

132

ВЫПУСК

132



# Библиотечка КВАНТ

Библиотечка КВАНТ

КОЛМОГОРОВСКОЙ  
ШКОЛЕ – ПЯТЬДЕСЯТ

СБОРНИК СТАТЕЙ  
Часть 2



БИБЛИОТЕЧКА  
**КВАНТ**  
вывпуск

**132**

Приложение к журналу  
«Квант» №4/2014

**Колмогоровской школе –  
пятьдесят**

**Сборник статей**

**Часть 2**

*Составители*

*В.В.Вавилов, В.Н.Дубровский, А.А.Егоров*

**Москва**  
**Издательство МЦНМО**  
**2014**

УДК 51(07)  
ББК 22.1-22.3ж 74.04(2)  
К60

Серия «Библиотечка «Квант»  
основана в 1980 году

**Редакционная коллегия:**

Б.М.Болотовский, А.А.Варламов, Г.С.Голицын, Ю.В.Гуляев,  
М.И.Каганов, С.С.Кротов, С.П.Новиков, В.В.Производов, Н.Х.Розов,  
А.Л.Стасенко, В.Г.Сурдин, В.М.Тихомиров, А.Р.Хохлов,  
А.И.Черноуцан

**К60 Колмогоровской школе – пятьдесят. Сборник статей.**  
**Часть 2.** Составители В.В.Вавилов, В.Н.Дубровский, А.А.Егоров. – М.: Издательство МЦНМО, 2014. – 160 с. (Библиотечка «Квант». Вып. 132. Приложение к журналу «Квант» №4/2014.)

ISBN 978-5-4439-0623-2

Книга представляет собой сборник статей математиков – учеников Андрея Николаевича Колмогорова, работавших вместе с ним как в области науки, так и в области образования. Тематика статей сборника разнообразна, в них обсуждаются и базовые математические понятия, и более узкие вопросы, подходящие для рассмотрения на занятиях математического кружка.

Книга предназначена для старшеклассников, учителей, руководителей математических кружков, а также для всех, кто интересуется математикой.

ББК 22.1-22.3ж 74.04(2)

ISBN 978-5-4439-0623-2



9 785443 906232 >

6+

## ОТ СОСТАВИТЕЛЕЙ

---

Шестидесятые годы прошлого века ознаменованы бурным развитием математического просвещения в СССР. Во многих городах появились математические классы и школы с математическим уклоном, начали работать математические кружки при университетах и других вузах. Возникла огромная армия энтузиастов – студентов, аспирантов, учителей, преподавателей вузов и ученых, бескорыстно работавших со школьниками. Во всех областях и крупных городах проводились математические олимпиады. В 1961 году прошла первая Всероссийская математическая олимпиада, превратившаяся в 1967 году во Всесоюзную олимпиаду.

Наконец, по предложению крупнейших математиков и физиков – академиков А.П.Александрова, А.Н.Колмогорова, М.А.Лаврентьева, И.К.Кикоина, И.Г.Петровского – было издано постановление правительства СССР об образовании физико-математических школ-интернатов при Московском, Ленинградском, Новосибирском и Киевском государственных университетах.

Весной 1963 года началась подготовка к созданию физико-математической школы (ФМШ) при МГУ. Победители заключительного тура Всероссийской математической олимпиады были приглашены в летнюю школу. Руководил работой летней школы Андрей Николаевич Колмогоров – величайший ученый XX века. После трех недель активных занятий в летней школе 19 лучших ее учеников были зачислены в ФМШ.

Осенью того же года в 49 областях центра России и Белоруссии прошли вступительные экзамены в ФМШ по математике и физике. Для их проведения в областные центры были направлены экзаменационные комиссии, составленные из преподавателей и аспирантов МГУ и МФТИ. Экзаменаторы должны были проверять не формальные умения и навыки школьников, а сообразительность, смекалку, умение решать простые, но нестандартные задачи.

Наконец, после зачисления школьников, рекомендованных экзаменационными комиссиями, в декабре 1963 года начала работу физико-математическая школа-интернат №18 при МГУ – так школа называлась до 1988 года. Ныне это Специализирован-

ный учебно-научный центр (СУНЦ) – физико-математическая школа-интернат имени А.Н.Колмогорова при МГУ или, как ее называют ученики и преподаватели, Колмогоровская ФМШ.

В первые десятилетия существования школы, пока позволяло здоровье, А.Н.Колмогоров руководил ее работой, читал лекции, посещал семинары и другие занятия, устраивал музыкальные вечера, водил школьников и молодых преподавателей в походы, принимал участие в работе летних школ.

В ФМШ работали многие выдающиеся педагоги и ученые. За 50 лет ее закончили более десяти тысяч учеников. Среди выпускников два академика РАН, десять членов-корреспондентов РАН, около 300 докторов наук и более 1000 кандидатов наук.

Часто выпускники становились преподавателями школы, иногда еще до окончания обучения в университете или аспирантуре.

Сейчас Колмогоровская ФМШ – подразделение МГУ на правах факультета. Обучение устроено так же, как и в вузах: лекции, семинары, экзамены. Ученики успешно выступают на математических и физических олимпиадах самого высокого уровня – от Московских и Всероссийских до Международных. Некоторые еще на школьной скамье делают свои первые научные работы.

Выпускники и преподаватели школы – авторы многих научных монографий, учебников для вузов и средней школы, научно-популярных книг и статей во всех существующих в нашей стране научных и научно-популярных журналах. Только в «Кванте» ими опубликовано больше 200 статей. Многие сотрудники Колмогоровской ФМШ в течение сорока лет тесно взаимодействовали с журналом – работали в редакции, состояли и состоят в его редакционной коллегии.

Настоящий сборник (состоящий из двух частей) знакомит вас с очень небольшой частью статей по математике, написанных преподавателями и выпускниками Колмогоровской ФМШ. Статьи содержат сюжеты из самых разных областей математики. Для их понимания, как правило, достаточно познаний в рамках программы средней школы.

Надеемся, что сборник заинтересует и школьников, и учителей, а также просто любителей красивой математики.

# ЧИСЛА ПИЗО

А.Егоров

В этой статье пойдет речь о числах с одним удивительным свойством – их степени «почти целые». Это такие числа  $\alpha > 1$ , для которых расстояние от  $\alpha^n$  ( $n$  – натуральное число) до ближайшего целого числа стремится к нулю. Мы постараемся понять, почему это происходит. Но сначала рассмотрим примеры.

## Несколько примеров

В 50–60-х годах прошлого века на математических олимпиадах различных уровней и в разных странах были популярны задачи, сходные со следующей.

**Задача 1.** Найдите первые  $n$  знаков после запятой в десятичной записи числа  $(5 + \sqrt{26})^n$ .

Задача эта удивительно просто решается. Пусть  $\alpha = 5 + \sqrt{26}$ , а  $\beta = 5 - \sqrt{26}$  – сопряженное с  $\alpha$  число.

Слово «сопряженное» в данном случае означает, что  $\alpha$  и  $\beta$  – корни уравнения  $x^2 - 10x - 1 = 0$  с целыми коэффициентами.

Заметим, что если

$$\alpha^n = a_n + b_n\sqrt{26}, \text{ то } \beta^n = a_n - b_n\sqrt{26},$$

где  $a_n$  и  $b_n$  – натуральные числа. Но тогда

$$\alpha^n + \beta^n = 2a_n = A_n$$

– натуральное число. При этом

$$A_n - \alpha^n = \beta^n$$

и

$$|A_n - \alpha^n| = |\beta^n| = \frac{1}{(5 + \sqrt{26})^n} < \frac{1}{10^n}.$$

Таким образом, разность между  $\alpha^n$  и  $A_n$  отличается от нуля меньше чем на  $\frac{1}{10^n}$ . Поэтому ( $\beta < 0$ ) при четном  $n$  первые  $n$  цифр после запятой – девятки, а при нечетном  $n$  – нули. Задача решена.

---

Опубликовано в «Кванте» №5, 6 за 2005 г.

Пока запомним, что  $\alpha$  и  $\beta$  – иррациональные корни квадратного уравнения с целым коэффициентами, причем  $\alpha > 1$ , а  $|\beta| < 1$ .

### Упражнения

1. Существует ли такое  $n$ , что в десятичной записи числа  $(2 + \sqrt{2})^n$  окажется 1000 одинаковых цифр после запятой? Какие это цифры?
2. Те же вопросы для чисел  $(8 + \sqrt{65})^n$ .
3. Докажите, что при  $n \geq 5k$  расстояние от числа  $\left(\frac{1+\sqrt{5}}{2}\right)^n$  до ближайшего к нему целого числа меньше, чем  $\frac{1}{10^k}$ .

Введем теперь некоторые обозначения. Для данного числа  $a$  будем обозначать через  $(a)$  ближайшее к  $a$  целое число, а через  $\{\{a\}\}$  – расстояние от  $a$  до  $(a)$ , т.е.  $\{\{a\}\} = |(a) - a|$ .

**Определение 1.** Будем говорить, что число  $\alpha$  обладает свойством Пизо, если расстояние от  $\alpha^n$  до ближайшего целого числа стремится к нулю, т.е. если  $\{\{\alpha^n\}\} \rightarrow 0$  при  $n \rightarrow \infty$ .

Вот еще один пример. Рассмотрим теперь кубический многочлен

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

Поскольку  $p(3) < 0$ ,  $p(4) > 0$ , многочлен  $p$  имеет корень на промежутке  $(3; 4)$ . Аналогично, так как  $p(1) < 0$  и  $p(0) > 0$ , у него есть корень на интервале  $(0; 1)$ , а так как  $p(-1) < 0$ , то и на интервале  $(-1; 0)$  также имеется корень этого многочлена. Итак, многочлен  $p$  имеет в точности 3 корня  $\gamma \in (-1; 0)$ ,  $\beta \in (0; 1)$ ,  $\alpha \in (3; 4)$ .

**Упражнение 4.** Докажите, что числа  $\alpha, \beta, \gamma$  иррациональные.

**Задача 2.** Докажите, что число  $\alpha$  обладает свойством Пизо.

Рассмотрим сумму

$$S_n = \alpha^n + \beta^n + \gamma^n.$$

Докажем, что  $S_n$  – целое число при любом  $n$ . Для этого сначала вычислим  $S_1$ ,  $S_2$  и  $S_3$ .

По теореме Виета для кубических многочленов

$$\alpha + \beta + \gamma = 3, \quad \alpha\beta + \alpha\gamma + \beta\gamma = -2, \quad \alpha\beta\gamma = -1,$$

поэтому

$$S_2 = \alpha^2 + \beta^2 + \gamma^2 = (\alpha + \beta + \gamma)^2 - 2(\alpha\beta + \alpha\gamma + \beta\gamma) = 13.$$

Вычислить  $S_3$  несколько сложнее. Для этого воспользуемся известной формулой

$$x^3 + y^3 + z^3 - 3xyz = (x + y + z)(x^2 + y^2 + z^2 - xy - xz - yz).$$

Подставляя вместо  $x$ ,  $y$  и  $z$  числа  $\alpha$ ,  $\beta$ ,  $\gamma$ , получим, что

$$S_3 = \alpha^3 + \beta^3 + \gamma^3 = 3(13 + 2) + 3 = 48.$$

Запишем очевидные равенства

$$\alpha^{n+3} - 3\alpha^{n+2} - 2\alpha^{n+1} + \alpha^n = 0,$$

$$\beta^{n+3} - 3\beta^{n+2} - 2\beta^{n+1} + \beta^n = 0,$$

$$\gamma^{n+3} - 3\gamma^{n+2} - 2\gamma^{n+1} + \gamma^n = 0.$$

Сложив их, получим

$$S_{n+3} - 3S_{n+2} - 2S_{n+1} + S_n = 0. \quad (1)$$

Далее рассуждаем по индукции. При  $n = 1$  из того, что  $S_1$ ,  $S_2$ ,  $S_3$  – целые числа, следует, что и  $S_4$  – целое число.

Пусть  $S_{n+2}$ ,  $S_{n+1}$  и  $S_n$  – целые числа. Но тогда из (1) следует, что  $S_{n+3}$  тоже целое.

Итак,  $S_n$  целое при любом  $n$ . Из неравенства

$$|S_n - \alpha^n| = |\beta^n + \gamma^n| \leq |\beta|^n + |\gamma|^n$$

и того, что  $|\beta| < 1$  и  $|\gamma| < 1$ , вытекает, что

$$|S_n - \alpha^n| \rightarrow 0.$$

Но ведь это и есть свойство Пизо.

**Упражнение 5.** Найдите первые 100 знаков после запятой у числа  $\alpha^{2000}$ ,  $\alpha^{2001}$ .

Как видите, мы довольно легко управились с поведением чисел  $\alpha^n$  для кубического многочлена, все три корня которого действительны, причем один из них больше 1, а два оставшихся меньше единицы по модулю. Правда, пока мы еще никак явно не пользовались иррациональностью этих корней.

А что делать с кубическим многочленом, который имеет лишь один действительный корень? Рассмотрим соответствующий пример.

**Задача 3.** Пусть  $\alpha$  – действительный корень многочлена  $p(x) = x^3 - x - 1$ . Докажите, что число  $\alpha$  обладает свойством Пизо.

**Решение.** Исследование с помощью производной функции  $y = x^3 - x - 1$  показывает, что этот многочлен имеет в точности один действительный корень  $\alpha$ , причем  $1 < \alpha < 2$ . Пусть  $\beta$  и  $\gamma$

– комплексные корни этого многочлена. Можно доказать, что если  $\beta = a + bi$ , где  $a$  и  $b$  действительные, то  $\gamma = a - bi$ , так что  $|\beta| = |\gamma| = \sqrt{a^2 + b^2}$ . По теореме Виета

$$\begin{cases} \alpha + \beta + \gamma = 0, \\ \alpha\beta + \alpha\gamma + \beta\gamma = 0, \\ \alpha\beta\gamma = 1. \end{cases}$$

Из последнего равенства следует, что

$$|\beta|^2 = |\gamma|^2 = |\beta\gamma| = \frac{1}{\alpha}.$$

Поскольку свойства действий над комплексными числами такие же, как и над действительными, дословно повторяя рассуждения, проведенные при решении задачи 2, получим, что и здесь

$$S_n = \alpha^n + \beta^n + \gamma^n$$

– целое число. При этом

$$|S_n - \alpha^n| \leq |\beta|^n + |\gamma|^n < \frac{2}{(\sqrt{\alpha})^n}.$$

Но  $\sqrt{\alpha} > 1$ , так что разность  $S_n - \alpha^n$  стремится к 0 и  $\alpha$  обладает свойством Пизо.

**Упражнение 6.** Докажите, что первые 1000 знаков после запятой в десятичной записи числа  $\alpha^{25000}$  из задачи 3 одинаковы.

Во всех рассмотренных нами примерах мы имели дело с иррациональными корнями квадратичных или кубических многочленов с целыми коэффициентами и старшим коэффициентом, равным 1. При этом все остальные корни этих многочленов были по модулю меньше 1. Не в этом ли причина явления, названного нами свойством Пизо?

### Алгебраические числа и неприводимые многочлены

Здесь мы познакомимся с простейшими понятиями теории алгебраических чисел и неприводимых многочленов.

**Определение 2.** Число  $\alpha$  (действительное или комплексное) называется алгебраическим, если существует многочлен

$$p(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$$

с целыми коэффициентами  $a_0, a_1, \dots, a_n$  такой, что  $p(\alpha) = 0$ .

В этом определении мы можем считать, во-первых, что  $a_0 > 0$  и, во-вторых, что числа  $a_0, a_1, \dots, a_n$  в совокупности взаимно просты.

Таким образом, алгебраические числа – это корни многочленов с целыми коэффициентами.

Ясно, что существует многочлен наименьшей степени, имеющий данное алгебраическое число  $\alpha$  своим корнем. Такой многочлен называется *минимальным многочленом* числа  $\alpha$ , а его степень – *степенью* числа  $\alpha$ .

Рассмотренные нами ранее числа были алгебраическими числами второй и третьей степени.

**Определение 3.** Говорят, что многочлен  $p(x)$  с целыми коэффициентами неприводим, если он не делится ни на какой многочлен с целыми коэффициентами, имеющий меньшую ненулевую степень.

Во множестве  $\mathbb{Z}[x]$  всех многочленов с целыми коэффициентами неприводимые многочлены играют ту же роль, что и простые числа во множестве всех целых чисел. В частности, всякий многочлен из  $\mathbb{Z}[x]$  может быть разложен в произведение неприводимых множителей и это разложение с точностью до постоянных сомножителей единственно.

**Определение 4.** Многочлен с целыми коэффициентами

$$p(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$$

называется *примитивным*, если его коэффициенты в совокупности взаимно просты.

Важное свойство примитивных многочленов выражает следующая лемма.

**Лемма Гаусса.** Произведение двух примитивных многочленов – примитивный многочлен.

Лемма Гаусса доказывается вполне элементарно. Мы предлагаем вам доказать ее самостоятельно.

**Упражнение 7.** Проделайте это.

Из леммы Гаусса вытекает важное для нас в дальнейшем

**Утверждение 1.** Если многочлен  $p(x) \in \mathbb{Z}[x]$  делится на примитивный неприводимый многочлен, то частное – многочлен с целыми коэффициентами.

Это значит, что если  $p(x) = h(x)q(x)$ , где многочлен  $q(x)$  примитивен и неприводим, то  $h(x)$  – многочлен с целыми коэффициентами.

Важным для нас будет и еще одно свойство.

**Утверждение 2.** Если два многочлена  $f(x) \in \mathbb{Z}[x]$  и  $p(x) \in \mathbb{Z}[x]$  имеют общий корень  $\alpha$ :  $f(\alpha) = 0$ ,  $p(\alpha) = 0$ , а многочлен  $p(x)$  неприводим, то  $f(x)$  делится на  $p(x)$ , т.е.  $f(x) = h(x)p(x)$ , где  $h(x)$  – некоторый многочлен с рациональными коэффициентами. А если  $p(x)$  еще и примитивен, то  $h(x)$  имеет целые коэффициенты.

Верно также следующее утверждение.

**Утверждение 3.** Всякий неприводимый многочлен степени  $n$  имеет в точности  $n$  различных ( действительных или комплексных ) корней.

Доказательства всех перечисленных утверждений вы можете найти в любой книге, где подробно изложена теория многочленов, или же провести их самостоятельно.

И, наконец, приведем очень важное для дальнейшего

**Утверждение 4.** Минимальный многочлен всякого алгебраического числа  $\alpha$  неприводим.

Это утверждение почти очевидно. В самом деле, если минимальный многочлен  $p(x)$  числа  $\alpha$  раскладывается на множители  $p(x) = p_1(x)p_2(x)$ , где  $p_1(x), p_2(x)$  – многочлены меньшей степени, то  $p(\alpha) = p_1(\alpha)p_2(\alpha)$ . И либо  $p_1(\alpha) = 0$ , либо  $p_2(\alpha) = 0$ . Но это противоречит минимальности многочлена  $p(x)$ .

**Определение 5.** Пусть  $\alpha$  – алгебраическое число степени  $r$  и  $p(x) = a_0x^r + \dots + a_r$  – минимальный многочлен числа  $\alpha$ . Пусть  $\alpha = \alpha_1, \alpha_2, \dots, \alpha_r$  – его корни. Числа  $\alpha_2, \dots, \alpha_r$  называются сопряженными с числом  $\alpha$ .

### Теорема Виета

Пусть  $p(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$  ( $n \geq 1$ ) – произвольный многочлен с действительными или комплексными коэффициентами.

По основной теореме алгебры, он имеет действительный или комплексный корень (см., например, статью В.Тихомирова «Теоремы существования и основная теорема алгебры» в «Кванте» №4 за 2005 г.). Из основной теоремы следует, что

$$p(x) = a_0(x - x_1)(x - x_2)\dots(x - x_r),$$

где  $x_1, x_2, \dots, x_r$  – корни многочлена  $p(x)$ . Раскрывая скобки и приравнивая коэффициенты, получим равенства

$$a_1 = -a_0(x_1 + x_2 + \dots + x_r),$$

$$a_2 = a_0(x_1x_2 + x_1x_3 + \dots + x_{r-1}x_r),$$

...

$$a_k = (-1)^k a_0 (x_1 x_2 \dots x_k + x_1 x_2 \dots x_{k-1} x_{k+1} + \dots + x_{r-k+1} \dots x_r),$$

...

$$a_n = (-1)^n a_0 x_1 x_2 \dots x_r.$$

Здесь в  $k$ -й строчке стоят суммы всевозможных произведений по  $k$  сомножителей.

Выписанные соотношения называются теоремой Виета. Они обобщают хорошо известную школьникам теорему Виета для квадратного уравнения и несколько менее известную теорему Виета для кубического уравнения.

И, наконец, сформулируем еще одно утверждение, ранее доказанное нами для многочленов второй и третьей степеней.

**Утверждение 5.** Если  $p(x) = x^r + a_1 x^{r-1} + \dots + a_r$  многочлен с целыми коэффициентами,  $x_1, x_2, \dots, x_r$  – его корни,  $S_n = x_1^n + \dots + x_r^n$ , то  $S_n$  – целое число при любом натуральном  $n$ .

Это утверждение мы также не будем здесь доказывать.

### Целые алгебраические числа

**Определение 6.** Алгебраическое число  $\alpha$  называется целым алгебраическим числом, если оно – корень неприводимого многочлена с целыми коэффициентами и старшим коэффициентом, равным 1.

Как видите, понятие целого алгебраического числа существенно отличается от привычного понятия целого числа. Прежде всего, сами целые числа – это целые алгебраические числа степени 1: они корни многочленов вида  $p(x) = x - n$ , где  $n \in \mathbb{Z}$ .

Целые же алгебраические числа более высоких степеней иррациональны. Корни многочленов из задач 1–3 – целые алгебраические числа степеней 2 и 3. Числа вида  $\sqrt[3]{2}$ ,  $\sqrt[5]{2}$ ,  $\sqrt[4]{3}$  – это тоже целые алгебраические числа. Корни многочленов  $2x^3 - 3x - 1$ ,  $3x^3 + 3x + 1$  не являются целыми алгебраическими.

Пусть  $\alpha$  – целое алгебраическое число степени  $r$ , а числа  $\alpha_2, \alpha_3, \dots, \alpha_r$  – сопряженные с ним числа, т.е. остальные корни минимального многочлена  $p(x)$  числа  $\alpha$ . Из утверждения 4 следует, что

$$\alpha^n + \alpha_2^n + \dots + \alpha_r^n = S_n$$

– целое число при любом  $n$ .

## Числа Пизо

Теперь приступим к главному в этой статье.

**Определение 7.** Действительное целое алгебраическое число  $\alpha > 1$  называется числом Пизо, если все сопряженные с ним числа по модулю меньше 1.

Докажем, что числа Пизо обладают свойством Пизо.

**Теорема 1.** Если  $\alpha$  – число Пизо степени  $r > 1$ , то

$$\lim_{n \rightarrow \infty} \{ \{\alpha^n\} \} = 0 .$$

**Доказательство.** Пусть  $\alpha = \alpha_1, \alpha_2, \dots, \alpha_r$  – корни минимального многочлена  $p(x) = x^r + a_1x^{r-1} + \dots + a_r$ .

Из утверждения 5 следует, что

$$S_n = \alpha^n + \alpha_2^n + \dots + \alpha_r^n$$

– целое число. Рассмотрим разность  $S_n - \alpha^n = \alpha_2^n + \dots + \alpha_r^n$  и оценим ее модуль:

$$|S_n - \alpha^n| = |\alpha_2^n + \dots + \alpha_r^n| \leq |\alpha_2|^n + \dots + |\alpha_r|^n .$$

Поскольку  $|\alpha_i| < 1$  при  $i = 2, \dots, r$ , правая часть последнего неравенства стремится к нулю и потому  $S_n - \alpha^n \rightarrow 0$ . Но это в точности значит, что  $\{ \{\alpha^n\} \} \rightarrow 0$ .

Итак, мы обнаружили целый класс алгебраических чисел, обладающих свойством Пизо – это числа Пизо, т.е. действительные алгебраические числа, большие 1, все сопряженные которых по модулю меньше 1.

В дальнейшем мы увидим, что верна и обратная теорема: если действительное алгебраическое число  $\alpha > 1$  обладает свойством Пизо, то оно является числом Пизо.

Однако сначала выясним, какие *целые* алгебраические числа обладают свойством Пизо.

Пусть  $\alpha > 1$  – целое действительное алгебраическое число, а  $p(x) = x^r + a_1x^{r-1} + \dots + a_r$  – его минимальный многочлен. Предположим, что

$$\{ \{\alpha^n\} \} \rightarrow 0 .$$

Пусть  $A_n$  – ближайшее к  $\alpha^n$  целое число и

$$\{ \{\alpha^n\} \} = |A_n - \alpha^n| \rightarrow 0 .$$

Тогда  $\alpha^n = A_n + \delta_n$ , где  $\delta_n \rightarrow 0$  при  $n \rightarrow \infty$ . Теперь умножим равенство  $\alpha^r + a_1\alpha^{r-1} + \dots + a_r = 0$  на  $\alpha^n$ :

$$\alpha^{n+r} + a_1\alpha^{n+r-1} + \dots + a_r\alpha^n = 0 .$$

Подставим в полученное равенство  $\alpha^{n+r} = A_{n+r} + \delta_{n+r}$ ,  $\alpha^{n+r-1} = A_{n+r-1} + \delta_{n+r-1}, \dots, \alpha^n = A_n + \delta_n$ . Тогда

$$A_{n+r} + a_1 A_{n+r-1} + \dots + a_r A_n + \gamma_n = 0, \quad (2)$$

где  $\gamma_{n+r} = \delta_{n+r} + a_1 \delta_{n+r-1} + \dots + a_r \delta_n$ . Ясно, что  $\gamma_n \rightarrow 0$  при  $n \rightarrow \infty$ . Но число  $A_{n+r} + a_1 A_{n+r-1} + \dots + a_r A_n$  целое и потому при всех достаточно больших  $n$

$$A_{n+r} + a_1 A_{n+r-1} + \dots + a_r A_n = 0. \quad (3)$$

В самом деле, из того, что  $\gamma_n \rightarrow 0$ , следует, что существует такое  $N$ , что  $|\gamma_n| < 1$  при всех  $n \geq N$ . Но при таких  $n$  равенство (2) возможно только при выполнении равенства (3).

Итак, числа  $A_n$  при  $n \geq N$  удовлетворяют рекуррентному уравнению (3). С этого момента линейные рекурренты становятся основным аппаратом наших исследований. Но сначала выясним, как ведут себя степени рациональных чисел.

### Рациональные числа

Если  $\alpha = \frac{p}{q} > 1$  – несократимая дробь ( $q > 1$ ), то  $\alpha$  – корень уравнения первой степени  $qx - p = 0$ .

Докажем, что  $\{\alpha^n\}$  не стремится к нулю. Предположим, что  $\lim_{n \rightarrow \infty} \{\alpha^n\} = 0$ . Пусть  $(\alpha^n) = A_n$ , так что  $\alpha^n = A_n + \delta_n$  и  $\delta_n \rightarrow 0$  при  $n \rightarrow \infty$ ,  $|\delta_n| = \{\alpha^n\}$ . Но  $q\alpha - p = 0$ , поэтому  $q\alpha^{n+1} - p\alpha^n = 0$ , т.е.  $q(A_{n+1} + \delta_{n+1}) - p(A_n + \delta_n) = qA_{n+1} - pA_n + \gamma_n = 0$ , где  $\gamma_n = q\delta_{n+1} - p\delta_n$ , так что  $\lim_{n \rightarrow \infty} \gamma_n = 0$ . Но число  $qA_{n+1} - pA_n$  целое. Поэтому при достаточно больших  $n$  оно равно нулю, т.е.

$$A_{n+1} = \frac{p}{q} A_n \text{ при } n \geq N.$$

Но тогда

$$A_{N+k} = \left(\frac{p}{q}\right)^k A_N$$

при всех  $k > 1$ . Это невозможно, так как  $A_N$  не делится на  $q^k$  при достаточно больших  $k$ , т.е.  $A_{N+k}$  не целое.

Итак, не целые рациональные числа свойством Пизо не обладают.

В то же время обычные целые положительные числа, очевидно, являются числами Пизо.

Таким образом, все рациональные числа со свойством Пизо – целые.

**Упражнение 7.** Существует ли предел  $\lim_{n \rightarrow \infty} \left\{ \left\{ \left( \frac{3}{2} \right)^n \right\} \right\}$ ?

## Квадратичные иррациональности

Рассмотрим квадратное уравнение

$$x^2 + px + q = 0$$

с целыми коэффициентами и иррациональными корнями  $\alpha > 1$  и  $\beta$ .

Если  $|\beta| < 1$ , то  $\alpha$  – число Пизо и, как нам уже известно,  $\{\{\alpha^n\}\} \rightarrow 0$  (это, напомним, следует из того, что  $S_n = \alpha^n + \beta^n$  целое при любом  $n$ ).

Пусть  $\{\{\alpha^n\}\} \rightarrow 0$ . Докажем, что  $|\beta| < 1$ . Запишем для нашего случая уравнение (3). Если  $A_n = (\alpha^n)$ , то при всех достаточно больших  $n$  будет

$$A_{n+2} + pA_{n+1} + qA_n = 0.$$

Теперь нам предстоит немного отвлечься от основной темы и поговорить о линейных рекуррентных уравнениях.

## Линейные рекурренты

**Определение 8.** Последовательность  $x_0, x_1, \dots, x_n, \dots$  называется линейной рекуррентой (ЛР) порядка  $k$ , если существуют такие числа  $a_0 \neq 0, a_1, \dots, a_k \neq 0$ , что при любом натуральном  $n$

$$a_0 x_{n+k} + a_1 x_{n+k-1} + \dots + a_k x_n = 0. \quad (4)$$

Иногда линейные рекурренты называются также возвратными последовательностями. Числа  $A_n$ , появившиеся выше, образуют линейную рекурренту порядка  $r$ . Также являются ЛР известные многим нашим читателям числа Фибоначчи, т.е. последовательность  $f_0, f_1, \dots, f_n, \dots$ , для которой  $f_0 = 1, f_1 = 1, f_{n+2} = f_{n+1} + f_n$ .

Любая геометрическая прогрессия  $u_n$  образует ЛР порядка 1:  $u_{n+1} = qu_n$ .

Всякая арифметическая прогрессия  $a_n$  является ЛР порядка 2, ибо для нее  $a_{n+2} - a_{n+1} = a_{n+1} - a_n$ , т.е.  $a_{n+2} - 2a_{n+1} + a_n = 0$ .

Заметим, что если две последовательности  $u_n$  и  $v_n$  удовлетворяют уравнению (4), то все последовательности вида  $w_n = \lambda u_n + \mu v_n$  также удовлетворяют этому уравнению.

В самом деле,

$$\begin{aligned} a_0 w_{n+k} + a_1 w_{n+k-1} + \dots + a_k w_n &= \\ = a_0 (\lambda u_{n+k} + \mu v_{n+k}) + a_1 (\lambda u_{n+k-1} + \mu v_{n+k-1}) + \dots + a_k (\lambda u_n + \mu v_n) &= \\ = \lambda (a_0 u_{n+k} + \dots + a_k u_n) + \mu (a_0 v_{n+k} + \dots + a_k v_n) &= 0. \end{aligned}$$

Будем искать решения уравнения (4) в виде комбинации простейших решений. А поскольку простейшей линейной рекуррентой является геометрическая прогрессия, сначала будем искать решения в виде  $x_n = \lambda^n$ , где  $\lambda \neq 0$ .

Подставляя  $\lambda^n$  вместо  $x_m$  в (4), получаем

$$a_0 \lambda^{n+k} + \dots + a_k \lambda^n = 0.$$

После сокращения на  $\lambda^n$  приходим к уравнению

$$a_0 \lambda^k + \dots + a_k = 0,$$

которое называется *характеристическим уравнением* линейной рекурренты.

Сделаем некоторые упрощающие дело предположения. Именно, будем считать, что характеристическое уравнение имеет  $k$  различных ненулевых корней (этот случай и будет нужен в дальнейшем). Ясно, что все прогрессии вида  $\lambda_i^n$  ( $i = 1, 2, \dots, k$ ) – решения ЛР.

Докажем теперь, что всякое решение ЛР можно «скомбинировать» из прогрессий.

**Утверждение 6.** *Всякое решение ЛР (4) представляется в виде*

$$x_n = c_1 \lambda_1^n + c_2 \lambda_2^n + \dots + c_k \lambda_k^n.$$

Прежде всего заметим, что ЛР однозначно определяется заданием первых  $k$  членов последовательности:  $x_0, x_1, \dots, x_{k-1}$ .

Будем искать решение в виде

$$x_n = c_1 \lambda_1^n + c_2 \lambda_2^n + \dots + c_k \lambda_k^n.$$

Подставляя сюда  $n = 0, 1, \dots, k-1$ , получим систему уравнений

$$\begin{cases} c_1 + c_2 + \dots + c_k = x_0, \\ c_1 \lambda_1 + c_2 \lambda_2 + \dots + c_k \lambda_k = x_1, \\ \dots \\ c_1 \lambda_1^{k-1} + c_2 \lambda_2^{k-1} + \dots + c_k \lambda_k^{k-1} = x_{k-1}. \end{cases}$$

Вспомним, что все числа  $\lambda_1, \lambda_2, \dots, \lambda_k$  различны, и рассмотрим

многочлен

$$f_i(\lambda) = \frac{f(\lambda)}{\lambda - \lambda_i} = a_0 (\lambda - \lambda_i) \dots (\lambda - \lambda_{i-1})(\lambda - \lambda_{i+1}) \dots (\lambda - \lambda_k) = \\ = a_0 \lambda^{k-1} + \tilde{a}_1 \lambda^{k-1} + \tilde{a}_2 \lambda^{k-2} + \dots + \tilde{a}_{k-1,i}.$$

Корни этого многочлена – это все числа  $\lambda_1, \lambda_2, \dots, \lambda_k$ , кроме  $\lambda_i$ . Умножив первое уравнение на  $\tilde{a}_{k-1,i}$ , второе – на  $\tilde{a}_{k-2,i}$ , ..., последнее на  $a_0$  и сложив полученные уравнения, придем к равенству

$$c_1 f_i(\lambda_1) + c_2 f_i(\lambda_2) + \dots + c_k f_i(\lambda_k) = a_{k-1,i} x_0 + \dots + a_0 x_{k-1} = M_i.$$

Но  $f_i(\lambda_j) = 0$  при  $i \neq j$ , а  $f_i(\lambda_i) \neq 0$ . Поэтому

$$c_i = \frac{M_i}{f_i(\lambda_i)}.$$

Тем самым утверждение 6 доказано.

Найдем, например, формулу для  $n$ -го числа Фибоначчи. Характеристическое уравнение имеет вид

$$\lambda^2 - \lambda - 1 = 0.$$

Его корни  $\lambda_1 = \frac{1+\sqrt{5}}{2}$ ,  $\lambda_2 = \frac{1-\sqrt{5}}{2}$ . Ищем решение в виде

$$f_n = c_1 \lambda_1^n + c_2 \lambda_2^n.$$

При  $n = 0$  и  $n = 1$  имеем

$$\begin{cases} c_1 + c_2 = 0, \\ c_1 \lambda_1 + c_2 \lambda_2 = 1, \end{cases} \text{ откуда } c_1 = \frac{1}{\sqrt{5}}, \quad c_2 = -\frac{1}{\sqrt{5}}.$$

Мы получили формулу Бине:

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

Любопытно, что в записи целых чисел Фибоначчи участвуют иррациональности.

Вот еще один пример ЛР:

$$x_0 = 0, \quad x_1 = 1, \quad x_{n+2} = 2x_{n+1} - 3x_n.$$

Характеристическое уравнение  $\lambda^2 - 2\lambda + 3 = 0$  имеет комплексные корни:  $\lambda_1 = 1 + i\sqrt{2}$  и  $\lambda_2 = 1 - i\sqrt{2}$ , а последовательность, состоящая из целых чисел, записывается так:

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

т.е. в записи целочисленной последовательности участвуют даже мнимые числа.

### Квадратичные иррациональности (завершение)

Вернемся к рассмотрению квадратичных иррациональностей.

Мы видели, что числа  $A_n$  при  $n \geq N$  образуют ЛР:

$$A_{n+2} + pA_{n+1} + qA_n = 0.$$

Это значит, что

$$A_n = c_1\alpha^n + c_2\beta^n,$$

где  $c_1, c_2$  – некоторые действительные числа, причем  $c_1 \neq 0$  и  $c_2 \neq 0$ . (Если, например,  $c_2 = 0$ , то  $A_{n+1}/A_n = \alpha$ , что противоречит иррациональности числа  $\alpha$ .)

Рассмотрим разность

$$A_n - \alpha^n = (c_1 - 1)\alpha^n + c_2\beta^n. \quad (5)$$

Нам известно, что  $A_n - \alpha^n \rightarrow 0$  при  $n \rightarrow \infty$ . Поэтому и  $(c_1 - 1)\alpha^n + c_2\beta^n \rightarrow 0$  при  $n \rightarrow \infty$ . Теперь докажем лемму из элементарного курса анализа.

**Лемма.** *Если числа  $c_1, c_2, \dots, c_r$  не равны нулю, числа  $\lambda_1, \lambda_2, \dots, \lambda_r$ , попарно различные, не равны нулю и все отличны от 1, а последовательность*

$$u_n = c_1\lambda_1^n + c_2\lambda_2^n + \dots + c_r\lambda_r^n$$

*имеет предел, то  $|\lambda_i| < 1$  при всех  $i = 1, 2, \dots, r$  и  $\lim_{n \rightarrow \infty} u_n = 0$ .*

**Доказательство.** Проведем индукцию по  $r$ . При  $r = 1$  утверждение леммы очевидно. Пусть лемма справедлива для  $r = k$ , и

$$u_n = c_1\lambda_1^n + c_2\lambda_2^n + \dots + c_{k+1}\lambda_{k+1}^n.$$

Рассмотрим последовательность

$$w_n = u_{n+1} - \lambda u_n = c_2(\lambda_2 - \lambda_1)\lambda_2^n + \dots + c_{k+1}(\lambda_{k+1} - \lambda_1)\lambda_{k+1}^n.$$

Из существования предела последовательности  $u_n$  следует существование предела последовательности  $w_n$ , а из предположения индукции – справедливость леммы.

Из леммы сразу следует, что в нашем случае  $c_1 = 1$ , а  $|\beta| < 1$ . (Если  $c_1 \neq 1$ , то из существования предела последовательности  $(c_1 - 1)\alpha^n + c_2\beta^n$  следовало бы, что  $|\alpha| < 1$ .)

Итак,  $A_n = \alpha^n + c_2\beta^n$ . Кроме того, разность  $A_n - S_n = (c_2 - 1)\beta^n$  тоже стремится к нулю (поскольку  $|\beta| < 1$ ) и потому  $A_n = S_n$  при достаточно больших  $n$  и  $c_2 = 1$ , т.е.  $A_n = S_n$ .

Мы доказали, что если иррациональный корень уравнения  $x^2 + px + q = 0$  обладает свойством Пизо, то второй корень  $\beta$  по модулю меньше 1, т.е.  $\alpha$  – число Пизо.

Итак, целое алгебраическое число второго порядка обладает свойством Пизо тогда и только тогда, когда оно является числом Пизо.

### Кубические иррациональности

Пусть  $\alpha > 1$  – целое алгебраическое число третьего порядка, т.е. корень уравнения

$$p(x) = x^3 + ax^2 + bx + c,$$

где  $a, b, c$  – целые числа, многочлен  $p(x)$  неприводим, а  $\beta$  и  $\gamma$  – числа (действительные или комплексные), сопряженные с  $\alpha$ .

Предположим, что  $\{\{\alpha^n\}\} \rightarrow 0$ , и докажем, что  $\alpha$  – число Пизо. Для этого мы должны доказать, что  $|\beta| < 1$ ,  $|\gamma| < 1$ .

Числа  $A_n$  в этом случае образуют ЛР, для которой

$$A_{n+3} + aA_{n+2} + bA_{n+1} + cA_n = 0$$

при достаточно больших  $n$ . Как мы видели,

$$A_n = c_1\alpha^n + c_2\beta^n + c_3\gamma^n$$

для некоторых  $c_1, c_2, c_3$ .

Покажем, что ни одно из чисел  $c_1, c_2, c_3$  не равно нулю. Предположим, например, что  $c_3 = 0$ . Тогда

$$\begin{aligned} A_n &= c_1\alpha^n + c_2\beta^n, \\ A_{n+1} &= c_1\alpha^{n+1} + c_2\beta^{n+1}, \\ A_{n+2} &= c_1\alpha^{n+2} + c_2\beta^{n+2}. \end{aligned} \tag{6}$$

Пусть  $q(x) = (x - \alpha)(x - \beta) = x^2 + kx + l$  – квадратичный многочлен с корнями  $\alpha$  и  $\beta$ .

Умножим первое из равенств (6) на  $l$ , второе на  $k$  и сложим все три равенства:

$$A_{n+2} + kA_{n+1} + lA_n = c_1\alpha^n q(\alpha) + c_2\beta^n q(\beta) = 0,$$

т.е. целые числа  $A_n$  образуют линейную рекурренту второго порядка.

Докажем, что числа  $k$  и  $l$  рациональны. В самом деле, так как

$$\begin{cases} kA_{n+1} + lA_n = -A_{n+2}, \\ kA_{n+2} + lA_{n+1} = -A_{n+3}, \end{cases}$$

пара  $(k, l)$  – решение системы из двух линейных уравнений с двумя неизвестными.

Вычислим определитель этой системы:

$$\begin{vmatrix} A_{n+1} & A_n \\ A_{n+2} & A_{n+1} \end{vmatrix} = A_{n+1}^2 - A_n A_{n+2} = -c_1 c_2 \alpha^n \beta^n (\alpha - \beta)^2 \neq 0.$$

Итак, решение системы единствено, а числа  $k$  и  $l$ , очевидно, рациональны.

Число  $\alpha$  при этом оказывается корнем квадратного трехчлена  $Mq(x)$  с целыми коэффициентами ( $M$  – общий знаменатель дробей  $k$  и  $l$ ). Но это противоречит неприводимости многочлена  $p(x)$ . Здесь у нас впервые явно «сработала» неприводимость. Аналогично доказывается, что  $c_1 \neq 0$ ,  $c_2 \neq 0$ .

Рассмотрим теперь разность

$$A_n - \alpha^n = (c_1 - 1)\alpha^n + c_2\beta^n + c_3\gamma^n.$$

Поскольку она стремится к нулю, а  $\alpha > 1$ , из леммы предыдущего пункта следует, что  $c_1 = 1$ ,  $|\beta| < 1$  и  $|\gamma| < 1$ . Итак,  $\alpha$  – число Пизо.

Подведем итоги. Мы доказали, что если целое алгебраическое число второго или третьего порядка обладает свойством Пизо, то оно является числом Пизо.

Продолжим изучение чисел Пизо. Мы доказали, что целые алгебраические числа второй и третьей степеней, обладающие свойством Пизо, являются числами Пизо. Аналогичное утверждение, естественно, справедливо и для произвольных степеней. Доказательством этого мы сейчас и займемся.

### Иrrациональности степени $r > 3$

Итак, пусть  $\alpha > 1$  – целое алгебраическое число степени  $r > 3$  и  $p(x) = x^r + a_1x^{r-1} + \dots + a_r$  – минимальный многочлен числа  $\alpha$ . Предположим, что  $\{\{\alpha^n\}\} \rightarrow 0$ . Оказывается, что справедлива следующая теорема.

**Теорема 2.** При выполнении перечисленных выше условий  $\alpha$  – число Пизо.

**Доказательство.** Мы должны доказать, что все числа  $\alpha_2, \dots, \alpha_r$ , сопряженные с  $\alpha$ , по модулю меньше 1.

Вспомним, что если  $A_n = (\alpha^n)$ , то, начиная с некоторого номера  $N$ , т.е. при  $n \geq N$ , целые положительные числа  $A_n$  образуют линейную рекурренту:

$$A_{n+r} + a_1 A_{n+r-1} + \dots + a_r A_n = 0.$$

Это, в свою очередь, означает, что

$$A_n = c_1\alpha_1^n + \dots + c_r\alpha_r^n,$$

где  $\alpha = \alpha_1$ ,  $n \geq N$ , а  $c_1, c_2, \dots, c_r$  – некоторые (вообще говоря, комплексные) числа.

Докажем, что, как и в случае квадратичных и кубических иррациональностей, числа  $c_1, c_2, \dots, c_r$  – не нули.

Предположим, например, что  $c_1, \dots, c_k$  – не нули, а  $c_{k+1} = \dots = c_r = 0$ . Тогда при  $n \geq N$

$$\begin{cases} A_n = c_1\alpha_1^n + \dots + c_k\alpha_k^n, \\ \dots \\ A_{n+k} = c_1\alpha_1^{n+k} + \dots + c_k\alpha_k^{n+k}. \end{cases}$$

Рассмотрим многочлен

$$q(x) = (x - \alpha_1) \dots (x - \alpha_k) = x^k + q_1x^{k-1} + \dots + q_k.$$

Числа  $\alpha_1, \dots, \alpha_k$  – корни многочлена  $q(x)$ . Умножим первое из уравнений системы на  $q_k$ , второе – на  $q_{k-1}, \dots$ , предпоследнее – на  $q_1$  и сложим все уравнения. Получим равенство

$$\begin{aligned} A_{n+k} + q_1A_{n+k-1} + \dots + q_kA_n &= \\ &= c_1\alpha_1^n q(\alpha_1) + c_2\alpha_2^n q(\alpha_2) + \dots + c_k\alpha_k^n q(\alpha_k). \end{aligned}$$

Но  $q(\alpha_1) = q(\alpha_2) = \dots = q(\alpha_k) = 0$ , так что

$$A_{n+k} + q_1A_{n+k-1} + \dots + q_kA_n = 0.$$

Следовательно, целые числа  $A_n$  образуют линейную рекурренту порядка  $k < r$ .

Докажем, что все коэффициенты  $q_1, \dots, q_k$  – рациональные числа. Ранее мы уже доказали аналогичное утверждение для случая  $r = 3$ . Будем действовать сходным образом.

Рассмотрим систему из  $k$  уравнений относительно  $q_1, \dots, q_k$ :

$$\begin{cases} A_{n+k} + q_1A_{n+k-1} + \dots + q_kA_n = 0, \\ A_{n+k} + q_1A_{n+k} + \dots + q_kA_{n+1} = 0, \\ \dots \\ A_{n+2k-1} + q_1A_{n+2k-2} + \dots + q_kA_{n+k-1} = 0. \end{cases}$$

Пользуясь теорией определителей (это единственное место, где мы вынуждены воспользоваться ссылкой на сравнительно неэлементарные понятия), можно доказать, что определитель системы не равен нулю (это следует из того, что числа  $\alpha_1, \alpha_2, \dots, \alpha_k$  попарно различны и не равны нулю). Отсюда и из известного

правила Крамера решения линейных систем следует, что числа  $q_1, \dots, q_k$  рациональны. Впрочем, если представить себе процедуру решения системы методом исключения неизвестных, не трудно понять, что при условии существования и единственности решения числа  $q_1, \dots, q_k$  будут рациональны.

Таким образом, числа  $\alpha_1, \dots, \alpha_k$  – корни многочлена  $q(x)$  с рациональными коэффициентами. Пусть  $M$  – общий наименьший знаменатель дробей  $q_1 = \frac{m_1}{n_1}$ ,  $q_2 = \frac{m_2}{n_2}$ , ...,  $q_k = \frac{m_k}{n_k}$ . Числа  $\alpha_1, \dots, \alpha_k$  являются корнями многочлена  $Q(x) = Mq(x)$  с целыми коэффициентами. Но степень многочлена  $Q(x)$  меньше степени многочлена  $p(x)$ , что противоречит минимальности многочлена  $p(x)$ . Вот где работает неприводимость минимального многочлена!

Итак, ни одно из чисел  $c_1, \dots, c_r$  не равно нулю. Теперь легко завершить доказательство теоремы. Рассмотрим разность

$$\delta_n = A_n - \alpha^n = (c_1 - 1)\alpha_1^n + c_2\alpha_2^n + \dots + c_r\alpha_r^n.$$

Поскольку  $\delta_n$  стремится к нулю,  $\alpha = \alpha_1 > 1$ , а числа  $c_2, \dots, c_r$  – не нули, то по лемме из раздела «Квадратичные иррациональности»

$$c_1 = 1, |\alpha_2| < 1, \dots, |\alpha_r| < 1.$$

Следовательно,  $\alpha$  – число Пизо, и теорема 2 доказана.

Теперь, мы полностью разобрались с целыми алгебраическими числами. Сформулируем окончательный результат.

**Теорема 2'.** Целое действительное алгебраическое число  $\alpha > 1$  обладает свойством Пизо тогда и только тогда, когда оно является числом Пизо, т.е. когда все числа, сопряженные с  $\alpha$ , по модулю меньше 1.

Мы уже видели, что не целые рациональные числа (т.е. алгебраические числа первого порядка) свойством Пизо обладать не могут. А как обстоят дела с не целыми алгебраическими иррациональностями? Например, как ведут себя степени  $\alpha^n$ , где  $\alpha = \frac{1+\sqrt{2}}{2}$  – корень уравнения  $4x^2 - 4x - 1 = 0$ ? Сопряженное с  $\alpha$  число  $\beta = \frac{1-\sqrt{2}}{2}$  по модулю меньше 1, так что  $\alpha$  очень похоже на число Пизо. Однако, как мы увидим дальше, такие числа свойством Пизо не обладают.

## Произвольные алгебраические числа

Пусть  $\alpha$  – не целое алгебраическое число степени  $r$ , т.е. корень неприводимого многочлена

$$p(x) = a_0x^r + a_1x^{r-1} + \dots + a_r,$$

где  $a_0 > 1$ ,  $a_1, \dots, a_r$  – целые числа. Далее мы будем считать, что многочлен  $p(x)$  примитивен, т.е. что его коэффициенты в совокупности взаимно просты.

Предположим, что  $\alpha$  обладает свойством Пизо, т.е.  $\alpha^n = A_n + \delta_n$ , где  $A_n$  – натуральное, а  $\delta_n \rightarrow 0$  при  $n \rightarrow \infty$ . Поскольку  $\alpha$  – корень многочлена  $p(x)$ , так же, как и раньше, получим, что

$$a_0A_{n+r} + a_1A_{n+r-1} + \dots + a_rA_n = 0$$

при всех достаточно больших  $n$ . Таким образом, целые числа  $A_n$  образуют ЛР с целыми коэффициентами, причем «старший коэффициент»  $a_0$  отличен от 1.

Нашей целью станет изучение таких линейных рекуррент, соответствующих *неприводимым* многочленам  $p(x) = a_0x^r + a_1x^{r-1} + \dots + a_r$ .

Сначала мы докажем, что существуют целые числа  $p_1, p_2, \dots, p_m$ , где  $m > r$ , для которых при всех достаточно больших  $n$  выполняются равенства

$$A_{n+m} + p_1A_{n+m-1} + \dots + p_mA_n = 0,$$

т.е. что *целочисленная* ЛР со старшим коэффициентом  $a_0 > 1$  является ЛР большего порядка со старшим коэффициентом 1. А уже отсюда и из неприводимости многочлена  $p(x)$  будет следовать, что характеристический многочлен

$$P(x) = x^m + p_1x^{m-1} + \dots + p_m$$

делится на многочлен  $p(x)$ . Но многочлен  $p(x)$  неприводим и примитивен, и поэтому, в силу утверждения 2,  $P(x) = h(x)p(x)$ , где  $h(x)$  – многочлен с целыми коэффициентами, а это возможно лишь при  $a_0 = 1$ . Таков наш план. К его осуществлению мы и приступаем.

Ограничимся случаем иррациональностей второй степени. Для более высоких степеней рассуждения отличаются лишь большей громоздкостью.

Итак, пусть  $\alpha > 1$  – иррациональный корень квадратного уравнения  $ax^2 + bx + c = 0$ , где  $a > 1$ , целые числа  $a, b, c$  взаимно просты, а  $\beta$  – сопряженное с  $\alpha$  число.

Предположим, что  $\alpha^n = A_n + \delta_n$ , где  $\delta_n \rightarrow 0$ . Тогда

$$aA_{n+2} + bA_{n+1} + cA_n + a\delta_{n+2} + b\delta_{n+1} + c\delta_n = aA_{n+2} + bA_{n+1} + cA_n + \gamma_n = 0.$$

Так как  $\gamma_n \rightarrow 0$  при  $n \rightarrow \infty$ , то при достаточно больших  $n \geq N_0$  числа  $A_n$  образуют линейную рекурренту:

$$aA_{n+2} + bA_{n+1} + cA_n = 0.$$

При этом

$$A_n = c_1\alpha^n + c_2\beta^n,$$

где  $c_1$  и  $c_2$  – отличные от нуля числа (убедитесь в этом).

Теперь на время прервемся, чтобы сформулировать и доказать одно чисто алгебраическое утверждение.

### Алгебраическая лемма

Назовем  $n$ -мерным целочисленным вектором произвольный упорядоченный набор  $\bar{a} = (x_1, x_2, \dots, x_n)$  целых чисел.

**Определение 9.** Пусть  $A$  – некоторое множество целочисленных  $n$ -мерных векторов. Набор векторов  $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_r$  из  $A$  называется базисным, если любой вектор  $\bar{x} \in A$  можно представить в виде  $\bar{x} = \lambda_1\bar{a}_1 + \lambda_2\bar{a}_2 + \dots + \lambda_r\bar{a}_r$ , где  $\lambda_1, \lambda_2, \dots, \lambda_r$  – некоторые целые числа.

**Алгебраическая лемма.** Во всяком множестве  $A$  целочисленных векторов имеется конечный базисный набор.

**Доказательство.** Рассмотрим сначала некоторое множество  $A$  одномерных векторов, т.е. целых чисел. Пусть  $x_1, x_2, \dots, x_n, \dots$  – эти числа. Будем (для определенности) считать, что  $x_1 > 0$ . Рассмотрим остатки от деления чисел  $x_2, x_3, \dots, x_n, \dots$  на  $x_1$ . (Если все числа делятся на  $x_1$ , то уже  $x_1$  есть базисный набор.) Поделим  $x_i$  на  $x_1$  с остатком:

$$x_i = q_i x_1 + r_i, \text{ где } 0 \leq r_i < x_1.$$

Заметим, что если  $x_j$  и  $x_i$  дают одинаковые остатки, то

$$x_j - x_i = (q_j - q_i)x_1,$$

и

$$x_j = (q_j - q_i)x_1 + x_i,$$

т.е.  $x_j$  – линейная комбинация чисел  $x_1$  и  $x_i$ .

Разобьем все множество  $A$  на классы чисел, дающих одинаковые остатки при делении на  $x_1$ . Пусть этих классов имеется  $r$

штук ( $r \leq x_1$ ). В каждом классе выберем какое-нибудь число. Множество выбранных чисел и даст нам, очевидно, требуемый базисный набор. Изменив, если надо, нумерацию чисел, можем считать, что это числа  $x_1, x_2, \dots, x_r$ .

В самом деле, для любого числа  $x_m$  среди чисел  $x_1, x_2, \dots, x_r$  есть число  $x_k$ , сравнимое с ним по модулю  $x_1$ , и тогда

$$x_m - x_k = q_m x_1,$$

т.е.

$$x_m = x_k + q_m x_1.$$

Далее применим индукцию по  $n$ . Предположим, что утверждение уже доказано для  $n$ -мерных векторов и  $A$  – некоторое множество целочисленных  $(n+1)$ -мерных векторов  $(x_1, x_2, \dots, x_n, x_{n+1})$ . Рассмотрим множество  $B$   $n$ -мерных векторов, получаемых из векторов множества  $A$  отбрасыванием последней координаты  $x_{n+1}$ . По предположению индукции, во множестве  $B$  имеется базисный набор векторов  $\bar{b}_1, \dots, \bar{b}_k$ . Пусть  $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$  – соответствующие им векторы из множества  $A$  (набор  $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$  восстанавливается неоднозначно, но это нас не должно беспокоить). Если  $\bar{a} \in A$  – произвольный вектор:  $a = (x_1, \dots, x_n, x_{n+1})$ , то для некоторых целых чисел  $\lambda_1, \lambda_2, \dots, \lambda_k$  будет

$$(x_1, x_2, \dots, x_n) = \sum_{i=1}^k \lambda_i \bar{b}_i,$$

но тогда первые  $n$  координат вектора

$$\bar{a}' = \bar{a} - \sum_{i=1}^k \lambda_i \bar{a}_i$$

равны нулю, т.е.

$$\bar{a}' = (0, 0, \dots, 0, \tilde{x}_{n+1}).$$

Рассмотрим множество  $A'$  всех векторов вида  $\bar{a}'$ . Во множестве  $A'$ , как было доказано раньше, имеется базисный набор векторов. Пусть это будут векторы  $\bar{e}_1', \bar{e}_2', \dots, \bar{e}_l'$ . Вспомним, что при любом  $1 \leq j \leq l$

$$\bar{e}_j' = \bar{e}_j - \sum_{i=1}^k \lambda_{ij} \bar{a}_i,$$

где  $\bar{e}_j$  – некоторый вектор из  $A$ , а  $\lambda_{ij} \in \mathbb{Z}$ ,  $1 \leq i \leq k$ .

Добавим теперь к векторам  $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$  векторы  $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_l$ . Это и будет требуемый базисный набор в  $A$ .

Действительно, пусть  $\bar{a} \in A$  – произвольный вектор. Тогда

$$\bar{a}' = \sum_{j=1}^l \mu_j \bar{e}_j', \text{ где } \mu_j \in \mathbb{Z},$$

т.е.

$$\bar{a} = \sum_{i=1}^k \lambda_i \bar{a}_i + \sum_{j=1}^l \mu_j \bar{e}_j'.$$

Выражая  $\bar{e}_j'$  через  $e_j$  и  $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_k$ , получим (после сокращений) представление  $\bar{a}$  в виде линейной комбинации векторов  $\bar{a}_1, \dots, \bar{a}_k, \bar{e}_1, \dots, \bar{e}_l$ . Лемма доказана.

### Завершающий этап

Рассмотрим последовательность  $A$  двумерных векторов  $\bar{a}_n = (A_n, A_{n+1})$ . По доказанному, в этой последовательности есть базисный набор. Пусть это векторы  $\bar{a}_{n_1}, \bar{a}_{n_2}, \dots, \bar{a}_{n_k}$ . Но тогда существуют целые числа  $p_1, \dots, p_k$  такие, что

$$\bar{a}_{n_k+1} = \sum_{i=1}^k p_i \bar{a}_{n_i}.$$

Обозначив  $n_k + 1$  через  $r$ , получаем равенство

$$(A_r, A_{r+1}) = \sum_{i=1}^k p_i (A_{n_i}, A_{n_i+1}) = \left( \sum_{i=1}^k p_i A_{n_i}, \sum_{i=1}^k p_i A_{n_i+1} \right).$$

Приравнивая координаты левой и правой частей, имеем

$$A_r = \sum_{i=1}^k p_i A_{n_i}, A_{r+1} = \sum_{i=1}^k p_i A_{n_i+1}.$$

Заменив все числа  $A_n$  их выражениями через  $\alpha$  и  $\beta$ :

$$A_n = c_1 \alpha^n + c_2 \beta^n,$$

получаем систему уравнений

$$\begin{cases} c_1 \left( \alpha^r - \sum_{i=1}^k p_i \alpha^{n_i} \right) + c_2 \left( \beta^r - \sum_{i=1}^k p_i \beta^{n_i} \right) = 0, \\ c_1 \left( \alpha^{r+1} - \sum_{i=1}^k p_i \alpha^{n_i+1} \right) + c_2 \left( \beta^{r+1} - \sum_{i=1}^k p_i \beta^{n_i+1} \right) = 0. \end{cases}$$

Пусть  $P(x) = x^r - \sum_{i=1}^k p_i x^{n_i}$ . Многочлен  $P(x)$  имеет целые коэффициенты, а старший его коэффициент равен 1. Систему пере-

пишем так:

$$\begin{cases} c_1 P(\alpha) + c_2 P(\beta) = 0, \\ c_1 \alpha P(\alpha) + c_2 \beta P(\beta) = 0. \end{cases}$$

Отсюда следует, что

$$\begin{cases} c_1 (\beta - \alpha) P(\alpha) = 0, \\ c_2 (\beta - \alpha) P(\beta) = 0. \end{cases}$$

Но  $c_1$  и  $c_2$  не равны нулю и  $\beta \neq \alpha$ , поэтому  $P(\alpha) = 0$  и  $P(\beta) = 0$ , т.е.  $\alpha$  и  $\beta$  – корни многочлена  $P(x)$ . А раз так, то

$$P(x) = Q(x)(ax^2 + bx + c).$$

Из неприводимости и примитивности квадратного трехчлена следует, что  $Q(x)$  – многочлен с целыми коэффициентами (см. утверждение 2). Но тогда, если  $q_0$  – старший коэффициент многочлена  $Q(x)$ , то старший коэффициент многочлена  $P(x)$  равен  $1 = a \cdot q_0$ . Но это возможно лишь при  $a = 1$  и  $q_0 = 1$ . А это мы и собирались доказать.

В общем случае все рассуждения аналогичны. Именно, рассматривая последовательность  $A$  целочисленных  $r$ -мерных векторов  $\bar{a}_n = (A_n, A_{n+1}, \dots, A_{n+r-1})$ , выбираем в  $A$  базисный набор. Исходя из базисного набора, строим многочлен  $P(x)$ . Доказываем, что все корни  $\alpha = \alpha_1, \dots, \alpha_r$  являются корнями  $P(x)$ , т.е.  $P(x)$  делится на неприводимый и примитивный многочлен  $p(x)$  и частное является многочленом с целыми коэффициентами. Откуда и следует, что старший коэффициент  $a_0$  многочлена  $p(x)$  равен 1.

Тем самым доказана замечательная теорема.

**Теорема 3.** Алгебраическое число  $\alpha$  обладает свойством Пизо тогда и только тогда, когда  $\alpha$  – число Пизо.

### Какие бывают числа Пизо

**Определение 10.** Многочлен  $p(x)$  с целыми коэффициентами и старшим коэффициентом, равным 1, называется многочленом Пизо, если один из его корней  $\alpha > 1$ , а остальные корни не равны нулю и по модулю меньше единицы.

Докажем, что всякий многочлен Пизо неприводим. Пусть  $p(x) = p_1(x)p_2(x)$ , где  $p_1(x)$  и  $p_2(x)$  – многочлены с целыми коэффициентами, и  $\alpha > 1$  – корень многочлена  $p(x)$ . Тогда  $p_1(\alpha)p_2(\alpha) = 0$ . Пусть, для определенности,  $p_1(\alpha) = 0$ . Но свободный член многочлена  $p_2(x)$ , равный по теореме Виета

произведению его корней, не равен нулю и по модулю меньше 1. Противоречие.

Докажем, что существуют числа Пизо любой наперед заданной степени  $r$ . Для этого достаточно построить многочлен Пизо степени  $r$ . Построение осуществим так.

Возьмем точки  $x_1 = \frac{1}{r}, x_2 = \frac{2}{r}, \dots, x_{r-1} = \frac{r-1}{r}, x_r = 1$  и построим многочлен  $f(x)$  с рациональными коэффициентами степени  $r - 1$  такой, что

$$\begin{aligned} f(1) &= -2, f\left(\frac{r-1}{r}\right) = \\ &= 2, \dots, f(x_k) = (-1)^{k+1} \cdot 2, \dots, f(x_1) = f\left(\frac{1}{r}\right) = (-1)^r \cdot 2. \end{aligned}$$

Вот один из способов такого построения (для знатоков – мы строим так называемый интерполяционный многочлен Лагранжа). Сначала строим многочлены  $\varphi_i(x)$  такие, что  $\varphi_i(x_i) = 1$ , а  $\varphi_i(x_k) = 0$  при  $i \neq k$ :

$$\begin{aligned} \varphi_1(x) &= \frac{(x - x_2)(x - x_3)\dots(x - x_r)}{(x_1 - x_2)(x_1 - x_3)\dots(x_1 - x_r)}, \\ \varphi_2(x) &= \frac{(x - x_1)(x - x_3)\dots(x - x_r)}{(x_2 - x_1)(x_2 - x_3)\dots(x_2 - x_r)}, \\ &\dots \\ \varphi_r(x) &= \frac{(x - x_1)\dots(x - x_{r-1})}{(x_r - x_1)\dots(x_r - x_{r-1})}. \end{aligned}$$

Все многочлены  $\varphi_i(x)$  имеют рациональные коэффициенты и степень  $r - 1$ .

Пусть

$$g(x) = -2\varphi_r(x) + 2\varphi_{r-1}(x) + \dots + (-1)^r \cdot 2\varphi_1(x).$$

Коэффициенты многочлена  $g(x)$  рациональны, а

$$g(x_i) = (-1)^{r-i+1} \cdot 2 \text{ при } i = 1, 2, \dots, r.$$

Предположим, что  $M$  – наименьший общий знаменатель его коэффициентов. Многочлен  $P(x) = Mg(x)$  имеет целые коэффициенты. При этом

$$P(x_i) = (-1)^{r-i+1} \cdot 2M.$$

Знаки чисел  $P(x_i)$  чередуются, и каждое из этих чисел по модулю не меньше чем 2.

Пусть теперь

$$f(x) = x^r + P(x).$$

Степень многочлена  $f(x)$  равна  $r$ , его коэффициенты – целые числа, старший коэффициент равен 1, а знаки чисел  $f(x_i) = x_i^r + P(x_i)$  чередуются. Поэтому  $f(x)$  имеет корень на каждом из промежутков  $(x_1; x_2), (x_2; x_3), \dots, (x_{r-1}; x_r)$ . А так как  $f(x_r) = f(1) < 0$ , то и на промежутке  $(1; +\infty)$  тоже есть корень.

Таким образом, многочлен  $f(x)$  степени  $r$  имеет в точности  $r$  действительных корней. Один из них больше 1, остальные положительны и меньше 1. По доказанному ранее, многочлен  $f(x)$  неприводим и больший его корень – число Пизо.

Итак, чисел Пизо достаточно много. Однако при любом натуральном  $k$  в промежутке  $(k; k+1)$  содержится лишь *конечное* количество чисел Пизо данной степени  $r$ . Это следует из того, что если  $P(x) = x^r + a_1x^{r-1} + \dots + a_r$  – многочлен Пизо и  $k < \alpha < k+1$  – его корень, то  $a_k$  – целое число, равное сумме произведений корней многочлена  $p(x)$  по  $k$ , умноженной на  $(-1)^k$ . Всего таких произведений  $C_r^k$ , а каждое из них по модулю меньше чем  $k+1$  (один корень – между  $k$  и  $k+1$ , а остальные – по модулю меньше 1), так что  $|a_k| < (k+1)C_r^k$ .

Но уже таких многочленов конечное число, так как все они имеют ограниченные в совокупности целые коэффициенты.

В то же время, в каждом промежутке  $(k; k+1)$  есть числа Пизо любой наперед заданной степени.

Мы приведем без доказательства соответствующие многочлены Пизо.

Для *второй* степени – это многочлен

$$x^2 - kx - 1,$$

где  $k$  – натуральное число. Один из его корней  $\alpha = \frac{k + \sqrt{k^2 + 4}}{2}$  лежит в интервале  $(k; k+1)$ , второй корень отрицателен и по модулю меньше 1.

• Для *третьей* степени это многочлен

$$x^3 - x - 1,$$

имеющий корень в интервале  $(1; 2)$ , и многочлен

$$x^3 - (k+1)x^2 + 1$$

при натуральном  $k \geq 2$  для промежутка  $(k; k+1)$ .

Наконец, для произвольной степени  $r$  это многочлен

$$x^r - x^{r-1} - \dots - 1$$

для промежутка  $(1; 2)$  и многочлен

$$x^r - (k+1)x^{r-1} + 1$$

для промежутка  $(k; k+1)$ .

Все перечисленные многочлены неприводимы и являются многочленами Пизо.

Известно и наименьшее не целое число Пизо. Это корень уравнения

$$x^3 - x - 1 = 0,$$

равный (по формуле Кардано)

$$x = \frac{1}{\sqrt[3]{18}} \left( \sqrt[3]{9 + \sqrt{69}} + \sqrt[3]{9 - \sqrt{69}} \right).$$

Доказано также, что множество чисел Пизо замкнуто, т.е. если последовательность  $\alpha_n$ , состоящая из чисел Пизо, имеет предел  $\lim_{n \rightarrow \infty} \alpha_n = \beta$ , то число  $\beta$  является числом Пизо.

Вопрос о существовании неалгебраических чисел  $\alpha > 1$ , обладающих свойством Пизо, открыт. Однако доказано, что если  $\{\{\alpha^n\}\}$  стремится к нулю достаточно быстро, точнее – если ряд  $\sum_{n=1}^{\infty} \{\{\alpha^n\}\}^2$  сходится, то  $\alpha$  является алгебраическим числом и, следовательно (в силу теоремы 3), – числом Пизо.

### Заключение

Впервые публикации о числах Пизо появились в научной литературе в 1946 году. Авторами их были французский математик Пизо и индийский математик Виджаярагхаван. В последующие годы неоднократно появлялись статьи о числах Пизо. В самое последнее время обнаружена связь чисел Пизо с теорией квазикристаллов в физике.

В 2000 году задача о числах Пизо была предложена школьникам – участникам конференции Турнира городов. Почти элементарное доказательство свойств чисел Пизо и подборка задач об этих числах были разработаны А.Канелем-Беловым, С.Дориченко и автором этой статьи.

# МАТЕМАТИКА БИЛЬЯРДА

А.Земляков

*Посвящается светлой памяти Александра Абрамовича Шершевского, Учителя математики.*

Бильярды, о которых рассказывается в этой статье, не очень похожи на обычные. Отличий много: бильярдный стол не всегда прямоугольный, шар только один и нет ни луз, ни кия. Но не расстраивайтесь: мы постараемся показать, что и один шарик на бильярдном поле может задать немало головоломок любителям математики.

## § 1. Общая проблема бильярда

Согласно законам механики, при отражении абсолютно упругого бильярдного шара от прямолинейного борта  $MN$  «угол падения» шара равен «углу его отражения» – более точно,

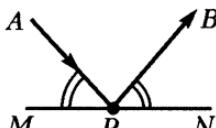


Рис. 1

угол  $APM$  равен углу  $BPN$  (рис.1). Ниже мы будем ссылаться на этот факт как на **закон упругого отражения**. Уже здесь мы предположили, что бильярдный шарик *точечный*, т.е., как говорят физики, «размерами шарика в данных рассмотрениях можно пренебречь». Таким образом, мы считаем, что бильярдный шар – это *движущаяся точка*, и поэтому можно говорить о *траектории* бильярдного шара – о ломаной линии, по которой движется шар в соответствии с законом упругого отражения (рис.2,а).

Точки излома этой ломаной лежат на «бортах бильярдного стола». Мы будем рассматривать бильярдные столы произвольной формы. В принципе борты бильярда могут быть и криволинейными – тогда закон упругого отражения формулируется, как и раньше, только под углами падения и отражения следует понимать углы между соответствующими («падающим» и «отраженным») звеньями траектории и *касательной*  $MN$  к криволинейному борту в точке  $P$  соударения шара с бортом. Мы не даем общего определения касательной к кривой – достаточно интуитивно понимать, что такое касательная. В точке излома самого бильярдного борта (на рисунке 2,а

---

Опубликовано в «Кванте» №5 за 1976 г.

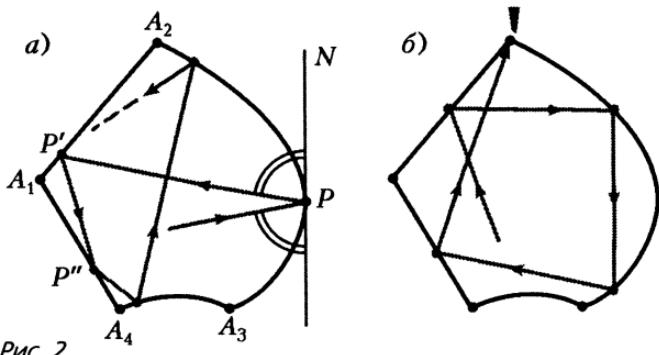


Рис. 2

это точки  $A_1, A_2, \dots$ ) касательной к борту нет, и если бильярдный шар (точка!) попадает в точности в одну из таких точек, то закон упругого отражения не дает возможности выяснить, как траектория шара должна выходить из этой точки. Поэтому траектории, попадающие в одну из точек излома борта, будем считать *заканчивающимися* в этих точках (рис.2,б).

Итак, пусть на плоскости задана произвольная область  $Q$  (бильярдный стол), ограниченная некоторой кривой  $\Gamma$  (бортом), возможно, с точками излома. Примеры бильярдных столов – круг, круговой сектор, многоугольник. Для простоты будем считать бильярдный стол *выпуклым*. Через заданную в области  $Q$  точку  $P$  в заданном направлении  $\vec{v}$  можно провести единственную бильярдную траекторию – ломаную  $PP_1P_2P_3\dots$  (рис.3). В общем случае эта траектория никогда не попадет точно в точки излома борта, и бильярдный шар будет двигаться по ней неограниченно долго. Может случиться так, что при построении траектории, проходящей через точку  $P$  в направлении  $\vec{v}$ , через некоторое время эта траектория снова пройдет через ту же точку  $P$  в том же самом направлении. Очевидно, что далее траектория опять пойдет по тому же пути. Такая ситуация соответствует повторяющемуся, периодическому движению бильярдного шарика. Отвечающие подобному движению траектории называются *периодическими*.

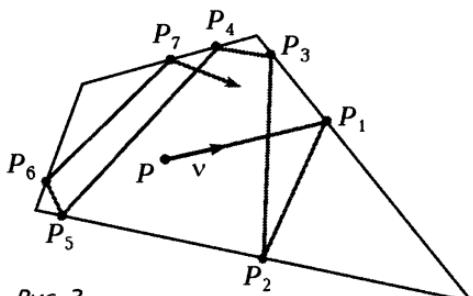


Рис. 3

Геометрически периодическая бильярдная траектория – это замкнутая вписанная в кривую  $\Gamma$  ломаная, обладающая

свойством «равенства углов падения и отражения». Примеры периодических траекторий в некоторых областях изображены на рисунке 4.

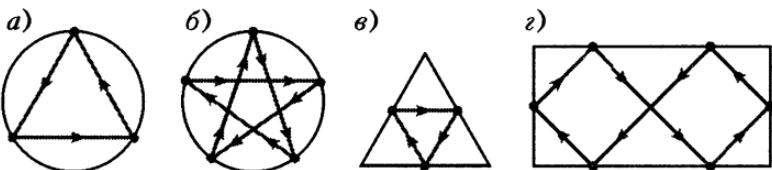


Рис. 4

Большой интерес представляют собой не периодические бесконечные траектории, мы их будем называть просто *непериодическими*. Шарик движется по любой непериодической траектории неограниченно долго и при этом никогда не проходит через одну и ту же точку дважды в одинаковом направлении (разумеется, в разных направлениях шарик может проходить через одну и ту же точку и несколько раз). Примеры непериодических траекторий будут приведены ниже.

*Проблема бильярда в области Q* заключается в том, чтобы найти ответы на такие вопросы: *существуют ли периодические траектории? много ли их? как они устроены? как узнать, будет ли траектория, выходящая из данной точки в данном направлении, периодической или непериодической? в каких частях области Q побывает шарик, двигаясь по конкретной непериодической траектории?* и т.д. В § 2 мы решим проблему бильярда в круге – это простейший из бильярдов, поддающихся полному исследованию. Но сначала – несколько задач, над которыми полезно подумать до того, как читать дальше.

**Задача 1.** Пусть  $Q$  – круг. Как узнать по начальному звену бильярдной траектории, будет ли эта траектория периодической или нет? Как близко может подойти к центру круга эта траектория? В каких частях круга побывает шарик, двигаясь по одной из непериодических траекторий?

**Задача 2.** Укажите хотя бы одну непериодическую траекторию бильярда в прямоугольнике или в квадрате. Укажите несколько типов периодических траекторий бильярда в квадрате. Как по углу  $\alpha$ , образованному начальным звеном траектории бильярда в квадрате с одной из сторон квадрата, узнать, будет эта траектория периодической или нет?

**Задача 3.** Укажите несколько периодических траекторий бильярда в равностороннем треугольнике (не таких, как на рисунке 4,в).

## § 2. Бильярд в круге

Рассмотрим бильярд в круге  $Q$ , ограниченном окружностью  $\Gamma$ . Его траекториями являются вписанные в  $\Gamma$  ломаные  $P_0P_1P_2P_3\dots$ , обладающие свойством равенства углов, падения и отражения. Из этого свойства следует, что все звенья траектории равны между собой, как равны друг другу и опирающиеся на них центральные углы  $P_0OP_1$ ,  $P_1OP_2$ ,  $P_2OP_3$ , ... (докажите это – например, с помощью рисунка 5). Пусть  $\alpha$  – радианная мера этих углов.

Ясно, что каждая вершина  $P_k$  траектории  $P_0P_1P_2\dots$  получается из предыдущей вершины  $P_{k-1}$  поворотом на угол  $\alpha$  радиан относительно центра окружности  $\Gamma$ , поэтому вершина  $P_n$  получается из начальной вершины  $P_0$  поворотом на угол  $n\alpha$  радиан. Вид бильярдной траектории в круге полностью определяется числом  $\alpha$ .

*a) Если число  $\alpha$  соизмеримо с  $2\pi$ , то отвечающая  $\alpha$  бильярдная траектория периодична.* (Напомним, что числа  $a$  и  $b$  называются соизмеримыми, если их отношение является рациональным числом, т.е.  $a/b = m/n$ , где  $m$  и  $n$  – целые числа.)

*б) Если отношение  $\alpha/\pi$  иррационально (т.е.  $\alpha$  и  $\pi$  несоизмеримы), то отвечающая углу  $\alpha$  траектория непериодична.*

**Доказательство.** а) Пусть  $\alpha$  соизмеримо с  $2\pi$ ; тогда  $\alpha = (m/n) \cdot 2\pi$ , где  $m$  и  $n$  – целые. Следовательно,  $n\alpha = m \cdot 2\pi$ , и поэтому при повороте около центра окружности на угол  $n\alpha$  радиан каждая точка окружности переходит в себя. В частности, для вершин рассматриваемой траектории получаем:  $P_n = P_0$ ,  $P_{n+1} = P_1$ ,  $P_{n+2} = P_2$ , ... Таким образом, вершины траектории  $P_0P_1P_2P_3\dots$ , начиная с  $n-й}, повторяются, т.е. траектория периодична.$

Заметим, что если дробь  $m/n = \alpha/2\pi$  несократима, то отвечающая  $\alpha$  периодическая траектория – это замкнутая ломаная  $P_0P_1P_2\dots P_{n-1}P_0$ , состоящая в точности из  $n$  звеньев. При  $m=1$  это будет правильный  $n$ -угольник, вписанный в окружность  $\Gamma$ , а при  $m \geq 2$  траектория представляет собой правильную самопересе-

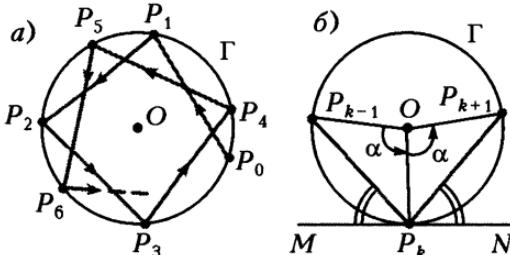


Рис. 5

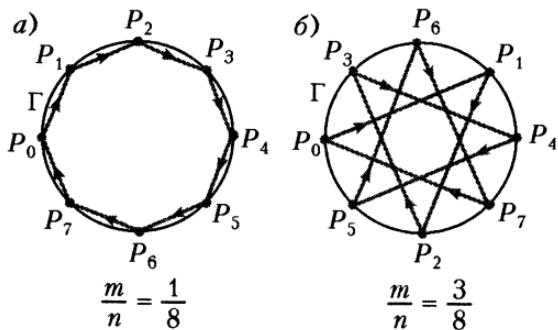


Рис. 6

кающуюся замкнутую (звездчатую) ломаную, вписанную в  $\Gamma$  (рис.6).

б) Рассуждая от противного, видим, что достаточно доказать следующее: если траектория  $P_0P_1P_2P_3\dots$  периодична, то  $\alpha$  и  $\pi$  соизмеримы. Из периодичности траектории вытекает, что, начиная с некоторого номера  $n$ , вершины траектории повторяются:  $P_n = P_0$ ,  $P_{n+1} = P_1$ ,  $P_{n+2} = P_2$ , ... Но это означает, что при повороте на угол  $n\alpha$  радиан точка  $P_0$  переходит сама в себя; следовательно, угол  $n\alpha$  есть целое кратное полного угла:  $n\alpha = m \cdot 2\pi$ . Таким образом,  $\alpha/\pi = 2m/n$ .

Итак, если  $\alpha$  несоизмеримо с  $\pi$ , то отвечающая углу  $\alpha$  бильярдная траектория есть незамкнутая ломаная  $P_0P_1P_2P_3\dots$

с равными по длине звенями, вписанная в окружность  $\Gamma$ . Поскольку звенья этой ломаной равны по длине, их середины удалены от центра окружности  $\Gamma$  на одинаковое расстояние, т.е. лежат на некоторой окружности  $\Gamma_1$ , концентрической с  $\Gamma$  (рис.7). Все звенья рассматриваемой траектории касаются этой меньшей окружности  $\Gamma_1$ , и траектория никогда не заходит внутрь  $\Gamma_1$ , т.е. целиком лежит в кольце  $K$  между окружностями  $\Gamma_1$  и  $\Gamma$ .

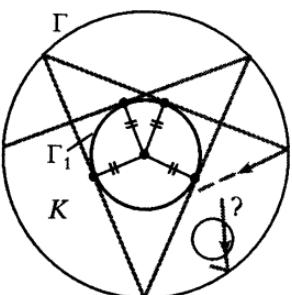


Рис. 7

**Теорема 1.** Если  $\alpha$  и  $\pi$  несоизмеримы, то любая траектория бильярда в круге, отвечающая углу  $\alpha$ , всюду плотно заполняет кольцо  $K$ .

**Совет.** Прежде чем читать дальше, продумайте хорошенко формулировку теоремы, поймите ее и попробуйте доказать.

[Самое главное место в формулировке – слова «траектория всюду плотно заполняет кольцо  $K$ ». Их точный смысл можно пояснить так:

бильярдный шарик, двигаясь по непериодической траектории, рано или поздно побывает в любом кусочке кольца  $K$ . Если считать, что шарик «чернильный» и оставляет после себя след, то он со временем *обязательно закрасит все кольцо целиком*, каким бы тонким ни был чернильный след (но имеющим все-таки ненулевую «толщину»). Периодическая траектория может заполнить кольцо «очень плотно», но не «всюду плотно» — на кольце обязательно останутся такие участки (например, кружки), которые периодическая траектория не пересечет никогда.]

Пусть  $P_0P_1P_2P_3\dots$  — рассматриваемая непериодическая траектория. Утверждение теоремы следует из того факта, что точки  $P_0, P_1, P_2, P_3, \dots$  расположены на окружности  $\Gamma$  *всюду плотно*, а это, в свою очередь, вытекает из такого утверждения.

**Теорема Якоби.** *Пусть последовательность  $\{P_k\} = \{P_0, P_1, P_2, \dots\}$  точек окружности  $\Gamma$  такова, что каждая следующая ее точка  $P_{k+1}$  получается из предыдущей точки  $P_k$  поворотом около центра окружности на один и тот же угол  $\alpha$  радиан. Тогда, если  $\alpha$  и  $\pi$  несоизмеримы, то последовательность точек  $\{P_k\}$  заполняет окружность  $\Gamma$  всюду плотно, т.е. для любой дуги  $\Delta$  окружности  $\Gamma$  найдется номер  $n$  такой, что точка  $P_n$  с этим номером лежит на дуге  $\Delta$ .*

Попробуйте доказать теорему Якоби самостоятельно. Проведите подробный вывод теоремы 1 из теоремы Якоби. Чтобы дать возможность подумать, мы этот вывод и доказательство теоремы Якоби вынесли в специальное дополнение к §2.

Теперь бильярд в круге можно считать полностью исследованным. Резюмируем полученные результаты. *Траектория бильярда в круге является либо периодической (если число  $\alpha/\pi$  рационально), либо всюду плотной в некоторой области — в кольце  $K$  между бортом бильярда  $\Gamma$  и концентрической с  $\Gamma$  окружностью  $\Gamma_1$  (если число  $\alpha/\pi$  иррационально).*

**Задача 4.** Исследуйте бильярды в полукруге и в секторе с прямым центральным углом. (Попробуйте свести эти два бильярда к бильярду в круге!)

Возникает естественный вопрос: а как ведут себя траектории бильярда в произвольной выпуклой области  $Q$ , ограниченной кривой  $\Gamma$  без точек излома? Такие бильярды изучались многими математиками, в том числе известным американским математиком Дж.Биркгофом, советскими математиками В.Ф.Лазуткиным, Л.А.Бунимовичем. Эти исследования, мягко говоря, не совсем элементарны. Мы сформулируем два наиболее интересных факта, относящихся к таким бильярдам.

**I (Анри Пуанкаре, Дж.Биркгоф).** Бильярд в произвольной выпуклой области имеет бесконечно много периодических траекторий. Более того, для любого  $n \geq 3$  существуют периодические траектории, состоящие в точности из  $n$  звеньев. (В конце §3 мы приведем указанное Биркгофом изящное доказательство этого утверждения.)

**II (В.Ф.Лазуткин).** В общем случае внутри выпуклой области  $Q$  с границей  $\Gamma$  можно указать семейство кривых  $\{\Gamma_c\}$ , обладающих следующим свойством: если начальное звено какой-нибудь траектории касается одной из кривых  $\Gamma_c$ , то и все остальные звенья этой траектории касаются *той же* кривой  $\Gamma_c$  (рис.8). Кривые с такими свойствами

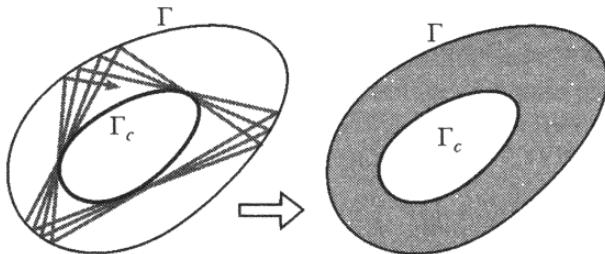


Рис. 8

называются *бильярдными каустиками в области Q*. Когда  $Q$  – круг, каустиками являются окружности, концентрические с  $\Gamma$ . Доказать существование каустик для произвольной области  $Q$  весьма сложно. [Попробуйте найти каустики для бильярда в эллипсе (см. задачи 12 и 13).] Совсем трудно доказать в общем случае аналог теоремы 1: *почти любая непериодическая траектория, касающаяся каустики  $\Gamma_c$ , всюду плотно заполняет область между бортом бильярда  $\Gamma$  и этой каустикой  $\Gamma_c$  («кривое кольцо»)*.

Оказывается, что периодические траектории бильярда в области  $Q$  и его каустики тесно связаны с так называемыми собственными колебаниями упругой мембранны, имеющей форму области  $Q$ . Грубо говоря, если сделать барабан в форме области  $Q$  (с каркасом в форме кривой  $\Gamma$ ), то издаваемые этим барабаном музыкальные звуки как раз и связаны с соответствующим бильярдом! К сожалению, чтобы объяснить этот поразительный факт, нам пришлось бы углубиться в слишком серьезные для «Кванта» разделы математики. Заинтересовавшемуся читателю можно только посоветовать изучить математику в таком объеме, чтобы самому объяснить связь между бильярдом и барабаном.

### §3. Экстремальные свойства траекторий бильярда

Способ построения бильярдной траектории, ведущей из одной данной точки  $A$  в другую данную точку  $B$  после одного отражения от прямолинейного борта  $MN$ , вытекает из закона упругого отражения: строим точку  $B'$ , симметричную точке  $B$  относительно прямой  $MN$ , и соединяем  $A$  с  $B'$  (рис.9); точка  $P$  пересечения прямых  $AB'$  и  $MN$  и будет точкой соударения шара с бортом  $MN$ , а ломаная  $APB$  – это искомая траектория (поясните). Из построения очевидно, что бильярдная траектория  $APB$  – кратчайшая из всех ломанных вида  $AP'B$  с точкой  $P'$  на прямой  $MN$ , ведущих из  $A$  в  $B$ ! (Дайте подробное доказательство.) Это утверждение есть частный и простейший случай фундаментального принципа классической механики – *принципа наименьшего действия* (в оптике он называется *принципом Ферма*)<sup>1</sup>.

В нашей ситуации этот принцип можно сформулировать так: бильярдный шарик «умный» и из всех возможных путей выбирает самый короткий. Пока «принцип наикратчайшего пути» установлен нами лишь в случае одного отражения от прямолинейного борта. В задачах 5–7 предлагается проверить его в тех случаях, когда шар отражается не один раз, а несколько, но от прямолинейных бортов.

**Задача 5.** Внутри острого угла  $MON$  даны точки  $A$  и  $B$ . Постройте кратчайшую ломаную  $AP_1P_2B$  с вершинами  $P_1$  на луче  $OM$  и  $P_2$  на луче  $ON$ . Докажите, что эта ломаная является траекторией бильярда внутри угла  $MON$ , ведущей из  $A$  в  $B$ .

**Задача 6. а)** В данный острый угол  $MON$  впишите треугольник  $ABC$  наименьшего периметра, вершина  $A$  которого находилась бы в данной внутрь угла точке, а две другие лежали бы на сторонах угла.

**б)** «Задача Фаньяно». В данный остроугольный треугольник  $MLN$  впишите треугольник  $ABC$  наименьшего периметра. Докажите, что этот треугольник  $ABC$  дает периодическую траекторию бильярда внутри треугольника  $MLN$ .

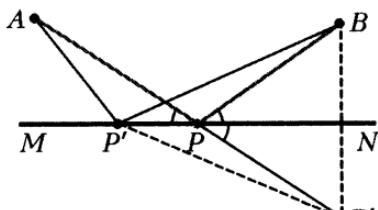


Рис. 9

<sup>1</sup> Подробно об этом см., например, в статье С.Г. Верова «Брахистохрона, или еще одна тайна циклоиды» в «Кванте» № 12 за 1975 г.

**Задача 7.** На прямоугольном бильярдном столе  $KLMN$  лежат два шара  $A$  и  $B$ . В каком направлении нужно толкнуть шар  $A$  для того, чтобы после четырех последовательных отражений от бортов  $KL$ ,  $LM$ ,  $MN$ ,  $NK$  он ударился бы о шар  $B$ ? При любых ли положениях шаров  $A$  и  $B$  задача имеет решение? Будет ли эта траектория  $AP_1P_2P_3P_4B$  кратчайшей из всех ломаных такого вида (начинающихся в точке  $A$ , заходящих на борты стола в указанной последовательности и кончающихся в точке  $B$ )?

Оказывается, что замена прямолинейного борта криволинейным настолько «сбивает с толку» бильярдный шарик, что он иногда движется не самым коротким, а, напротив, самым длинным путем! Рассмотрим эту ситуацию подробно.

**Теорема 2.** *Если ломаная  $APB$  с концами  $A$  и  $B$  и с точкой излома  $P$  на гладкой кривой  $\Gamma$  является локально минимальной или локально максимальной, т. е. имеет наименьшую или наибольшую длину среди тех ломанных  $AP'B$ , вершины  $P'$  которых лежат на кривой  $\Gamma$  в некоторой окрестности точки  $P$ , то эта ломаная  $APB$  является бильярдной траекторией, ведущей из  $A$  в  $B$  после отражения от кривой  $\Gamma$  в точке  $P$  (т.е. звенья  $PA$  и  $PB$  образуют равные углы с касательной к кривой  $\Gamma$  в точке  $P$ ).*

**Замечание.** Далее для краткости слово «локальный» мы опускаем. При доказательстве сформулированного «принципа экстремального» (т.е. минимального или максимального) пути для бильярдных траекторий мы воспользуемся распространенным приемом исследования задач на максимум и минимум; его описание в общем случае можно найти в книжке Н.Б.Васильева и В.Л.Гутенмакера «Прямые и кривые», §7, или в статье В.Г.Болтянского и И.М.Яглома «Геометрические задачи на максимум и минимум» в V томе «Энциклопедии элементарной математики».

**Доказательство теоремы 2.** Для каждого числа  $c > AB$  рассмотрим множество  $\Gamma_c$ , состоящее из тех точек  $M$ , для которых  $AM + MB = c$ . Как известно (см., например, упомянутую выше книгу «Прямые и кривые», §5, или статью И.Н.Бронштейна «Эллипс», «Кvant» №1 за 1975 год), множества  $\Gamma_c$  – это эллипсы с фокусами в точках  $A$  и  $B$ ; они образуют семейство софокусных эллипсов  $\{\Gamma_c\}$ . Нетрудно видеть, что если ломаная  $APB$  минимальна или максимальна, то точка  $P$  является точкой касания кривой  $\Gamma$  с одним из эллипсов семейства  $\{\Gamma_c\}$  – докажите это с помощью рисунков 10,а и б. (Отметим, что не все точки касания дают максимум или минимум – см. рис.10,в.) В такой точке  $P$  кривая  $\Gamma$  и касающийся ее эллипс  $\Gamma_c$  имеют общую

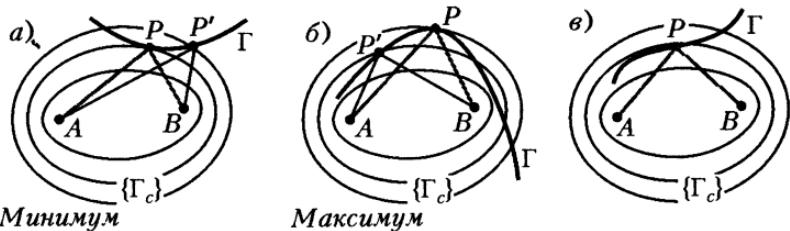


Рис. 10

касательную  $MN$ . Согласно *оптическому свойству эллипса*<sup>2</sup>, отрезки  $PA$  и  $PB$ , соединяющие точку  $P$  с фокусами эллипса  $\Gamma_c$ , образуют равные углы с касательной  $MN$  к  $\Gamma_c$  в точке  $P$  (рис.11). Поскольку прямая  $MN$  одновременно является и касательной к кривой  $\Gamma$  в точке  $P$ , ломаная  $APB$  – бильярдная траектория, идущая из  $A$  в  $B$  после отражения от  $\Gamma$  в точке  $P$ ! Теорема 2 доказана.

#### Обращение к придиличным читателям.

Если наши рассуждения показались вам слишком нестрогими – например, из-за отсутствия точных определений касания кривых, локальных экстремумов и т.д., – то мы советуем обратиться к упоминавшейся выше статье В.Г.Болтянского и И.М.Яглома в V томе «Энциклопедии элементарной математики» (с. 270–346). В свое оправдание заметим, что строгое обсуждение всех этих понятий уело бы нас слишком далеко от бильярдов.

**Пример.** Пусть  $\Gamma$  – окружность,  $A$  и  $B$  – точки внутри  $\Gamma$ . Рассматривая семейство  $\{\Gamma_c\}$  софокусных эллипсов с фокусами  $A$  и  $B$ , можно показать, что существует ровно четыре траектории, ведущих из  $A$  в  $B$  после одного отражения от  $\Gamma$ , причем две из них отвечают минимумам, а две другие – максимумам суммы  $AP + PB$  (см. рис.12).

**Задача 8.** Утверждение, обратное принципу экстремального пути, гласит: *бильярдная траектория  $APB$ , идущая из точки  $A$  в точку  $B$  после отражения от кривой  $\Gamma$ , является либо минимальной, либо максимальной ломаной*. Приведите пример, по-

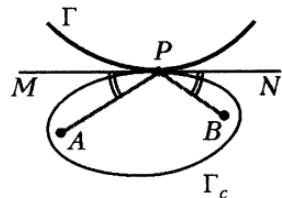


Рис. 11

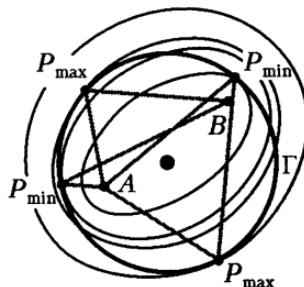


Рис. 12

<sup>2</sup> См., например, статью В.Г.Болтянского в «Кванте» № 12 за 1975 г.

казывающий, что для невыпуклых кривых  $\Gamma$  это утверждение неверно. Будет ли оно верно для выпуклых кривых  $\Gamma$  (т.е. для кривых, лежащих целиком по одну сторону от любой своей касательной)? Докажите этот «обратный принцип» в том случае, когда кривая  $\Gamma$  выпуклая и обращена своей выпуклостью к точкам  $A$  и  $B$  (как на рисунке 10, $a$ ).

Применим теперь принцип экстремального пути к упомянутой в §2 теореме Биркгофа:

*если  $Q$  – выпуклая область, ограниченная кривой  $\Gamma$  без точек излома, то для произвольного  $n \geq 3$  существует хотя бы одна периодическая траектория бильярда о области  $Q$ , состоящая ровно из  $n$  звеньев.*

**Доказательство.** Рассмотрим всевозможные несамопересекающиеся замкнутые  $n$ -звенные ломаные  $P_1P_2\dots P_nP_1$ , вписанные в область  $Q$  (т.е. с вершинами  $P_i$  на кривой  $\Gamma$ ). Выберем среди них ломаную наибольшей длины (подумайте, почему такая ломаная существует?).

**Задача 9.** Докажите с помощью принципа экстремального пути, что эта самая длинная ломаная является бильярдной траекторией в области  $Q$ .

Из задачи 9 следует, что «наи длиннейшая» ломаная  $P_1P_2\dots P_nP_1$  и будет замкнутой периодической траекторией, и теорема Биркгофа доказана (мы не хотим лишать читателя удовольствия от самостоятельного решения задачи 9).

**Задача 10.** а) Проанализируйте утверждение теоремы Биркгофа при  $n = 2$ .

б) Подумайте, будет ли настоящей  $n$ -звенной периодической траекторией ломаная, «наи длиннейшая» среди всех (в том числе и самопересекающихся)  $n$ -звенных ломаных, вписанных в кривую  $\Gamma$ ? (Рассмотрите случаи  $n = 4$  и  $n = 5$ .)

в) Какие затруднения возникнут при доказательстве теоремы Биркгофа, если вместо самой длинной ломаной  $P_1P_2\dots P_nP_1$  взять *самую короткую*?

**Задача 11.** Какой из вписанных в окружность  $n$ -угольников имеет наибольший периметр?

Сформулируем две задачи про бильярд внутри эллипса  $\Gamma$  с фокусами  $A$  и  $B$ . Согласно оптическому свойству эллипса (см. выше), если выпустить бильярдную траекторию из фокуса, то ее звенья будут поочередно проходить через фокусы  $A$  и  $B$ , причем эта траектория будет неограниченно приближаться к большой оси эллипса – к отрезку  $A_1B_1$  (рис.13, $a$ ).

**Задача 12.** Пусть начальное звено бильярдной траектории в эллипсе  $\Gamma$  не пересекает отрезок  $AB$  между фокусами. Докажите, что и все следующие звенья этой траектории не будут пересекать отрезок  $AB$ .

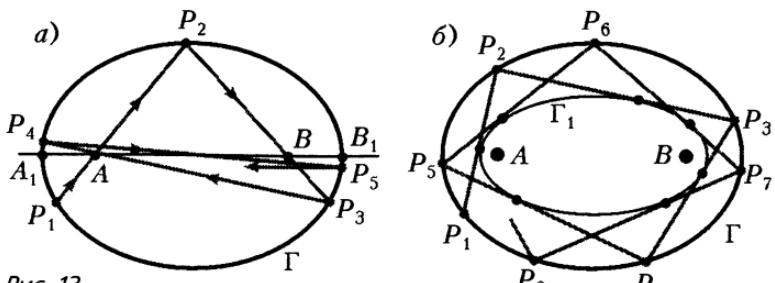


Рис. 13

Более того, если построить меньший эллипс  $\Gamma_1$  с теми же фокусами, касающийся начального звена траектории, то  $\Gamma_1$  будет касаться и всех остальных звеньев (см. рис. 13, б).

**Задача 13.** Докажите, что если начальное звено бильярдной траектории в эллипсе  $\Gamma$  пересекает отрезок  $AB$  между фокусами, то и все следующие ее звенья пересекают отрезок  $AB$ . Более того, если построить гиперболу с теми же фокусами  $A$  и  $B$ , касающуюся начального звена  $P_1P_2$  (или прямой  $P_1P_2$ ), то ветви этой гиперболы будут касаться и всех остальных звеньев  $P_kP_{k+1}$  (или их продолжений – прямых  $P_kP_{k+1}$ )<sup>3</sup>.

Задачи 12 и 13 показывают, что для бильярда в эллипсе с фокусами  $A$  и  $B$  *каустиками* (см. § 2) являются гиперболы и меньшие эллипсы с теми же фокусами  $A$  к  $B$ .

В заключение приведем две замечательные задачи, относящиеся к общей проблеме бильярдов в многоугольниках.

**Задача 14.** Докажите, что любая непериодическая траектория бильярда в прямоугольнике всюду плотно заполняет весь (!) прямоугольник (сравните с задачей 2; отметим, что отнюдь не всегда непериодическая траектория всюду плотно заполняет некоторую область – например, траектория бильярда в эллипсе, проходящая через фокусы, этим свойством не обладает!).

**Задача 15.** Докажите то же самое для бильярда в равностороннем треугольнике.

## Дополнение к §2

**(А) Теорема Якоби.** Напомним, как она формулируется.

Пусть  $\alpha$  – несоизмеримое с  $\pi$  число,  $\{P_0, P_1, P_2, \dots\} = \{P_k\}$  – последовательность точек окружности  $\Gamma$  такая, что каждая следующая

<sup>3</sup> Определение и нужные свойства гиперболы см. в статье И.Н.Бронштейна «Гипербола» в «Кванте» № 3 за 1975 г., а также в вышеупомянутой статье В.Г.Болтнисского.

точка последовательности  $P_{k+1}$  получается из предыдущей ее точки  $P_k$  поворотом на угол  $\alpha$  радиан. Тогда для любой дуги  $\Delta$  окружности  $\Gamma$  хотя бы одна точка последовательности  $\{P_k\}$  лежит на этой дуге  $\Delta$ .

**Доказательство.** Пусть  $\Delta$  – произвольная дуга,  $\varepsilon$  – ее радианная мера. Выберем натуральное число  $N$  такое, что  $2\pi/N < \varepsilon$ . Разобьем окружность  $\Gamma$  на  $N$  равных по длине дуг  $\Delta_1, \Delta_2, \dots, \Delta_N$ ; радианная мера каждой из них равна  $2\pi/N$  и меньше  $\varepsilon$ . Рассмотрим первые  $N + 1$  точек последовательности  $\{P_k\}$ , т.е. точки  $P_0, P_1, P_2, \dots, P_N$ . Согласно принципу Дирихле, хотя бы на одной из  $N$  дуг  $\Delta_1, \dots, \Delta_N$  лежат по крайней мере две из этих точек. Пусть, например, точки  $P_n$  и  $P_m$ , где  $m = n + l > n$ , лежат на дуге  $\Delta_i$ . Так как  $\alpha$  несоизмеримо с  $\pi$ , то  $P_m \neq P_n$  (поясните!). Пусть радианная мера дуги  $P_n P_m$  равна  $\delta$ ; тогда  $\delta < 2\pi/N$  и тем более  $\delta < \varepsilon$ .

Рассмотрим теперь точки последовательности  $\{P_k\}$  с номерами  $n, n + l, n + 2l, \dots$  (рис. 14). Они образуют подпоследовательность

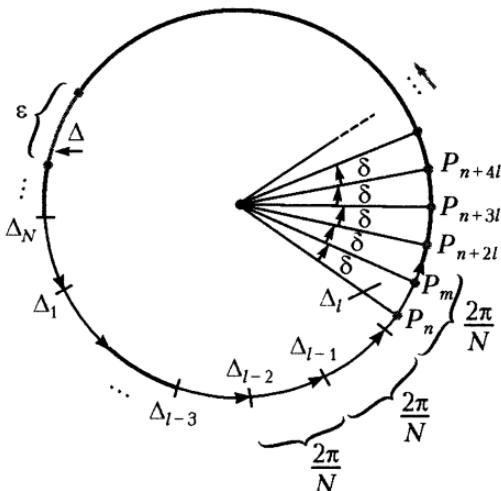


Рис. 14

$\{P_n, P_{n+l}, P_{n+2l}, P_{n+3l}, \dots\}$ , каждая следующая точка которой  $P_{n+(r+1)l}$  получается из предыдущей точки  $P_{n+rl}$  поворотом на угол  $l\alpha$ , т.е. точно так же, как точка  $P_{n+l} = P_m$  получается из точки  $P_n$ . Таким образом, соседние точки выделенной подпоследовательности отстоят друг от друга на дугу радианной меры  $\delta$ . Так как  $\delta < \varepsilon$ , хотя бы одна из точек  $P_n, P_{n+l}, P_{n+2l}, \dots$  обязана попасть на дугу  $\Delta$ . Тем самым теорема Якоби доказана.

**(Б) Вывод теоремы 1 из теоремы Якоби.** Пусть  $\alpha$  – несоизмеримое с  $\pi$  число,  $\Gamma_1$  – окружность, которой касаются звенья каждой отвечающей углу  $\alpha$  траектории бильярда внутри окружности  $\Gamma$ . Пусть  $K$  –

кольцо между  $\Gamma_1$  и  $\Gamma$ , а  $D$  – произвольный кружок внутри кольца  $K$ . Проведем через его центр хорду  $MN$  окружности  $\Gamma$ , касающуюся окружности  $\Gamma_1$  (рис.15). Очевидно, что если взять достаточно малую дугу  $\Delta$  окружности  $\Gamma$  с серединой в точке  $M$ , то для любой другой точки  $M'$  дуги  $\Delta$  касательная к окружности  $\Gamma_1$ , проведенная из точки  $M'$  по ту же сторону от  $\Gamma_1$ , что и  $MN$ , обязательно пересекает кружок  $D$  (поясните!).

Рассмотрим теперь любую траекторию  $P_0P_1P_2\dots$  бильярда внутри окружности  $\Gamma$ , отвечающую углу  $\alpha$ . Требуется доказать, что хотя бы одно из ее звеньев пересекает кружок  $D$ . Согласно теореме Якоби, точки  $P_0, P_1, P_2, \dots$  всюду плотно заполняют окружность  $\Gamma$ , и поэтому хотя бы одна из них, скажем  $P_n$ , лежит на выбранной выше дуге  $\Delta$ . В силу выбора этой дуги  $\Delta$  одно из двух звеньев траектории, выходящих из точки  $P_n$ , обязательно пересекает кружок  $D$ . Тем самым теорема 1 доказана.

Кроме теоремы 1 теорема Якоби имеет и много других интересных следствий<sup>4</sup>. Мы сформулируем два из них как задачи.

**Задача 16.** Докажите, что если число  $d$  иррационально, то для любого числа  $a$  последовательность дробных частей членов арифметической прогрессии  $a, a+d, a+2d, \dots$  т. е. последовательность чисел  $a_n = (a+nd)$ , *всюду плотно* заполняет отрезок  $0 \leq x \leq 1$  на числовой оси. ( $\{x\}$  – обозначение для дробной части числа  $x$ :  $\{x\} = x - [x]$ ).

**Задача 17.** Докажите, что если десятичный логарифм числа  $c$  иррационален, то десятичная запись натуральных степеней  $c$ , т.е. чисел  $c, c^2, c^3, \dots, c^n, \dots$  может начинаться с любого наперед заданного набора цифр. (Например, найдется такая натуральная степень числа 17, которая начинается с семнадцати идущих подряд наборов цифр 1976.)

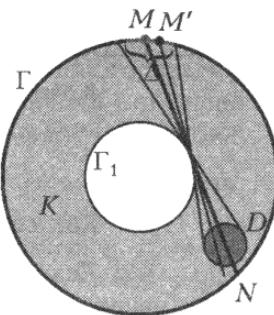


Рис. 15

---

<sup>4</sup> Некоторые приложения теоремы Якоби можно найти в статье А.А.Егорова «Решетки и правильные многоугольники» в «Кванте» №12 за 1974 г. (см. также «Библиотечку «Квант», вып.131).

# КВАТЕРНИОНЫ

А.Мищенко, Ю.Соловьев

## Как сделать из точек числа?

Если речь идет о точках на прямой – это просто. Выбрав начало отсчета («нуль») и масштаб с направлением («единицу»), можно получить из прямой числовую ось и тем самым превратить каждую точку в действительное число – ее координату (рис.1).

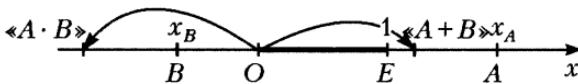


Рис. 1

С точками на плоскости сложнее. Выбрав начало отсчета («нуль») и пару перпендикулярных осей, можно сопоставить каждой точке на плоскости пару ее координат  $(x; y)$ . Чтобы

каждую такую пару – *дуплет* – сделать числом, нужно научиться «складывать» и «умножать» эти дуплеты, причем так, чтобы сохранялись привычные свойства сложения и умножения (переместительный, сочетательный и распределительный законы, наличие обратных операций – деления и вычитания).

Со сложением просто. Дуплеты естественно складывать как векторы – покоординатно (рис.2):

$$(x; y) + (x'; y') = (x + x'; y + y'). \quad (1)$$

С умножением же дело обстоит хитрее<sup>1</sup>. Однако и здесь не очень сложная формула дает выход из положения:

$$(x; y) \cdot (x'; y') = (xx' - yy'; xy' + x'y). \quad (2)$$

Опубликовано в «Кванте» №9 за 1983 г.

<sup>1</sup> Покоординатное умножение  $(x; y)(x'; y') = (xx'; yy')$  ничего хорошего не дает: для такой операции нет обратной операции (невозможно, к примеру, деление на ненулевой дуплет  $(0; 2)$ ).

Нетрудно проверить, что такое умножение дуплетов вместе со сложением (1) обладают выше упомянутыми «привычными свойствами». Таким образом, множество дуплетов с операциями (1), (2) можно считать полноценным числовым множеством.

На самом деле дуплеты – не что иное, как комплексные числа. Их чаще записывают не в виде  $(x; y)$ , а как  $x + yi$ , где  $i$  – мнимая единица (дуплет  $(0; 1)$ ), обладающая замечательным свойством  $i^2 = i \cdot i = -1$ , позволяющим (в области комплектных чисел) извлекать корни из отрицательных чисел. Подробнее о комплексных числах можно прочитать в статье Л.С.Понtryгина в «Кванте» №2 за 1983 г.

А как превратить точки пространства в числа? Здесь снова, вводя систему координат, можно записывать точки в виде наборов их координат, но уже не двух, а трех:  $(x; y; z)$ . Такие тройки, или *триплеты*, естественно складывать покоординатно:

$$(x; y; z) + (x'; y'; z') = (x + x'; y + y'; z + z'). \quad (3)$$

Триплеты можно будет считать числами, если найдется способ их умножения, обладающий, вместе со сложением (3), обычными свойствами этих операций. В частности, обратной к умножению операцией (делением на ненулевые элементы). Так все-таки, как умножаются триплеты? В 1833 г. этой задачей заинтересовался ирландский математик Уильям Роан Гамильтон (1805–1865). Но об этом незаурядном человеке стоит рассказать особо.

### У.Р.Гамильтон

Гамильтон обладал блестящими и многосторонними способностями. В десять лет он знал наизусть много стихов Гомера, в четырнадцать лет владел девятью языками, в 1824 г. опубликовал в трудах Королевской Ирландской Академии работу, посвященную геометрической оптике, в 1827 г. получил звание королевского астронома Ирландии.

К 1833 г. Гамильтон занимал пост директора обсерватории в Денсингке (около Дублина) и был известен как автор ряда работ по оптике и аналитической механике. Исходя из своих работ по геометрической оптике, Гамильтон предсказал эффект двойной конической рефракции в двуосных кристаллах, который вскоре был обнаружен его коллегой Ллойдом.

В течение долгих десяти лет Гамильтон безуспешно пытался придумать правило умножения триплетов. Позже в письме к сыну он вспоминал: «Каждое утро..., когда я спускался к завтраку, ты и твой брат Уильям Эдвин обычно спрашивали меня: «Ну как, папа, ты уже умножаешь триплеты?» На что я

всегда был вынужден печально отвечать: «Нет, я умею лишь складывать и вычитать их.»

## Векторное произведение

Задача, которую решал Гамильтон, поначалу казалась несложной. Как складывать векторы – ясно (по формуле (3)), остается «только» найти формулу их умножения – что-нибудь вроде формулы (2) для умножения дуплетов. Но все формулы, которые перепробовал Гамильтон, упорно не подходили – то нарушалось одно из обычных свойств, то другое.

Уже тогда была хорошо известна операция векторного произведения: *векторным произведением*  $\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$  ненулевых векторов  $\vec{v}_1, \vec{v}_2$  называется вектор, перпендикулярный плоскости, проходящей через векторы  $\vec{v}_1, \vec{v}_2$ , имеющий направление,

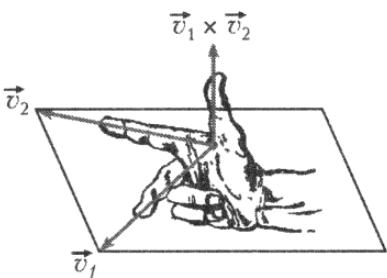


Рис. 3

определяемое правилом правой руки» (рис.3), и длину  $|\vec{v}_1| \cdot |\vec{v}_2| \cdot \sin(\vec{v}_1, \vec{v}_2)$ . Для дальнейшего заметим, что если векторы  $\vec{v}_1, \vec{v}_2$  заданы своими координатами в прямоугольной системе координат:

$$\vec{v}_1 = (\alpha_1; \beta_1; \gamma_1),$$

$$\vec{v}_2 = (\alpha_2; \beta_2; \gamma_2),$$

то

$$\vec{v}_1 \times \vec{v}_2 = (\beta_1 \gamma_2 - \beta_2 \gamma_1; \gamma_1 \alpha_2 - \gamma_2 \alpha_1; \alpha_1 \beta_2 - \alpha_2 \beta_1). \quad (4)$$

Но операция векторного произведения не годилась Гамильтону, поскольку она не имеет обратной. Например, если  $\vec{v}_1 = \vec{v}_2 \neq \vec{0}$ , то угол  $(\vec{v}_1, \vec{v}_2)$  между этими векторами равен нулю. Значит, длина векторного произведения  $\vec{v}_3 = \vec{v}_1 \times \vec{v}_2$  равна нулю, т.е. и сам вектор  $\vec{v}_3$  нулевой. Если бы операция деления на ненулевой вектор существовала всегда, то мы имели бы  $(\vec{v}_1 \times \vec{v}_2) : \vec{v}_2 = \vec{v}_1 \neq 0$ , в то время как  $\vec{v}_1 \times \vec{v}_2 = 0$  и, значит,  $(\vec{v}_1 \times \vec{v}_2) : \vec{v}_2 = 0$ . Получившееся противоречие показывает, что деление на  $\vec{v}_2$  невозможно.

Не может не вызвать уважения и восхищения то, что, несмотря на разочарование и неудачи, Гамильтон не оставлял надежды и в течение десяти лет с завидным упорством пытался решить поставленную перед собой задачу. И хотя задача так и не была решена (и не могла быть решена – почему, мы объясним позже), десятилетний труд не пропал даром. В один прекрасный

день 1843 года Гамильтон вдруг решил для определения умножения рассматривать не триплеты (тройки чисел), а четверки, или, как он их тут же окрестил, кватернионы. Вот как это произошло.

### Случай на Брокемском мосту

В одном из писем к своему сыну, написанном в свойственном тому времени несколько высокопарном стиле, Гамильтон вспоминал: «Это был 16-й день октября, который случился в понедельник, в день заседания Совета Королевской Ирландской Академии, где я должен был председательствовать. Я направлялся туда с твоей матерью вдоль Королевского канала; и, хотя она говорила мне какие-то отдельные фразы, я их почти не воспринимал, так как в моем сознании подспудно что-то творилось. Неожиданно как будто бы замкнулся электрический контур; блеснула искра, предвещающая многие длинные годы определенно направленной мысли и труда, моего – если доведется, или труда других, если мне будет даровано достаточно сознательной жизни, чтобы сообщить о своем открытии. Я оказался не в состоянии удержаться от желания высечь ножом на мягком камне Брокемского моста фундаментальную формулу о символах  $i, j, k$ ,

$$i^2 = j^2 = k^2 = ijk = -1,$$

содержащую решение проблемы, но, конечно, эта запись с тех пор стерлась. Однако более прочное упоминание осталось в Книге записей Совета Академии за этот день, где засвидетельствовано, что я попросил и получил разрешение на доклад о кватернионах на первом заседании сессии, который и был прочитан соответственно в понедельник 13-го следующего месяца – ноября».

### Определение кватернионов

*Кватернионы* – это четверки действительных чисел  $(x; y; u; v)$ , которые удобно записывать в виде

$$q = x + yi + uj + vk,$$

где  $i, j, k$  – новые числа, являющиеся аналогом мнимой единицы в комплексных числах. Требуется, чтобы числа  $i, j, k$  удовлетворяли следующим соотношениям:

$$i^2 = j^2 = k^2 = -1, \tag{5}$$

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j, \tag{6}$$

$x$	$i$	$j$	$k$
$i$	-1	$k$	$j$
$j$	$-k$	-1	$i$
$k$	$-j$	$-i$	-1

Рис. 4

которые удобно записать в виде «таблицы умножения» (рис.4).

По определению операции сложения и умножения кватернионов производятся по обычным правилам раскрытия скобок и приведения подобных членов с учетом правил (5)–(6).

Согласно этому определению, если  $q_1$  и  $q_2$  – два кватерниона, то

$$\begin{aligned}
 q_1 + q_2 &= (x_1 + y_1i + u_1j + v_1k) + (x_2 + y_2i + u_2j + v_2k) = \\
 &= x_1 + y_1i + u_1j + v_1k + x_2 + y_2i + u_2j + v_2k = \\
 &= (x_1 + x_2) + (y_1i + y_2i) + (u_1j + u_2j) + (v_1k + v_2k) = \\
 &= (x_1 + x_2) + (y_1 + y_2)i + (u_1 + u_2)j + (v_1 + v_2)k . \quad (7)
 \end{aligned}$$

Это, разумеется, привычное нам «покоординатное» сложение. Далее, произведение кватернионов  $q_1$  и  $q_2$  вычисляется так:

$$\begin{aligned}
 q_1q_2 &= (x_1 + y_1i + u_1j + v_1k)(x_2 + y_2i + u_2j + v_2k) = \\
 &= x_1x_2 + x_1y_2i + x_1u_2j + x_1u_2k + y_1x_2i + y_1y_2i^2 + y_1u_2ij + \\
 &\quad + y_1v_2ik + u_1x_2j + u_1y_2ji + u_1u_2j^2 + u_1v_2jk + \\
 &\quad + v_1x_2k + v_1y_2ki + v_1u_2kj + v_1v_2k^2 = \\
 &= x_1x_2 + x_1y_2i + x_1u_2j + x_1v_2k + y_1x_2i - y_1y_2 + y_1u_2k - y_1v_2j + \\
 &\quad + u_1x_2j - u_1y_2k - u_1u_2 + u_1v_2i + v_1x_2k + v_1y_2j - v_1u_2i - v_1v_2 = \\
 &= (x_1x_2 - y_1y_2 - u_1u_2 - v_1v_2) + (x_1y_2 + y_1x_2 + u_1v_2 - v_1u_2)i + \\
 &\quad + (x_1u_2 - y_1v_2 + u_1x_2 + v_1y_2)j + (x_1v_2 + y_1u_2 - u_1y_2 + v_1x_2)k . \quad (8)
 \end{aligned}$$

Длинная, но совершенно автоматическая проверка показывает, что умножение кватернионов обладает сочетательным свойством:

$$(q_1q_2)q_3 = q_1(q_2q_3) .$$

Естественно считать, что действительные и комплексные числа являются частным случаем кватернионов. Так, действительное число  $x$  – это кватернион вида

$$x = x + 0 \cdot i + 0 \cdot j + 0 \cdot k .$$

Комплексное число  $z = x + iy$  представляется как кватернион

$$z = x + iy = x + iy + 0 \cdot j + 0 \cdot k . \quad (9)$$

Читатель, не знакомый с комплексными числами, не должен смущаться: он может считать формулу (9), вместе с формулами (7) и (8),

определенением комплексного числа. Ему также будет полезно написать формулу умножения (8) для случая (9) и сравнить с формулой (2).

У операции сложения кватернионов, очевидно, имеется обратная операция – вычитание. Именно, разность двух кватернионов  $q_1$  и  $q_2$  определяется формулой

$$q_1 - q_2 = (x_1 - x_2) + (y_1 - y_2)i + (u_1 - u_2)j + (v_1 - v_2)k.$$

Если  $q_1 = q_2$ , то разность  $q_1 - q_2$  – это нулевой кватернион, равный

$$q_1 - q_2 = 0 + 0 \cdot i + 0 \cdot j + 0 \cdot k = 0.$$

### Деление кватернионов

Перейдем теперь к операции деления кватернионов, обратной к операции умножения. Вообще, что мы понимаем под частным от деления числа  $a$  на число  $b \neq 0$ ? Это такое число  $c$ , что

$$bc = a. \quad (10)$$

Так определяется частное от деления для действительных и комплексных чисел. К сожалению, для кватерниона применить непосредственно это определение мы не можем. И дело здесь вот в чем. Для того чтобы формула (10) «корректно» определяла частное, нужно, чтобы произведение не зависело от порядка сомножителей. В противном случае наряду с частным  $c = b^{-1}a$ , определенным формулой (10), существует вполне равноправное «левое частное»  $c'$ , определяемое формулой

$$c'b = a,$$

которое может отличаться от «правого частного»  $c$  из (10). Вот здесь, кроме необходимости выйти за пределы трехмерного пространства, Гамильтону пришлось принести еще одну жертву.

Оказывается, определенные им новые числа – кватернионы – потеряли еще одно привычное качество: произведение кватернионов зависит от порядка сомножителей. Действительно, уже в формулах (6) при изменении порядка множителей произведение меняет знак.

Таким образом, можно говорить лишь о «делении слева» и «делении справа». Как реально найти, скажем, «левое частное» от деления кватерниона  $q_1$ , на кватернион  $q_2 \neq 0$ ?

Обозначим искомое частное через  $q = x + yi + uj + vk$ . Тогда, используя правило умножения для кватернионов и определение

левого частного, получим следующее равенство кватернионов:

$$qq_2 = q_1,$$

или

$$\begin{aligned} (xx_2 - yy_2 - uu_2 - vv_2) + (xy_2 + yx_2 + uv_2 - vu_2)i + \\ + (xu_2 + ux_2 - yv_2 - vy_2)j + (xv_2 + vx_2 + yu_2 - uy_2)k = \\ = x_1 + y_1i + u_1j + v_1k. \end{aligned}$$

Полученное равенство равносильно системе четырех линейных уравнений с переменными  $x, y, u, v$ :

$$\begin{aligned} x_2x - y_2y - u_2u - v_2v &= x_1, \\ y_2x + x_2y + v_2u - u_2v &= y_1, \\ u_2x - v_2y + x_2u - y_2v &= u_1, \\ v_2x + u_2y - y_2u + x_2v &= v_1. \end{aligned}$$

Предлагаем читателю в качестве упражнения решить эти систему и тем самым найти «левое частное» от деления  $q_1$  на  $q_2$ . Аналогичным образом находится «правое частное» от деления  $q_1$  на  $q_2$ .

Рассмотрим частный случай, когда делимое  $q_1$  равно действительному числу 1. В этом случае частное от деления  $q_1 = 1$  на кватернион  $q_2$  (и «слева» и «справа») равно одному и тому же кватерниону

$$p = \frac{x_2 - y_2i - u_2j - v_2k}{x_2^2 + y_2^2 + u_2^2 + v_2^2}$$

(докажите!). Поэтому кватернион  $p$  обозначается через

$$q^{-1} = \frac{x_2 - y_2i - u_2j - v_2k}{x_2^2 + y_2^2 + u_2^2 + v_2^2}.$$

Тогда «правое частное» от деления кватерниона  $q_1$  на ненулевой кватернион  $q_2$  выражается формулой

$$q = q_2^{-1} \cdot q_1,$$

а «левое частное» от деления кватерниона  $q_1$  на  $q_2$  – формулой

$$q = q_1 \cdot q_2^{-1}.$$

Практически частное от деления двух кватернионов ищется другим путем. Для этого нам потребуются

## Скалярные и векторные кватернионы

Так же как комплексные числа разлагаются в сумму своей действительной и мнимой частей, кватернион

$$q = x + yi + uj + vk$$

тоже можно разложить в сумму

$$q = x + (\bar{r}i + \bar{v}j + \bar{u}k).$$

Первое слагаемое в этом разложении называется *скалярной частью* кватерниона, а второе слагаемое – *векторной частью*. Скалярная часть  $x$  – это просто действительное число, а векторная часть  $\bar{r}i + \bar{v}j + \bar{u}k$  может быть изображена вектором

$$\bar{r} = \bar{y}i + \bar{v}j + \bar{u}k$$

в трехмерном пространстве, где  $i, j, k$  мы теперь рассматриваем как единичные векторы прямоугольной системы координат (рис.5).

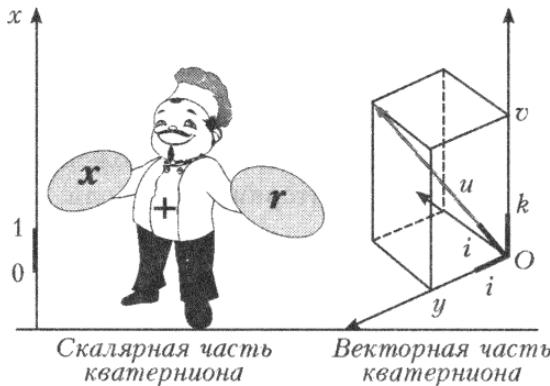


Рис. 5

Таким образом, каждый кватернион  $q$  представляется в виде суммы

$$q = x + \bar{r},$$

где  $x$  – скалярная часть кватерниона  $q$ , а  $\bar{r}$  – векторная часть. Если  $\bar{r} = 0$ , то  $q = x$  и кватернион  $q$  называется *скаларным кватернионом*. Если же  $x = 0$ , то  $q = \bar{r}$  и  $q$  называется *векторным кватернионом*.

При сложении кватернионов независимо складываются их скалярные и векторные части.

При умножении кватернионов дело обстоит сложнее. Если  $q_1$  и  $q_2$  – скалярные кватернионы, то их произведение  $q_1q_2$  тоже

скалярный кватернион. В случае, когда  $q_1 = x$  – скалярный кватернион, а  $q_2 = \vec{r}$  – векторный кватернион, произведение

$$q_1 q_2 = x \cdot (yi + uj + vk) = (xy)i + (xu)j + (xv)k$$

является векторным кватернионом и операция умножения совпадает с умножением вектора  $\vec{r}$  в пространстве на действительное число  $x$ .

И наконец, если оба кватерниона векторные:

$$q_1 = \vec{r}_1 = y_1 i + u_1 j + v_1 k,$$

$$q_2 = \vec{r}_2 = y_2 i + u_2 j + v_2 k,$$

то

$$q_1 q_2 = -(y_1 y_2 + u_1 u_2 + v_1 v_2) +$$

$$+ (u_1 v_2 - v_1 u_2) i + (v_1 y_2 - y_1 v_2) j + (y_1 u_2 - u_1 y_2) k.$$

Как видно из последней формулы, скалярная часть произведения  $q_1 q_2$  равна скалярному произведению  $(\vec{r}_1, \vec{r}_2)$  векторов  $\vec{r}_1$ , и  $\vec{r}_2$  с обратным знаком. Векторная же часть  $q_1 q_2$  – это наш старый знакомый: векторное произведение  $\vec{r}_1 \times \vec{r}_2$ , записанное в координатах (см. (4)).

Объединяя все рассмотренные случаи, получим общую формулу для умножения кватернионов. Если  $q_1 = x_1 + \vec{r}_1$  и  $q_2 = x_2 + \vec{r}_2$ , то

$$q_1 q_2 = (x_1 x_2 - \vec{r}_1 \cdot \vec{r}_2) + (x_1 \vec{r}_2 + x_2 \vec{r}_1 + \vec{r}_1 \times \vec{r}_2).$$

### **А как же триплеты?**

Почему Гамильтону все-таки не удалось найти удовлетворительного способа для умножения триплетов? Не из-за недостаточного остроумия или трудолюбия – выше мы уже отмечали, что задачу построения «трехмерных чисел» решить нельзя. Действительно, доказано, что попросту не существует способа умножения точек пространства, удовлетворяющего нашим требованиям (ассоциативности, дистрибутивности относительно покоординатного сложения, возможности деления на ненулевые элементы). Более того, сейчас известны все случаи, когда можно ввести такое умножение. Как доказал немецкий математик Ф.Г.Фробениус (1849–1917), этих случаев три: в размерности один (обычные действительные числа), в размерности два (комплексные числа) и в «размерности четыре» (кватернионы).

## Что было дальше

Гамильтон и его последователи возлагали большие надежды на кватернионы. От кватернионов ожидали таких же результатов, как от комплексных чисел, и даже больше. И действительно, с помощью исчисления кватернионов были обнаружены совершенные в их математической красоте формулы, описывающие ряд важных физических явлений. Но дальнейшие надежды на развитие алгебраического и функционального исчисления кватернионов не оправдались.

Для кватернионов не имеет места основная теорема алгебры о существовании корней у многочлена с кватернионными коэффициентами, а, с другой стороны, существует такой многочлен с кватернионными коэффициентами от одной переменной, для которого любой кватернион является корнем.

Оптимизм сменился скепсисом. В начале нашего века математики перестали интересоваться кватернионами. Но время шло, и физики упорно искали математический формализм для некоторых эффектов, связанных с так называемым спином элементарных частиц. Кватернионы снова получили признание, когда была понята их роль в построении различных геометрических преобразований пространства, используемых в квантовой физике. Геометрические свойства кватернионов – это особая большая тема, о которой мы надеемся рассказать в дальнейшем.

# ДЕМОКРАТИЯ С ТОЧКИ ЗРЕНИЯ МАТЕМАТИКИ

---

**В.Пахомов**

Считается, что демократия – это искусство политического маневрирования. Наверное, во многом это так, но далеко не во всем. Цель этой статьи – показать, что и математика играет не последнюю роль в политике. Обсудим проблему выбора, которую мы будем понимать достаточно широко. Это может быть выбор должностного лица, лауреата какого-нибудь конкурса, выбор проекта (конституции, закона, освоения нового региона, стратегии политической или экономической деятельности и т.п.) и так далее. Главное условие – выбирается один вариант из нескольких альтернативных и обязательно путем голосования. В основе демократических принципов и лежит это правило – наиболее важные решения принимаются путем прямого голосования самых широких слоев общества. Казалось бы, при таком подходе к делу все должны быть удовлетворены, однако на практике получается, что каждый раз после очередного голосования число недовольных вполне сравнимо (если не превышает) с числом удовлетворенных его результатом. Случайно ли это? Попробуем рассмотреть этот вопрос детально.

## Некоторые правила голосования

Несколько слов о терминологии. Пусть имеется  $n$  избирателей и  $m$  кандидатов, например, в депутаты, президенты, мэры и т.д. Каждый из избирателей упорядочивает кандидатуры, определяя для себя наиболее предпочтительного кандидата, затем занимающего второе, третье и т.д. места. Например, если избиратель с номером 6 считает, что наилучшим является кандидат  $a$ , затем идет  $b$ , а затем  $c$ , то мы это запишем в виде его *системы индивидуальных предпочтений* так:

$$a \succ b \succ c$$

и будем говорить, что  $a$  лучше  $b$ ,  $b$  лучше  $c$  с точки зрения 6-го избирателя. Сведя эти системы индивидуальных предпочтений в одну итоговую таблицу, мы получим *профиль голосования*. Например, пусть число избирателей  $n = 17$ , а кандидатов четверо:  $a$ ,  $b$ ,  $c$  и  $d$ ; и пусть пять избирателей упорядочивают кандидатов

---

Опубликовано в «Кванте» №9, 10 за 1992 г.

так:  $a \succ d \succ c \succ b$ , трое других:  $a \succ d \succ b \succ c$ , еще пятеро:  $b \succ c \succ d \succ a$ , а оставшиеся четверо так:  $c \succ d \succ b \succ a$ . В таком случае профиль голосования выглядит следующим образом:

Профиль A

Количество голосов	5	3	5	4
Кандидаты	$d$	$d$	$b$	$c$
	$c$	$b$	$c$	$d$
	$b$	$c$	$d$	$b$

Рассмотрим теперь некоторые правила голосования.

1. *Правило относительного большинства.* Каждый избиратель отдает ровно один голос за своего кандидата. Побеждает тот, кто получит наибольшее количество голосов.

Для профиля A это означает, что при голосовании  $a$  получит 8 голосов,  $b$  – 5 голосов,  $c$  – 4 голоса. Следовательно, по этому правилу побеждает  $a$ .

2. *Правило абсолютного большинства.* Каждый избиратель отдает ровно один голос за своего кандидата. Набравший более половины голосов побеждает. Если никто не набрал более половины голосов, то проводится второй тур голосования. При этом во второй тур выходят два кандидата, набравшие наибольшее количество голосов. Во втором туре побеждает тот, кто набрал большинство голосов (а следовательно, и более половины числа отданных голосов).

Что это означает для нашего профиля A?

В первом туре число голосов, поданное за кандидатов, равно  $n_a = 8$ ,  $n_b = 5$ ,  $n_c = 4$ . Следовательно, во второй тур проходят  $a$  и  $b$ . Вычеркнув из профиля A кандидатов  $c$  и  $d$ , не прошедших во второй тур, мы получим следующий профиль второго тура:

5	3	5	4
$a$	$a$	$b$	$b$
$b$	$b$	$a$	$a$

Это означает, что во втором туре выигрывает кандидат  $b$ , за которого подано 9 голосов, тогда как за  $a$  – лишь 8.

К сожалению, это правило не всегда определяет победителя.

3. *Правило Бордá.* Каждый избиратель дает нуль очков кандидату, находящемуся на последнем месте, одно очко – предпоследнему, два очка – находящемуся на третьем месте с конца и т.д. Побеждает кандидат, набравший наибольшую сумму очков.

(Это правило часто используется с небольшой модификацией на спортивных соревнованиях: каждый судья определяет место спортсмена, и побеждает набравший наименьшую сумму мест.)

Для профиля  $A$  имеем:

Число избирателей		5	3	5	4
Очки за места					
3		$a$	$a$	$b$	$c$
2		$d$	$d$	$c$	$d$
1		$c$	$b$	$d$	$b$
0		$b$	$c$	$a$	$a$

Это означает, что число очков  $n_a$ , набранных кандидатом  $a$ , равно  $n_a = 8 \times 3 + 9 \times 0 = 24$ , аналогично  $n_b = 5 \times 3 + 7 \times 1 + 5 \times 0 = 22$ ,  $n_c = 4 \times 3 + 5 \times 2 + 5 \times 1 + 3 \times 0 = 27$  и  $n_d = 12 \times 2 + 5 \times 1 = 29$ .

Таким образом, побеждает  $d$ , на втором месте оказывается  $c$ , а победители по предыдущим правилам занимают наихудшие места!

Взглянем теперь на голосование с другой стороны. Что, например, представляет собой с точки зрения избирателей кандидат  $d$ ? Из профиля  $A$  видно, что 8 избирателей считают его хуже  $a$ , а 9 – лучше  $a$ . Иначе говоря,  $d$  выигрывает дуэль с  $a$  девятью голосами против восьми, что мы запишем так:  $d \succ a - 9:8$ .

Аналогично в сравнении с  $b$  имеем – 12:5, т.е. еще больше  $d$  выигрывает в сравнении с победителем голосования по правилу абсолютного большинства.

Результаты других дуэлей:

$$\begin{aligned} b \succ a & - 9:8; \quad c \succ a - 9:8; \\ c \succ b & - 9:8, \quad c \succ d - 9:8. \end{aligned}$$

4. *Правило Кондорсе.* Победителем по Кондорсе называется такой кандидат  $x$ , который выигрывает в парных сравнениях у всех остальных кандидатов.

В профиле  $A$  победитель по Кондорсе – кандидат  $c$ . Отметим, что победителя по Кондорсе может и не оказаться.

Рассмотрим, наконец, еще одно правило, обобщающее два из уже перечисленных.

5. *Правило с подсчетом очков.* Пусть число кандидатов равно  $m$ . Фиксируем числа  $s_1, s_2, \dots, s_m$  такие, что  $0 = s_1 \leq s_2 \leq \dots \leq s_m$ , где  $s_m > 0$ . Избиратели дают  $s_1$  очков своим последним кандидатам,  $s_2$  – предпоследним,  $s_3$  – третьим с конца, ...,  $s_m$  – первым кандидатам. Побеждает кандидат, набравший максимальную сумму очков.

При  $s_1 = s_2 = \dots = s_{m-1} = 0$  и  $s_m = 1$  мы получим правило относительного большинства, а при  $s_1 = 0, s_2 = 1, \dots, s_m = m$  – правило Борда. Таким образом, разные шкалы голосования  $s_1, s_2, \dots, s_m$  могут выводить разных кандидатов в победители.

### Парадоксы голосования

Первые четыре правила дают четыре различных представления о понятии «наилучшего» (с точки зрения коллектива) выбора». Приводят они, как мы уже видели, к различным результатам. Причем победитель по одному из правил может оказаться наихудшим по другому. Например, для профиля  $A$  победитель по правилу относительного большинства является наихудшим по Кондорсе, а победитель по правилу абсолютного большинства – наихудшим по Борда. Поэтому при нашем профиле любой способ голосования даст результат, неудовлетворительный более чем для половины избирателей (которые, естественно, будут анализировать его с разных точек зрения). Правило 5 достаточно гибкое, и может показаться, что при надлежащем подборе шкалы оно сможет заменить любое из первых четырех. Действительно, правило относительного голосования и правило Борда следуют из него как частные случаи. Далее, пусть  $s_1 = 0, s_2 = 4, s_3 = 9, s_4 = 16$ , тогда победителем для  $A$  по этому правилу становится  $c$  – победитель по Кондорсе. Тем не менее это правило существенно отличается от правил абсолютного большинства и Кондорсе. Именно, верны следующие два утверждения:

**Лемма 1°.** *Существуют такие профили голосования, что победитель по правилу абсолютного большинства не может быть победителем ни при каком подсчете очков.*

**Лемма 2°.** *Существуют такие профили голосования, что победитель по Кондорсе не может быть победителем ни при каком подсчете очков.*

**Доказательство леммы 1°.** Его дает профиль  $A$ . Здесь победитель по правилу абсолютного большинства – кандидат  $b$ . Подсчитаем очки каждого кандидата для шкалы  $0 = s_1 \leq s_2 \leq$

$\leq s_3 \leq s_4$ ,  $s_4 > 0$ :

$$n_a = 8s_4,$$

$$n_b = 5s_4 + 7s_2,$$

$$n_c = 4s_4 + 5s_3 + 5s_2,$$

$$n_d = 12s_3 + 5s_2.$$

Если  $b$  – победитель, то

$$n_b - n_a = -3s_4 + 7s_2 > 0,$$

$$n_b - n_c = s_4 - 5s_3 + 2s_2 > 0.$$

Из этих неравенств получаем

$$\frac{7}{3}s_2 > s_4 > 5s_3 - 2s_2,$$

откуда вытекает, что

$$s_2 > \frac{15}{13}s_3.$$

С другой стороны,  $s_2 \leq s_3$ , так что

$$s_3 \geq s_2 > \frac{5}{13}s_3,$$

а значит,

$$\frac{2}{13}s_3 < 0.$$

Последнее неравенство противоречит условию  $s_3 \geq 0$ . Выходит,  $b$  не может быть победителем по правилу 5.

Для доказательства леммы 2° рассмотрим такой профиль  $B$ :

Профиль  $B$

	3	6	4	4
$s_3$	$c$	$a$	$b$	$b$
$s_2$	$a$	$b$	$a$	$c$
0	$b$	$c$	$c$	$a$

Здесь  $a \succ b = 9:8$  и  $a \succ c = 10:7$ , т.е.  $a$  – победитель по Кондорсе. Подсчитаем очки каждого кандидата:

$$n_a = 6s_3 + 7s_2,$$

$$n_b = 8s_3 + 6s_2,$$

$$n_c = 3s_3 + 4s_2,$$

откуда видно, что  $n_b - n_a = 2s_3 - s_2 = s_3 + (s_3 - s_2) > 0$ , так как  $s_3 \geq s_2$  и  $s_3 > 0$ . Значит,  $a$  проигрывает  $b$  по очкам при любом подборе шкалы.

Рассмотрим еще несколько парадоксов. Как мы уже знаем, при голосовании по правилу абсолютного большинства для профиля  $A$  побеждает  $b$ . Кандидат  $a$ , не имеющий шансов выиграть, может снять свою кандидатуру. В таком случае мы получим следующий профиль голосования:

Профиль  $C$

5	3	5	4
$d$	$d$	$b$	$c$
$c$	$b$	$c$	$d$
$b$	$c$	$d$	$b$

Теперь во второй тур выходят  $d$  и  $b$  (8 и 5 голосов соответственно), и во втором туре:

12	5
$d$	$b$
$b$	$d$

побеждает  $d$ ! Итак,  $a$ , отказавшись от участия в выборах, тем самым проваливает своего соперника  $b$ .

Аналогичный пример для правила Борда дает такой профиль:

Профиль  $D$

	2	1
4	$b$	$d$
3	$a$	$c$
2	$d$	$e$
1	$c$	$b$
0	$e$	$a$

Здесь выигрывает  $b$  с 9 очками, за ним следует  $d$  с 8 очками. Пусть теперь  $a$  снимает свою кандидатуру, тогда получаем

следующий профиль:

Профиль  $E$

	2	1
3	$b$	$d$
2	$d$	$c$
1	$c$	$e$
0	$e$	$b$

где побеждает  $d$  с 7 очками (у  $b$  теперь всего 6 очков, у  $c$  – 4, у  $e$  – 1 очко).

И наконец, еще более эффектный парадокс. Рассмотрим два профиля:

Профиль  $F$

6	5	4	2
$a$	$c$	$b$	$b$
$b$	$a$	$c$	$a$
$c$	$b$	$a$	$c$

Профиль  $G$

6	5	4	2
$a$	$c$	$b$	$a$
$b$	$a$	$c$	$b$
$c$	$b$	$a$	$c$

У этих двух профилей три первые колонки одинаковы, а последние отличаются тем, что в профиле  $G$  положение  $a$  улучшается по сравнению с профилем  $F$ , а положение  $b$  – ухудшается. Можно себе представить, что второй профиль получился из первого после того, как два избирателя, сдленные в последнюю колонку, поменяли свое мнение в пользу  $a$  против  $b$ .

При голосовании по правилу абсолютного большинства для профиля  $F$  в первом туре побеждают  $a$  и  $b$ , а во втором – выигрывает  $a$ . Для профиля  $G$  имеем: после первого тура остаются  $a$  и  $c$ , а во втором туре выигрывает  $c$ . Таким образом, голосуя за  $a$  вместо  $b$ , упомянутые два избирателя проваливают  $a$ !

Можно было бы привести еще массу правил голосования и парадоксов, чтобы убедиться в отсутствии абсолютно хорошей демократической системы. Все это наводит на печальную для многих мысль, что демократии как волеизъявления большинства не существует, ибо не существует, как мы видели, самого понятия «мнение большинства»...

## **Функция коллективного предпочтения**

Поставим несколько более общую задачу. Пусть  $n$  членов парламента определяют систему приоритетов государственного финансирования  $m$  различных программ. Скажем, в первую очередь средства выделяются на здравоохранение, из оставшейся части бюджета – на программу образования, затем из остатка – на программу строительства жилья и т.д. У каждого члена парламента есть свое мнение по поводу упорядочения этих статей расхода. Необходимо построить их упорядочение, отражающее мнение парламента в целом.

Поскольку между упорядочением государственных программ и упорядочением кандидатов нет никакой разницы с математической точки зрения, мы по-прежнему будем говорить для удобства об избирателях и кандидатах вместо парламентариев и программ. Итак, имеется  $n$  избирателей и  $m$  кандидатов. Требуется построить функцию, определяющую коллективный порядок на множестве кандидатов, т.е. правило, которое для любых заданных порядков  $\succ^1, \succ^2, \dots, \succ^n$  определяет коллективный порядок  $\succeq$ :

$$\succeq = f(\succ^1, \succ^2, \dots, \succ^n).$$

Здесь мы предполагаем три возможности при парных сравнениях кандидатов, скажем,  $a$  и  $b$ :

$$a \succ b \quad (a \text{ лучше } b),$$

$$a \overset{s}{=} b \quad (a \text{ и } b \text{ одинаковы}),$$

$$a \prec b \quad (a \text{ хуже } b),$$

и их комбинации типа

$$a \succeq b \quad (a \text{ не хуже } b).$$

Таких функций может существовать много. Например,  $f$  может объявлять всех кандидатов равными. Поэтому необходимо наложить на  $f$  какие-то ограничения. Прежде всего она должна быть определена для любого (конечного) числа кандидатов. Например, для двух кандидатов все перечисленные правила голосования совпадают, поэтому функция  $f$  в этом случае строится легко: либо кто-то из двух кандидатов выигрывает и, значит, занимает первое место (а противник – второе), либо

выясняется, что они равны. Например, для профиля

Избиратель	1	2	3	4	5	6
Кандидаты	$a$ $b$	$a$ $b$	$b$ $a$	$a$ $b$	$b$ $a$	$a$ $b$

эквивалентного профилю

Группа избирателей	$\{1, 2, 4, 6\}$	$\{3, 5\}$
Кандидаты	$a$ $b$	$b$ $a$

имеем  $a \overset{s}{\succ} b$  по любому из правил голосований. Для профиля

Группа избирателей	$\{1, 2, 5\}$	$\{3, 4, 6\}$
Кандидаты	$a$ $b$	$b$ $a$

ни одно из наших правил не дает победителей, поэтому следует считать  $a \overset{s}{=} b$ .

Однако эти правила следует еще доопределять для случаев, когда число кандидатов более двух. Простейший способ заключается в следующем.

Пусть  $M$  – множество кандидатов, а  $p$  – правило выборов. Обозначим победителей (их может быть несколько) символом  $p(M)$ . Они займут первое место. Затем исключим из  $M$  этих победителей, в результате получим новое множество  $M_1 = M \setminus p(M)$ . Применяя к нему правило  $p$ , получим множество  $p(M_1)$  победителей из  $M_1$ , они займут второе место. Исключая из  $M_1$  этих победителей, получим множество  $M_2 = M_1 \setminus p(M_1)$  и т.д. В результате получаем функцию  $f$ .

Разберем несколько примеров.

Пусть  $p$  – правило относительного большинства. Рассмотрим профиль  $H$  ( $n = 3, M = \{a, b, c, d\}$ ):

Профиль  $H$

Избиратель	1	2	3
Кандидаты	$a$	$d$	$b$
	$b$	$a$	$c$
	$c$	$b$	$d$
	$d$	$c$	$a$

Правило относительного большинства дает по одному голосу для  $a, d, b$ , следовательно,  $p(M) = \{a, b, d\}$ . Они занимают первое место, а оставшийся кандидат  $c$  займет второе. Значит,  $f$  сопоставляет профилю  $H$  порядок

$$a \stackrel{s}{=} b \stackrel{s}{=} d \succ c$$

Рассмотрим эту же функцию на профиле  $A$ . Имеем  $p(A) = \{a\}$ ,  $A_1 = \{b, c, d\}$ , а соответствующий профиль  $A_1$  даст  $p(A_1) = \{d\}$  и  $A_2 = \{b, c\}$ :

Профиль  $A_1$

Количество голосов	5	3	5	4
Кандидаты	$d$	$d$	$b$	$c$
	$c$	$b$	$c$	$d$
	$b$	$c$	$d$	$b$

Профиль  $A_2$

Количество голосов	5	3	5	4
Кандидаты	$c$	$b$	$b$	$c$
	$b$	$c$	$c$	$b$

Наконец, из профиля  $A_2$  видно, что  $p(A_2) = \{c\}$ , т.е.  $f$  задает порядок

$$a \succ d \succ c \succ b.$$

Правило абсолютного большинства порождает свою функцию  $f$ , которая, правда, не является определенной для всех профилей (например, для профиля  $H$ ), а для профиля  $A$  дает (проверьте!) порядок  $b \stackrel{s}{>} c \stackrel{s}{>} d \stackrel{s}{>} a$ , в точности противоположный порядку по правилу относительного большинства.

Правило Кондорсе также определено не для всех профилей, но для профиля  $A$  дает (проверьте!) порядок  $c \succ^S d \succ^S b \succ^S a$ .

Наконец, правило Борда при нашем способе определяет функцию  $f$ , которая профилю  $A$  сопоставляет порядок  $d \succ^S c \succ^S b \succ^S a$ .

Заметим, что правило Борда может и непосредственно определять функцию  $f$ : достаточно расставить кандидатов в порядке убывания их очков. Это дает для профиля  $A$  порядок  $d \succ^S c \succ^S a \succ^S b$ . Как видите, при таком определении порядка худшим оказался  $b$ .

Итак, все правила достаточно разумны, но приводят к совершенно различным результатам, вплоть до противоположных. Чтобы добиться однозначного результата, попробуем предъявить некоторые требования к функции коллективного предпочтения, которые сформулируем в виде аксиом. Обозначим для удобства множество избирателей через  $S$ .

**Аксиома 1** (полнота). Для любых кандидатов  $a$  и  $b$  коллективный порядок устанавливает, что либо  $a \succ^S b$ , либо  $a = b$ , либо  $a \prec^S b$ .

**Аксиома 2** (транзитивность). Если  $a \succ b$  и  $b \succ c$ , то  $a \succ c$ .

Первая из аксиом говорит о том, что мы можем сравнивать любых двух кандидатов, а вторая естественна для любого порядка, без нее может возникнуть ситуация, в которой  $a$  лучше  $b$ ,  $b$  лучше  $c$  и  $c$  лучше  $a$ .

**Аксиома 3** (единогласие). Если все избиратели считают, что  $a$  лучше  $b$ , то и в коллективном предпочтении  $a$  лучше  $b$ .

Эта аксиома не вызывает сомнения и не позволяет строить такое правило, при котором «все равны».

**Аксиома 4** (независимость). Положение любых двух кандидатов  $a$  и  $b$  зависит только от их взаимного расположения в индивидуальных предпочтениях избирателей и не зависит от расположения других кандидатов.

Например, профили

Избиратели	1	2	3	4
Кандидаты	$a$	$c$	$b$	$a$
	$b$	$a$	$c$	$b$
	$c$	$b$	$a$	$c$

Избиратели	1	2	3	4
Кандидаты	$a$	$a$	$b$	$c$
	$c$	$c$	$a$	$a$
	$b$	$b$	$c$	$b$

вообще говоря, различны, но относительное положение  $a$  и  $b$  у всех избирателей одинаково: избиратели 1, 2 и 4 считают  $a$  лучше  $b$ , и только третий избиратель считает  $b$  лучше  $a$ . В этом случае функция коллективного предпочтения должна и для одного и для другого профиля расположить  $a$  и  $b$  одинаково. Это могут быть, скажем, порядки  $a > b > c$  и  $a > c > b$  соответственно (они даются здесь правилом Борда). Отказ от аксиомы независимости приводит к возможности манипуляций: как мы уже видели, снятием (а значит, и введением) кандидатуры можно изменить взаимное расположение остальных кандидатов для правила Борда (сравните профили  $D$  и  $E$ ).

Приведем один пример функции коллективного предпочтения, удовлетворяющей всем этим аксиомам:

$$f\left(\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\right) = \frac{1}{n}.$$

Эта функция задает *правило диктатора*: независимо от предпочтения других избирателей она устанавливает тот порядок, который определяется первым избирателем. Разумеется, вместо первого можно было бы взять второго, третьего,  $n$ -го избирателя – суть дела не меняется. Меняется лишь диктатор.

### Теорема Эрроу

Неожиданный для многих результат Эрроу заключается в том, что при всей разумности наших аксиом *единственное правило, которое им удовлетворяет, и есть правило диктатора*.

Для доказательства этой теоремы введем несколько важных понятий. Подмножество  $A$  множества избирателей  $S$  назовем *коалицией*. Будем называть коалицию *f-решающей для кандидата a против кандидата b* тогда и только тогда, когда из того, что все члены коалиции ставят  $a$  выше  $b$ , а все, не входящие в нее, ставят  $b$  выше  $a$ , следует  $a \stackrel{f}{>} b$ . Другими словами, для профилей вида

Группа избирателей	Кандидаты
$A$	$\dots a \dots b \dots$
$S \setminus A$	$\dots b \dots a \dots$

где точки означают произвольное расположение других кандидатов, функция  $f$  определяет порядок типа

$$\dots \succ^S a \succ^S \dots \succ^S b \succ^S \dots$$

Этот факт будем кратко записывать так:  $A = f(a, b)$ . Отметим, что  $A$  зависит от конкретных кандидатов  $a$  и  $b$  и не обязана быть  $f$ -решающей для  $c$  против  $d$ . Наконец, если для любых двух кандидатов  $x$  и  $y$  коалиция  $A$  является  $f$ -решающей для  $x$  против  $y$ , будем называть ее  $f$ -решающей.

Для  $f$ -решающей коалиции теперь не важны имена кандидатов: если члены коалиции ставят кого-либо выше другого, а все остальные избиратели располагают этих кандидатов в точности наоборот (а не как-нибудь!), то в групповом предпочтении выше стоит тот, кого выше поставила коалиция.

Отметим, что все множество избирателей  $S$  является «пределным» случаем  $f$ -решающей коалиции, ибо для него это определение, безусловно, верно в силу аксиомы единогласия.

По этой же причине пустое множество не может быть  $f$ -решающим для  $a$  против  $b$  ни при каких  $a$  и  $b$ : если никто не ставит  $a$  выше  $b$ , т.е. все ставят  $b$  выше  $a$ , то в силу аксиомы единогласия  $a \prec^S b$ , и, значит, не может быть  $a \succ^S b$ .

**Лемма 1.** Существует пара  $(a, b)$ , для которой найдется коалиция  $D$ , состоящая из одного избирателя  $\{d\}$ , такая, что  $d$  является  $f$ -решающей для  $a$  против  $b$ .

**Доказательство.** Обозначим через  $F$  множество таких коалиций  $A$ , для каждой из которых существует такая пара  $(a, b)$ , что  $A = f(a, b)$  (для каждой коалиции из  $F$  – своя).

Множество  $F$  не пусто, так как включает множество всех избирателей  $S$ . Возьмем коалицию  $D$  из множества  $F$ , содержащую наименьшее число избирателей. Так как  $F$  не пусто, то  $D$  содержит не менее одного избирателя. Требуется доказать, что он там и будет ровно один. Таким образом, если предположение о том, что их там не менее двух, приведет к противоречию, то лемма будет доказана.

Итак, предположим, что  $D$  можно разбить на два непересекающихся непустых подмножества: одно  $\{d\}$ , состоящее из одного элемента, а другое  $E$  – из всех остальных.

Рассмотрим теперь профиль

Группа избирателей	$\{d\}$	$E$	$S \setminus D$
Кандидаты	$a$	$c$	$b$
	$b$	$a$	$c$
	$c$	$b$	$a$

Так как коалиция  $D$  –  $f$ -решающая для  $a, b$ , то  $a \succ^s b$ . Если  $c \succ^s b$ , то это означает, что коалиция  $E$  является  $f$ -решающей для  $c$  против  $b$ . Но в  $E$  меньше избирателей, чем в  $D$ , которая по определению минимальна. Полученное противоречие показывает, что  $c$  не может быть не хуже  $b$ . Значит, в силу аксиомы полноты,  $b$  лучше  $c$ .

Теперь имеем  $a \succ^s b$ ,  $b \succ^s c$ , откуда, в силу аксиомы транзитивности, получаем  $a \succ^s c$ . Но это значит, что  $\{d\}$  является  $f$ -решающей для  $a$  против  $c$ , что также противоречит минимальности  $D$ .

Таким образом, мы пришли к противоречию; и лемма доказана.

**Лемма 2.** Коалиция  $D = \{d\}$ , существование которой доказано в предыдущей лемме, является  $f$ -решающей.

**Доказательство.** Пусть  $c$  – произвольный кандидат. Рассмотрим профиль

Группа избирателей	Кандидаты
$\{d\}$	$\dots a \dots b \dots c \dots$
$S \setminus \{d\}$ ( $S$ без $d$ )	$\dots b \dots c \dots a \dots$

( $d$  ставит  $a$  выше  $b$ , а  $b$  – выше  $c$ , все остальные ставят  $b$  выше  $c$ , а  $c$  выше  $a$ ). Так как  $\{d\} = f(a, b)$ , то  $a \succ^s b$ ; в силу аксиомы единогласия, имеем  $b \succ^s c$ , а значит, по аксиоме транзитивности,  $a \succ^s c$ . В соответствии с аксиомой независимости наш результат не зависит от  $b$ , т.е. для всякого профиля вида

Группа избирателей	Кандидаты
$\{d\}$	$\dots a \dots c \dots$
$S \setminus \{d\}$	$\dots c \dots a \dots$

имеем  $a \succ^s c$ , значит,  $\{d\} = f(a, c)$ , где кандидат  $c$  уже произведен.

Аналогично для произвольного кандидата  $e$ , взяв профиль вида

Группа избирателей	Кандидаты
$\{d\}$	$\dots e \dots a \dots c \dots$
$S \setminus \{d\}$ ( $S$ без $d$ )	$\dots c \dots e \dots a \dots$

и повторив все рассуждения, получим, что  $e \succ^s c$ . В конечном счете,  $\{d\} = f(c, e)$  для произвольных  $e$  и  $c$ , т. е. является  $f$ -решающей.

**Лемма 3.** *Описанный выше избиратель  $d$  – диктатор.*

К этому моменту мы получили, что  $d$  может навязывать свое мнение по поводу любых двух кандидатов  $a$  и  $b$  при условии, что мнение остальных избирателей противоположно. В этом пока проявляется зависимость и от мнения других. Мы должны доказать, что, как бы ни голосовали остальные избиратели, коллективное мнение совпадает с мнением  $d$ .

**Доказательство леммы 3.** Рассмотрим такие профили голосования, в которых у избирателя  $d$  порядок вида  $\dots \succ^d a \succ^d \dots \succ^d c \succ^d \dots \succ^d b \succ^d \dots$ , все остальные кандидаты ставят с выше  $a$  и  $b$ , а в остальном расположение их кандидатов произвольно.

Так как коалиция  $\{d\}$  –  $f$ -решающая, то  $a \succ^s c$ ; в силу аксиомы единогласия,  $c \succ^s b$ , значит, по аксиоме транзитивности,  $a \succ^s b$ . Исключая из соотношений  $c$  (в силу аксиомы независимости), получаем, что если  $d$  предпочитает кандидату  $b$  кандидата  $c$ , то ему и будет оказано коллективное предпочтение:

$$a \succ^d b \Rightarrow a \succ^s b$$

(независимо от расположения  $a$  и  $b$  у других избирателей!).

Меняя ролями  $a$  и  $b$ , получаем

$$a \overset{d}{\prec} b \Rightarrow a \overset{d}{\prec} b.$$

Итак, порядки  $\overset{d}{\succ}$  и  $\overset{s}{\succeq}$  совпадают:

$$a \overset{d}{\succ} b \Leftrightarrow a \overset{s}{\succ} b.$$

Тем самым теорема Эрроу полностью доказана.

### **Некоторые житейские выводы из теоремы Эрроу**

Что же мы получили в итоге? Выходит, введя четкие и разумные правила, обеспечивающие, на первый взгляд, выбор действительно того кандидата, которого предпочитает большинство, мы получили ...диктатуру. Четкая аксиоматика погубила демократию.

Итак, диктатура – понятие достаточно ясное, удовлетворяющее простым принципам, понятие же демократии невозможно сформулировать иначе, как «альтернатива диктатуре». Несмотря на это, демократия привлекает людей, являемся естественной политической средой для развития общества. Только в ней может происходить естественный отбор «сильнейших» и «мудрейших», ибо победа в выборах или в проведении своего решения зачастую требует, кроме умения убеждать, умения рассчитывать. Но в этом и ее бич. С помощью расчета слишком часто можно манипулировать результатами.

Непротиворечивая же самодостаточная система диктатуры – мертва: ей развиваться некуда.

И еще один житейский вывод. Вы не задумывались, почему настолько популярна во многих демократических государствах именно двухпартийная система? Одну из причин этому мы упоминали: для двух кандидатур различные правила голосования дают один и тот же результат, и большинство противоречий тем самым снимается.

### **И наконец, комментарий...**

В этой статье упоминаются три человека, о которых мы хотим сказать несколько слов.

Сведения о Жане Шарле Борда, которыми мы располагаем, таковы. Родился в 1733, умер в 1799 году. Следовательно, пережил Французскую революцию и видел восхождение Бонапарта. Был физиком и геодезистом, членом французской Академии. Участвовал в разработке метрической системы мер. В 1792 г. определил длину секундного

маятника в Париже и нашел способ точного определения периода колебаний маятника. Занимался также гидравликой; доказал теорему, носящую теперь его имя, об ударе струи жидкости или газа. Оставил математические работы по дифференциальным уравнениям и вариационному исчислению. Политикой занимался во время революции; именно тогда возникла настоятельная необходимость разработать демократические процедуры, и Борда предложил свою систему голосования.

Гораздо более известен Жан Антуан Никола де Кондорсе (1745–1794), маркиз, энциклопедист, философ, математик, социолог, политик. Друг Даламбера и Вольтера. Разумеется, член французской Академии. В 1776 году был избран членом Петербургской академии наук, но в 1792 году исключен повелением Екатерины II, не одобравшей (мягко говоря) его политических взглядов. Погиб Кондорсе трагически. Его миновала гильотина, скосившая столько его политических единомышленников-жирондистов, но, вынужденный скрываться от нее, Кондорсе не в добрый час встретился в лесу с волками.

Еще одно имя, встретившееся вам,— Кеннет Эрроу. Он — наш современник, известный американский экономист, нобелевский лауреат 1972 года по экономике. Работает в области эконометрии, теории общего экономического равновесия, экономической теории благосостояния, теории управления.

# ДИАГОНАЛИ ПРАВИЛЬНОГО 18-УГОЛЬНИКА

В.Прасолов

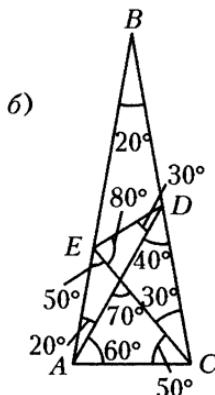
Почти во всех геометрических теоремах речь идет о достаточно широком классе фигур. Если даже теорема относится не к любому треугольнику, то, по крайней мере, к прямоугольному треугольнику. На этом фоне резко выделяются задачи, относящиеся к треугольникам с вполне конкретными углами, причем эти углы используются вовсе не для того, чтобы подставить их в какую-то общую для всех треугольников формулу. Вот два примера таких задач:

1. В равнобедренном треугольнике  $ABC$  с основанием  $BC$  угол при вершине  $A$  равен  $80^\circ$ . Внутри треугольника взята точка  $M$  так, что  $\angle MBC = 30^\circ$  и  $\angle MCN = 10^\circ$ . Докажите, что  $\angle AMC = 70^\circ$ .

2. В равнобедренном треугольнике  $ABC$  с основанием  $AC$  угол при вершине  $B$  равен  $20^\circ$ . На сторонах  $BC$  и  $AB$  взяты точки  $D$  и  $E$  соответственно так, что  $\angle DAC = 60^\circ$  и  $\angle ECA = 50^\circ$ . Докажите, что  $\angle ADE = 30^\circ$ .

Эти задачи не простые; прямым вычислениям они не поддаются. Их решения обычно используют замысловатые дополнительные построения. Такие решения оставляют ощущение недовлетворенности. Сомневаться в их правильности нет оснований, но совершенно непонятно, откуда они берутся.

Попробуем, однако, сформулировать обе задачи одинаковым образом. На рисунке 1 изображены треугольники, о которых идет речь. Из этого



a)

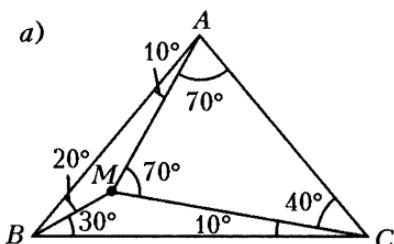


Рис. 1

Опубликовано в «Кванте» №5 за 1995 г.

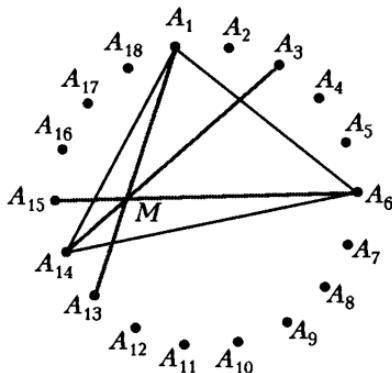


Рис. 2

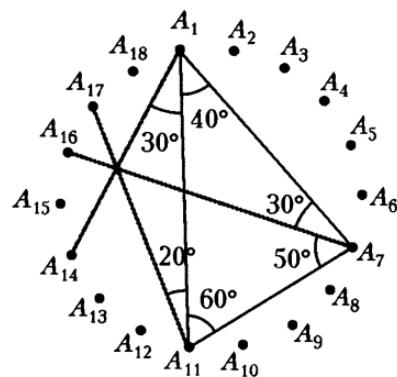


Рис. 3

рисунка можно понять, что задача 1 эквивалентна следующему утверждению: в правильном 18-угольнике диагонали  $A_1A_{13}$ ,  $A_3A_{14}$  и  $A_6A_{15}$  пересекаются в одной точке (рис.2). В самом деле, если эти диагонали пересекаются в одной точке, то

$$\angle A_1MA_6 = \frac{1}{2}(\angle A_1A_6 + \angle A_{13}A_{15}) = 50^\circ + 20^\circ = 70^\circ.$$

Задача 2 эквивалентна такому утверждению: в правильном 18-угольнике диагонали  $A_1A_{14}$ ,  $A_7A_{16}$  и  $A_{11}A_{17}$  пересекаются в одной точке.

Тройки диагоналей получились совсем разные, поэтому задачи, казалось бы, независимые. Так ли это? Нет, это вовсе не так. Задачу 2 очень легко решить, если воспользоваться тем, что первая тройка диагоналей пересекается в одной точке. Обратимся для этого к рисунку 4. Диагонали  $A_4A_{13}$  и  $A_9A_{15}$  симметричны относительно диаметра  $A_5A_{14}$ , поэтому точка их пересечения лежит на нем. Значит, утверждение задачи 1 для треугольника  $OA_{14}A_{15}$  заключается в том,

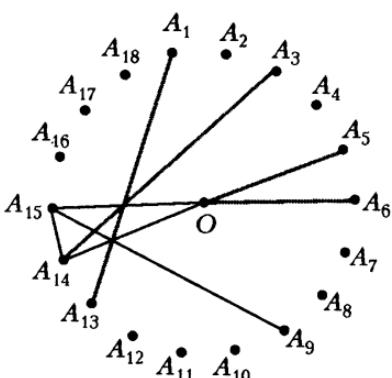


Рис. 4

что угол между диагоналями  $A_1A_{13}$  и  $A_3A_{14}$  равен  $30^\circ$ . Но этот угол легко вычисляется: он равен  $(\angle A_1A_3 + \angle A_{13}A_{14})/2 = 20^\circ + 10^\circ$ . Совсем разные точки пересечения диагоналей оказались странным образом связаны друг с другом.

Взаимосвязей между точками пересечения троек диагона-

лей очень много, особенно если рассматривать еще и точки пересечения продолжений диагоналей. Например, тройку пересекающихся диагоналей, изображенную на рисунке 2, мы получили, рассмотрев описанную окружность треугольника  $ABC$  (рис. 1, а). Но ведь мы можем рассмотреть и описанные окружности треугольников  $ABM$ ,  $BCM$  и  $CAM$ . В этих случаях мы, правда, получим не точки пересечения диагоналей, а точки пересечения продолжений диагоналей, но все они разные. Есть еще один способ получать по тройке пересекающихся диагоналей другие тройки. Но нам уже пора научиться доказывать, что все тройки диагоналей, о которых шла речь, действительно пересекаются в одной точке.

Проверять, пересекаются ли тройки диагоналей в одной точке, очень удобно с помощью следующего утверждения.

**Теорема.** На сторонах треугольника  $ABC$  взяты точки  $A_1$ ,  $B_1$  и  $C_1$  ( $A_1$  на  $BC$  и т. д.). Отрезки  $AA_1$ ,  $BB_1$  и  $CC_1$  пересекаются в одной точке тогда и только тогда, когда

$$\frac{\sin \angle BAA_1}{\sin \angle CAA_1} \cdot \frac{\sin \angle ACC_1}{\sin \angle BCC_1} \cdot \frac{\sin \angle CBB_1}{\sin \angle ABB_1} = 1.$$

**Доказательство.** Предположим сначала, что отрезки  $AA_1$ ,  $BB_1$  и  $CC_1$  пересекаются в точке  $O$  (рис. 5). Тогда

$$2S_{AOB} : 2S_{AOC} = (AB \cdot AO \sin \angle BAO) : (AC \cdot AO \sin \angle CAO).$$

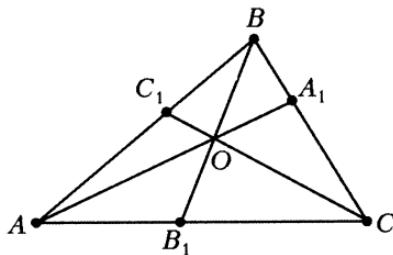


Рис. 5

Следовательно,

$$\begin{aligned} 1 &= \frac{S_{AOB}}{S_{AOC}} \cdot \frac{S_{COA}}{S_{COB}} \cdot \frac{S_{BOC}}{S_{BOA}} = \\ &= \left( \frac{AB}{AC} \cdot \frac{CA}{CB} \cdot \frac{BC}{BA} \right) \cdot \frac{\sin \angle BAO}{\sin \angle CAO} \cdot \frac{\sin \angle ACO}{\sin \angle BCO} \cdot \frac{\sin \angle CBO}{\sin \angle ABO}. \end{aligned}$$

Предположим теперь, что для точек  $A_1$ ,  $B_1$  и  $C_1$  выполняется указанное соотношение. Пусть  $O$  – точка пересечения отрез-

ков  $AA_1$  и  $BB_1$ . Нужно доказать, что отрезок  $CC_1$  проходит через точку  $O$ . Иными словами, если  $C'_1$  – точка пересечения прямых  $CO$  и  $AB$ , то  $C'_1 = C_1$ . Отрезки  $AA_1$ ,  $BB_1$  и  $CC'_1$  пересекаются в одной точке, поэтому, как только что было доказано,

$$\frac{\sin \angle BAA_1}{\sin \angle CAA_1} \cdot \frac{\sin \angle CBB_1}{\sin \angle ABB_1} \cdot \frac{\sin \angle ACC'_1}{\sin \angle BCC'_1} = 1.$$

Сравнив эту формулу с условием теоремы, получим

$$\sin \angle ACC_1 : \sin \angle BCC_1 = \sin \angle ACC'_1 : \sin \angle BCC'_1.$$

Остается доказать, что при движении точки  $X$  по отрезку  $AB$  величина  $\sin \angle ACX : \sin \angle BCX$  изменяется монотонно. Сами углы  $ACX$  и  $BCX$  изменяются монотонно, но их синусы могут не быть монотонными для тупого угла  $C$ . Это не беда. В любом треугольнике есть острый угол, и мы с самого начала могли бы взять в качестве угла  $C$  острый угол. Доказательство теоремы завершено.

Теперь проверка того, что тройка диагоналей, изображенная на рисунке 2, пересекается в одной точке, сводится к проверке тождества

$$\frac{\sin 10^\circ}{\sin 70^\circ} \cdot \frac{\sin 30^\circ}{\sin 20^\circ} \cdot \frac{\sin 40^\circ}{\sin 10^\circ} = 1.$$

Доказательство несложно:

$$\sin 30^\circ \sin 40^\circ = \frac{1}{2} \sin 40^\circ = \sin 20^\circ \cos 20^\circ = \sin 20^\circ \sin 70^\circ.$$

Тройка диагоналей, изображенная на рисунке 3, соответствует тождеству

$$\sin 20^\circ \sin 40^\circ \sin 20^\circ = \sin 30^\circ \sin 60^\circ \sin 10^\circ.$$

Есть еще три тождества, приводящие к тройкам пересекающихся диагоналей:

$$\sin 10^\circ \sin 20^\circ \sin 80^\circ = \sin 20^\circ \sin 20^\circ \sin 30^\circ,$$

$$\sin 20^\circ \sin 30^\circ \sin 30^\circ = \sin 10^\circ \sin 40^\circ \sin 50^\circ,$$

$$\sin 10^\circ \sin 20^\circ \sin 30^\circ = \sin 10^\circ \sin 10^\circ \sin 100^\circ.$$

Проверку этих тождеств мы оставляем читателю.

От перестановки сомножителей произведение не изменяется. Все мы так к этому привыкли, что даже не ожидаем извлечь из этого что-либо содержательное. Но, изменения в наших тождествах порядок сомножителей, мы изменяем порядок дуг и

получаем совсем другие тройки диагоналей! Это и есть обещанный способ получения новых троек пересекающихся диагоналей.

Наш интерес к 18-угольнику, а не к какому-либо другому  $n$ -угольнику, связан с тем, что именно к нему приводят треугольники с углами, кратными  $10^\circ$ . По счастливой случайности у него оказался обильный набор троек пересекающихся диагоналей. Среди всех правильных многоугольников, число вершин которых меньше 18, интересные наборы пересекающихся диагоналей есть у 12-угольника. Например, диагонали  $A_1A_5$ ,  $A_2A_6$ ,  $A_3A_8$  и  $A_4A_{11}$  правильного 12-угольника пересекаются в одной точке (рис.6). Это утверждение эквивалентно следующей хорошо известной задаче.

**3. Внутри квадрата  $ABCD$  взята точка  $P$  так, что треугольник  $ABP$  равносторонний. Докажите, что  $\angle PCD = 15^\circ$ .**

Чтобы у читателя не сложилось мнение, будто с пересечением троек диагоналей связаны лишь разобранные выше задачи, а другие задачи такого рода этим способом решить нельзя, разберем еще два примера. В обоих случаях рассматривается треугольник  $ABC$  с углами  $\angle A = 50^\circ$ ,  $\angle B = 60^\circ$ ,  $\angle C = 70^\circ$ .

**4. На сторонах  $BA$  и  $AC$  взяты точки  $D$  и  $E$  так, что  $\angle DCB = \angle EBC = 40^\circ$ . Тогда  $\angle BED = 30^\circ$ .**

**5. На сторонах  $BA$  и  $BC$  взяты точки  $D$  и  $E$  так, что  $\angle DCA = 50^\circ$  и  $\angle EAC = 40^\circ$ . Тогда  $\angle AED = 30^\circ$ .**

Эти утверждения легко доказать, воспользовавшись рисунками 7 и 8.

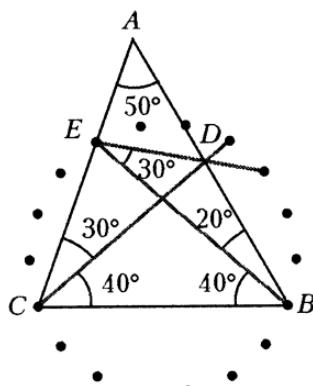


Рис. 7

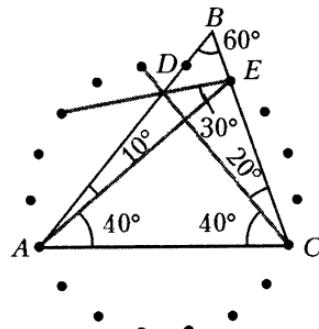


Рис. 8

Задачи о треугольниках, связанные с углами, кратными  $10^\circ$ , можно решить как геометрически, так и с помощью троек пересекающихся диагоналей правильного 18-угольника, т.е., по сути дела, тригонометрически. По крайней мере, это справедливо для всех задач, известных автору статьи. Что же касается задач о треугольниках, связанных с углами, кратными  $1^\circ$ , то для одной из них известно лишь тригонометрическое решение.

Читатель может попытаться найти геометрическое доказательство этого утверждения.

**6.** В треугольнике  $ABC$  углы  $A$ ,  $B$  и  $C$  равны  $14^\circ$ ,  $62^\circ$  и  $104^\circ$ . На сторонах  $AC$  и  $AB$  взяты точки  $D$  и  $E$  соответственно так, что  $\angle DBC = 50^\circ$  и  $\angle ECB = 94^\circ$ . Тогда  $\angle CED = 34^\circ$  (рис. 9).

Тригонометрическое доказательство этого утверждения сводится к проверке тождества

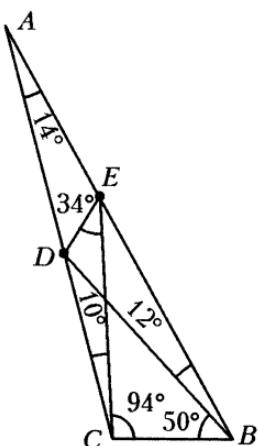
$$\sin 10^\circ \sin 16^\circ \sin 24^\circ = \sin 2^\circ \sin 34^\circ \sin 94^\circ.$$

Рис. 9

Наборы пересекающихся диагоналей правильных многоугольников были одной из тем, обсуждавшихся на летней конференции математического Турнира городов (Челябинск, июль 1991 г.).

### Упражнения

- Докажите, что диагонали  $A_1A_{n+2}$ ,  $A_{2n-1}A_3$  и  $A_{2n}A_5$  правильного  $2n$ -угольника пересекаются в одной точке.
- Докажите, что диагонали  $A_1A_7$ ,  $A_3A_{11}$  и  $A_5A_{21}$  правильного 24-угольника пересекаются в точке, лежащей на оси симметрии  $A_4A_{16}$ .
- Докажите, что диагонали  $A_1A_7$ ,  $A_2A_9$  и  $A_4A_{17}$  правильного 30-угольника пересекаются в одной точке.
- Докажите, что в правильном 36-угольнике семь диагоналей  $A_1A_{13}$ ,  $A_2A_{17}$ ,  $A_3A_{21}$ ,  $A_4A_{24}$ ,  $A_5A_{26}$ ,  $A_8A_{29}$  и  $A_{10}A_{30}$  пересекаются в одной точке.



# УМЕЕТЕ ЛИ ВЫ СЧИТАТЬ?

В.Рождественский

Последние десятилетия характеризуются бурным развитием и распространением вычислительной техники. Электронные вычислительные машины, настольные, карманные и прочие счетные устройства позволяют значительно сократить время всевозможных расчетов. И порой от школьников можно услышать, что «уметь считать» сейчас вообще не нужно. Но как это ни парадоксально, наиболее эффективную помочь технике оказывает именно тому, кто умеет грамотно вычислять.

Мы попробуем показать читателю трудности, возникающие при практических расчетах, а также дать некоторые советы, как с ними справляться.

## Озадачивающие примеры

**Пример 1.** При вычислении на настольной вычислительной машинке выражения  $x = 1/3 \cdot 3/5 \cdot 5/7 \cdot 7/8 \cdot 8$  с точностью до двух цифр после запятой вместо ожидаемой единицы получается  $x = 0,88$ . Если же вычислять  $1/8 \cdot 8/7 \cdot 7/5 \cdot 5/3 \cdot 3$ , то получится  $0,90$ .

**Пример 2.** При сложении на ЭВМ чисел  $1 + 1/2 + 1/3 + 1/4 + \dots + 1/250000$  с пятью знаками после запятой слева направо получается  $12,88777$ , а справа налево получается  $12,95821$ ; мы видим, что разница в ответах больше, чем каждое из большинства слагаемых.

**Пример 3.** Приближенное значение  $\sqrt{290}$  равно  $17,029$  (ошибка меньше  $0,0004$ ). При вычислении дроби  $1\left(\sqrt{290} - 17\right)$  с тремя цифрами после запятой получается  $34,483$ . Однако если дробь преобразовать к виду  $\sqrt{290} + 17$ , то результат сложения будет  $34,029$ , с ошибкой меньшей  $0,0004$ . Стало быть, ответ  $34,483$  гораздо грубее.

**Пример 4.** Решением системы

$$\begin{cases} x + 5y = 17, \\ 1,5x + 7,501y = 25,503 \end{cases}$$

<sup>1</sup> Опубликовано в «Кванте» №4 за 1984 г.

является пара чисел (2; 3). Если изменить число 25,503 всего на 0,003 –

$$\begin{cases} x + 5y = 17, \\ 1,5x + 7,501y = 25,5. \end{cases}$$

то решением будет пара (17; 0). Если же в исходной системе число 7,501 заменить на 7,499 –

$$\begin{cases} x + 5y = 17, \\ 1,5x + 7,499y = 25,503, \end{cases}$$

то решением будет (32; -3).

Подобные примеры могут смутить начинающего вычислителя. Чтобы правильно ориентироваться в похожих ситуациях, нужно знать:

- 1) какие бывают ошибки и откуда они берутся;
- 2) как количественно характеризовать ошибку;
- 3) как грамотно организовать вычисления.

### Классификация ошибок

Ошибки можно разделить на три типа<sup>1</sup>.

**1) Ошибки, содержащиеся в исходной информации.** Исходными данными во всякой задаче обычно служат либо часто встречающиеся константы, либо величины, полученные в результате измерений. Но любой измерительный прибор обладает какой-то определенной разрешающей способностью. Например, на торговых весах невозможно уловить разницу в 1 г, так что, уплатив за 100 г масла, вряд ли стоит верить, что вы получите ровно 100 г, а не 101 г или 99 г. Ошибки измерений могут возникнуть также из-за неверного снятия показания прибора либо в результате грубого просмотра.

Исключая два последних источника ошибок, вызванных субъективными причинами, остается, тем не менее, принять тот факт, что ошибки в исходных данных неизбежны и обусловлены точностью измерительного прибора. *Никакие манипуляции с исходными данными не позволяют гарантировать, что результат обладает большей точностью, чем данные.* Утверждения вроде следующего: «площадь нарисованного на бумаге круга,

<sup>1</sup> О различных источниках ошибок, возникающих при обработке экспериментальных данных, и о том, что получается, если ими пренебречь, рассказывалось также в статье А.Н.Зайделя «Обман или заблуждение?» («Квант» № 5 за 1983 г.)

радиус которого измерялся линейкой, равна  $3206,9569 \text{ мм}^2$ , – по меньшей мере наивны.

**2) Ошибки округления.** Эти ошибки тесно связаны со спецификой приближенного счета. Дело в том, что никакое вычислительное «средство», будь то ЭВМ или человек, не может производить арифметические операции с бесконечными последовательностями. Таким образом, все данные и промежуточные результаты должны быть записаны при помощи ограниченного набора цифр и, следовательно, как правило, неточны.

**3) Ошибки ограничения.** Очень многие вычисления описываются как бесконечная или астрономически большая последовательность уточнений результата. Естественно, приходится обрывать такой процесс в какой-то момент и довольствоваться последним полученным приближением. Примером может служить процесс вычисления  $\sqrt{a}$  методом последовательных приближений:

$$x_1 = (a + 1)/2, \dots, x_{n+1} = (a/x_n + x_n)/2, \dots$$

Итак, мы видим, что ошибки неизбежны. Значит, необходимо уметь их количественно оценивать, чтобы получать результаты с необходимой точностью.

### Распространение ошибок

Поскольку каждый вычислительный процесс в конечном счете сводится к выполнению четырех арифметических действий, посмотрим, что происходит с ошибками при выполнении этих действий.

Для простоты будем считать все рассматриваемые величины (которые могут быть исходными данными, промежуточными или конечными результатами вычислений) положительными. Их точные значения будем обозначать  $\underline{x}, \underline{y}, \dots$ , соответствующие приближенные –  $x, y, \dots$ .

Характеристиками точности приближения служат *абсолютная ошибка* (или погрешность)  $\Delta_x = \underline{x} - x$  и *относительная ошибка*  $\omega_x = \Delta_x/\underline{x}$ . В таблице 1 собраны формулы, выражющие зависимость абсолютной и относительной ошибок результатов арифметических действий от ошибок данных (проверьте эти формулы самостоятельно!).

При практических измерениях и расчетах точные значения абсолютной и относительной ошибок бывают неизвестны (потому что неизвестны точные значения рассматриваемых величин). Однако, зная точность измерительных приборов, правила, по которым производилось округление промежуточных результа-

Таблица 1

Абсолютная ошибка	Относительная ошибка
$\Delta_{x+y} = \Delta_x + \Delta_y$	$\omega_{x+y} = \frac{x}{x+y} \omega_x + \frac{y}{x+y} \omega_y$
$\Delta_{x-y} = \Delta_x - \Delta_y$	$\omega_{x-y} = \frac{x}{x-y} \omega_x - \frac{y}{x-y} \omega_y$
$\Delta_{x \cdot y} = x\Delta_y + y\Delta_x + \Delta_x\Delta_y$	$\omega_{x \cdot y} = \omega_x + \omega_y + \omega_x\omega_y$
$\Delta_{x/y} = \left( \frac{\Delta x}{y} - \frac{x}{y} \frac{\Delta y}{y} \right) : \left( 1 + \frac{\Delta y}{y} \right)$	$\omega_{x/y} = (\omega_x - \omega_y) \frac{1}{1 + \omega_y}$

тов и т.п., обычно можно гарантировать, что абсолютная или относительная ошибки не превосходят по модулю некоторых величин — границ этих ошибок. Если заданы границы ошибок для величин  $x$  и  $y$ , то из формул таблицы 1 нетрудно получить границы ошибок для  $x + y$ ,  $x - y$ ,  $xy$  и  $x/y$  (см. таблицу 2, в

Таблица 2

Граница абсолютной ошибки	Граница относительной ошибки
$ \Delta_{x+y}  \leq \Delta_1 + \Delta_2$	$ \omega_{x+y}  \leq \frac{x}{x+y} \omega_1 + \frac{y}{x+y} \omega_2$
$ \Delta_{x-y}  \leq \Delta_1 + \Delta_2$	$ \omega_{x-y}  \leq \frac{x}{x-y} \omega_1 + \frac{y}{x-y} \omega_2$
$ \Delta_{x \cdot y}  \leq x\Delta_2 + y\Delta_1 + \Delta_1\Delta_2$	$ \omega_{x \cdot y}  \leq \omega_1 + \omega_2 + \omega_1\omega_2$
$ \Delta_{x/y}  \leq \frac{y\Delta_1 + x\Delta_2}{y^2} : \left  1 + \frac{\Delta_y}{y} \right $	$ \omega_{x/y}  \leq (\omega_1 + \omega_2) \times \left  \frac{1}{1 + \omega_y} \right $

которой предполагается, что  $|\Delta_x| \leq \Delta_1$ ,  $|\Delta_y| \leq \Delta_2$ ,  $|\omega_x| \leq \omega_1$ ,  $|\omega_y| \leq \omega_2$ .

В этой таблице особняком стоит формула для абсолютной ошибки частного  $\Delta_{x/y}$ : в ней входит сама ошибка  $\Delta_y$ , а не ее граница  $\Delta_2$ , поэтому  $|\Delta_{x/y}|$  может оказаться большим при  $\Delta_y < 0$ , даже если  $\Delta_1$  и  $\Delta_2$  малы.

## **Анализ примеров**

Попробуем сейчас понять, чем объясняются «странныости», с которыми мы столкнулись в приведенных в начале статьи примерах.

**Пример 1.** По правилам обычной арифметики в произведении дробей  $\frac{1}{3} \cdot \frac{3}{5} \cdot \frac{5}{7} \cdot \frac{7}{8} \cdot \frac{8}{1}$  можно, не меняя результата, как угодно переставить числители, а также знаменатели. Из нашего примера видно, что в арифметике приближенных вычислений это уже неверно. Чтобы понять, как возникает расхождение в результатах, возьмем совсем простой пример – вычислим двумя способами выражение  $(1 \cdot 3)/3$  (с точностью до двух знаков после запятой). Имеем

$$1,0 \times 3,0 = 3,0, \quad 3,0/3,0 = 1,0,$$

но, с другой стороны,

$$1,0/3,0 = 0,33, \quad 0,33 \times 3,0 = 0,99.$$

Если арифметических действий много, то разница в ответах может быть еще большей.

**Пример 2** показывает, что результат приближенного сложения нескольких чисел может зависеть от порядка, в котором оно производится. Рассмотрим более наглядный пример. Пусть надо вычислить сумму

$$100 + \underbrace{0,1 + 0,1 + \dots + 0,1}_{2000 \text{ раз}},$$

округляя все результаты до трех значащих цифр. (Напомним, что *значающими* называются все цифры в записи числа, начиная с первой слева ненулевой; скажем, в числах 0,017 и 1,7 – по две значащие цифры, в числе 170 – три.) Если вести сложение слева направо, то сначала получим  $100 + 0,1 = 100,1$  и после округления опять 100; этот результат, очевидно, не изменится и в дальнейшем. При сложении справа налево получим последовательно:  $0,1 + 0,1 = 0,2$ ,  $0,2 + 0,1 = 0,3$ , ...  $99,9 + 0,1 = 100$ . Следующая тысяча слагаемых результата не изменит, а при последнем сложении получим  $100 + 100 = 200$ . Это уже лучше, но все еще далеко от точного ответа 300, получить который можно сложив тысячу чисел 0,1, потом еще тысячу 0,1 и затем три полученные сотни. Этот пример делает достаточно понятным следующее *практическое правило*:

*суммирование большого числа слагаемых должно быть организовано так, чтобы на каждом шагу складывались по возможности близкие числа.*

Чтобы скрупулезно следовать этому правилу, пришлось бы перебирать очень много разных вариантов суммирования. Обычно поступают проще: располагают все слагаемые по возрастанию и складывают их последовательно от наименьшего к наибольшему. Можно доказать, что этот способ лучше любого другого, если не расставлять скобок (не разбивать слагаемые на группы).

**Пример 3.** В этом примере первый, «плохой», способ счета дает значительную ошибку по двум причинам: сначала мы находили разность близких чисел  $\sqrt{290}$  и 17, а затем делили (единицу) на эту малую разность. Очевидно, относительная ошибка разности близких чисел  $x$  и  $y$  может оказаться гораздо больше, чем их собственные ошибки  $\omega_x$  и  $\omega_y$ , а при делении на малое число может сильно вырасти абсолютная ошибка (см. формулы для  $\omega_{x-y}$  и  $\Delta_{x/y}$  в таблице 1). Поэтому

следует по возможности избегать вычитания близких чисел и деления на малые числа.

Пренебрежение этими правилами порой приводит к неприятностям даже в самых, казалось бы, безобидных ситуациях. Попробуем, например, решить квадратное уравнение  $x^2 + 9,9x - 1 = 0$ , округляя все результаты до двух значащих цифр. Найдем дискриминант:  $D = 9,9^2 + 4 = 102,01$ ; после округления получим  $D = 100$ . Если вычислять корни по обычной формуле для решения квадратного уравнения  $x^2 + px + q = 0$  ( $x_{1,2} = (-p \pm \sqrt{D})/2$ ), мы получим  $x_1 = (-9,9 + 10)/2 = 0,05$ ,  $x_2 = -9,95 \approx -10$ . Корень  $x_2$  (после округления) совпал со своим точным значением, но  $x_1$  отличается от точного значения 0,1 в 2 раза! При вычислении  $x_1$  вычитались близкие числа 10 и 9,9 и их верные цифры просто взаимно уничтожались. Избежать ошибки здесь можно, заменив выражение  $(-p + \sqrt{D})/2$  равным ему выражением  $-2q/(p + \sqrt{D}) = -2/(9,9 + 10) \approx 0,1$ , не содержащим разности близких чисел. (Но заметим, что при вычислении по этой формуле корня  $x_2$  получается  $x_2 = -20$ !)

**Пример 4** иллюстрирует ситуацию, когда очень маленькая разница в исходных данных (коэффициентах системы линейных уравнений) ведет к очень большой разнице в ответах. Если в рассмотренных системах мы не можем гарантировать правильность всех цифр коэффициентов, то нельзя доверять ни одной цифре ответов. Как возникла эта ситуация? Дело в том, что вторые уравнения всех трех систем с высокой точностью совпадают с уравнением  $1,5x + 7,5y = 25,5$ , которое просто эквивалентно первому уравнению  $x + 5y = 17$ . С геометрической точки

зрения все эти уравнения задают близкие прямые. Но ясно, что одна прямая может быть очень близка к другой и при этом пересекать ее в любой заданной точке. Поэтому при «малом шевелении» одной из двух «почти совпадающих» прямых точка их пересечения может сдвинуться весьма значительно. С алгебраической же точки зрения неустойчивость решений наших систем относительно изменений коэффициентов объясняется тем, что при нахождении неизвестных нам неизбежно придется делить на малый знаменатель. А это, как мы видели, ведет к большим ошибкам.

Какие же можно сделать выводы?

Прежде всего, нужно ясно понимать, что даже хорошая точность исходных данных еще не гарантирует столь же хорошую точность результата. При этом результат может зависеть от способа вычисления и даже от порядка выполнения операций. Таким образом,

*одной из главных задач при приближенных вычислениях является задача о выборе наилучшей последовательности вычислений.*

Продумывая последовательность вычислений, кроме уже приведенных правил, нужно учитывать и то, что каждая операция может вносить свою ошибку округления. Поэтому дадим еще один совет:

*старайтесь свести к минимуму число операций.*

## Две практические задачи

**Задача 1.** На бумаге нарисован параллелограмм. Требуется найти его площадь и оценить ошибку. (В нашем распоряжении имеются линейка, циркуль и транспортир.)

Известны две формулы для площади параллелограмма:  $S = ab \sin \alpha$  и  $S = ah$  ( $a, b$  – его стороны,  $\alpha$  – угол между ними,  $h$  – высота, опущенная на сторону  $a$ ). Ясно, что лучше взять вторую формулу, потому что она требует на одно измерение и на одну арифметическую операцию меньше, чем первая, к тому же в первой формуле нужно находить  $\sin \alpha$ . (Правда, построение высоты циркулем и линейкой тоже вносит некоторую погрешность, но на общую ошибку она влияет мало.) Допустим, что измерения дали такие результаты:  $a = 20,2$  см,  $h = 11,3$  см. Поскольку они проводились линейкой,  $|\Delta_a| < 0,1$  см,  $|\Delta_h| < 0,1$  см, а относительные ошибки не превосходят 0,01.

Умножая  $a$  на  $h$ , получаем  $S = 228,26$  см<sup>2</sup>. При этом можно гарантировать, что найденное число отличается от истинного не

более чем на  $20,2 \cdot 0,1 + 11,3 \cdot 0,1 + 0,1^2 = 3,16$  ( $\text{см}^2$ ) (см. формулу для  $|\Delta_{xy}|$  в таблице 2). Это может показаться удивительным – ошибка в 0,1 см в исходных данных приводит к ошибке, большей 3  $\text{см}^2$ , в результате! Но именно такова разница между площадями в двух предельных случаях:  $a = 20$  см,  $h = 11,2$  см и  $a = 20,3$  см,  $h = 11,4$  см.

**Задача 2.** Имея в распоряжении циркуль и линейку (с делениями), найти площадь поверхности бильярдного шара.

Найдем сначала радиус шара  $R$ . Для этого нарисуем на шаре циркулем окружность и отметим на ней три точки  $A, B, C$ . Измеряя циркулем попарные расстояния между этими точками, перенесем этот треугольник на плоскость и опишем около него окружность. Пусть  $r$  – радиус этой окружности, а  $l$  –

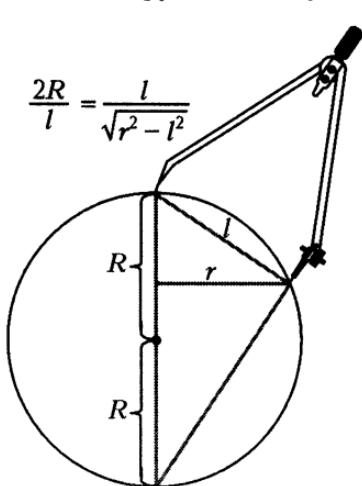
исходный раствор циркуля (см. рисунок). Легко видеть, что тогда  $R = l^2 / 2\sqrt{l^2 - r^2}$ , т.е. искомая площадь равна

$$S = 4\pi R^2 = \frac{\pi^2}{1 - \left(\frac{r}{l}\right)^2}.$$

Проанализируем эту формулу. Если  $r$  близко к  $l$  (т.е. раствор циркуля мал по сравнению с размерами шара), в знаменателе будет стоять разность близких чисел. Поэтому ошибка результата может быть очень большой. Значит, нужно постараться, чтобы

радиус окружности как можно больше отличался от  $l$ . Этого легко добиться для бильярдного шара (особенно если циркуль с кривыми ножками). Но если шар большой, получить заслуживающий доверия результат обычным циркулем практически невозможно.

Например, при  $l = 6,0$  см,  $r = 3,0$  см получаем  $S = 48$   $\text{см}^2$ , причем  $|\Delta_S| < 2,2$   $\text{см}^2$ , если граница абсолютной ошибки измерений составляет 0,1 см, – это вполне удовлетворительная точность. Но допустим, что измерения на большом шаре дали результаты  $l = 6,3$  см,  $r = 6,0$  см и, соответственно,  $S = 427$   $\text{см}^2$ . Тогда при изменении данных на 0,1 см получим для  $l = 6,2$  см,  $r = 6,1$  см –  $S = 1201$   $\text{см}^2$ , а для  $l = 6,4$  см,  $r = 5,9$  см –  $S = 273$   $\text{см}^2$ . Следовательно, приближение  $S = 427$   $\text{см}^2$  в этом



случае очень и очень грубое и находить таким способом площадь поверхности большого шара нельзя. Нужно либо найти циркуль побольше, либо заменить линейку и циркуль более точными инструментами, либо... придумать другой способ вычисления радиуса.

### Упражнения

1. Всегда ли верны следующие равенства в арифметике приближенных вычислений:

- а)  $(a + b) - c = a + (b - c) = (a - c) + b$ ; б)  $(ab)/c = a(b/c) = (a/c)b$ ;  
в)  $(a - b)c = ac - bc$ ; г)  $(a + b)^2 = a^2 + 2ab + b^2$ ?

2. Решите систему уравнений

$$\begin{cases} 11,04x - 10,49y = 1, \\ 10,48x - 10,04y = 0, \end{cases}$$

округлив коэффициенты: а) до двух; б) до трех значащих цифр. Сравните результаты с ответом, полученным при вычислениях с четырьмя значащими цифрами.

3. Придумайте в примере 1 наихудшую последовательность вычислений.

### Литература

1. Иванова Т. П., Пухова Г.В. Программирование и вычислительная математика. – М.: Просвещение, 1978.

2. Мак-Кракен Д., Дорн У. Численные методы и программирование на фортране. – М.: Мир, 1977.

## ПОКРЫТИЯ ПОЛОСКАМИ

М.Смурров, А.Спивак

*Полоса – часть плоскости, заключенная между двумя параллельными прямыми.*

На Московской олимпиаде 1997 года одиннадцатиклассники решали задачу, вошедшую в «Задачник «Кванта»:

**M1600.** На плоскости даны: конечное число полос, сумма ширин которых равна 100, и круг радиусом 1. Докажите, что каждую из полос можно параллельно перенести так, чтобы все они вместе покрыли круг.

С ней справился только один из 410 участвовавших в олимпиаде одиннадцатиклассников. Между тем при обсуждении варианта многие члены жюри, даже не желая слушать условие до конца, заявляли, что задача им известна. Они путали M1600 со знаменитой задачей, о которой будет рассказано во второй части статьи.

Число 100 играло в условии роль «большого числа». Мы докажем, что достаточно меньшей суммарной ширины полос, равной  $\pi + 2$ .

В то же время, 100 нельзя заменить ни на какое число, меньшее  $\pi$ . Чтобы доказать это, впишем в круг с радиусом 1 правильный  $2n$ -угольник. Через концы каждой его стороны проведем перпендикулярные ей прямые (рис.1). Получим  $n$  полос.

Сумма их ширин равна половине периметра  $2n$ -угольника. Она стремится к  $\pi$  при возрастании  $n$ . Во второй части статьи мы докажем, что если сузить полосы (т.е. уменьшить их ширины), то никакими их сдвигами рассматриваемый  $2n$ -угольник не покроешь.

Впрочем, основное содержание статьи – рассказ о некоторых трудных и интересных проблемах комбинаторной геометрии. Их формулировки привлекательны и просты. Но на многие вопросы еще нет ответа.

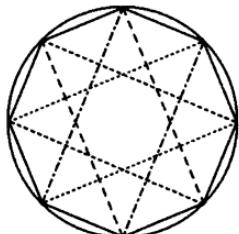


Рис. 1

---

Опубликовано в «Кванте» №4, 5 за 1998 г.

## Что такое ширина?

### Ширина по направлению

Начнем с простой ситуации. Пусть даны фигура  $F$  и одна полоса (рис.2). Можно ли полосу параллельно перенести («сдвинуть») так, чтобы накрыть  $F$ ?

Разумеется, некоторые фигуры (например, угол) вообще не помещаются ни в какую полосу. Поэтому дальше будем предполагать, что фигура  $F$  является ограниченной, т.е. содержитя в некотором круге. Для таких фигур ответ определяется тем, что больше – ширина полосы или ширина фигуры в соответствующем направлении, т.е. ширина самой узкой полосы, в которой содержится  $F$  и края которой параллельны заданному направлению.

Чтобы получить такую полосу, можно взять сначала полосу соответствующего направления, достаточно широкую для того, чтобы в ней содержалась фигура  $F$  (на рисунке 3 такая полоса ограничена прямыми  $m$  и  $n$ ). Затем края полосы надо двигать друг к другу до тех пор, пока они не упрются в  $F$ . На рисунке 3 эти прямые –  $a$  и  $b$ . Они называются опорными прямыми фигуры  $F$ .

Каждому направлению соответствует своя пара опорных прямых. Как уже было сказано, расстояние между ними, т.е. ширина получившейся полосы, называется шириной фигуры  $F$  в соответствующем направлении. Например, если  $F$  – квадрат со стороной 1, то ширина, в зависимости от направления, меняется от 1 до  $\sqrt{2}$ .

Если  $F$  – треугольник, то ширина меняется от минимальной, равной наименьшей из высот треугольника, до максимальной, равной наибольшей его стороне.

Доказать последнее утверждение о максимальной ширине совсем просто: с одной стороны, расстояние между опорными прямыми, перпендикулярными стороне треугольника, не может быть меньше длины этой стороны. С другой стороны, расстояние между параллельными опорными прямыми треугольника не превосходит длины стороны, концы которой лежат на этих прямых (например, ширина горизонталь-

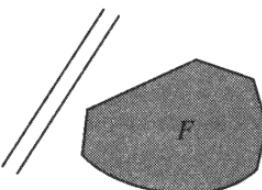


Рис. 2

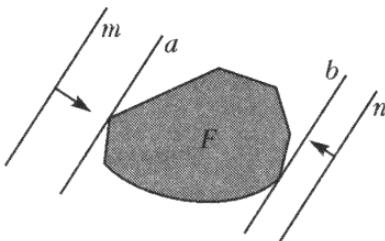


Рис. 3

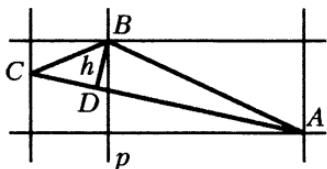


Рис. 4

2. Какой должна быть ширина полосы бумаги, чтобы из нее можно было вырезать треугольник со сторонами 15 см, 20 см и 25 см?

### Фигуры постоянной ширины

Ширина круга во всех направлениях одинакова и равна диаметру круга. Хотя это и не связано прямо с темой статьи, отметим, что существуют отличные от круга выпуклые фигуры постоянной ширины.

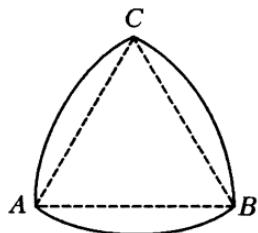


Рис. 5

Пример – *треугольник Рело*. Он получается, если нарисовать правильный треугольник  $ABC$  и дуги с центрами  $A, B, C$  (рис.5). Есть и другие примеры. Так, вместо правильного треугольника можно было взять правильный пятиугольник (или даже выпуклый пятиугольник, все диагонали которого одной длины).

### Определения

Если ширина фигуры зависит от направления, то естественно рассмотреть наименьшую и наибольшую из ширин. (В *Приложении* будет объяснено, почему эти величины действительно существуют. Пока лучше на это не отвлекаться.)

**Определение 1.** *Шириной* фигуры называется наименьшая из ее ширин по направлениям, т.е. ширина самой узкой полосы, которой можно покрыть эту фигуру (или, что то же самое, ширина наименьшей полосы, из которой фигуру можно вырезать).

**Определение 2.** *Диаметром* замкнутой<sup>1</sup> ограниченной фигуры называется наибольшее из расстояний между ее точками.

### Упражнения

3. Найдите диаметр а) треугольника; б) прямоугольника; в) полукруга.

<sup>1</sup> Замкнутой называется фигура, которая содержит все точки своей границы.

**4.** Для каких треугольников диаметр совпадает с диаметром описанной окружности?

**5.** Докажите, что а) ширина прямоугольника равна его наименьшей стороне; б) ширина параллелограмма равна его наименьшей высоте.

**6.** Докажите, что ширину выпуклого многоугольника можно найти следующим образом: сначала для каждой стороны найти расстояние до наиболее удаленной от нее вершины, а затем из этих расстояний выбрать наименьшее.

**7.** Докажите, что диаметр фигуры совпадает с максимальной из ее ширин по направлениям.

**8.** а) Может ли опорная прямая не иметь с ограниченной фигурой ни одной общей точки? б) Может ли опорная прямая неограниченной замкнутой фигуры не иметь с фигурой ни одной общей точки?

\* \* \*

Подведем итог в задаче о покрытии фигуры одной полосой: если о направлении полосы ничего не известно, то для того, чтобы ее параллельным сдвигом можно было покрыть фигуру, достаточно потребовать, чтобы ширина полосы была не меньше диаметра фигуры.

Покрытия квадрата двумя полосами обсуждены в разделе «Гипотеза о покрытиях». Покрытия тремя и более полосами примененным там методом не изучишь. Сейчас перейдем к задаче, в условии которой нет слов «полоса» и «ширина». Но именно ширина – ключ к решению.

### Разделяющая прямая

#### Правильные треугольники

**Задача 1.** Если в правильном треугольнике со стороной 1 расположены не налегающие друг на друга правильные треугольники со сторонами  $a$  и  $b$ , то  $a + b \leq 1$  (рис.6).

**Решение.** Как бы ни были расположены не пересекающие друг друга правильные треугольники, можно провести такую прямую  $l$  (рис.7), что один треугольник окажется по одну сторону от  $l$ , а другой – по другую

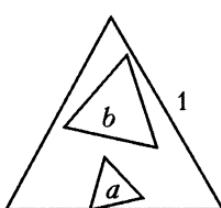


Рис. 6

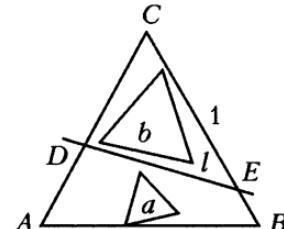


Рис. 7

Прямая  $l$  разобьет треугольник  $ABC$  на две части – треугольник и четырехугольник (который выродится в треугольник, если  $l$  пройдет через вершину  $\Delta ABC$ ). Пусть, для определенности,  $AD \geq BE$ .

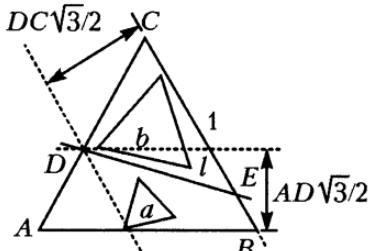


Рис. 8

Проведем через точку  $D$  прямые параллельно сторонам  $AB$  и  $BC$  треугольника  $ABC$ . Получим две полосы (рис.8), одна из которых содержит четырехугольник  $ABED$ , а другая – треугольник  $CDE$ .

Поскольку четырехугольник  $ACED$  расположен в полосе шириной  $AD\sqrt{3}/2$ , для ширины  $a\sqrt{3}/2$  правильного треугольника со стороной  $a$  выполняется неравенство  $a\frac{\sqrt{3}}{2} \leq AD\frac{\sqrt{3}}{2}$ , т.е.  $a \leq AD$ . Аналогично, правильный треугольник со стороной  $b$  лежит в полосе шириной  $DB\sqrt{3}/2$ . Значит,  $b\frac{\sqrt{3}}{2} \leq DB\frac{\sqrt{3}}{2}$ , т.е.  $b \leq DB$ .

Итак,  $a \leq AD$ ,  $b \leq DC = AC$ , откуда  $a + b \leq AD + DC = AC$ .

**Упражнение 9.** Дан треугольник, величина одного из углов которого равна  $60^\circ$ . Разместите в нем равносторонний треугольник наибольшей возможной площади.

**Замечания.** В решении задачи 1 использовано очевидное (и важное!) свойство: если некоторая фигура  $A$  является подмножеством фигуры  $B$ , то ширина  $A$  не может быть больше ширины  $B$ .

Между прочим, ширина фигуры может совпадать с шириной ее собственной части. Иными словами, от некоторых фигур можно отрезать кусочек, не уменьшив при этом их ширину. Например, ширина прямоугольника  $a \times b$ , где  $a < b$ , равна  $a$ . Такова же ширина содержащегося в нем квадрата со стороной  $a$ . (Ширина вписанного в этот квадрат круга тоже равна  $a$ .)

Ширины треугольника  $CDE$  и четырехугольника  $ABED$  равны ширинам полос рисунка 8 (в решении задачи 1 мы использовали только то, что они не превосходят ширин этих полос). Следовательно, как ни режь правильный треугольник произвольной прямой на две части, сумма ширин частей будет равна ширине треугольника.

### Упражнения

**10.** Если треугольник не разносторонний, то его можно разрезать прямой на две части, сумма ширин которых больше ширины исходного треугольника.

**11.** Из любого ли неравнобедренного треугольника можно вырезать равнобедренный треугольник той же ширины, что и исходный треугольник?

**12.** а) Если круг разрезан прямой на две части, то сумма ширин этих частей равна ширине (диаметру) круга. б) Сумма диаметров любых двух кругов, расположенных без пересечения в некотором круге, не превосходит его диаметра.

**13.** Квадрат разрезан на две части прямой, не параллельной его сторонам. Докажите, что сумма ширин частей больше ширины (стороны) квадрата.

### Квадраты

**Задача 2.** Внутри квадрата разместили два непересекающихся квадрата (рис.9). Докажите, что сумма их сторон не превосходит стороны квадрата.

**Решение.** Проведем разделяющую прямую. Она разрежет квадрат на треугольник и пятиугольник (рис.10) или на две трапеции (рис.11)<sup>2</sup>. Достроим эти фигуры до прямоугольных треугольников и поставим задачу:

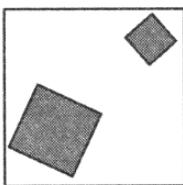


Рис. 9

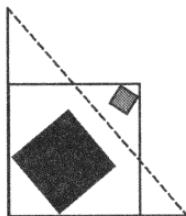


Рис. 10

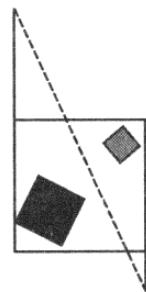


Рис. 11

**Задача 3.** В данный прямоугольный треугольник вписать квадрат с наибольшей возможной стороной.

**Решение задачи 3.** Интуитивно ясно<sup>3</sup>, что все вершины наибольшего квадрата должны лежать на сторонах треугольника. Таким образом, надо рассмотреть два случая: 1) квадрат упирается одним углом в прямой угол треугольника, а противоположным – в гипотенузу (рис.12); 2) две вершины квадрата лежат на гипотенузе, а две другие – на катетах (рис.13).

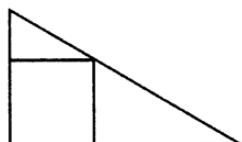


Рис. 12

<sup>2</sup> Возможны еще случаи, когда разделяющая прямая проходит через вершину квадрата. Разберите их самостоятельно.

<sup>3</sup> На самом деле это – тема для отдельной заметки.

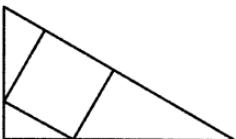


Рис. 13

**Упражнение 14.** а) Выразите длину стороны квадрата рисунка 12 через катеты  $a$  и  $b$  рассматриваемого прямоугольного треугольника.

б) Выразите длину стороны квадрата рисунка 13 через гипотенузу  $c$  и опущенную на нее высоту  $h$ .

в) Сравните значения выражений – ответов пунктов а) и б).

Из последнего упражнения получаем ответ задачи 3: наибольший квадрат, вписанный в данный прямоугольный треугольник, изображен на рисунке 12 (одна из диагоналей этого квадрата – биссектриса прямого угла треугольника). Теперь в задаче 2 все ясно: сумма сторон квадратиков достигает наибольшего значения, когда их диагонали составляют диагональ квадрата (рис.14).

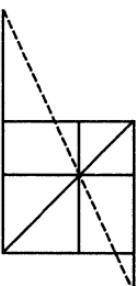


Рис. 14

### Упражнения

**15\***. Докажите, что если внутри правильного  $2n$ -угольника лежат два правильных не налагающих друг на друга  $2n$ -угольника, то сумма периметров последних не превосходит периметра объемлющего многоугольника.

*Замечание.* Доказывать аналогичное утверждение для правильных многоугольников с нечетным числом сторон мы не умеем.

**16\* (М935).** Если внутри правильного  $2n$ -угольника со стороной  $a$  и центром  $O$  поместить правильный  $2n$ -угольник, стороны которого больше  $a/2$ , то он накроет точку  $O$ .

*Указание.* Если выпуклый многоугольник не содержит точку  $O$ , то найдется такая его сторона  $AB$ , что точка  $O$  и многоугольник лежат по разные стороны от  $AB$ . Примените центральную симметрию и результат предыдущего упражнения.

### Гипотеза Эрдёша

До сих пор не решена поставленная более 50 лет назад венгерским математиком Эрдёшем задача: доказать, что для любой расположенной в единичном квадрате системы не налагающих друг на друга  $k^2 + 1$  квадратиков сумма их периметров не превосходит  $4k$ , т.е. суммы периметров  $k^2$  одинаковых квадратиков со стороной  $1/k$ , на которые можно разбить единичный квадрат.

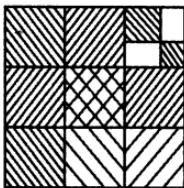


Рис. 15

Например, при  $k = 3$  гипотеза Эрдёша состоит в том, что расположение 10 квадратиков рисунка 15 дает наибольшую возможную сумму периметров.

**Упражнение 17.** Как связана задача 2 с гипотезой Эрдёша?

### Решение задачи M1600

Наивная попытка решения M1600 могла бы состоять в следующем: сумма ширин полос настолько велика, что можно вырезать из каждой полосы по квадрату и потом покрывать такими квадратами единичный круг.

К сожалению, если полосы очень узкие, например, если дана система 1000000 полос с одинаковой шириной  $1/10000$ , то сумма площадей квадратиков равна

$$1000000 \times (1/10000)^2 = 0,01,$$

что гораздо меньше площади единичного круга.

К дню олимпиады мы знали довольно длинное решение. Потом было придумано короткое решение, дающее более точную оценку, так что можете сразу переходить к следующему подразделу.

#### «Длинное» решение

Каждой полосе сопоставим перпендикулярный вектор (рис.16), длина которого равна ширине этой полосы (какое именно из двух возможных направлений придать такому вектору, пока не существо-

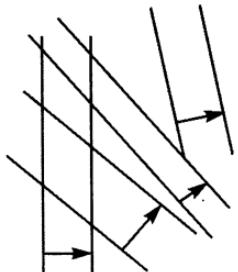


Рис. 16

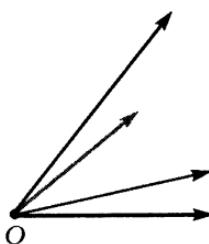


Рис. 17

ственno – следующий абзац все прояснит). Отложим векторы от одной и той же точки  $O$  (рис.17).

Разобьем плоскость на 12 углов величиной  $30^\circ$  с вершиной  $O$ . Объединим каждый из этих углов в пару с вертикальным ему углом. Для каждой из полученных 6 пар вертикальных углов подсчитаем сумму длин векторов, лежащих внутри или на границе этих углов. Хотя бы одна из сумм не меньше  $100/6$  – в противном случае сумма длин всех векторов была бы меньше 100.

Выберем такую пару вертикальных углов. Заменив, если потребуется, некоторые векторы на противоположные, добьемся, чтобы все они попали в один и тот же угол величиной  $30^\circ$ .

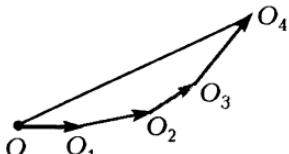


Рис. 18

Теперь упорядочим векторы по направлению и, прикладывая начало каждого следующего вектора к концу предыдущего, составим из попавших в угол векторов выпуклую ломаную  $OO_1O_2\dots O_n$  (рис.18). Длина ее, как уже было сказано, не меньше  $100/6$ .

Поскольку направления всех векторов  $\overrightarrow{OO_1}, \overrightarrow{O_1O_2}, \dots, \overrightarrow{O_{n-1}O_n}$  заключены между направлениями векторов  $\overrightarrow{OO_1}$  и  $\overrightarrow{O_{n-1}O_n}$ , направление суммы  $\overrightarrow{OO_1} + \overrightarrow{O_1O_2} + \dots + \overrightarrow{O_{n-1}O_n} = \overrightarrow{OO_n}$  тоже заключено между этими направлениями. Следовательно,  $\overrightarrow{OO_1}, \overrightarrow{O_1O_2}, \dots, \overrightarrow{O_{n-1}O_n}$  наклонены к вектору  $\overrightarrow{OO_n}$  под углами, не превосходящими  $30^\circ$ . Поэтому длина  $OO_n$  не меньше  $(100/6) \cdot \cos 30^\circ = 25/\sqrt{3}$ .

Перенесем полосы параллельно так, чтобы концы отрезков  $[OO_1], [O_1O_2], \dots, [O_{n-1}O_n]$  лежали бы на краях перпендикулярных им полос (рис.19). Обозначим через  $M$  пересечение перпендикуляров к отрезкам  $[OO_1]$  и  $[O_{n-1}O_n]$ , восстановленных в точках  $O$  и  $O_n$  соответственно.

Индукцией по числу полос легко доказать, что многоугольник  $MOO_1O_2\dots O_n$  полностью покрыт полосами. Поскольку величины углов  $MOO_n$  и  $MO_nO$  не меньше  $90^\circ - 30^\circ = 60^\circ$ , в треугольнике  $MOO_n$  содержится правильный треугольник со стороной  $OO_n$ , в который, в свою очередь, вписан

круг радиусом

$$OO_n/(2\sqrt{3}) = (25/\sqrt{3})/(2\sqrt{3}) = \frac{25}{6} > 1.$$

*Замечание.* Жюри не догадалось использовать индукцию для доказательства того, что многоугольник  $MOO_1O_2\dots O_n$  покрыт полосами. Предполагалось следующее рассуждение.

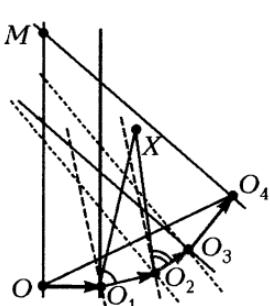


Рис. 20

Рассмотрим любую точку  $X$  многоугольника  $MOO_1O_2\dots O_n$ . Если перпендикулярная отрезку  $[OO_1]$  полоса не покрывает точку  $X$ , то угол  $XO_1O_2$  острый (рис.20). Если точку  $X$  не покрывает и полоса, перпендикулярная отрезку  $[O_1O_2]$ , то угол  $XO_2O_3$  острый. Продолжая в том же духе, мы поймем, что если точка  $X$  не покрыта ни одной из первых  $n - 1$  полос, то все углы  $XO_1O_2, XO_2O_3, \dots, XO_{n-1}O_n$  острые. Но тогда точка  $X$  как раз покрыта полосой, перпендикулярной отрезку  $O_{n-1}O_n$ !

**Упражнение 18.** Из точки, расположенной внутри выпуклого многоугольника, опустили перпендикуляры на стороны или их продолжения. Докажите, что хотя бы один из перпендикуляров попал именно на сторону, а не на ее продолжение. (Другими словами, полосы, построенные перпендикулярно сторонам выпуклого многоугольника, покрывают его.)

**Указание.** Физик сделал бы из данного многоугольника колесо, поместив в интересующую его точку грузик, и позволил бы колесу перекатываться (рис. 21).

Математик сделает по сути то же самое, если рассмотрит кратчайшее из расстояний от данной точки до прямых, на которых лежат стороны данного многоугольника.

**Упражнение 19.** Дан выпуклый многогранник и точка внутри него. Докажите, что хотя бы один из перпендикуляров, опущенных из этой точки на плоскости граней, пересекается с соответствующей гранью, а не только с ее плоскостью.

**Замечание.** Если вам все еще кажется, что утверждение задачи М1600 тривиально для «очень большой» суммарной ширины полос, обдумайте стереометрический вариант:

*В пространстве даны: шар радиуса 1 и несколько «слоев»<sup>4</sup>, сумма толщин которых равна 1000000. При любом ли расположении слоев можно каждый из них параллельно перенести так, чтобы они покрыли шар?*

Эта задача – нерешенная проблема, даже если вместо 1000000 написать сколь угодно большое число.

### Более точная оценка

После олимпиады было придумано простое решение задачи М1600. Позже выяснилось, что такое же решение опубликовал в 1984 году американский математик Грёмер.

Рассмотрим систему полос и выпуклую фигуру  $F$ , граница которой состоит из кривой  $AB$  и отрезков  $BO$ ,  $OA$ . На рисунке 22 полоса  $p$  параллельна лучу  $OB$ , пересечение полосы  $q$  с углом  $AOB$  неограничено, а полосы  $r$  – ограничено.

Пусть в данной системе полос нет полос типа  $r$ . Другими словами, проведя через вершину  $O$  угла прямые парал-



Рис. 21

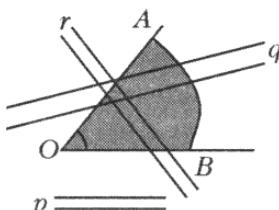


Рис. 22

<sup>4</sup> Слоем мы называем здесь часть пространства, заключенную между двумя параллельными плоскостями.

лько границам полос, потребуем, чтобы все проведенные прямые имели по лучу внутри или на границе угла  $AOB$ . Докажем, что если сумма ширин полос больше длины кривой  $AB$ , то  $F$  можно покрыть сдвигами этих полос.

Для этого упорядочим направления полос по часовой стрелке и возьмем крайнее из них. Перенесем соответствующую полосу так, чтобы одна из ограничивающих ее прямых стала опорной прямой фигуры  $F$  (рис.23, 24). Поскольку от кривой  $AB$  будет

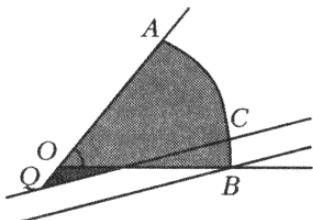


Рис. 23

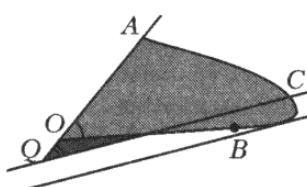


Рис. 24

отрезан кусочек не меньший, чем ширина полосы, то задача будет сведена к покрытию оставшимися полосами фигуры, ограниченной отрезками  $AQ$ ,  $QC$  и кривой  $CA$ .

Решение задачи М1600 теперь очевидно (рис.25): верхняя полуокружность и два вертикальных отрезка образуют кривую длиной  $\pi + 2$ . Круг радиуса 1 заключен между этой кривой и горизонтальной прямой, которую мы представим как две стороны развернутого угла с вершиной  $O$ .

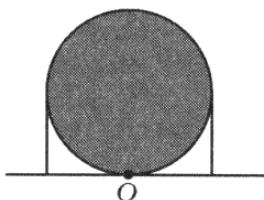


Рис. 25

### Гипотеза о покрытиях

Итак, число 100 в формулировке М1600 можно заменить на  $2 + \pi$ . Было бы интересно выяснить, можно ли заменить его на число  $\pi$ . Еще интереснее узнать, верна ли следующая гипотеза.

**Гипотеза.** Любую выпуклую фигуру с периметром  $P$  можно покрыть параллельными сдвигами любых полос, сумма ширин которых равна половине периметра фигуры.

Рассмотрим два частных случая: квадрат и шестиугольник.

### Покрытия квадрата двумя полосами

Квадрат со стороной 1 невозможно покрыть вертикальной и горизонтальной полосами, ширины которых меньше 1 (рис.26).

**Задача 4.** Если сумма ширин двух полос равна 2, то любой квадрат со стороной 1 можно покрыть параллельными сдвигами этих полос.

**Решение.** Обозначим большую из ширин полос буквой  $w$ . Очевидно,  $w \geq 1$  – в

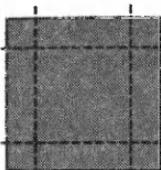


Рис. 26

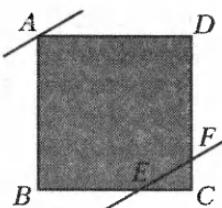


Рис. 27

противном случае сумма ширин рассматриваемых полос была бы меньше 2. Перенесем эту полосу так, чтобы один из ее краев стал опорной прямой квадрата (рис.27).

Докажем, что треугольник  $EFC$  можно покрыть второй полосой, т.е. что ширина этого треугольника не превосходит ширины полосы:

$$EF \leq 2 - w. \quad (*)$$

Рассмотрим  $EF$  как функцию от  $w$ . Легко понять, что эта функция линейная:  $EF = kw + b$  при некоторых не зависящих от  $w$  величинах  $k$  и  $b$ .

Значит, неравенство  $(*)$  достаточно проверить в крайних точках. При  $w = 1$  рассмотрим окружность радиусом 1 с центром  $A$  (рис.28). По равенству отрезков касательных,  $BE = EK$ ,  $KF = FD$ . Следовательно,

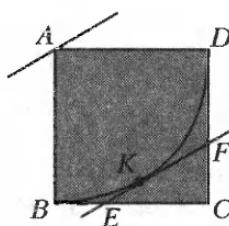


Рис. 28

$$2EF < (EK + KF) + (EC + CF) = BE + DF + EC + CF = 2,$$

откуда  $EF < 1$ .

В другом крайнем случае, когда первая полоса проходит через точку  $C$  и целиком покрывает квадрат,  $EF = 0$  и неравенство  $(*)$  верно.

### Покрытие шестиугольника

Полосы рисунка 29 не покрывают шестиугольник. Тем не менее, его можно покрыть их сдвигами (рис.30). Вообще, пусть  $A_1A_2A_3A_4A_5A_6$  – выпуклый центрально-симметричный шестиугольник.

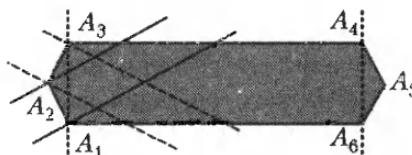


Рис. 29

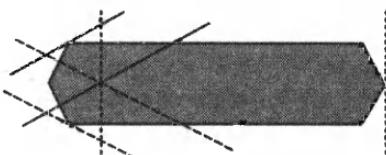


Рис. 30

Оказывается, чтобы покрыть его полосами, образованными перпендикулярами к сторонам  $A_1A_2$ ,  $A_2A_3$  и  $A_3A_4$  достаточно эти полосы

приложить не к «своим» вершинам, а к вершинам  $A_3$ ,  $A_1$  и  $A_5$  соответственно.

**Упражнение 20.** Докажите это.

*Замечание.* Было бы интересно узнать ответ на следующий частный случай гипотезы: если через концы каждой стороны выпуклого центрально-симметричного многоугольника  $S$  проведем перпендикулярные этой стороне прямые и из каждого двух образовавшихся полос одного направления оставим только одну, то всегда ли  $S$  можно покрыть сдвигами полученных полос?

\* \* \*

Во второй части статьи доказана теорема Банга–Тарского: если выпуклая фигура  $F$  покрыта полосами, то ширина фигуры не превышает суммы ширин полос. Сформулированы гипотезы Банга и Дэвенпорта, рассказано о некоторых других проблемах комбинаторной геометрии.

### Покрытия круга и выход в пространство

Очевидно, круг диаметра  $d$  можно покрыть параллельными друг другу полосами, сумма ширин которых равна  $d$  (рис.31).

**Задача 5.** Нельзя ли так расположить полосы разных направлений, чтобы они покрыли круг большего диаметра, чем сумма ширин полос?

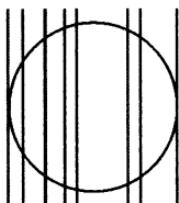


Рис. 31

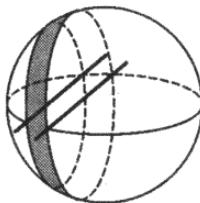


Рис. 32

**Решение.** Нельзя. Доказательство использует, как ни странно, выход в пространство.

Пусть круг диаметра  $d$  покрыт  $n$  полосами ширин  $w_1, w_2, \dots, w_n$ . Рассмотрим сферу, экватором которой служит этот круг, т.е. сферу, центр и радиус которой совпадают с центром и радиусом круга (рис.32). Проведя через границы полос плоскости, перпендикулярные экватору, сопоставим каждой полосе сферическую полоску (или шапочку, если круг пересечен только одной из ограничивающих полосу прямых).

Теперь используем замечательный (и известный уже Архиме-

ду) факт: *площадь сферической полоски, высекаемой на сфере двумя параллельными плоскостями, выражается через расстояние  $w$  между плоскостями и диаметр  $d$  сферы формулой  $\pi d w$ .* (Обе плоскости, разумеется, должны пересекать сферу или хотя бы касаться ее.) Поскольку площадь сферы равна  $\pi d^2$  и поскольку сфера должна быть полностью покрыта, имеем

$$\pi d w_1 + \pi d w_2 + \dots + \pi d w_n \geq \pi d^2,$$

откуда

$$w_1 + w_2 + \dots + w_n \geq d. \quad (1)$$

### Упражнения

**21.** В круге радиусом  $R$  расположен кружок радиусом  $r$ . Сколько прямых необходимо провести, чтобы хотя бы одна из них пересекла кружок? (Подразумевается, что мы видим круг, но не видим кружок.)

*Указание.* Прямая пересекает кружок в том и только том случае, когда центр кружка удален от нее не более чем на расстояние  $r$ , т.е. лежит в соответствующей полосе шириной  $2r$ .

**22.** Квадрат разрезали на прямоугольники. Может ли сумма меньших сторон прямоугольников оказаться меньше стороны квадрата?

*Замечание.* Это упражнение предлагалось десятиклассникам на московской олимпиаде 1998 года.

**23.** Докажите, что если квадрат покрыт полосами, то сумма ширин полос не может быть меньше стороны квадрата.

*Указание.* Впишите в квадрат круг.

### Теорема Банга–Тарского

В 1932 году польский математик А. Тарский поставил задачу: доказать неравенство (1) для любой ограниченной выпуклой фигуры  $F$ .

В 1950 году датский математик Т. Банг решил эту задачу и поставил проблему, не решенную по сей день. Чтобы сформулировать ее, запишем неравенство (1) в виде

$$\frac{w_1}{d} + \frac{w_2}{d} + \dots + \frac{w_n}{d} \geq 1,$$

где  $w_1, w_2, \dots, w_n$  – ширины полос, покрывающих выпуклую фигуру  $F$ ,  $d$  – ширина фигуры  $F$ . Банг предложил делить ширину  $w_i$  полосы не на (наименьшую) ширину  $d$ , а на ширину  $d_i$  фигуры  $F$  в направлении  $i$ -й полосы:

$$\frac{w_1}{d_1} + \frac{w_2}{d_2} + \dots + \frac{w_n}{d_n} \geq 1. \quad (2)$$

Поскольку  $\frac{w_i}{d} \geq \frac{w_i}{d_i}$ , из неравенства (2) следует неравенство (1).

В 1991 году американский математик К.Болл доказал неравенство Банга (2) в случае, когда фигура  $F$  имеет центр симметрии.

**Упражнение 24.** В кубическом куске сыра имеется дырка – сферическая пустота диаметром в  $1/1998$  длины ребра. Каким минимальным числом плоских разрезов ее можно обнаружить? (Считайте, что кусок сыра после разрезания не распадается.)

### План доказательства теоремы Банга–Тарского

Доказательство будем вести «от противного». Предположим, что выпуклая фигура  $F$  покрыта полосами, сумма ширин  $w_1 + \dots + w_n$  которых меньше ширины  $d$  фигуры  $F$ .

Построим векторы  $\vec{v}_1, \dots, \vec{v}_n$ , каждый из которых перпендикулярен границе соответствующей полосы и равен по длине ее ширине. Как и в решении задачи M1600 в первой части статьи (см. рис.16), выбор одного из двух возможных (противоположных) направлений каждого из векторов  $\vec{v}_i$  произволен.

Поскольку сумма  $w_1 + \dots + w_n$  ширин полос строго меньше ширины  $d$  фигуры, можно увеличить длину каждого из векторов  $\vec{v}_i$  таким образом, чтобы сумма длин полученных векторов по-прежнему была меньше  $d$ . Теперь длины векторов будут больше, чем ширины соответствующих полос:  $|\vec{v}_i| > w_i$  при  $i = 1, \dots, n$ , но  $|\vec{v}_1| + \dots + |\vec{v}_n| < d$ .

Чтобы доказать теорему Банга–Тарского, мы введем понятие остова  $S(\vec{v}_1, \dots, \vec{v}_n)$ , построенного по системе векторов  $\vec{v}_1, \dots, \vec{v}_n$ . Мы докажем, что остов можно расположить внутри  $F$ . Для этого будет использована лемма о ширине пересечения выпуклой фигуры со своим сдвигом. Завершит доказательство теоремы применение леммы Банга.

**Лемма Банга.** Пусть каждому из векторов  $\vec{v}_i$ , где  $i = 1, \dots, n$ , поставлена в соответствие перпендикулярная полоса  $p_i$ , края которой перпендикулярны вектору  $\vec{v}_i$ , а ширина меньше  $|\vec{v}_i|$ . Тогда система полос  $p_1, \dots, p_n$  не покрывает остова  $S(\vec{v}_1, \dots, \vec{v}_n)$ . (Более того, обязательно найдется вершина остова, не лежащая ни в одной из полос.)

Тот, кто преодолеет весь этот путь, будет приятно удивлен: все леммы не только являются этапами доказательства теоремы Банга–Тарского, но и представляют самостоятельный интерес.

Вот пример применения леммы Банга.

### Оценка снизу в задаче М1600

В первой части статьи решена задача М1600 – доказано, что если на плоскости дано конечное число полос, сумма ширин которых равна  $\pi + 2$ , то параллельными сдвигами этих полос можно покрыть круг единичного радиуса. Оказывается, число  $\pi + 2$  нельзя заменить никаким числом, меньшим  $\pi$ . Более того, частным случаем леммы Банга является следующее утверждение:

*Если к каждой стороне вписанного в единичный круг правильного  $2n$ -угольника построить перпендикулярную полосу, то получим  $n$  полос, параллельными сдвигами которых нельзя покрыть никакой  $2n$ -угольник, гомотетичный исходному с коэффициентом гомотетии, большим 1.*

Полосы, образованные перпендикулярами к сторонам правильного  $n$ -угольника, покрывают не только этот многоугольник, но и больший правильный многоугольник (рис.33). Его стороны не параллельны сторонам исходного многоугольника, так что никакого противоречия нет. Мы только видим, что слово «гомотетичный» нельзя заменить словом «подобный».

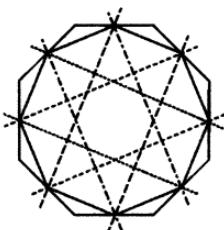


Рис. 33

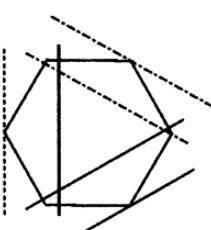


Рис. 34

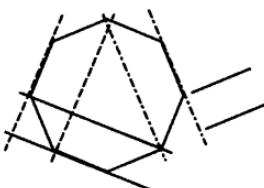


Рис. 35

Важно, что покрыты должны быть не только граничные, но и внутренние точки многоугольника. Например, на рисунке 34 граница шестиугольника покрыта полосами даже более узкими, чем сторона шестиугольника. А для восьмиугольника (рис.35) одна полоса оказалась лишней!

**Упражнение 25.** Перпендикулярные сторонам правильного  $2n$ -угольника полосы покрывают не только описанную окружность, но даже изображенный на рисунке 33 описанный вокруг этой окружности  $2n$ -угольник. Докажите, что никакой больший правильный  $2n$ -угольник сдвигами рассматриваемых полос покрыть невозможно.

## Гомотетичные образы

Подготовку к доказательству теоремы Банга–Тарского начнем издалека. Пусть выпуклая фигура  $F$  разрезана на выпуклые фигуры. Впишем в них фигуры, гомотетичные  $F$  (рис.36).

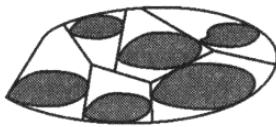


Рис. 36

**Гипотеза.** При любом разрезании гомотетичные  $F$  фигуры можно выбрать настолько большими, чтобы сумма коэффициентов гомотетии была не меньше 1.

Мы умеем доказывать следующий частный случай гипотезы:

*Если выпуклая фигура  $F$  разрезана прямой  $I$  на две части, то в эти части можно вписать гомотетичные образы фигуры  $F$  так, чтобы сумма соответствующих коэффициентов гомотетии равнялась 1.*

Для доказательства проведем опорные прямые  $m$  и  $k$ , параллельные прямой  $I$  (рис.37). Отметим на прямых  $m$  и  $k$  точки  $M$  и  $K$  «соприкосновения» с фигурой  $F$ , т.е. точки, принадлежащие

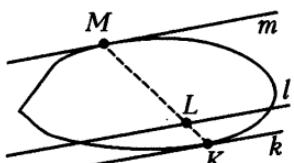


Рис. 37

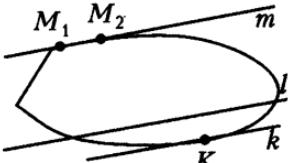


Рис. 38

одновременно и соответствующей прямой, и границе фигуры  $F$ . (Эти точки не всегда определены однозначно – см. рис.38.)

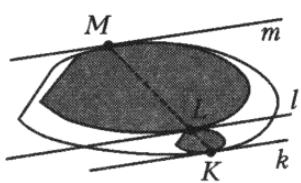


Рис. 39

Обозначим буквой  $L$  точку пересечения отрезка  $KM$  с прямой  $I$ . Образы  $H_K^{KL/KM}(F)$  и  $H_M^{ML/MK}(F)$  фигуры  $F$  при гомотетиях с центрами  $K$  и  $M$  и коэффициентами  $KL/KM$  и  $ML/MK$  содержатся в частях, на которые прямая  $I$  делит фигуру  $F$  (рис.39). Сумма

коэффициентов гомотетии

$$\frac{KL}{KM} + \frac{ML}{MK} = \frac{KL + LM}{KM}$$

равна 1.

### Упражнения

**26.** Дан прямоугольник, не являющийся квадратом. Разместите внутри него два непересекающихся подобных ему с коэффициентами  $k_1$

и  $k_2$  прямоугольника так, чтобы сумма  $k_1 + k_2$  была больше 1.

**Решение.** Выберем точку на диагонали прямоугольника вблизи его вершины<sup>5</sup> и проведем через нее прямые параллельно сторонам (рис.40).

Прямоугольник окажется разбит на четыре части, две из которых содержат отрезки его диагонали и являются прямоугольниками, гомотетичными исходному. Сумма коэффициентов гомотетии равна 1.



Рис. 40

Отразив меньший из прямоугольничков относительно биссектрисы его угла, являющегося заодно и углом исходного прямоугольника, видим, что образовался зазор, который позволяет увеличить прямоугольничек.

**27.** Ограниченнная фигура  $F$  разрезана прямой на две части. Докажите, что нельзя вписать в них гомотетичные образы фигуры  $F$  так, чтобы сумма коэффициентов гомотетии была больше 1.

**28.** Между параллельными опорными прямыми фигуры  $F$  провели на равных расстояниях одну от другой  $n$  параллельных прямых (рис.41). Докажите, что ни в одну из  $n + 1$  частей, на которые оказалась разбита фигура  $F$ , нельзя поместить ее гомотетичный образ с коэффициентом гомотетии, превосходящим  $1/(n+1)$ .

**Замечание.** Американский математик Дэвенпорт выдвинул гипотезу: как бы ни были расположены  $n$  прямых, хотя бы в одной из частей, на которые они делят выпуклую ограниченную фигуру  $F$ , можно разместить гомотетичный образ фигуры  $F$  при гомотетии с коэффициентом  $1/(n+1)$ .

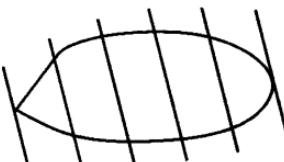


Рис. 41

**29.** Если описанный около окружности радиусом  $r$  многоугольник разбит на треугольники, то сумма  $r_1 + \dots + r_n$  радиусов их вписанных окружностей не меньше  $r$  (рис.42).

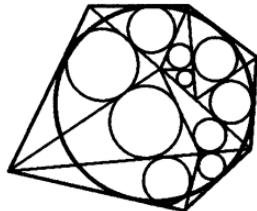


Рис. 42

### Свойство наибольшей хорды

Прежде чем применять гомотетии для доказательства теоремы Банга–Тарского, рассмотрим всевозможные хорды выпуклой фигуры  $F$ , параллельные некоторому фиксированному направлению. Опорные прямые, проведенные в концах

<sup>5</sup> Это нужно, чтобы при описываемой далее симметрии прямоугольничек не вылез из границ исходного прямоугольника.

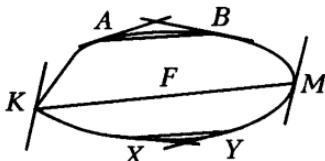


Рис. 43

опорные прямые окажутся параллельными.

**Лемма.** Чрез концы  $K$  и  $M$  самой длинной<sup>6</sup> из семейства параллельных хорд выпуклой фигуры  $F$  можно провести параллельные опорные прямые.

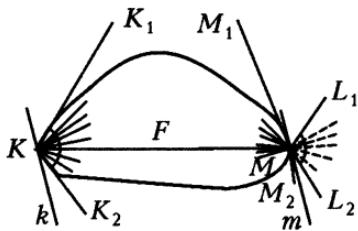


Рис. 44

хорды  $AB$  рисунка 43, не параллельны одна другой. Не параллельны и опорные прямые, проведенные в концах хорды  $XY$ , но пересекаются они «с другой стороны». Интуиция подсказывает, что для какой-то промежуточной между  $AB$  и  $XY$  хорды

параллельные опорные прямые.

**Доказательство.** Проведем всевозможные лучи с началом  $K$ , пересекающие фигуру  $F$  (рис.44). Образуется угол  $K_1KK_2$ . Аналогично строим угол  $M_1MM_2$ . Проведем лучи  $ML_1$  и  $ML_2$  параллельно лучам  $KK_1$  и  $KK_2$  соответственно.

Отделим углы  $M_1MM_2$  и  $L_1ML_2$  один от другого прямой  $m$  – и останется лишь провести через точку  $K$  прямую  $k$ , параллельную  $m$ . Параллельные опорные прямые построены!

### Упражнения

**30.** Мы использовали не доказанный (хотя и очевидный) факт, что углы  $M_1MM_2$  и  $L_1ML_2$  можно отделить один от другого прямой, т. е. что  $\angle K_1KM + \angle KMM_1 \leq 180^\circ$  и  $\angle K_2KM + \angle KMM_2 \leq 180^\circ$ . Докажите эти неравенства.

**31.** Провели две опорные прямые выпуклой замкнутой фигуры, расстояние между которыми равно ширине фигуры. Докажите, что фигура имеет хорду с концами на этих прямых, которая перпендикулярна этим прямым.



Рис. 45

**Указание.** Проведите хорды, перпендикулярные опорным прямым (рис.45). Выберите из них самую длинную. Если ее длина меньше, чем расстояние между опорными прямыми, то проведите через концы этой хорды параллельные опорные прямые.

<sup>6</sup> Существование такой хорды мы не доказываем, поскольку соответствующие теоремы математического анализа не имеют прямого отношения к теме статьи.

## Ширина пересечения фигуры со своим сдвигом

Следующая лемма потребуется нам для доказательства теоремы Банга–Тарского.

**Лемма.** Пусть  $F$  – выпуклая фигура,  $w$  – ее ширина,  $F'$  – фигура, полученная из  $F$  параллельным переносом на вектор  $\vec{v}$  (рис. 46). Тогда ширина пересечения  $F \cap F'$  не меньше чем  $w - |\vec{v}|$ .

В частности, если  $w > |\vec{v}|$ , то пересечение  $F \cap F'$  не пусто.

**Доказательство.** Рассмотрим хорды фигуры  $F$ , параллельные вектору  $\vec{v}$ . Наибольшая из них  $KM$  обладает, как сказано в предыдущем разделе, тем свойством, что через ее концы можно провести параллельные опорные прямые  $k$  и  $m$  (рис. 47).

Обозначим буквой  $L$  точку пересечения с отрезком  $KM$  прямой  $l$ , которая получается из опорной прямой  $k$  параллельным переносом на вектор  $\vec{v}$ . Рассмотрим фигуру  $\tilde{F} = H_M^{LM/KM}(F)$ . Она содержится в  $F$  (рис. 48). Та же самая фигура  $\tilde{F}$  есть образ фигуры  $F'$  при гомотетии с центром  $L$  и тем же коэффициентом  $LM/KM$ . (Это видно из подобия треугольников  $MYL$  и  $XYZ$  на рисунке 49, где  $Z$  – образ точки  $X$  при сдвиге на  $\vec{v}$ .) Следовательно,  $\tilde{F}$  содержится не только в  $F$ , но и в  $F'$ , а значит – в пересечении  $F \cap F'$ . Ширина фигуры  $\tilde{F}$  – гомотетичного образа фигуры  $F$  – равна ширине  $w$  фигуры  $F$ , умноженной на коэффициент гомотетии  $(KM - |\vec{v}|)/KM$ .

**Упражнение 32.** Проверив неравенство  $w - w \frac{|\vec{v}|}{KM} \geq w - |\vec{v}|$ , завершите доказательство леммы.

## Остовы

При доказательстве леммы Банга мы будем следить не за всеми точками многоугольника, а только за вершинами его остова. На рисунках 50, 51 и 52 изображены остовы правильных шести-, восьми- и десятиугольника.



Рис. 46

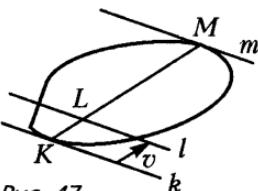


Рис. 47

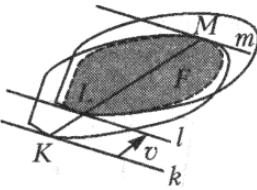


Рис. 48

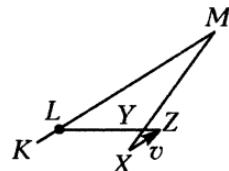


Рис. 49

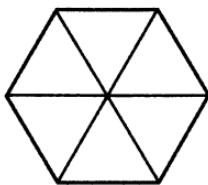


Рис. 50

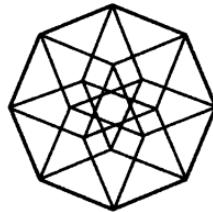


Рис. 51



Рис. 52

### Центрально-симметричные многоугольники

Точное определение остова мы дадим чуть позже, а пока взгляните на рисунок 53. На нем изображены центрально-симметричные многоугольники – четырехугольник (параллелограмм), шестиугольник и восьмиугольник. Любой многоугольник, имеющий центр симметрии, имеет четное число вершин, а его противоположные стороны равны и параллельны.

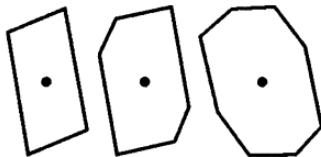


Рис. 53

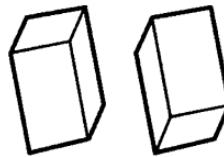


Рис. 54

На рисунке 54 центрально-симметричный шестиугольник разрезан на параллелограммы, .стороны которых равны и параллельны сторонам шестиугольника. На рисунке 55 на параллелограммы разными способами разрезан восьмиугольник.

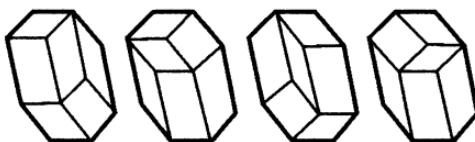


Рис. 55

Изобразим на одном рисунке все разбиения выпуклого центрально-симметричного многоугольника  $F$  на параллелограммы

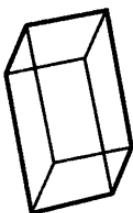


Рис. 56



Рис. 57

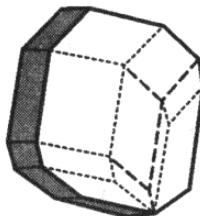


Рис. 58

со сторонами, параллельными и равными сторонам многоугольника (рис.56, 57). Возникнет система точек и отрезков, которую будем называть *остовом* многоугольника  $F$ . Хотя нам это и не потребуется для доказательства теоремы Банга–Тарского, отметим, что любой центрально-симметричный выпуклый многоугольник можно разрезать на параллелограммы со сторонами, равными и параллельными сторонам многоугольника. Идея доказательства показана на рисунке 58.

### Определение остова

Рассмотрим систему векторов  $\vec{v}_1, \dots, \vec{v}_n$  и точку  $O$  (рис.59). Понять, что такое остов, поможет следующее упражнение.

**Упражнение 33.** От незагашенного окурка в одной точке загорелся лес. Ветер дул один час со скоростью  $\vec{v}_1$ , второй час – со скоростью  $\vec{v}_2, \dots, n$ -й час – со скоростью  $\vec{v}_n$ . Пожар распространялся от загоревшихся участков со скоростью ветра (причем загоревшиеся участки продолжали гореть). Какой участок выгорел за  $n$  часов?

Точное определение остова таково. Сдвинем точку  $O$  на вектор  $\vec{v}_1$ . Соединив полученную точку с исходной, получим отрезок  $S_1$  (рис.60). Сдвинем  $S_1$  на вектор  $\vec{v}_2$ . Получим параллелограмм  $S_2$ . Из  $S_2$  при помощи вектора  $\vec{v}_3$  получаем систему отрезков  $S_3$ . Вообще, чтобы получить  $S_{k+1}$ , мы к  $S_k$  добавляем образ  $S_k$  при параллельном переносе на вектор  $\vec{v}_{k+1}$ . Кроме того, добавляем отрезки, соединяющие вершины остова  $S_k$  с вершинами, в которые они переходят при сдвиге на  $\vec{v}_{k+1}$ . Таким образом по системе векторов  $\vec{v}_1, \dots, \vec{v}_n$  мы строим остов  $S_n = S(\vec{v}_1, \dots, \vec{v}_n)$ . Произвольная вершина этого остова получается из начальной точки  $O$  сдвигом на вектор  $a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n$ , где коэффициенты  $a_i$  суть 0 или 1. Напри-

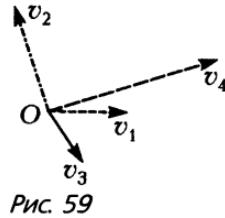


Рис. 59

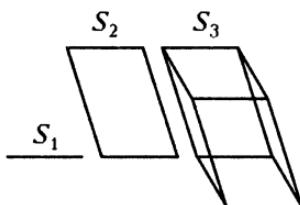


Рис. 60

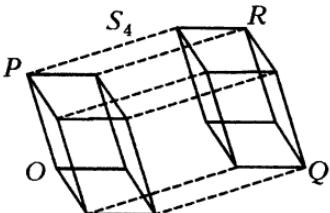


Рис. 61

употребили выражение «вообще говоря» из-за того, что какие-то вершины могли совпасть, как это произошло, например, на рисунке 50, где в центре правильного шестиугольника совпали две вершины его остова.

Остов назовем *невырожденным*, если все его вершины различны. Из каждой вершины невырожденного остова  $S_n$  исходит ровно  $n$  ребер. (Некоторые ребра могут, как это показано на рисунке 57, лежать на одной прямой. На существование дела это не влияет. Надо только не дать себя запутать случайными совпадениями.) При доказательстве теоремы Банга–Тарского будем рассматривать только невырожденные остовы. Мы имеем право это делать, поскольку длины векторов  $\vec{v}_i$ , как вы помните, можно чуточку варьировать.

**Упражнение 34.** Докажите, что набор векторов  $\vec{v}_1, \dots, \vec{v}_n$  дает невырожденный остов в том и только том случае, когда никакой вектор этого набора не представим в виде  $\vec{v}_k = a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_{k-1}\vec{v}_{k-1}$ , где коэффициенты  $a_i$  суть 0, 1 или  $-1$ .

Итак, для того чтобы построить остов выпуклого центрально-симметричного  $2n$ -угольника  $A_1A_2 \dots A_nA_{n+1} \dots A_{2n}$ , достаточно взять  $\vec{v}_1 = \overrightarrow{A_1A_2}$ ,  $\vec{v}_2 = \overrightarrow{A_2A_3}$ , ...,  $\vec{v}_n = \overrightarrow{A_nA_{n+1}}$ , а в качестве начальной точки  $O$  – точку  $A_1$ .

### Почему остов можно разместить в фигуре?

Сдвинем фигуру  $F$  на вектор  $\vec{v}_1$  и пересечение полученной фигуры  $F'$  с  $F$  обозначим через  $F_1$  (рис. 62). Как пересечение выпуклых фигур, фигура  $F_1$  выпуклая. Ее ширина не меньше чем  $d - |\vec{v}_1|$ .

Сдвинем  $F_1$  на вектор  $\vec{v}_2$  и обозначим через  $F_2$  пересечение полученной фигуры с  $F_1$ . Ширина фигуры  $F_2$  не меньше чем  $d - |\vec{v}_1| - |\vec{v}_2|$ .

Будем сдвигать и пересекать фигуры до того момента, когда из  $F_{n-1}$  получим  $F_n$  (на рисунке 63  $n = 4$ ). Каждое из множеств

мер, точка  $P$  рисунка 61 получена из  $O$  сдвигом на  $\vec{v}_2$ , точка  $Q$  получена сдвигом на  $\vec{v}_1 + \vec{v}_3 + \vec{v}_4$ , а  $\overline{OR} = \vec{v}_1 + \vec{v}_2 + \vec{v}_4$ .

Так как число вершин системы  $S_{k+1}$ , вообще говоря, вдвое больше числа вершин системы  $S_k$ , то число вершин остова  $S_n$  равно  $2^n$ . Мы

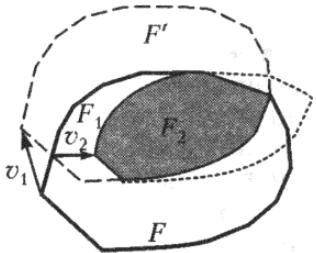


Рис. 62

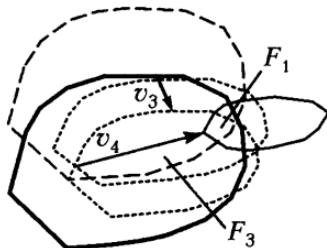


Рис. 63

$F_i$ , где  $i = 1, \dots, n$ , содержится в  $F_{i-1}$ . (Чтобы это утверждение имело смысл при  $i = 1$ , пусть  $F_0 = F$ .) Более того, вместе с фигурой  $F_i$  в  $F_{i-1}$  содержится ее сдвиг на вектор  $-\vec{v}_i$ .

Ширина фигуры  $F_n$  не меньше чем  $d - |\vec{v}_1| - \dots - |\vec{v}_n| > 0$ . Поэтому существует точка  $O \in F_n$ . Если начать построение с нее, то легко понять, что остав  $S(-\vec{v}_n)$  будет содержаться в  $F_{n-1}$ , остав  $S(-\vec{v}_{n-1}, -\vec{v}_n)$  – в  $F_{n-2}, \dots, S(-\vec{v}_1, \dots, \vec{v}_n)$  – в  $F$  (рис. 64).

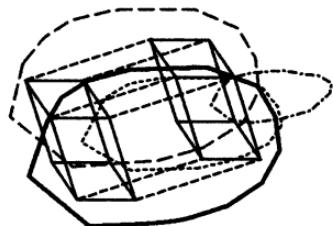


Рис. 64

### Доказательство леммы Банга

Возьмем наугад произвольную вершину остава. Укажем способ, двигаясь в соответствии с которым по ребрам остава, мы дойдем до вершины, не лежащей ни в одной из данных полос.

Если вершина не лежит ни в одной из полос, то двигаться nowhere не надо. Если же она содержится в некоторой полосе  $p_i$ , рассмотрим выходящее из этой вершины ребро остава, соответствующее вектору  $\vec{v}_i$  (рис. 65). Длина этого ребра равна  $|\vec{v}_i|$ . Она больше, чем ширина  $w_i$  полосы  $p_i$ , поэтому второй конец ребра расположен вне полосы  $p_i$ . Туда и надо идти. (Если вершина принадлежала сразу нескольким полосам, то можно идти по любому из соответствующих ребер.)

Применяя к новой вершине остава то же правило, мы перейдем в следующую вершину, и т.д.

К сожалению, чтобы доказать, что мы убежим из всех полосок, недостаточно сказать, что можно выбежать из любой

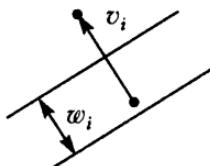


Рис. 65

полоски. Дело в том, что, выбежав на некотором шаге из полосы, мы не имеем гарантии, что не вернемся в нее же через несколько шагов. И все-таки процесс убегания завершится (причем менее чем за  $2^n$  шагов).

Чтобы доказать это, рассмотрим произвольное ребро  $PQ$  остова  $S_n$  (рис.66). Оно соответствует некоторому вектору  $\vec{v}_i$  и полосе  $p_i$ . Ребру  $PQ$  сопоставим разность квадратов  $QM^2 - PM^2$  расстояний от точек  $Q$  и  $P$  до оси полосы  $p_i$  (ось полосы – это прямая, параллельная ее краям и делящая ее на две полосы одинаковой ширины). По теореме Пифагора,

$$QN^2 - PN^2 = (QM^2 + MN^2) - (PM^2 + MN^2) = QM^2 - PM^2$$

для любой точки  $N$  оси полосы, так что рассматриваемая величина есть разность расстояний от точек  $Q$  и  $P$  до любой точки  $N$  оси полосы.

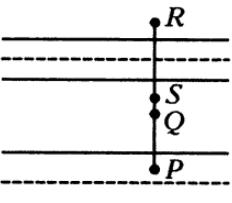
Ломаной  $A_1A_2\dots A_n$ , состоящей из ребер остова, сопоставим сумму чисел, сопоставленных ее звеньям  $A_1A_2, A_2A_3 \dots A_{n-1}A_n$ . Отметим, что если ломаная возникла в процессе убегания от полосок, то соответствующее ломаной число положительно (начало всякого ребра  $A_iA_{i+1}$  лежит в соответствующей этому ребру полосе, а конец – не лежит).

**Лемма.** *Любому замкнутому (т.е. заканчивающемуся там же, где начался) маршруту по вершинам остова сопоставлено число 0.*

Рис. 67

**Доказательство.** Рассмотрим сначала короткий замкнутый маршрут – параллелограмм  $PQRS$  (рис.67). Ребрам  $PQ, QR, RS$  и  $SP$  сопоставлены числа  $QL^2 - PL^2, RL^2 - QL^2, SL^2 - RL^2$  и  $PL^2 - SL^2$ . Их сумма равна нулю.

**Упражнение 35.** Разберите самостоятельно изображенный на рисунке 68 случай, когда оси полос параллельны.



Если маршрут состоит более чем из четырех ребер, то, выбрав любую его точку в качестве начальной, запишем последовательность ребер, из которой маршрут состоит. То, что по любому параллелограмму сумма сопоставленных его сторонам величин равна

0, означает, что можно, не меняя сопоставленной всему маршруту величины, менять маршрут, переставляя соседние векторы. Рассмотрев любой входящий в маршрут вектор и противоположный ему, перестановками будем приближать их друг к другу. Когда они окажутся рядом, эти два вектора исключим из маршрута. Лемма доказана. (Тем, кто знаком с понятием  $n$ -мерного куба, проще всего понять это доказательство, если вообразить, что «стягивается в точку» замкнутый путь, состоящий из ребер  $n$ -мерного куба.)

Из этой леммы как раз следует, что в процессе убегания от полосок вершины не повторяются. (Последнее утверждение могло бы оказаться неверным, если бы мы рассматривали не только невырожденные остовы. Но утверждение все же было бы верно и для вырожденных остовов, если бы мы следили не только за положением вершины остова на плоскости, но и за набором коэффициентов  $a_1, \dots, a_n$ , задающих вершину по формуле  $\overrightarrow{OA} = \overrightarrow{a_1v_1} + \dots + \overrightarrow{a_nv_n}$ .) Поскольку число вершин остова конечно, лемма Банга доказана.

Есть и другой способ доказательства леммы Банга. Скалярное произведение  $\bar{v}_i \cdot \overrightarrow{OM}$  не меняется, когда точка  $M$  перемещается параллельно оси полосы  $p_i$ . Выбрав на оси полосы любую точку  $M_i$ , положим  $b_i = \bar{v}_i \cdot \overrightarrow{OM_i}$ . Прямые, ограничивающие полосу  $p_i$ , можно задать уравнениями

$$\bar{v}_i \cdot \overrightarrow{OM} - b_i = \pm \frac{1}{2} |\bar{v}_i| w_i.$$

Рассмотрим для каждого вектора  $\bar{v} = a_1\bar{v}_1 + a_2\bar{v}_2 + \dots + a_n\bar{v}_n$ , где коэффициенты  $a_i$  суть 0 или 1, величину

$$f(a_1, \dots, a_n) = \bar{v}^2 - 2a_1b_1 - 2a_2b_2 - \dots - 2a_nb_n.$$

Выберем набор  $(a_1, \dots, a_n)$ , для которого величина  $f(a_1, \dots, a_n)$  наибольшая. Тогда при любом  $i = 1, \dots, n$  должно выполняться неравенство

$$f(a_1, \dots, a_i, \dots, a_n) \geq f(a_1, \dots, 1 - a_i, \dots, a_n),$$

которое можно записать в виде

$$\begin{aligned} \bar{v}^2 - 2a_1b_1 - 2a_2b_2 - \dots - 2a_nb_n &\geq \\ &\geq (\bar{v} + (1 - 2a_i)\bar{v}_i)^2 - 2a_1b_1 - 2a_2b_2 - \dots - 2a_nb_n - 2(1 - 2a_i)b_i. \end{aligned}$$

Раскрыв скобки и упростив, получаем неравенство

$$\bar{v}_i^2 \leq 2(2a_i - 1)(\bar{v}_i \bar{v} - b_i),$$

из которого, поскольку  $|2a_i - 1| = 1$  и  $\bar{v}_i^2 > |\bar{v}_i|w_i$ , следует, что

$$\frac{1}{2}|\bar{v}_i|w_i < |\bar{v}_i\bar{v} - b_i|,$$

т.е. соответствующая набору  $(a_1, \dots, a_n)$  вершина остова не лежит в  $i$ -й полосе. Что и требовалось.

## Приложение

### Определения

Фигура называется выпуклой, если вместе с любыми двумя своими точками  $A$  и  $B$  она содержит все точки отрезка  $AB$ .

Точка  $A$  плоскости называется граничной точкой фигуры, если любой круг с центром  $A$  содержит как точки фигуры, так и точки, ей не принадлежащие. Множество граничных точек фигуры называется ее границей.

Точка называется внутренней точкой фигуры, если некоторый круг с центром в этой точке целиком содержится в фигуре. На рисунке 69

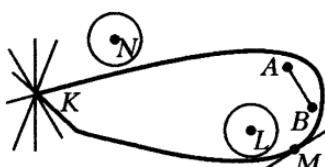


Рис. 69

точки  $K, M$  – граничные точки фигуры  $F$ , точки  $A, B, L$  – внутренние. Точка  $N$  не является ни внутренней, ни граничной точкой, поскольку существует круг с центром  $N$ , в котором нет ни одной точки фигуры  $F$ . В точке  $M$  к фигуре  $F$  можно провести единственную опорную прямую, а в точке  $K$  – бесконечно много.

**Теорема.** Через любую граничную точку выпуклой фигуры проходит хотя бы одна опорная прямая.

**Доказательство.** Пусть  $K$  – граничная точка выпуклой фигуры  $F$ . Проведем всевозможные лучи с началом  $K$  через точки фигуры  $F$

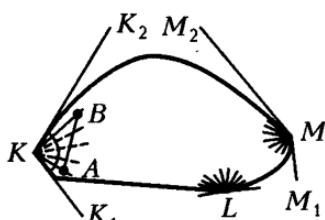


Рис. 70

(рис.70). По определению выпуклой фигуры, вместе с любыми проведенными лучами  $KA$  и  $KB$ , где  $A, B \in F$ , окажутся проведены (через точки отрезка  $AB$ ) все лучи угла  $AKB$ . Значит, образуется угол  $K_1KK_2$ . Он может оказаться развернутым (в роли точки  $K$  могла выступить точка  $L$ ). Ограничивающие этот угол лучи могут войти в построенное множество лучей {как, например, луч  $KK_1$ }, а могут и не войти в него (как  $MM_2$ ).

Но в любом случае через точку  $K$  можно провести опорную прямую к фигуре  $F$ .

## Разделяющая прямая

**Теорема.** *Если  $F$  и  $G$  – выпуклые замкнутые ограниченные непересекающиеся фигуры, то существует прямая, относительно которой  $F$  и  $G$  лежат по разные стороны (в разных полуплоскостях).*

**Доказательство.** Рассмотрим кратчайший отрезок  $AB$ , где  $A \in F$ ,  $B \in G$ . (Длина этого отрезка называется расстоянием между фигурами  $F$  и  $G$ .) Перпендикуляр к отрезку  $AB$ , как легко доказать методом «от противного», разделяет фигуры  $F$  и  $G$ .

### Упражнения

**36.** Докажите, что для любых двух непересекающихся выпуклых многоугольников в качестве разделяющей прямой годится одна из их сторон (точнее, прямая, на которой она лежит).

**37.** Докажите, что если  $F$  и  $G$  – выпуклые непересекающиеся фигуры, то существует прямая, относительно которой  $F$  и  $G$  лежат по разные стороны.

### Почему ширина существует?

Мы считали очевидным, что ограниченная фигура  $F$  имеет ширину по любому направлению. Чтобы прояснить ситуацию, спроектируем фигуру вдоль данного направления на перпендикулярную прямую. Мы приходим к известной теореме о существовании точной верхней (нижней) грани ограниченного сверху (соответственно, снизу) множества. Эта теорема – один из краеугольных камней учения о вещественных числах. Ее доказательство можно найти в учебниках математического анализа.

Если каждой прямой  $l$ , проходящей через начало координат, сопоставить ширину  $w(l)$  ограниченной фигуры  $F$  по направлению  $l$ , мы получим непрерывную функцию (докажите!). По теореме Вейерштрасса, доказательство которой тоже можно найти в учебниках, определенная на ограниченном замкнутом множестве непрерывная функция принимает и наибольшее, и наименьшее значения. Значит, среди всех ширин  $w(l)$  существуют наименьшая и наибольшая.

# ГЕОМЕТРИЯ СКОЛЬЗЯЩИХ ВЕКТОРОВ

---

Ю.Соловьев, А.Сосинский

*...Да Лебедь рвется в облака.  
Рак пятится назад, а Щука тянется в воду.*

Вы, конечно, помните вынесенные в эпиграф строчки замечательной басни И.А.Крылова. А задумывались ли вы над тем, как рассчитать движение воза, если известно, с какими силами его тащат лебедь, рак и щука? Задача эта совершенно естественная, можно сказать, типичная: к некоторому твердому массивному телу в определенных местах приложены силы – что произойдет? Как ни странно, привычный векторный аппарат, изучаемый в школе на уроках физики и математики, совсем непригоден для решения подобных задач. Когда мы должны принимать в расчет не только массу, но и размеры тела, когда мы имеем дело не с абстрактной материальной точкой, а с настоящим предметом, то не всегда ясно, как складывать векторы, приложенные к предмету в разных местах, что можно и чего нельзя делать с этими векторами и, собственно говоря, что представляют собой такие «настоящие» векторы.

Ответ на эти вопросы – в виде небольшой математической теории (*теории скользящих векторов*) – и составляет предмет нашей статьи.

## Какие бывают векторы?

Вектор в пространстве или на плоскости принято изображать направленным отрезком  $\overrightarrow{AB}$ . Он задается двумя точками: своим *началом* (или *точкой приложения*) *A* и *концом* *B*. Если отрезок *AB* неограниченно продолжить в обе стороны, получится прямая, которая называется *линией действия* вектора *AB*.

Все векторы, в зависимости от того, какие геометрические или физические величины они представляют, могут быть разбиты на три следующих типа.

1°. Может случиться, что два *геометрически равных* вектора изображают одну и ту же физическую или механическую

---

Опубликовано в «Кванте» №8 за 1985 г.

величину. (Напомним, что в геометрии два вектора (направленных отрезка) называются *равными*, если их линии действия параллельны, длины равны и порядок точек задает одно и то же направление на их линиях действия, или, другими словами, если они совмещаются параллельным переносом.) Такого рода векторы, не имеющие ни определенной линии действия, ни определенной точки приложения, называют свободными. Например, вектор магнитной индукции постоянного магнитного поля или вектор скорости одной инерциальной системы относительно другой – свободные векторы; их можно считать приложенными к любой точке. Математики, как правило, тоже изучают свободные векторы; простейший пример – вектор, задающий параллельный перенос: приложенный к любой точке, он указывает своим концом образ этой точки при данном переносе.

2°. С другой стороны, встречаются такие физические величины, что изображающие их векторы не могут быть отделены от своей точки приложения. Такого рода векторы называются связанными (или закрепленными). Так, вектор мгновенной скорости движущейся точки в данный момент времени – пример связанного вектора.

Его нельзя отделить от движущейся точки (если, конечно, все остальные точки пространства не перемещаются с той же скоростью).

3°. Наконец, может случиться, что два геометрически равных вектора изображают равные физические величины лишь при условии, что эти векторы имеют одну и ту же линию действия, и изображают не равные физические величины, если они имеют различные линии действия. Таковы, например, векторы, изображающие силы, действующие на твердое тело. Такие неотделимые от линии действия векторы называются скользящими. Это, так сказать, «настоящие» силовые векторы, реально действующие не на абстрактные «бесконечно малые» тела, а на жесткие предметы, имеющие определенные размеры и форму. Заметим, что здесь условие жесткости тела существенно: тело не растягивается и не сжимается, без потерь передает усилие вдоль линии действия, поэтому не важно, к какой именно точке на линии действия приложена сила; точка приложения может как бы скользить по линии действия.

О скользящих векторах и пойдет наш рассказ. В дальнейшем слово «вектор» без определения следует понимать именно как «скользящий вектор»; а когда речь пойдет о свободных или закрепленных векторах, это будет оговорено.

Мы будем обозначать (скользящие) векторы либо одной полужирной буквой, например  $\mathbf{v}$ , либо двумя буквами с чертой сверху, например  $\overline{AB}$ . Свободные же векторы мы будем обозначать буквами со стрелками, например  $\vec{v}$  или  $\overrightarrow{AB}$ .

### Системы скользящих векторов

Итак, (скользящий) вектор  $\mathbf{v} = \overline{AB}$  задается прямой  $l = AB$  – своей линией действия – и направленным отрезком  $AB$ , лежащим на этой прямой. Два вектора  $\mathbf{v} = \overline{AB}$  и  $\mathbf{u} = \overline{CD}$  считаются совпадающими, если они задают один и тот же свободный вектор ( $\overline{AB} = \overline{CD}$ ) и, кроме того, у них одна и та же линия действия (прямые  $AB$  и  $CD$  совпадают).

В дальнейшем нас будут интересовать не отдельно взятые векторы, а конечные системы векторов ( $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ), ведь именно такие системы отвечают системам сил, действующим на твердое тело. Порядок перечисления векторов, разумеется, не существует; однако один и тот же вектор может встречаться в системе несколько раз, и его нужно считать столько раз, сколько он в ней встречается.

Наша цель – научиться преобразовывать системы друг к другу, сводить – если это возможно – произвольную систему к системам простейшего вида. Для этого нам потребуются

### Элементарные операции и эквивалентность систем векторов

Когда линии действий двух векторов  $\mathbf{v}$  и  $\mathbf{u}$  пересекаются, эти векторы можно сложить естественным образом: если  $\mathbf{u} = \overline{AB}$ ,  $\mathbf{v} = \overline{AC}$  ( $A$  – общая точка линий действия), то

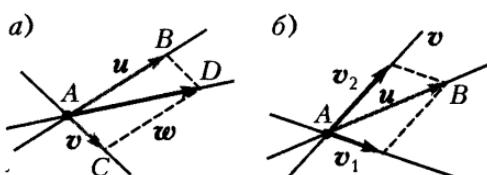


Рис. 1

сложение «пересекающихся» векторов будет нашей первой элементарной операцией; она сводит систему из двух векторов ( $\mathbf{u}, \mathbf{v}$ ) к системе из одного вектора ( $\mathbf{w}$ ).

У этой операции есть обратная: если дан любой вектор  $\mathbf{v} = \overline{AB}$ , то можно выбрать две произвольные прямые, проходящие через любую точку линии действия (скажем, через точку  $A$ ),

суммой  $\mathbf{w} = \mathbf{u} + \mathbf{v}$  векторов  $\mathbf{u}$  и  $\mathbf{v}$  называется вектор  $\mathbf{w} = \overline{AD}$ , где точка  $D$  получена сложением  $\overline{AB} + \overline{AC} = \overline{AD}$  по правилу параллелограмма (рис.1,а).

Такое

и разложить вектор  $\overrightarrow{AB}$  по правилу параллелограмма (рис.1,б). Из одного вектора ( $v$ ) мы получим систему из двух векторов ( $v_1, v_2$ ).

Обратите внимание, что складывать можно только векторы, чьи линии действия пересекаются или совпадают. Особенно просто складываются векторы с общей линией действия. В частности (рис.2), так складываются два противоположных вектора, т.е. векторы вида  $v = \overrightarrow{AB}$  и  $-v = \overrightarrow{BA}$ ; их суммой будет нуль-вектор  $0 = \overrightarrow{AA}$ .

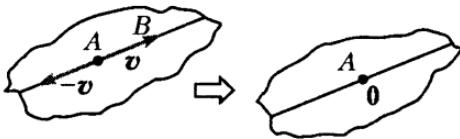


Рис. 2

Наша вторая элементарная операция состоит в *уничтожении нуль-вектора*, т.е. в переходе от системы вида  $(v_1, v_2, \dots, v_k, v, -v)$  к системе  $(v_1, v_2, \dots, v_k)$ . В частности (при  $k = 0$ ), простейшая система  $(v_1, -v)$  сводится к пустой или нулевой системе, которую мы будем обозначать так же, как нуль-вектор — через 0.

У второй элементарной операции тоже есть обратная — *рождение нуль-вектора*, т.е. переход от системы  $(v_1, v_2, \dots, v_k)$  к системе  $(v_1, v_2, \dots, v_k, v, -v)$ , где  $v$  — произвольный вектор.

Мы будем говорить, что две системы векторов  $(u_1, u_2, \dots, u_k)$  и  $(v_1, v_2, \dots, v_k)$  *эквивалентны*, или *приводятся друг к другу*, если от одной системы к другой можно перейти с помощью конечной последовательности элементарных операций.

С точки зрения механики понятно, почему мы интересуемся элементарными операциями и эквивалентными системами. Ведь две эквивалентные системы сил всегда оказывают одинаковое воздействие на твердое тело, в чем можно убедиться не только экспериментально, а просто продумав физический смысл элементарных операций. Приводя же сложные системы к более простым, мы получаем возможность легко разобраться в том, какое именно результирующее воздействие данная сложная система сил оказывает на твердое тело.

### Простейшие системы на плоскости; пары

Рассмотрим теперь подробнее приведение систем векторов к более простым для случая систем векторов на плоскости.

На рисунке 3 показано, как с помощью элементарных операций осуществляется приведение некоторых простейших систем векторов к другим таким системам. Мы советуем читателю

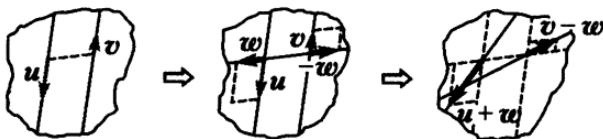
Простейшие преобразования плоских систем:



а) два пересекающихся вектора  $\rightarrow$  один вектор;



б) два сонаправленных вектора  $\rightarrow$  два пересекающихся вектора;



в) два противоположных вектора разной длины  $\rightarrow$  два пересекающихся вектора;



г) два противоположных вектора равной длины (пара)  $\rightarrow$  другая пара.

Рис. 3

внимательно проследить за всеми операциями и даже проделать их самостоятельно на бумаге.

Особое внимание следует обратить на рисунок 3, г. На нем показаны системы из двух противоположных векторов равной длины на параллельных линиях действия; такие системы коротко называют *парами*. Из рисунка видно, что пары не удается упростить, а можно только «повернуть», превратить в другую пару.

На рисунке 4 показана еще одна серия элементарных операций. Здесь происходит не упрощение исходной системы (из одного вектора), а ее усложнение: из данного вектора мы получаем вектор с другой линией действия и пару; впрочем, если

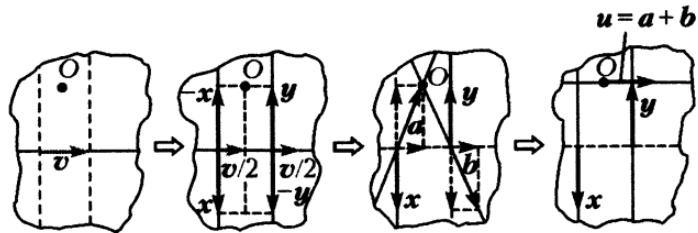


Рис.4. Один вектор  $\rightarrow$  вектор, проходящий через данную точку, и пара

этот рисунок прочитать в обратном порядке (а это осмысленно, ибо каждая элементарная операция имеет обратную), мы получим упрощение (система, состоящая из пары и вектора, приводится к системе из одного вектора). Переход  $(v) \rightarrow (u, x, y)$  стоит запомнить, он нам потребуется несколько позже.

Заметим еще, что приемы приведения плоских систем бывают полезны и при рассмотрении пространственных задач. Так, пространственная система  $(l, p, \varpi, t)$  ( $l = -t$ ,  $|p| = |\varpi|$ ) применением операции разложения вектора приводится (рис.5, а) к паре  $(u, v)$ :  $l = a + v$ ,  $\varpi = b + u$ , а плоская система  $\{a, b, l, t\}$ ,

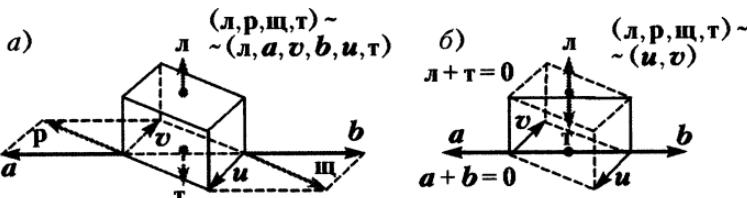


Рис. 5

остающаяся после выделения из системы  $(l, p, \varpi, t)$  пары  $(u, v)$  (рис.5, а), как легко видеть, приводится к нулю (рис.5, б).

### Момент вектора и момент пары

Выше мы видели, что пару векторов нельзя привести к одному скользящему вектору. Однако пара допускает удобную векторную характеристику. Чтобы получить ее, нам придется начать издалека.

Пусть дан вектор  $v = \overline{AB} \neq 0$  и точка  $O$ , не лежащая на его линии действия; тогда *моментом вектора  $v$  относительно точки  $O$*  называется (закрепленный!) вектор  $OM$  (приложенный к точке  $O$ ) с линией действия, перпендикулярной плоскости  $OAB$ , длины  $OM = r \cdot |v|$  (где  $r$  – плечо вектора  $v$  относительно  $O$ , т.е. расстояние от  $O$  до прямой  $AB$ ) и направленный так, чтобы направление вращения  $v$  вокруг  $O$ , наблюдаемое из точки

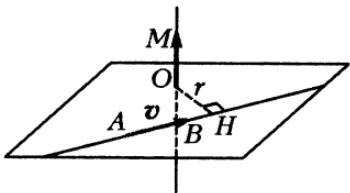


Рис. 6

$M$ , было положительным (против часовой стрелки; рис. 6); если точка  $O$  лежит на линии действия вектора  $v$ , или  $v = 0$ , то момент принимается равным нулю. Момент вектора играет важную роль в механике при изучении вращательного движения и имеет красивые приложения в геометрии.

Пусть теперь дана пара векторов  $v = \overline{AB}$ ,  $v' = \overline{A'B'}$  и точка  $O$ ; пусть  $OM$  и  $OM'$  – моменты  $v$  и  $v'$  относительно  $O$ . Тогда легко проверить (сделайте это!), что длина и направление суммы моментов относительно точки  $O$  векторов  $v$  и  $v'$  не зависят от выбора этой точки, т.е. эта сумма является свободным вектором. Его называют (*векторным*) *моментом пары*  $(v, v')$ . Легко проверяется, что длина векторного момента пары  $(v, v')$  равна  $d \cdot |v|$ , где  $d$  – расстояние между параллельными прямыми  $AB$  и  $A'B'$ . Пользуясь этим, а также утверждением задачи 3, нетрудно доказать, что *пара приводится к нулю тогда и только тогда, когда ее векторный момент равен нулю*.

### Приведение плоских систем к паре или вектору

Мы сейчас докажем следующую замечательную теорему.

*Всякая (конечная) система (скользящих) векторов на плоскости приводится либо к одному вектору, либо к паре.*

**Доказательство.** Если данная система состоит из одного вектора – теорема доказана. Если в системе есть векторы с пересекающимися линиями действия, то их можно попарно упрощать (см. рис. 3,а), пока не останутся лишь векторы с параллельными линиями действия. Если таких векторов три или более, то по крайней мере два из них сонаправлены, и мы можем упрощать систему дальше (см. рис. 3,б). Таким образом, переходы 3,б и 3,а позволяют нам свести доказательство теоремы к случаю двух противоположно направленных векторов с параллельными линиями действий. Если длины векторов различны, то переход 3,в с последующим переходом 3,а даст нам один вектор, а если длины равны, то мы получили пару; теперь теорема доказана полностью.

Заметим, во-первых, что доказательство не только устанавливает справедливость теоремы, но и дает эффективный способ для

того, чтобы найти вектор (или пару), эквивалентный исходной системе. Во-вторых, если заранее выбрать некоторую фиксированную точку, то можно так проводить наши элементарные операции, чтобы привести данную систему к системе из одной пары и одного вектора, чья линия действия проходит через эту точку. Действительно, это верно, если система приводится к одной паре. Если же она приводится к одному вектору, то операция, показанная на рисунке 4, позволяет привести этот вектор к паре и вектору, чья линия действия проходит через заданную точку.

### Исчисление систем векторов; базис

Мы научились приводить системы плоских векторов к простейшим системам. Однако оказывается, что системы можно не только приводить одну к другой, но и выражать одну через другие. При этом возникает исчисление систем векторов, очень похожее на обычное исчисление свободных векторов в пространстве.

Чтобы описать это исчисление, определим сумму двух систем  $N = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$  и  $M = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l)$  как их «свободное объединение», т.е. систему

$$L = N + M = N = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_l)$$

(равенство систем векторов понимается здесь и ниже как их эквивалентность), и произведение системы  $N$  на число  $\lambda$  – как новую систему

$$\lambda N = (\lambda \mathbf{u}_1, \lambda \mathbf{u}_2, \dots, \lambda \mathbf{u}_k),$$

где  $\lambda \mathbf{u}$  обозначает вектор с той же линией действия, что  $\mathbf{u}$ , длины  $|\lambda| \cdot |\mathbf{u}|$ , направленный так же, как  $\mathbf{u}$ , если  $\lambda > 0$ , и в противоположную сторону, если  $\lambda < 0$ .

Скажем, что система  $N$  линейно выражается через системы  $M_1, \dots, M_n$ , если существуют числа  $\lambda_1, \lambda_2, \dots, \lambda_n$ , такие что

$$N = \lambda_1 M_1 + \lambda_2 M_2 + \dots + \lambda_n M_n.$$

Рассмотрим теперь случай систем на плоскости. Пусть  $O$  – произвольная точка плоскости,  $I$  и  $J$  – одновекторные системы, состоящие из ненулевых неколлинеарных векторов, приложенных к точке  $O$ , а  $K$  – любая (ненулевая) пара. Тогда системы  $I, J, K$  образуют базис в множестве систем плоских векторов, т.е. любая система  $N$  однозначно линейно выражается через  $I, J, K$ :

$$N = \alpha I + \beta J + \gamma K$$

(числа,  $\alpha, \beta, \gamma$  однозначно определены системой  $N$  при фиксированных  $I, J, K$ ).

Читатель, знакомый с понятием векторного пространства, конечно, понял, что это утверждение означает, что *множество классов эквивалентных систем (скользящих) векторов на плоскости относительно операций, определенных выше, образует векторное пространство размерности 3*.

\* \* \*

Дойдя до этого места, авторы серьезно задумались, а потом заспорили: о каких именно приложениях теории скользящих векторов рассказать? Где только эта теория не используется! В статике (теория ферм и веревочных многоугольников), в строительной механике, в кинематике и динамике твердого тела, а ее обобщения используются в некоторых новейших разделах геометрии. Но нельзя объять необъятное. Здесь мы приведем только один пример – из кинематики твердого тела, а также, в конце статьи, ряд задач.

### **Вращение твердого тела**

Здесь нас интересует такая *кинематическая задача*: как описать движение твердого тела, вращающегося около некоторой оси, если эта ось в свою очередь вращается вокруг неподвижной оси? Мы будем считать задачу решенной, если дан способ нахождения вектора скорости любой наперед заданной точки тела (относительно неподвижной системы отсчета).

Ответ красиво формулируется в терминах скользящих векторов. Однако прежде чем его привести, мы немного расскажем о том, как вообще может двигаться твердое тело. Начнем с простейших примеров.

При *равномерном прямолинейном движении* твердое тело перемещается с постоянной скоростью в фиксированном направлении. При этом векторы скорости во всех точках одинаковы, их направление и величина неизменны во времени. Все точки движутся по прямым. Движение характеризуется одним свободным вектором скорости  $\vec{v}$ .

При *равномерном вращательном движении* тело движется с неизменной угловой скоростью около неподвижной оси. При этом вектор скорости любой точки на оси вращения – нулевой, а вектор скорости любой другой точки перпендикулярен плоскости, проходящей через эту точку и ось, и по величине пропорционален расстоянию от точки до оси. Точки оси остаются на месте, а остальные точки двигаются по окружностям с центрами на оси. Такое движение характеризуется одним (скользящим) вектором

$\omega$  – угловой скоростью вращения, линия действия которого совпадает с осью вращения.

При равномерном винтовом движении тело равномерно вращается около оси (называемой винтовой осью) и одновременно равномерно перемещается вдоль нее. Вектор скорости каждой точки равен сумме вектора вращательного движения и вектора прямолинейного движения (рис.7). Точки тела описывают винтовые линии, лишь точки винтовой оси двигаются прямолинейно вдоль нее самой. Движение характеризуется двумя векторами ( $\vec{v}$ ,  $\omega$ ): (свободным) вектором перемещения  $\vec{v}$  и (скользящим) вектором угловой скорости  $\omega$  (рис.7).

Возможные движения твердого тела в пространстве не исчерпываются, конечно, приведенными примерами, хотя бы потому, что в реальной ситуации векторы  $\vec{v}$  и  $\omega$  могут меняться во времени. Однако имеет место вот какой замечательный факт (см. задачу 10): сколь бы сложное движение ни совершало твердое тело, мгновенное распределение скоростей его точек будет таким же, как при одном из трех перечисленных выше типов движения. Отметим, кстати, что прямолинейное и вращательное движения можно считать частными случаями винтового. Два конкретных примера таких более сложных движений показаны на рисунках 8 и 9.

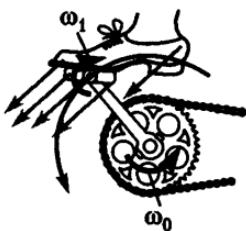


Рис. 8

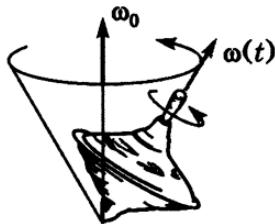


Рис. 9

Вернемся теперь к нашей изначальной кинематической задаче: предположим, что подвижная ось  $M_tN_t$  твердого тела вращается с постоянной угловой скоростью  $\omega_0$  около неподвижной оси  $AB$ , а само тело к тому же вращается вокруг оси  $M_tN_t$  с постоянной по величине угловой скоростью  $\omega_1$ . (Подчеркнем, что как скользящий вектор эта угловая скорость переменная,  $\omega_1 = \omega_1(t)$ , – меняется положение ее линии действия  $M_tN_t$ ) Как описать результирующее движение? Ответ зависит от взаимного расположения осей  $AB$  и  $M_tN_t$ .

1°. Оси совпадают. Тогда тело вращается около неподвижной оси  $AB = M_t N_t$  с постоянной угловой скоростью  $\omega = \omega_0 + \omega_1$ . (Пример: движение часовой или минутной стрелки часов, лежащих на Северном полюсе.)

2°. Оси параллельны, причем  $\omega_0 = -\omega_1$ . Тогда в любой момент  $t$  скорости всех точек тела такие же, как если бы оно равномерно вращалось с угловой скоростью  $\omega = \omega_0 + \omega_1$  (здесь имеется в виду сумма скользящих векторов — см. рис.3 и рис.10) вокруг линии действия  $P_t Q_t$  вектора  $\omega$ . Говорят, что тело имеет *мгновенную угловую скорость  $\omega$  и мгновенную ось вращения  $P_t Q_t$* . В данном случае мгновенная ось вращается вокруг неподвижной оси  $AB$ , оставаясь ей параллельной.

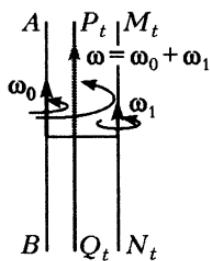


Рис. 10

3°. Оси параллельны, причем  $\omega_0 = -\omega_1$ . Тогда тело совершают *поступательное движение* с вектором скорости  $\vec{v}$ , равным моменту пары  $(\omega_0, \omega_1)$ , т.е. в любой момент  $t$  скорости всех точек тела одинаковы и равны  $\vec{v}$ . (Пример: движение подвижной части педали; см. рис.8.)

4°. Оси пересекаются. Этот случай аналогичен 2°: в любой момент  $t$  тело имеет мгновенную ось вращения и мгновенную угловую скорость  $\omega = \omega_0 + \omega_1$ . Разница в том, что здесь мгновенная ось описывает не цилиндр, а конус с вершиной в неподвижной точке тела — точке пересечения осей. (Пример: прецессионное движение волчка; рис. 9.)

5°. Оси скрещиваются. В этом случае тело имеет, так сказать, «*мгновенную винтовую ось*»; подробнее об этом см. в задачах 7, 8.

Мы видим, таким образом, что решение нашей кинематической задачи коротко и просто формулируется в терминах сложения скользящих векторов. Несложные доказательства утверждений 1° — 4° мы оставляем читателю в качестве упражнений.

### Разные задачи

1. Назовем *главным вектором* системы скользящих векторов  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  сумму равных им свободных векторов  $\vec{v}_1 + \dots + \vec{v}_n$ , а *главным моментом* этой системы относительно точки  $O$  — сумму моментов векторов  $\mathbf{v}_1, \dots, \mathbf{v}_n$  относительно  $O$ . Докажите, что у любых двух эквивалентных систем главные векторы и главные моменты относительно одной и той же точки  $O$  совпадают.

**2.** Покажите, что любая система векторов, расположенных в одной плоскости, эквивалентна трем векторам, направленным вдоль сторон треугольника, произвольно выбранного в той же плоскости.

**3.** Докажите, что плоская система векторов, перпендикулярных к сторонам выпуклого  $n$ -угольника в их серединах, эквивалентна нулю, если длины векторов пропорциональны соответствующим сторонам и все векторы обращены внутрь многоугольника (или все наружу).

**4.** Сформулируйте и докажите утверждение, аналогичное утверждению предыдущей задачи, для тетраэдра.

**5\*.** Покажите, что любая система векторов пространства эквивалентна шести векторам, направленным по ребрам произвольно выбранного тетраэдра.

**6\*.** Покажите, что любая система векторов пространства эквивалентна системе, состоящей из одного вектора (проходящего через произвольно выбранную точку) и из одной пары. Выведите отсюда, что системы скользящих векторов в пространстве, рассматриваемые с точностью до эквивалентности, образуют шестимерное векторное пространство.

*Указание.* Сначала приведите систему к трем скользящим векторам, проходящим через три произвольно выбранные точки (раскладывая каждый вектор системы по трем направлениям), затем к двум векторам, один из которых проходит через данную точку, и, наконец, примените построение из рисунка 4.

**7.** Твердое тело вращается с угловой скоростью  $\omega$  около оси, которая в свою очередь движется равномерно и прямолинейно со скоростью  $\vec{v}$ . Докажите, что тело: а) совершает винтовое движение, если  $\omega \parallel \vec{v}$ ; б) имеет мгновенную ось вращения, если  $\omega \perp \vec{v}$ ; в) имеет мгновенную винтовую ось, если  $\vec{v}$  не параллельно и не перпендикулярно  $\omega$  (это значит, что все точки некоторой прямой – «винтовой оси» – имеют одинаковые скорости  $\vec{v}_1$ , направленные вдоль этой оси, а векторы скорости остальных точек равны сумме вектора  $\vec{v}_1$  и вектора скорости мгновенного вращения данной точки около винтовой оси).

**8\*.** Покажите, что в случае скрещивающихся осей  $AB$  и  $M_t N_t$  в кинематической задаче, рассмотренной выше, тело имеет мгновенную винтовую ось, параллельную вектору  $\vec{\omega}_0 + \vec{\omega}_1$  и пересекающую общий перпендикуляр осей  $AB$  и  $M_t N_t$ . Уточните положение мгновенной винтовой оси и вектор скорости вдоль оси.

*Указание.* Воспользуйтесь задачей 7.

**9\*.** Опишите результирующее движение тела, совершающего мгновенное винтовое движение относительно оси, которая в свою очередь совершает винтовое движение относительно некоторой неподвижной оси.

**10\***. Докажите, что как бы ни двигалось твердое тело, в любой данный момент распределение скоростей его точек такое же, как если бы оно совершало равномерное (прямолинейное, вращательное или винтовое) движение.

**11.** Пользуясь тем, что момент вектора  $\overline{AB}$  относительно точки  $O$  равен по величине удвоенной площади треугольника  $OAB$ , докажите, что множество точек внутри выпуклого многоугольника, для которых постоянна сумма расстояний до его сторон (точнее, прямых, содержащих стороны), есть либо отрезок, либо весь многоугольник, либо пустое множество.

### **Когда в товарищах согласья нет...**

Вернемся теперь к басне. Предположим, что тяжелый воз, по форме напоминающий прямоугольный параллелепипед, стоит на дороге. На него действует сила тяжести  $t$ , которую можно считать приложенной к центру тяжести, но она компенсируется реакцией опоры: воз стоит на месте. Но вот явились преисполненные благих намерений товарищи: Лебедь, Рак и Щука и стали тянуть воз с силами  $l$ ,  $r$  и  $Щ$  (рис.5). Что же произойдет?

Вооруженные теорией, мы теперь можем ответить на этот вопрос. Система сил ( $l$ ,  $r$ ,  $Щ$ ,  $t$ ) – как и любая другая – приводится к паре и вектору (см. задачу 6). Как их найти? А мы это уже сделали, когда разбирали пример в разделе «Простейшие системы на плоскости; пары». У нас получилась пара  $(u, v)$  в горизонтальной плоскости. И теперь ясно, что будет: не сходя с места, воз начнет вращаться там, где стоял. Мы можем поэтому с уверенностью подтвердить слова И. А. Крылова:

...Да только воз и ныне там!

## ВЫЗОВ ВАН РОУМЕНА

---

Ю.Соловьев

В начале октября 1594 года король Франции Генрих IV принимал в Фонтенбло посланника Республики Соединенных Нидерландов, которую по имени самой крупной ее провинции чаще всего называли Голландией. Возникшая в результате длительной и упорной борьбы против испанского владычества, республика была очень молода — ей шел всего второй десяток лет. Война с Испанией не прекращалась — голландское правительство настойчиво искало союзников. А во Франции только что погасли пожары многолетней междоусобицы, и Генрих, преодолев бешеное сопротивление оппозиции, поддерживаемой испанцами, стал французским королем. Генрих не скрывал своего интереса к Голландии как к союзнику в борьбе с Испанией, но более всего его интересовали бурно развивающиеся голландская промышленность, торговля и мореплавание. Поэтому, прогуливаясь по парку в Фонтенбло, он внимательно слушал рассказ посланника о новых шелковых мануфактурах Роттердама, бумажных фабриках Уtrechtа, корабельных верфях Зандама.

— В Голландии сейчас много талантливых инженеров и ученых, — говорил посланник. — Математик и механик Симон Стевин разрабатывает новые системы шлюзов и плотин, по проектам математика Лудольфа Цейлена возводятся крепости, математик Андриен Ван Роумен славится своими головоломными вычислениями. Кстати, — продолжал посланник, — не так давно Ван Роумен сделал вызов математикам всего мира. Он разоспал во многие страны письмо, в котором предлагает решить придуманную им задачу. Но пока это никому не удалось.

— Победителем непременно будет француз, — засмеялся король.

— Ваше величество, — заметил посланник, — я привез это письмо, но по-видимому, Франция не имеет выдающихся математиков, поскольку Ван Роумен среди тех, кому он адресовал свой вызов, не упомянул ни одного француза.

– И все же у меня есть математик, и весьма выдающийся, –  
ответил Генрих.– Позовите Виета.

Так в этот осенний день столкнулись судьбы двух очень непохожих людей.

Андрисен Ван Роумен – родился в 1561 году в г.Лувене в Испанских Нидерландах (ныне Бельгия). Изучал медицину и математику в Лувенском университете. В Лувене же получил степень доктора. Преподавал математику в Лейдене и Вюрцбурге. Занимался исследованиями по геометрии и тригонометрии, а также практическими вопросами астрономии и навигации. Получил ряд частных результатов о разложении функций  $\sin nx$  и  $\cos nx$  по степеням  $\sin x$  и  $\cos x$ , определил значение числа  $\pi$  с семнадцатью десятичными знаками, т.е. с наивысшей точностью для Европы того времени. При жизни Ван Роумен был очень знаменит в Голландии и Германии, но со временем его работы утратили свое значение, и сейчас упоминания о нем можно разыскать лишь в самых толстых энциклопедиях.

Франсуа Виет – родился в 1540 году в г.Фонтене. С 1559 года занимался адвокатской деятельностью, серьезно интересуясь при этом математикой и астрономией. В 1571 г. переехал в Париж, где продолжал адвокатскую деятельность и завязал знакомства с парижскими математиками. В 1573 году стал советником парламента в Бретани, затем частным советником короля Генриха III. В 1580 году получил должность королевского докладчика по ходатайствам. В последние годы правления Генриха III занимался расшифровкой переписки между противниками короля и испанцами. Нашел ключ к сложному шифру, который использовался испанским королем Филиппом II и его генералами. Филипп II, узнав из перехваченных французских депеш, что его секретную корреспонденцию читают при французском дворе, в гневе принес жалобу Папе Римскому, указывая, что расшифровка явно проводилась с помощью колдовства и черной магии. После убийства Генриха III в августе 1589 г. перешел на службу к Генриху Наваррскому, будущему французскому королю Генриху IV. Автор большого числа работ по алгебре, геометрии, тригонометрии, астрономии. Установил зависимость между корнями и коэффициентами алгебраического уравнения, нашел разложения функций  $\sin nx$  и  $\cos nx$  по степеням  $\sin x$  и  $\cos x$ , создал современную алгебраическую символику. Работы Виета, написанные тяжелым, усложненным языком, остались непонятными его современниками и лишь спустя почти полвека после его смерти оказали огромное влияние на развитие алгебры и геометрии.

Но вернемся в Фонтенбло. Когда появился Виет, посланник достал письмо Ван Роумена. В нем предлагалось решить урав-

нение

$$45x - 3795x^3 + 95634x^5 - 1138500x^7 + 7811375x^9 - 34512075x^{11} + \\ + 105306075x^{13} - 232676280x^{15} + 384942375x^{17} - 488494125x^{19} + \\ + 483841800x^{21} - 378658800x^{23} + 236030652x^{25} - 117679100x^{27} + \\ + 46955700x^{29} - 14945040x^{31} + 3764565x^{33} - 740259x^{35} + \\ + 111150x^{37} - 12300x^{39} + 945x^{41} - 45x^{43} + x^{45} = a,$$

в частности, при

$$a = \sqrt{1\frac{3}{4} - \sqrt{\frac{5}{16}}} - \sqrt{1\frac{7}{8} - \sqrt{\frac{45}{64}}}.$$

Для облегчения задачи Ван Роумен сообщал ответы, которые получаются при двух других значениях  $a$  (записываемых еще более громоздко).

Виет прочел письмо и сразу же написал решение. Посланник сказал, что в резиденции у него имеется запечатанный конверт с решением самого Ван Роумена. В присутствии нотариуса он вскроет его и завтра ответит, прав ли Виет. На следующий день голландец подтвердил, что решение Виета верно, а Виет, в свою очередь, прислал еще двадцать два других решения, не известных Ван Роумену. Кроме того, Виет указал на ошибку в условии задачи (сделанную либо переписчиками, либо самим Ван Роуменом).

Попробуем разобраться, как удалось Виету необыкновенно быстро решить столь чудовищное на первый взгляд уравнение. Обратимся для этого к некоторым его математическим работам. Главными тригонометрическими результатами, полученными Виетом, являются выражения для синусов и косинусов кратных дуг. Виет получает эти выражения в форме правила, указывающего, как чисто механически можно получить их. Это правило напоминает известное правило, использующее треугольник Паскаля, разница лишь в том, что вместо треугольника Паскаля Виет пользуется следующей таблицей:

1	1	1	1	1	1	1	...
1	2	3	4	5	6	...	
1	3	6	10	15	21	...	
1	4	10	20	35	56	...	
1	5	15	35	70	126	...	
1	6	21	56	126	252	...	
1	7	28	84	210	462	...	

Каждое число здесь получается посредством сложения чисел, стоящих перед ним и над ним. Стоит отметить, что Виет выражает не  $\sin nx$  и  $\cos nx$  через  $\sin x$  и  $\cos x$ , как это делаем мы, а  $2\sin nx$  и  $2\cos nx$  через  $2\sin x$  и  $2\cos x$ . Если в образованных таким образом выражениях рассматривать величины  $2\sin nx$  или  $2\cos nx$  как известные, то мы получим уравнения  $n$ -й степени относительно неизвестных  $2\sin x$  и  $2\cos x$ .

Первоначальная цель, ради которой Виет выводил формулы синусов кратных дуг, заключалась в вычислении синусов дуг через синусы малых дуг, т.е. в составлении таблицы синусов. Затем эти формулы нашли применение в алгебре и геометрии. В частности, геометрическую задачу о трисекции угла  $\alpha$  Виет связал с уравнением  $3x - x^3 = a$ , которому удовлетворяют значения  $a = 2\sin \alpha$ ,  $x = 2\sin(\alpha/3)$ . Два положительных решения этого уравнения Виет трактует как хорды, отвечающие дугам  $2\alpha/3$  и  $(360^\circ - 2\alpha)/3$ . Отрицательный корень он, по обыкновению своего времени, вообще не принимает во внимание. Точно так же задачу о делении угла на пять равных частей Виет связывает с уравнением  $5x - 5x^3 + x^5 = a$ , которому удовлетворяют значения  $a = 2\sin \alpha$ ,  $x = 2\sin(\alpha/5)$ .

Из сказанного ясно, в какой степени был подготовлен Виет, чтобы мгновенно решить задачу Ван Роумена. Он сразу же увидел, что предложенное значение  $a$  является длиной стороны правильного пятнадцатигольника, вписанного в единичный круг (проверьте это!), или, что то же самое, хордой, отвечающей дуге  $24^\circ$ . Коэффициенты при первом и предпоследнем членах левой части позволили ему предположить, что левая часть есть не что иное, как выражение  $2\sin 45\alpha$  через  $2\sin \alpha$ . Но поскольку  $a = 2\sin 12^\circ$ , то  $\alpha = 12^\circ/45 = 4^\circ/15$ , и значит,  $x = 2\sin(4^\circ/15)$ . Именно это решение и передал Виет голландскому посланнику.

После аудиенции Виет проверил свое предположение. Однако, проделав необходимые вычисления, он обнаружил, что левая часть предложенного уравнения не совпадает с разложением  $2\sin 45\alpha$  по степеням  $2\sin \alpha$ !

Наверное, в этот момент он чувствовал себя не очень уютно, скорее всего, просто скверно.

Что произошло? Может быть, это всего лишь ошибка в весьма утомительных вычислениях? Видимо, в этот момент Виет и сумел найти другой – геометрический – способ выразить  $2\sin 45\alpha$  через  $2\sin \alpha$ : для того чтобы разделить дугу на 45 частей, нужно разделить ее сначала на пять частей, затем каждую часть на три и снова на три. Другими словами, левую часть уравнения Ван

Роумена можно получить из системы

$$\begin{cases} 3z - z^3 = a, \\ 3y - y^3 = z, \\ 5x - 5x^3 + x^5 = y. \end{cases}$$

Лишь проанализировав те ответы, которые сообщал Ван Роумен для двух других значений  $a$ , Виет убедился, что речь действительно идет о делении дуги на сорок пять частей, и смело исправил ошибку в условии. Но Виет не ограничился этим решением. Двадцать два других решения, сообщенных им на следующий день, имели вид

$$2 \sin \frac{360^\circ k + 12}{45} = 2 \sin \frac{120^\circ k + 4^\circ}{15}, \quad k = 1, 2, \dots, 22.$$

Таким образом, Виету удалось найти все положительные корни (напомним, что только они считались решениями в его время).

Здесь можно было бы поставить точку, однако математическое соревнование между Виетом и Ван Роуменом на этом не кончилось. Спустя короткое время Виет предложил Роумену построить циркулем и линейкой окружность, касающуюся трех данных окружностей (задача Аполлония). Вскоре Виет указал изящное геометрическое построение, выполнимое только циркулем и линейкой.

Рассказывают, что, потерпев второе поражение, Ван Роумен стал ревностным почитателем Виета и даже приезжал к нему учиться.

# АРИФМЕТИКА ЭЛЛИПТИЧЕСКИХ КРИВЫХ

---

Ю.Соловьев

Античная математика оставила в наследство несколько великих сочинений. Одно из них – «Арифметика» Диофанта Александрийского. Удивительна судьба этой книги. Написанная в III веке н.э., она исчезла более чем на тысячелетие и считалась утраченной. Лишь в 1464 г. немецкий ученый Региомонтан случайно обнаружил 6 из 13 книг «Арифметики». В первый раз она была напечатана в латинском переводе в 1575 г. После ее издания 1621 г., подготовленного Баше де Мезириаком, она стала настольной книгой многих математиков, например, П.Ферма (1601–1665) и Р.Декарта (1596–1650).

За тысячелетие книга совсем не устарела – она сильно опережала уровень лучших алгебраических исследований XVI века. Посудите сами: в отличие от европейских алгебраистов этого времени, Диофант свободно оперировал отрицательными и рациональными числами, владел буквенной нотацией для уравнений, а самое главное – умел находить решение в целых и рациональных числах линейных, квадратных, кубических уравнений и систем с двумя и более неизвестными с целыми коэффициентами. Решение таких уравнений – они теперь называются диофантовыми – с тех пор находится в центре внимания математиков.

О решении некоторых диофантовых уравнений – по-моему, самых красивых – я и собираюсь рассказать вам. Для этого придется не только бросить пристальный взгляд на сочинения великого Диофанта, но и коснуться самых последних событий современной математической жизни.

## Метод секущих Диофанта

Проиллюстрируем этот метод на конкретном примере – частном случае одного из тех, которые Диофант разбирает в своей «Арифметике». Пусть дано уравнение

$$x^2 - y^2 = 1 \quad (1)$$

и требуется найти все его рациональные решения, т.е. найти все

---

Опубликовано в «Кванте» №7 за 1987 г.

пары

$$(x; y) = \left( \frac{a}{b}; \frac{c}{d} \right), \quad a, b, c, d \in \mathbb{Z},$$

обращающие уравнение (1) в числовое тождество.

Уравнение (1), как и любое уравнение от переменных  $x, y$ , можно рассматривать как кривую на плоскости  $Oxy$ . В данном случае это гипербола (рис.1). Сразу бросается в глаза решение  $(1; 0)$  – точка пересечения  $P$  кривой с осью  $Ox$ . Проведем через эту точку секущую

$$y = k(x - 1) \quad (2)$$

и найдем ее вторую точку пересечения с кривой (1). Для этого подставим выражение (2) для  $y$  в уравнение (1) и решим получившееся квадратное уравнение относительно  $x$ . Получим

$$x_{1,2} = \frac{-k^2 \pm 1}{1 - k^2}.$$

Корень  $x_1 = 1$  нам и так известен (он относится к точке  $(1; 0)$ ), а второй корень  $x_2 = (k^2 + 1)/(k^2 - 1)$  дает нам искомую вторую точку

$$(x_2; y_2) = \left( \frac{k^2 + 1}{k^2 - 1}; \frac{2k}{k^2 - 1} \right). \quad (3)$$

Для любого рационального  $k$  ( $k \neq \pm 1$ ) эта формула определяет точку на нашей кривой, а значит и рациональное решение данного уравнения. (При  $k = \pm 1$  секущая пересекает кривую только в точке  $P$  (см. рис.1).) Обратно, для любого рационального решения, т.е. рациональной точки  $M$  на кривой, секущая  $PM$  задается уравнением (2) с рациональным  $k$  (ибо тогда катеты прямоугольного треугольника  $PMH$  рациональны).

Таким образом, формула (3) при всевозможных рациональных  $k \neq \pm 1$  дает все решения в рациональных числах уравнения (1).

Сам Диофант, конечно, не вводил в рассмотрение систему координат  $Oxy$ , не рассматривал кривую данного уравнения – метод координат появился лишь в работах Декарта в XVII веке. Диофант делал подстановку (1) чисто алгебраически и получил – разумеется, в другой записи – формулу (3). Более того, он понимал, что продемонстрированный метод с успехом применим

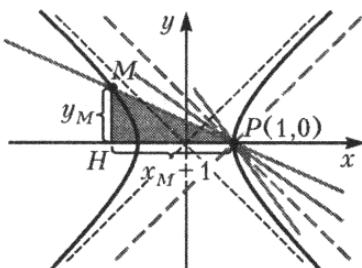


Рис. 1

не только к многочлену  $x^2 - y^2 - 1$ , но и вообще к многочлену второй степени от двух переменных общего вида:

$$p(x, y) = ax^2 + bxy + cy^2 + dx + ey + f,$$

где  $a, b, \dots, f$  целые (или рациональные) числа, при условии, что у многочлена удалось найти хотя бы один рациональный корень.

Не на всякой кривой второй степени имеются рациональные точки; например, их нет на окружности  $x^2 + y^2 = 3$  или на эллипсе  $x^2 + 82y^2 = 3$ .<sup>1</sup> Задача о существовании хотя бы одной рациональной точки на кривой второй степени оказалась очень трудной. Первые нетривиальные продвижения в ее решении получили индийские математики Брахмагупта (VII век) и Бхаскара (XII век), а окончательный ответ был найден лишь в 1768 г. французским математиком Ж.-Л.Лагранжем (1736–1813).

Диофант не ограничился уравнениями второй степени. Он с успехом берется и за третью степень, демонстрируя общий прием в одной конкретной задаче.

### Одна задача «Арифметики» Диофанта. Касательная

В этой задаче требуется найти рациональное решение уравнения

$$y(6-y) = x^3 - x. \quad (4)$$

Короткое решение, содержащее в зародыше замечательную идею, Диофант излагает с незаурядным мастерством. Попробуем, пишет он, замену  $x = 2y - 1$  (разумеется, его обозначения совсем другие). Тогда получим

$$6y - 6y^2 = 8y^3 - 12y^2 + 4y.$$

Если бы 6 равнялось 4, как хорошо бы сократились члены с  $y$  в первой степени! Но число 4 взялось из двойки в замене  $x = 2y - 1$ . Так заменим его тройкой, т.е. возьмем  $x = 3y - 1$ . Тогда линейные члены сокращаются, и мы находим

$$y^2(9y - 7) = 0, \quad (5)$$

откуда  $y = 7/9$  и  $x = 16/9$ . Получено рациональное решение  $(16/9; 7/9)$  кубического уравнения (4).

<sup>1</sup> Рациональные точки  $(a/b; b/c)$  имеются, однако, на окружности  $x^2 + y^2 = 1$ . Для таких точек тройка целых чисел  $(a, b, c)$  называется *пифагоровой* – она удовлетворяет соотношению  $a^2 + b^2 = c^2$ . Все пифагоровы тройки могут быть найдены, например, методом секущих (см. «Квант» №1 за 1987 г.).

На первый взгляд здесь нет ничего особенного – просто удачная замена  $x = 3y - 1$  позволила найти решение. В чем же «глубокая идея»? Чтобы ответить на этот вопрос, вновь воспользуемся методом координат и построим график кривой (4) – см. рисунок 2.<sup>2</sup> На этом рисунке показана прямая  $x - 3y + 1 = 0$ . Она касается нашей кривой в точке  $P(-1; 0)$  (действительно, уравнение (5) имеет, кроме корня  $y = 7/9$ , еще и «два слившихся корня»  $y^2 = 0$ ).

По этому пути можно было бы пойти дальше: через полученную рациональную точку  $(16/9; 7/9)$  построить еще одну касательную к кривой (4) до пересечения с ней в третьей рациональной точке и т.д. Но Диофант не сделал этого шага. И потребовалось более 1500 лет, прежде чем математики сумели до конца воспользоваться идеями Диофанта.

### Кривые третьей степени

Оставаясь верными геометрическому подходу, рассмотренному выше, мы сосредоточим наше внимание не на решении уравнений третьей степени, а на следующем эквивалентном вопросе: *каковы рациональные точки кривой на плоскости, задаваемой уравнением третьей степени*

$$f(x, y) = ax^3 + bx^2y + \dots + hx + iy + j = 0$$

с целочисленными коэффициентами?

Все такие кривые можно разбить на два больших класса. К первому классу отнесем те кривые, у которых имеются точки заострения (как точка  $(0; 0)$  кривой  $y^2 = x^2$ ), самопересечения (рис.3), а также кривые, для которых  $f(x, y)$  представляется в виде

$$f(x, y) = f_1(x, y) \cdot f_2(x, y), \quad \text{Рис. 3}$$

где  $f_1(x, y)$ ,  $f_2(x, y)$  – многочлены меньших степеней (рис.4). Назовем такие кривые *вырожденными*. Второй класс образуют

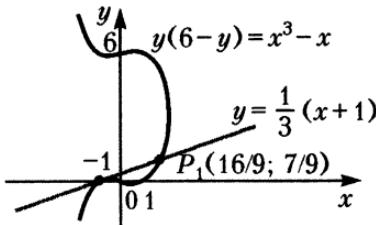
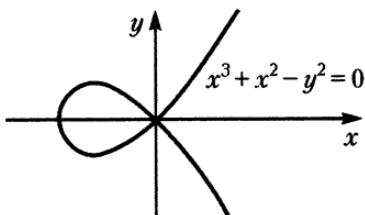


Рис. 2



<sup>2</sup> О том, как строить графики подобных кривых, мы подробно расскажем ниже.

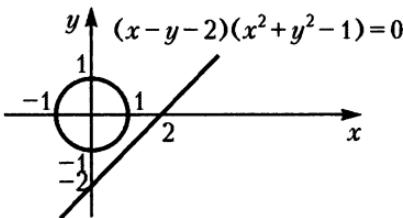


Рис. 4

невырожденные кривые третьей степени с целочисленными коэффициентами – такие кривые называются **эллиптическими**. Именно этот (наиболее общий) класс и будет нас интересовать. Мы будем рассматривать эллиптические кривые, заданные в так называемой **канонической форме**, т.е. уравнением

$$y^2 = x^3 + ax^2 + bx + c \quad (6)$$

с целыми коэффициентами  $a$ ,  $b$  и  $c$ , в котором многочлен  $P(x) = x^3 + ax^2 + bx + c$  не имеет кратных корней.

Это не нарушает общности: любую неособую кривую  $f(x, y) = 0$  третьей степени можно некоторым преобразованием привести к виду (6); при этом, если коэффициенты  $f(x, y)$  были целыми, задачу отыскания рациональных точек на кривой  $f(x, y) = 0$  можно свести к аналогичной задаче для кривой (6) с целыми  $a$ ,  $b$  и  $c$ .

### Графическое изображение эллиптических кривых

Прежде всего выясним, как выглядит кривая (6). Чтобы получить ее **графическое изображение**, нужно нарисовать график функции  $y = \sqrt{x^3 + ax^2 + bx + c}$  и симметрично отразить его на оси  $Ox$ . Для построения графика  $y = \sqrt{x^3 + ax^2 + bx + c}$  построим вначале график функции  $y = x^3 + ax^2 + bx + c$ . Известно, что у многочлена третьей степени (без кратных корней) могут быть либо один, либо три вещественных корня. По предположению все эти корни различны. Поэтому график  $y = x^3 + ax^2 + bx + c$  выглядит так, как показано на рисунке 5,

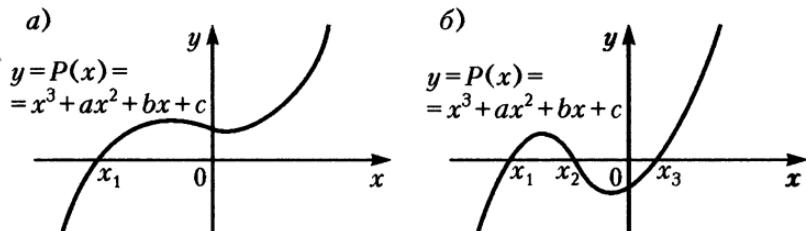


Рис. 5

*а и б.* А теперь уже нетрудно получить график функции

$$y = \sqrt{x^3 + ax^2 + bx + c}$$

(рис.6,*a*) и, тем самым, вид эллиптической кривой  $y^2 = x^3 + ax^2 + bx + c$  (рис.6,*б*) для случая кривой 5,*а*. Случай

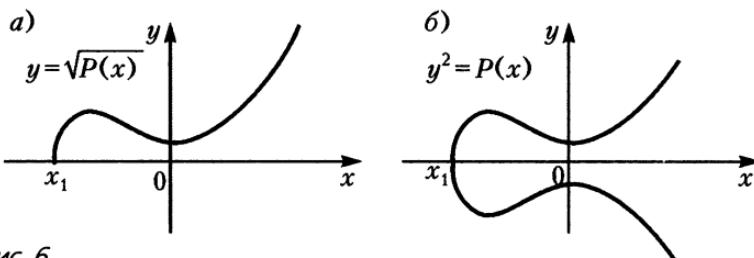


Рис. 6

кривой 5,*б* мы оставляем читателю; результирующая кривая состоит из двух кусков (см. рис.9).

Отметим следующее обстоятельство. Графики функций  $y = \sqrt{P(x)}$  и  $y = -\sqrt{P(x)}$  «склеиваются» в точках  $x_1, x_2, x_3$  «гладко», т.е. без углов. Это происходит потому, что касательная к графику  $y = \sqrt{P(x)}$  в точках  $x_1, x_2, x_3$  вертикальная; другими словами, ее угловой коэффициент обращается в бесконечность. Это легко доказать подсчетом производной функции  $y = \sqrt{P(x)}$ .

### Сложение точек на эллиптической кривой

Метод секущих, примененный к эллиптической кривой  $C$ , приводит к неожиданному результату: оказывается, точки на ней можно «складывать». Мы определим операцию сложения точек на  $C$ , отправляясь от ее графического изображения (рис.7). Возьмем на  $C$  две точки  $P$  и  $Q$  и проведем через них прямую. Эта прямая имеет третью точку пересечения с кривой  $C$ . Отразим эту точку от оси  $Ox$  и назовем получившуюся точку *суммой*  $P$  и  $Q$  (обозначение  $P + Q$ , рис.7). Не всегда прямая, проходящая через две точки, пересекает кривую  $C$  в третьей – например, вертикальная прямая. Далее мы более подробно рассмотрим эту ситуацию.

Исследуем свойства сложения точек на эллиптической кривой. За образец примем свойства операции сложения чисел. Эта

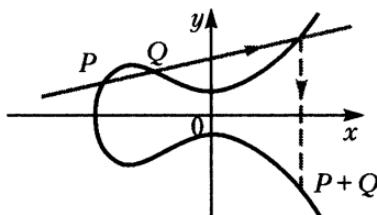


Рис. 7

операция *коммутативна*, т. е.  $a + b = b + a$ , и *ассоциативна*, т. е.  $(a + b) + c = a + (b + c)$ . Кроме того, у этой операции существует *нуль*, т.е. такое число  $O$ , что  $a + O = a$  для любого  $a$ , и, наконец, для каждого числа  $a$  имеется *противоположное* ему, т.е. такое число  $(-a)$ , что  $a + (-a) = 0$ .

А как обстоит дело на эллиптической кривой? Прежде всего, *операция сложения точек коммутативна*. В самом деле, для вычисления  $Q + P$  мы используем ту же самую прямую, что и для  $P + Q$ . Следовательно,  $Q + P = P + Q$ .

*Ассоциативность* для сложения точек на эллиптической кривой также выполняется, но доказать это непросто. Попробуйте осмыслить этот замечательный факт геометрически, сделав чертеж.

Займемся теперь существованием нуля. Нуль – это такая точка  $E$  на кривой, что  $P + E = P$ . Как ее найти? Посмотрим на рисунок 8. Пусть на кривой дана точка  $P$ . Мы хотим найти что-

то такое, что если провести прямую через  $P$  и это «что-то такое», пересечь получившуюся прямую с кривой, а потом отразить точку пересечения от оси  $Ox$ , то вновь получится  $P$ . Обозначим через  $R$  точку, симметричную  $P$  относительно оси  $Ox$ . Из сказанного вытекает, что прямая должна проходить через точки  $P$  и  $R$ ,

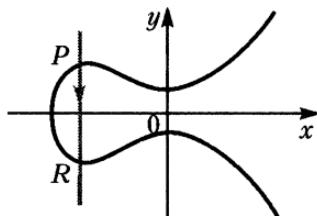


Рис. 8

т.е. должна быть вертикальной. Следовательно, если имеется точка  $E$ , для которой  $P + E = P$ , то эта точка не может находиться в плоскости, поскольку она должна лежать и на кривой и на вертикальной прямой.

Раз точки  $E$  в плоскости нет, а она нам очень нужна, то мы добавим ее к плоскости и назовем *бесконечно удаленной точкой*. Каким требованиям она должна удовлетворять? Любая вертикальная прямая стремится к бесконечности сверху и снизу. Потребуем, чтобы все эти бесконечности были одной и той же точкой  $E$ , т.е. будем считать, что  $E$  есть *точка пересечения всех вертикалей*. Это требование корректно определяет точку  $E$  – нулевую точку относительно нашей операции сложения. В силу нашего соглашения вертикальная прямая, проходящая через точку  $P$ , проходит через  $P$  и  $E$ . Поэтому точка  $R$  пересечения этой прямой с эллиптической кривой удовлетворяет соотношению  $P + R = E$ , т.е. является противоположной к  $P$ . А с другой стороны,  $R$  – это точка, симметричная к  $P$  относительно оси  $Ox$ . Значит, любая точка  $P$  имеет противоположную  $-P = R$ . Тем

самым мы убедились, что сложение точек на эллиптической кривой обладает теми же свойствами, что и сложение чисел.

Как вычислить точку  $P + P$ ? Когда точки были различны, мы проводили секущую. Раз они слились, понятно, что нужно провести касательную (рис.9).

А что делать, чтобы найти  $3P$ ? Очень просто, берем сумму  $2P$  и  $P$ . Подобно этому  $4P = 3P + P$ ,  $5P = 4P + P$  и т.д.

### Поиск рациональных точек

Вооружившись операцией сложения, займемся теперь рациональными точками. Пусть  $P = (x_1, y_1)$ ,  $Q = (x_2, y_2)$  – две рациональные точки на эллиптической кривой  $y^2 = x^3 + ax^2 + bx + c$ , где  $a, b, c$  – целые числа, и прямая, проходящая через  $P$  и  $Q$ , пересекает эту кривую еще в одной точке  $R = (x_3; y_3)$ . Тогда  $R$  также является рациональной точкой.

Доказывается это утверждение довольно просто. Если

$$y = kx + d \quad (7)$$

– уравнение прямой, проходящей через точки  $P$  и  $Q$ , то  $k$  и  $d$  – рациональные числа, поскольку их можно выразить через координаты  $(x_1; y_1)$  и  $(x_2; y_2)$  точек  $P$  и  $Q$  по формулам

$$k = \frac{y_1 - y_2}{x_1 - x_2}, \quad d = y_1 - kx_1 = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}.$$

Подставив (7) в уравнение эллиптической кривой, получим для  $x$  уравнение третьей степени с рациональными коэффициентами

$$(kx + d)^2 = x^3 + ax^2 + bx + c,$$

т.е.

$$x^3 + (a - k^2)x^2 + (b - 2kd)x + c - d^2 = 0.$$

По теореме Виета

$$x_1 + x_2 + x_3 = k^2 - a.$$

Так как  $x_1$  и  $x_2$  рациональны, то рациональным будет  $x_3$ , а значит, и  $y_3 = kx_3 + d$ .

Из этого доказательства сразу же следует формула для вычисления координат точки  $P + Q$ . По определению  $P + Q$

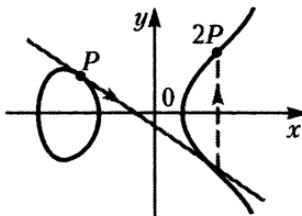


Рис. 9

получается из  $R$  отражением от оси  $Ox$ , значит, координаты  $(u; v)$  точки  $P + Q$  могут быть найдены по формулам

$$u = k^2 - a - x_1 - x_2,$$

$$v = -ku - d = -[k(u - x_1) + y_1].$$

Подставив сюда значения  $k$  и  $d$ , получим

$$u = \frac{(y_1 - y_2)^2}{(x_1 - x_2)^2} - (a_1 + x_1 + x_2), \quad v = \frac{y_1 - y_2}{x_1 - x_2} (x_1 - u) - y_1. \quad (8)$$

Ясно, что если  $x_1 = x_2$ , эти формулы не имеют смысла. В этом случае уравнение секущей (7) нужно заменить уравнением касательной и действовать по прежней схеме. В результате получим

$$\begin{aligned} u &= -2x_1 + a - \left( \frac{3x_1^2 + 2ax_1 + b}{2y_1} \right)^2, \\ v &= y_1 + \frac{3x_1^2 + 2ax_1^2 + b}{2y_1} (u - x_1). \end{aligned} \quad (9)$$

Таким образом, зная хотя бы одну рациональную точку  $P$  на эллиптической кривой, мы можем найти по указанным формулам точки  $2P, 3P$  и т.д. Рассмотрим пример. Пусть кривая задана уравнением  $y^2 = x^3 - 2$  и  $P(3; 5)$ . Тогда  $2P = \left( \frac{129}{100}; -\frac{383}{1000} \right)$  – новая рациональная точка. Теперь можно вычислить  $3P, 4P$  и т.д. Заметим, что объем вычислений с каждым шагом стремительно растет. Если обозначить через  $u_n$  первую координату точки  $nP$ , то

$$u_1 = 3, \quad u_2 = \frac{129}{100}, \quad u_3 = \frac{164323}{29241},$$

$$u_4 = \frac{2340922881}{58675600}, \quad u_5 = \frac{307326105747363}{160280942564521}.$$

Далее еще быстрее. Например, у  $u_{11}$  в числителе 71 знак.

В настоящее время неизвестно никакой общей процедуры для нахождения всех рациональных решений уравнения  $y^2 = x^3 + ax^2 + bx + c$ . В разобранном примере  $y^2 = x^3 - 2$  одно решение  $(3; 5)$  мы просто подобрали; в общем же случае неизвестно никакого метода, который позволил бы найти это первое решение. Нахождение рационального решения эллиптического уравнения с помощью эффективной процедуры является одной

из крупнейших проблем теории чисел. Однако, если одно решение есть, можно найти другие по формулам (8) и (9).

### Порядок точек на эллиптической кривой

При получении точек  $nP$  из данной точки  $P$  возможны два случая. В первом случае на шаге  $n$  получается нуль, т.е. существует такое число  $n$ , что  $nP = E$ . Если для всех  $m < n$   $mP \neq E$ , то говорят, что точка  $P$  имеет *конечный порядок*  $n$ . Например, на кривой  $y^2 = x^3 + 4$  точка  $P = (0; 2)$  имеет порядок 3, на кривой  $y^2 = x^3 + 1$  точка  $P = (2; 3)$  имеет порядок 6, на кривой  $y^2 = x^3 - 43x + 166$  точка  $P = (3; 8)$  имеет порядок 7. Можно поставить вопрос: сколько существует рациональных точек конечного порядка и каковы эти порядки?

В 1976 году американский математик Б.Мазур получил выдающийся результат, показав, что *если  $P$  – рациональная точка порядка  $n$ , то  $n \leq 10$  или  $n = 12$ ; с другой стороны, на эллиптической кривой существует самое большое 16 рациональных точек конечного порядка.*

Второй случай – это когда все точки  $2P, 3P, 4P$  и т.д. различны. В 1901 году французский математик А.Пуанкаре (1854–1912) высказал гипотезу о том, что *всегда можно найти такое конечное число рациональных точек  $P_1, \dots, P_r$ , что всякая рациональная точка  $P$  выражается через них, т.е. представляется в виде*

$$P = n_1 P_1 + \dots + n_r P_r + Q,$$

где  $n_1, \dots, n_r$  – целые числа, однозначно определяемые точкой  $P$ , а  $Q$  – точка конечного порядка. Сами же точки  $P_1, \dots, P_r$  не выражаются друг через друга. Число  $r$  называется *рангом* кривой.

Гипотезу Пуанкаре в 1922 году доказал англичанин Л.Мордэлл, но его доказательство не дает никакого способа для вычисления ранга. До сих пор неясно, существуют ли эллиптические кривые сколь угодно большого ранга. Известно лишь, что ранг оценивается через коэффициенты  $a, b, c$  уравнения  $y^2 = x^3 + ax^2 + bx + c$ , поэтому кривые большого ранга должны иметь большие коэффициенты. Например, к числу кривых ранга  $r \geq 8$  относится недавно найденная кривая, у которой

$$a = -3^2 \cdot 1487 \cdot 1873, \quad b = 2^5 \cdot 3^2 \cdot 5 \cdot 151 \cdot 14551 \cdot 33353,$$

$$c = 2^8 \cdot 3^4 \cdot 5^2 \cdot 7 \cdot 151^2 \cdot 193 \cdot 273 \cdot 156307.$$

## **Заключение: кривые произвольных степеней**

Мы здесь ограничились кривыми (а значит, диофантовыми уравнениями) степеней 2 и 3. А как обстоит дело со степенями  $n \geq 4$ ? В этом случае тоже естественно выделить класс «невырожденных» кривых степени  $n$  (типичный представитель – кривая  $x^n + y^n = 1$ ). При  $n > 3$  картина разительно меняется. Еще в 1931 году английский математик Л. Морделл выдвинул гипотезу – на таких кривых *число рациональных точек всегда конечно*. Гипотеза Морделла более полувека была в центре внимания ведущих математиков всего мира. Существенный вклад в ее решение внесли советские математики И.Р.Шафаревич, Ю.И.Манин, С.Ю.Аракелов, А.Н.Паршин, Ю.Г.Зархин, но честь ее решения в 1983 г. выпала молодому ученному из ФРГ Герду Фальтингсу. За это в 1986 г. на Конгрессе математиков в Беркли (США) он был награжден медалью Филдса – высшей наградой для математиков.

## ОТКРЫТИЕ ВСЕЛЕННОЙ

---

Ю.Соловьев

Из всех картин природы, развертывающихся перед человеческими глазами, самая величественная – картина звездного неба. С древнейших времен эта картина будила человеческое воображение, вызывая к жизни то могучее интеллектуальное течение, которое мы теперь называем наукой. О том, как человек проникал в тайну мироздания, в тайну движения небесных тел – наш рассказ.

Первые воззрения древних на мироздание основывались на непосредственно видимом. Для древних египтян и вавилонян Вселенная отождествлялась с Землей, которая представлялась громадным диском, плавающим по беспредельному океану. Небо они представляли опрокинутой чашей, опирающейся на плоскость, с вкрапленными во внутреннюю поверхность чаши звездами.

Упорядоченные конфигурации звезд называются созвездиями. Созвездия почти не меняются день ото дня, год от года и даже век от века. За ночь звезды поворачиваются относительно некоторой неподвижной точки, находящейся ныне вблизи Полярной звезды, будто чаша вращается как единое целое вокруг оси, проходящей через эту точку и глаз наблюдателя. Тщательные наблюдения показывают, что чаша совершает один оборот за 23 часа 56 минут.

Из-за вращения чаши некоторые звезды уходят за горизонт на западе, тогда как другие восходят на востоке. Это наблюдение наводит на мысль, что чаша представляет собой часть полной сферы, причем звезды располагаются на ее поверхности. Эта сфера называется небесной сферой или сферой неподвижных звезд. Две неподвижные точки, в которых небесная сфера пересекает ось вращения, называются полюсами мира. Воображаемая окружность на небесной сфере, точки которой равноудалены от обоих полюсов, называется небесным экватором.

## Шарообразная Земля

Итак, наблюдения за суточным движением звезд привели древних астрономов к представлению о небесной сфере. Сделать вывод о том, что Земля имеет форму шара, оказалось трудней. Мысль о шарообразности Земли высказали древнегреческие ученые в начале V века до новой эры. К этой мысли они

пришли на основании рассказов путешественников, которые заметили, что высота северного полюса мира возрастает по мере перемещения к северу и уменьшается при движении в южном направлении (рис.1).

Первое настоящее измерение радиуса Земли было проведено египетским греком Эратосфеном (276–195 гг. до н.э.), уроженцем города Сиена (ныне Асуан) на юге Египта. Еще будучи подростком он заметил, что в Сиене ежегодно 21 июня в полдень Солнце находится точно над головой, и вертикальные стволы деревьев не отбрасывают теней. Позже в Александрии, расположенной на севере Египта, он обнаружил, что тени здесь в указанное время

Рис.1. 1 – направление на северный полюс мира; 2 – направление на горизонт в точке В; 3 – направление на горизонт в точке А

не исчезают, и в гениальном озарении понял, что причина этому – кривизна поверхности Земли. Александрия находится на 480 миль севернее Сиены, и когда Солнце в зените над Сиеной, над Александрией оно должно располагаться на некотором угловом расстоянии от зенита. Этот угол можно измерить по тени

вертикального ствола дерева или колонны в Александрии (рис.2). Угол определяется по высоте дерева или колонны и измеренной длине тени в Александрии в момент, когда в Сиене тень не отбрасывается (полдень 21 июня). Углы  $\alpha$  и  $\beta$  равны как внутренние накрест лежащие углы при пересечении параллельных прямых. Измеренное значение составляло  $\alpha = 7^\circ$ , следовательно, угол в  $7^\circ$  с вершиной в центре Земли стяги-

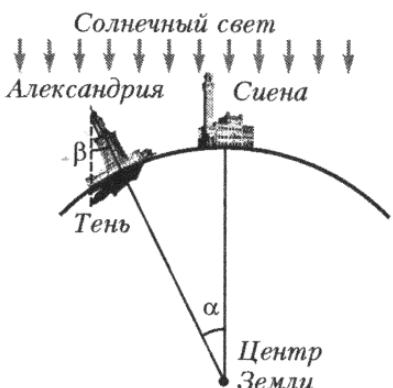


Рис. 2

вался на поверхности дугой в 480 миль. Поскольку в окружности  $360^\circ$ , длина окружности Земли по меридиану должна составлять 24 000 миль, а ее радиус – около 4000 миль (точные современные значения составляют 24 989 миль для длины окружности и 3950 миль для радиуса).

Идея шарообразной Земли позволила упростить геометрию мироздания. Естественно было считать, что земная и небесная сферы концентричны, а ось вращения небесной сферы служит продолжением полярной оси Земли.

### **Блуждающие звезды**

Кроме неподвижных звезд, на небесной сфере можно наблюдать тела, положение которых меняется день ото дня. Такие тела называются планетами. В переводе с древнегреческого планета – «блуждающая звезда». С древнейших времен было известно семь таких «блуждающих звезд»: Луна, Меркурий, Венера, Солнце, Марс, Юпитер и Сатурн.

Чтобы понять, как движется Солнце по небесной сфере, вспомним, что сутки (24 часа) – это время между двумя его последовательными кульминациями. Тот факт, что небесная сфера совершают полный оборот вокруг оси за время чуть меньшее суток (23 часа 56 минут), означает, что Солнце перемещается по небесной сфере в направлении, противоположном вращению небесной сферы. Поэтому восход Солнца ежедневно запаздывает по сравнению с восходом звезд на 4 минуты. Отмечая ежедневное положение Солнца относительно звезд в момент восхода, можно проследить его траекторию на небесной сфере. Оказывается, что эта траектория представляет собой еще одну окружность, центр которой совпадает с центром Земли, а плоскость наклонена под углом  $23^\circ 30'$  к плоскости небесного экватора. По этой окружности, называемой эклиптикой, Солнце движется в направлении с запада на восток с почти постоянной угловой скоростью, равной примерно  $1^\circ$  в сутки, совершая полный оборот за 365 дней и примерно 6 часов.

Луна также непрерывно перемещается относительно звезд. Ее траектория представляет собой окружность, в центре которой находится Земля. Плоскость этой окружности наклонена к эклиптике под углом  $5^\circ$ . Луна почти равномерно движется вдоль своей траектории в том же направлении, что и Солнце, т.е. обратно направлению суточного вращения небесной сферы, совершая полный оборот за время немногим более 27 суток. Подобно движению Солнца, движение Луны приводит к тому, что ее восход ежедневно запаздывает относительно восхода

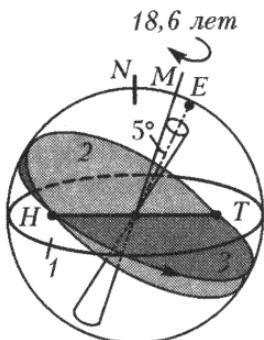
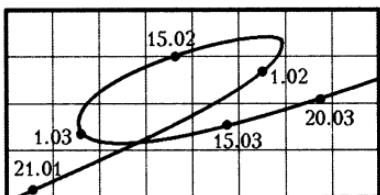


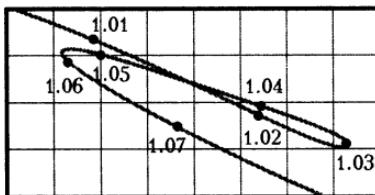
Рис.3. 1 – экватор; 2 – эклиптика; 3 – путь Луны за 27 суток. Траектория Луны представляет собой большой круг на небесной сфере, наклоненный под углом  $5^{\circ}$  к эклиптике. Полярная ось  $OM$ , перпендикулярная к лунной орбите, описывает конус вокруг оси  $OE$ , перпендикулярной к плоскости эклиптики, за 18,6 лет

звезд, но уже не на четыре минуты, как у Солнца, а почти на целый час (рис.3).

Пять остальных «блуждающих звезд» также движутся по небесной сфере, однако их перемещение гораздо сложнее, чем движение Солнца и Луны (рис.4). Древние астрономы разделили эти пять планет по своим видимым движениям на две группы: нижний (Меркурий, Венера) и верхние (Марс, Юпитер и



а) Видимое движение Меркурия



б) Видимое движение Марса

Рис. 4

Сатурн). Движения нижних и верхних планет по сфере неподвижных звезд существенно различаются.

Меркурий и Венера находятся на небе, не удаляясь слишком далеко от Солнца ( $28^{\circ}$  для Меркурия и  $47^{\circ}$  для Венеры). Наибольшее угловое удаление планеты от Солнца к востоку называется ее наибольшей восточной элонгацией, к западу – наибольшей западной элонгацией. При наибольшей восточной элонгации планета видна на западе в лучах вечерней зари вскоре после захода Солнца и заходит через некоторое время после него. Затем, попутным движением (т.е. с востока на запад), сначала медленно, а потом все быстрее, планета начинает приближаться к Солнцу, скрывается в его лучах и перестает быть видимой. В это время наступает нижнее соединение планеты с Солнцем. Спустя некоторое время после нижнего соединения планета вновь становится видимой, но теперь уже на востоке, незадолго перед восходом Солнца. В это время, продолжая

попятное движение, она постепенно удаляется от Солнца. Замедлив скорость попятного движения и достигнув наибольшей западной элонгации, планета останавливается и меняет направление своего движения на прямое (с запада на восток). Вначале она движется медленно, затем ее движение становится более быстрым. Удаление ее от Солнца уменьшается и, наконец, она скрывается в утренних лучах Солнца – наступает ее верхнее соединение с Солнцем. Спустя некоторое время она снова видна на западе в лучах вечерней зари. Продолжая прямое движение, планета замедляет свою скорость. Достигнув предельного восточного удаления, т. е. наибольшей восточной элонгации, планета останавливается, меняет направление своего движения на попятное, и все повторяется сначала. Период одного такого «колебания» составляет для Меркурия 88 суток, а для Венеры – 225 суток.

Видимые движения верхних планет происходят иначе. Когда верхняя планета видна после захода Солнца в западной части неба, она перемещается среди звезд прямым движением, т. е. с запада на восток, как и Солнце. Но скорость ее движения меньше, чем у Солнца, которое постепенно нагоняет планету, и она на некоторое время перестает быть видимой, т. е. восходит и заходит почти одновременно с Солнцем. Затем, когда Солнце обгонит планету, она становится видимой на востоке перед восходом Солнца. Скорость ее прямого движения постепенно уменьшается, планета останавливается и затем начинает перемещаться среди звезд попятным движением, с востока на запад. Через некоторое время планета снова останавливается, меняет направление своего движения на прямое, снова ее с запада нагоняет Солнце и она опять перестает быть видимой – и все явления повторяются в том же порядке.

### **Первые модели мира**

**Эвдокс.** Первоначальная модель мироздания была очень проста. Длительные наблюдения привели древнегреческих ученых к убеждению, что наряду с Землей шарообразными являются и другие планеты. Более того, с течением времени для двух небесных светил, которые казались самыми большими – Солнца и Луны – было получено так много данных, что их стали считать телами, до известной степени родственными Земле. Не было никаких причин считать другие «блуждающие звезды» непохожими на Солнце и Луну. Стало быть, все они более-менее подобны Земле, а различия в их наблюдаемых размерах объясняются различным их удалением от Земли.

Но ведь эти громадные тела, совершающие вечный полет над нашими головами, должны быть достаточно прочно укреплены. Небесная сфера для этого не подходила: планеты движутся независимо от нее. Поэтому греки вообразили семь новых сфер, по одной для каждой планеты. Эти сферы меньше, чем сфера неподвижных звезд, и концентричны с ней. Все семь планетных сфер участвуют в движении сферы неподвижных звезд, обращаясь раз в сутки и, кроме того, совершают еще и независимые вращения.

В дальнейшем из этой модели Пифагор развил идею о музыке сфер: каждая из семи планетных сфер сравнивалась с одним из тонов октавы; восьмой тон представляла сфера неподвижных звезд. Пифагор считал, что колоссальные полые шары, к которым прикреплены такие большие тела, как Солнце и Луна (да и другие планеты), при вращении должны издавать звук, как вращающиеся колеса какого-нибудь механического аппарата. Получающиеся при этом различные тона соединяются в чудесную мелодию, могущие звуки которой наполняют всю Вселенную. И только мы, несовершенные жители Земли, не можем слышать этой небесной музыки, составляющей вечное наслаждение олимпийских богов.

После того как были приобретены дальнейшие знания относительно движения небесных светил, потребовалось соответствующим образом пополнить идею о сферах, составляющих основу порядка Вселенной. В мировоззрении древних философов Греции считалось непоколебимым фактом, что Земля есть центр мира, главное тело мироздания. Поэтому все многочисленные усложнения, которые обнаружили древние ученые, — неравномерные движения планет, попятные движения и т.д., — требовалось объяснить только введением новых сфер, которые общим движением влияли бы на одно и то же небесное светило. Знаменитый древнегреческий ученый Эвдокс (408–355 гг. до н.э.) построил систему мира, состоящую из 27 сфер: по три для Солнца и Луны, по четыре для пяти остальных планет и большой сферы неподвижных звезд. Однако вскоре выяснилось, что и 27 сфер недостаточно для описания видимого движения планет. Поэтому вскоре последователь Эвдокса афинянин Калипп присоединил к эвдоксовым еще 22 сферы. С течением времени небесная машина становилась все сложнее и сложнее, так что в конце концов настойчиво выступила необходимость более простого и ясного описания.

**Гиппарх.** В основании всех древних воззрений на мир лежал принцип равномерного движения по окружностям. Впер-

вые этот принцип был поколебленalexандрийским ученым Гиппархом (II век до н.э.), открывшим, что длины времен года неодинаковы. Гиппарх первым нашел перигей и апогей Солнца и установил, что вблизи первого оно движется быстрее, чем вблизи второго. Но аксиома о равномерном движении слишком глубоко вошла в плоть и кровь античной науки, и Гиппарх не решился уничтожить ее.

Чтобы объяснить свои открытия, Гиппарх избрал иной выход. Допустим, что Солнце движется с постоянной скоростью по окружности, но центр этой окружности не совпадает с центром Земли, а лежит вне его, где-нибудь в свободном пространстве. Тогда, действительно, нам должно казаться, что Солнце движется неравномерно – быстрее, когда оно идет по части круга, к которой Земля расположена ближе, и медленнее – в противоположной части. Рисунок 5 поясняет этот механизм. Центр движения Солнца находится здесь в точке пересечения пунктирных линий, сплошные линии сходятся в центре Земли. Путем проб можно найти подходящее место для точки, находясь в которой наблюдатель будет видеть упомянутые выше особенности движения Солнца, хотя на самом деле оно совершается равномерно по кругу. Линию, соединяющую перигей и апогей, Гиппарх назвал линией апсид. Отношение измеренного по ней расстояния от центра солнечной орбиты до центра Земли к радиусу орбиты было названо им эксцентриситетом орбиты. Эти названия удержались в астрономическом языке до настоящего времени.

По аналогии с солнечной, центр лунной орбиты Гиппарх также поместил вне Земли. Он вычислил для нее направление линии апсид, эксцентриситет, перигей и апогей. Движения Солнца и Луны Гиппарх определил с удивительной для его времени точностью. Например, по данным, полученным Гиппархом, можно было бы с точностью до одного дня вычислить даты наступления полнолуний даже для нашего времени (более чем на 2000 лет вперед!). Великий alexандрийский астроном начал также исследовать движения других планет, что представляло более значительные трудности. Но огромный шаг

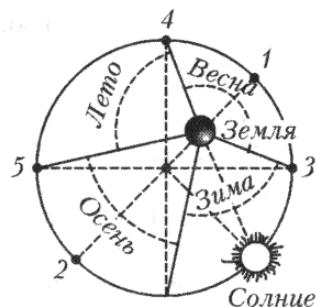


Рис.5. 1 – перигей; 2 – апогей; 3 – весеннее равноденствие; 4 – летнее солнцестояние; 5 – осеннее равноденствие

вперед в этом направлении удалось сделать только его земляку и последователю Клавдию Птолемею (II век новой эры).

## Система мира Птолемея

Птолемеева система мира, которая царила, не оспариваемая никем, около полутора тысяч лет, была основана на наблюдениях и вычислениях Гиппарха. Птолемей изложил свою систему мироздания в знаменитом сочинении «Мегале синтаксис» («Великое построение»), более известном под арабским названием «Альмагест». До конца средневековья этот труд почитался почти наравне с божественным откровением. Сомнение в словах «Альмагеста» считалось преступлением.

В основе системы мира Птолемея лежат четыре постулата:

- I. Земля находится в центре Вселенной.
- II. Земля неподвижна.
- III. Все небесные тела движутся вокруг Земли.
- IV. Движение небесных тел происходит по окружностям с постоянной скоростью, т.е. равномерно.

Птолемей положил в основу своей системы эксцентрические круги Гиппарха. Однако, согласно Птолемею, все планеты, за исключением Солнца, не обращаются непосредственно по этим кругам – по ним движется центр другого круга, и только по нему обращается планета (рис.6). Круги, по которым обращаются планеты, называются эпициклами; круги, по которым движутся центры эпициклов – деферентами. Деферент Солнца, деференты и эпициклы других планет лежат внутри сферы, на поверхности которой расположены «неподвижные звезды».

Суточное движение всех светил объяснялось вращением всей Вселенной как единого целого вокруг неподвижной Земли.



Рис.6. Система мира Птолемея

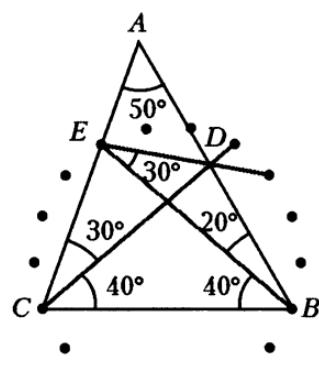


Рис. 7

Прямые и попятные движения планет объяснялись следующим образом.

Пусть в определенный момент времени планета находится на эпицикле в точке  $P_1$  (рис.7), а центр эпицикла – на деференте в точке  $N_1$ . В процессе равномерного кругового движения обеих точек – планеты с угловой скоростью  $\alpha$  вокруг точки  $N_1$  и самой точки  $N_1$  как центра эпицикла с угловой скоростью  $\omega$  вокруг Земли – планета описывает петлю, которую наблюдатель видит в проекции на небесную сферу. Причина образования петли очевидна: в точке  $P_1$  движения по эпициклу и по деференту направлены в одну сторону – справа налево. Описав дугу в  $180^\circ$ , планета движется по эпициклу слева направо. При определенном соотношении между  $\alpha$  и  $\omega$  направление видимого движения в положении, близком  $P_2$ , изменяется – планета здесь совершают попятное движение.

Для каждой планеты Птолемей подобрал относительные размеры радиусов эпицикла и деферента, положения центров деферентов и скорости движения планет по эпициклам и деферентам так, чтобы движение получалось близким к реально наблюдаемому. Это оказалось возможным при выполнении некоторых условий, которые Птолемей принял в качестве постулатов. Они сводились к следующему:

- 1) центры эпициклов нижних планет лежат на направлении от Земли к Солнцу;
- 2) у всех верхних планет этому направлению параллельны радиусы эпициклов, проведенные в точку положения планеты.

Таким образом, направление на Солнце в системе мира Птолемея оказывалось преимущественным.

Система мира Птолемея не только давала качественное объяснение видимым движениям планет, но и позволяла вычислить их положение на будущее с довольно высокой точностью. По мере повышения точности наблюдений возникали разногласия их с теорией, которые устраивались путем усложнения системы. Например, некоторые «неправильности» в видимых движениях планет, открытые позднейшими наблюдениями, объяснялись тем, что вокруг центра первого эпицикла обращается не планета, а центр второго эпицикла, по окружности которого уже движется планета. Когда и такое построение для какой-либо планеты оказывалось недостаточным, вводили третий, четвертый и т. д. эпициклы, пока положение планеты на окружности последнего из них не давало более или менее сносного согласия с наблюдением. К началу XVI столетия система Птолемея содержала в общей сложности 40 кругов.

Возвратимся еще раз к «Альмагесту» Птолемея и в небольшой таблице приведем числа, которые дал этот великийalexандрийский ученый для планетных движений по эпициклам и для движений самих эпициклов по деферентам:

Планеты	Суточное движение по эпициклам	Суточное движение центра эпицикла по деференту	Суммарное движение
Солнце	0°0'00,0"	0°59'8,3"	0°59'8,3"
Меркурий	3°6'21,4"	0°59'8,3"	4°5'32,4"
Венера	0°36'59,4"	0°59'8,3"	1°36'7,7"
Марс	0°27'41,7"	0°31'26,6"	0°59'8,3"
Юпитер	0°54'9,0"	0°4'58,3"	0°59'8,3"
Сатурн	0°57'7,7"	0°2'0,6"	0°59'8,3"

В этой таблице бросается в глаза тот поразительный факт, что центр эпицикла у нижних планет (Меркурия и Венеры) движется с той же скоростью, что и Солнце вокруг Земли. Для верхних планет – Марса, Юпитера и Сатурна – эти числа различаются, но суммы обоих движений дают ту самую величину движения Солнца. Следовательно, движение Солнца заключается во всех планетных движениях. Конечно, такое явление должно было казаться чрезвычайно странным. Сам собою возникает вопрос, не лежит ли в основе этих одинаковых числовых величин какая-нибудь общая причина?

Нет сомнения, что многие мыслители древности и средневековья ставили такой вопрос. Например, античный астроном Аристарх утверждал центральное положение Солнца в системе мира. Однако первым человеком, который дерзнул строго математически разработать идею о том, что все планеты обращаются вокруг Солнца, был гениальный польский астроном Николай Коперник (1473–1543).

### Система мира Коперника

Книга Коперника «Об обращениях небесных сфер» – труд всей его жизни – была опубликована в 1543 году незадолго до смерти ученого. В этом сочинении Коперник разработал идею о движении Земли и положил начало новой астрономии. Созданная им система мира называется гелиоцентрической. В ее основе лежат следующие утверждения:

- 1) в центре мира находится Солнце, а не Земля;
- 2) шарообразная Земля вращается вокруг своей оси, и это вращение объясняет кажущееся суточное движение всех светил;

3) Земля, как и все другие планеты, обращается вокруг Солнца, и это обращение объясняет видимое движение Солнца среди звезд;

4) все движения представляются в виде комбинации равномерных круговых движений;

5) видимые прямые и попятные движения планет объясняются движением Земли.

Кроме того, Коперник считал, что Луна движется вокруг Земли и как спутник вместе с Землей – вокруг Солнца.

Постулат о равномерном движении по окружностям заставил Коперника подобно Птолемею ввести эпициклы и сместить центры окружностей-деферентов относительно центра Солнца. В итоге модель Коперника выглядела не менее сложно, чем модель Птолемея: достаточно сказать, что она содержала 48 кругов вместо 40 кругов в геоцентрической модели. Не отличалась она и большей точностью. Однако в ней было зерно научной истины, сделавшее ее фундаментом новой астрономии.

По иронии судьбы, подтвердить вывод Коперника о том, что Земля вращается вокруг Солнца, выпало на долю датского ученого Тихо Браге (1546–1601), который имел весьма веские основания не принимать гелиоцентрическую систему. Главный довод Тихо Браге против системы Коперника заключался в следующем: если бы Земля вращалась вокруг Солнца, то и Венера, и Меркурий имели бы фазы, как Луна, чего никто из серьезных астрономов никогда не наблюдал. Доводы Браге звучали убедительно, и хотя предсказываемые фазы, как теперь хорошо известно, действительно существуют, несовершенство оптических инструментов не позволяло их обнаружить. И все же именно точные наблюдения Браге за положением планет в конце концов подтвердили точку зрения Коперника. Данные, полученные Браге, позволили его ученику Иоганну Кеплеру (в итоге восьмилетней работы) объявить, что орбиты планет представляют собой эллипсы, в одном из фокусов которых находится Солнце, и что в каждую единицу времени радиус-вектор планеты заметает одинаковую площадь. Так рухнула Пифагорова гармония круговых орбит вокруг богом данного особого положения нашей планеты. Законы Кеплера в свою очередь (в большей степени, чем легендарное падающее яблоко) явились фундаментом, на котором покоятся закон всемирного тяготения Ньютона, ставший почти на три столетия основой физики и космологии.

## СОДЕРЖАНИЕ

---

От составителей	3
Числа Пизо. <i>A.Егоров</i>	5
Математика бильярда. <i>A.Земляков</i>	30
Кватернионы. <i>А.Мищенко, Ю.Соловьев</i>	44
Демократия с точки зрения математики. <i>В.Пахомов</i>	54
Диагонали правильного 18-угольника. <i>В.Прасолов</i>	71
Умеете ли вы считать? <i>В.Рождественский</i>	77
Покрытия полосками. <i>M.Смурров, A.Спивак</i>	86
Геометрия скользящих векторов. <i>Ю.Соловьев, A.Сосинский</i>	114
Вызов Ван Роумена. <i>Ю.Соловьев</i>	127
Арифметика эллиптических кривых. <i>Ю.Соловьев</i>	132
Открытие вселенной. <i>Ю.Соловьев</i>	143

**КОЛМОГОРОВСКОЙ ШКОЛЕ – ПЯТЬДЕСЯТ**  
**Сборник статей**

**Часть 2**

*Составители В.В.Вавилов, В.Н.Дубровский, А.А.Егоров*

*Библиотечка «Квант». Выпуск 132*

*Приложение к журналу «Квант» №4 / 2014*

*Редактор А.Ю.Котова*

*Обложка А.Е.Пацхверия*

*Макет и компьютерная верстка Е.В.Морозова*

*Компьютерная группа М.Н.Грицук, Е.А.Митченко*

*Формат 84×108 1/32. Бум. офсетная. Гарнитура кудряшевская*

*Печать офсетная. Объем 4,5 печ.л. Тираж: 1-й завод 900 экз.*

*Заказ № 12491*

**119296 Москва, Ленинский пр., 64-А, «Квант»**

**Тел.: (495)930-56-48, e-mail: math@kvant.ras.ru, phys@kvant.ras.ru**

**Отпечатано «ТДДС-СТОЛИЦА-8»**

**Тел.: 8(495)363-48-86, <http://capitalpress.ru>**

## ВЫШЛИ ИЗ ПЕЧАТИ КНИГИ СЕРИИ «БИБЛИОТЕЧКА «КВАНТ»

---

1. *М.П.Бронштейн.* Атомы и электроны
2. *М.Фарадей.* История свечи
3. *О.Оре.* Приглашение в теорию чисел
4. Опыты в домашней лаборатории
5. *И.Ш.Слободецкий, Л.Г.Асламазов.* Задачи по физике
6. *Л.П.Мочалов.* Головоломки
7. *П.С.Александров.* Введение в теорию групп
8. *В.Г.Штейнгауз.* Математический калейдоскоп
9. Замечательные ученые
10. *В.М.Глушков, В.Я.Валах.* Что такое ОГАС?
11. *Г.И.Копылов.* Всего лишь кинематика
12. *Я.А.Смородинский.* Температура
13. *А.Е.Карпов, Е.Я.Гик.* Шахматный калейдоскоп
14. *С.Г.Гиндикян.* Рассказы о физиках и математиках
15. *А.А.Боровой.* Как регистрируют частицы
16. *М.И.Каганов, В.М.Цукерник.* Природа магнетизма
17. *И.Ф.Шарыгин.* Задачи по геометрии: планиметрия
18. *Л.В.Тарасов, А.Н.Тарасова.* Беседы о преломлении света
19. *А.Л.Эфрос.* Физика и геометрия беспорядка
20. *С.А.Пикин, Л.М.Блинов.* Жидкие кристаллы
21. *В.Г.Болтянский, В.А.Ефремович.* Наглядная топология
22. *М.И.Башмаков, Б.М.Беккер, В.М.Гольховой.* Задачи по математике: алгебра и анализ
23. *А.Н.Колмогоров, И.Г.Журбенко, А.В.Прохоров.* Введение в теорию вероятностей
24. *Е.Я.Гик.* Шахматы и математика
25. *М.Д.Франк-Каменецкий.* Самая главная молекула
26. *В.С.Эдельман.* Вблизи абсолютного нуля
27. *С.Р.Филонович.* Самая большая скорость
28. *Б.С.Бокштейн.* Атомы блуждают по кристаллу
29. *А.В.Бялко.* Наша планета – Земля
30. *М.Н.Аршинов, Л.Е.Садовский.* Коды и математика
31. *И.Ф.Шарыгин.* Задачи по геометрии: стереометрия
32. *В.А.Займовский, Т.Л.Колупаева.* Необычные свойства обычных металлов
33. *М.Е.Левинштейн, Г.С.Симин.* Знакомство с полупроводниками

34. *В.Н.Дубровский, Я.А.Смородинский, Е.Л.Сурков.* Релятивистский мир
35. *А.А.Михайлов.* Земля и ее вращение
36. *А.П.Пурмаль, Е.М.Слободецкая, С.О.Травин.* Как превращаются вещества
37. *Г.С.Воронов.* Штурм термоядерной крепости
38. *А.Д.Чернин.* Звезды и физика
39. *В.Б.Брагинский, А.Г.Полнарев.* Удивительная гравитация
40. *С.С.Хилькевич.* Физика вокруг нас
41. *Г.А.Звенигородский.* Первые уроки программирования
42. *Л.В.Тарасов.* Лазеры: действительность и надежды
43. *О.Ф.Кабардин, В.А.Орлов.* Международные физические олимпиады школьников
44. *Л.Е.Садовский, А.Л.Садовский.* Математика и спорт
45. *Л.Б.Окунь.  $\alpha$ ,  $\beta$ ,  $\gamma$  ...  $Z$ :* элементарное введение в физику элементарных частиц
46. *Я.Е.Гегузин.* Пузыри
47. *Л.С.Марочник.* Свидание с кометой
48. *А.Т.Филиппов.* Многоликий солитон
49. *К.Ю.Богданов.* Физик в гостях у биолога
50. Занимательно о физике и математике
51. *Х.Рачлис.* Физика в ванне
52. *В.М.Липунов.* В мире двойных звезд
53. *И.К.Кикоин.* Рассказы о физике и физиках
54. *Л.С.Понtryгин.* Обобщения чисел
55. *И.Д.Данилов.* Секреты программируемого микрокалькулятора
56. *В.М.Тихомиров.* Рассказы о максимумах и минимумах
57. *А.А.Силин.* Трение и мы
58. *Л.А.Ашкинази.* Вакуум для науки и техники
59. *А.Д.Чернин.* Физика времени
60. Задачи московских физических олимпиад
61. *М.Б.Балк, В.Г.Болтянский.* Геометрия масс
62. *Р.Фейнман.* Характер физических законов
63. *Л.Г.Асламазов, А.А.Варламов.* Удивительная физика
64. *А.Н.Колмогоров.* Математика – наука и профессия
65. *М.Е.Левинштейн, Г.С.Симин.* Барьера: от кристалла до интегральной схемы
66. *Р.Фейнман.* КЭД – странная теория света и вещества
67. *Я.Б.Зельдович, М.Ю.Хлопов.* Драма идей в познании природы
68. *И.Д.Новиков.* Как взорвалась Вселенная
69. *М.Б.Беркинблит, Е.Г.Глаголева.* Электричество в живых организмах
70. *А.Л.Стасенко.* Физика полета

71. *А.С.Штейнберг*. Репортаж из мира сплавов
72. *В.Р.Полищук*. Как исследуют вещества
73. *Л.Кэрролл*. Логическая игра
74. *А.Ю.Гросберг, А.Р.Хохлов*. Физика в мире полимеров
75. *А.Б.Мигдал*. Квантовая физика для больших и маленьких
76. *В.С.Гетман*. Внуки Солнца
77. *Г.А.Гальперин, А.Н.Земляков*. Математические бильярды
78. *В.Е.Белонучкин*. Кеплер, Ньютон и все-все-все...
79. *С.Р.Филонович*. Судьба классического закона
80. *М.П.Бронштейн*. Солнечное вещество
81. *А.И.Буздин, А.Р.Зильberman, С.С.Кротов*. Раз задача, два задача...
82. *Я.И.Перельман*. Знаете ли вы физику?
83. *Р.Хонсбергер*. Математические изюминки
84. *Ю.Р.Носов*. Дебют оптоэлектроники
85. *Г.Гамов*. Приключения мистера Томпкинса
86. *И.Ш.Слободецкий, Л.Г.Асламазов*. Задачи по физике (2-е изд.)
87. Физика и...
88. *А.В.Спивак*. Математический праздник
89. *Л.Г.Асламазов, И.Ш.Слободецкий*. Задачи и не только по физике
90. *П.Гнэдиг, Д.Хоньек, К.Райли*. Двести интригующих физических задач
91. *А.Л.Стасенко*. Физические основы полета
92. Задачник «Кванта». Математика. Часть 1
93. Математические турниры имени А.П.Савина
94. *В.И.Белотелов, А.К.Звездин*. Фотонные кристаллы и другие метаматериалы
95. Задачник «Кванта». Математика. Часть 2
96. Олимпиады «Интеллектуальный марафон». Физика
97. *А.А.Егоров, Ж.М.Работин*. Олимпиады «Интеллектуальный марафон». Математика
98. *К.Ю.Богданов*. Прогулки с физикой
99. *П.В.Блиох*. Радиоволны на земле и в космосе
100. *Н.Б.Васильев, А.П.Савин, А.А.Егоров*. Избранные олимпиадные задачи. Математика
101. У истоков моей судьбы...
102. *А.В.Спивак*. Арифметика
103. *Я.А.Смородинский*. Температура (3-е изд.)
104. *А.Н.Васильев*. История науки в коллекции монет
105. *И.Ф.Акулич*. Королевские прогулки
106. Исаак Константинович Кикоин в жизни и в «Кванте»
107. *Г.С.Голицын*. Макро- и микромир и гармония

108. *П.С.Александров*. Введение в теорию групп (2-е изд.)
109. *А.В.Спивак*. Арифметика-2
110. *П.Г.Крюков*. Лазер – новый источник света
111. *А.Б.Сосинский*. Узлы. Хронология одной математической теории
112. *А.П.Пятаков, П.П.Григал*. Лаборатория на коленке
113. *А.А.Заславский*. Олимпиады имени И.Ф.Шарыгина
114. *С.В.Коновалихин*. Сборник качественных задач по физике
115. *Е.Я.Гик*. Математика и шахматы
116. *Л.К.Белопухов*. Физика внезапного
117. *Н.Б.Васильев, А.А.Егоров*. Задачи всесоюзных математических олимпиад. Часть 1
118. Задачник «Кванта». Физика. Часть 1
119. *Н.Б.Васильев, А.А.Егоров*. Задачи всесоюзных математических олимпиад. Часть 2
120. Задачник «Кванта». Физика. Часть 2
121. *Н.Б.Васильев, В.Л.Гутенмакер, Ж.М.Раббот, А.Л.Тоом*. Заочные математические олимпиады
122. *А.З.Долгинов*. Строение материи: от атомов до Вселенной
123. Задачник «Кванта». Физика. Часть 3
124. *А.Толпыго*. 130 нестандартных задач
125. *Н.Б.Васильев*. Статьи из журнала «Квант». Часть 1
126. *Н.Б.Васильев*. Статьи из журнала «Квант». Часть 2
127. *Г.Е.Горелик*. Новые слова науки – от маятника Галилея до квантовой гравитации
128. *Е.Я.Гик*. Компьютерные шахматы
129. *М.И.Каганов*. Физика глазами физика. Часть 1
130. *М.И.Каганов*. Физика глазами физика. Часть 2
131. Колмогоровской школе – пятьдесят. Сборник статей. Часть 1

Индекс 90964



# Библиотечка КВАНТ

Колмогоровской школе —  
пятьдесят

СБОРНИК СТАТЕЙ  
ЧАСТЬ 2

ВЫПУСК

132