$C_n^k = C_n^{n-k}.$$

Чуть менее очевидно другое интересное свойство: сумма чисел n-й строки треугольника Паскаля есть степень двойки:

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^{n-1} + C_n^n = 2^n.$$

Дело в том, что при переходе от одной строки треугольника Паскаля к следующей сумма возрастает ровно вдвое: каждое число участвует в образовании двух чисел, расположенных под ним слева и справа. Треугольник Паскаля обладает и другими интересными свойствами: например, если заменим каждое его четное число нулем, а каждое нечетное единицей, то получим красивый узор (рис. 2). Видно, что все числа $C_{2^n-1}^k$ нечетны, а числа $C_{2^n}^k$, где $0<k<2^n$, наоборот, все четны. Красиво устроены и остатки от деления на 3 (рис. 3).

Суммируя числа треугольника Паскаля вдоль синих линий рисунка 1, получаем числа Фибоначчи. (Докажите!) Известны десятки других интересных свойств чисел сочетаний. Но мы сейчас расскажем о самом важном. ■

Пусть m и n — натуральные числа. Рассмотрим уравнение $x_1+x_2+\ldots+x_m=n$. Каждое его решение в натуральных числах можно изобразить в виде разбиения n лежащих в ряд шариков на группы, считая слева направо, из $x_1, x_2, \ldots, x_m$ шариков (на рисунке $m=4, x_1=1, x_2=3, x_3=3, x_4=2$ и $n=1+3+3+2=9$).

Разбиение задается положениями $m-1$ красных разделителей, которые могут располагаться на любых из $n-1$ возможных позиций. Поэтому уравнение имеет C_{n-1}^{m-1} решений. ■

Выяснить, сколько решений имеет уравнение

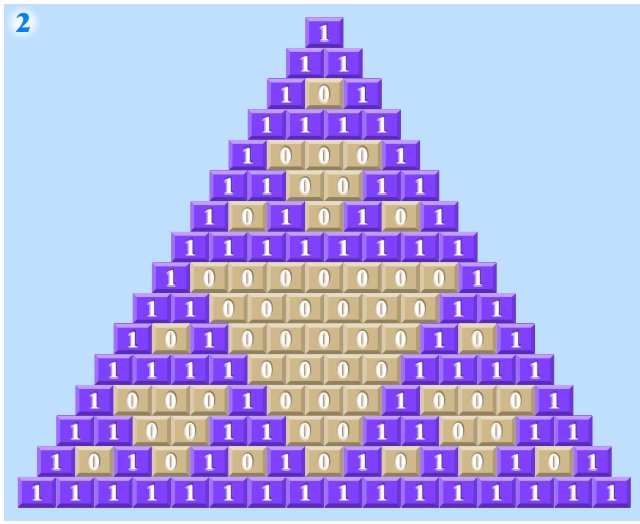
$$y_1+y_2+\ldots+y_m=n$$

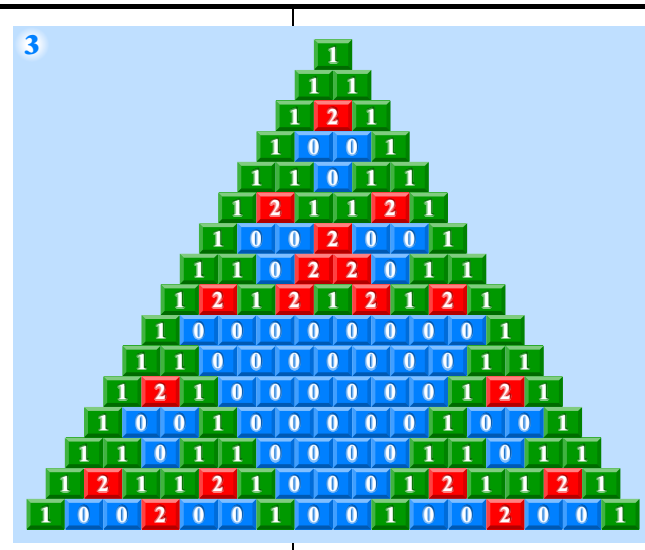
в неотрицательных целых числах $y_1, y_2, \ldots, y_m$, помогает замена $y_1=x_1-1$, $y_2=x_2-1$, ... , $y_m=x_m-1$, приводящая к уравнению

$$x_1+x_2+\ldots+x_m=n+m.$$

Ответ: C_{m+n-1}^{m-1}. ■

1





Бином Ньютона. Как известно,

$$(a+b)^2 = a^2 + 2ab + b^2,$$
$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3.$$

Можно возвести сумму $a+b$ и в четвертую степень:

$$\begin{aligned}(a+b)^4 &= (a^3 + 3a^2b + 3ab^2 + b^3)(a+b) = \\ &= a^4 + 3a^3b + 3a^2b^2 + ab^3 + \\ &\quad + a^3b + 3a^2b^2 + 3ab^3 + b^4 = \\ &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.\end{aligned}$$

Легко заметить закон образования коэффициентов: коэффициент 4 при a^3b есть сумма коэффициентов 3 и 1 при a^2b и a^3. Аналогично, коэффициент 6 при a^2b^2 является суммой $3+3$ коэффициентов при ab^2 и a^2b.
Далее, домножив $(a+b)^4$ на $a+b$, получим следующую формулу:

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5.$$

Числа 1, 5, 10, 10, 5, 1, как мы помним, образуют 5-ю строку треугольника Паскаля. Вообще,

$$(a+b)^n = a^n + C_n^1 a^{n-1}b + C_n^2 a^{n-2}b^2 + C_n^3 a^{n-3}b^3 + \ldots + C_n^{n-1} ab^{n-1} + b^n. \qquad (**)$$

Поэтому числа сочетаний C_n^k также называют *биномиальными коэффициентами*. Первое дошедшее до нас описание формулы бинома Ньютона содержится в появившейся в 1265 г. книге среднеазиатского математика ат-Туси, где дана таблица биномиальных коэффициентов C_n^k до $n=12$ включительно. В 1664—1665 гг. И. Ньютон обобщил формулу (**) на случай произвольных (дробных и отрицательных) показателей, но при этом получаются ряды — суммы бесконечного множества слагаемых. А именно, при $|x|<1$ имеем

$$(1+x)^n = 1 + nx + \frac{n\cdot(n-1)}{1\cdot 2}x^2 + \frac{n\cdot(n-1)\cdot(n-2)}{1\cdot 2\cdot 3}x^3 + \ldots + \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{1\cdot 2\cdot\ldots\cdot k}x^k + \ldots$$

При $n=-1$ эта формула превращается в формулу суммы бесконечной геометрической прогрессии:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots + (-1)^{n-1}x^n + \ldots \blacksquare$$

Блез Паскаль (1623—1662) — французский математик, механик, физик и философ. Занимался математикой под руководством отца. Трактат «Опыт теории конических сечений» написал в 16-летнем возрасте. В нем содержится одна из главных теорем проективной геометрии — теорема Паскаля: во всяком шестиугольнике, вписанном в окружность (или другое коническое сечение: эллипс, гиперболу или параболу), точки пересечения трех пар противоположных сторон (или их продолжений) лежат на одной прямой. При этом шестиугольник может быть самопересекающимся. Частный случай теоремы Паскаля для конического сечения, распадающегося на две прямые, известен с древних времен как теорема Паппа. ■

Докажем формулу (*) для чисел треугольника Паскаля.
I способ. Индукция:

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

II способ. Число C_n^k треугольника Паскаля равно количеству способов сделать k шагов вправо-вниз и $n-k$ шагов влево-вниз. Это количество равно количеству способов выбрать k элементов из n данных. Осталось заметить, что A_n^k есть число способов выбрать упорядоченный набор из k элементов n-элементного множества; разделив на $k!$, перейдем от упорядоченных наборов к неупорядоченным. ■

Цепи и антицепи

Для любого частично упорядоченного множества наибольшее количество элементов его цепи равно наименьшему количеству антицепей, на которые можно разбить это частично упорядоченное множество. Верно и утверждение, получаемое переменой местами слов «цепь» и «антицепь».

Из любых ли пяти различных чисел, выписанных в ряд, можно выбрать три, которые расположены в этом ряду в порядке убывания или в порядке возрастания?

Да, из любых. Обозначим числа, слева направо, через x_1, x_2, x_3, x_4 и x_5. Не ограничивая общности, можно считать, что $x_1 < x_2$. (Иначе можно в дальнейшем рассуждении заменить везде знаки «<» на знаки «>» и наоборот.) Если при этом еще и $x_2 < x_3$, то первые три числа расположены в порядке возрастания. Значит, надо разобрать случай $x_1 < x_2 > x_3$. Если $x_3 > x_4$, то годится тройка $x_2 > x_3 > x_4$. Поэтому можно считать, что $x_1 < x_2 > x_3 < x_4$. Далее, если $x_4 < x_5$, то годится тройка $x_3 < x_4 < x_5$. Таким образом, надо разобрать случай $x_1 < x_2 > x_3 < x_4 > x_5$. Сравним числа x_2 и x_4. Если $x_2 < x_4$, то годится тройка $x_1 < x_2 < x_4$. Если же $x_2 > x_4$, то годится $x_2 > x_4 > x_5$. Существование трехэлементной монотонной подпоследовательности доказано.

Есть и более короткое рассуждение. Пусть a и b — наибольшее и наименьшее из данных чисел. Если между ними есть какое-то число, то утверждение верно. Если же они стоят рядом, то либо слева, либо справа от них есть еще хотя бы два числа. Они и образуют нужную тройку либо с числом a, либо с числом b. ■

Из любых ли девяти различных чисел, выписанных в ряд, можно выбрать четыре, стоящие в этом ряду в порядке убывания или в порядке возрастания? Нет. Пример — последовательность 3, 2, 1, 6, 5, 4, 9, 8, 7. А вот из любых десяти различных чисел, выписанных в ряд, четыре числа, расположенные в порядке убывания или в порядке возрастания, выделить можно. Почему? Скоро узнаете.

Из любых ли $24 = 4 \cdot 6$ разных чисел, выписанных в ряд, можно выбрать 5 $(= 4 + 1)$ чисел в порядке возрастания или 7 $(= 6 + 1)$ в порядке убывания? Нет. Пример — 6, 5, 4, 3, 2, 1, 12, 11, 10, 9, 8, 7, 18, 17, 16, 15, 14, 13, 24, 23, 22, 21, 20, 19. А вот из любых 25 — можно. Почему? Скоро узнаете!

Вообще, как бы мы ни переставляли $mn + 1$ различных чисел, среди них обязательно найдется или возрастающая последовательность длины $m + 1$, или убывающая последовательность длины $n + 1$. Почему? Скоро узнаете!!! А пока познакомьтесь с проблемами одного сказочного короля. ■

Король пригласил на пир всех людоедов своей страны. Среди них есть людоеды, которые хотят съесть других людоедов. Известно, что наидлиннейшая цепочка, в которой первый людоед хочет съесть второго, второй — третьего и так далее, состоит из n людоедов. Докажем, что король может так рассадить людоедов за n столов, что ни за каким столом никто не будет желать скушать никого из сидящих за тем же столом. ■

Как это связано с четырехчленными монотонными подпоследовательностями? А вот как. Будем говорить, что число b хочет съесть число a,

Вообразите, что после уроков вы с приятелем задержались в школе. Никого уже не было, и вы радовались тишине. Вдруг из ему одному ведомого места выбежал некто и начал крушить все, что может. Сначала вы растерялись, но потом, чтобы его остановить, схватили что под руку попалось, — и тут некстати видите спешащего на звук битого стекла директора школы.

Хулигана не догонишь. Директор видел только вас. С опущенными от незаслуженных обвинений головами вы сидите в кабинете директора: «Я своими ушами слышал, как били стекла. Если вы оба откажетесь признать свою вину, то получите по замечанию в дневник за то, что бегали по школе. Если один из вас признает вину, а другой нет, я прощу признавшегося и даже дам ему конфетку, а другому придется заплатить за весь ремонт и расстаться с нашей школой. Признаетесь оба — пополам оплатите ремонт и забудем эту историю».

«Понятно», — отвечаете вы хором, и директор разводит вас по разным классам, чтобы вы обдумали его предложение. Посоветоваться друг с другом вы не можете, и решение надо принимать самостоятельно. Признать или отрицать вину?

Предавать не годится. Поэтому постараемся отвлечься от моральных соображений, «выключим» голос совести, займемся только задачей. Пусть вы оба заинтересованы отделаться как можно легче. Рассуждаем логически: «Если приятель решит признать вину, то признавать нужно и мне — иначе выгонят из школы и заставят одного меня платить за все. Если он не признает вину, то мне тоже лучше признаться — тогда вообще не накажут и даже конфетку дадут. В любом случае — лучше сознаться!» Вы идете к директору, придумывая, как именно били стекло. А ваш приятель тоже пришел к выводу, что лучше признаться. В результате — платите за ремонт.

А если бы не сознались, то отделались бы замечанием. Правильные рассуждения привели к нежелательному исходу. Это и есть знаменитый парадокс, известный как «дилемма заключен-

ного» (обычно рассказывают историю про двух заключенных, сидящих в разных камерах, а роль директора играет следователь).

Если вам кажется, что вопрос этот не очень серьезный — какое дело нам, свободным людям, до каких-то заключенных! — подумайте о другом варианте этого парадокса. На сей раз в нем участвуют две враждующие державы, которые могут нанести ракетно-ядерный удар. Если нападают одновременно, обе несут огромный ущерб. Если одна страна нападает, а другая не готова к атаке, то начавшая войну первой побеждает и (через несколько тысяч лет, когда спадет радиоактивность) сможет воспользоваться богатствами проигравшей страны. Наконец, если обе страны не нападают, то несут не очень значительный ущерб — всего лишь надо платить зарплату военным и модернизировать вооружения. Неужели и в этом случае логика обеим сторонам подсказывает выбор агрессивного поведения? Об этом даже думать страшно. . .

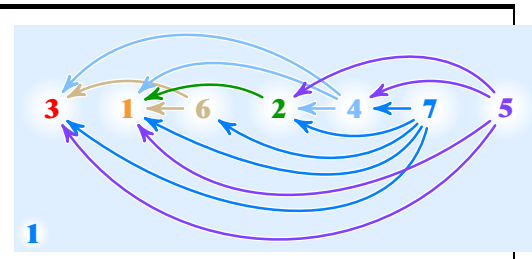
Парадокс заключенного можно сформулировать и на языке теории игр. Два игрока одновременно делают один из двух возможных ходов: «признать» (коротко «**П**») или «отрицать» («**О**»). Если оба делают ход **П**, то получают мало; если оба выбирают ход **О**, то получают существенно больше; если же один игрок делает ход **П**, а другой — **О**, то первый получает еще больше, а второй не получает ничего хорошего. Все это можно записать в виде таблицы. Величины выигрышей (числа 1, 3, 5) выбраны более или менее произвольно. Из таблицы видно, что при любом ходе противника игрок получит больше, если сделает ход **П**, а не **О** — это и обеспечивает парадокс.

Ходы		Очки	
I	II	I	II
О	**О**	3	3
О	**П**	0	5
П	**О**	5	0
П	**П**	1	1

В 1979 г. мичиганец Р. Аксельрод исследовал дилемму заключенного с помощью компьютера.

(*Продолжение на с. 550.*)

если $a<b$ и в рассматриваемом ряду a расположено левее, чем b (на рисунке 1 это показано для последовательности 3, 1, 6, 2, 4, 7, 5; желание съесть показано стрелочкой). Тогда цепочка чисел, в которой каждое следующее хочет съесть предыдущее, — это возрастающая подпоследовательность. А сидящие за одним столом числа, ни одно из которых не хочет съесть никакое другое, — это убывающая подпоследовательность.



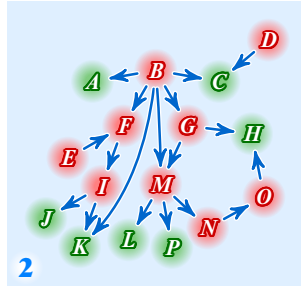
1

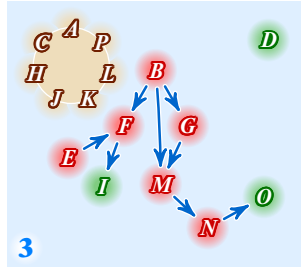
Разумеется, число n в условии задачи о людоедах надо не забыть заменить на число 3. И если нет убывающей подпоследовательности длины 4, то есть, если самая длинная цепочка, в которой каждое число хочет съесть следующее за ним, состоит не более чем из 3 чисел, то мы сможем рассадить 10 имеющихся чисел за тремя столами так, что ни за каким столом никто никого не будет хотеть скушать. Поскольку $3<10/3$, хотя бы за одним столом окажутся не менее чем 4 числа. Они образуют искомую возрастающую последовательность. ■

Приступим к доказательству теоремы о людоедах.

I способ. Обозначим их кружками и проведем стрелки, обозначающие, кто кого хочет съесть (на рисунке 2 изображена одна из возможных ситуаций). Если бы существовал цикл, то можно было бы, двигаясь вдоль стрелок, прийти из некоторой точки в нее саму, и длина цепочки людоедов, о которой сказано в условии, не была бы ограничена сверху: «крутясь по циклу», мы смогли бы построить сколь угодно длинную такую цепочку.

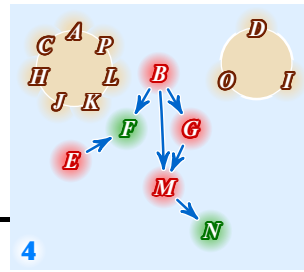
Значит, циклов нет. Теперь для каждого людоеда посмотрим, какие цепочки начинаются с него. (Быть может, его тоже кто-то хочет съесть, но мы на это не обращаем внимания и начинаем цепочку именно с него. Если же людоед — вегетарианец, то цепочка состоит только из него самого.) Количество людоедов в самой длинной из таких цепочек — это и есть номер стола, за который мы посадим этого людоеда. (Вы поняли, почему посаженные нами за один стол людоеды не хотят съесть один другого? Если нет, перечитайте этот абзац!)

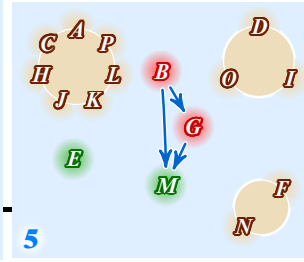

2

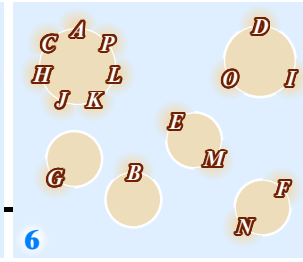

3

Возможно, однако, кому-то по душе больше придется другой, более формальный способ изложения (по сути того же самого) решения.

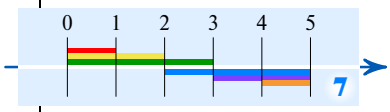
II способ. Всех людоедов-вегетарианцев, никого не желающих скушать (они отмечены на рисунке 2 зелеными кружками), посадим за первый стол (рис. 3). Исключим их из рассмотрения. Среди оставшихся людоедов есть свои вегетарианцы, которых мы посадим за второй стол (рис. 4). Исключив из рассмотрения и этих людоедов, обнаружим новых вегетариацев и посадим их за третий стол (рис. 5), и так далее. В конце концов все окажутся рассажены за n столов (рис. 6). ■


4


5


6

Теперь для разнообразия — задачи об отрезках прямой. Можно ли разместить на прямой а) 6, б) 7 отрезков так, чтобы из любых трех отрезков нашлись два пересекающихся и никакая точка прямой не принадлежала четырем или более отрезкам?



Ответ пункта а) очевиден: достаточно рассмотреть отрезки [0; 1], [0; 2], [0; 3], [2; 5], [3; 5] и [4; 5] (рис. 7).

Намного интереснее пункт б). Сначала рассмотрим решение, не использующее теорему о людоедах. Рассуждаем «от противного». Пусть удалось расположить семь отрезков так, что из любых трех отрезков некоторые два пересекаются и никакая точка не принадлежит четырем отрезкам сразу. Обозначим буквой B самый левый из правых (да-да, именно так!) концов рассматриваемых семи отрезков. Пусть AB — один из этих отрезков. Поскольку точка B принадлежит не более чем трем из семи отрезков, то отрезок AB пересекается не более чем с двумя другими отрезками. (Подумайте, почему это так! Суть в том, что никакой отрезок не лежит целиком левее точки B.)

Значит, существуют четыре отрезка, целиком расположенные правее точки B. Каждые два из них пересекаются. (Иначе вместе с отрезком AB два непересекающихся отрезка образовывали бы тройку отрезков, никакие два из которых не пересекаются.) Но если каждые два из четырех отрезков пересекаются, то все эти четыре отрезка имеют общую точку — таковой точкой является, например, самый левый из их правых концов.

Задача решена. А теперь смотрите. Пусть один отрезок хочет съесть другой, если тот лежит целиком левее. Тогда любая система отрезков, ни один из которых не хочет съесть ни один из отрезков этой системы, обязательно имеет общую точку — самый левый из их правых концов. ■

Частично упорядоченное множество M — это множество, для любых двух элементов a, b которого известно, находятся они в некотором отношении $<$ или нет. При этом должны быть выполнены следующие аксиомы:

— если $a<b$ и $b<c$, то $a<c$;

— если $a<b$, то $a\neq b$;

— неравенства $a<b$ и $b<a$ не могут быть выполнены одновременно.

(Впрочем, третья аксиома, очевидно, следует из первых двух.) Множество, любые два элемента которого сравнимы, то есть множество, для любых элементов a и b которого выполнено одно из соотношений $a<b$, $a=b$ или $b<a$, называют линейно упорядоченным или, коротко, цепью. Элемент a называют максимальным (соответственно, минимальным), если неравенство $a<b$ (соответственно, $b<a$) не выполнено ни для какого другого элемента b.

Приведем несколько примеров частично упорядоченных множеств: множество всех вещественных чисел (оно не только частично, но и даже линейно упорядочено, то есть любые два числа можно сравнить); множество всех отрезков прямой, где один отрезок больше другого, если все точки первого лежат правее точек второго; множество всех натуральных чисел, упорядоченное по делимости (то есть одно число больше или равно другому, если первое число делится на второе без остатка); множество всех подмножеств данного множества, упорядоченное по включению (то есть одно множество больше или равно другому, если все элементы второго множества являются и элементами первого множества).

Обобщением доказанных выше утверждений о числах и отрезках является следующая теорема.

Теорема. ***В частично упорядоченном множестве из $mn+1$ элементов есть либо цепь из идущих в порядке возрастания $m+1$ элементов, либо $n+1$ попарно несрав-***

Он обратился с предложением провести турнир компьютерных программ в описанную выше игру. На его призыв откликнулись 14 человек. Турнир был организован так: каждая программа играла 5 серий по 200 игр с другими программами, в том числе сама с собой. Победителем турнира объявлялась программа, набравшая в сумме наибольшее количество очков (заметьте: цель не в том, чтобы победить другие программы, а в том, чтобы набрать в сумме побольше очков).

Принципы игры программ могли быть любыми (например, очередной ход **П** или **О** можно выбирать случайно), но большинство программ помнили все предыдущие ходы противника и выбирали очередной ход, анализируя эту информацию. Авторы старались составить программы так, чтобы поощрять «сотрудничество» противника (то есть выбор им хода **О**), да к тому же зарабатывать очки, «предавая».

Прежде чем сказать, чем закончилось это состязание алгоритмов, напомним рассказ Э. По «Украденное письмо»: «Я знал одного восьмилетнего мальчика, который изумлял всех своим искусством играть в «чет и нечет». Игра эта очень простая: один из играющих зажимает в руке несколько шариков, а другой должен угадать, четное их число или нечетное. Если угадает — получит один шарик. Если нет — должен отдать шарик противнику. Мальчик, о котором я говорю, обыгрывал всех в школе. Разумеется, у него был свой метод отгадывания, основанный на простой наблюдательности и оценке сообразительности партнеров. Например, играет с ним какой-нибудь простофиля, зажал в руке шарики и спрашивает: «Чет или нечет?» Наш игрок отвечает: «Нечет» — и проигрывает. Но в следующий раз уже выигрывает, ибо он рассуждает так: «Простофиля взял в первый раз четное число, хитрости у него хватит как раз настолько, чтобы взять теперь нечет, — поэтому я должен сказать нечет». Он говорит: «Нечет» — и выигрывает. Имея дело с партнером немного поумнее, он рассу-

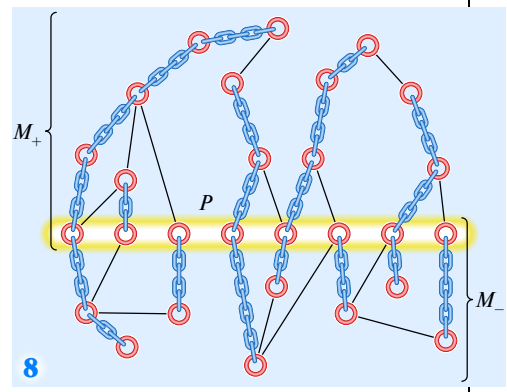
ждал так: «В первый раз я сказал нечет; помня это, он будет рассчитывать (как и первый), что в следующий раз я скажу чет, и, стало быть, ему следует взять нечет. Но он тотчас сообразит, что это слишком простая хитрость, и решится взять чет».
Результат турнира оказался не в пользу византийских хитрецов: выиграла самая простая из программ. Ее придумал А. Раппопорт из Торонто. Играет она так:
— первый ход — **О**;
— каждый следующий ход — это предыдущий ход противника.
Вот и все!
Программа «Ты — мне, я — тебе» не способна выиграть ни одной игры; лучшее, на что она может надеяться, — это ничья. Действительно, пока противник отвечает на ее ход **О** своими **О**, обе программы набирают по 3 очка. Как только противник «обманывает» программу «Ты — мне, я — тебе», сделав ход **П**, она мстит за это ходом **П**, в лучшем случае возвращая себе потерянные очки. Но программа «Ты — мне, я — тебе» незлопамятна: как только противник «раскаялся», сделав ход **О**, она отвечает тем же «доброжелательным» ходом.
Почему же «Ты — мне, я — тебе» побеждает в турнире? Рассмотрим мини-соревнование, в котором участвует еще одна программа «Предатель», всегда делающая ход **П**. Пусть каждая игра состоит из 10 ходов. Всего проведем три игры: «Ты — мне, я — тебе» против «Предателя», «Ты — мне, я — тебе» против самой себя и, наконец, «Предатель» сам с собой.
«Ты — мне, я — тебе» набирает в первой и второй играх, соответственно, 9 и 30 очков. А «Предатель» в первой и третьей играх набирает 14 и 10 очков. Итого: «Ты — мне, я — тебе» побеждает с результатом 39 очков против 24 у «Предателя».
Если бы программа «Ты — мне, я — тебе» была человеком, ее можно было бы назвать доброжелательной, готовой к сотрудничеству, не прощающей предательства и незлопамятной. Локально эти свойства невыгодны. Глобально — именно они приводят к триумфу. Поучительно и для нас с вами, не правда ли? ■

нимых элементов (так называемая антицепь). Более того, если d — наибольшее количество элементов цепи конечного частично упорядоченного множества M, то M можно разбить на d антицепей. ■

Если поменять слова «цепь» и «антицепь» местами, то мы получим гораздо более трудно доказываемое свойство частично упорядоченных множеств: если n — наибольшее количество элементов антицепи данного конечного частично упорядоченного множества M, то M можно разбить на n цепей.

Теорема Р. П. Дилворта. ***Если N — наименьшее количество цепей, на которые можно разбить данное конечное частично упорядоченное множество M, а n — наибольшее количество его попарно несравнимых элементов (то есть наибольшая мощность антицепи), то $n = N$.***

Доказательство. Неравенство $n \leqslant N$ очевидно: никакие два элемента одной антицепи не могут войти в одну цепь.
Докажем неравенство $n \geqslant N$ индукцией по числу элементов множества M. База очевидна. Пусть неравенство верно для всех частично упорядоченных множеств, содержащих менее m элементов. Рассмотрим частично упорядоченное множество M, состоящее из m элементов.



8

Предположим, в нем есть антицепь P, не содержащая ни некоторого минимального элемента $a \in M$, ни некоторого максимального элемента $b \in M$. Обозначим

$$M_+ = \{s \in M \mid \exists\, p \in P\ (p \leq s)\},$$

$$M_- = \{s \in M \mid \exists\, p \in P\ (s \leq p)\}.$$

Эти две формулы означают, что M_+ — множество элементов $s \in M$, для которых (в формуле последние два слова обозначены вертикальной чертой) существует (а это слово принято обозначать квантором $\exists$ — это перевернутая первая буква слова Exist — существовать) $p \in P$ такой, что $p < s$ или $p = s$; аналогично M_- — это множество элементов $s \in M$, для которых существует $p \in P$ такой, что $s < p$ или $p = s$. Учитывая предположение об антицепи P, имеем

$$M_+ \neq M, \qquad M_- \neq M, \qquad M = M_- \cup M_+.$$

По предположению индукции, каждое из множеств M_+ и M_- можно разложить на n цепей; «склеивая» эти цепи в точках, принадлежащих P, получаем разложение множества P на n цепей. (На рисунке 8 показана идея этого рассуждения: множество M разбито на «верхнюю» и «нижнюю» части, пересекающиеся в точности по максимальной антицепи P. Каждая из этих частей разбита на цепи, которые, будучи «склеены» между собой, дают разбиение множества M на цепи.)
Предположим теперь, что каждая антицепь содержит либо все максимальные элементы, либо все минимальные элементы множества M. Поскольку содержащая все минимальные (или все максимальные) элементы антицепь не может содержать ни одного другого элемента, то множество M имеет не более двух антицепей (и если их две, то одна антицепь состоит из всех минимальных, а другая — из всех максимальных элементов множества M). Пусть a и b — минимальный и максимальный элементы множества M, причем $a \leq b$. По индуктивному предположению, множество $M \setminus \{a, b\}$ можно разложить не более чем на $n - 1$ цепей. Добавляя цепь $a \leq b$, получаем разложение множества M не более чем на n цепей. ■

Какое наименьшее число гирь нужно иметь, чтобы можно было взвесить любой груз от 1 до 40 г?

Ответ — 4 гири. Оптимальный набор гирь — 1, 3, 9, 27 г. Чтобы взвесить груз в n граммов, надо представить число n в виде суммы

$$a_0+3a_1+9a_2+27a_3,$$

где $a_k=0$, 1 или -1 при $k=0$, 1, 2 или 3.

Как найти такую сумму? Один из возможных способов основан на сведении ее к представлению числа $n+40$ в троичной системе, то есть в виде суммы

$$b_0+3b_1+9b_2+27b_3,$$

где $b_k=a_k+1$, $k=0$, 1, 2 или 3.

Уравновешенная троичная система конкурирует с двоичной как по простоте арифметических алгоритмов, так и по количеству применений в математике. Для изменения знака числа достаточно изменить знаки у всех его цифр. Положительное число или отрицательное, определяет знак старшей ненулевой цифры, поэтому не нужен специальный знаковый бит, как в двоичной или десятичной системах. Для округления достаточно отбросить «лишние» цифры. В неуравновешенной системе, даже двоичной, округление выполняется не столь просто. При сравнении чисел по величине тоже не нужно отдельно рассматривать знак числа. Вычитание сводится к сложению сменой знака у вычитаемого. При записи чисел удобно вместо -1 писать $\overline{1}$.

$+$	$\overline{1}$	0	1
$\overline{1}$	$\overline{1}1$	$\overline{1}$	0
0	$\overline{1}$	0	1
1	0	1	$1\overline{1}$

$\times$	$\overline{1}$	0	1
$\overline{1}$	1	0	$\overline{1}$
0	0	0	0
1	$\overline{1}$	0	1

Троичная уравновешенная система была положена в основу советского компьютера «Сетунь», построенного в конце 1950-х гг. Однако широкого распространения он не получил, так как элементная база компьютеров (как того времени, так и современных) скорее двоичная, чем троичная. ■

Тождества и биекции

Во многих комбинаторных задачах оказываются полезны бесконечные ряды — производящие функции последовательностей. Л. Эйлер во «Введении в анализ бесконечных» — исторически первом учебнике математического анализа — писал: «...я вывел из того же источника [бесконечных рядов] решения многих вопросов, которые возникают при разбиении чисел [на слагаемые]; вопросы подобного рода без помощи этих приемов, видимо, превышают силы анализа».

В настоящее время в этих и многих других задачах, научились обходиться без бесконечных рядов. Почти все результаты доказаны в статье дважды: комбинаторно и при помощи производящих функций.

Раскрывая скобки, легко проверить тождества $(1+x)(1+x^2)=1+x+x^2+x^3$, $(1+x)(1+x^2)(1+x^4)=(1+x+x^2+x^3)(1+x^4)=1+x+x^2+x^3+x^4+x^5+x^6+x^7$ и вообще,

$$(1+x)(1+x^2)\dots(1+x^{2^n})=1+x+x^2+x^3+\dots+x^{2^{n+1}-1}.$$

Бесконечное произведение $(1+x)(1+x^2)(1+x^4)(1+x^8)(1+x^{16})\dots$ после раскрытия (бесконечного!) количества скобок превращается в ряд $1+x+x^2+x^3+x^4+x^5+x^6+x^7+\dots$, сходящийся при $|x|<1$.

Устремив n к бесконечности, получаем тождество

$$\prod_{k=0}^{\infty}(1+x^{2^k})=\sum_{m=0}^{\infty}x^m.$$

Оно означает, что любое натуральное число единственным образом представимо в двоичной системе счисления. Аналог для десятичной системы счисления выглядит так:

$$\prod_{k=0}^{\infty}(1+x^{10^k}+x^{2\cdot10^k}+x^{3\cdot10^k}+x^{4\cdot10^k}+x^{5\cdot10^k}+x^{6\cdot10^k}+x^{7\cdot10^k}+x^{8\cdot10^k}+x^{9\cdot10^k})=\sum_{m=0}^{\infty}x^m.$$

Это тождество следует из того, что для любого натурального n имеем

$$(1+x+x^2+\dots+x^9)(1+x^{10}+x^{20}+\dots+x^{90})\dots(1+x^{10^n}+x^{2\cdot10^n}+\dots+x^{9\cdot10^n})=1+x+\dots+x^{10^{n+1}-1}.$$ ■

Для троичной системы счисления аналогичное тождество можно записать в довольно неожиданном виде

$$(x^{-1}+1+x)(x^{-3}+1+x^3)\dots(x^{-3^n}+1+x^{3^n})=x^{-1}\,\frac{x^3-1}{x-1}\,x^{-3}\,\frac{x^9-1}{x^3-1}\dots$$

$$\dots x^{-3^n}\,\frac{x^{3^{n+1}}-1}{x^{3^n}-1}=x^{-N}\,\frac{x^{3^{n+1}}-1}{x-1}=x^{-N}+x^{-N+1}+\dots+x^{N-1}+x^N,$$

где $N=1+3+9+\dots+3^n=(3^{n+1}-1)/2$.

Мы доказали, что при помощи гирь 1, 3, 9, 27, 81, ... можно, пользуясь обеими чашками весов, составить любой целый положительный вес, и притом лишь одним способом:

$1=1$, $4=3+1$, $7=9-3+1$, $10=9+1$,
$2=3-1$, $5=9-3-1$, $8=9-1$, $11=9+3-1$,
$3=3$, $6=9-3$, $9=9$, $12=9+3$,
. . . . . .

Ниже пойдет речь о подобных равенствах, только менее очевидных. Мы придадим многим из них «комбинаторный смысл». Очень полезно, вооружившись ручкой и бумагой, повторять выкладки: это поможет ощутить степень нетривиальности результатов. ■

Количество разбиений числа n на попарно различные слагаемые равно количеству разбиений на нечетные слагаемые, например,

$13=13$	$13=13$
$=12+1$	$=3+3+3+3+1$
$=11+2$	$=11+1+1$
$=10+3$	$=5+5+3$
$=10+2+1$	$=5+5+1+1+1$
$=9+4$	$=9+1+1+1+1$
$=9+3+1$	$=9+3+1$
$=8+5$	$=1+1+1+1+1+1+1+1+5$
$=8+4+1$	$=1+1+1+1+1+1+1+1+1+1+1+1+1$
$=8+3+2$	$=1+1+1+1+1+1+1+1+3+1+1$
$=7+6$	$=7+3+3$
$=7+5+1$	$=7+5+1$
$=7+4+2$	$=7+1+1+1+1+1+1$
$=7+3+2+1$	$=7+3+1+1+1$
$=6+5+2$	$=3+3+5+1+1$
$=6+4+3$	$=3+3+1+1+1+1+3$
$=6+4+2+1$	$=3+3+1+1+1+1+1+1+1$
$=5+4+3+1$	$=5+1+1+1+1+3+1$

Каждому разбиению на попарно различные слагаемые сопоставим разбиение на нечетные слагаемые, представив каждое слагаемое в виде произведения $2^a(2b+1)$, где a, b — целые неотрицательные числа, и заменив его суммой 2^a слагаемых, равных $2b+1$. Например, разбиению

$$4069=1+4+5+6+20+48+1992+1993$$

сопоставим

$$4069=1+\underbrace{1+1+1+1}+5+\underbrace{3+3}+\underbrace{5+5+5+5}+$$
$$+\underbrace{3+3+3+3+3+3+3+3+3+3+3+3+3+3+3+3}+$$
$$+\underbrace{249+249+249+249+249+249+249+249}+1993.$$

Обратно, каждому разбиению на нечетные слагаемые сопоставим разбиение на попарно различные слагаемые, подсчитав для каждого нечетного слагаемого $2b+1$, сколько раз оно входит в сумму, представив это количество N в двоичной системе счисления (то есть в виде суммы различных степеней числа 2)

$$N=2^{a_1}+2^{a_2}+\ldots+2^{a_m}$$

и заменив N слагаемых $(2b+1)$ на сумму

$$2^{a_1}(2b+1)+2^{a_2}(2b+1)+\ldots+2^{a_m}(2b+1).$$

Например, разбиению

$$6087=\underbrace{1+1+1+1+1+1+1}+\underbrace{3+3}+\underbrace{7+7+7+7+7+7}+\underbrace{17+17}+19+$$
$$+\underbrace{1993+1993+1993}$$

сопоставим

$$6087=\underbrace{4+2+1}+\underbrace{6}+\underbrace{28+14}+\underbrace{34}+19+\underbrace{3986+1993}.$$

Мы построили биекцию (взаимно-однозначное соответствие) между разбиениями на нечетные слагаемые и разбиениями на попарно различные слагаемые. (Впервые это сделал английский математик Дж. Глэйшер в конце XIX в.) ■

Придворный астролог называет момент времени хорошим, если часовая, минутная и секундная стрелки часов находятся по одну сторону от какого-нибудь диаметра циферблата. (Стрелки вращаются на общей оси без скачков.)

Рассмотрим моменты времени, различающиеся на 6 часов. Показания часов отличаются лишь положением часовой стрелки, которая меняет свое направление на противоположное. Поэтому в сутках хорошего и плохого времени поровну. ■

Коэффициент при ζ^n, образующийся после раскрытия скобок бесконечного произведения

$$(1+q\zeta)(1+q\zeta^2)(1+q\zeta^4)\times$$
$$\times(1+q\zeta^8)(1+q\zeta^{16})\ldots,$$

равен $q^{E(n)}$, где $E(n)$ — количество единиц в двоичной записи числа n.

Знак $(n+1)$-го члена в разложении произведения

$$(1-a)(1-b)(1-c)(1-d)\ldots=$$
$$=1-a-b+ab-c+ac+bc-$$
$$-abc-d+\ldots$$

равен $(-1)^{E(n)}$. ■

Каждое натуральное n ровно 2^{n-1} способами представимо в виде суммы меньших натуральных слагаемых, если представления, отличающиеся хотя бы порядком слагаемых, считать различными и разрешить сумме состоять лишь из одного слагаемого. Например, для $n=4$ имеем $2^{4-1}=8$ способов: $1+1+1+1$, $1+1+2$, $2+2$, $1+3$, $1+2+1$, $3+1$, $2+1+1$, 4.

Доказательство. Написав сумму n единиц, сопоставим каждому подмножеству множества использованных знаков разбиение числа n на слагаемые, выполнив все вошедшие в это подмножество действия сложения. ■

Равенство

$$\sum_{k=0}^{m} C_r^k\cdot C_s^{m-k}=C_{r+s}^m,$$

где $m\leqslant r\leqslant s$, можно вывести алгебраически из равенства $(1+x)^r\times(1+x)^s=(1+x)^{r+s}$, а можно комбинаторно, рассмотрев количество способов выбора m человек из r мужчин и s женщин.

В частности, $\sum\limits_{k=0}^{m}(C_s^k)^2=C_{2s}^s$. ■

Джеймс Джозеф Сильвестр (1814—1897) в 1878 г. основал и поныне существующий «The American Journal of Mathematics». Изобрел пантограф — приспособление в виде шарнирного параллелограмма для перечерчивания (копирования) чертежей в измененном масштабе. В 1852 г. в статье «Доказательство теоремы о том, что всякий однородный квадратичный полином приводится вещественной ортогональной подстановкой к сумме положительных или отрицательных квадратов» доказал закон инерции квадратичных форм — независимость количеств знаков + и знаков − в диагональном виде квадратичной формы. (Несколько ранее это доказал К. Г. Я. Якоби, но доказательство не опубликовал.)
Занимался теорией инвариантов — наукой, отвечающей на следующий вопрос: «Геометрический объект задан уравнениями в некоторой системе координат. Какие алгебраические характеристики уравнений, например, какие функции их коэффициентов — инвариантны (то есть независимы) от выбора системы координат?» Сильвестру принадлежат основные термины этой теории: инвариант, дискриминант, ковариант. Роль Дж. Буля, Дж. Сальмона и Сильвестра в создании теории инвариантов столь велика, что Ш. Эрмит называл их «инвариантной троицей». Уточнил постулат Бертрана, доказав, что для всех достаточно больших n между n и $1{,}092n$ есть хотя бы одно простое число; усилил и результаты П. Л. Чебышёва об отношении $\frac{\pi(x)}{x/\ln x}$, доказав, что при достаточно больших x оно заключено между 0,95695 и 1,04423. ■

Коэффициент при x^n в степенном ряде $\prod_{k=1}^{\infty}(1+x^k)$ равен количеству разбиений n на попарно различные слагаемые. (Если это утверждение неочевидно, подумайте, как раскрываются «все скобки сразу».) Поскольку

$$\frac{1}{1-x}=1+x^1+x^{1+1}+x^{1+1+1}+x^{1+1+1+1}+x^{1+1+1+1+1}+\dots,$$
$$\frac{1}{1-x^3}=1+x^3+x^{3+3}+x^{3+3+3}+x^{3+3+3+3}+x^{3+3+3+3+3}+\dots,$$
$$\frac{1}{1-x^5}=1+x^5+x^{5+5}+x^{5+5+5}+x^{5+5+5+5}+x^{5+5+5+5+5}+\dots,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

в степенном ряде $\prod_{k=1}^{\infty}\frac{1}{1-x^{2k-1}}$ коэффициент при x^n равен количеству разбиений n на нечетные слагаемые. Значит, $\prod_{k=1}^{\infty}\frac{1}{1-x^{2k-1}}=\prod_{k=1}^{\infty}(1+x^k)$. ■

Это тождество можно доказать и алгебраически. Поскольку $(1-x)\times$ $\times(1+x+\dots+x^n)=1-x^{n+1}$, то $(1-x)(1+x+\dots+x^n+x^{n+1}+\dots)=1$. (Это равенство выполнено не только в смысле равенства формальных рядов, то есть совпадения соответствующих коэффициентов после раскрытия скобок; при $|x|<1$ имеем $\lim\limits_{n\to\infty}x^{n+1}\to 0$, поэтому формула суммы бесконечной геометрической прогрессии верна для любого $|x|<1$.) Следовательно,

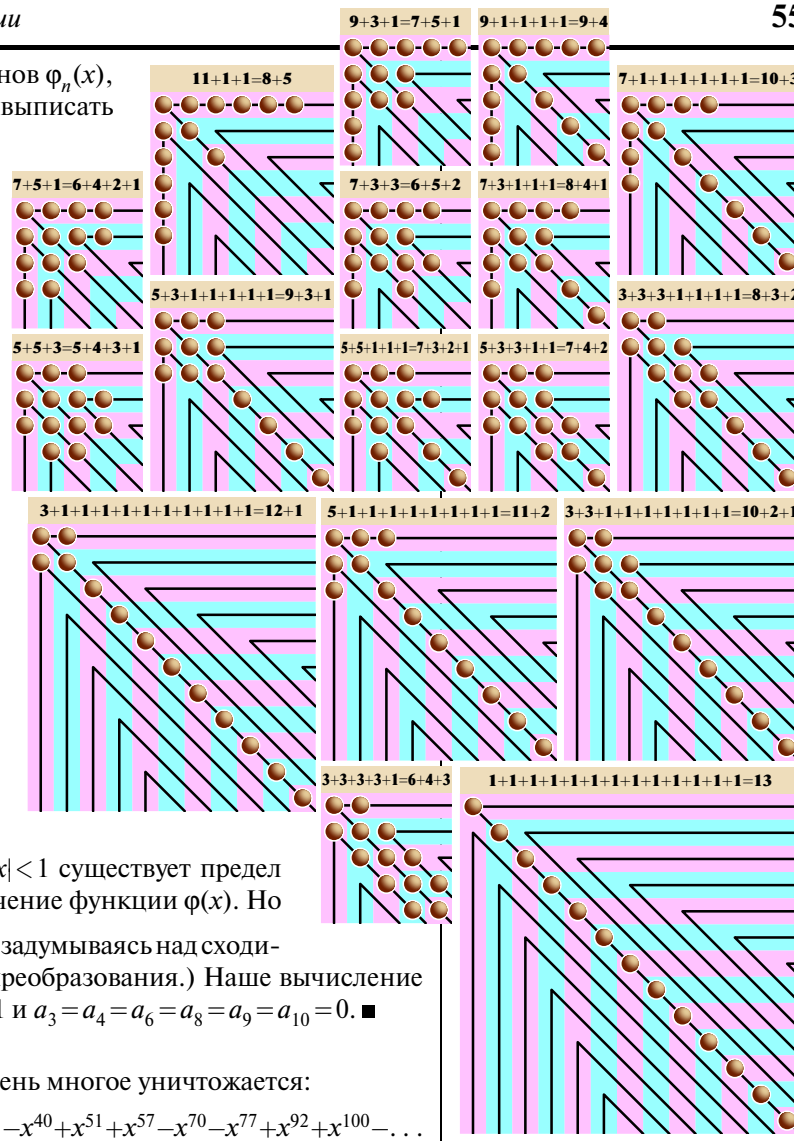
$$\frac{1}{1-x}=(1+x)(1+x^2)(1+x^4)(1+x^8)(1+x^{16})\dots,$$
$$\frac{1}{1-x^3}=(1+x^3)(1+x^6)(1+x^{12})(1+x^{24})(1+x^{48})\dots,$$
$$\frac{1}{1-x^5}=(1+x^5)(1+x^{10})(1+x^{20})(1+x^{40})(1+x^{80})\dots,$$
$$\frac{1}{1-x^7}=(1+x^7)(1+x^{14})(1+x^{28})(1+x^{56})(1+x^{112})\dots,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

Перемножая, получаем в левой части $\prod_{n=1}^{\infty}\frac{1}{1-x^{2n-1}}$. А в правой части — $\prod_{n=1}^{\infty}(1+x^n)$, поскольку каждое натуральное число n можно представить — причем единственным образом — в виде произведения нечетного числа и степени двойки. ■

А вот более короткое доказательство: $(1+x)(1+x^2)(1+x^3)(1+x^4)\dots=$ $=\frac{1-x^2}{1-x}\cdot\frac{1-x^4}{1-x^2}\cdot\frac{1-x^6}{1-x^3}\cdot\frac{1-x^8}{1-x^4}\cdot\dots=\frac{1}{(1-x)(1-x^3)(1-x^5)\dots}$. ■

В произведении $\varphi_n(x)=(1-x)(1-x^2)(1-x^3)\dots(1-x^n)$ Эйлер раскрыл скобки и получил поразительный результат:

$$\begin{aligned}
\varphi_1&=1-x,\\
\varphi_2&=1-x-x^2+x^3,\\
\varphi_3&=1-x-x^2\quad+x^4+x^5-x^6,\\
\varphi_4&=1-x-x^2\quad+2x^5\quad-x^8-x^9+x^{10},\\
\varphi_5&=1-x-x^2\quad+x^5+x^6+x^7-x^8-x^9-x^{10}\dots,\\
\varphi_6&=1-x-x^2\quad+x^5\quad+2x^7\quad-x^9-x^{10}\dots,\\
\varphi_7&=1-x-x^2\quad+x^5\quad+x^7+x^8\quad-x^{10}\dots,\\
\varphi_8&=1-x-x^2\quad+x^5\quad+x^7\quad+x^9\quad\dots,\\
\varphi_9&=1-x-x^2\quad+x^5\quad+x^7\quad+x^{10}\dots,\\
\varphi_{10}&=1-x-x^2\quad+x^5\quad+x^7\quad\dots\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}$$



Многоточия обозначают части многочленов $\varphi_n(x)$, содержащие x в степенях, бóльших 10 (выписать эти многочлены полностью не позволяет формат бумаги: многочлен $\varphi_{10}(x)$, например, имеет степень 55).

Начнем с очевидного, но важного наблюдения: коэффициенты многочлена $\varphi_n(x)$ с ростом n стабилизируются, то есть каждый из них начиная с некоторого n не меняется. Это легко понять: переход от $\varphi_{n-1}(x)$ к $\varphi_n(x)$, состоящий в умножении на $1-x^n$, не меняет коэффициенты при $1, x, \ldots, x^{n-1}$, так что при $n \geqslant k$ коэффициент при x^k в многочлене $\varphi_n(x)$ от n не зависит (например, вычисленная часть многочлена $\varphi_{10}(x)$ не изменится, если вместо φ_{10} взять φ_{11}, φ_{12} и так далее.). Поэтому можно говорить о бесконечном произведении

$$\varphi(x) = (1-x)(1-x^2)(1-x^3)(1-x^4)\ldots,$$

понимая под ним, конечно, не многочлен, а степенной ряд, то есть выражение вида $a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots$, где $a_0, a_1, a_2, a_3, \ldots$ — числа, в нашем случае это стабилизировавшиеся коэффициенты. (При $|x|<1$ существует предел $\lim\limits_{n\to\infty} \varphi_n(x)$, так что можно определить значение функции $\varphi(x)$. Но нам это не понадобится, так что лучше, не задумываясь над сходимостью рядов, выполнять формальные преобразования.) Наше вычисление показывает, что $a_0 = a_5 = a_7 = 1$, $a_1 = a_2 = -1$ и $a_3 = a_4 = a_6 = a_8 = a_9 = a_{10} = 0$. ■

После раскрытия скобок очень многое уничтожается:

$$\varphi(x) = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - x^{35} - x^{40} + x^{51} + x^{57} - x^{70} - x^{77} + x^{92} + x^{100} - \ldots$$

Надо полагать, не боявшийся длинных выкладок Эйлер примерно столько членов ряда $\varphi(x)$ и вычислил. Он не мог не заметить, что ненулевые коэффициенты расположены в строгом порядке: две единицы, две минус единицы, две единицы, две минус единицы и так далее. В таблице

показатели	1, 2	5, 7	12, 15	22, 26	35, 40	51, 57	70, 77	92, 100
коэффициенты	−1	1	−1	1	−1	1	−1	1

выписаны показатели степеней, коэффициенты при которых ненулевые.

Легко угадать, что в n-м столбце таблицы в верхней строке стоят числа $(3n^2 \pm n)/2$, в нижней — число $(-1)^n$. Если это так при всех n, то

$$(1-x)(1-x^2)(1-x^3)\ldots = 1 - x - x^2 + x^5 + x^7 - \ldots + (-1)^n x^{(3n^2-n)/2} + (-1)^n x^{(3n^2+n)/2} + \ldots$$

или, короче, $\prod\limits_{n=1}^{\infty}(1-x^n) = \sum\limits_{n=-\infty}^{+\infty}(-1)^n x^{(3n^2+n)/2}$. Это — пентагональное тождество Эйлера: число n столькими же способами представимо в виде суммы четного числа различных слагаемых, сколькими и в виде суммы нечетного числа различных слагаемых, если только n не имеет вид $(3k^2 \pm k)/2$, где k — натуральное число. Если $n = (3k^2 \pm k)/2$, то разложений в сумму четного числа различных слагаемых на $(-1)^k$ больше.

Биекцию между разбиениями на нечетные слагаемые и разбиениями на различные слагаемые, отличную от биекции Дж. Глэйшера, придумал Дж. Дж. Сильвестр.

Нарисуем каждое нечетное число $2n+1$ как «уголок». Разбиение на нечетные слагаемые нарисуем, упорядочив их по величине и нарисовав сначала самое большое, затем — со сдвигом на 1 вниз и на 1 вправо — следующее по величине и так далее. Затем проведем ломаные, как показано на рисунках, и подсчитаем количества кружочков на них.

Нетрудно доказать, что эти количества на разных линиях разные. Можно научиться и восстанавливать по любому разбиению числа на различные слагаемые соответствующее разбиение на нечетные слагаемые. ■

На окружности отмечены 100 точек $A_1, A_2, \ldots, A_{100}$. Каких выпуклых многоугольников с вершинами в отмеченных точках больше: тех, у которых A_1 является вершиной, или остальных? На сколько?

К множеству вершин любого из многоугольников, у которого A_1 не является вершиной, можно добавить точку A_1. Поскольку треугольник с вершиной в точке A_1 так получить нельзя, получаем ответ: на $C_{99}^2 = 99\cdot 98/2$ больше многоугольников, у которых A_1 — вершина. ■

Докажем тождество

$$\prod_{k=1}^{\infty}(1-x^k) = \exp\left(-\sum_{n=1}^{\infty}\frac{\sigma(n)}{n}t^n\right),$$

где $\sigma(n)$ — сумма делителей числа n.

Поскольку

$$\ln(1-x) = -x-\frac{x^2}{2}-\frac{x^3}{3}-\frac{x^4}{4}-\ldots,$$

имеем:

$$\ln\prod_{k=1}^{\infty}(1-x^k) = \sum_{k=1}^{\infty}\ln(1-x^k) =$$

$$= -\sum_{k=1}^{\infty}\sum_{m=1}^{\infty}\frac{x^{km}}{m} = -\sum_{n=1}^{\infty}\sum_{km=n}\frac{kx^n}{n} =$$

$$= -\sum_{n=1}^{\infty}\left(\frac{x^n}{n}\sum_{n:k}k\right) =$$

$$= -\sum_{n=1}^{\infty}\frac{\sigma(n)x^n}{n}. \blacksquare$$

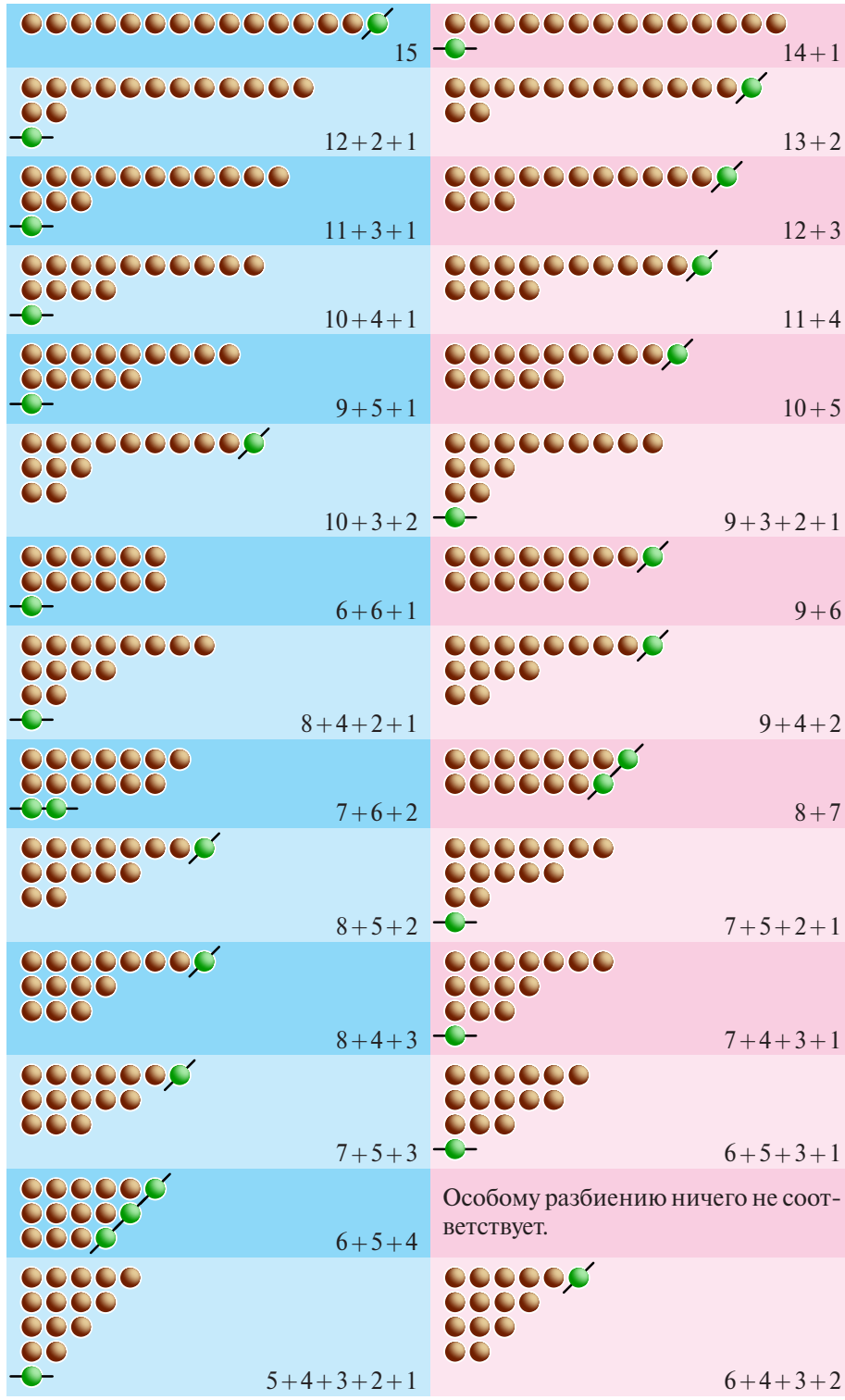
Как известно, любое натуральное число является числом Фибоначчи или суммой двух или более разных чисел Фибоначчи. Верны следующие утверждения.

а) Любое натуральное число $n \geqslant 3$ можно, причем единственным образом, представить в виде суммы различных чисел Фибоначчи, которая вместе с каждым слагаемым φ_k $(k \geqslant 4)$ содержит хотя бы одно из двух предыдущих чисел Фибоначчи: φ_{k-1} или φ_{k-2}.

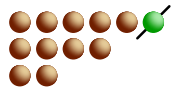
б) Обозначим количество представлений числа n в виде суммы четного количества чисел Фибоначчи через K_n, а в виде суммы нечетного количества — через H_n. Верно неравенство $|K_n - H_n| \leqslant 1$.

в) Более того, если перемножить несколько подряд стоящих двучленов из последовательности $1-x$, $1-x^2$, $1-x^3$, $1-x^5$, $\ldots$ $\ldots$, $1-x^{\varphi_n}$, $\ldots$, то в полученном многочлене все коэффициенты равны -1, 0 или 1. ■

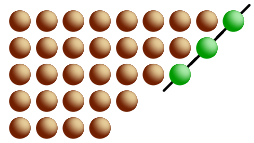
Для доказательства из крайней справа точки верхней строки проведем через точки нашей диаграммы насколько возможно длинную прямую с угловым коэффициентом 1; может случиться, разумеется, что она будет содержать только одну точку. Дальнейшее зависит от того, больше или меньше точек содержит нижняя строка, и ясно из таблицы разбиений числа 15.



Для описания этого соответствия на языке разбиений $n=n_1+n_2+\ldots+n_k$, $n_1>n_2>\ldots>n_k$, обозначим через s наибольшее такое число, что $n_s-n_1=s-1$, то есть s чисел n_1, n_2, …, n_s идут подряд. Например, для разбиения $12=6+4+2$ имеем $s=1$:



А для разбиения $33=9+8+7+5+4$ имеем $s=3$:



Мы скажем, что у разбиения $(n_1,\ldots,n_k)$
— *короткая наклонная*, если $s<n_k$, исключая случай $n_k=s+1$, $s=k$;
— *длинная наклонная*, если $n_k\leqslant s$, исключая случай $n_1=s=k$.
Особыми назовем исключенные разбиения, то есть те, где $s=k$ и $n_1=s$ или $n_1=s+1$.
Поставим теперь в соответствие разбиению $n_1+n_2+\ldots+n_k$ с короткой наклонной разбиение

$$(n_1-1)+(n_2-1)+\ldots+(n_s-1)+n_{s+1}+\ldots+n_k+s.$$

Разбиению $n_1+n_2+\ldots+n_k$ с длинной наклонной сопоставим

$$(n_1+1)+(n_2+1)+\ldots+(n_{n_k}+1)+n_{n_k+1}+\ldots+n_{k-1}.$$

Заметьте: сопоставления согласованы, а количества слагаемых в соответствующих друг другу разбиениях различаются на 1.
Осталось заметить, что особые разбиения из s слагаемых имеют вид

$$\frac{3s^2-s}{2}=(2s-1)+\ldots+(s+1)+s,$$

$$\frac{3s^2+s}{2}=(2s)+\ldots+(s+2)+(s+1).\ \blacksquare$$

Обозначим через $p(n)$ количество способов, которыми n можно представить в виде суммы натуральных слагаемых (при этом слагаемые в суммах могут повторяться, а представления, различающиеся лишь порядком слагаемых, считаем одинаковыми).
С каждым разбиением связана диаграмма Юнга: каждое слагаемое изображаем строкой из соответствующего количества точек, верхняя строка — самая длинная, ниже — остальные в порядке убывания.
Как вычислять $p(n)$? Повозившись, можно найти $p(10)=42$. А если нужно знать, скажем, $p(50)$? На помощь приходит тождество Эйлера.
Пусть

$$f(x)=1+p(1)x+p(2)x^2+p(3)x^3+\ldots=1+x+2x^2+3x^3+5x^4+7x^5+11x^6+\ldots$$

(Как и $\varphi(x)$, функция f определена при $|x|<1$. Но она интересует нас только как степенной ряд.)
Оказывается, ряды $\varphi(x)$ и $f(x)$ взаимно обратны, то есть

$$\varphi(x)f(x)=1.$$

Вы понимаете, что это значит? Степенные ряды можно перемножать:

$$(a_0+a_1x+a_2x^2+\ldots)(b_0+b_1x+b_2x^2+\ldots)=$$
$$=a_0b_0+(a_1b_0+a_0b_1)x+(a_0b_2+a_1b_1+a_2b_0)x^2+\ldots$$

Раскроем все скобки бесконечного произведения $G(z)=(1+qz)(1+q^2z)(1+q^3z)\ldots$ и приведем подобные слагаемые. Получим степенной ряд $G(z)=A_0+A_1z+A_2z^2+A_3z^3+\ldots$ Определим значения A_n и их комбинаторный смысл.
Очевидно,

$$G(z)=(1+qz)G(qz).$$

Сравнением коэффициентов при z^n в обеих частях этого функционального уравнения получаем для любого натурального n равенство

$$A_n(1-q^n)=A_{n-1}q^n,$$

откуда, так как $A_0=1$, находим

$$A_n=\frac{q^{1+2+\ldots+n}}{(1-q)(1-q^2)\ldots(1-q^n)}.$$

Следовательно,

$$\prod_{n=1}^{\infty}(1+q^nz)=$$
$$=\sum_{n=0}^{\infty}\frac{q^{n(n+1)/2}z^n}{(1-q)(1-q^2)\ldots(1-q^n)}.$$

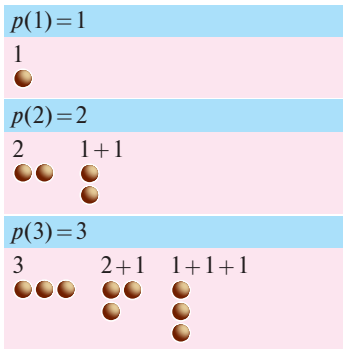
Заменив z на z/q и q на q^2, получим тождество

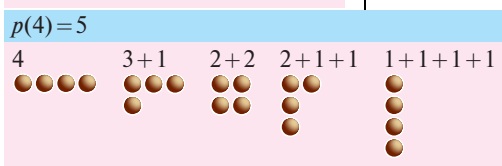
$$\prod_{k=1}^{\infty}(1+q^{2k-1}z)=$$
$$=\sum_{n=0}^{\infty}\frac{q^{n^2}z^n}{(1-q^2)(1-q^4)\ldots(1-q^{2n})}.$$

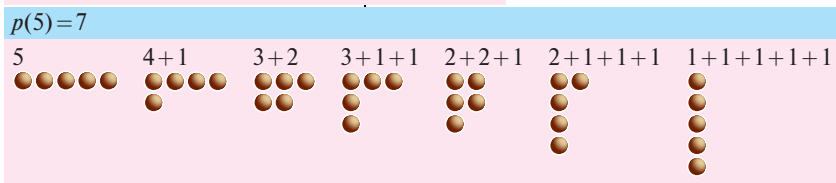
Комбинаторный смысл коэффициента при q^mz^n после раскрытия скобок левой части — количество разбиений числа m на n различных нечетных слагаемых. Чтобы понять смысл правой части, заметим: коэффициент при q^k после раскрытия скобок в выражении

$$\frac{1}{(1-q)(1-q^2)\ldots(1-q^n)}=$$
$$=(1+q+q^{1+1}+\ldots)\times$$
$$\times(1+q^2+q^{2+2}+\ldots)\ldots$$
$$\ldots(1+q^n+q^{n+n}+q^{n+n+n}+\ldots)$$

— это количество разбиений числа k на слагаемые, не превосходящие числа n.
Разбиение числа k вместе с сопряженным разбиением и квадратом $n\times n$ образуют самосопряженное (то есть симметричное относительно диагонали) разбиение числа n^2+2k. Поэтому коэффициент при q^mz^n правой части — это количество самосопряженных разбиений числа m с диагоналями длины n.
Биективное соответствие между левой и правой частями получаем, «распрямляя» уголок в нечетное число. $\blacksquare$







Наше утверждение означает, что если перемножить таким образом ряды $\varphi(x)$ и $f(x)$, то полученное произведение сведется к 1: коэффициенты при x, x^2, x^3, ... будут равны нулю. Докажем это:

$$\frac{1}{\varphi(x)}=\frac{1}{1-x}\cdot\frac{1}{1-x^2}\cdot\ldots\cdot\frac{1}{1-x^k}\cdot\ldots=$$
$$=(1+x+x^2+x^3+\ldots)(1+x^2+x^4+x^6+\ldots)\ldots(1+x^k+x^{2k}+x^{3k}+\ldots)\ldots$$

При раскрытии скобок в этом произведении получаем сумму всевозможных выражений вида $x^{a_1}x^{2a_2}\ldots x^{ka_k}$, где a_1, a_2,..., a_k — целые неотрицательные числа. Таким образом, x^n войдет в сумму столько раз, сколькими способами n можно представить в виде $a_1+2a_2+\ldots+ka_k$. Но это представление можно переписать так:

$$\underbrace{1+\ldots+1}_{a_1}+\underbrace{2+\ldots+2}_{a_2}+\ldots+\underbrace{k+\ldots+k}_{a_k}.$$

Видно, что представлений n в виде $a_1+2a_2+\ldots+ka_k$ столько же, сколько представлений n в виде суммы натуральных слагаемых, то есть $p(n)$. Таким образом, коэффициент при x^n равен $p(n)$, то есть $f(x)=1/\varphi(x)$.

Положив для удобства $p(0)=1$, напишем

$$(1-x-x^2+x^5+x^7-\ldots)(p(0)+p(1)x+p(2)x^2+\ldots)=1$$

(коэффициенты в первом сомножителе пишутся согласно тождеству Эйлера!). Раскроем скобки и приравняем коэффициенты при $x, x^2, \ldots, x^n$ в левой части:

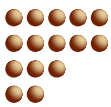
$$p(1)-p(0)=0;$$
$$p(2)-p(1)-p(0)=0;$$
$$p(3)-p(2)-p(1)=0;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$p(n)-p(n-1)-p(n-2)+p(n-5)+p(n-7)-\ldots=0$$

(в левой части последней формулы нужно писать слагаемые до тех пор, пока аргумент у p остается неотрицательным). Итак,

$$p(n)=p(n-1)+p(n-2)-p(n-5)-p(n-7)+\ldots$$

Эта формула позволяет быстро составить довольно длинную таблицу чисел $p(n)$. Вот практический совет, как это сделать. Возьмите лист клетчатой бумаги — лучше двойной тетрадный лист. Отрежьте вдоль его длинной стороны полоску шириной 3—4 клетки. Положите эту полоску перед собой вертикально и у левого среза в нижней клетке поставьте какой-нибудь знак, скажем, звездочку. Затем, двигаясь вверх, поставьте в первой клетке +, во второй +, в пятой −, в седьмой −, в двенадцатой +, в пятнадцатой + и так далее, насколько хватит длины полоски. Оставшуюся часть листа также положите перед собой вертикально и, отступив 10—15 клеток от ее левого среза, проведите вертикальную черту — сверху донизу. В клетки, прилегающие к черте слева, двигаясь сверху вниз, выпишите уже известные нам числа $p(n)$, начиная с $p(0)$: 1, 1, 2, 3, 5, 7. Чтобы найти следующее значение, приложите отрезанную полоску справа к вертикальной черте, чтобы звездочка оказалась против первой пустой клетки. Теперь из суммы чисел, стоящих против плюсов, вычтите сумму чисел, стоящих против минусов. Результат впишите в клетку против звездочки: это — следующее значение функции $p(n)$. Опустите полоску на одну клетку вниз и повторите то же самое. И так далее. Через несколько минут вы получите колонку чисел $p(n)$ высотой в лист. ■

Разбиению $15=5+5+3+2$ соответствует диаграмма



Начала строк образуют один столбец. Диаграмму можно читать также по столбцам, и тогда разбиение имеет вид $15=4+4+3+2+2$. Разбиения, связанные таким образом, назовем сопряженными. Очевидно, отношение сопряженности симметрично. Сопряженное разбиение для разбиения на k частей — это разбиение, наибольшая часть которого есть k; верно и обратное. Поэтому количество разбиений числа n на k частей равно количеству разбиений числа n на части, наибольшая из которых есть k. ■

Разность $p(n+1)-p(n)$ — это количество разбиений числа $n+1$ на части, ни одна из которых не равна 1.

Докажем неравенство

$$p(n+2)+p(n)\geqslant 2p(n+1).$$

Запишем его в виде

$$p(n+2)-p(n+1)\geqslant p(n+1)-p(n).$$

Прибавляя 1 к наибольшей части разбиения числа $n+1$ на части, бóльшие 1, получаем разбиение числа $n+2$ на части, бóльшие 1, не правда ли? ■

n	$p(n)$	n	$p(n)$	n	$p(n)$	n	$p(n)$
1	1	26	2436	51	239 943	76	9 289 091
2	2	27	3010	52	281 589	77	10 619 863
3	3	28	3718	53	329 931	78	12 132 164
4	5	29	4565	54	386 155	79	13 848 650
5	7	30	5604	55	451 276	80	15 796 476
6	11	31	6842	56	526 823	81	18 004 327
7	15	32	8349	57	614 154	82	20 506 255
8	22	33	10 143	58	715 220	83	23 338 469
9	30	34	12 310	59	831 820	84	26 543 660
10	42	35	14 883	60	966 467	85	30 167 357
11	56	36	17 977	61	1 121 505	86	34 262 962
12	77	37	21 637	62	1 300 156	87	38 887 673
13	101	38	26 015	63	1 505 499	88	44 108 109
14	135	39	31 185	64	1 741 630	89	49 995 925
15	176	40	37 338	65	2 012 558	90	56 634 173
16	231	41	44 583	66	2 323 520	91	64 112 359
17	297	42	53 174	67	2 679 689	92	72 533 807
18	385	43	63 261	68	3 087 735	93	82 010 177
19	490	44	75 175	69	3 554 345	94	92 669 720
20	627	45	89 134	70	4 087 968	95	104 651 419
21	792	46	105 558	71	4 697 205	96	118 114 304
22	1002	47	124 754	72	5 392 783	97	133 230 930
23	1255	48	147 273	73	6 185 689	98	150 198 136
24	1575	49	173 525	74	7 089 500	99	169 229 875
25	1958	50	204 226	75	8 118 264	100	190 569 292

Существует много тождеств типа тождеств Эйлера—Гаусса—Якоби. Например,

а) $\dfrac{(\varphi(t^2))^2}{\varphi(t)}=\sum\limits_{k=-\infty}^{+\infty} t^{2k^2+k}$ (Гаусс);

б) $\dfrac{(\varphi(t^2))^5}{(\varphi(t))^2}=\sum\limits_{k=-\infty}^{+\infty}(-1)^k\times$
$\times(3k+1)t^{3k^2+2k}$ (Гордон);

в) $\dfrac{(\varphi(t))^5}{(\varphi(t^2))^2}=\sum\limits_{k=-\infty}^{+\infty}(6k+1)t^{(3k^2+k)/2}$
(Гордон);

г) $\dfrac{\varphi(t^2)(\varphi(t^3))^2}{\varphi(t)\varphi(t^6)}=\sum\limits_{k=-\infty}^{+\infty} t^{(3k^2+k)/2}$;

д) $\dfrac{\varphi(t)\varphi(t^6)}{\varphi(t^2)\varphi(t^3)}=\sum\limits_{k=-\infty}^{+\infty}(-1)^k t^{3k^2+2k}$;

е) $\dfrac{(\varphi(t))^2\varphi(t^6)}{\varphi(t^2)\varphi(t^3)}=\sum\limits_{k=-\infty}^{+\infty}\left(\dfrac{k+1}{3}\right)\times$
$\times t^{(k^2+k)/2}$;

ж) $\dfrac{(\varphi(t^2))^2\varphi(-t^3)}{\varphi(-t)\varphi(t^6)}=\sum\limits_{k=-\infty}^{+\infty}\left(\dfrac{k+1}{3}\right)t^{k^2}$,

где $\left(\dfrac{k+1}{3}\right)$ — символ Лежандра, то есть $\left(\dfrac{k+1}{3}\right)=0$, 1 или -1 в зависимости от того, дает k при делении на 3 в остатке 2, 0 или 1. ■

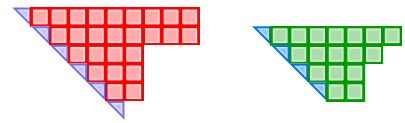
Более общим, чем пентагональное тождество Эйлера, является тождество Гаусса—Якоби

$$\prod_{k=1}^{\infty}(1-q^{2k})(1+q^{2k-1}z)(1+q^{2k-1}z^{-1})=\sum_{n=-\infty}^{+\infty}z^n q^{n^2}.$$

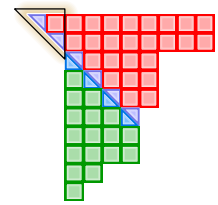
Запишем его в виде

$$\prod_{k=1}^{\infty}(1+q^{2k-1}z)(1+q^{2k-1}z^{-1})=\sum_{n=-\infty}^{+\infty}z^n q^{n^2}\times$$
$$\times\prod_{k=1}^{\infty}(1+q^{2k}+q^{4k}+q^{6k}+\ldots).$$

Будем рисовать диаграммы, считая, что половина клетки изображает единицу, а клетка — двойку. Тогда разбиения $a=19+17+9+7+5++1$ и $b=15+11+7+5$ изобразятся следующими рисунками:



Приложив их, получаем такой рисунок:



Отрезав треугольник (сторона которого n равна разности количеств слагаемых в разбиениях), получим диаграмму разбиения числа $a+b-n^2=16+16++10+10+10+8+8+8+4+2$ на четные слагаемые. ■

Тождество Гаусса—Якоби можно доказать и алгебраически. Найдем такие $c_0, c_1, c_2, \ldots$, чтобы при любом $z\neq 0$ было верно равенство

$$(1+qz)(1+qz^{-1})(1+q^3z)(1+q^3z^{-1})\ldots(1+q^{2n-1}z)(1+q^{2n-1}z^{-1})=$$
$$=c_0+c_1(z+z^{-1})+c_2(z^2+z^{-2})+\ldots+c_n(z^n+z^{-n}).$$

Обозначим левую часть через $F(z)$. Очевидно, $c_n=q^{1+3+\ldots+(2n-1)}=q^{n^2}$ и

$$F(q^2z)=F(z)\,\frac{1+q^{2n+1}z}{1+qz}\cdot\frac{1+q^{-1}z^{-1}}{1+q^{2n-1}z^{-1}}=F(z)\,\frac{1+q^{2n+1}z}{qz+q^{2n}}.$$

Поэтому $F(q^2z)(qz+q^{2n})=F(z)(1+q^{2n+1}z)$. Приравнивая коэффициенты при z^{k+1} в левой и правой частях этого равенства, получаем при $k=0, 1, \ldots, n-1$ равенство

$$q^{2k+1}c_k+q^{2(k+1)+2n}c_{k+1}=c_{k+1}+c_kq^{2n+1},$$

то есть $c_k=c_{k+1}\,\dfrac{1-q^{2n+2k+2}}{q^{2k+1}(1-q^{2n-2k})}$. Следовательно,

$$c_k=\frac{(1-q^{2n+2k+2})(1-q^{2n+2k+4})\ldots(1-q^{4n-2})(1-q^{4n})}{(1-q^{2n-2k})(1-q^{2n-2k-2})\ldots(1-q^4)(1-q^2)}q^{k^2}$$

при $k=n-1, n-2, \ldots, 2, 1, 0$. Для доказательства тождества Гаусса—Якоби осталось при каждом фиксированном k перейти к пределу при $n\to\infty$: числитель стремится к единице, а знаменатель — к бесконечному произведению $(1-q^2)(1-q^4)(1-q^6)(1-q^8)\ldots$ ■

Подставим в тождество Гаусса—Якоби $u=v=-t$. Левая часть превратится в

$$\prod_{k=1}^{\infty}(1-t^{2k})(1-t^{2k-1})^2 = \prod_{k=1}^{\infty}(1-t^{2k-1})\prod_{k=1}^{\infty}(1-t^k).$$

Вспоминая тождество Эйлера $\prod_{k=1}^{\infty}(1-t^{2k-1})=\prod_{k=1}^{\infty}(1+t^k)^{-1}$, получаем

$$\prod_{k=1}^{\infty}(1+t^k)^{-1}\prod_{k=1}^{\infty}(1-t^k) = \frac{(1-t)(1-t^2)(1-t^3)\ldots}{(1+t)(1+t^2)(1+t^3)\ldots}.$$

Правая часть тождества Гаусса—Якоби при подстановке $u=v=-t$ превращается в $\sum_{n=-\infty}^{+\infty}(-t)^{n^2} = 1-2t+2t^4-2t^9+2t^{16}-2t^{25}+\ldots$

Таким образом,

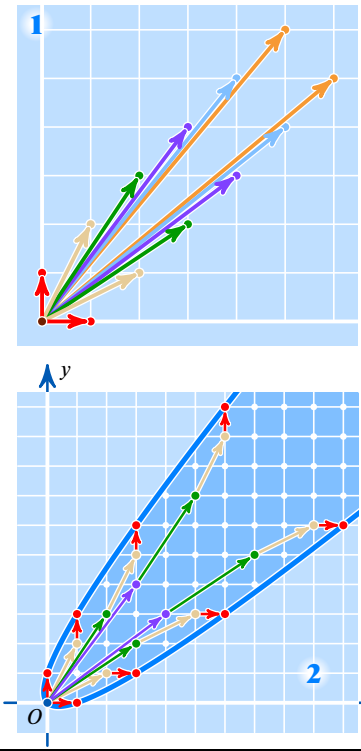
$$\frac{(1-t)(1-t^2)(1-t^3)\ldots}{(1+t)(1+t^2)(1+t^3)\ldots} = 1-2t+2t^4-2t^9+2t^{16}-2t^{25}+\ldots$$

Очевидно, левая часть равна

$$(1-t)^2(1-t^2)(1-t^3)^2\times\times(1-t^4)(1-t^5)^2\ldots$$

Подстановка $u=t$ и $v=1$ аналогичным образом приводит к еще одному тождеству Гаусса

$$\frac{(1-t^2)(1-t^4)(1-t^6)\ldots}{(1-t)(1-t^3)(1-t^5)\ldots} = 1+t+t^3+t^6+t^{10}+t^{15}+\ldots\ \blacksquare$$



Заменяя в тождестве Гаусса—Якоби z на $-z\sqrt{q}$ и q на $\sqrt{q}$, получаем

$$\prod_{k=1}^{\infty}(1-q^k)(1-q^kz)(1-q^{k-1}z^{-1}) = \sum_{n=-\infty}^{+\infty}(-1)^n z^n q^{(n^2+n)/2}.$$

Разделим на $1-z^{-1}$ обе части тождества:

$$\prod_{k=1}^{\infty}(1-q^k)(1-q^kz)(1-q^kz^{-1}) = (1+z^{-1}+z^{-2}+z^{-3}+\ldots)\sum_{n=-\infty}^{+\infty}(-1)^n z^n q^{(n^2+n)/2} =$$
$$= \sum_{n=0}^{\infty}(-1)^n q^{(n^2+n)/2}(z^n+z^{n-1}+z^{n-2}+\ldots)+$$
$$+\sum_{n=0}^{\infty}(-1)^{-n-1}q^{((-n-1)^2-n-1)/2}(z^{-n-1}+z^{-n-2}+z^{-n-3}+\ldots) =$$
$$= \sum_{n=0}^{\infty}q^{(n^2+n)/2}(-1)^n(z^n+z^{n-1}+\ldots+z^{-n}).$$

Таким образом,

$$\prod_{k=1}^{\infty}(1-q^k)(1-q^kz)(1-q^kz^{-1}) = \sum_{n=0}^{\infty}q^{(n^2+n)/2}(-1)^n(z^n+z^{n-1}+\ldots+z^{-n}).$$

Подставив $z=1$, получаем тождество Гаусса

$$\prod_{k=1}^{\infty}(1-q^k)^3 = \sum_{k=0}^{\infty}(-1)^k(2k+1)q^{k(k+1)/2}.$$

Другую форму тождество Гаусса приобретает под действием подстановки $q=\sqrt{uv}$ и $z=\sqrt{v/u}$: очевидно, $\prod_{k=1}^{\infty}(1-q^{2k})(1+q^{2k-1}z)(1+q^{2k-1}z^{-1}) = \prod_{k=1}^{\infty}(1-u^kv^k)\times\times(1+u^{k-1}v^k)(1+u^kv^{k-1})$ и $\sum_{n=-\infty}^{+\infty}z^nq^{n^2} = \sum_{n=-\infty}^{+\infty}v^{(n^2+n)/2}u^{(n^2-n)/2}$. Следовательно, тождество Гаусса—Якоби можно записать в виде

$$\prod_{k=1}^{\infty}(1-u^kv^k)(1+u^{k-1}v^k)(1+u^kv^{k-1}) = \sum_{n=-\infty}^{+\infty}v^{(n^2+n)/2}u^{(n^2-n)/2}.\ \blacksquare$$

В силу тождества Эйлера $f(uv)=1/\varphi(uv)$, так что тождество Гаусса—Якоби можно записать в виде

$$\prod_{k=1}^{\infty}(1+u^{k-1}v^k)(1+u^kv^{k-1}) = \sum_{n=-\infty}^{+\infty}v^{(n^2+n)/2}u^{(n^2-n)/2}\sum_{m=0}^{\infty}p(m)u^mv^m.$$

Обозначим через $t(x, y)$ коэффициент, возникающий при u^xv^y после раскрытия скобок и последующего приведения подобных слагаемых в бесконечном произведении $\prod_{k=1}^{\infty}(1+u^{k-1}v^k)(1+u^kv^{k-1})$. Рассмотрим векторы $(k-1;\,k)$ и $(k;\,k-1)$, где $k=1,\,2,\,3,\ldots$ (рис. 1). Очевидно, $t(x, y)$ — количество способов представить вектор $(x;\,y)$ в виде суммы нуля (и тогда сумма равна нулю, а такой способ единственный), одного (и тогда сумма равна своему единственному слагаемому) или нескольких рассматриваемых векторов, то есть

$$(x;\,y) = (r_1;\,r_1-1)+\ldots+(r_a;\,r_a-1)+(s_1;\,s_1+1)+\ldots+(s_b;\,s_b+1),$$

где $0<r_1<r_2<\ldots<r_a$ и $0\leqslant s_1<s_2<\ldots<s_b$, где возможны случаи $a=0$ или $b=0$. На рисунке 2 показаны точки $(x;\,y)$, которые можно представить в таком виде. Очевидно, эти точки — точки синей параболы и точки, лежащие внутри нее.

Лемма 1. ***$t(x;\,y)>0$ тогда и только тогда, когда $x+y\geqslant(x-y)^2$.***

Доказательство. Выясним, для каких неотрицательных x и y верно неравенство $x+y-(x-y)^2\geqslant 0$. Числа $x+y$ и $x-y$ одной четности, поэтому для любого неотрицательного числа m и любого целого числа q достаточно решить систему уравнений

$$\begin{cases} x-y=q,\\ x+y-q^2=2m. \end{cases}$$

Очевидно, $(x;y)=\left(\frac{q^2+q}{2}+m;\ \frac{q^2-q}{2}+m\right)$. Геометрический смысл этой формулы таков: $\left(\frac{q^2+q}{2};\ \frac{q^2-q}{2}\right)$ — точка на параболе, а вектор $(m;m)$ сдвигает эту точки внутрь параболы при $m>0$ и оставляет на месте при $m=0$.

Поскольку задача симметрична относительно перестановки координат векторов (геометрически этому соответствует симметрия рисунков 1 и 2 относительно биссектрисы первого координатного угла), достаточно рассмотреть случаи, когда $x\geqslant y$. Очевидно,

$$1+2+\ldots+(q-1)+q+m=\frac{q(q+1)}{2}+m,$$

$$0+1+\ldots+(q-2)+(q-1)+m=\frac{q(q-1)}{2}+m,$$

так что при $m>0$ имеем $(x;y)=(1;0)+(2;1)+\ldots+(q-1;q-2)+(q;q-1)+(1;0)+(m-1;m)$, а при $m=0$ последние два слагаемых излишни.

Обратно, для каждого вектора $(x;y)=(r_1+r_2+\ldots+r_a+s_1+s_2+\ldots+s_b;\ r_1+r_2+\ldots+r_a+s_1+s_2+\ldots+s_b-a+b)$ имеем $x-y=a-b$ и

$$(x;y)=\left(\frac{q(q+1)}{2};\ \frac{q(q-1)}{2}\right)+(m;m),$$

где $m=r_1-1+r_2-2+\ldots+r_q+r_{q+1}+\ldots+r_a+s_1+s_2+\ldots+s_b$. Поскольку $r_1\geqslant 1$, $r_2\geqslant 2,\ \ldots,\ r_q\geqslant q$, то $m\geqslant 0$, что и требовалось доказать.

Лемма 2. ***t*(*x*; *y*) *зависит только от*** $m=(x+y-(x-y)^2)/2$, ***а не от*** $q=x-y$.

Доказательство. Рассмотрим отображение, определенное правилом

$$\psi(r_1,r_2,\ldots,r_a;s_1,s_2,\ldots,s_b)=$$
$$=\begin{cases}(r_1-1,r_2-1,\ldots,r_a-1;0,s_1+1,s_2+1,\ldots,s_b+1), & \text{если } r_1>1,\\ (r_2-1,\ldots,r_a-1;s_1+1,s_2+1,\ldots,s_b+1), & \text{если } r_1=1.\end{cases}$$

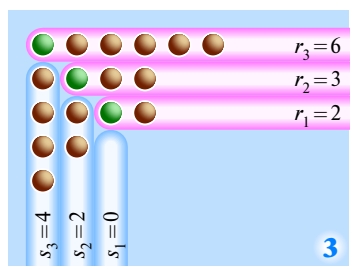
Очевидно, отображение ψ не меняет величину m, но уменьшает на 1 величину q. Чтобы доказать, что ψ — биекция, рассмотрим следующее отображение:

$$\Psi(r_1,r_2,\ldots,r_a;s_1,s_2,\ldots,s_b)=$$
$$=\begin{cases}(1,r_1+1,r_2+1,\ldots,r_a+1;s_1-1,s_2-1,\ldots,s_b-1), & \text{если } s_b>0,\\ (r_1+1,r_2+1,\ldots,r_a+1;s_1+1,s_2+1,\ldots,s_b+1), & \text{если } s_1=0.\end{cases}$$

Очевидно, ψ и Ψ взаимно обратны; следовательно, ψ — биекция.

Теорема (тождество Гаусса—Якоби). ***t*(*x*; *y*)=*p*(*m*).**

Доказательство. Достаточно доказать равенство $t(m;m)=p(m)$. Каждому разбиению числа m сопоставим некоторое представление вектора $(m;m)$ в виде суммы нуля (при $m=0$) или нескольких векторов, координаты которых неотрицательны и отличаются одна от другой на 1. Проведем диагональ на диаграмме Юнга (рис. 3). Пусть $a=b$ — ее длина. Пусть r_a — количество точек первой строки, лежащих на диагонали или справа от нее, r_{a-1} — количество таких точек второй строки и так далее; s_a — количество точек первого столбца, лежащих под диагональю, s_{a-1} — второго столбца и так далее. Например, на рисунке $r_3=6$, $r_2=3$, $r_1=2$ и $s_3=4$, $s_2=2$ и $s_1=0$; соответствующее представление вектора (17; 17) имеет вид $(2;1)+(3;2)+(6;5)+(0;1)+(2;3)+(4;5)$. ■



Обозначим через d_k количество тех домов некоторого города, в которых живет не меньше k жителей, а через c_m — количество жителей в m-м по величине населения доме. Рассматривая диаграммы Юнга, докажите равенства

а) $c_1+c_2+c_3+\ldots=d_1+d_2+d_3+\ldots$;

б) $c_1^2+c_2^2+c_3^2+\ldots=d_1+3d_2+5d_3+\ldots+(2k-1)d_k+\ldots$;

в) $d_1^2+d_2^2+d_3^2+\ldots=c_1+3c_2+5c_3+\ldots+(2k-1)c_k+\ldots$ ■

Количество различных слагаемых в разбиении натурального числа n назовем его разбросом (например, разброс разбиения 2+2+2+5+6+6=23 равен 3). Докажем, что а) разброс любого разбиения не превосходит $\sqrt{2n}$; б) сумма разбросов всех $p(n)$ разбиений числа n равна

$$1+p(1)+p(2)+\ldots+p(n-1).$$

Доказательство. а) Если $a_1<a_2<\ldots<a_k$ — все различные слагаемые, входящие в разбиение числа n, то $a_1\geqslant 1$, $a_2\geqslant 2,\ \ldots,\ a_k\geqslant k$ и

$$n\geqslant a_1+a_2+\ldots+a_k\geqslant 1+2+\ldots+k=k(k+1)/2;$$

следовательно,

$$k^2<k(k+1)\leqslant 2n.$$

б) В каждом из $p(n)$ разбиений числа n выделить слагаемое можно столькими способами, каков разброс этого разбиения. Вычеркнем выделенное слагаемое — получим разбиение некоторого неотрицательного числа, меньшего n. ■

Тождества Роджерса—Рамануджана. 1) Количество разбиений натурального числа, в которых нет слагаемых, различающихся менее чем на 2, равно количеству разбиений этого же числа на слагаемые, дающие при делении на 5 остаток 1 или 4.

2) Количество разбиений натурального числа, в которых нет слагаемых, различающихся менее чем на 2, причем все слагаемые больше 1, равно количеству разбиений этого же числа на слагаемые, дающие при делении на 5 остаток 2 или 3.

Эти тождества С. Рамануджан в 1913 г. сообщил в письме английскому математику Г. Х. Харди. Прозрачные комбинаторные доказательства этих тождеств до сих пор не найдены. ■

Числа Каталана

Расстановки скобок, деревья, разрезания выпуклого многоугольника на треугольники, перестановки, никакие три числа которых не идут в порядке возрастания, и многие другие задачи приводят к одной и той же последовательности $C_1=1$, $C_2=1$, $C_3=2$, $C_4=5$, $C_5=14$, . . . , *удовлетворяющей рекуррентному соотношению*

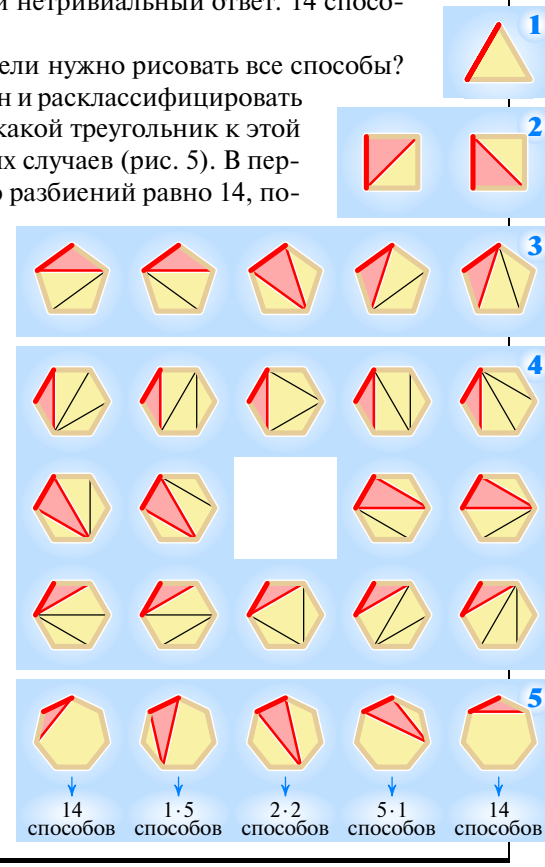
$$C_n = C_1C_{n-1} + C_2C_{n-2} + \ldots + C_{n-1}C_1. \qquad (*)$$

Тремя разными способами — при помощи польской записи арифметических выражений, леммы об отражении и бинома Ньютона — мы выведем явную формулу

$$C_n = \frac{1}{2n-1}C_{2n-1}^{n-1}. \qquad (**)$$

Разрезания на треугольники. Л. Эйлер был первым, кто столкнулся с этой последовательностью. Он спросил себя, сколькими способами выпуклый n-угольник можно разрезать на треугольники диагоналями, не пересекающимися внутри него. При $n=3$ никаких диагоналей проводить не надо, способ единственен (рис. 1). Для $n=4$ можно провести любую из двух диагоналей (рис. 2). Чтобы разрезать диагоналями выпуклый пятиугольник на треугольники, нужно из некоторой вершины провести обе диагонали (рис. 3). При $n=6$ получаем первый нетривиальный ответ: 14 способов (рис. 4).

Как быть с семиугольником? Неужели нужно рисовать все способы? Нет, можно выделить одну из сторон и расклассифицировать разрезания в зависимости от того, какой треугольник к этой стороне примыкает. Имеем 5 разных случаев (рис. 5). В первом и последнем из них количество разбиений равно 14, поскольку после отрезания треугольника остается шестиугольник. Во втором и четвертом случаях при вырезании треугольника семиугольник распадается на треугольник и пятиугольник. Треугольник резать не надо, а пятиугольник, как мы знаем, дает 5 способов. В третьем случае от семиугольника остаются два четырехугольника; каждый из них можно разбить двумя способами, получаем $2\cdot2=4$ варианта. Итак, семиугольник можно разбить $14+5+2\cdot2+5+14=42$ способами. Рассматривая восьмиугольник, аналогично получаем $42+14+2\cdot5+5\cdot2+14+42=132$ способа. Для девятиугольника — $132+42+2\cdot14+5\cdot5+14\cdot2+42+132=429$ способов, а для десятиуголь-



C_1 =	1
C_2 =	1
C_3 =	2
C_4 =	5
C_5 =	14
C_6 =	42
C_7 =	132
C_8 =	429
C_9 =	1430
C_{10} =	4862
C_{11} =	16 796
C_{12} =	58 786
C_{13} =	208 012
C_{14} =	742 900
C_{15} =	2 674 440
C_{16} =	9 694 845
C_{17} =	35 357 670
C_{18} =	129 644 790
C_{19} =	477 638 700
C_{20} =	1 767 263 190
C_{21} =	6 564 120 420
C_{22} =	24 466 267 020
C_{23} =	91 482 563 640
C_{24} =	343 059 613 650
C_{25} =	1 289 904 147 324
C_{26} =	4 861 946 401 452
C_{27} =	18 367 353 072 152
C_{28} =	69 533 550 916 004
C_{29} =	263 747 951 750 360
C_{30} =	1 002 242 216 651 368
C_{31} =	3 814 986 502 092 304
C_{32} =	14 544 636 039 226 909
C_{33} =	55 534 064 877 048 198

Из формулы (**) следует равенство

$$C_{n+1} = \frac{4n-2}{n+1}C_n.$$

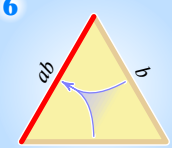
Убедитесь в этом! ■

Число C_n нечетно тогда и только тогда, когда $n=2^k$, где k — целое неотрицательное число. Дело в том, что формула (*) одинаково читается как слева направо, так и справа налево. Поэтому четны все числа C_{2n+1}, где $n>0$; а число C_{2n} четно тогда и только тогда, когда четно C_n. ■

ника — $429+132+2\cdot42+5\cdot14+14\cdot5+42\cdot2+132+429=1430$ способов. Такие вычисления можно проводить и дальше, но мы благоразумно остановимся и займемся другой — алгебраической — задачей. ■

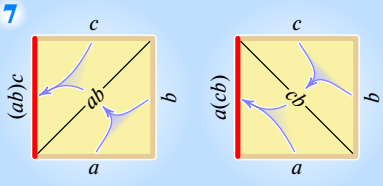
Произведение abc можно понимать двояко: $(ab)c$ и $a(bc)$. (Конечно, результат не зависит от порядка умножений. Но промежуточные результаты — зависят!) Произведение $abcd$ можно понимать пятью способами: $((ab)c)d$, $(a(bc))d$, $a((bc)d)$, $a(b(cd))$ и $(ab)(cd)$. Произведение $abcde$ — четырнадцатью способами: есть 5 способов вида $a(bcde)$, 2 способа вида $(ab)(cde)$, 2 способа вида $(abc)(de)$ и 5 способов вида $(abcd)e$.

6



Число Каталана C_n — это количество способов расставить скобки в произведении n множителей. Выведем рекуррентную (то есть выражающую очередной член последовательности через предыдущие) формулу. Для этого рассмотрим знак умножения, которое будет выполнено в последнюю очередь. Произведение $x_1x_2\ldots x_n$ получается в конечном счете как произведение некоторого произведения первых нескольких символов на некоторое произведение остальных:

7

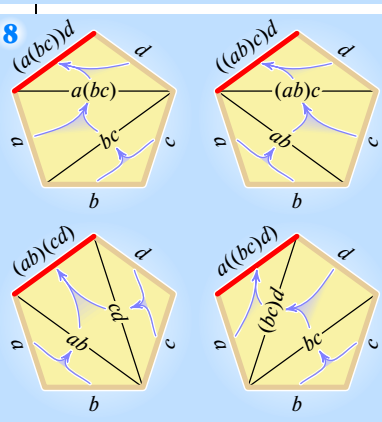


$$x_1x_2\ldots x_n=(x_1\ldots x_r)\cdot(x_{r+1}\ldots x_n).$$

Первые r символов могут быть скомбинированы C_r способами, последние $n-r$ символов — C_{n-r} способами. Таким образом, верна формула $(*)$.

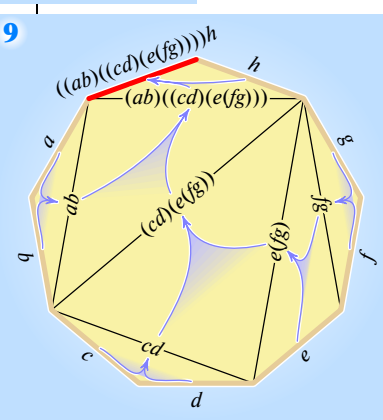
Этой же рекуррентной зависимости удовлетворяет рассмотренная в предыдущем разделе статьи последовательность количеств способов разрезания на треугольники. Поэтому разрезать выпуклый n-угольник на $n-2$ треугольника непересекающимися диагоналями можно C_{n-1} способами. ■

8



Между рассмотренными задачами есть более явная связь: взаимно-однозначное соответствие (коротко говоря, биекция) между разбиениями на треугольники и способами подсчета произведений. Как заметил в 1961 г. Фордер, можно выделить одну сторону $(n+1)$-угольника и написать сомножители около других его сторон, по одной букве у каждой стороны, а затем «стягивать» треугольники, на двух сторонах которых уже что-то написано, записывая произведение на третью сторону. На рисунках 6, 7 и 8 это показано для $n=2$, 3 и 4 соответственно, а на рисунке 9 изображен один из 429 случаев для $n=8$. ■

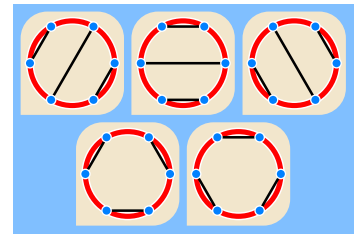
9



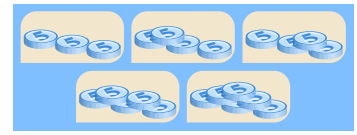
Расстановки скобок. Рассмотрим какое-нибудь арифметическое выражение и сотрем все, кроме скобок. Получим некоторую систему открывающих и закрывающих скобок. Какими свойствами она обладает?

Несколько десятков конструкций приводят к числам Каталана. Рассмотрим несколько из них. Краткости ради их описания не вполне точны. Рисунки (а они выполнены для $n=4$) помогут восстановить пропущенные детали. Итак, n-е число Каталана равно количеству

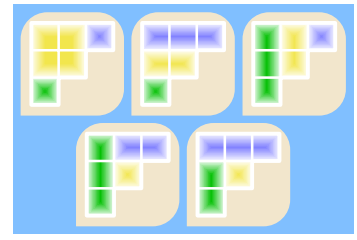
• способов так разбить $2(n-1)$ точек окружности на пары, чтобы соответствующие хорды не пересекались;



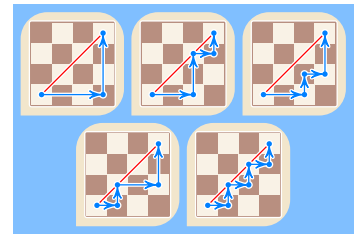
• способов расположить на плоскости стопку из нескольких одинаковых монет, где в нижнем ряду — $n-1$ монета;



• разбиений состоящей из $n(n-1)/2$ клеток изображенной на рисунке фигуры на $n-1$ прямоугольников, каждый из которых содержит по одной клетке правой границы фигуры;

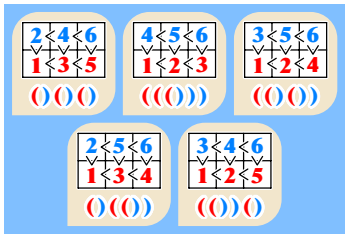


• способов пройти из левого нижнего угла доски размером $n\times n$ в правый верхний угол, сдвигаясь каждым ходом на одну клетку вправо или вверх и ни разу не оказавшись выше диагонали, соединяющей левый нижний угол с правым верхним;

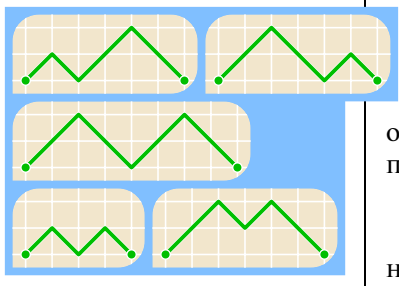


(*Продолжение на с. 564.*)

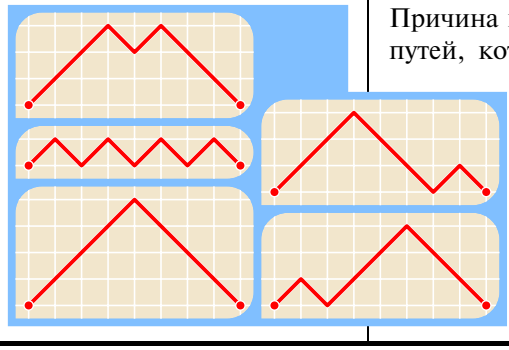
• способов расставить числа от 1 до $2(n-1)$ в таблицу из 2 строк и $n-1$ столбцов таким образом, чтобы в каждой из строк числа слева направо возрастали и каждое число нижней строки было меньше стоящего над ним соответствующего числа верхней строки (на рисунке каждой правильной расстановке $n-1$ пар скобок сопоставлена таблица, в ее нижней строке — номера открывающих скобок, а в верхней — номера закрывающих);



• путей Дика (то есть путей, начинающихся в начале координат и не спускающихся ниже оси абсцисс, где каждый шаг — перемещение на единицу вправо и на единицу вверх или вниз), оканчивающихся на оси абсцисс, обладающих $n-1$ точками локального минимума (другими словами, $n-2$ пиками — локальными максимумами) и не имеющих трех подряд идущих шагов в одном и том же направлении;

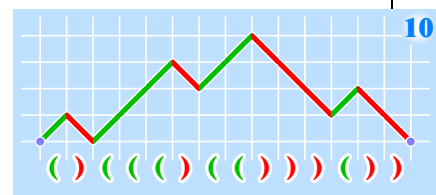


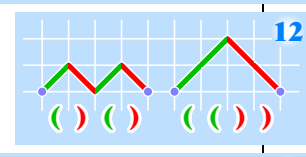
• путей Дика, оканчивающихся в точке $(2n; 0)$, ни один из пиков которых не расположен на высоте 2;

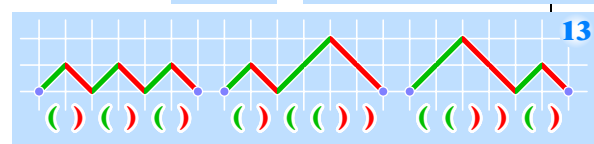


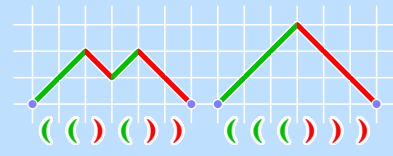
Во-первых, открывающих скобок ровно столько же, сколько и закрывающих. Во-вторых, ни в каком начальном отрезке количество закрывающих скобок не может оказаться больше количества открывающих скобок. Например, расстановки)(и ((()))(неправильные. Нетрудно доказать, что эти два условия не только необходимы, но и достаточны.

Рассмотрим несколько примеров. Одна пара скобок может выглядеть единственным способом: (). Две пары — двумя способами: ()() или (()). Три пары — пятью способами: ()()(), ()(()), (())(), (()()) или ((())). Четыре пары, как нетрудно проверить, — четырнадцатью способами. Чтобы понять, сколькими способами могут выглядеть правильно расставленные пять пар скобок, рассмотрим закрывающую скобку, парную к первой открывающей скобке. Остальные четыре пары тогда разделятся на две группы: расположенные внутри рассмотренной пары и расположенные справа от нее. (Разумеется, любая из этих групп может состоять из 0 скобок.) Способов, когда все четыре пары внутри или все четыре справа, — по 14 в каждом случае. Когда три пары внутри, а одна справа, — 5 способов. Столько же — когда одна внутри, а три справа. Наконец, когда две пары внутри, а две справа, имеем $2\cdot 2=4$ способа.









Итого, $14+5+2\cdot 2+5+14=42$ способа.

Сопоставим открывающей скобке сдвиг направо-вверх, а закрывающей — направо-вниз. Путь возвращается на исходную горизонталь, когда количества открывающих и закрывающих скобок уравниваются. Путь спускается ниже нее, как только количество закрывающих скобок превышает количество открывающих. Правильные расстановки скобок взаимно однозначно соответствуют путям Дика — путям, возвращающимся на ось абсцисс и не спускающимся ниже нее (рис. 10). На рисунках 11, 12 и 13 проиллюстрированы случаи 1, 2 и 3 пар скобок соответственно. ■

Треугольник Каталана. Треугольник Паскаля получается, если начать с верхней единицы и затем действовать по правилу «каждое очередное число равно сумме чисел, расположенных над ним справа и слева» (а на краях — единицы). Проведем вертикальную линию, левее которой заходить нельзя. И будем выписывать числа так, словно мы составляем треугольник Паскаля. На вертикальной линии — числа Каталана (рис. 14)!

Причина в том, что любое число треугольника Каталана равно количеству путей, которыми можно прийти в соответствующую точку из вершины, двигаясь вниз-влево или вниз-вправо. Поверните треугольник Каталана на 90° против часовой стрелки, вспомните рисунки 10—13 — и все станет понятно!

Сумма чисел, выделенных на рисунке 14 красным цветом, равна числу, обведенному красным кружком:

$$275+110+35+8+1=429.$$

А сумма чисел, выделенных зеленым цветом, равна числу, обве-

14

денному зеленым кружком:

$$14+28+48+75=165.$$

(Обратите внимание: начальное число 5 этой диагонали в сумму не входит.) Идея доказательства этого свойства (как и предыдущего) такова:

$$14+28+48+75=42+48+75=90+75=165.\ \blacksquare$$

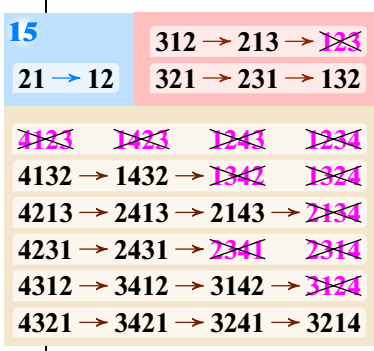
Перестановки Д. Кнута.

Сколькими способами можно так переставить первые n натуральных чисел, чтобы никакие три из чисел полученной последовательности не шли в порядке возрастания?

При $n=1$ перестановка единственна. При $n=2$ годятся обе перестановки: и 21, и 12. При $n=3$ не годится лишь перестановка 123, остальные годятся. При $n=4$ годятся все перестановки, оставленные невычеркнутыми на рисунке 15. Заметьте: в четырех столбцах одного рисунка остались невычеркнутыми соответственно 1 и 1 перестановки, в столбцах другого — 2, 2 и 1, третьего — 5, 5, 3 и 1 перестановок. Очевидно, $1+1=2=C_3$, $1+2+2=5=C_4$ и $5+5+3+1=14=C_5$.

Возникает гипотеза: для любого натурального n количество правильных по Д. Кнуту перестановок, в которых n расположено на k-м месте, равно числу, расположенному в k-м столбце треугольника Каталана на n-й линии, идущей параллельно красной линии рисунка 14.

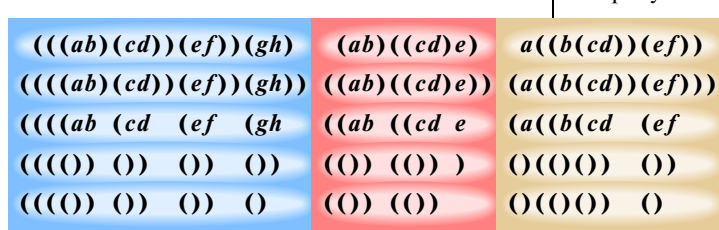
15



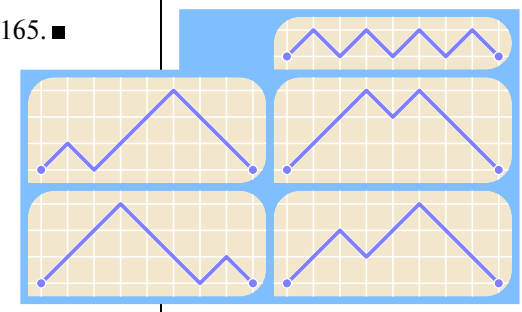
Докажем эту гипотезу. Начнем с простого замечания. Первые $k-1$ чисел правильной перестановки, на k-м месте которой стоит n, образуют убывающую последовательность — иначе вместе с числом n числа a и b, где $a<b$ и a левее, чем b, а b левее, чем n, образовывали бы тройку расположенных в порядке возрастания чисел $a<b<n$.

Пусть непосредственно справа от числа n расположено число c. Можно ли поменять местами n и c? Иногда можно, иногда нельзя (посмотрите на стрелочки рисунка 15!). Зависит это от того, останется ли убывающей последовательность чисел, расположенных справа от n. Проще говоря, если слева от n стояло число, которое больше числа c, то менять местами числа n и c можно. В противном случае — нельзя.

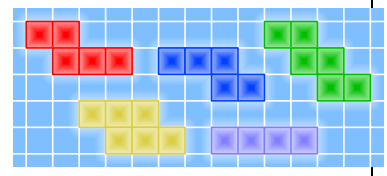
В этом самом противном случае заметим, что если справа от c окажется какое-то число $d>c$, то расположенное непосредственно слева от n число вместе с числами c и d образует тройку чисел в порядке возрастания. Значит, числа от $c+1$ до $n-1$ расположены левее n. Поскольку последовательность чисел, расположенных левее n, убывающая, она начинается с чисел $n-1$, $n-2$, . . . , $c+1$.



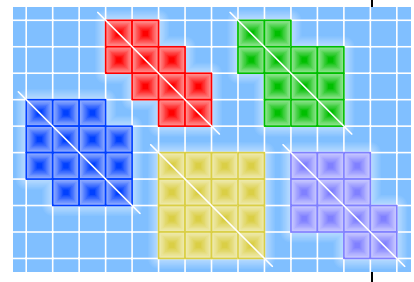
• путей Дика, оканчивающихся в точке $(2n; 0)$ и обладающих тем свойством, что длина любой максимальной последовательности подряд идущих шагов «направо-вниз» нечетна;



• горизонтальных полимино ширины n (то есть полимино, состоящих из нескольких горизонтальных рядов, каждый из которых расположен строго правее всех рядов над ним);



• симметричных полимино периметра $4n$;



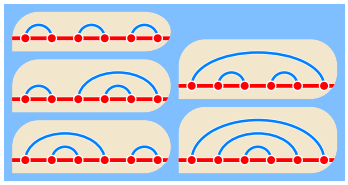
(*Продолжение на с. 566.*)

Рассмотрим произведение $(((ab)(cd))(ef))(gh)$. Заключим его в скобки. Сотрем все закрывающие скобки. Заменим все буквы на закрывающие скобки и последнюю из них сотрем: ((((())(()))()). Получили правильную расстановку скобок. На рисунке показаны и другие примеры. Немного подумав, вы докажете, что это — биекция между расстановками скобок и способами подсчета произведений. ■

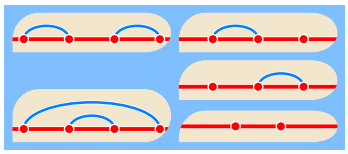
Числа Белла пересчитывают разбиения n-элементного множества на классы (на рисунках разобраны случаи $n=2$, 3 и 4). Для поэта число Белла — количество различных рифмовок. Например, четверостишие имеет 15 возможных рифмовок (одна из них — отсутствие какой бы то ни было рифмы). А 14-строчное стихотворение можно срифмовать 190 899 322 способами (именно таково 14-е число Белла).

А количество рифмовок, не нуждающихся в пересечении дуг, — число Каталана! ■

• способов разбить на пары $2(n-1)$ точек, расположенных на горизонтальной прямой, при помощи непересекающихся линий, расположенных в верхней полуплоскости (форма линий не имеет значения, важно лишь то, пересекаются линии или нет и какие точки они соединяют);



• способов соединить некоторые из расположенных на одной горизонтальной прямой точек таким образом, чтобы дуги лежали в верхней полуплоскости и не пересекали одна другую, их количество в сумме с количеством изолированных точек равно $n-2$, причем ни одна изолированная точка не лежит под дугой;



• способов соединить n точек, расположенных на горизонтальной прямой, дугами, лежащими в верхней полуплоскости и не пересекающими одна другую во внутренних точках таким образом, чтобы из любой точки можно было дойти по дугам в любую другую, причем единственным способом (другими словами, дуги должны образовывать дерево),

Вычеркнем эти числа вместе с числом n — получим последовательность из первых c ратуральных чисел, в которой c расположено на одну позицию ближе к правому краю, чем раньше располагалось число n.

Рассмотрим для примера $n=9$, $k=4$. Первые три числа любой такой правильной перестановки образуют убывающую последовательность. Сдвинуть число 9 на пятое место (стоявшее на пятом месте число перемещая на четвертое место) можно в 275 случаях — именно столько существует правильных перестановок на 9 элементах, в которых 9 расположено на 4-м месте. Если же сдвиг невозможен, то справа от 9 должно стоять число большее, чем слева.

Случай $c=1$ невозможен. Если $c=2$, то все числа от 3 до 8 должны стоять на первых трех местах, но им там не хватает места. То же — для $c=3$, 4 или 5.

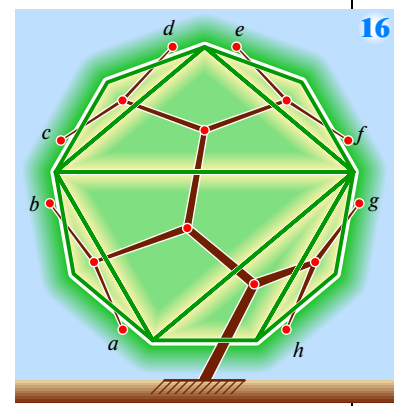
А вот случай $c=6$ возможен:

$$87*96****.$$

Вычеркивая цифры 7, 8 и 9, получаем последовательность вида $*6****$, а таких (правильных по Кнуту) последовательностей всего 42. Далее, при $c=7$ имеем

$$8**97****.$$

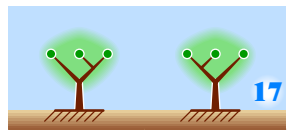
Вычеркивая цифры 8 и 9, получаем $**7****$; таковых 90 штук. Наконец, при $c=8$ надо вычеркнуть только цифру 9, а правильных по Д. Кнуту последовательностей вида $***8****$ существует 165 штук. Итого: $275+(42+90+165)=275+297=572$. ■

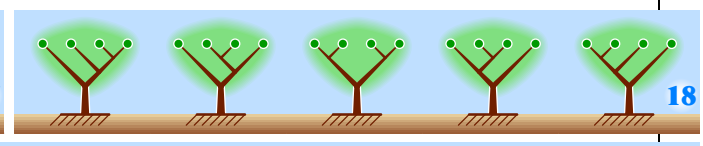


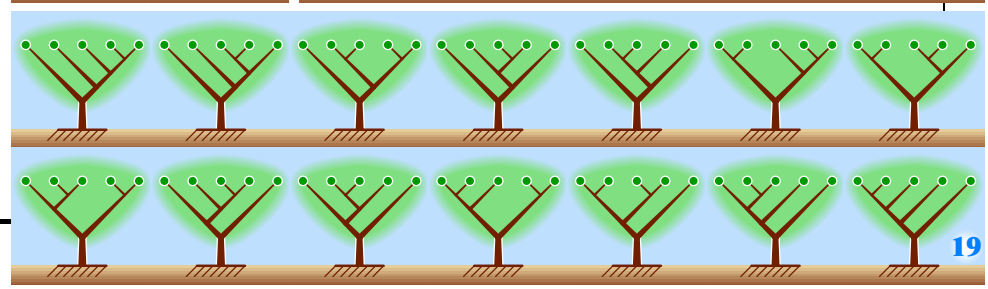
Деревья с n листьями. На рисунке 16 разбиению выпуклого девятиугольника сопоставлено корневое дерево (дерево — это связный граф без циклов; слово «корневое» означает, что одна из вершин выделена — названа корнем дерева) с 8 листьями (лист — это отличная от корня вершина дерева, из которой выходит только одно ребро; на рисунке 16 листы помечены буквами).

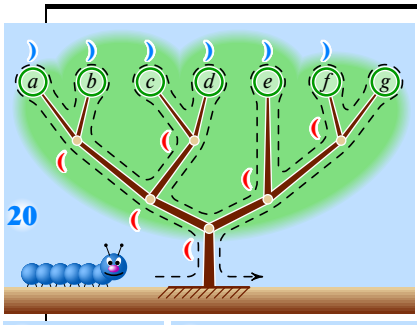
А. Кэли заметил, что C_n есть количество корневых деревьев с n листьями, степень любой вершины которого (количество сходящихся в ней ребер) равна 1 или 3. На рисунках 17, 18 и 19 изображены все интересующие нас деревья с 3, 4 и 5 листьями соответственно.

Я. Лукасевич предложил несложный способ нахождения расстановки скобок, соответствующего дереву. Он пометил все листья, кроме последнего, закрывающими скобками, а вершины степени 3 — открывающими (рис. 20).





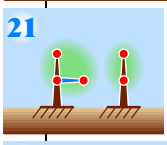


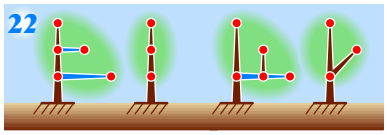


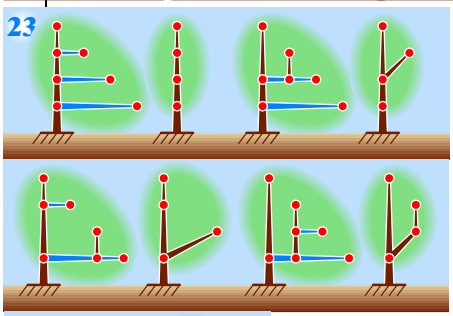
20

А затем вообразил, что гусеница оползает вокруг всего дерева вдоль пунктирной линии, собирая все скобки. Нетрудно доказать, что гусеница получит правильную расстановку скобок, соответствующую способу вычисления произведения. Например, на рисунке 20 она получит расстановку (((())))()(), соответствующую произведению ((*ab*)(*cd*))(*e*(*fg*)).

Ф. Бернхарт придумал биекцию между корневыми деревьями с n листьями, степени вершин которых могут равняться только 1 или 3, и деревьями с $n+1$ вершинами, одна из которых — корень. Это взаимно однозначное соответствие для $n=2$, 3 и 4 показано на рисунках 21, 22 и 23 соответственно. Идея в том, что все горизонтальные (синие) ребра стягиваем в точки. Гусеница, оползающая любое из вновь нарисованных деревьев, даст ту же самую правильную расстановку скобок, что и на исходном дереве, если изменит свой алгоритм следующим образом: начнет не с корня, а с нижней вершины; каждый раз, когда поползет вверх, напишем открывающую скобку, а на спуске по ребру вниз — закрывающую. ■


21


22


23

Явная формула. Рекуррентная формула (∗) требует для вычисления числа C_n знать все предыдущие значения C_1, C_2, …, C_{n-1}. Хотелось бы найти формулу, выражающую C_n непосредственно через n. Сделаем это тремя способами: при помощи 1) польской записи; 2) леммы об отражении; 3) производящих функций. ■

Польская запись хорошо известна программистам, которым приходится учить машину вычислять значения арифметических выражений или производить какие-то другие операции. Рассмотрим, например, арифметическое выражение

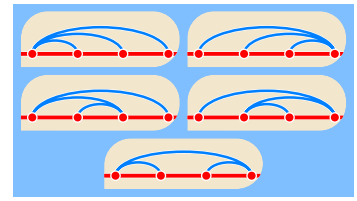
$$((1-2)+(3+4)):((5-6)\cdot 7-8\cdot 9).$$

Убрать скобки, очевидно, нельзя: прорядок действий пострадает. Но давайте вместо $1-2$ писать $1\ 2-$. И вообще, для любой бинарной операции $*$ вместо $a*b$ будем писать $a\ b\ *$. Тогда вместо $(1-2)+(3+4)$ получим $1\ 2-3\ 4++$, а вместо всего выражения получим

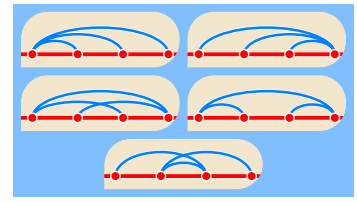
$$1\ 2-3\ 4++5\ 6-7\cdot 8\ 9\cdot -:.$$

Заметьте: символ выполняемой в последнюю очередь операции деления оказался на последнем месте. Что, согласитесь, логично! Немного подумав, вы поймете, как при помощи стека компьютер может вычислять (бесскобочное!) польское выражение. ■

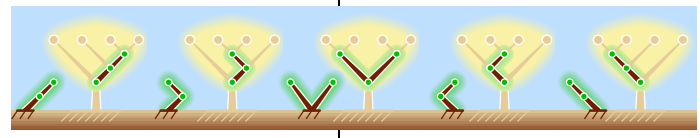
Перейдем к числам Каталана. Каждому произведению с правильно расставленными скобками сопоставим слово из букв *a* и *p* по следующему правилу: если выражение состоит лишь из одной буквы, пишем *a*; если да к тому же в любой точке сходящиеся в ней дуги должны идти либо все направо, либо все налево;



• то же условие, но запрет на пересечение во внутренних точках заменен на запрет дуге лежать всеми своими точками строго под другой дугой;



• деревьев с $n-1$ вершинами, «вырастающих» так: из каждой почки дерева и из его корня вырастает побег либо вправо, либо влево, либо два побега — и влево, и вправо (вообразите, что у корневого дерева с n листьями, степени вершин которого могут равняться только 1 или 3, отсохли все листья и отвалился корень);



• последовательностей

$$a_1\leqslant a_2\leqslant\ldots\leqslant a_{n-1}$$

натуральных чисел, удовлетворяющих неравенствам $a_k\leqslant k$, где $1\leqslant k<n$, а именно, при $n=4$ это последовательности

$$1\leqslant 1\leqslant 1,\quad 1\leqslant 1\leqslant 2,\quad 1\leqslant 1\leqslant 3,$$
$$1\leqslant 2\leqslant 2,\quad 1\leqslant 2\leqslant 3;$$

• последовательностей

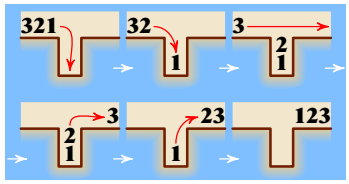
$$a_1<a_2<\ldots<a_{n-2}$$

натуральных чисел, удовлетворяющих неравенствам $a_k\leqslant 2k$, где $1\leqslant k<n-1$, при $n=4$ это последовательности

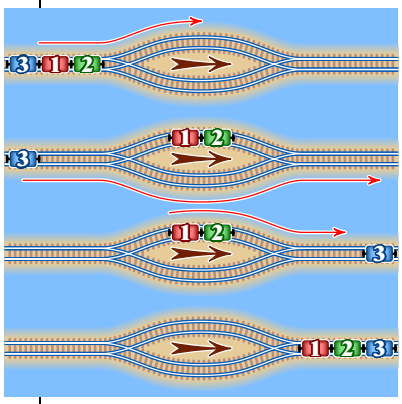
$$1<2,\quad 1<3,\quad 1<4,\quad 2<3,\quad 2<4;$$

• перестановок σ множества первых $n-1$ натуральных чисел, для которых нет таких чисел $i<j<k$, что $\sigma(j)<\sigma(k)<\sigma(i)$ (для $n=4$ это перестановки 123, 132, 213, 231 и 321);

(*Продолжение на с. 568.*)



• перестановок множества первых $n-1$ натуральных чисел, которые могут быть превращены в тождественную при помощи стека (при $n=4$ это перестановки 123, 132, 213, 231 и 321);



• перестановок множества первых $n-1$ натуральных чисел, которые можно преобразовать в тождественную при помощи минимальной маневровой горки (при $n=4$ это перестановки 123, 132, 213, 231 и 312). ■

Если для набора равномерно распределенных на одном и том же отрезке независимых случайных величин $x_1, x_2, \ldots, x_n$ рассмотреть точки

$$A_1(1; x_1),$$
$$A_2(2; x_2),$$
$$\ldots\ldots$$
$$A_n(n; x_n),$$

то вероятность того, что ломаная $A_1A_2\ldots A_n$ выпукла вверх, равна

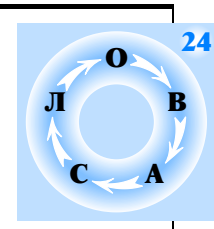
$$C_n/(n!)^2.$$

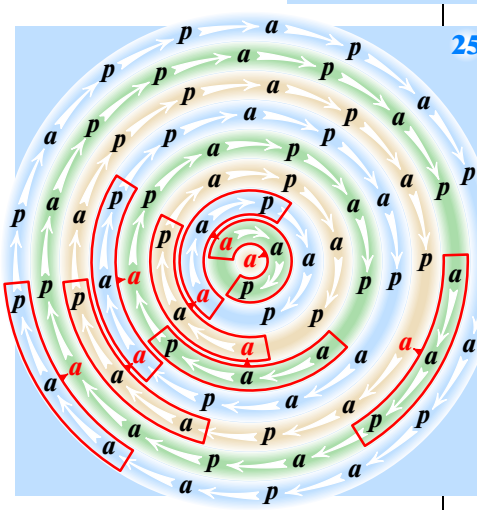
Рассмотрим внутри данного квадрата n точек, выбранных случайно и независимо. Пусть для любой из этих точек и для любой области квадрата вероятность того, что точка попадает в эту область, равна отношению площади области к площади квадрата. Оказывается, вероятность того, что все n точек являются вершинами своей выпуклой оболочки, равна

$$(C_n/(n-1)!)^2. \blacksquare$$

выражениям U и V уже сопоставлены слова u и v, то произведению $U\cdot V$ сопоставляем uvp. Например,

$$ab \mapsto aap,$$
$$(ab)c \mapsto aapap,$$
$$a(bc) \mapsto aaapp,$$
$$((ab)c)d \mapsto ((aap)c)d \mapsto (aapap)d \mapsto aapapap,$$
$$a(b(cd)) \mapsto aaaappp,$$
$$(a(bc))(de) \mapsto aaappaapp.$$





Как видите, каждому способу вычисления произведения n сомножителей соответствует слово из n букв a и $n-1$ букв p. Обратное неверно: например, слова *pppaa* или *apaap* не могут быть получены с помощью описанной процедуры. Ситуацию проясняет понятие циклической перестановки. Начнем с примера. Записав буквы СЛОВА по кругу, по часовой стрелке сможем прочитать одно из слов: СЛОВА, ЛОВАС, ОВАСЛ, ВАСЛО, АСЛОВ (рис. 24).

Вообще слово $a_{k+1}\ldots a_n a_1\ldots a_k$ будем называть циклической перестановкой слова $a_1\ldots a_k a_{k+1}\ldots a_n$. Оказывается, если W — слово из n букв a и $n-1$ букв p, то одно и только одно из циклически сравнимых с W слов получается из некоторой расстановки скобок вышеописанной конструкцией.

Докажем это индукцией по n. Среди циклически сравнимых с A слов имеет смысл рассматривать лишь те, которые не начинаются с буквы p, но кончаются ей. Такие, конечно, существуют. В таком слове W — это тоже легко понять — найдется подслово *aap*. Заменим его буквой a (рис. 25). Мы получили слово, в котором $n-1$ букв a и $n-2$ букв p, а для него применимо индукционное предположение.

Теперь выведем формулу для C_n. Среди $2n-1$ мест мы можем в точности C_{2n-1}^{n-1} способами выбрать $n-1$ место для букв p. Разделив C_{2n-1}^{n-1} на $2n-1$, получаем ответ:

$$C_n = \frac{1}{2n-1}C_{2n-1}^{n-1} = \frac{(2n-2)!}{(n-1)!\,n!}. \blacksquare$$

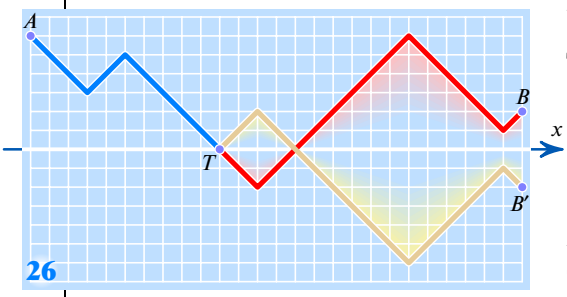
Лемма об отражении. Как помните, C_n равно количеству способов так расположить $n-1$ открывающих и $n-1$ закрывающих скобок в ряд, чтобы при чтении слева направо ни в какой момент число закрывшихся скобок не превосходило числа открывшихся. Заменив «(» на «+1», а «)» на «−1», получим, что C_n есть количество последовательностей из $n-1$ единиц и $n-1$ минус единиц, суммы всех начальных отрезков которых неотрицательны.

При этом каждой последовательности $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$, где $\varepsilon_k = \pm 1$, сопоставляем путь, выходящий из начала координат, k-й отрезок которого $(i=k, \ldots, n)$ является отрезком прямой с угловым коэффициентом ε_k, соединяющим точку $(k-1; \varepsilon_1+\ldots+\varepsilon_{k-1})$ с точкой $(k; \varepsilon_1+\ldots+\varepsilon_{k-1}+\varepsilon_k)$.

Через $N_{n,s}$ обозначим количество путей, начинающихся в начале координат и оканчивающихся в точке $(n; s)$. Очевидно, если среди $(\varepsilon_1, \ldots, \varepsilon_n)$ имеется p единиц и q минус единиц, то $s=p-q$. Поскольку p мест для положительных ε_k выбираются из $n=p+q$ имеющихся мест, $N_{p+q,p-q} = C_{p+q}^{p}$.

Пусть A и B — точки с целыми координатами, причем B лежит выше оси абсцисс, а B' симметрична B относительно оси абсцисс (рис. 26).

Лемма об отражении. ***Количество начинающихся в точке A и оканчивающихся в точке B путей, которые касаются оси абсцисс или пересекают ее, равно числу путей, оканчивающихся в точке B′.***



Доказательство. Если T — самая левая точка пути, попавшая на ось абсцисс, отразим участок AT относительно оси (см. рис. 26)!

Следующая теорема доказана в 1878 г. У. Уитвортом и в 1887 г. Ж. Бертраном.

Теорема о баллотировке. ***Если на выборах кандидат P набрал p голосов, а кандидат Q набрал q голосов, где $q \leqslant p$, то вероятность того, что при последовательном подсчете голосов P все время опережал Q, равна*** $\dfrac{p-q}{p+q}$.

Иными словами, если n и s — натуральные числа, то существует ровно $\dfrac{s}{n}N_{n,s}$ путей из начала координат в точку $(n; s)$, все точки которых, кроме начала координат, расположены выше оси абсцисс.

Доказательство. По принципу отражения, путей из (1; 1) в $(n; s)$, не задевающих ось абсцисс, имеется ровно

$$N_{n-1,s-1} - N_{n-1,s+1} = C_{p+q-1}^{p-1} - C_{p+q-1}^{p}.$$

Простая выкладка показывает, что правая часть равна $N_{n,s}\dfrac{p-q}{p+q}$.

Очевидно, число Каталана C_n получается, если в теореме о баллотировке положить $p=n$ и $q=n-1$. ■

Производящие функции — одно из мощных орудий комбинаторики. Идея состоит в том, чтобы «запаковать» всю бесконечную последовательность в одно выражение. Производящая функция для последовательности Каталана 1, 1, 2, 5, 14, 42, 132, 429, . . . — это функция

$$f(x) = x + x^2 + 2x^3 + 5x^4 + 14x^5 + 42x^6 + 132x^7 + 429x^8 + \ldots = \sum_{n=1}^{\infty} C_n x^n.$$

При этом пока нас даже не интересует, для каких x этот степенной ряд сходится: математики говорят в таких случаях, что мы рассматриваем формальный степенной ряд. Поскольку

$$f(x)\cdot f(x) = C_1^2 x^2 + (C_1C_2 + C_2C_1)x^3 + (C_1C_3 + C_2C_2 + C_3C_1)x^4 + \ldots$$
$$\ldots + (C_1C_{n-1} + C_2C_{n-2} + \ldots + C_{n-1}C_1)x^n + \ldots,$$

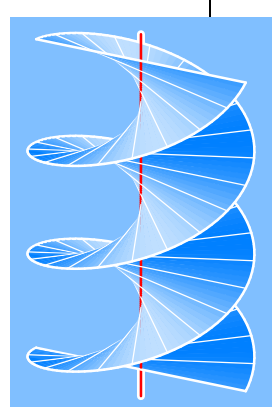
то $f^2(x) = -x + f(x)$. (Проверьте!) Решая квадратное уравнение относительно $f(x)$, получаем

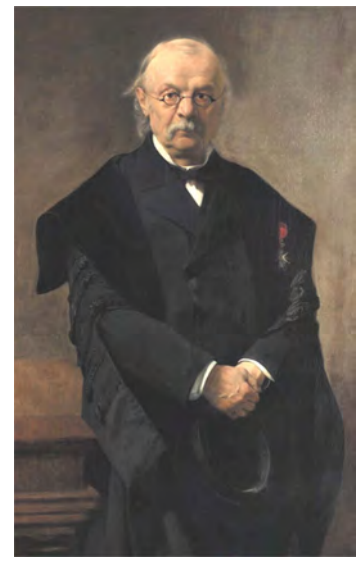
$$f(x) = \frac{1-\sqrt{1-4x}}{2}.$$

(Теперь ясно, что ряд для $f(x)$ сходится при $|x| < 1/4$. Впрочем, нам это не важно.) Здесь взят знак минус, так как ряд $f(x)$ не имеет свободного члена. Разложив правую часть в ряд по степеням x по биному Ньютона, получаем

$$C_n = -\frac{1}{2}\cdot\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)\cdots\left(\frac{3-2n}{2}\right)}{n!}\cdot(-4)^n.$$

Упрощая это выражение, получаем: $C_n = \dfrac{(2n-2)!}{n!\,(n-1)!}$. ■





Эжен Шарль Каталан (1814—1894) — бельгийский математик. Окончил Политехническую школу в Париже. Преподавал там же и в Сорбонне. В 1849 г. отказался присягнуть Наполеону III и был лишен права преподавания. С 1865 г. — профессор Льежского университета. Одновременно с К. Г. Я. Якоби и М. В. Остроградским предложил метод замены переменных в кратных интегралах. Каталану принадлежит постановка проблемы о том, что уравнение

$$x^y - z^t = 1$$

имеет единственное решение в натуральных числах, больших единицы: $3^2 - 2^3 = 1$.

Каталан доказал, что геликоид является единственной линейчатой минимальной поверхностью. Минимальная поверхность — это поверхность наименьшей площади с данным краем (такова, например, мыльная пленка). Линейчатая поверхность — это поверхность, образованная движением прямой линии. Геликоид (винтовая поверхность) — это поверхность, описываемая прямой, которая равномерно вращается вокруг пересекающейся с ней под углом α неподвижной оси и одновременно с постоянной скоростью движется вдоль этой оси. Параметрическое уравнение изображенного на рисунке геликоида, у которого $\alpha = 90°$, таково: $x = u\cos v$, $y = u\sin v$, $z = v$. ■

Графы без запрещенных подграфов

Обозначим через $f_k(n)$ наибольшее число отрезков, соединяющих n точек пространства так, чтобы не образовалось ни одного k-угольника. Оказывается, $f_3(n)=[n^2/4]$ и $\lim\limits_{n\to\infty}\dfrac{f_4(n)}{n\sqrt{n}}=\dfrac{1}{2}$. Первое равенство мы докажем тремя разными способами, а второе — при помощи конечной проективной плоскости.

Граф — это набор точек (вершин), соединенных линиями (ребрами). Не важно, где расположены и линиями какой формы соединены точки — важно лишь, какие точки соединены между собой, а какие не соединены. Например, на рисунке 1 тремя разными способами изображен один и тот же граф. У него 8 вершин — случайно возникающие на рисунках пересечения ребер вершинами не считаются.

Обозначим для краткости $f_3(n)=f(n)$. Очевидно, $f(2)=1$, $f(3)=2$ и $f(4)=4$ (рис. 2). На рисунке 3 показано, что $f(5)\geqslant 6$, $f(6)\geqslant 9$, $f(7)\geqslant 12$ и $f(8)\geqslant 16$.

Вообще, разобьем вершины графа на две примерно равные доли — по k вершин в каждой, если $n=2k$, и по k и $k+1$ вершин, если $n=2k+1$. Соединив каждую вершину одной доли с каждой вершиной другой доли, получим граф без треугольников. У него k^2 ребер при $n=2k$ и $k(k+1)$ ребер при $n=2k+1$. Поскольку $[(2k)^2/4]=k^2$ и $[(2k+1)^2/4]=[(4k^2+4k+1)/4]=k^2+k$, получаем неравенство $f(n)\geqslant[n^2/4]$.

Теорема П. Турана. $f_3(n)=[n^2/4]$.

Доказательство. I способ. Пусть в графе n вершин и ни одного треугольника. Рассмотрим вершину A наибольшей степени d. Если бы какие-то соединенные с ней вершины были соединены между собой, то образовался бы треугольник. Значит, каждое из ребер графа исходит либо из вершины A, либо из одной из $n-d-1$ вершин, не соединенных с A. Поэтому

$$f(n)\leqslant d+(n-d-1)d.$$

Следовательно, $f(n)\leqslant(n-d)d=\dfrac{n^2}{4}-\dfrac{n^2}{4}+nd-d^2=\dfrac{n^2}{4}-\left(\dfrac{n}{2}-d\right)^2\leqslant\dfrac{n^2}{4}$.

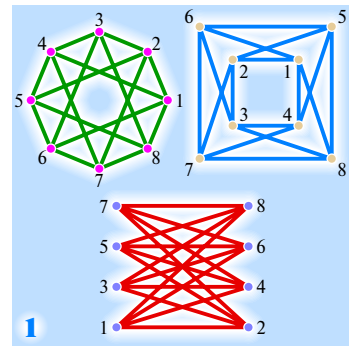
II способ — индукция по n, причем переход не от n к $n+1$, как обычно бывает, а от n к $n+2$. База состоит из двух (совершенно очевидных!) случаев $n=1$ и 2. Рассмотрим граф с $n+2$ вершинами, не содержащий треугольников. Пусть две его вершины A и B соединены ребром. Поскольку с любой из оставшихся n вершин может быть соединена не более чем одна из вершин A и B, то

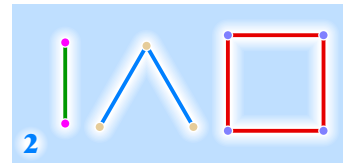
$$f(n+2)\leqslant 1+n+f(n)$$

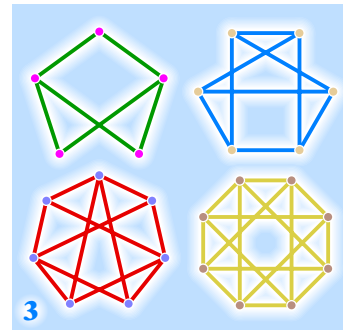
и, следовательно, $f(n+2)\leqslant 1+n+[n^2/4]=[(4+4n+n^2)/4]=[(n+2)^2/4]$.

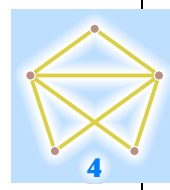
III способ — индукция от n к $n+1$. Доказательство состоит из двух лемм довольно общего характера.

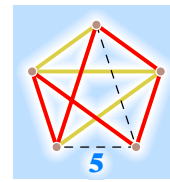
Пусть зафиксирован некоторый набор «запрещенных» подграфов. Обозначим через $f(n)$ наибольшее возможное число ребер графа, в котором n вершин и ни одного запрещенного подграфа.

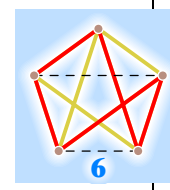


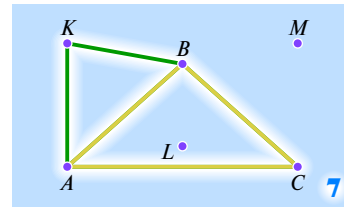


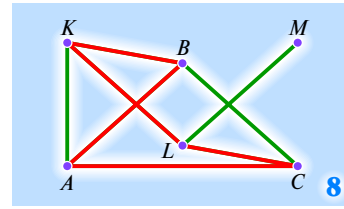


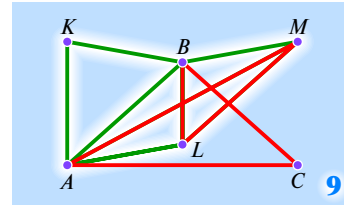












Если каждую из n вершин соединить со всеми остальными вершинами, то всего будет проведено $C_n^2 = n(n-1)/2$ ребер. Рассмотрим долю $f(n)/(n(n-1)/2)$ ребер, вошедших в граф. Оказывается, с возрастанием n эта доля не возрастает, то есть $\dfrac{f(n)}{n(n-1)/2} \geqslant \dfrac{f(n+1)}{(n+1)n/2}$.

Лемма 1. $\boldsymbol{\dfrac{f(n)}{n-1} \geqslant \dfrac{f(n+1)}{n+1}}$.

Доказательство. I способ. Рассмотрим граф без запрещенных подграфов, в котором $n+1$ вершин и $f(n+1)$ ребер. Отбросим любую из $n+1$ вершин — получим граф без запрещенных подграфов, у которого n вершин и не более $f(n)$ ребер.

Всего таких графов $n+1$ (мы могли выбросить любую вершину первоначального графа). Каждое ребро исходного графа мы посчитали $n-1$ раз. Следовательно,

$$(n+1)f(n) > (n-1)f(n+1).$$

II способ. Поскольку каждое ребро соединяет две вершины, то сумма степеней всех вершин графа с $f(n+1)$ ребрами равна $2f(n+1)$. Следовательно, если в графе $n+1$ вершин, то хотя бы из одной из них выходит не более $\dfrac{2f(n+1)}{n+1}$ ребер. Остальные n вершин соединены между собой не более чем $f(n)$ ребрами:

$$f(n+1) \leqslant f(n) + \frac{2f(n+1)}{n+1},$$

то есть $\dfrac{n-1}{n+1} f(n+1) \leqslant f(n)$. Лемма 1 доказана. ■

Лемма 2. ***Если*** $\boldsymbol{f(n) \leqslant \left[\dfrac{n^2}{4}\right]}$, ***то и*** $\boldsymbol{f(n+1) \leqslant \left[\dfrac{(n+1)^2}{4}\right]}$.

Доказательство. Если $n = 2k$, то

$$f(2k+1) \leqslant \frac{2k+1}{2k-1} f(2k) = \frac{2k+1}{2k-1} k^2 = k^2 + k + \frac{k}{2k-1} \leqslant k^2 + k + 1.$$

Если же $n = 2k-1$, то $f(2k) \leqslant \dfrac{2k}{2k-2} f(2k-1) = \dfrac{k}{k-1} k(k-1) = k^2$. ■

Очевидно, $f_5(5) \geqslant 7$ (рис. 4). Если у графа 5 вершин и 8 ребер, то два «отсутствующих» ребра или выходят из одной вершины (рис. 5), или не имеют общих вершин (рис. 6). В обоих случаях есть пятиугольник. Значит, $f_5(5) = 7$.

Лемма 3. $\boldsymbol{f_5(6) \leqslant 9}$.

Доказательство. Пусть существует граф с 6 вершинами и 10 ребрами. Поскольку $f_3(6) < 10$, есть хотя бы один треугольник. Обозначим его вершины буквами A, B и C. Поскольку остальные три вершины K, L и M соединены не более чем 3 ребрами, то ребер, один из конец которых — A, B или C, а другой — K, L или M, не менее чем $10-3-3=4$. Значит, достаточно разобрать случай, когда K соединена с A и B (рис. 7).

Если бы степень вершины L была меньше 3, то равенство $f_5(5) = 7$ дало бы противоречие. Вершина L не может быть соединена более чем с двумя из вершин A, B, C и K. Поэтому L соединена с M и с двумя из вершин A, B, C и K. Соединение L с K и C (рис. 8) дает пятиугольник $ABKLC$. Значит, L соединена с A и B. Поскольку вершина M ничем не хуже вершины L, то и M соединена с A и B (рис. 9) и очевиден пятиугольник $ACBLM$.

Лемма 3 доказана. Поскольку в двудольном графе нет не только треугольников, но и пятиугольников, то $f_5(n) \geqslant [n^2/4]$. В частности, $f_5(6) = 9$. Теперь, применяя лемму 2, получаем формулу $f_5(n) = [n^2/4]$ для любого $n \geqslant 6$. ■

В стране n городов и несколько авиалиний. Никакие четыре города не соединены между собой более чем четырьмя авиалиниями. Какое наибольшее количество авиалиний может быть в этой стране?

Ответ: $[n^2/4]$. **Указание.** Заметьте, что ответ верен при $n=4$, и примените лемму 1. ■

В пространстве расположено $2n$ точек, $n \geqslant 2$. Никакие четыре точки не лежат в одной плоскости. Проведено n^2+1 отрезков с концами в этих точках. Докажем, что возникло не менее n треугольников.

При $n=2$ утверждение легко проверить непосредственно. Докажем утверждение для $2n+2$ точек, считая его верным для $2n$ точек. Согласно теореме 1, проведенные отрезки образуют хотя бы один треугольник ABC. Обозначим количества отрезков, выходящих из вершин треугольника ABC (не считая его сторон), через k_a, k_b и k_c соответственно и рассмотрим два случая.

1) Если $k_a + k_b + k_c \leqslant 3n-2$, то среди точек A, B, C найдутся такие две точки, что сумма соответствующих чисел k не превосходит $2n-2$. Выбросив эти две точки и все выходящие из них отрезки, получаем $2n$ точек, соединенных не менее чем $(n+1)^2+1-(2n-2)-3 = n^2+1$ отрезками, что вместе с $\triangle ABC$ дает не менее n треугольников.

2) Если $k_a + k_b + k_c \geqslant 3n-1$, то обозначим через N_k, где $k=0$, 1, 2 или 3, количество тех из остальных $2n-1$ точек, которые соединены ровно с k вершинами треугольника ABC. Тогда

$$N_1 + N_2 + N_3 \leqslant 2n-1,$$
$$N_1 + 2N_2 + 3N_3 = k_a + k_b + k_c.$$

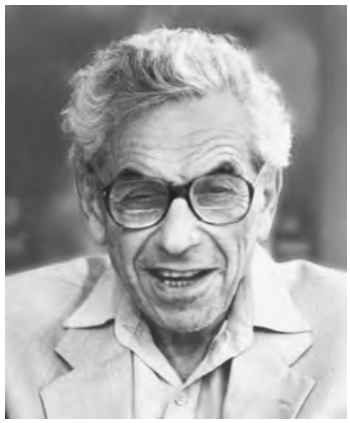
Следовательно, кроме $\triangle ABC$, есть еще не менее

$$N_2 + 3N_3 \geqslant N_2 + 2N_3 = (N_1 + 2N_2 + 3N_3) - (N_1 + N_2 + N_3) \geqslant 3n - 1 - (2n-1) = n$$

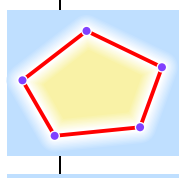
треугольников. ■

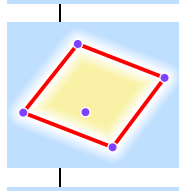
В 1972—1977 гг. Д. Р. Вудал, Г. Н. Копылов и П. Эрдёш доказали, что

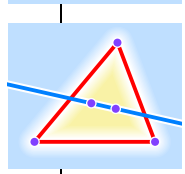
$$f_{2k+1}(n) = \begin{cases} C_n^2, & \text{если } n \leqslant 2k, \\ C_{2k}^2 + C_{n-2k+1}^2, & \\ & \text{если } 2k < n \leqslant 4k-1, \\ [n^2/2], & \text{если } n > 4k-1. \end{cases}$$ ■



Паль Эрдёш (1913—1996) — венгерский математик. С наукой его познакомили родители, оба — учителя математики. Отец был призван в армию, попал в плен и пробыл в Сибири 6 лет, где изучил английский язык по словарю, с неправильным произношением слов. Вернувшись, научил английскому сына, передав ему свой неповторимый акцент. Почти вся семья Эрдёша погибла в 1940-е гг. во время холокоста, сам он бежал в Англию, затем в США. В 1954 г. не получил разрешения вернуться в США после поездки на одну из научных конференций; визу в США получил лишь в 1964 г. В 1933 г. выдвинул и по сей день недоказанную гипотезу, что из любых $2^{n-2}+1$ точек плоскости, никакие три из которых не лежат на одной прямой, можно выбрать n вершин выпуклого n-угольника. Для $n=4$ доказательство состоит в том, что выпуклая оболочка $2^{4-2}+1=5$ точек (наименьший выпуклый многоугольник, содержащий их) является пятиугольником, четырехугольником или треугольником. Первые два случая очевидны, а в третьем надо провести прямую через две точки, попавшие внутрь выпуклой оболочки. Автор более 1400 статей и целого ряда интригующих гипотез. Называл себя машиной по переработке кофе в математические задачи. Говорил, что на небесах есть Книга, где записаны божественно красивые доказательства теорем; призвание математика — прочитать хотя бы одну страницу или абзац этой Книги. ■







Рассмотрим конечный набор графов $G_1, G_2, \ldots, G_r$ и через $f(n)$ обозначим наибольшее число ребер графа на n вершинах, не содержащего ни одного из графов $G_1, G_2, \ldots, G_r$ в качестве подграфа.

Теорема Эрдёша—Шимоновича.

$$f(n)=\left(1+\frac{1}{1-\min\limits_{1\leqslant k\leqslant r}\chi(G_k)}\right)\frac{n^2}{2}+o(n^2),$$

где $\chi(G)$ — хроматическое число графа G, то есть наименьшее количество красок, необходимое для того, чтобы можно было покрасить все вершины графа, не выкрасив никакие две соединенные ребром вершины в один цвет.

Хроматическое число треугольника и пятиугольника равно 3, а четырехугольника — 2. Для функции f_4 теорема, доказанная Эрдёшем и М. Шимоновичем, утверждает лишь, что $f_4(n)=o(n^2)$, то есть $\lim\limits_{n\to\infty}\dfrac{f_4(n)}{n^2}=0$. Мы докажем более сильный результат (хотя точную формулу для $f_4(n)$ не найдем, да и вряд ли это возможно).

Теорема 2. $\lim\limits_{n\to\infty}\dfrac{f_4(n)}{n\sqrt{n}}=\dfrac{1}{2}$.

Доказательство. Пусть в графе без четырехугольников n вершин и f ребер. Занумеруем вершины числами от 1 до n. Пусть из первой, второй, ..., n-й его вершины выходит, соответственно, $d_1, d_2, \ldots, d_n$ ребер.

Назовем пару вершин занятой, если существует вершина, соединенная с ними обеими. Если бы пара была занята дважды, то образовался бы четырехугольник. Вершина, из которой выходит d ребер, создает C_d^2 занятых пар. Следовательно,

$$\frac{d_1(d_1-1)}{2}+\frac{d_2(d_2-1)}{2}+\ldots+\frac{d_n(d_n-1)}{2}\leqslant\frac{n(n-1)}{2},$$

то есть $d_1^2+d_2^2+\ldots+d_n^2-2f\leqslant n^2-n$. В силу неравенства о среднем арифметическом и среднем квадратичном имеем

$$d_1^2+d_2^2+\ldots+d_n^2\geqslant\frac{(d_1+d_2+\ldots+d_n)^2}{n}.$$

Следовательно,

$$\frac{4f^2}{n}-2f\leqslant n^2-n,$$

откуда $f\leqslant\dfrac{2+\sqrt{4+16(n-1)}}{8/n}=\dfrac{n}{4}(1+\sqrt{4n-3})<\dfrac{n}{4}(1+2\sqrt{n})=\dfrac{n\sqrt{n}}{2}+\dfrac{n}{4}$. ■

Оценку снизу получить труднее. При этом нам потребуется знать, что отношение p_{n+1}/p_n последовательных простых чисел стремится к 1 при $n\to\infty$. Это следует из доказанного Ж. Адамаром (1865—1963) и Ш. Ж. де ла Валле Пуссеном (1866—1962) в 1896 г. асимптотического закона распределения простых чисел. Доказательство использует методы теории функций комплексного переменного. А. Сельберг (р. 1917) и П. Эрдёш в 1949 г. предложили элементарное — то есть не использующее эти методы — доказательство закона распределения простых чисел. Но оно тоже очень сложное. ■

Для получения оценки снизу рассмотрим проективную плоскость над полем из p элементов, где p — простое число. Точка этой плоскости — это тройка $(x:y:z)$ вычетов по модулю p, не все из которых равны 0.

При этом тройки $(x:y:z)$ и $(kx:ky:kz)$, где $k\neq 0$, задают одну и ту же точку.

Проективная плоскость состоит из p^2+p+1 точек и имеет столько же прямых. Соединим точки $(x:y:z)$ и $(a:b:c)$ ребром в том и только том случае, когда

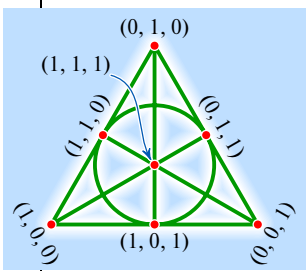
$$ax+by+cz\equiv 0 \bmod p,$$

то есть точку $(a:b:c)$ соединяем со всеми точками двойственной этой точке прямой. Вообще говоря, точка может оказаться на этой прямой. Но даже в этом случае она будет соединена не менее чем с p другими точками, так что

$$f_4(p^2+p+1)\geqslant \frac{p(p^2+p+1)}{2}.$$

Значит, при $n=p^2+p+1$ имеем $f_4(n)>\dfrac{(\sqrt{n}-1)n}{2}$.

Поскольку отношение двух последовательных простых чисел стремится к 1, сравнивая оценки сверху и снизу, получаем: $\dfrac{f_4(n)}{n\sqrt{n}}\to\dfrac{1}{2}$ при $n\to\infty$. ■



Проективная плоскость над полем $\mathbb{Z}_2$ состоит из $2^2+2+1=7$ точек

На Всесоюзной олимпиаде в 1975 г. школьникам предлагали такую задачу:

а) В квадрате 7×7 нужно отметить центры k клеток так, чтобы никакие четыре отмеченные клетки не являлись вершинами прямоугольника со сторонами, параллельными сторонам квадрата. При каком наибольшем k это возможно?

б) Решите аналогичную задачу для квадрата 13×13 клеток.

Оценки сверху — 21 и 52 — получаются тем же способом, что оценка сверху функции f_4. Примеры, реализующие эти оценку, участники олимпиады искали подбором, из соображений симметрии. А для нас это не проблема!

В примере а) в качестве «номеров» строк и столбцов используем тройки из чисел 0 и 1, отличные от (0; 0; 0). Их как раз 7: (1; 1; 1), (1; 1; 0), (1; 0; 1), (0; 1; 1), (1; 0; 0), (0; 1; 0), (0; 0; 1). Клетку на пересечении строки $(a; b; c)$ и столбца $(x; y; z)$ отмечаем, если $ax+by+cz$ четно (рис. 10).

10	111	110	101	011	100	010	001
111		●	●	●			
110	●	●					●
101	●		●			●	
011	●			●	●		
100				●		●	●
010			●		●		●
001		●			●	●	

В примере б) в качестве «номеров» строк и столбцов используем тройки из чисел 0, 1 и 2, отличные от (0; 0; 0), причем из двух троек, получающихся одна из другой заменой 1 на 2 и 2 на 1 (то есть умножением на -1 по модулю 3), используем лишь одну. Их как раз 13 штук: (1; 1; 1) и по три перестановки каждой из троек (1; 1; 0), (1; 1; 2), (1; 2; 0) и (1; 0; 0). Клетку на пересечении строки $(a; b; c)$ и столбца $(x; y; z)$ отмечаем, если $ax+by+cz$ делится на 3 (рис. 11). ■

11	111	110	101	011	112	121	211	120	201	012	001	010	100
111	●							●	●	●			
110						●	●	●			●		
101					●		●		●			●	
011					●	●				●			●
112				●	●	●			●				
121		●		●			●			●			
211		●	●					●			●		
120	●	●			●							●	
201	●		●			●							●
012	●			●			●						●
001		●						●				●	●
010			●						●		●		●
100				●						●	●	●	

Клетки таблицы заполнены буквами В, З, М, Ш, покрашены четырьмя цветами и обведены рамками четырех типов так, что в каждой строке и в каждом столбце встречаются все буквы, цвета и рамки; каждая буква покрашена по одному разу каждым цветом; рамка каждого типа содержит каждую букву и каждый цвет тоже по одному разу.

Вообще, когда таблица $n\times n$ заполнена n знаками так, что в каждой строке и в каждом столбце каждый знак встречается по одному разу, то говорят, что задан латинский квадрат. Два латинских квадрата называют ортогональными, если в тех n клетках, где в первом квадрате стоит некоторый знак z, в другом квадрате все знаки различны (и так для каждого z).

Рисунок — это три ортогональных друг другу латинских квадрата размером 4×4. (Один дает букву, другой — цвет, третий — форму рамки.)

При построении таких таблиц (их используют, например, при планировании многоцелевых экспериментов) полезны конечные аффинные плоскости.

Пусть F — конечное поле из q элементов; пары $(x; y)$ его элементов будем называть точками конечной аффинной плоскости, а множества решений уравнения $ax+by=c$, где хотя бы один из элементов a и b поля F отличен от 0, — прямыми. Всего получается q^2 точек и $q(q+1)$ прямых, причем они разбиваются на $q+1$ семейств, каждое из которых содержит q параллельных между собой прямых (их уравнения получаются из уравнения $ax+by=c$ одной из них варьированием c).

Для $q=4$ (как и для любого q, являющегося простым числом или степенью простого числа) существует конечное поле из q элементов. Поставим в соответствие каждому из $q+1=5$ семейств параллельных прямых определенное свойство: «номер строки», «номер столбца», «буква», «цвет», «форма рамки». Каждые две непараллельные прямые пересекаются в одной точке, поэтому для каждых двух заданных свойств есть ровно одна клетка — точка аффинной плоскости — с нужной парой свойств. ■

Таблицы сложения и умножения в поле из 4 элементов.

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$\times$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

Аксель Туэ (1863—1922) — норвежский математик. Доказал, что уравнение

$$a_0x^n+a_1x^{n-1}y+a_2x^{n-2}y^2+\ldots \\ \ldots+a_ny^n=b,$$

где $n>2$ и $a_0, a_1, \ldots, a_n$ — целые числа, причем левая часть не разложима на множители с целыми коэффициентами, имеет лишь конечное (быть может, пустое) множество решений в целых числах x и y. ■

Слово называют словом без перекрытий, если в нем нет перекрывающихся вхождений одного и того же слова, то есть нет подслова вида $xy=zx$, где слово y короче слова x.

Теорема. ***Слово свободно от перекрытий тогда и только тогда, когда оно сильно бескубно.***

Доказательство. Рассмотрим слово, не являющееся сильно бескубным, и его подслово вида $\alpha y\alpha y\alpha$. Очевидны два перекрывающиеся вхождения слова $\alpha y\alpha$. Обратно, если слово не свободно от перекрытий, то в нем есть подслово $xy=zx$, где z короче, чем x. Пусть α — первая буква слова z. Тогда $x=zt$, где первая буква слова t — тоже α. Следовательно, $zz\alpha$ — подслово рассматриваемого слова. ■

Рассмотрим слова

А,
АБА,
АБАВАБА,
АБАВАБАГАБАВАБА,
. . . . . .

вообще, написав очередное слово v, возьмем новую букву алфавита λ и напишем слово $v\lambda v$. Так для любого алфавита из n букв можно построить бесквадратное слово из 2^n-1 букв, которое перестает быть бесквадратным при приписывании к нему любой буквы этого алфавита. ■

0	А	10000	Б
1	Б	10001	А
10	Б	10010	А
11	А	10011	Б
100	Б	10100	А
101	А	10101	Б
110	А	10110	Б
111	Б	10111	А
1000	Б	11000	А
1001	А	11001	Б
1010	А	11010	Б
1011	Б	11011	А
1100	А	11100	Б
1101	Б	11101	А
1110	Б	11110	А
1111	А	11111	Б

Бесповторные последовательности

В начале XX в. А. Туэ рассмотрел бесконечное слово

АББАБААББААБАББАБААБАББААББАБААБ. . . ,

n-я буква которого А *или* Б *в зависимости от того, нечетно или четно количество единиц в двоичной записи числа* $n-1$. *Это слово бескубно: никакое подслово не повторяется в нем три раза подряд. Расширив алфавит до трех букв, он построил бесконечное бесквадратное слово.*

Начав с одной буквы А, будем заменять А↦АБ и Б↦БА. Получим последовательность слов

$w_0=$А,
$w_1=$АБ,
$w_2=$АББА,
$w_3=$АББАБААБ,
$w_4=$АББАБААББААБАББА,
. . . . . .

Теорема 1. ***Для любого n слово w_{n+1} получается из w_n приписыванием к нему слова $\overline{w_n}$, полученного из w_n заменой всех* А *на* Б *и* Б *на* А.**

Доказательство. Обозначив операцию замены А↦АБ и Б↦БА буквой h, применим индукцию. При $n=0$ утверждение верно. Переход от n к $n+1$ тоже очевиден:

$$w_{n+2}=h(w_{n+1})=h(w_n\overline{w_n})=h(w_n)\,h(\overline{w_n})=h(w_n)\,\overline{h(w_n)}=w_{n+1}\overline{w_{n+1}}.$$

Итак, w_n является началом w_{n+1}, так что существует бесконечное слово Туэ, началами которого являются все w_n.

Теорема 2. ***Сопоставив каждому целому неотрицательному числу n букву* А *или* Б *в зависимости от того, четна или нечетна сумма цифр двоичной записи числа n, получим слово Туэ.***

Доказательство очевидно из таблицы, показывающей переход от w_4 к w_5.

Определение. Слово называют бесквадратным (бескубным), если оно не содержит двух (соответственно трех) подряд идущих одинаковых подслов.

Например, слово АНАПРАНА бесквадратное: подслова АНА разделены буквами ПР. А слово АНАРРКУВ не является бесквадратным, ибо содержит две подряд идущие буквы Р. Доказать бескубность слова Туэ — то же самое, что для любого натурального n доказать бескубность слова w_n. Обычно такого рода утверждения доказывают по индукции. К сожалению, наивная попытка ничего не дает: непонятно, как из бескубности слова w_n можно вывести бескубность слова w_{n+1}. Придется доказывать более сильное утверждение: не бескубность, а сильную бескубность!

Определение. Слово называют сильно бескубным, если в нем нет подслов вида $xx\alpha$, где x — непустое (то есть состоящее не менее чем из одной буквы) слово, α — его первая буква.
Иначе говоря, слово сильно бескубное, если в нем нет подслов вида $\alpha y\alpha y\alpha$, где α — буква, y — слово (возможно, пустое).

Теорема Туэ. ***Слово Туэ сильно бескубно.***

Доказательство проведем методом бесконечного спуска: предположив, что в слове Туэ w есть подслово вида $\alpha y\alpha y\alpha$, докажем, что в нем есть и более короткое подслово того же типа. Этого, поскольку длина слова не может бесконечно убывать, достаточно для получения противоречия.
Как вы помните, $w=h(w)$. Посмотрим, из какого слова применением h получается слово, содержащее $\alpha y\alpha y\alpha$.
Поскольку длина слова $\alpha y\alpha y\alpha$ нечетна, одна из двух крайних букв α находится в четном разряде (считаем, что слово Туэ начинается с буквы номер 0). Для определенности, пусть такова левая буква α.
Если длина слова y нечетна, то $\alpha y\alpha y\alpha$ получается из $h^{-1}(\alpha y)\,h^{-1}(\alpha y)\alpha$, и все доказано. (Операция h^{-1} обратна операции h: она не увеличивает, а уменьшает длины слов вдвое.)
Если же длина y четна, то y должно начинаться и кончаться на отличную от α букву β, то есть $y=\beta z\beta$. Рассматривая $\alpha\beta z\beta\alpha\beta z\beta\alpha$, видим, что z начинается и кончается на α. Продолжая это бесконечно, получаем противоречие. ■

Построим бесквадратное бесконечное слово, в написании которого использованы лишь три буквы. Мы получим его из слова Туэ, применяя весьма употребительный в теории языков прием, состоящий в группировке нескольких букв в одну.
Пройдем вдоль слова, вставляя между двумя соседними буквами А единицу, между буквами А и Б — двойку, между Б и А — тройку, а между соседними

буквами Б — четверку. Если бы в новом слове встретились два одинаковых подряд идущих подслова, то, как нетрудно сообразить, слово Туэ не было бы сильно бескубным. Значит, мы построили бесквадратное слово в алфавите, состоящем из цифр 1, 2, 3, 4.
Заметьте: перед единицей не может стоять никакая цифра, кроме тройки (иначе в слове Туэ оказались бы три буквы А подряд). А после единицы непременно идет двойка. Перед любой четверкой — двойка, а после четверки — тройка. Поэтому мы можем везде заменить цифру 4 на 1 — слово останется бесквадратным! ■

Мы привели явный пример бесквадратного бесконечного слова. Единственно ли оно? Конечно нет. Построим бесквадратное бесконечное слово (правда, не в трехциферном, а в десяти- или хотя бы четырехциферном алфавите) совсем другим способом.
Для этого оценим снизу количество $f(n)$ бесквадратных (не содержащих ни одной пары рядом стоящих одинаковых групп цифр) последовательностей длины n, составленных из цифр 0, 1, . . . , 9, а затем из существования сколь угодно длинных бесквадратных слов выведем существование бесконечного бесквадратного слова.

Назовем два слова u и v коммутирующими, если слова uv и vu, получающиеся приписыванием одного к другому в разном порядке, совпадают. Например, слова $u=$ ПАПА и $v=$ ПА коммутируют: при любом приписывании получается ПАПАПА. А вот МЯУ и КИС не коммутируют:

$$\text{МЯУКИС} \neq \text{КИСМЯУ}.$$

Нетрудно доказать, что два слова коммутируют тогда и только тогда, когда они являются степенями одного и того же слова. Чуть труднее — что если даны три слова, первое из которых коммутирует со вторым, а второе с третьим, то все три являются степенями одного и того же слова. ■

Доля бесквадратных слов длины n среди всех слов длины n стремится к 0 при $n\to\infty$. Точнее говоря, для любого $\varepsilon>0$ для всех достаточно больших n имеем $f(n)<\varepsilon\cdot 10^n$.
В самом деле, поскольку одинаковые цифры в бесквадратной последовательности не могут стоять рядом и поскольку пар различных цифр 90, разбив последовательность длины $2n$ на n пар, имеем: $f(2n)\leqslant 90^n$. ■

Среди обычных правил игры в шахматы есть правила о ничьей, обеспечивающие, в частности, конечность числа ходов любой игры. Рассмотрим следующую их модификацию. Начнем с очевидного: фигуры и пешки на доске образуют различные конфигурации, причем возможно лишь конечное число различных конфигураций. (В конфигурацию включаем информацию о том, чей сейчас ход.) Пусть ничью объявляют, когда некоторая последовательность конфигураций повторилась два раза подряд. (Отменим правило о том, что после определенного числа ходов должна быть сдвинута пешка или взята фигура.) Покажите, что при таких правилах возможна бесконечная игра, даже если у каждого из игроков остался лишь король! ■

Лемма 1. $f(n+1) > 10f(n) - f(n) - f(n-1) - \ldots - f(1)$.

Доказательство. Допустим, что выписаны все бесквадратные последовательности длины n — всего $f(n)$ штук, и мы хотим выписать бесквадратные последовательности длины $n+1$.

Возьмем произвольную бесквадратную последовательность $a_1a_2 \ldots a_n$ и напишем 10 последовательностей длины $n+1$:

$$a_1a_2 \ldots a_nk, \qquad \text{где } k = 0, 1, \ldots, 9.$$

Через f_r обозначим количеств тех из них, у которых в конце два раза повторяется последовательность длины r. Очевидно,

$$f(n+1) \geqslant 10f(n) - f_1 - f_2 - \ldots - f_{[(n+1)/2]}.$$

При этом $f_r \leqslant f(n+1-r)$. Действительно, последовательности, количество которых мы обозначили f_r, имеют вид $a_1a_2 \ldots a_1b_1 \ldots b_rb_1 \ldots b_r$, где $l + 2r = n+1$ и последовательность $a_1a_2 \ldots a_1b_1 \ldots a_kb_1 \ldots b_{r-1}$ бесквадратная. Осталось сопоставить каждой такой последовательности бесквадратную последовательность $a_1a_2 \ldots a_1b_1 \ldots b_r$ длины $n+1-r$.

Теорема 2. $f(n+1) > 8f(n)$.

Следствие. $f(n) > 8^n$.

Доказательство теоремы — индукция по n. База: $f(2) = 90 > 8f(1)$. Переход. Пусть уже известно, что $f(n) > 8f(n-1) > 8^2f(n-2) > 8^3f(n-3) > \ldots$ Тогда

$$f(n+1) > 10f(n) - f(n)\left(1 + \frac{1}{8} + \frac{1}{8^2} + \ldots + \frac{1}{8^{n-1}}\right) > 10f(n) - 2f(n) = 8f(n).$$

Аналогично можно доказать, что для любого n количество бесквадратных последовательностей длины n, составленных из цифр 1, 2, 3 и 4, больше 2^n.

Теорема 3. ***Если имеется некоторый список запрещенных слов и сколь угодно длинные слова без запрещенных подслов, то существует и бесконечно длинное слово без них.***

Доказательство. Выпишем слова U_1, U_2, U_3, ..., не содержащие запрещенных подслов, позаботившись, чтобы длины этих слов становились все больше и больше. Посмотрим, какие буквы находятся на первом месте. Поскольку алфавит конечен, а слов бесконечно много, с какой-то буквы слова начнутся бесконечно много раз. (Таких букв может быть несколько, но мы сосредоточим внимание на какой-нибудь одной.) Вычеркнув все слова с другими первыми буквами, рассмотрим вторую букву всех оставшихся слов. Какая-то буква во второй позиции встречается бесконечно много раз. Вычеркнув все слова с иными вторыми буквами, перейдем к третьей букве и так далее. ■

Аналогично нетрудно доказать и следующие утверждения.

Если человечество бессмертно, а люди смертны, то существует бесконечная цепочка женщин, где каждая — дочь предыдущей.

Если имеется некоторый конечный список запрещенных слов и сколь угодно длинные слова без запрещенных подслов, то существует и бесконечно длинное периодическое слово без них.

Теорема Ч. Арцелы (1847—1912). ***Из любой равномерно ограниченной и равностепенно непрерывной последовательности определенных на компакте функций можно выбрать равномерно сходящуюся подпоследовательность.*** ■

По городу с прямоугольной планировкой ездит велосипедист. Попав на перекресток, он либо продолжает ехать прямо, либо поворачивает налево или направо (поворачивать назад нельзя!) — с учетом того, что правилами уличного движения запрещены некоторые комбинации этих поворотов

В одном из 1000 окопов, расположенных в ряд, спрятался пехотинец. Пушка может одним выстрелом накрыть любой окоп. В каждом промежутке между выстрелами пехотинец (если уцелел) обязан перебежать в соседний окоп (даже если этот окоп только что обстрелян или в него полетит следующий снаряд). Сможет ли пушка попасть в пехотинца, если он заранее знает расписание стрельбы?

Занумеруем окопы слева направо числами от 1 до 1000. Предположим сначала, что к началу обстрела пехотинец сидит в окопе с четным номером. Выстрелим во второй окоп. Если не попали, пехотинец перебежит в окоп с нечетным номером. Выстрелим в третий окоп. Если не попали, то пехотинец перебежит в окоп с четным номером, не меньшим 4. Выстрелим в окоп номер 4, затем в окоп номер 5 и так далее, оттесняя пехотинца. Во время выстрела в 998-й окоп он будет в 1000-м окопе и вынужден будет перейти из него в 999-й окоп под очередной выстрел.

Если же вначале пехотинец находился в окопе с нечетным номером, то после 998 выстрелов он вновь — в окопе с нечетным номером. Повторно стреляем в 999-й, 998-й, 997-й, ..., 2-й окопы. Всего 1996 выстрелов.

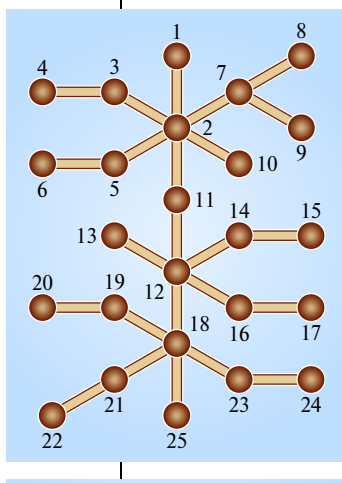
Теорема. ***Для цепочки из n окопов необходимы $2n-4$ выстрела.***

Доказательство. Если выстрелов $2n-5$, то четных (по номеру выстрела, а не окопа) среди них $n-3$. Следовательно, минимум в 3 окопа четные выстрелы не попадали. Среди этих окопов есть хотя бы один не крайний окоп A. Пехотинец может пережидать четные выстрелы в A, а любой нечетный выстрел — в том из соседних с A окопов, куда этот выстрел не направлен.

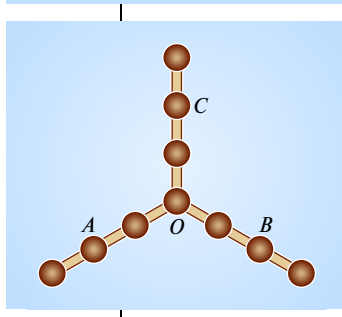
Является ли линейная цепочка окопов единственной, которая не позволяет пехотинцу спастись? Очевидно, если в графе есть хотя бы один цикл, то пехотинец уцелеет, перемещаясь вдоль этого цикла. Поэтому надо исследовать графы без циклов — деревья.

Если дерево можно представить в виде «ствола», из некоторых вершин которого выходят «отростки» длиной 1 или 2, то у пушки есть следующая стратегия. Как обычно, будет два прохода, чтобы угадать четность. По стволу

«огневой вал» пойдет подряд от одного края к другому. После выстрела в вершину A, из которой выходят один или несколько отростков, следует выстрел в соседнюю с A вершину первого отростка, затем снова в A, затем в соседнюю с A вершину второго отростка, и так пока не обстреляны все отростки. Затем снова выстрел в A, и обстрел идет по стволу дальше.



Например, для изображенного на рисунке дерева последовательность выстрелов такова: 2, 3, 2, 5, 2, 7, 2, 11, 12, 14, 12, 16, 12, 18, 19, 18, 21, 18, 23, 18 и в обратном порядке — 18, 23, 18, 21, 18, 19, 18, 12, 16, 12, 14, 12, 11, 2, 7, 2, 5, 2, 3, 2.



Теорема. ***На дереве, из некоторой вершины которого выходят три линии не менее чем по три окопа в каждой, пехотинец может спастись.***

Доказательство. Пусть пехотинец находится в одном из окопов O, A, B, C, причем об одном из этих окопов мы знаем, что пехотинца там нет. Сначала пусть его нет в окопе O. Окопы A, B, C равноправны. Выстрелим в A. После этого пехотинец сможет перебежать в один из четырех окопов, соседних с B и C. Куда ни стреляй, пехотинец сможет оказаться в одном из окопов O, B, C, — мы вернулись к первоначальной ситуации!

Если же изначально пехотинца нет в окопе A, то стрелять в окоп O, как мы выяснили, не имеет смысла. Значит, стрелять надо в B или C, для определенности — в B. После этого пехотинец может перебежать в один из четырех окопов, соседних с O или C. Куда ни стреляй, он затем может оказаться в вершинах O, C и в одной из вершин A и B.

Итак, какова бы ни была конечная последовательность выстрелов, пехотинец может найти себе такой первоначальный окоп и так перебегать между выстрелами, что останется жив. Осталось применить теорему 3. ■

(налево или направо) и проезда прямо. Известно, что все запрещенные комбинации имеют длину не меньше двух и что нет различных запрещенных комбинаций одинаковой длины. Оказывается, велосипедист может разъезжать по городу, не нарушая правил, сколь угодно долго.

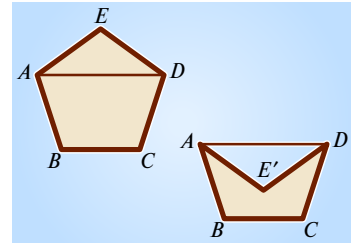
Математическая формулировка этой задачи такова. Рассмотрим алфавит из трех букв А, Б, В. Некоторые буквосочетания (длины два и более) запретим. Пусть в списке запрещенных буквосочетаний все слова разной длины. Докажем, что существование сколь угодно длинных слов, не содержащих запрещенных подслов. (А значит, в силу теоремы 3, и существование бесконечного слова без запрещенных подслов.)

Слово будем называть хорошим, если оно не содержит ни одного запрещенного подслова. Число хороших слов длины n обозначим $g(n)$. Для единообразия пусть $g(0)=1$.

Лемма 2. $\boldsymbol{g(n+1)\geqslant 3g(n)-g(n-1)-g(n-2)-\ldots-g(1)-g(0).}$

Доказательство. Заметим, что все хорошие слова длины $n+1$ содержатся среди слов, получающихся приписыванием справа одной из букв А, Б, В к хорошим словам длины n. При таком приписывании получается $3g(n)$ слов, среди которых могут быть и нехорошие. Запрещенное буквосочетание может находиться только в конце, то есть это слово состоит из хорошего слова длины $n+1-k$ и запрещенного слова длины k, где $k\geqslant 2$. Поскольку запрещенное слово длины k, самое большее, одно, количество нехороших слов этого типа не превышает $g(n+1-k)$.

Отразим вершину E правильного пятиугольника $ABCDE$ относительно диагонали AD. Получим невыпуклый пятиугольник, копиями которого можно покрыть плоскость как периодически, так и не периодически.

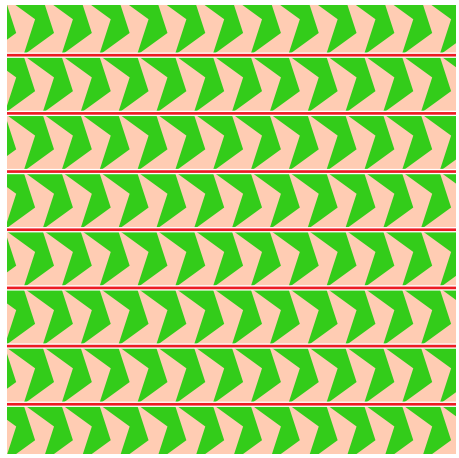


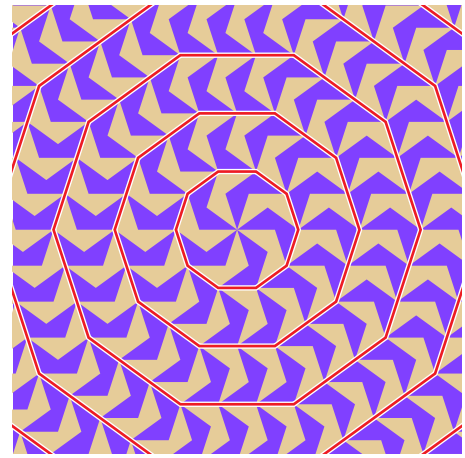
Теорема 4. $\boldsymbol{g(n+1)>2g(n).}$

Доказательство аналогично доказательству теоремы 2.

Следствие. $\boldsymbol{g(n)>2^n.}$

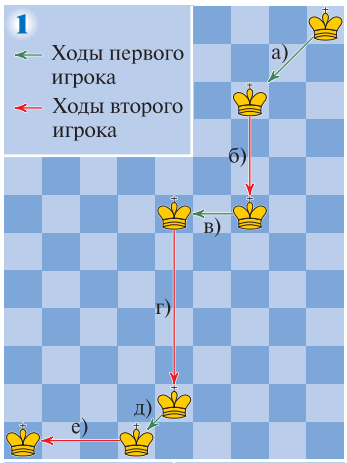
Доказательство — индукция по n. ■

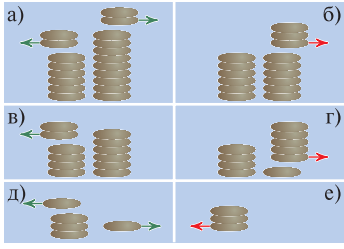


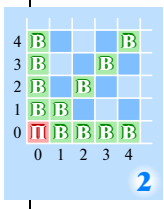


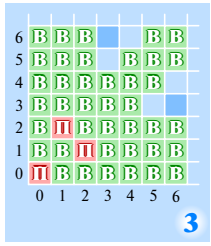
Игра цзяньшицзы

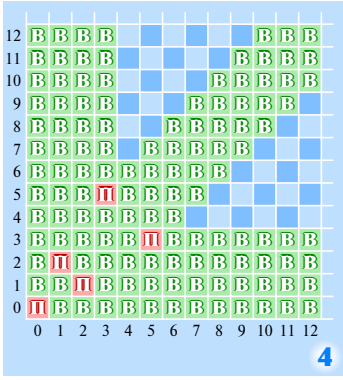
Правила старинной китайской игры цзяньшицзы (выбирание камней) таковы. Имеются две кучи камней. Два игрока по очереди берут камни. За один ход можно взять один или несколько камней из одной (какой угодно) кучи, а можно поровну из обеих куч. Побеждает тот, кто взял последний камень. Как играть наилучшим образом? Ответ сформулируем тремя весьма непохожими способами: при помощи 1) бесконечного фибоначчиева слова, 2) формулы лорда Рэлея и 3) фибоначчиевой системы счисления.

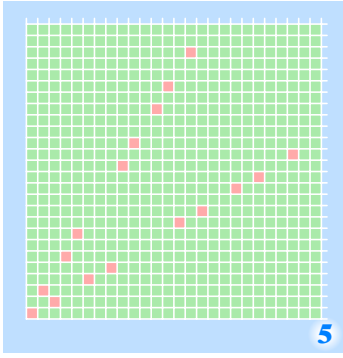












Разрешим ферзю ходы только вниз, налево и по диагонали налево-вниз (рис. 1). Очевидно, рано или поздно он окажется в левом нижнем углу доски. Тот из двух играющих одним ферзем игроков, который поставил ферзя на это поле, — победитель, а его соперник — проигравший.

Эта игра «ферзя в угол» ничем по сути не отличается (изоморфна, сказал бы математик) от цзяньшицзы. Говорить о ферзе чуть удобнее потому, что на полях доски можно расставлять буквы П и В. Позицию называем выигрышной, если у игрока, который из этой позиции делает ход, есть выигрышная стратегия — способ так играть, так отвечать на действия противника, что тот проиграет. Остальные позиции — те, где выигрыш невозможен вовсе или возможен в случае ошибки соперника — называем проигрышными.

На рисунке 2 буквой П отмечено поле (0; 0), а буквами В — поля, с которых можно за один ход поставить противника в проигрышную позицию (0; 0). Заметьте: «пробиты» горизонталь $x=0$, вертикаль $y=0$ и диагональ $x=y$. Обозначим $a_0=b_0=0$.

С полей (1; 2) и (2; 1) можно сделать ход только в выигрышные позиции, поэтому сами они проигрышные. А поля, с которых за один ход можно попасть на (1; 2) или (2; 1) — выигрышные (рис. 3). Обозначим $a_1=1$ и $b_1=2$. Теперь пробиты горизонтали $x=0, 1, 2$, вертикали $y=0, 1, 2$ и диагонали $y-x=0, \pm1$. Следующие красные клетки — (3; 5) и (5; 3) (рис. 4). Обозначаем $a_2=3$ и $b_2=5$. Действуя в таком духе, нетрудно заполнить буквами В и П весьма большую по размеру доску (рис. 5).

Выпишем координаты проигрышных полей. Из полей $(a; b)$ и $(b; a)$, симметричных относительно биссектрисы первого координатного угла, указываем только то, где $a\leqslant b$. Полю (0; 0) присвоим номер 0, а остальные нумеруем в порядке «проявления» буквы П при нашей процедуре расстановки букв.

n	0	1	2	3	4	5	6	7	8	9
a_n	0	1	3	4	6	8	9	11	12	14
b_n	0	2	5	7	10	13	15	18	20	23

Какие позиции проигрышные? Например, чему равно a_{1000}?

Проигрышные позиции $(a_n; b_n)$ довольно ровно лежат на некоторой прямой, проходящей через начало координат. Как это доказать или хотя бы точно сформулировать?

Не будем пока отвечать на эти вопросы. Резко изменим точку зрения — сначала даже покажется, что занялись совершенно другой задачей. Затем, обогащенные новыми знаниями, мы легко построим теорию игры цзяньшицзы. Вернее, и строить не придется, настолько все прояснится! ■

Рассмотрим последовательность слов, первое из которых состоит из одной буквы А, второе — АБ, третье — АБА, четвертое — АБААБ, пятое — АБААБАБА, и так далее: очередное слово получаем из предыдущего, заменяя каждую букву А на АБ, а Б — на А.

Теорема 1. ***Каждое слово этой последовательности, начиная с третьего, получается приписыванием предпредыдущего слова к предыдущему.*** **(*Например*, АБААБАБА — *это* АБААБ *плюс* АБА.)**

Доказательство. Внимательно посмотрев на процесс построения нового очередного слова, вы поймете, что утверждение теоремы 1 очевидно. Но строгости ради докажем его по индукции. Обозначим $w_1 = А$, $w_2 = АБ$, $w_3 = АБА$, и вообще, $w_{n+1} = h(w_n)$, где h обозначает одновременную замену всех букв А на АБ, а Б — на А. Мы должны доказать соотношение

$$w_{n+2} = w_{n+1} w_n.$$

При $n = 1$ оно верно: $w_3 = АБА = w_2 w_1$. Пусть оно верно при некотором n. Тогда

$$w_{n+3} = h(w_{n+2}) = h(w_{n+1} w_n) = h(w_{n+1})\, h(w_n) = w_{n+2} w_{n+1}.$$

Следствие. ***Существует бесконечное слово Фибоначчи***

АБААБАБААБААБАБААБАБА. . . ,

начальными отрезками которого являются все слова w_n. Слово w_n состоит из Φ_{n-1} букв.

Обозначим $a_1 = 1$, $b_1 = 2$, $a_2 = 3$, $a_3 = 4$, $b_2 = 5$, $a_4 = 6$, $b_3 = 7$, $a_5 = 8$, $a_6 = 9$, $b_4 = 10$, и вообще, пусть a_n и b_n — номера мест, на которых стоят n-е буквы А и Б в слове Фибоначчи.

Теорема 2. $\boldsymbol{b_n = n + a_n}$.

Доказательство. n-я буква Б получается из n-й буквы А. Операция h преобразует каждую букву А в две буквы (А и Б), а Б — в одну букву (А). Каждая из первых n букв А даст «лишнюю» букву, то есть $b_n = a_n + n$. ■

Именно такие пары $(a_n; b_n)$, как вы помните, задавали проигрышные поля игры цзяньшицзы. Это настолько замечательно, что стоит дать вам время это обдумать (даже если ничего не придумаете, смело читайте статью дальше!), рассказав о другом способе получения этих пар.

Рассмотрим разбиение натурального ряда на две возрастающие непересекающиеся последовательности $a_1 < a_2 < a_3 < \ldots$ и $b_1 < b_2 < b_3 < \ldots$, которые при любом натуральном n удовлетворяют условию $b_n = a_n + n$. Двигаясь по натуральному ряду, можно последовательно вычислять члены обеих последовательностей.

	b_1			b_2		b_3			b_4			b_5		b_6			b_7		b_8			b_9			b_{10}		b_{11}	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	. . .
a_1		a_2	a_3		a_4		a_5	a_6		a_7	a_8		a_9		a_{10}	a_{11}		a_{12}		a_{13}	a_{14}		a_{15}	a_{16}		a_{17}			

А именно, поскольку $a_n < b_n$, наименьшее натуральное число, то есть 1, должно равняться a_1, а поэтому $b_1 = 1 + 1 = 2$. Теперь наименьшее свободное число — 3. Поэтому $a_2 = 3$ и $b_2 = 3 + 2 = 5$. Сейчас наименьшее неиспользованное число — это $a_3 = 4$, откуда $b_3 = 4 + 3 = 7$. Так можно действовать бесконечно, каждый раз выбирая наименьшее неиспользованное натуральное число и именно его полагая равным a_n, а затем вычисляя $b_n = a_n + n$.

Не получится ли так, что очередное вычисленное значение a_n или b_n окажется уже занято каким-то ранее определенным a_m или b_m, где $m < n$?

Рассмотрим последовательность слов

А,
АБ,
АБАА,
АБААААБАБ,
АБАААБАБАБАААБАА,
. . . . . .

(Очередное слово получается из предыдущего заменой А на АБ, а Б — на АА.) Каждое слово этой последовательности является началом следующего ее слова, а номер места, на котором в соответствующем бесконечном слове

АБААААБАБАБАААБАААБАААБАБАБАБААААБАБ. . .

стоит n-я буква Б, в два раза больше номера места, на котором стоит n-я буква А. ■

Рассмотрим последовательность $a_1 = 1$, $a_2 = 2$, $a_3 = 2$, $a_4 = 3$, $a_5 = 3$, $a_6 = 3$, $a_7 = 4$, $a_8 = 4$, $a_9 = 4$, $a_{10} = 4$, . . . (последовательно выписаны единица, две двойки, три тройки, четыре четверки, пять пятерок и так далее).

Пусть $a_n = k$. До первого появления числа k выписано $1 + 2 + \ldots + (k-1) = k(k-1)/2$ чисел. Последнее раз число k стоит на месте с номером $k(k+1)/2$. Поэтому

$$\frac{k(k-1)}{2} < n \leqslant \frac{k(k+1)}{2},$$

то есть $k^2 - k < 2n \leqslant k^2 + k$. Поскольку числа n и k целые, это неравенство можно записать в виде

$$k^2 - k + \frac{1}{4} < 2n < k^2 + k + \frac{1}{4},$$

то есть $\left(k - \frac{1}{2}\right)^2 < 2n < \left(k + \frac{1}{2}\right)^2$.

Таким образом,

$$k - \frac{1}{2} < \sqrt{2n} < k + \frac{1}{2},$$

то есть $a_n = k = \left[\sqrt{2n} + \frac{1}{2}\right]$. ■

Двое по очереди обрывают лепестки у ромашки: 1 или 2 рядом растущих лепестка за один ход. Выигрывает тот, кто сделает последний ход. Кто выиграет при правильной игре?

Указание. Второй может разбить ромашку на две одинаковые части и поддерживать это состояние. ■

Натуральные числа m и n взаимно просты. Отрезок $[0;1]$ разбит на $m+n$ равных отрезков. Докажем, что в каждом из этих отрезков, кроме двух крайних, лежит ровно одно из $m+n-2$ чисел $\frac{1}{m}, \frac{2}{m}, \ldots, \frac{m-1}{m}$, $\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}$. (Это — вариант теоремы 4 для рациональных чисел $\alpha=\frac{m+n}{m}$ и $\beta=\frac{m+n}{n}$.)

I способ. Если точки $\frac{a}{m}$ и $\frac{b}{n}$ попали в один отрезок, то расположенная между ними их медианта $\frac{a+b}{m+n}$ лежит внутри этого отрезка.

II способ. Нарисуем на клетчатой бумаге прямоугольник размерами $m\times n$ клеток и проведем его диагональ (на рисунке $m=5$ и $n=8$) — она и будет играть роль отрезка $[0;1]$. Взаимная простота чисел m и n означает, что диагональ не проходит через узлы сетки, кроме ее концов. Вертикали делят диагональ на n равных частей, горизонтали — на m. Проведенные под углом 45° черные прямые делят диагональ на $m+n$ равных отрезков. На диагонали между любыми соседними синей и красными точками обязательно лежит черная точка, поскольку, пересекая какую-то клетку, диагональ пересекает и ее черную диагональ. ■

В 1877 г. в «Теории звука» лорд Рэлей (до получения титула — Дж. У. Стретт) писал: «Если x есть некоторое положительное иррациональное число, меньшее единицы, то можно взять два ряда величин $\frac{n}{x}$ и $\frac{n}{1-x}$, где $n=1, 2, \ldots$; каждое число, принадлежащее к тому или иному ряду и только оно одно, будет заключено между двумя последовательными натуральными числами». Другими словами, последовательности $a_n=\left[\frac{n}{x}\right]$ и $b_n=\left[\frac{n}{1-x}\right]$ заполняют без пропусков и перекрытий весь натуральный ряд, если $0<x<1$ и $x\notin\mathbb{Q}$.

Формулы теоремы 3 получаются из формул Рэлея при $x=\frac{2}{1+\sqrt{5}}$, поскольку при этом $1-x=\frac{2}{3+\sqrt{5}}$. ■

Нет, не получится: a_n, как сказано выше, есть наименьшее натуральное число, отличное от $a_1<a_2<\ldots<a_{n-1}$ и от $b_1<b_2<\ldots<b_{n-1}$, и потому a_n не может совпадать ни с одним из них; число $b_n=a_n+n>a_{n-1}+n-1=b_{n-1}$ тоже не может ни с чем совпасть.

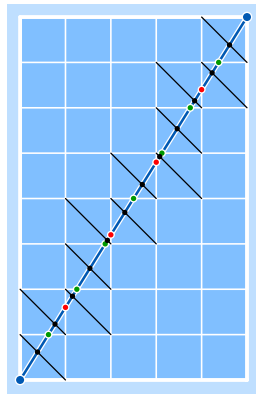
Теорема 3. $a_n=\left[\frac{(1+\sqrt{5})n}{2}\right]$ *и* $b_n=a_n+n=\left[\frac{(1+\sqrt{5})n}{2}\right]+n=\left[\frac{(3+\sqrt{5})n}{2}\right]$.

Сначала докажем более общую теорему.

Теорема 4. *Если* α *и* β *— положительные иррациональные числа, связанные соотношением* $\frac{1}{\alpha}+\frac{1}{\beta}=1$, *то среди чисел вида* $[\alpha n]$ *и* $[\beta n]$, *где* $n\in\mathbb{N}$, *каждое натуральное число встречается ровно один раз.*

Доказательство. I способ. Поскольку $\alpha>1$, в последовательности $[\alpha]$, $[2\alpha]$, $[3\alpha]$, ... никакое число не повторяется. Аналогично, вследствие неравенства $\beta>1$, строго возрастает и последовательность $[\beta]$, $[2\beta]$, $[3\beta]$, ...

Дальше доказательство проведем методом «от противного». Предположим сначала, что некоторое натуральное число k вошло в обе последовательности, то есть $k=[\alpha m]=[\beta n]$, где m, n — натуральные числа. Тогда должны быть выполнены неравенства $k<\alpha m<k+1$ и $k<\beta n<k+1$, то есть $\frac{m}{k+1}<\frac{1}{\alpha}<\frac{m}{k}$ и $\frac{n}{k+1}<\frac{1}{\beta}<\frac{n}{k}$. Сложим эти неравенства, не забыв использовать условие $\frac{1}{\alpha}+\frac{1}{\beta}=1$. Получим $\frac{m+n}{k+1}<1<\frac{m+n}{k}$, откуда



$$k<m+n<k+1.$$

Такого для натуральных чисел не бывает. Значит, число k не могло войти в обе рассматриваемые последовательности.

Теперь предположим, что натуральное число k не вошло ни в одну из последовательностей, то есть отрезок $[k;k+1]$ не содержит ни одного из чисел вида αm или βn. Тогда для некоторых натуральных чисел m и n должны быть выполнены неравенства

$$\alpha m<k<k+1<\alpha(m+1), \qquad \beta n<k<k+1<\beta(n+1),$$

которые мы преобразуем к виду

$$\frac{m}{k}<\frac{1}{\alpha}<\frac{m+1}{k+1}, \qquad \frac{n}{k}<\frac{1}{\beta}<\frac{n+1}{k+1}.$$

Складывая, получаем $\frac{m+n}{k}<1<\frac{m+n+2}{k+1}$, откуда $m+n<k$ и $k+1<m+n+2$, что невозможно для натуральных чисел. Получили желанное противоречие. Теорема доказана.

II способ. Левее любого натурального числа N лежат $[N/\alpha]$ членов первой последовательности и $[N/\beta]$ членов второй. Поскольку α иррационально, числа N/α и N/β имеют ненулевые дробные части. Далее, сумма $\frac{N}{\alpha}+\frac{N}{\beta}=N$ является целым числом, так что дробные части слагаемых дополняют друг друга, то есть в сумме дают в точности 1. Значит, сумма целых частей $[N/\alpha]+[N/\beta]$ равна $N-1$, то есть левее числа N лежит в точности $N-1$ членов этих по-

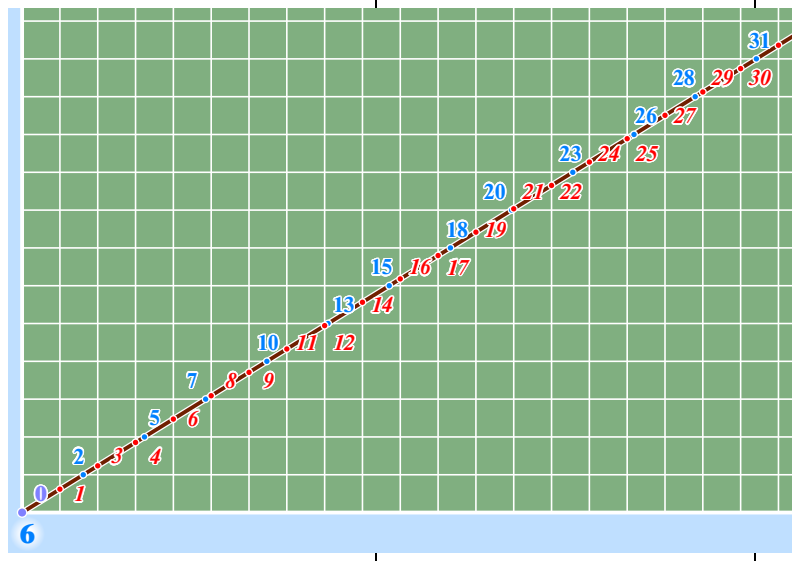
следовательностей. Как легко понять, просматривая натуральный ряд слева направо (любитель строгости сказал бы: применяя индукцию), это как раз означает, что рассматриваемые последовательности однократно покрывают натуральный ряд.

III способ — геометрический. Пусть, как и ранее, α и β — положительные иррациональные числа, причем $\frac{1}{\alpha}+\frac{1}{\beta}=1$.

Тогда $\beta+\alpha=\alpha\beta$, откуда

$$(\alpha-1)(\beta-1)=1.$$

Нарисуем на клетчатой бумаге как на координатной плоскости прямую (рис. 6), заданную уравнением $y=(\alpha-1)x$, которое можно записать также в виде $x=(\beta-1)y$. Занумеруем подряд все клетки, которые пересекает прямая, начиная с нулевой клетки, которой принадлежит начало координат (на рисунке 6 взято $\alpha=(1+\sqrt{5})/2$). Поскольку число α иррационально, прямая не проходит через узлы сетки (кроме, разумеется, начала координат). Значит, она входит в очередную клетку либо слева, пересекая вертикальную линию сетки, либо снизу, пересекая горизонтальную линию.



Если прямая вошла в клетку слева и пересекла при этом вертикаль $x=n$, то номер клетки, в которую при этом она вошла, равен $n+[(\alpha-1)n]=[\alpha n]$. (В n из этих клеток прямая вошла слева, а в остальные — снизу.) Если же прямая пересекла горизонталь $y=m$, то номер соответствующей клетки равен $[(\beta-1)m]+m=[\beta m]$. ■

Поскольку $\frac{1}{\sqrt{2}}+\frac{1}{2+\sqrt{2}}$, последовательности, заданные формулами $a_n=[n\sqrt{2}]$ и $b_n=a_n+2n$, заполняют весь натуральный ряд без пропусков и перекрытий. ■

Если иррациональные числа $\alpha>1$ и $\beta>1$ таковы, что ни для каких натуральных чисел m и n не выполнено равенство $[\alpha m]=[\beta m]$, то, как доказали в 2005 г. А. А. Заславский и А. В. Спивак, существуют такие натуральные числа r и s, что $\frac{r}{\alpha}+\frac{s}{\beta}=1$. (В силу теоремы Рэлея, верно и обратное.) ■

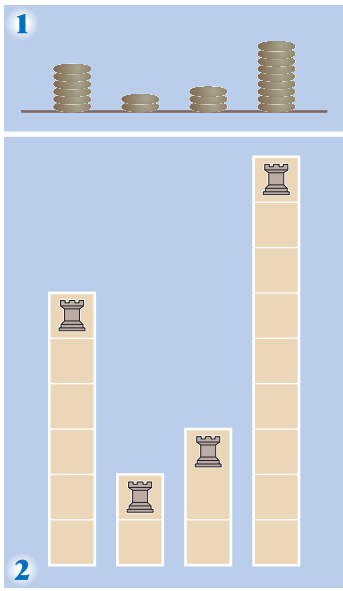
Фибоначчиева система записи натуральных чисел похожа на десятичную систему, только вместо степеней числа 10 используют числа Фибоначчи 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, ... Разложим, например, число 210. При вычитании числа 233 из числа 210 получаем отрицательный результат. Поэтому вычтем 144. Разность равна 66. Поскольку $66<89$, вычтем из 66 число 55. Разность равна 11. Очевидно, $11=8+3$. Таким образом можно каждое натуральное число представить в виде суммы (возможно, состоящей из одного слагаемого) чисел Фибоначчи, причем соседние числа Фибоначчи не могут одновременно войти в такую сумму: если из числа, не превращая его в отрицательное число, можно вычесть φ_{n+1} и φ_n, то можно вычесть и их сумму $\varphi_{n+1}+\varphi_n=\varphi_{n+2}$. Например, 210 в фибоначчиевой системе записывается как $\overline{10100010100}$.

Запишем числа a_n и b_n в фибоначчиевой системе счисления, не забыв и само число n написать не только в десятичной, но и в фибоначчиевой системе.

Например, $a_8=1+3+8=12$ и $b_8=2+5+13=20$. Правило образования четвертого столбца таблицы из третьего очевидно: надо приписать цифру 0. Числа b_n оканчиваются нечетным числом нулей, а a_n — четным. Чуть хитрее правило получения третьего столбца из второго: если фибоначчиева запись числа n оканчивается на цифру 0, то надо приписать 0; если же на 1, то надо заменить эту последнюю цифру 1 на 01. (И в случае $n=1$ не забыть отбросить этот ноль, добавил бы любитель абсолютной строгости.) ■

n дес.	n фиб.	a_n	b_n
1	1	1	10
2	10	100	1000
3	100	101	1010
4	101	1001	10010
5	1000	10000	100000
6	1001	10001	100010
7	1010	10100	101000
8	10000	10101	101010
9	10001	100001	1000010
10	10010	100100	1001000
11	10100	100101	1001010
12	10101	101001	1010010
13	100000	1000000	10000000
14	100001	1000001	10000010
15	100010	1000100	10001000

Игра ним



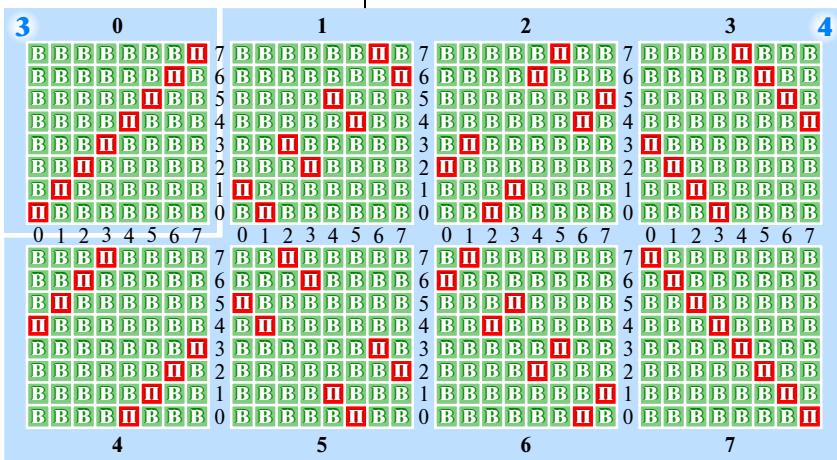
Алеша Попович и Добрыня Никитич воюют с девятиглавым змеем. По очереди богатыри ходят к его пещере и отрубают 1, 2 или 3 головы. Как начавшему бой Алеше обрести славу победителя змея (то есть отрубить последнюю голову)? А если змей двенадцатиглавый?
Ответ. Если у змея 1, 2 или 3 головы, то выигрывает Алеша Попович. Если голов 4, то выигрывает Добрыня. Далее, если у змея 5, 6 или 7 голов, то выигрывает Алеша, если голов 8, то выигрышная стратегия есть у Добрыни. Таким образом, проигрышные для Алеши количество голов — это числа, кратные 4. ■

На столе лежат несколько куч камней. Два игрока по очереди берут камни, причем за один ход игрок может взять один или несколько камней из одной (обязательно одной!) кучи. Победитель — тот, кто взял последний камень.

Играть в ним удобно у классной доски — без камней, но с мелом и тряпкой в руках. Или можно нарисовать несколько вертикальных полос и на каждую из них поставить по ладье, разрешив им двигаться только вниз. Каждым ходом игрок должен передвинуть одну из ладей. Кто не сможет сделать ход — по той причине, что все ладьи уже в крайних нижних положениях, — тот и проиграл.
Расстановка ладей, соответствующая кучам камней рисунка 1, показана на рисунке 2. Как играть? Из какой кучи и сколько взять камней? Ситуация слишком сложная, чтобы можно было догадаться «просто так». ■

Исследуем задачу, начав с простейших случаев. Если куча всего одна, то выигрышная стратегия — взять все камни. Если куч две и в них камней не поровну, помогает идея симметрии: можно уравнивать количества камней в кучах. Например, пусть в одной куче 19 камней, а в другой — 11. Первый игрок возьмет 8 камней из первой кучи. Второй игрок будет вынужден нарушить равенство. Он возьмет, например, 2 камня, — и первый сразу восстановит симметрию, взяв 2 камня из другой кучи. Так они и будут играть: второй всякий раз вынужденно нарушает равенство, а первый его восстанавливает, повторяя ход соперника, но с другой кучей.
Таким образом в случае двух куч тот, кто делает первый ход, выиграет — кроме случаев, когда он не сможет уравнять количества камней в кучах по той причине, что они уже равны; в этих случаях уравнивать может второй, он и выиграет. ■

Сложнее играть, когда куч не 2, а 3. Даже отмечать позиции надо не на плоскости, а в пространстве! Не беда. Будем считать, что левый верхний квадрат рисунка 3 изображает ситуации, когда в первых двух кучах не более 7 камней, а в третьей — ни одного. Нарисуем еще 7 таких же квадратов — по одному квадрату для ситуаций, когда в третьей куче 1, 2, . . . , 7 камней.



Начнем закрашивать клеточки. Принцип такой: если из некоторой позиции все ходы ведут только в зеленые (выигрышные) клетки, то противник сможет победить при любом нашем ходе; такие клетки красим красным цветом. Если же из некоторой позиции можно за один ход попасть на красную клетку, то противник при безошибочной нашей игре проиграет; такие клетки красим зеленым цветом. Получили очень красивый рисунок 4. Прояснить суть дела помогают числа Гранди и ним-суммы. ■

Рассмотрим более общую, чем игра ним, ситуацию. Двое на одной «игровой площадке» ходят по очереди. Кто не может сделать ход — проиграл. Такой игре сопоставим ориентированный граф, вершины которого — позиции игры; если из одной позиции можно сходить в другую, рисуем соответствующую стрелочку — направленное ребро графа. Игру называем финитной, если при любой начальной позиции она закончится через конечное число ходов (то есть нельзя бесконечно долго двигаться по стрелкам). ■

Число Гранди позиции финитной игры — это наименьшее целое неотрицательное число, не являющееся числом Гранди никакой из позиций, в которые можно пойти непосредственно из данной позиции. Это определение рекурсивное: используя его, для любой финитной игры можно одно за одним найти все числа Гранди. Например, рассмотрим игру ним с одной кучкой камней. Очевидно, число Гранди кучи из n камней равно n.

Теорема 1. ***Проигрышные позиции — это позиции, числа Гранди которых равны* 0; *все другие позиции — выигрышные.***

Доказательство. Из определения следует, что
- число Гранди любой тупиковой вершины равно 0;
- из любой позиции с ненулевым числом Гранди можно пойти в позицию, число Гранди которой равно 0;
- из позиции, число Гранди которой равно 0, нельзя пойти в позицию с числом Гранди 0.

Таким образом, множество вершин графа, числа Гранди которых равны 0, удовлетворяют требованиям, предъявляемым к определению множества проигрышных позиций и однозначно характеризующим это множество. ■

Сумма $A \oplus B$ игр A и B — это игра, позиции которой — пары вида $(a; b)$, где a — позиция игры A, а b — позиция игры B. Сделать ход из позиции $(a_1; b_1)$ в позицию $(a_2; b_2)$ можно, если либо $a_1 = a_2$ и в игре B можно пойти из b_1 в b_2, либо же $b_1 = b_2$ и в игре A можно пойти из a_1 в a_2. Другими словами, каждым своим ходом игрок делает ход не в обеих разворачивающихся играх, а только — по своему выбору — в одной из них. Очевидно, сумма любых двух финитных игр финитна.

Определение. Расположив двоичные записи целых неотрицательных чисел x и y одну под другой так, как это принято для сложения (разряд единиц числа y под разрядом единиц числа x, разряд двоек — под разрядом двоек и так далее), выполним в каждом разряде сложение по модулю 2, не выполняя никаких переносов. Получили двоичную запись ним-суммы — числа $x \oplus y$.

Теорема 2. ***Число Гранди позиции* $(a; b)$ *игры* $A \oplus B$ *равно ним-сумме их чисел Гранди:* $G(a \oplus b) = G(a) \oplus G(b)$.**

Доказательство. Поскольку числа Гранди определены единственным образом, достаточно доказать, что заявленные значения удовлетворяют определению чисел Гранди. Очевидно, при изменении любого из слагаемых x и y ним-сумма $x \oplus y$ меняется. Осталось доказать, что если x, y, z — целые неотрицательные числа и $x \oplus y = z$, то для любого целого неотрицательного числа $w < z$ можно так уменьшить одно из слагаемых ним-суммы $x \oplus y$, что она станет равна w.

Рассмотрим самый старший двоичный разряд, в котором двоичные записи чисел x и y различаются. Так как $w < z$, в этом разряде в числе w цифра 0, а в числе z — цифра 1. Следовательно, в одном из слагаемых x и y в этом разряде цифра 1, это слагаемое и будем уменьшать. Поменяем в рассматриваемом разряде 1 на 0, а цифры младших (правее данного) разрядов сделаем такими, чтобы ним-сумма стала равна d (каждую цифру подбираем отдельно). ■

Два мудреца играют в такую игру. Выписаны числа 0, 1, 2, 3, . . . , 1024. Первый вычеркивает по своему выбору некоторые 512 чисел, второй вычеркивает 256 из оставшихся чисел, затем снова первый вычеркивает еще 128, потом второй — еще 64 числа и так далее. Своим последним пятым ходом второй вычеркивает одно число. Остаются два числа, и второй платит первому разницу между этими числами. Сколько уплатит второй первому, если оба будут играть наилучшим образом?

Ответ: 32. **Решение.** Первый игрок может каждым своим ходом вычеркивать каждое второе из оставшихся к этому моменту чисел. Тогда после первого его хода разность любых двух невычеркнутых чисел не меньше 2, после второго его хода любая такая разность не меньше 4, после третьего — не меньше 8, после четвертого — не меньше 16, после пятого — не меньше 32.

А второй игрок может первым своим ходом вычеркнуть все числа, меньшие 512 или все числа, большие 512, в зависимости от того, каких осталось меньше (ведь не может же и тех, и других остаться больше 256 штук). Таким образом, после первого хода второго игрока разность между крайними из невычеркнутых чисел не больше 512. Аналогично вторым своим ходом второй игрок может добиться того, что все оставшиеся невычеркнутыми числа будут находиться в одном из отрезков [0; 256], [256; 512], [512; 768], [768; 1024], то есть уменьшить разность между крайними числами по крайней мере до 256. Точно так же третьим своим ходом может уменьшить эту разность до 128, четвертым — до 64 и пятым — до 32. ■

Игра начинается с числа 60. За ход разрешается уменьшить имеющееся число на любой из его делителей. Проигрывает тот, кто получит 0. Кто выигрывает при правильной игре?

Ответ. Проигрышные позиции — нечетные числа. ■

Из металлических прутьев сварен октаэдр. Двое по-очереди перепиливают прутья. Тот, после чьего хода конструкция распадется на части, проигрывает. Кто выигрывает при правильной игре? ■

Латинские квадраты и устойчивые браки

В 1978 г. Дж. Диниц выдвинул следующую гипотезу. Рассмотрим множества A_{kj}, где $1 \leqslant k, j \leqslant n$, каждое из которых состоит из n элементов. Тогда можно выбрать из каждого множества A_{kj} по одному элементу a_{kj} таким образом, что получится нечто вроде латинского квадрата, то есть в любой строке все выбранные элементы будут разные и в любом столбце все выбранные элементы будут разные. В 1995 г. ее доказал Ф. Гальвин.

Если все множества A_{kj} совпадают, то утверждение очевидно: на рисунке 1 изображен случай $n=7$; такой же латинский квадрат существует для любого натурального n. Если же множества A_{kj} не обязательно совпадают, то, неформально говоря, свободы тем больше, чем сильнее отличаются они друг от друга. Будем рассматривать элементы множеств A_{kj} как краски и говорить о раскрасках клеток квадрата, при которой в каждой строке, как и в каждом столбце, все цвета разные. Чем сильнее различаются «палитры» A_{kj}, тем, казалось бы, проще красить! Но как превратить эту идею в математическое рассуждение? ■

Построим ориентированный граф: каждую клетку квадрата — или, если угодно, каждую пару $(k;j)$ — превратим в вершину графа. Соединим $(k;j)$ с другой вершиной $(r;s)$ ребром, если $k=r$ или $j=s$, то есть если они находятся в одной строке или в одном столбце. Обозначим через $f(k;j)$ число, расположенное в клетке $(k;j)$ латинского квадрата размером $n \times n$; например, если строить латинский квадрат, как показано на рисунке 1, то $f(k;j)=k-j+1 \bmod n$. Поставим стрелочку от $(k;j)$ к $(k;m)$, если $f(k;j) > f(m;j)$; и от $(k;j)$ к $(m;j)$, если $f(k;j)<f(m;j)$. При этом из каждой из n^2 вершин построенного ориентированного графа выходят $n-1$ ребер и в каждую вершину входят столько же, $n-1$ ребер (на рисунке 2 изображен такой граф для $n=3$).

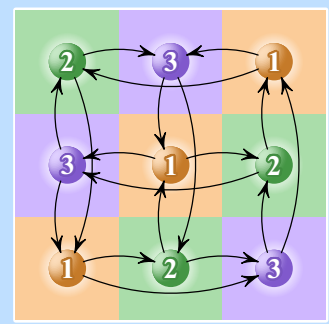
Докажем по индукции, что если в каждой из вершин этого ориентированного графа или некоторого меньшего графа, полученного из него вычеркиванием некоторых вершин вместе со всеми выходящими и входящими в них ребрами, имеется палитра, в которой больше красок, чем выходит из этой вершины ребер, то можно раскрасить вершины так, чтобы каждая вершина была покрашена одной из красок ее палитры и чтобы для любого ребра его концы были покрашены разными красками.

Заметьте: в начале в каждой вершине — палитра из n красок. И из каждой вершины выходит $n-1$ ребер; выходящих ребер меньше, чем красок в палитре! ■

Приступим к доказательству. Будем рассматривать цвета по очереди и каждый раз красить несколько вершин и вычеркивать их вместе со всеми выходящими и входящими ребрами. Рассмотрим цвет a, присутствующий хотя бы в одной из палитр. Пусть S — множество всех вершин графа, в палитрах которых присутствует цвет a. Если никакие элементы S не соединены ребрами между собой, то можно покрасить их все в цвет a и вычеркнуть все элементы множества S (со всеми выходящими и входящими ребрами).

Вот еще одна формулировка теоремы Диница—Гальвина. Пусть на конкурсе бальных танцев каждый из n мальчиков должен станцевать перед жюри один танец с каждой из n девочек. Если на репетициях каждый мальчик разучил с каждой из девочек n танцев, то можно устроить так, что будет всего n^2 выступлений и каждый из $2n$ человек выступит с разными n танцами. ■

1



2

В одной деревне жили-были юноши и девушки. Задумали девицы все сразу выйти замуж. Каждая составила список юношей, которые ей любы. При каких условиях девушки смогут сговориться так, что каждая предложит руку и сердце одному из милых ей юношей и никто из юношей не получит более одного предложения?

Одно требование очевидно: каждой девушке должен нравиться хотя бы кто-нибудь. Далее, если некоторые две девушки хотят выйти замуж только за одного и того же юношу, то беда. Поэтому для любых двух девушек объединение их списков должно состоять не менее чем из двух юношей. Вообще, для любых k девушек множество юношей, которые нравятся хотя бы одной из них, должно состоять не менее чем из k юношей.

Докажем, что эти условия не только необходимы, но и достаточны индукцией по количеству n девушек в деревне. При $n=1$ утверждение очевидно. Пусть оно верно для любого количества девушек, меньшего некоторого натурального числа n. Докажем его для n девушек. Возможны два случая.
1) Для некоторого $k<n$ есть k девушек, в объединении списков которых ровно k юношей. По предположению индукции, для этих юношей и девушек полная свадьба возможна. Отпразднуем ее. Если бы для некоторых m незамужних девушек в объединении списков их женихов оказалось меньше m человек, то, отменив все прошедшие свадьбы, мы получили бы $m+k$ девушек, объединение списков женихов которых состоит менее чем из $m+k$ человек. Следовательно, по предположению индукции, все $n-k$ незамужних девушек могут одновременно выйти замуж!
2) Для любого $k<n$ у любых k девушек объединение списков женихов состоит более чем из k человек. Этот случай очевиден: любая девушка выходит замуж за кого хочет, а к остальным применяем предположение индукции. ■

Следствия из теоремы о свадьбах. 1) Если каждый школьник решил n задач и каждую задачу решили n школьников, то можно организовать разбор задач так, что каждый расскажет одну задачу и каждая задача будет рассказана один раз.
2) В кубе размером $n\times n\times n$ надо поставить n^2 ладей так, чтобы они не били друг друга. Если в первых k слоях, где $k<n$, уже расставили kn ладей, соблюдая это условие, то и в $(k+1)$-й слой можно добавить n ладей, не нарушив его. ■

Будем использовать слово «линия» для обозначения строки или столбца прямоугольной таблицы. Теорема о свадьбах по существу равносильна следующей теореме Дж. Кёнига.
Если прямоугольная таблица составлена из нулей и единиц, то минимальное число линий, которые содержат все единицы, равно максимальному числу единиц, которые можно выбрать так, чтобы никакие две из них не лежали на одной линии. ■

Труднее дело обстоит в случае, когда некоторые элементы множества S соединены ребрами между собой. Тогда нельзя красить их все в цвет a. Но можно — при помощи доказываемой ниже теоремы об устойчивых браках — найти такое подмножество $P\subset S$, что никакие две вершины множества P не соединены ребром и для каждой вершины из $S\setminus P$ хотя бы одно выходящее из нее ребро ведет в одну из вершин множества P.
Понимаете, для чего это нам нужно? Мы можем вычеркнуть из графа все вершины множества P, а из палитр всех вершин множества $S\setminus P$ можем при этом вычеркнуть цвет a. Главное условие — то, что в палитре любой вершины больше красок, чем из этой вершины выходит ребер, — при этом останется верным! ■

Поговорим о свадьбах. Пусть в деревне живут n невест (раньше это были строки квадратной таблицы или, если угодно, числа $k=1, 2, \ldots, n$) и n женихов (раньше это были столбцы таблицы или числа $j=1, 2, \ldots, n$). Пусть каждая девушка упорядочила всех женихов некоторым — совершенно любым! — способом. И пусть каждый юноша упорядочил всех девушек своим — опять-таки любым! — способом. (Эти системы предпочтений в доказательстве соответствуют стрелочкам между вершинами графа.) Пусть мы как-то образовали n супружеских пар.
Если A и C — юноши, B и D — девушки и если мы образовали пары $(A; B)$ и $(C; D)$, то соблазн для B и C возникает, если выполнены следующие два условия: во-первых, B любит своего мужа A меньше, чем C; во-вторых, C любит свою жену меньше, чем B. Почему это соблазн? Понятно: B и C оба выигрывают, разрушая свои пары и создавая пару $(B; C)$.

Теорема об устойчивых бракосочетаниях. ***При любой системе списков предпочтений можно образовать n супружеских пар так, чтобы не было ни одного соблазна.***

Ее доказательство довольно короткое и красивое. Но прежде чем его рассказывать, заметим, что для теоремы Диница нужна не сама теорема об устойчивых браках, а некоторое ее (впрочем, точно так же доказываемое) обобщение. А именно, во время доказательства гипотезы Диница мы вычеркивали вершины графа, так что в начале рассуждения граф состоял из n^2 вершин, а в процессе доказательства количество вершин могло уменьшиться. Поэтому в обобщении теоремы об устойчивых бракосочетаниях нам нужно учесть, что браки между некоторыми юношами и девушками запрещены (а разрешены — только те, что принадлежат S). Другими словами, у каждого юноши (девушки) список желанных девушек (юношей) может состоять не из всех n девушек (юношей). Немного подумав, вы поймете, что этот вариант теоремы об устойчивых браках — в точности нужное нам утверждение о существовании множества P. ■

Теперь — обещанное доказательство теоремы об устойчивых браках. Рассмотрим холостяка. Женим его на лучшей из его девушек. Дальше будем каждый раз рассматривать холостого к этому моменту юношу и женить его на лучшей из следующего списка девушек: всех незамужних его подруг и всех тех замужних, которые любят своего мужа меньше, чем рассматриваемого юношу. (В последнем случае придется один брак расторгнуть.) Доказательсво того, что этот алгоритм рано или поздно остановится, очевидно: раз выйдя замуж, девушка может менять мужей, но уже никогда не станет незамужней. И даже если девушка меняет мужей, то каждый раз с ее точки зрения муж становится все лучше и лучше. Но бесконечно улучшать качество мужа нельзя! ■

Одноцветные прогрессии

Б. Л. Ван дер Варден в 1926 г. доказал, что если множество натуральных чисел разбито на k непересекающихся подмножеств, то хотя бы в одном из них есть сколь угодно длинная арифметическая прогрессия. Другими словами, для любого натурального l существует такое число $n=f(k, l)$, что при любой раскраске n подряд идущих натуральных чисел k цветами (каждое число — одним цветом, использовать все цвета не обязательно) найдутся l одноцветных чисел, образующих арифметическую прогрессию.

Рассмотрим пример: $k=2$ и $l=3$. Докажем, что в качестве $f(2, 3)$ можно взять число 9.

Всего существует $2^9=512$ способов раскрасить 9 чисел синей и красной красками (включая случаи, когда все числа красные или все синие). Можно перебрать все эти случаи и убедиться в справедливости утверждения. Но мы докажем его, рассмотрев только два случая.

1) Пусть числа 4 и 6 одного цвета, для определенности — синие. Во избежание прогрессии 4, 5, 6, красим 5 в красный цвет. Чтобы избежать прогрессий 2, 4, 6 и 4, 6, 8, красим 2 и 8 в красный цвет. Получили красную прогрессию 2, 5, 8.

2) Пусть числа 4 и 6 разного цвета, для определенности 4 — синее, 6 — красное. Из симметрии следует, что достаточно рассмотреть случай, когда число 5 — красное. Чтобы избежать прогрессии 5, 6, 7, красим 7 в синий цвет. Чтобы избежать 1, 4, 7, число 1 красное; далее из-за прогрессии 1, 3, 5 число 3 — синее; из-за 2, 3, 4 число 2 — красное; из-за 2, 5, 8 число 8 — синее; из-за 7, 8, 9 число 9 — красное. Получили красную прогрессию 1, 5, 9.

Конечно, такой метод работает только при очень маленьких числах k и l. Общий случай перебору неподвластен. ■

При $l=2$ и любом k теорема Ван дер Вардена тривиальным образом верна, $f(k, 2)=k+1$. В самом деле, при раскраске $k+1$ чисел в k цветов найдутся два числа одного цвета; эта пара чисел и образует арифметическую прогрессию длины 2.

База индукции доказана. Выполним индукционный переход: считая, что теорема доказана для любого k и некоторого l, докажем ее для $l+1$ (и, разумеется, для любого k). Рассмотрим последовательности $q_0, q_1, q_2, \ldots$ и $n_0, n_1, n_2, \ldots$, определенные своими начальными членами

$$q_0=1 \qquad \text{и} \qquad n_0=f(k, l)$$

и рекуррентными соотношениями

$$q_{s+1}=2n_sq_s \qquad \text{и} \qquad n_{s+1}=f(k^{n_s}, l),$$

где $s=0, 1, 2, \ldots$ Докажем, что в качестве $f(k, l+1)$ можно взять число q_k. Другими словами, если числа отрезка Δ натурального ряда, состоящего из q_k чисел, раскрашены в k цветов, то существует одноцветная арифметическая прогрессия длины $l+1$.

При палитре в k красок количество способов раскрасить два числа равно k^2, три числа — k^3, а количество способов раскрасить m чисел равно k^m.

Бартел Лендерт Ван дер Варден (1903—1996) — голландский математик. Автор монографии «Алгебра» (1931), завершившей создание «общей» алгебры. Книга и по сей день является лучшим изложением основ теории групп, полей и колец. Занимался историей математики и астрономии («Пробуждающаяся наука. Математика древнего Египта, Вавилона и Греции»).

К. Т. В. Вейерштрасс (1815—1897) придумал всюду непрерывную, но нигде не дифференцируемую функцию:

$$f(x)=\sum_{n=0}^{\infty} b^n \cos(a^n\pi x),$$

где a — целое нечетное число, $0<b<1$ и $ab>1+\frac{3}{2}\pi$. В 1916 г. Г. Х. Харди (1877—1947) доказал, что достаточно потребовать $0<b<1$ и $ab>1$.

Ван дер Варден в 1930 г. построил более простой пример такой функции. Обозначим через $\langle x\rangle$ расстояние от числа x до ближайшего целого числа. Другими словами, при $-1/2\leqslant x\leqslant 1/2$ пусть $\langle x\rangle=|x|$, а дальше продолжим эту функцию периодически. Для любого целого неотрицательного числа n обозначим $f_n(x)=\frac{\langle 4^nx\rangle}{4^n}$ и

$$w(x)=\sum_{n=0}^{\infty} f_n(x).$$

Функция $w(x)$ непрерывна как сумма равномерно сходящегося ряда непрерывных функций.

Пусть a — произвольное вещественное число. Для любого натурального n выберем $h_n=1/4^{n+1}$

или $h_n=-1/4^{n+1}$ так, что

$$|f_n(a+h_n)-f_n(a)|=|h_n|.$$

Тогда разность $f_m(a+h_n)-f_m(a)$ равна 0 при $m>n$ и равна $\pm h_n$ при $m\leqslant n$. Следовательно, отношение $(w(a+h_n)-w(a))/h_n$ является целым числом, которое четно при нечетном n и нечетно при четном n. Значит, предел

$$\lim_{n\to+\infty}\frac{w(a+h_n)-w(a)}{h_n}$$

не существует. Функция w не дифференцируема.

Функция w обладает еще одним интересным свойством. Пусть $x=k/4^n$, где k — целое число. Обозначим $h=1/4^{2n+1}$. Тогда

$$w(x)=f_0(x)+f_1(x)+\ldots+f_{n-1}(x)$$

и

$$\begin{aligned}w(x+h)-w(x)&=(f_0(x+h)-f_0(x))+\\&+(f_1(x+h)-f_1(x))+\ldots\\&\ldots+(f_{n-1}(x+h)-f_{n-1}(x))+\\&+f_n(x+h)+f_{n+1}(x+h)+\\&+f_{n+2}(x+h)+\ldots+f_{2n}(x+h)\geqslant\\&\geqslant -nh+(n+1)h>0.\end{aligned}$$

Аналогично,

$$w(x-h)-w(x)\geqslant -hn+(n+1)h>0.$$

Итак, $w(x-h)>w(x)<w(x+h)$. Поскольку точки вида $x=k/4^n$ всюду плотны, то не существует интервала, на котором функция w монотонна. ■

Обозначим в честь Ван дер Вардена через $W(l)$ такое наименьшее натуральное число, что при любом разбиении множества первых $W(l)$ натуральных чисел на два подмножества хотя бы в одном из них найдется арифметическая прогрессия длины l. Известны следующие значения: $W(2)=3$, $W(3)=9$, $W(4)=35$, $W(5)=178$, а $W(6)$ не удается найти даже при помощи современных компьютеров.

Интересно было бы выяснить, насколько быстро растет функция W. Р. Грэхем выдвинул гипотезу, что $W(l)\leqslant w_l$, где

$$w_l=2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}\Big\}\,l$$

— «башня» из l двоек, точнее говоря, $w_1=2$ и $w_{n+1}=2^{w_n}$ для любого натурального n. Но есть основания сомневаться даже в этой оценке: в некоторых похожих комбинаторных задачах (так называемых задачах Рамсея) встречаются функции, растущие еще быстрее. ■

Поскольку $q_k=2n_{k-1}q_{k-1}$, то левая половина отрезка Δ состоит из $n_{k-1}=f(k^{q_{k-1}},l)$ отрезков длины q_{k-1} каждый. По смыслу числа $f(k^{q_{k-1}},l)$, рассматриваемая половина отрезка Δ содержит арифметическую прогрессию из l одинаково раскрашенных отрезков $\Delta_0,\Delta_1,\ldots,\Delta_{l-1}$ длины q_{k-1} каждый. При этом мы для краткости говорим, что отрезки образуют арифметическую прогрессию, если таковую образуют их первые числа; одинаково раскрашенными называем два отрезка, один из которых переходит в другой при параллельном переносе, при котором цвет каждого из чисел первого отрезка совпадает с цветом соответствующего числа второго отрезка.

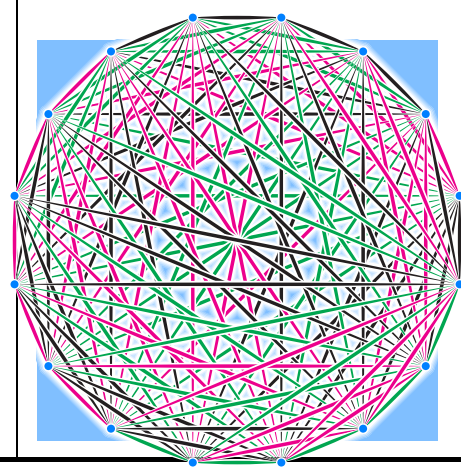
К прогрессии $\Delta_0,\Delta_1,\ldots,\Delta_{l-1}$ присоединим следующий, $(l+1)$-й отрезок Δ_l. Возможно, он выйдет за пределы левой половины отрезка Δ, но за пределы отрезка Δ — не выйдет!

Таким образом, на отрезке Δ мы построили арифметическую прогрессию $\Delta_0,\Delta_1,\ldots,\Delta_{l-1},\Delta_l$ длиной q_{k-1} каждый, первые l отрезков раскрашены одинаково, а про раскраску отрезка Δ_l мы ничего не знаем. Обозначим разность построенной прогрессии через d_1. ■

Отрезок Δ_0 состоит из $q_{k-1}=2n_{k-2}q_{k-2}$ чисел. Поэтому его левую половину можно разбить на $n_{k-2}=f(k^{q_{k-2}},l)$ отрезков длиной q_{k-2} каждый. По смыслу величины $f(k^{q_{k-2}},l)$ существует арифметическая прогрессия $\Delta_{00},\Delta_{01},\ldots,\Delta_{0,l-1}$ из l одинаково раскрашенных отрезков длины q_{k-2} каждый. Обозначим разность этой прогрессии через d_2 и присоединим к ней $(l+1)$-й отрезок Δ_{0l}, про раскраску которого мы ничего не знаем. Отрезок Δ_{0l} может выходить за границу левой половины отрезка Δ_0, но не может — за пределы всего отрезка Δ_0.

Мы провели построение пока только в одном отрезке Δ_0. Перенесем его при помощи параллельных переносов во все другие отрезки $\Delta_1,\ldots,\Delta_l$. Получим семейство отрезков с двумя индексами Δ_{ij}, где $0\leqslant i,j\leqslant l$. Отрезки с индексами, меньшими l, раскрашены одинаково. ■

Одна из задач теории Рамсея — найти наибольшее количество точек, которые можно так соединить отрезками трех цветов, чтобы не образовалось ни одного треугольника с одноцветными сторонами и каждая точка была соединена с каждой. Ответ в этой задаче — 16. На рисунке показано, как их можно соединить. 17 точек так соединить нельзя: непременно возникнет треугольник, все стороны которого одного цвета!



Проведем такое построение k раз. Результатом первого шага были отрезки длины q_{k-1}, второго — длины q_{k-2} и так далее. После k-го шага получим отрезки длины $q_0=1$, то есть числа. Тем не менее, будем обозначать их по-прежнему: $\Delta_{i_1i_2\ldots i_k}$, где $0\leqslant i_1,i_2,\ldots,i_k\leqslant l$. Отрезки с индексами, меньшими l, раскрашены одинаково. ■

Среди $k+1$ чисел $a_0=\Delta_{00\ldots00}$, $a_1=\Delta_{00\ldots0l}$, $\ldots$, $a_{k-1}=\Delta_{0l\ldots ll}$, $a_k=\Delta_{ll\ldots ll}$ найдутся два одного цвета; пусть это будут a_r и a_s, где $r<s$. Рассмотрим числа

$$c_t=\Delta_{\underbrace{00\ldots0}_{k-s}\underbrace{tt\ldots t}_{s-r}\underbrace{ll\ldots l}_{r}},$$

где $0\leqslant t\leqslant l$. Очевидно, числа $c_0,c_1,\ldots,c_{l-1}$ одного цвета. Поскольку $c_0=a_r$ и $c_l=a_s$, числа c_0 и c_l тоже одного цвета. Таким образом, $c_0,c_1,\ldots,c_l$ — одноцветная арифметическая прогрессия с разностью $d_{k-s+1}+d_{k-s+2}+\ldots+d_{k-r}$. ■

История математики

Древняя история математики очень интересна, но установить точные факты и даты весьма трудно из-за скудости источников. Краткости и достоверности ради мы начнем с XII—XIII вв., когда на латинский язык были переведены арабские и греческие сочинения и в Европе начала распространяться десятичная система счисления. Многие из указанных ниже дат условны: большинство открытий не могут быть приписаны одному человеку и определенному моменту времени.

1202 — Леонардо Пизанский (Фибоначчи) (ок. 1180—ок. 1240) написал «Книгу об абаке», которая стала основным учебником арифметики и алгебры в Европе. В частности, он рассказал о десятичной нумерации, числах Фибоначчи, извлечении кубического корня.

Ок. 1260 — А. Насирэддин ат-Туси (1201—1274) выделил тригонометрию из астрономии в отдельную науку, дал концепцию положительного действительного числа.

Ок. 1325 — Т. Брадвардин (ок. 1290—1349) написал «Теоретическую геометрию» и «Трактат о континууме», где ввел понятие иррациональности.

Ок. 1360 — Н. Орем (ок. 1323—1382) пользовался координатным методом, применял дробные и иррациональные показатели степени, ряды, различая ряды сходящиеся и расходящиеся.

XIV—XV вв. — Происходит совершенствование алгебраической символики, введение обозначений для операций возведения в степень, извлечения корня и для степени неизвестной.

Ок. 1425 — Аль-Каши (умер ок. 1436 г.) сформулировал правила извлечения корней любой степени из целых чисел, основанные на формуле бинома Ньютона, выписал таблицу биномиальных коэффициентов, вычислил 16 знаков числа π.

1482 — В Венеции напечатаны «Начала» Евклида.

1544 — М. Штифель (1487—1567) знал правила образования биномиальных коэффициентов, близко подошел к идее логарифмов.

XVI в. — Крупный успех европейской математики: С. Ферро (1465—1526) решил уравнение $x^3+mx=n$, где m и $n>0$. Благодаря Н. Тарталье (ок. 1499—1557) и Дж. Кардано (1501—1576) решение уравнений третьей степени общего вида приобрело известность. Ученик Кардано Л. Феррари (1522—1565) решил уравнение четвертой степени.

1572 — Р. Бомбелли (ок. 1530—ок. 1572) в «Алгебре» впервые рассмотрел числа вида $a+bi$ и сформулировал правила действий над ними. Эти числа он трактовал как символы, удобные для получения результатов относительно действительных чисел. Разлагал квадратные корни в цепные дроби.

1583 — Т. Финке (1561—1656) в «Геометрии сферы» ввел термин «тангенс».

1585 — С. Стевин (1548—1620) ввел десятичные дроби.

Конец XVI в. — Ф. Виет (1540—1603) ввел буквенные обозначения для неизвестных и постоянных величин. На Руси появляются многочисленные рукописи по арифметике и геометрии с различными практическими примерами.

1614 — Дж. Непер (1550—1617) опубликовал первые таблицы логарифмов. Чуть позже таблицы логарифмов опубликовал Й. Бюрги (1552—1632).

1615 — Опубликована «Новая стереометрия винных бочек» И. Кеплера (1571—1630).

1625 — А. Жирар (1595—1633) вычислил площадь сферического треугольника. В 1629 г. он сформулировал основную теорему алгебры и дал геометрическую интерпретацию отрицательного корня уравнения.

1636—1637 — Р. Декарт (1596—1650) и П. Ферма (1601—1665) начали разрабатывать аналитическую геометрию, сводя при помощи метода координат геометрические задачи к алгебраическим.

1639 — Ж. Дезарг (1593—1662), разрабатывая начертательную и проективную геометрию и теорию перспективы, ввел понятия «бесконечно удаленная точка», «инволюция», «полярное преобразование».

Первая половина XVII в. — Развитие анализа бесконечно малых (вычисления объемов, площадей, центров тяжести, скоростей, ускорений, минимумов и максимумов) в трудах И. Кеплера, Б. Кавальери (1598—1647), Э. Торричелли (1608—1647), П. Ферма, Б. Паскаля (1623—1662), Дж. Валлиса (1616—1703), И. Барроу (1630—1677).

Середина XVII в. — П. Ферма сформулировал задачи, на столетия определившие развитие теории чисел. Б. Паскаль и Х. Гюйгенс (1629—1695) исследовали свойства циклоиды и других линий, образующихся при качении одной линии по другой.

1665 — Б. Паскаль в «Трактате об арифметическом треугольнике», изучая свойства биномиальных коэффициентов, сформулировал и применил метод математической индукции.

1660—1680 — И. Ньютон (1643—1727) и Г. В. Лейбниц (1646—1716) создали дифференциальное и интегральное исчисления. Ньютон ввел в математику

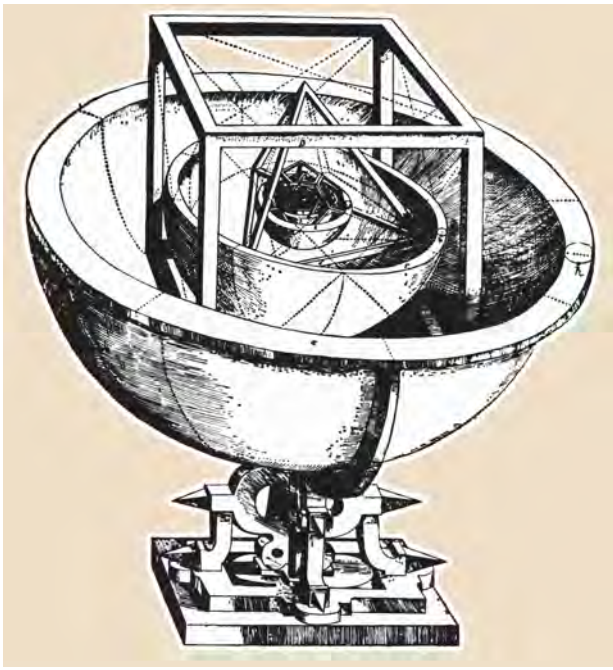
степенные ряды и распространил формулу возведения бинома в степень на случай, когда показатель — любое рациональное число. Дж. Грегори (1638—1675) разложил в степенные ряды синус и косинус.
1667 — Г. В. Лейбниц изобретает счетную машину.
1687 — Опубликованы «Математические начала натуральной философии» И. Ньютона.
1696 — Г. Ф. А. Лопиталь (1661—1704) опубликовал первый учебник анализа, основанный на лекциях И. Бернулли (1667—1748).
Ок. 1700 — И. Ньютон, Я. Бернулли (1654—1705) и И. Бернулли положили начало развитию вариационного исчисления, рассматривая задачу о брахистохроне (линии наискорейшего спуска, циклоиде).
1704 — И. Ньютон классифицировал алгебраические кривые третьей степени.
1713 — Опубликовано сочинение Я. Бернулли, содержащее простейшую форму закона больших чисел.
1733 — А. К. Клеро (1713—1765) ввел понятие «аффинное преобразование».
1730—1770 — Л. Эйлер (1707—1783) рассмотрел функцию $\zeta(s)$, заложив основы аналитической теории чисел. Он открыл основные теоремы элементарной теории чисел, квадратичный закон взаимности, но не доказал его. При исследовании частного случая великой теоремы Ферма (доказательства неразрешимости уравнения $x^3+y^3=z^3$ в натуральных числах) использовал числа вида $m+n\sqrt{-3}$, где m, n — целые числа; это было первым обобщением понятия целого числа.
Ок. 1740 — Г. Крамер (1704—1752) заложил основы теории определителей.
1748 — Л. Эйлер опубликовал «Введение в исчисление бесконечно малых».
1766 — И. Г. Ламберт (1728—1777) доказал иррациональность числа π.
1770—1771 — Ж. Л. Лагранж (1736—1813) проанализировал методы решения в радикалах уравнений первых четырех степеней и объяснил, почему они не годятся для уравнения пятой степени. Он открыл, что разрешимость уравнений связана со свойствами группы перестановок корней уравнения.
1794 — В Париже создана Политехническая школа.
1795 — Г. Монж (1746—1818) опубликовал курс начертательной геометрии и «Приложение анализа к геометрии», где изложил исследования по дифференциальной геометрии и теории поверхностей.
1796 — К. Ф. Гаусс (1777—1885) изучал многочлены деления круга и доказал, что если число $n=2^{2^p}+1$ простое, то правильный n-угольник можно построить циркулем и линейкой. В 1801 г. он опубликовал «Арифметические исследования», где развил теорию сравнений, доказал квадратичный закон взаимности, глубоко развил теорию квадратичных форм.
1821—1823 — О. Л. Коши (1789—1857) развил теорию пределов, определил понятие суммы ряда и непрерывности функции. Он разрабатывал и теорию функций комплексной переменной (1825).
1827 — К. Ф. Гаусс изучает внутреннюю геометрию поверхностей.
1829—1831 — Н. И. Лобачевский (1792—1856) опубликовал первые работы по неевклидовой геометрии. Ранее этими идеями владел К. Ф. Гаусс, а в 1831 г. к ним пришел Я. Бойаи (1802—1860).
1847 — Э. Э. Куммер (1810—1893) развил при помощи так называемых идеальных чисел арифметику целых чисел поля деления круга и таким образом доказал великую теорему Ферма для $3\leqslant n\leqslant 100$.
1849 — П. Л. Чебышёв (1821—1894) получил первые после Евклида результаты о распределении простых чисел.
1871—1889 — Р. Дедекинд (1831—1916), Е. И. Золотарев (1847—1878) и Л. Кронекер (1823—1891) построили теорию идеалов любого поля алгебраических чисел. Дедекинд ввел понятия кольца, идеала и модуля над кольцом.
1872 — Ф. Клейн (1849—1925) опубликовал свою Эрлангенскую программу, в которой изложил общий принцип построения геометрий при помощи теории групп.
1873 — Ш. Эрмит (1822—1901) доказал трансцендентность числа e.
1874 — Г. Кантор (1845—1918) доказал несчетность континуума.
1881—1882 — Р. Дедекинд, Г. Кантор (1845—1918) и К. Т. В. Вейерштрасс (1815—1897) построили, каждый своим способом, теорию действительных чисел. Вскоре в работах Дедекинда и Кантора возникла теория множеств.
1884 — К. Вейерштрасс доказал возможность разложения любой непрерывной на отрезке функции в равномерно сходящийся ряд многочленов.
1896 — Г. Минковский (1864—1909) разработал «геометрию чисел».
1889 — Д. Гильберт (1862—1943) построил систему аксиом геометрии Евклида. С тех пор важную роль в математике играет аксиоматический метод.

Математика в XX в. обогатилась не только новыми результатами, но и многими разделами, которых не было в XIX в. Были решены почти все проблемы, сформулированные Д. Гильбертом, и многие другие трудные задачи (великая теорема Ферма, возможность окраски любой географической карты четырьмя красками и огромное количество менее элементарно формулируемых, но гораздо более важных для науки и практики задач).

Иоганн Кеплер (1571—1630)

— немецкий астроном и математик, открывший законы движения планет. Составил таблицу логарифмов (1624). Происходил из дворянской семьи, обедневшей настолько, что добывали средства к существованию мелкой торговлей. В сочинении «Новые космографические исследования или космографическая тайна» в 1597 г. развивал теорию Коперника. Кеплер предположил, что в центре Вселенной находится Солнце, ближайшая к нему планета — Меркурий. Вокруг соответствующей сферы описал правильный 8-гранник (октаэдр). Радиус его описанной сферы — расстояние от Солнца до Венеры. Продолжая таким же образом, но пользуясь последовательно 20-, 12-, 4- и 6-гранниками, Кеплер получает расстояния до Земли, Марса, Юпитера и Сатурна (другие планеты не были известны).

В 1598 г. терпимое отношение к протестантам сменилось религиозными притеснениями, завершившимися их изгнанием за пределы Штирии.



Кеплеру было предложено сохранить должность профессора астрономии при условии перемены религии, но он с негодованием отверг это предложение. После двухлетних странствований с семьей (он женился в 1597 г. на вдове, имевшей ребенка от первого мужа) Кеплер поступил на службу к Тихо Браге (1546—1601), а после его смерти был назначен на должность имперского астронома, которую до этого занимал Браге.

В руки Кеплера перешли дневники наблюдений Браге, который в обсерватории «Ураниборг» в течение 20 лет скрупулезно измерял положения планет Солнечной системы. Кеплер принялся за обработку результатов наблюдений за движением Марса и увидел, что его прежние догадки о гармонии мира ложны: расстояние от Марса до Солнца непостоянно.

Открытие первых двух законов Кеплера движения планет потребовало 8 лет работы: в 1609 г. в Праге Кеплер опубликовал книгу «Astronomia nova» («Новая астрономия»). Одним из важных источников погрешностей наблюдений Тихо Браге являлось преломление лучей в атмосфере, поэтому параллельно с основной работой Кеплер изучал оптические явления. В 1604 г. вышла книга по оптике «Дополнение к Вителию». В 1611 г. — «Диоптрика», посвященная исследованию оптических стекол. Там дана теория зрительной трубы Галилея, которую сконструировал в 1608 г. голландский мастер оптических стекол Липпершей. «Галилеева труба» была прототипом нашего бинокля и состояла из двояковыпуклого и двояковогнутого стекол. В «Диоптрике» изложена и конструкция трубы Кеплера, состоящей из двух двояковыпуклых стекол и дающей перевернутое изображение объекта.

Кеплер был прекрасным вычислителем и немало способствовал развитию вычислительной техники. Прежние таблицы Птолемея, Регионтана, Ретика и других не давали нужной точности. В 1627 г. вышли составленные Кеплером и названные в честь императора Рудольфа «Рудольфовы таблицы».

Обсерватория Тихо Браге «Ураниборг» на острове Вен близ Копенгагена.

Кеплер был одним из первых математиков, высоко оценивших изданную в 1614 г. Непером таблицу логарифмов, и опубликовал в 1624 г. свою таблицу, устроенную более удобно, чем таблица Непера. В приближенных вычислениях чрезвычайно важную роль играет умение правильно пренебрегать: нужно знать, какие величины можно, а какие нельзя отбрасывать. Это чутье было в высшей степени развито у Кеплера и послужило ему основой при создании основ новой науки — исчисления бесконечно малых. Инфинитезимальными рассуждениями он пользовался уже в «Новой астрономии», но только в книге «Стереометрия винных бочек. . .» (1615) дано систематическое изложение этих идей.

В 1610 г. умерла первая жена Кеплера. Трехлетний сын и восьмилетняя дочь остались на попечении отца, а третий их ребенок, сын, умер незадолго до смерти матери. Чтобы обеспечить более стабильный доход, Кеплер оставил пражскую резиденцию императора и переселился в Линц, где работал школьным учителем математики и вторично женился (1613). Как писал Кеплер, «Господь благословил этот брак» и даровал 8 детей.

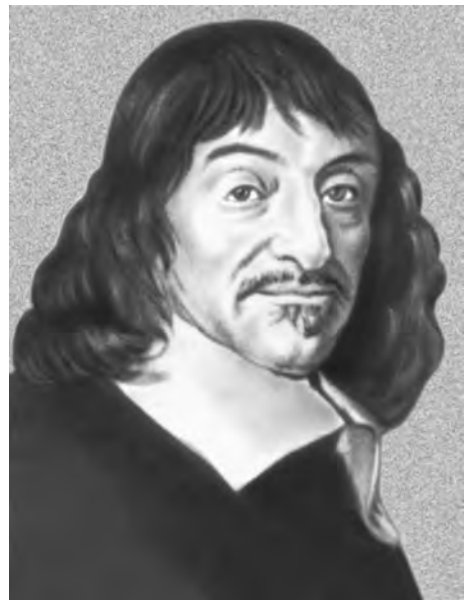
В 1615 г. мать Кеплера Екатерину обвинили в колдовстве и посадили в тюрьму в Леонберге сроком на год. После освобождения она переехала в Линц к сыну Иоганну вместе с другим своим сыном — оловянных дел мастером Христофором. Затем мать Кеплера опять посадили в тюрьму и приговорили к казни без пролития крови, то есть к сожжению. Процесс тянулся 5 лет. Кеплер употреблял все усилия, чтобы спасти мать: писал письма влиятельным лицам, ездил в Регенсбург хлопотать перед имперскими властями. В 1620 г. он добился ее освобождения. Жизнь сына ведьмы в Линце была трудна, поэтому он 8 лет скитался по германским городам. Семья оставалась все это время в Линце. В это время Кеплер получил несколько почетных приглашений в Англию и Италию. Он их не принял, желая жить на родине.

В 1619 г. в книге «Harmonia mundi» («Мировая гармония») Кеплер опубликовал третий из своих законов.

В 1628 г. Валленштейн, организатор и предводитель наемных армий, сражавшихся на полях Германии под знаменем императора и католической церкви, потребовал от императора Фердинанда, чтобы тот сделал его владетельным князем. Император отдал Мекленбургское герцогство, поставив при этом ряд условий, в числе которых новоиспеченному герцогу было вменено в обязанность выплатить ряд императорских долгов, в том числе долг Кеплеру.

Кеплер никак не мог заставить всесильного кондотьера выполнить такое предписание. Поэтому и у Валленштейна он занимался составлением гороскопов. Они не понравились герцогу. Кеплер был отставлен от двора и возвратился в Линц. Оттуда он вновь обратился к имперским властям об уплате жалованья и несколько раз ездил в Регенсбург. Последнее такое путешествие состоялось осенью 1630 г. Кеплеру было уже 59 лет; здоровье было подорвано годами невзгод и путешествий. По дороге в Регенсбург он заболел и вскоре умер. Семья — жена и четверо малолетних детей — не получили после его смерти ни гроша из причитавшегося Кеплеру жалованья; за тридцатилетнюю службу он получил лишь тысячу флоринов — восьмимесячное жалованье.

Рене Декарт (1596—1650)



— в латинском написании Ренатий Картезий — философ и математик. Крайне отрицательно относился к составителям гороскопов, даже точная дата его рождения была опубликована лишь после смерти, во втором издании латинского перевода «Геометрии». Как некогда семь городов Греции оспаривали честь быть родиной Гомера, так четыре французские провинции — Турень, Пуату, Бретань и графство Блуа — боролись за честь быть родиной Декарта. Скорее всего он родился в Турени, в городке Лаэ, ныне переименованном в Лаэ-Декарт.

В роду Декартов были весьма образованные люди. Среди его предков по мужской линии — медики Пьер Декарт и Жан Ферран. Дед его дружил с Гаспаром д'Овернье, переводчиком Макиавелли. Мать происходила по женской линии из семьи Созэ, хранителей королевской библиотеки университета в Пуатье. Не очень понятно, как эти семейные традиции отразились на Декарте, поскольку мать умерла, когда Рене было чуть больше года, а отец не занимался ни наукой, ни литературой — его больше заботили приумножение поместий и судебная карьера.

В раннем детстве Декарт был слаб здоровьем, и отец стремился прежде всего укрепить его. Но ребенок был любопытным, и отец отвечал на его вопросы. Когда мальчик подрос, его отдали в иезуитский колледж Ла Флеш в провинции Анжу.

Иезуиты, в 1594 г. изгнанные из страны Генрихом IV, вернулись, и хотя их деятельность в Париже была по-прежнему запрещена, открыли в 1604 г. в провинции Анжу и нескольких других провинциях свои школы.

Ректор коллежа Этьен Шарле был дальним родственником Декартов. В отступление от (довольно суровых) школьных правил мальчик спал один и мог не присутствовать на утренних занятиях. На всю жизнь у Декарта осталась привычка по утрам, не вставая с постели, предаваться размышлениям; эти часы навсегда стали наиболее плодотворным рабочим временем.

Ученики Ла Флеш изучали литературу, латинский и греческий языки, поэзию и риторику; курс философии включал логику, физику, математику, этику и метафизику. Математика тогда подразделялась на арифметику, геометрию, музыку и астрономию. Среди учебников была «Алгебра» Х. Клавдия (1537—1603), широко известного тогда ученого; это был труд, обобщавший основные результаты алгебраистов XVI в. Сочинений Ф. Виета (1540—1603) в коллеже не знали (видимо, дело в том, что Виет был близок гугенотам).

Школьники ставили спектакли — комедии и балеты. Фехтовали, играли в кегли. Многие, в том числе Декарт, увлекались поэзией. Открытия Г. Галилея, сделанные при помощи телескопа, произвели неизгладимое впечатление на учеников и преподавателей (кто мог представить, что через пару десятилетий он вступит в конфликт с церковью?). Школа была воистину замечательной, Декарт с благодарностью вспоминал своих учителей. Это не мешало ему сомневаться в самых основах философии, которую они преподавали: «Я вижу, как она разрабатывалась в течение многих веков превосходнейшими умами и тем не менее не имеет ни одного пункта, который не вызывал бы споров и, следовательно, не был бы сомнительным».

Самостоятельность мышления, проявленная Декартом уже в школьные годы, часто приводила в замешательство его учителей. Он строил свои рассуждения как геометр: начинал с определений используемых терминов, стремился свести доказываемые предложения к лежащим в их основе основным принципам. «Признаюсь, — писал он в «Правилах для руководства ума», — я родился с таким умом, что главное удовольствие при научных занятиях для меня заключалось не в том, что я выслушивал чужие мнения, а в том, что всегда стремился создать свои собственные. Это — единственное, что в молодости привлекало меня к наукам, и всякий раз, когда какая-нибудь книга сулила в своем заглавии открытие, я пытался, прежде чем приступить к ее чтению,

узнать, не могу ли я достичь чего-либо подобного с помощью своей природной проницательности, и исправно старался не лишать себя невинного удовольствия поспешным чтением».

После завершения образования были две традиционные карьеры: священника и военного. Декарт избрал военную службу. Она сама по себе мало его привлекала, но позволяла «путешествовать, увидеть дворы и армии, встречаться с людьми разных нравов и положений и собрать разнообразный опыт, испытать себя во встречах, которые пошлет судьба, и повсюду поразмыслить над встречающимися предметами так, чтобы извлечь какую-нибудь пользу из таких занятий». В 1618 г. Декарт вступил добровольцем в голландскую протестантскую армию, воевавшую с испано-австрийскими войсками. Военные действия к тому моменту приостановились, воевать ему не пришлось. До 1621 г. он кочевал с армией по Европе. Именно в армии Декарт пришел к мысли, что все науки, кроме математики, базируются не на строгих доказательствах, а скорее на предположениях.

В 1622 г. Декарт приехал в Ренн повидать отца, а затем отправился в Италию. По дороге он посетил Париж, где и познакомился с М. Мерсенном (1588—1648). Есть предположение, что они были знакомы и раньше — Мерсенн тоже учился в Ла Флеш — но из-за разницы в возрасте вряд ли тогда Мерсенн тесно общался с младшим учеником. В те времена сообщить результат Мерсенну — значило сообщить его всем заинтересованным ученым.

Слава Декарта как создателя новой философской системы, базирующейся на математике, быстро распространялась. Декарт решил, не откладывая далее, заняться совершенствованием метода: «Я, может быть, долго еще не решился бы приступить к немцу [этому труду], если бы до меня не дошли слухи, что я его успешно завершил. Не знаю, что дало повод к такому утверждению. Если я и содействовал немного этому своими речами, то лишь признаваясь в своем незнании более откровенно, чем это обыкновенно делают люди, чему-нибудь учившиеся, а может быть, и указывая основания, почему сомневался во многих вещах, считавшихся другими достоверными, но уже никак не похвальбой своего учения. Но имея достаточно совести, чтобы не желать быть быть принятым за большее, чем я есть на самом деле, я считал, что должен приложить все усилия, чтобы сделаться достойным сложившейся репутации».

Декарт писал: «Чтобы решить какую-либо задачу, нужно сначала считать ее как бы решенной и обозначить буквами все, как данные, так и неизвестные, величины. Затем, не делая различия между данными и искомыми величинами, заметить зависимость между ними так, чтобы получить два выражения для одной и той же величины; это приводит к уравнению, служащему для решения задачи, ибо можно приравнять одно выражение другому». При помощи системы координат каждому алгебраическому уравнению от двух переменных можно сопоставить кривую, координаты точек которой удовлетворяют этому уравнению. Таким образом, переходя с языка алгебры на язык геометрии и обратно, можно пользоваться преимуществами обоих методов, обходя трудности. Декарт надеялся, что к алгебраическим уравнениям можно свести все математические задачи. В этом он ошибался: важны и другие методы и теории. Но аналитическая геометрия, начало которой положили Декарт и П. Ферма, является важной областью математики.

В 1648 г. французское правительство в знак признания заслуг назначило Декарту солидную пенсию. Но когда он добрался до Парижа, во Франции наступил политический кризис и правительству стало не до философа. В Париже начались волнения. Люди, позвавшие его во Францию, проявили неприятие и равнодушие.

Королева Христина (1626—1689) мечтала превратить свою столицу в выдающийся научный центр и приглашала в Стокгольм наиболее выдающихся ученых Европы. Друг Декарта дипломат Пьер Шаню, представлявший при шведском дворе Францию, заинтересовал королеву новой философией. Христина пригласила Декарта. К маю 1649 г. Декарт решил уступить настойчивым просьбам, а осенью двинулся в путь.

В Стокгольм он прибыл 1 октября 1649 г. Шаню радушно принял его, королева была любезна; но положение при дворе оказалось неопределенным, среди приближенных к королеве ученых многие завидовали славе Декарта и его милости у королевы. Привычный режим дня оказался нарушен, пришлось отказаться от долгих утренних размышлений и погрузиться в светскую жизнь.

Декарт пытался работать, тосковал по уединению. Тем временем королева решила, что пора начать занятия философией, и назначила их три раза в неделю, начало — в 5 часов утра. Была зима, на редкость холодная даже для Швеции; вынужденный вставать до рассвета и по морозу добираться до дворца, Декарт расшатал свое здоровье. 1 февраля 1650 г. он почувствовал недомогание. От лечения он отказался, прописав сам себе табачную настойку и отвергнув кровопускание. 11 февраля Декарт умер от пневмонии.

Леонард Эйлер (1707—1783)

— математик, механик, физик и астроном. Родился в швейцарском городе Базеле. Отец Эйлера был пастором и хотел, чтобы сын тоже стал священником. В Базельском университете изучал богословие и древние языки, но слушал и лекции И. Бернулли (1667—1748), который занимался с одним Эйлером дополнительно.

В 1727 г. по рекомендации братьев Бернулли переехал в Санкт-Петербург, где нашел весьма благоприятные условия для научной деятельности. За 14 лет своего первого петербургского периода жизни Эйлер подготовил к печати около 80 трудов и опубликовал свыше 50. В Петербурге он изучил русский язык. Читал лекции студентам. Работал над усовершенствованием карт России. Создал двухтомный труд по теории кораблестроения, книгу по теории музыки и общедоступное «Руководство к арифметике».

В 1733 г. Эйлер женился на Екатерине Гзель — дочери академического живописца родом из Швейцарии, вывезенного Петром I из Голландии. Из тринадцати их детей выжили три сына и две дочери. Никакие научные занятия не были для него поводом пренебречь семейными обязанностями: ему приписывают слова «Где больше дадут, туда и служить пойду».

Неустойчивое положение времен регентства Анны Леопольдовны заставило Эйлера принять в 1741 г. предложение прусского короля Фридриха II переехать в Берлин, где предстояла реорганизация почти бездействовавшего Общества наук в новую академию. За 25 лет жизни в Берлине полностью или вчерне подготовил около 300 работ, среди них ряд больших монографий. Сохранил связи с Россией: печатал в изданиях Петербургской академии примерно половину своих статей, редактировал математический отдел ее ученых записок, приобретал для академии научную литературу и оборудование, сообщал в частных письмах научные новости. Годами в берлинском доме Эйлера жили русские ученые, с которыми он вел занятия.

В 1766 г. вернулся в Петербург по приглашению Екатерины II. Он всерьез принял ее предложение участвовать в реорганизации Академии, причем стремился не к автономии науки, а к переплетению деятельности Академии и правительственных учреждений. Однако директором Академии Екатерина назначила младшего брата своего фаворита — графа В. Г. Орлова, а президентом тогда был К. Г. Разумовский, который, как командир Измайловского полка, помог Екатерине во время дворцового переворота, приведшего ее к власти.

Эйлеру она отказала в чине с обычным своим дипломатическим мастерством: «Я дала бы, когда он хочет, чин, если бы не опасалась, что этот чин сравняет его с множеством людей, которые не стоят г. Эйлера. Поистине его известность лучше чина для оказания ему должного уважения».

Правый глаз Эйлера ослеп в 1738 г., а левый почти не видел с осени 1766 г. Но это не лишило его работоспособности. Благодаря сохранившейся силе ума и памяти, при помощи учеников за 17 лет вторичного пребывания в Санкт-Петербурге подготовил около 400 работ, среди них несколько больших книг. Последние годы жизни академические издания не справлялись с потоком его рукописей, и он шутливо обещал Орлову, что его работы будут печатать в «Комментариях» Академии 20 лет после смерти. А на самом деле это длилось полвека!

Эйлер легко вступал в научные дискуссии, давал консультации, охотно думал над случайными задачами и вопросами.Может показаться, что он разбрасывался, проявляя всеядность, но это только на первый взгляд. Он умел своевременно останавливаться, если не видел реальной возможности двигаться вперед; умел организовать жизнь так, чтобы текущие дела не сильно отражались на основном направлении его работы.

По сути, всю жизнь он занимался математикой: его успехи в других науках (механике, астрономии) связаны именно с применением математических методов. В своей швейцарской диссертации 19-летний Эйлер писал: «Я не считаю необходимым подтвердить эту новую теорию опытом, потому что она

полностью выведена из самых надежных и неопровержимых принципов механики и, таким образом, сомнение в том, верна ли она и имеет ли место в практике, просто не может возникнуть». Даже законы Ньютона Эйлер пытался вывести из более общих принципов, а в небесной механике он стремился не получать эмпирические формулы из обработки результатов наблюдений, а делать выводы непосредственно из закона всемирного тяготения. Всюду Эйлер стремился двигаться от теории к практике.

Эйлер дал определение логарифмической функции, согласно которому функция $\ln z$ определена для любого комплексного числа $z \neq 0$ и принимает в каждой точке бесконечно много значений: $\ln z = \ln|z| + i \arg z$, где $\ln|z| = \int\limits_1^{|z|} \frac{dt}{t}$. Именем Эйлера называют и формулу $e^{iz} = \cos z + i \sin z$, связывающую между собой показательную и тригонометрические функции.

Эйлер заметил, что ортоцентр (точка пересечения высот), центр описанной окружности и центр тяжести (точка пересечения медиан) лежат на одной прямой — прямой Эйлера. Кажется, и теорему о пересечении высот в одной точке, пропущенную в «Началах» Евклида, до Эйлера никто не формулировал. Эйлеру принадлежит и открытие окружности девяти точек (на ней лежат основания высот треугольника, середины его сторон и середины отрезков, соединяющих вершины с ортоцентром).

В марте 1736 г. Эйлер писал: «Некогда мне была предложена задача об острове, расположенном в Кенигсберге и окруженном рекой, через которую перекинуто 7 мостов. Спрашивается, может ли кто-нибудь обойти их, переходя только однажды через каждый мост. И тут же мне было сообщено, что никто до сих пор не смог это сделать, но никто и не доказал, что это невозможно. Вопрос этот, хотя и банальный, показался мне, однако, достойным внимания тем, что для его решения недостаточны ни геометрия, ни алгебра, ни комбинаторное искусство».

Ответ получился весьма простой: связный граф можно обойти, пройдя по всем его ребрам по одному разу, если либо степени всех вершин четны — и тогда обход закончится в той же вершине, где он начался, либо же степени двух его вершин нечетны — и тогда обход надо начать в одной из них, а закончить в другой. Эйлер чувствовал, что задача о мостах — лишь начало нового раздела математики (топологии).

Он открыл многочлен $n^2 - n + 41$, значения которого — простые числа при $n = 0, 1, 2, \ldots, 40$. Заметил, что число $2^{2^5} + 1$ делится на 641, и тем самым опроверг предположение Ферма о том, что все числа вида $2^{2^n} + 1$ простые. (Непосредственная проверка всех простых чисел от 3 до 641 была бы непосильна даже для такого виртуозного вычислителя, как Эйлер. Он обнаружил, занимаясь малой теоремой Ферма, что любой делитель числа $2^{2^n} + 1$ дает остаток 1 при делении на 2^{n+2}.)

Занимался он и задачей об обходе шахматной доски конем, который ни на какой клетке не бывает дважды, и конечными аффинными плоскостями, и многими другими отдельными задачами.

Рассмотрев разность $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - \ln n$ частичной суммы гармонического ряда и логарифма, заметил, что она стремится при $n \to \infty$ к величине $0{,}577216\ldots$, ныне носящей имя Эйлера. Рассматривая функцию $\sin x$, Эйлер из того, что она обращается в нуль при $x = 0, \pm\pi, \pm 2\pi, \ldots$, сделал вывод:

$$\sin x = x\left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right)\left(1 - \frac{x^2}{16\pi^2}\right)\ldots$$

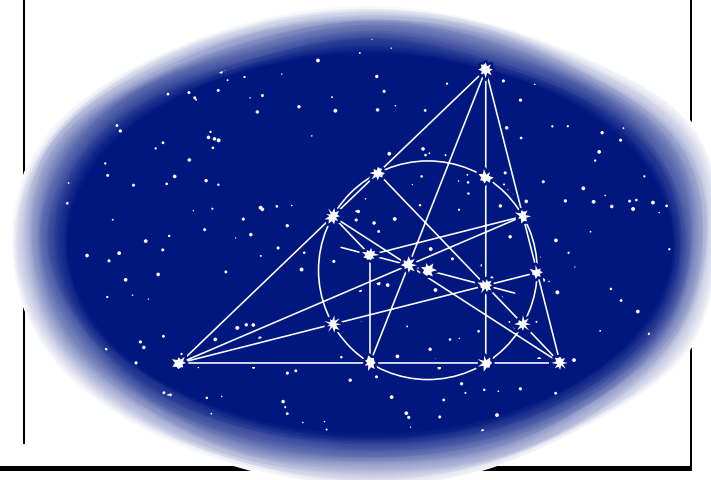
Если перемножить бесконечное множество скобок, то при x^3 получим коэффициент

$$\frac{1}{\pi^2} + \frac{1}{4\pi^2} + \frac{1}{9\pi^2} + \frac{1}{16\pi^2} + \ldots$$

Вспоминая ряд $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \ldots$, коэффициент при x^3 в котором равен $-1/6$, получаем равенство

$$\zeta(2) = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \ldots = \frac{\pi^2}{6}.$$

Аналогично, рассматривая коэффициенты при дальнейших степенях x, получаем равенства $\zeta(4) = \sum\limits_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}$, $\zeta(6) = \sum\limits_{k=1}^{\infty} \frac{1}{k^6} = \frac{\pi^6}{42 \cdot 6!}$ и так далее.



Карл Фридрих Гаусс (1777—1855)

Ему было 12 лет, когда разразилась Французская революция; 29, когда была распущена казавшаяся вечной Римская империя; 38, когда был разгромлен Наполеон; и за 70, когда в Германии произошла революция 1848 г. Его родственники, мелкие фермеры, переехали в Брауншвейг около 1740 г. Средневековые гильдии старались не допустить в город пришельцев из деревни, поэтому даже 30 лет спустя известный в городе своими способностями к счету отец Гаусса работал то садовником, то водопроводчиком, то уличным мясником, то бухгалтером похоронного общества. Главной заботой семьи было приобретение собственного дома: только владелец дома, расположенного в пределах города, мог стать полноправным горожанином.

Гаусс любил рассказывать, что научился считать раньше, чем говорить. По его словам, когда ему было 3 года, отец вычислял, сколько следует заплатить каменщикам, учитывая, что некоторые из них работали и в обеденные часы. Он собирался уже выплачивать деньги, когда сын заявил, что расчет неверен и должно быть столько-то: мальчик в уме повторял выкладки отца; велико же было удивление, когда вторичный расчет подтвердил правоту сына!

В 1784 г. Гаусс пошел в школу. Учитель Бюттнер среди 50 с лишним учеников разного возраста, сидевших одновременно в одной комнате, выделил его и уделил особое внимание. Ассистентом учителя был М. Бартельс (1769—1836), позже профессор математики Казанского университета, который учил математике и Н. И. Лобачевского (1792—1856). Вероятно, именно Бартельс зародил в Гауссе идеи неевклидовой геометрии.

В 1791 г. Гаусс был представлен герцогу Брауншвейг-Вольфенбюттельскому и получил ежегодную стипендию. Награды такого рода не были чем-то необычным в то время и являются прообразом современной системы финансирования образования и науки.

Благодаря покровительству герцога в 1795 г. поступил в Геттингенский университет. Ф. Клейн (1849—1825) в «Лекциях о развитии математики в XIX столетии» писал: «Естественный интерес, какое-то, я сказал бы, детское любопытство приводит впервые мальчика независимо от каких-либо внешних влияний к математическим вопросам. Первое, что его привлекает, это чистое искусство счета. Он беспрестанно считает с прямо-таки непреоборимым упорством и неутомимым прилежанием. Благодаря этим постоянным упражнениям в действиях над числами, например над десятичными дробями с невероятным числом знаков, он не только достигает изумительной виртуозности в технике счета, которой отличался всю жизнь, но его память овладевает таким колоссальным числовым материалом, он приобретает такой богатый опыт и такую широту кругозора в области чисел, какими навряд ли обладал кто-либо до или после него. Путем наблюдений над своими числами, стало быть, индуктивным, «экспериментальным» путем он рано постигает общие соотношения и законы. Этот метод, стоящий в резком противоречии с современными навыками математического исследования, был, однако, довольно распространен в XVIII столетии и встречается, например, также у Эйлера... Все эти ранние, придуманные только для своего удовольствия забавы ума являются подходами к значительной, лишь позже осознанной цели. В том-то именно и заключается подсознательная мудрость гения, что он уже при первых пробах сил, полуиграя, еще не сознавая всего значения своих действий, попадает, так сказать, своей киркой как раз в ту породу, которая в глубине своей таит золотоносную жилу». (Сам Гаусс утверждал, что отличается от других людей лишь прилежанием. Известны его слова «Nil actum reputans si quid superesset agendum» — «Что не сделано до конца, вообще не сделано».)

В 1799 г. Гаусс защитил диссертацию на тему «Новое доказательство теоремы о том, что каждая целая рациональная алгебраическая функция одной переменной может быть разложена на действительные множители первой или второй степени». Коротко говоря, тема — основная теорема алгебры: каждый непостоянный многочлен имеет хотя бы один комплексный корень. Заметьте: хотя Гаусс, несомненно, владеет идеей комплексного числа, он в названии диссертации остается в вещественной области! В этом проявляется желание выяснять суть дела, а не искать эффектные переформулировки уже известных результатов.

1 января 1801 г. Дж. Пиацци (1746—1826) открыл Цереру — один из сотен астероидов, рассеянных между Марсом и Юпитером. Наблюдение за Церерой было выполнено на интервале величиной 9°. Воспользоваться прежними методами определения орбит было невозможно: они требовали значительного по объему и подтвержденного повторными наблюдениями материала, который для больших планет, известных с древности, и в самом деле имелся. Гаусс составил уравнение восьмой степени и, проведя подробнейшие приближенные вычисления, нашел орбиту. Затем он учел результаты других наблюдений и уточнил ответ с помощью метода наименьших квадратов, известного ему с 1795 г. Планету нашли в указанном месте! Этот результат, оценить который могла и широкая публика, принес Гауссу первую славу: с 1807 г. и до конца жизни он был директором обсерватории Геттингена.

С 1820 г. руководил геодезической съемкой Ганноверского королевства, много работал в полевых условиях, измеряя длину дуги меридиана Геттинген—Альтона. Изобрел и многократно использовал гелиотроп — прибор, позволяющий при помощи концентрации солнечных лучей получать хорошо видимые точки визирования. Изучение формы земной поверхности потребовало общего геометрического метода исследования поверхностей («Общие исследования о кривых поверхностях», 1827 г.). Гаусс дал определение кривизны как величины $\frac{1}{r_1 r_2}$, где r_1 и r_2 — главные радиусы кривизны в рассматриваемой точке; доказал теорему о постоянстве кривизны при изгибаниях поверхности (без растяжений).

Он надеялся выяснить, не отклоняется ли сумма углов большого треугольника от 180°: его интересовал не философский вопрос о независимости аксиомы о параллельных, а свойства физического пространства. Полученное отклонение укладывалось в пределы ошибки эксперимента, а ответ на вопрос Гаусса по существу не получен и по сей день.

Первое математически небессмысленное изложение неевклидовой геометрии опубликовал в 1832 г. Я. Бойаи (1802—1860), сын друга Гаусса Ф. Бойаи (1775—1856). Гаусс признал храбрость Яноша и его математические достижения, но не захотел обсуждать вопрос, какова геометрия физического пространства. Поддержал Гаусс и публикации Н. И. Лобачевского, отметив, что давно знает все это.

В 1850 г. он писал: «. . .Вы совершенно не правы, если думаете, что я имею в виду лишь окончательную отделку языка и элегантность изложения. На это уходит сравнительно незначительная доля времени; но то, что я имею в виду, — это внутреннее совершенство. В иных из моих работ есть такие особые точки, которые стоили мне нескольких лет размышлений и где на маленьком пространстве сконцентрировано представление, про которое никто не замечает, какие трудности мне пришлось преодолеть!» А успеть придать всем своим мыслям законченную форму он не мог! Например, он планировал труд, который должен был содержать (выполненное в 1812 г.) исследование гипергеометрического ряда

$$F(\alpha,\beta,\gamma;x)=\sum_{k=0}^{\infty}\frac{\alpha^{\overline{k}}\beta^{\overline{k}}}{k!\cdot\gamma^{\overline{k}}}\cdot x^k=1+\frac{\alpha\beta}{1\cdot\gamma}x+\frac{\alpha(\alpha+1)\beta(\beta+1)}{1\cdot 2\cdot\gamma(\gamma+1)}x+\ldots,$$

теорию дифференциальных уравнений с рационально зависящими от x коэффициентами и теорию эллиптических функций. Но поскольку он так и не разобрался толком с многозначностью аналитических функций (и не мудрено, римановы поверхности открыты значительно позже!), то об эллиптических функциях Гаусс так никогда ничего и не опубликовал, и только после смерти из его бумаг выяснилось, что Якоби и Абель переоткрывали факты, известные ему с 1800 г.

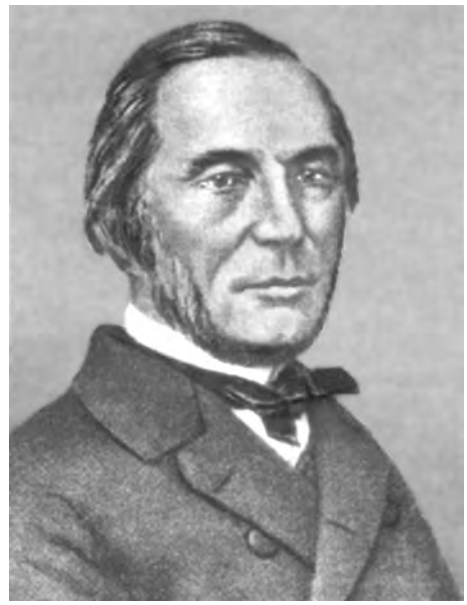
В 1811 г. (задолго до работ Коши!) он рассмотрел интеграл $\int\frac{dz}{z}$ по контуру, n раз обходящему вокруг начала координат, и сказал, что этот интеграл равен $2n\pi i$.

Исследования Гаусса по теоретической физике (1830—1840) являются результатом совместной работы с В. Э. Вебером (1804—1891). Они вдвоем создали абсолютную систему электромагнитных единиц (CGS) и сконструировали первый электромагнитный телеграф (1833). Гауссу принадлежит понятие потенциала электрического поля. В его честь названа единица измерения магнитной индукции.

В возрасте 62 лет Гаусс изучил русский язык. Кроме научной литературы, просил, например, прислать ему «Капитанскую дочку» А. С. Пушкина.

В 1855 г. была выпущена медаль с надписью «Mathematicorum princeps» («Король математиков»). Памятник Гауссу в Брауншвейге, согласно его завещанию, стоит на постаменте в виде правильного 17-угольника.

Пафнутий Львович Чебышёв (1821—1894)



родился в дворянской семье в селе Окатово Боровского уезда Калужской губернии. Получив домашнее образование, в 1837 г. поступил в Московский университет и в 1841 г. окончил его. В 1846 г. защитил магистерскую диссертацию «Опыт элементарного анализа теории вероятностей». В 1849 г. переехал в Петербург и защитил докторскую диссертацию «Теория сравнений». Потратил много сил, чтобы систематизировать и издать труды Л. Эйлера по арифметике.

Основал первую в России крупную математическую школу — Петербургскую. Чебышёв вполне осознавал свое значение в математике. Д. А. Граве (1863—1939) приводит его слова: «Разговаривали как-то мы втроем: Эрмит — крупнейший математик Франции, Сильвестр — крупнейший математик Англии, и я». Его учениками были Е. И. Золотарев (1847—1878), А. М. Ляпунов (1857—1918), А. А. Марков (1856—1922); он оказал существенное воздействие на творчество Г. Ф. Вороного (1868—1908), А. Н. Коркина (1837—1908), Ю. В. Сохоцкого (1842—1883) и других.

Одной из особенностей его творческой манеры была редкостная самостоятельность и независимость мышления. Он мало читал, но исключительно интенсивно размышлял как над проблемами, поставленными предшественниками, так и над теми, что ставили перед ним «требования практики». Одним из девизов его творчества было «удовлетворять требованиям практики». Ему принадлежат слова: «Сначала задачи ставили Боги, потом — полубоги. А нас заставляет ставить задачи нужда».

Отстраненность от достижений современников нередко заставляла преодолевать трансцендентные препятствия там, где это не вызывалось необходимостью. Штудируя его работы, трудно представить себе, как можно было прийти к полученным Чебышёвым формулам наилучшего приближения. Но комплексный анализ О. Коши, к творчеству которого Чебышёв питал резкую и малообоснованную неприязнь, давал ключ к пониманию сути дела.

Чебышёв был достаточно богатым человеком, но при этом весьма экономным и бережливым. У него было хобби: любил покупать имения. При этом он ни разу не осматривал имения до покупки, а только анализировал поведение тех, кто торговался с ним, и принимал решение. И ни разу не ошибся в выборе.

Чебышёв служил на многих поприщах — научном, педагогическом, инженерном. С 1856 г. работал в течение 13 лет в Артиллерийском отделении Военно-ученого комитета и 17 лет в Ученом комитете Министерства народного просвещения. Он сконструировал более 40 механизмов. Показ их в Чикаго произвел потрясающее впечатление на современников. Чебышёв построил стопоходящую машину, воспроизводящую движения животных при ходьбе, гребной механизм, имитирующий движения весел лодки, самокатное кресло и арифмометр-полуавтомат (хранится в Париже в Музее искусств и ремесел).

Говорят, как-то раз он объявил лекцию по раскрою одежды. Среди слушателей было много портных. Лекция началась словами: «Предположим, что человек имеет форму шара. . . » Дело в том, что Чебышёв был основоположником теории аппроксимации: ему принадлежит постановка задачи о наилучшем равномерном приближении функции, непрерывной на отрезке, алгебраическими полиномами. Он сформулировал необходимое условие полинома наилучшего приближения и нашел полиномы и рациональные функции для наилучшего приближения нескольких важных конкретных элементарных функций.

Теорема Чебышёва об альтернансе гласит: ***чтобы многочлен g степени n наименее уклонялся в метрике $C([a, b])$ от непрерывной функции f, необходимо и до-***

статочно, чтобы нашлись такие $n+2$ точек $a \leqslant \xi_1 < \xi_2 < \ldots < \xi_{n+2} \leqslant b$, что в них функция $f(x)-g(x)$ достигает, чередуя знаки, свои максимальные и минимальные значения, равные по модулю. (Поясним, что в метрике $C([a; b])$ расстояние между любыми двумя непрерывными функциями f и g определено как максимум модуля разности $f(x)-g(x)$, где x пробегает весь отрезок $[a; b]$.)

При этом говорят, что функция $f(x)-g(x)$ имеет $(n+2)$-альтернанс. Чебышёв в своем мемуаре 1854 г. сформулировал утверждение теоремы об $n+2$ точках (без условия чередования знаков), как необходимое условие экстремума, а в 1859 г. дал набросок доказательства. Элементарное доказательство теоремы об альтернансе восходит к Э. Борелю (1871—1956) и опубликовано лишь в 1905 г. Теперь оно содержится почти во всех учебниках математического анализа. Мы не будем его здесь воспроизводить, но все-таки расскажем о многочленах степени n со старшим коэффициентом 2^{n-1}, наименее уклоняющихся от нуля на отрезке $[-1; 1]$. Эту систему многочленов Чебышёв открыл в 1854 г.: $T_0(x)=1$, $T_1(x)=x$, $T_2(x)=2x^2-1$, $T_3(x)=4x^3-3x$, $T_4(x)=8x^4-8x^2+1$, и вообще, $T_n(x)=\cos(n \arccos x)$.

Идея в следующем. $\cos 2t = 2\cos^2 t - 1$. Обозначив $\cos t = x$, получаем $\cos(2 \arccos x) = 2x^2 - 1$. На отрезке $[-1; 1]$ график функции $y=2x^2-1$ дважды выходит на уровень $y=1$ и один раз — на уровень $y=-1$. Точки $x=-1$, 0 и 1 являются точками альтернанса. Нетрудно проверить, что $T_{n+1}(x)+T_{n-1}(x)=2xT_n(x)$ и

$$(1-x^2)T_n''(x)-xT_n'(x)+n^2T_n(x)=0.$$

Чебышёв является одним из создателей теории ортогональных полиномов: его многочлены ортогональны на отрезке $[-1; 1]$ с весом $\frac{1}{\sqrt{1-x^2}}$, то есть

$$\int_{-1}^{1} \frac{T_n(x)\,T_m(x)}{\sqrt{1-x^2}}\,dx = \begin{cases} 0, & m \neq n, \\ \pi/2, & m=n=1, \\ \pi, & m=n=0. \end{cases}$$

В 1853 г. он доказал, что неопределенный интеграл $\int x^m(a+bx^n)^p\,dx$, где a и b — действительные числа, m, n, p — рациональные числа, не выражается через элементарные функции ни при каких m, n, p, за исключением случаев, когда одно из чисел p, $\frac{m+1}{n}$ и $p+\frac{m+1}{n}$ — целое.

В статьях «Об определении числа простых чисел, не превосходящих данной величины» и «О простых числах» доказал для всех достаточно больших x неравенства

$$0{,}92 < \frac{\pi(x)\ln x}{x} < 1{,}06,$$

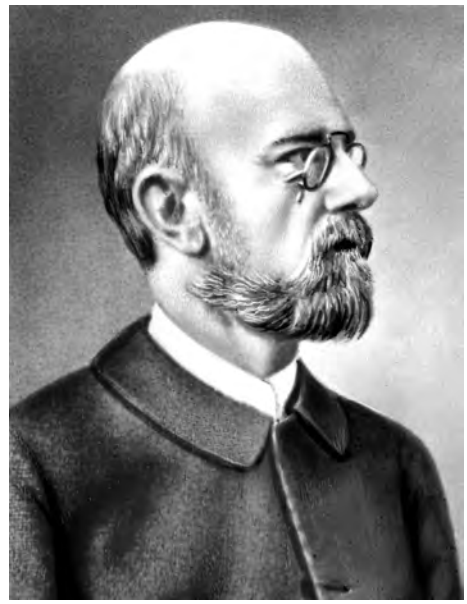
где $\pi(x)$ — количество простых чисел, не превосходящих числа x. Доказал он и постулат Ж. Л. Ф. Бертрана (1822—1900), согласно которому между числами n и $2n$ при $n>3$ всегда содержится хотя бы одно простое число.

Чрезвычайно важны работы Чебышёва по теории вероятностей. Это сравнительно молодая область математики возникла в XVI—XVII вв. Ее первые задачи были связаны с азартными играми, а впоследствии — с обработкой результатов наблюдений. И хотя многие факты ко времени Чебышёва были уже известны, методы этой теории были лишены должной строгости.

На практике вероятности находят, многократно повторяя опыт и вычисляя долю случаев («частоту»), в которых произошло интересующее нас событие. Например, если много раз подбросить монету, то она упадет цифрой вверх примерно в половине случаев. Такая устойчивость частоты при многократном повторении испытания наблюдается во многих ситуациях. Математическое объяснение этой устойчивости дал Я. Бернулли (1654—1705) в книге «Искусство предположения», опубликованной в 1713 г., доказавший закон больших чисел: если в каждом из n независимых испытаний с одной и той же вероятностью p может произойти некоторое событие A, то количество ξ появлений события A хотя и не обязано в точности равняться np и может сильно отклоняться от этой величины; но вероятности значительных отклонений малы. Точнее говоря, для любых положительных чисел ε и η вероятность $\mathsf{P}\left(\left|\frac{\xi}{n}-p\right|>\varepsilon\right)$ меньше η при всех достаточно больших n. Более простое, чем Бернулли, доказательство закона больших чисел дал Чебышёв.

А. Н. Колмогоров писал: «С методологической стороны основной переворот, совершенный Чебышёвым, заключается не только в том, что он впервые с полной настойчивостью выдвинул требование абсолютной строгости доказательства предельных теорем. . . , но главным образом в том, что Чебышёв всюду стремился получить точные оценки отклонений от предельных закономерностей, возможных при хотя бы и большом, но конечном числе испытаний, в виде безусловно правильных при любом числе испытаний неравенств». Он впервые понял всю силу понятий случайной величины, ее математического ожидания (среднего значения) и дисперсии (среднеквадратичного уклонения). Теория вероятностей, по Чебышёву, имеет целью нахождение вероятностей одних событий по известным вероятностям других событий, и тем самым является естественной частью математики.

Давид Гильберт (1862—1943)



родился в Велау близ Кенигсберга, в семье окружного судьи. С 1870 г. начал посещать приготовительную школу, а с 1872 г. — гимназию королевского Фридрихсколлежа, ориентированную на заучивание большого объема гуманитарных сведений, что было затруднительно для Давида, более склонного к логическим рассуждениям. Последний учебный год учился в государственной Вильгельм-гимназии, где математике уделяли больше внимания, что было облегчением для Гильберта, так как в математике ничего не надо бездумно заучивать. В 1880 г. поступил в Кенигсбергский университет. Отец желал, чтобы сын по семейной традиции стал юристом, но Давид записался на математический курс философского факультета: сказалось влияние матери — любительницы философии, астрономии и простых чисел.

Второй семестр, согласно немецкой традиции путешествовать по университетам, учился в Гейдельберге. Вернувшись в Кенигсберг, слушал, в частности, лекции по теории чисел и теории функций Г. Вебера (1842—1913), который познакомил его с теорией инвариантов. Одной из ее задач посвящена первая заметка Гильберта (1883). Дружил с А. Гурвицем и Г. Минковским: каждый вечер в пять часов они шли к яблоне и обменивались знаниями, мыслями и научными планами. По окончании университета в декабре 1884 г. сдал устный экзамен, в феврале 1885 г. — публичный выпускной экзамен, а в мае — государственный экзамен на право преподавания в гимназии. Затем, по старой университетской традиции, отправился в научное путешествие.

Вначале Гильберт посетил Лейпциг, где встретился с Феликсом Клейном (1849—1925), который в том же году представил к печати заметку Гильберта по теории инвариантов. В 1886 г. Гильберт посетил Париж и в конце июня возвратился в Кенигсберг. Ему предстояла хабилитация — серия испытаний на право преподавания в университете. Гильберт представил работу по теории инвариантов, прочитал лекцию «Самые общие периодические функции», сдал устный экзамен и стал доцентом, а спустя несколько лет и профессором Кенигсбергского университета.

Во время своей поездки в Лейпциг Гильберт произвел сильное впечатление на Клейна, который тогда же решил добиться, чтобы столь многообещающий математик работал в Геттингене. К 1895 г. его усилия увенчались успехом, и Гильберт с семьей перебрался в Геттинген. С тех пор его жизнь была неразрывно связана с Геттингеном.

Просто одетый, энергичный человек среднего роста с русой бородкой, манерами он напоминал учителя гимназии. В начале каждой лекции кратко повторял содержание прошлой, читал неторопливо, словно диктуя отдельные фразы, чтобы слушатели успевали записать за ним и все понять. К выбору тематики курсов лекций Гильберт подходил своеобычно. В самом начале своей карьеры преподавателя он решил объявлять лекции по тому разделу математики, который он сам намеревался изучить глубже. Он полагал, что лучше всего можно понять и усвоить предмет, если объяснять его другим. На его лекции собиралось много народа, иным приходилось устраиваться на подоконниках — в аудитории не хватало мест.

Научную деятельность Гильберта можно, в первом приближении, разбить на периоды, когда он занимался какой-либо одной областью математики:

- теория инвариантов (1885—1893);
- теория алгебраических чисел (1893—1898);
- основания геометрии (1898—1902);
- анализ (1902—1912);
- физика (1910—1922);
- основания математики (1922—1930).

Конечно, были и отклонения: например, в 1909 г., посреди занятий интегральными уравнениями, он при помощи преобразований интегралов решил задачу Варинга о том, что для каждого натурального n существует такое k, что любое натуральное число можно представить в виде суммы не более чем k слагаемых, каждое из которых — n-я степень.

К 1888 г. Гильберт решил проблему Гордана о существовании конечного базиса инвариантных форм. Работа эта настолько фундаментальная, что ее основную идею мы здесь обсудим.

Идеалом называют такое множество I элементов ассоциативного и коммутативного кольца K с единицей, что как сумма любых двух элементов идеала, так и произведение любого элемента идеала на любой элемент кольца принадлежит множеству I. Главным идеалом (a) называют множество всех кратных элемента a, то есть $\{ax \mid x \in K\}$. В кольце целых чисел, благодаря алгоритму Евклида, все идеалы главные. Не во всех кольцах все идеалы главные. Например, в кольце $\mathbb{Z}[x]$ многочленов одной переменной с целыми коэффициентами не является главным идеал $(2, x)$, состоящий из всех многочленов, свободные члены которых четны. Не является главным в кольце $\mathbb{R}[x, y]$ многочленов двух переменных идеал, состоящий из многочленов, свободные члены которых равны 0.

Определение. Кольцо K нётерово, если всякий его идеал I обладает конечным множеством образующих, то есть если для некоторого конечного подмножества $i_1, i_2, \ldots, i_n$ множества I имеем

$$I = \{i_1 a_1 + i_2 a_2 + \ldots + i_n a_n \mid a_1, a_2, \ldots, a_n \in K\}.$$

Нетрудно доказать, что нётеровы кольца — это в точности кольца, удовлетворяющие условию обрыва возрастающих цепочек идеалов.

Теорема Гильберта о базисе. ***Если кольцо K нётерово, то и кольцо K[x] многочленов одной переменной с коэффициентами из K тоже нётерово.***

В 1898 г. Гильберт объявил курс лекций по геометрии, а в следующем году опубликовал «Основания геометрии». В этой небольшой книге дана полная система аксиом геометрии Евклида; аксиомы распределены по группам, рассмотрены их следствия и исследованы «геометрии», получающиеся при изменении или изъятии некоторых аксиом.

Гильберт занимался обоснованием принципа Дирихле. Этот принцип широко применяли с начала XIX в. для доказательства существования и единственности решений краевых задач теории дифференциальных уравнений. К. Т. В. Вейерштрасс (1815—1897) указал пример, который заставил усомниться в границах применимости этого принципа. Гильберт решил эту проблему.

На Втором международном конгрессе математиков в августе 1900 г. он прочитал доклад «Математические проблемы», где сформулировал 23 проблемы, во многом определившие развитие математики XX в. Некоторые из них все еще не решены.

Гильберт говорил: «Математическую теорию можно считать совершенной только тогда, когда ты сделал ее настолько ясной, что берешься объяснить ее содержание первому встречному».

Зимой 1920—1921 гг. он читал четырехчасовые популярные лекции, которые обработали В. Роземан и С. Э. Кон-Фоссен (1902—1936). Результат этой работы — опубликованная в 1932 г. «Наглядная геометрия», надолго ставшая образцовым изложением многих разделов геометрии.

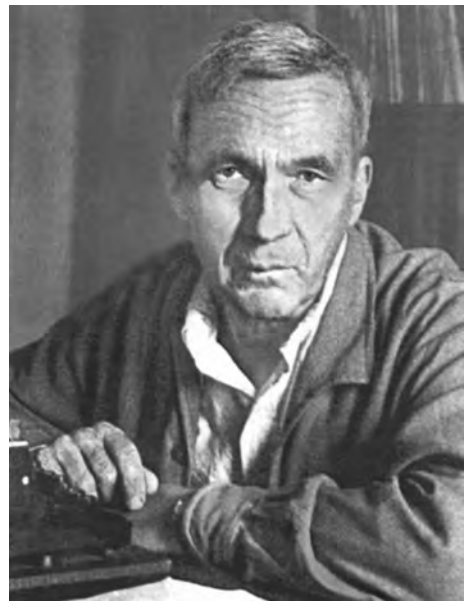
В 1904 г. на Третьем международном математическом конгрессе он выступил с докладом «Об основаниях логики и арифметики», в котором развиты его исследования в этом направлении, уже намеченные в «Основаниях геометрии». Затем были статьи «Логические основания математики», «Проблемы обоснования математики» и двухтомная монография «Основания математики», написанная совместно с И. П. Бернайсом (1888—1977).

Почему вообще перед математиками встали эти вопросы? В работах Коши, Вейерштрасса, Кантора и других математиков была построена теория множеств, ставшая естественной частью математического анализа. Но в ней к началу XX в. были обнаружены парадоксы; чтобы оградить математику от них, некоторые, например Л. Кронекер (1823—1891) и Л. Г. Я. Брауэр (1881—1966), считали неизбежным ограничиться «безопасными» разделами, в которых можно было не опасаться парадоксов. Гильберт был сторонником сохранения математики в полном объеме. Он предлагал аксиоматизировать каждый из разделов математики и потом исследовать полученные системы аксиом. Тем самым Гильберт внес значительный вклад в развитие математической логики. И хотя выяснилось, что в первоначальном оптимистическом варианте программа Гильберта невыполнима, сама постановка задачи способствовала прогрессу логики.

Не только математика захватывала Гильберта. В молодости он был заядлым танцором, охотно танцевал даже после 50 лет. Любил возиться в саду. Одно время он увлекался лыжами и тогда из дома в университет отправлялся не пешком, а на лыжах.

23 января 1930 г. Гильберту исполнялось 68 лет. Это был официальный возраст для ухода в отставку. Сограждане высоко оценили его деятельность. Одна из улиц Геттингена была названа Гильбертштрассе. Не обошел его вниманием и родной город. Городской совет Кенигсберга присвоил ему звание почетного гражданина. На устроенной в этой связи церемонии Гильберт выступил с речью «Познание природы и логика», которую закончил словами: «Мы должны знать. Мы будем знать».

Андрей Николаевич Колмогоров (1903—1987)

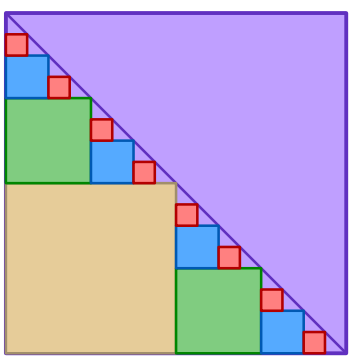


проявил интерес к математике очень рано: в возрасте четырехпяти лет он вел математический отдел в семейном детском журнале «Весенние ласточки», где публиковал придуманные им задачи и математические открытия, в частности, что сумма первых n нечетных чисел равна n^2. (Похожую историю о суммировании первых 20 натуральных чисел любил рассказывать про себя К. Ф. Гаусс.)

Так как мама умерла во время родов, Андрея воспитывала ее сестра. Его детство прошло под Ярославлем, в имении деда, предводителя угличского дворянства. В Москве он учился в одной из лучших частных московских гимназий. В 14 лет самостоятельно изучил дифференциальное и интегральное исчисление по энциклопедическому словарю Брокгауза и Ефрона.

В 1920 г. поступил в Московский университет и стал учеником Н. Н. Лузина (1883—1950). Своими учителями считал П. С. Александрова (1896—1982), П. С. Урысона (1898—1924), А. К. Власова (1868—1922) и В. В. Степанова (1889—1950). В 1929 г. окончил аспирантуру и до конца жизни работал на механико-математическом факультете МГУ.

На одной из лекций Н. Н. Лузин сослался на то, что если внутри квадрата K со стороной 1 расположено несколько не пересекающих друг друга квадратов, каждый из которых пересекает диагональ квадрата K, то сумма периметров этих квадратов не может быть слишком большой. На следующем занятии студент Колмогоров предъявил изображенную на рисунке конструкцию. На рисунке изображены 4 этапа, которые дают нам 1 желтый квадрат, 2 зеленых, 4 синих и 8 красных. На каждом следующем этапе длина стороны квадрата уменьшается вдвое; но вдвое увеличивается и количество квадратиков! Набрав таким образом сколь угодно большую сумму периметров, можно чуть уменьшить размеры всех квадратиков, чтобы они перестали иметь общие точки, а затем чуть сдвинуть их так, чтобы все они стали пересекаться с диагональю квадрата K, но по-прежнему не пересекались друг с другом.



Конечно, это — довольно простое, хотя и очень красивое замечание. Но важно, что после этого лектор не обиделся на студента, заметившего его ошибку, а постарался помочь ему.

В своей первой научной работе, написанной в 1920 г., Колмогоров изучал операции над множествами. Летом 1921 г. построил функцию, ряд Фурье которой почти всюду расходится. Работа двадцатых годов по математической логике (о принципе исключенного третьего) входит в число фундаментальных работ по математической логике.

Начиная с 1924 г., на протяжении последующих сорока лет Колмогоров работал в теории вероятностей и завоевал положение лидера этой науки, дав в 1933 г. в книге «Основные понятия теории вероятностей» общепринятую ныне ее аксиоматику.

В тридцатые годы XX в. изучал марковские случайные процессы; доказал теорему о скорости сходимости эмпирической функции распределения к истинной; построил теорию экстраполяции случайных процессов (параллельно, но чуть позже, эту теорию создал Н. Винер (1894—1964), который считал это одним из высших своих достижений); дал аксиоматическое построение многообразий постоянной кривизны; занимался топологической алгеброй.

Важную роль сыграл цикл его исследований по алгебраической топологии, где он одновременно с аме-

риканским математиком Дж. У. Александером (1888—1971) изобрел когомологические группы. Знаменитая работа по теории дифференциальных уравнений (совместная с И. Г. Петровским и Н. С. Пискуновым), в которой был получен принципиальный результат о бегущей волне, имеет широкую область применимости. В общей топологии построил первый пример открытого отображения, повышающего размерность.

В функциональном анализе Колмогоров определил понятие линейного топологического пространства и дал критерий его нормируемости, тем самым начав изучение топологических векторных пространств. В совместной с И. М. Гельфандом работе заложил первый камень в теорию банаховых алгебр.

В сороковые годы Колмогоров занялся локальной теорией турбулентности. Во время Второй мировой войны занимался статистической теорией стрельбы, а в первые послевоенные годы разрабатывал основы теории статистического контроля продукции; но к военным работам после войны привлечен не был. Это было счастьем для науки и просвещения и привело в пятидесятые годы к выдающимся результатам в небесной механике; решению (совместно с В. И. Арнольдом) 13-й проблемы Гильберта: всякую непрерывную функцию нескольких переменных удалось представить в виде композиции непрерывных функций одной переменной и функции $(x; y) \mapsto x + y$.

Он ввел понятие энтропии динамической системы, совершившее переворот в теории динамических систем, а также понятие ε-энтропии. При этом получил многие принципиальные результаты в теории аппроксимации и начал разработку концепции случайности как меры сложности объекта. Сейчас, когда широко распространены программы-архиваторы данных, определение сложности по Колмогорову как длины кратчайшего алгоритма, конструирующего этот объект, кажется самоочевидным. Но ведь кто-то должен был это «очевидное» определение сформулировать!

Колмогоров никогда не занимался состязательным спортом, но очень много внимания уделял физической культуре. В конце жизни жаловался не только на то, что стало трудно читать и говорить (каждое слово произносилось им с трудом, но — удивительное дело — если убрать паузы, получалась абсолютно грамотная и логичная речь!), но и на то, что он перестал видеть лыжню. В 1982 г., уже тяжело больной, он в апреле повел группу школьников во время Всесоюзной математической олимпиады на одесский пляж и — к ужасу присутствующих — разделся и поплыл в море! Если бы он начал тонуть, никто не решился бы броситься в ледяную воду. . .

Последние четверть века своей жизни Колмогоров посвятил проблемам школьного математического образования. Он руководил математическими олимпиадами, отделом математики журнала «Квант», создал школу-интернат (ныне СУНЦ при МГУ). В интернате и на мехмате создал систему математического практикума.

Он участвовал в школьной реформе конца 1960-х гг. Он взялся за осуществление этой реформы, понимая всю трудность такой работы и отдав ей все свои силы. Он говорил: «Не думайте. . . , что мой духовный мир целиком и полностью занимает математика. Я внутренне свободный человек и позволяю себе свободно размышлять надо всем и критически все оценивать».

Среди целей образования, по Колмогорову, должно присутствовать формирование научного мировоззрения. Он писал: «Вряд ли нужно доказывать, насколько желательно с общеобразовательной точки зрения достигнуть того, чтобы все учащиеся могли вполне конкретно понять хотя бы ньютоновскую концепцию математического естествознания».

Много энергии потратил Колмогоров на создание курса школьной геометрии. Замысел был таков: «постепенно подготовить материал для понимания возможностей разных «геометрий», отличных от евклидовой (как геометрия Лобачевского) или охватывающих евклидову в качестве частного случая (как «метрические пространства»)». Колмогорову не было позволено тронуть систему педагогических вузов, и он всегда ощущал мощное сопротивление всевозможных методистов и авторитетов, не являющихся квалифицированными математиками, но занимающих высокие посты в педагогической иерархии. Но даже они были вынуждены мириться с его деятельностью, пока высокий уровень образования был нужен стране. С исчезновением СССР и потребности в наукоемких производствах неизбежно произошла обратная реформа образования, которую Колмогоров тяжело переживал. Разрушено очень многое. Достаточно сказать, что тираж «Кванта», при Колмогорове менявшийся от 250 000 до 400 000 экземпляров, уменьшился в 100 раз. Дифференциальное и интегральное исчисление, по существу, вычеркнуты из программ общеобразовательных школ, произошел откат на десятилетия назад. Но это не значит, что деятельность Колмогорова пропала даром: остались написанные им книги, которые читают во всем мире. Работают его ученики и ученики его учеников. Математика как наука сейчас развивается удивительно быстро. Да и математическое образование в России рано или поздно возродится. В этом нам очень помогут труды и идеи Колмогорова.

Предметный указатель

Именной указатель