# МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ

Третья серия

выпуск 17

Москва
Издательство МЦНМО
2013

УДК 51.009
ББК 22.1
М34



Адрес редакции:
119002, Москва, Б. Власьевский пер., д. 11, к. 301
(с пометкой «Математическое просвещение»)
Email: matpros@mccme.ru  Web-page: www.mccme.ru/free-books

М34 **Математическое просвещение.** Третья серия, вып. 17. —
М.: МЦНМО, 2013. — 208 с.
ISBN 978-5-4439-0080-3

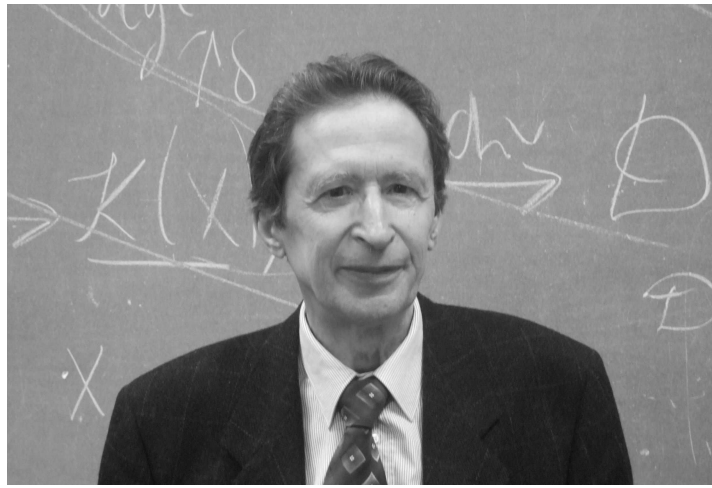
В сборниках серии «Математическое просвещение» публикуются материалы о проблемах современной математики, изложенные на доступном для широкой аудитории уровне, заметки по истории математики, обсуждаются проблемы математического образования.

УДК 51.009
ББК 22.1

Фотографии на с. 3 сделаны А. Зобниным (верхняя)
и С. Третьяковой (нижняя)

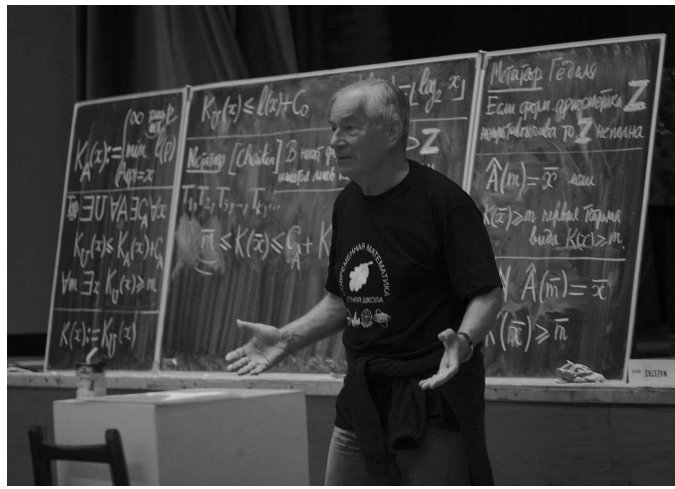
**ISBN 978-5-4439-0080-3**

*Поздравляем*



*Эрнеста Борисовича Винберга*

*и*



*Алексея Брониславовича Сосинского*

*с 75-летием!*

# Содержание

# Алан Тьюринг и теория чисел

Ю. В. Матиясевич

Настоящая публикация является слегка расширенным текстом ряда выступлений автора на конференциях The Alan Turing Centenary Conference (Manchester, UK, June 22–25 2012), The 7th International Computer Science Symposium in Russia (Нижний Новгород, 3 – 7 июля 2012 г.) и на заседании Санкт-Петербургского математического общества (9 октября 2012 г.) и не претендует на полноту освещения теоретико-числовых исследований Алана Тьюринга. Более подробные сведения можно найти, например, в [2, 3, 6, 7].

В 2012 году по всему миру отмечалось столетие со дня рождения Алана Матисона Тьюринга (Alan Mathison Turing, 1912–1954). Это стало большим событием как для учёных разных специальностей, так и для людей, далёких от науки.

Для кого-то Алан Тьюринг в первую очередь — один из основоположников теоретической информатики, по английски называемой computer science. Тезис Чёрча – Тьюринга открыл путь для математически строгих доказательств невозможности решения некоторых алгоритмических проблем, а про машину Тьюринга, по крайней мере слышали и очень многие из тех, кто весьма далёк от точных наук.

Для других Алан Тьюринг — пионер искусственного интеллекта, попытавшийся дать формализованный ответ на философский вопрос «Может ли машина думать?» и написавший первую программу для игры в шахматы в то время, когда ещё не было компьютера, способного выполнить эту программу.

Первоначальная публикация: «Математика в высшем образовании», №10, 2012. С. 111–134. Здесь публикуется с любезного разрешения редакции журнала «Математика в высшем образовании».

Широко известно про вклад Алана Тьюринга в расшифровку немецких радиограмм во время Второй мировой войны.

Оригинальные работы Алана Тьюринга по биологии намного опередили своё время и не были оценены по достоинству его современниками.

Но наряду с привнесением революционных идей в информатику, искусственный интеллект и биологию, Алан Тьюринг внёс существенный вклад и в такой традиционный раздел математики, как теория чисел. К сожалению, даже о самом существовании таких исследований Алана Тьюринга за пределами круга теоретико-числовиков известно немногим. Цель этой публикации — познакомить широкую аудиторию с основным вкладом Алана Тьюринга в теорию чисел.

Все опубликованные им работы по теории чисел связаны с одним, но фундаментальным вопросом этой области математики — распределением простых чисел. По традиции, количество таких чисел, не превосходящих некоторого $x$, обозначается $\pi(x)$. Уже Евклид знал, что эта функция возрастает неограничено, но математиков интересовал вопрос более точной оценки скорости роста $\pi(x)$.

Ответ, который нашли в 1896 году независимо друг от друга Жак Саломон Адамар (Jacques Salomon Hadamard) и Шарль Жан де ла Валле Пуссен (Charles Jean de la Vallée Poussin), известен как

Закон распределения простых чисел.

$$\pi(x) \sim \frac{x}{\ln(x)}\,. \tag{1}$$

Точный смысл символа $\sim$ в (1) таков:

$$\lim_{x\to\infty} \frac{\pi(x)}{\dfrac{x}{\ln(x)}} = 1. \tag{2}$$

Неформально, формулу (1) можно интерпретировать так: *вблизи числа $x$ среднее расстояние между простыми числами равно* $\ln(x)$. Другая интерпретация такова: *вероятность того, что некоторое число, близкое к $x$, будет простым, равна* $1/\ln(x)$.

Эта вторая интерпретация подсказывает рассмотрение следующей функции, называемой *сдвинутым интегральным логарифмом*:

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\ln(t)}\,\mathrm{d}t. \tag{3}$$

Легко проверить, что

$$\mathrm{Li}(x) \sim \frac{x}{\ln(x)} \tag{4}$$

и, соответственно, закон распределения простых чисел (1) можно записать

в эквивалентном виде как

$$\pi(x) \sim \mathrm{Li}(x) \tag{5}$$

или

$$\lim_{x\to\infty} \frac{\pi(x)}{\mathrm{Li}(x)} = 1. \tag{6}$$

При рассмотрении пределов отношений, как в (2) и (6), функция $\mathrm{Li}(x)$ не имеет никаких «преимуществ» перед «более простой» функцией $x/\ln(x)$. Ситуация становится другой, если вместо отношений правых и левых частей в (1) и (6) мы станем изучать их разности, некоторые из которых (округлённые до целых чисел) приведены в следующей таблице.

*Табл. 1. Приближения к $\pi(x)$*

| $x$ | $\pi(x) - x/\ln(x)$ | $\pi(x) - \mathrm{Li}(x)$ |
|---|---|---|
| $10^2$ | 3 | −4 |
| $10^3$ | 23 | −8 |
| $10^4$ | 143 | −16 |
| $10^5$ | 906 | −36 |
| $10^6$ | 611 | −128 |
| $10^7$ | 44158 | −338 |
| $10^8$ | 332773 | −753 |
| $10^9$ | 2592591 | −1699 |
| $10^{10}$ | 20758029 | −3102 |
| $10^{11}$ | 169923159 | −11586 |
| $10^{12}$ | 1416705192 | −38261 |
| $10^{13}$ | 11992858451 | −108970 |
| $10^{14}$ | 102838308635 | −314888 |

В этой таблице разности $\pi(x) - \mathrm{Li}(x)$ по абсолютной величине существенно меньше соответствующих разностей $\pi(x) - x/\ln(x)$. Кроме того, мы видим, что все приведённые там значения $\pi(x) - \mathrm{Li}(x)$ отрицательны. Было проверено, что неравенство $\pi(x) < \mathrm{Li}(x)$ справедливо для всех $x$, не превосходящих $10^{14}$. Возникает вопрос — будет ли это так всегда?

Некоторый неформальный аргумент в пользу этого даёт следующая формула, которую установил Георг Фридрих Бернхард Риман (Georg Friedrich Bernhard Riemann) в своей основополагающей работе [13]:

$$\pi(x) = \mathrm{Li}(x) - \frac{1}{2}\mathrm{Li}(x^{1/2}) + \sum_{\zeta(\rho)=0} \mathrm{Li}(x^{\rho}) + \text{малые члены}. \tag{7}$$

Суммирование в этой формуле идёт по невещественным комплексным нулям так называемой *дзета-функции Римана* $\zeta(s)$, о которой речь пойдёт дальше. Все слагаемые под знаком суммы в (7) осциллируют и можно ожидать, что в среднем они «гасят» друг друга, так что их сумма мала по сравнению с $\frac{1}{2}\mathrm{Li}(x^{1/2})$, абсолютной величиной второго члена в правой части (7). Оказалось, однако, что это не так — при некоторых $x$ многие слагаемые могут «попасть в резонанс», при котором их сумма станет больше $\frac{1}{2}\mathrm{Li}(x^{1/2})$. А именно, Джон Идензор Литлвуд (John Edensor Littlewood) доказал в 1914 году следующую теорему.

Теорема [10]. *Существует бесконечно много $x$ таких, что*

$$\pi(x) > \mathrm{Li}(x). \tag{8}$$

Доказательство, данное Литлвудом, было «чистым доказательством существования», не дающим ни конкретного значения $x$, удовлетворяющего (8), ни даже оценки сверху на величину такого $x$.

Этот «пробел» восполнил ученик Литлвуда Стенли Скьюз (Stanley Skewes), опубликовавший в 1933 году следующую оценку на наименьшее $x$, удовлетворяющее неравенству (8):

$$x < 10^{10^{10^{34}}}. \tag{9}$$

«Астрономическое число» в правой части (9) произвело в то время большое впечатление. Готфри Харолд Харди (Godfrey Harold Hardy) охарактеризовал его как «самое большое число, когда-либо использованное в математике для какой-либо конкретной цели» (см. [22]).

Правая часть неравенства (9) получила название *числа Скьюза*; когда впоследствии этот результат усиливался, новые оценки по-прежнему называли числами Скьюза. Сейчас числом Скьюза иногда называют наименьшее $x$, удовлетворяющее (8). Точное значение такого наименьшего $x$ до сих пор (2012 год) неизвестно, установлено однако, что оно не превосходит $10^{317}$.

В 1931 году Алан Тьюринг поступил в Кембриджский университет в Англии. Скьюз к тому времени уже окончил этот университет, но ещё там работал. Согласно [7], Алан и Стенли, занимаясь греблей, плавали в одной лодке, и весьма вероятно, что именно там, на реке Кем, Тьюринг узнал о числе Скьюза «из первых уст». Тьюринга привлекла идея получить меньшее значение, но как это сделать?

Найденная Скьюзом оценка (9) была условной, а именно, он получил следующий результат.

Теорема [14]. *Если гипотеза Римана верна, то существует число $x$, такое что выполнены неравенства* (8) *и* (9).

Использованная в доказательстве Скьюза гипотеза касается распределения комплексных нулей так называемой *дзета-функции Римана* $\zeta(s)$, определяемой *рядом Дирихле*

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}\,, \tag{10}$$

который сходится в полуплоскости $\mathrm{Re}(s) > 1$. Эту функцию при вещественных значениях аргумента изучал ещё Леонард Эйлер, и он указал её альтернативное определение в виде произведения

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ простое}} \frac{1}{1-p^{-s}}\,, \tag{11}$$

сходящегося при $s > 1$.

*Тождество Эйлера* (11) является, по существу, аналитической формой *основной теоремы арифметики*, гласящей, что каждое натуральное число представимо, и единственным образом, в виде произведения степеней простых чисел: достаточно заметить, что

$$\frac{1}{1-p^{-s}} \;=\; 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \ldots \tag{12}$$

подставить правую часть (12) в правую часть (11), раскрыть скобки и получить левую часть (11).

Тождество Эйлера объясняет, почему дзета-функция Римана играет такую важную роль в изучении простых чисел. В частности, Эйлер дал новое доказательство бесконечности количества простых чисел, по красоте сравнимое с классическим доказательством Евклида: *если бы количество простых чисел было конечно, то* (*расходящийся*) *гармонический ряд, получающийся из левой части* (11) *при* $s \to 1$, *имел бы конечную величину, равную произведению в правой части* (11) *при* $s = 1$.

В основе равенства (7) также лежит тождество Эйлера, но при этом надо рассматривать аналитическое продолжение дзета-функции на всю комплексную плоскость (за исключением точки $s = 1$, которая является полюсом этой функции). Риман доказал, что все отрицательные чётные числа, $-2, -4, \ldots, -2m, \ldots$, являются нулями дзета-функции, и эти нули принято называть *тривиальными*. Он также установил, что других вещественных нулей у дзета-функции нет, а остальные, *нетривиальные*, нули (по которым и ведётся суммирование в (7)) лежат в *критической* полосе $0 \leqslant \mathrm{Re}(s) \leqslant 1$. Эта полоса лежит вне полуплоскости сходимости ряда (10), что вызывает трудности в изучении нетривиальных нулей дзета-функции.

Доказательство закона распределения простых чисел, данное Адамаром и Валле Пуссеном, состояло, по сути, в усилении результата Римана

до строгих неравенств $0 < \mathrm{Re}(s) < 1$. Риман, однако, предполагал, что имеет место ещё более сильное утверждение, известное ныне как

Гипотеза Римана [13]. *Все нетривиальные нули дзета-функции лежат на* критической прямой $\mathrm{Re}(s) = 1/2$.

Эта гипотеза была опубликована Риманом в 1859 году. В 1900 году Давид Гильберт включил её в восьмую из 23 важнейших нерешённых математических проблем, которые уходящий девятнадцатый век оставлял в наследие наступающему двадцатому. Гипотеза Римана до сих пор (2012 год) остаётся недоказанной и неопровергнутой, и является одной из семи так называемых *проблем тысячелетия* [23].

Гипотеза Римана имеет много важных следствий как в самой теории чисел, так и вне её. Например, из гипотезы Римана следует, что

$$\pi(x) - \mathrm{Li}(x) = O(x^{1/2}\log(x)) \tag{13}$$

(и, наоборот, из (13) можно вывести справедливость гипотезы Римана).

То, что оценка (9) была получена Скьюзом в предположении справедливости гипотезы Римана, само по себе не является «грехом» — в современной теории чисел имеется огромное количество результатов, полученных только как следствия гипотезы Римана. Однако в данном случае ситуация была несколько иной — дело в том, что исходная теорема Литлвуда была получена безусловно. Точнее, её доказательство состояло из двух частей: в одной существование $x$, удовлетворяющего (8), устанавливалось в предположении истинности гипотезы Римана, в другой — в предположении её ложности. Скьюз вскоре получил и безусловное доказательство, но по какой-то причине только в 1955 году вышла из печати вторая часть его работы со следующим результатом:

Теорема [15]. *Если гипотеза Римана неверна, то существует число* $x$, *такое что выполнено неравенство* (8) *и*

$$x < 10^{10^{10^{10^3}}}. \tag{14}$$

Когда Алан Тьюринг учился в Кембридже, там одним из преподавателей математики (mathematics supervisors) был Альберт Эдвард Ингам (Albert Edward Ingham). В 1932 году вышло в свет первое издание его ставшей затем классической книги [8] «Распределение простых чисел» (книга была переиздана в 1964 и 1990 годах, её переводы на русский язык вышли в 1936 и в 2005 годах). В этой книге Ингам дал новое, более простое доказательство теоремы Литлвуда, состоящее также из рассмотрения двух случаев — справедливости и ложности гипотезы Римана.

Тьюринг много общался с Ингамом и во время обучения, и позднее по переписке. Когда Тьюринг сообщил Ингаму, что хочет уменьшить число

*Табл. 2. Проверка гипотезы Римана до А.Тьюринга*

| Год | Количество нулей | Автор |
|---|---|---|
| 1903 | 15 | J. P. Gram |
| 1914 | 79 | R. J. Backlund |
| 1925 | 138 | J. I. Hutchinson |
| 1936 | 1041 | E. C. Titchmarsh |

Скьюза, следуя, как и Скьюз, доказательству Литлвуда, Ингам ответил, что более перспективным для этой цели ему представляется его новое доказательство из [9], и на этом пути Скьюз уже улучшил свой результат.

В той части доказательства, в которой предполагается справедливость гипотезы Римана, для получения хорошей оценки требовалось знать не только то, что все нули лежат на критической прямой, но также и где именно лежит достаточно большое количество начальных нулей. Для первых 15 нулей точные значения вещественных частей (равные 1/2) и приблизительные значения мнимых частей были опубликованы Йоргеном Педерсеном Грамом (Jørgen Pedersen Gram) в 1903 году. К тому времени, когда Тьюринг вошёл в эту тематику, Эдвард Чарльз Титчмарш (Edward Charles Titchmarsh) довёл проверку гипотезы Римана до 1041 начального нуля (см. табл. 2).

Теория чисел была серьёзным увлечением Алана Тьюринга начиная от студенческих лет и до последних дней жизни, но, похоже, это увлечение никогда не было главным. В 1936 году Тьюринг опубликовал основополагающую работу [16], в которой ввёл свою знаменитую машину. В наши дни машину Тьюринга традиционно описывают как конечное устройство, работающее конечное время с конечным объёмом информации, но интересно отметить, что Тьюринг рассматривал (и это нашло своё отражение в названии [16]) введённые им вычислительные устройства как средство задания вещественных чисел — такие машины должны работать неограниченно долго, выписывая на ленте всё большее и большее количество десятичных знаков задаваемого вещественного числа. Не исключено, что Тьюринга привёл к этому его интерес к теоретико-числовым проблемам.

В 1938 году Тьюринг написал свою диссертацию (опубликована в [17]) по математической логике под руководством Алонзо Чёрча (Alonzo Church) в Принстоне, США. Однако возвратившись в том же году в Европу, Тьюринг снова стал заниматься дзета-функцией Римана.

Как было сказано выше, для улучшения оценки Скьюза на наименьшее $x$, удовлетворяющее (8), имелось два пути — или найти нуль

дзета-функции, лежащий вне критической прямой, или для возможно большего количества начальных нулей этой функции определить их положение на критической прямой. Однако даже просто вычисление значения дзета-функции в какой-либо точке в критической полосе было нетривиальной задачей.

С необходимостью вычислить значение дзета-функции встретился ещё Эйлер, взявшийся за решение так называемой *Базельской проблемы*, которую поставил Пьетро Меньоли (Pietro Mengoli) в 1644 году — чему равно значение суммы

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\ldots\ ? \tag{15}$$

Эйлер вычислил более дюжины десятичных знаков $\zeta(2)$ с помощью изобретённого им общего метода, который ныне носит название «суммирование Эйлера – Маклорена».

Риман нашёл другой метод, специфический для вычисления значений дзета-функции Римана. Этот метод не был им опубликован и только в 20-м столетии Карл Людвиг Зигель (Carl Ludwig Siegel), разбирая неопубликованные трудночитаемые рукописи Римана, смог сделать открытие Римана всеобщим достоянием.

В 1939 году Тьюринг подал заявку (см. рис. 1) на грант Королевского общества (the Royal Society), играющего в Великобритании ту же роль, что академии наук в других странах. В этой заявке Тьюринг просил средства для изготовления «аппарата» для вычисления значений дзета-функции. В качестве прототипа были взяты машины для расчётов высоты приливов. Такие машины использовались начиная с середины 19-го века и вплоть до появления ЭВМ. На рис. 2 изображена схема одной из таких машин, а на рис. 3 — фотография реальной машины.

Высота прилива в данный момент в данной точке зависит от нескольких периодических факторов: вращения Земли вокруг своей оси, вращения Луны вокруг Земли, вращения Земли вокруг Солнца, при этом для получения достаточной точности необходимо учитывать эксцентриситет орбит Луны и Земли. В итоге высота прилива определяется как сумма вида

$$\sum_{r=1}^{m} a_r \cos(\omega_r t-\theta_r). \tag{16}$$

Значение переменной $t$ — времени — задаётся поворотом рукоятки в левой части машины (см. рис. 2). Коэффициенты $\omega_k$ определяются соотношениями диаметров попарно сцепленных зубчатых колёс, расположенных в нижних рядах. На колёсах в точках с полярными координатами $\langle a_r, \theta_r\rangle$ имеются штырьки, вертикальные (декартовы) координаты которых, очевидно, равняются слагаемым из (16). Несложный механизм

3. A. M. Turing.....£40.

King's College,
Cambridge.
24 *March* 1939.

"1. It is proposed to make calculations of the Riemann zeta-function on the critical line for $1{,}450 < t < 6{,}000$ with a view to discovering whether all the zeros of the function in this range of $t$ lie on the critical line. An investigation for $0 < t < 1{,}464$ has already been made by Titchmarsh. The most laborious part of such calculations consists in the evaluation of certain trigonometrical sums

$$\sum_{r=1}^{m} \frac{1}{\sqrt{r}} \cos(t \log r - \vartheta) \qquad m = \left[\sqrt{\frac{t}{2\pi}}\right]$$

In the present calculation it is intended to evaluate these sums approximately in most cases by the use of apparatus somewhat similar to what is used for tide prediction. When this method does not give sufficient accuracy it will be necessary to revert to the straightforward calculation of the trigonometric sums, but this should be only very rarely necessary. I am hoping that the use of the tide-predicting machine will reduce the amount of such calculation necessary in a ratio of 50 : 1 or better. It will not be feasible to use already existing tide predictors because the frequencies occurring in the tide problem are entirely different from those occurring in the zeta-function problem. I shall be working in collaboration with D. C. MacPhail, a research student who is an engineer. We propose to do most of the machine-shop work ourselves, and are therefore applying only for the cost of materials, and some preliminary computation.

"2. Cost of materials for making tide predictor, estimated at £25, and not exceeding £35. Cost of preliminary computation, estimated at £3 10s., and not exceeding £5. Some further computation may be necessary after the work with the tide predictor, but the amount of this cannot be accurately estimated at this stage, and might be negligible. No application is being made on this account at present.

"4. Fellow of King's College, Cambridge.

"5. Apparatus would be of little permanent value. It could be added to for the purpose of carrying out similar calculations for a wider range of $t$, and might be used for some other investigations connected with the zeta-function. I cannot think of any applications that would not be connected with the zeta-function.

"6. At Cambridge University, principally in the Engineering Laboratories. Professor Hardy and Professor Titchmarsh are, I believe, willing to support this application."

*Рис. 1. Заявка на грант*

http://www.turing.org.uk/sources/zetamachine.html

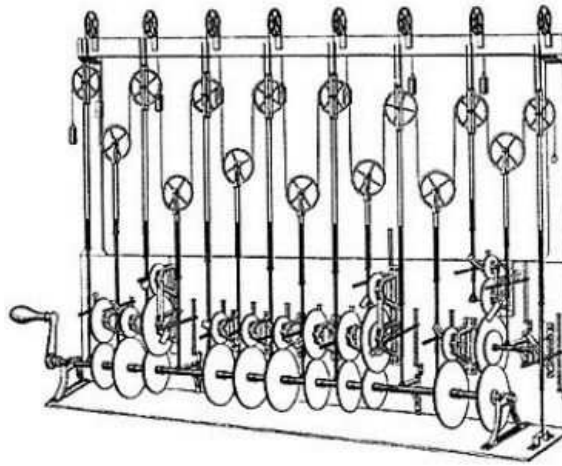


*Рис. 2. Третья машина для предсказания приливов сэра Уильяма Томсона, лорда Кельвина (William Thomson, Lord Kelvin), 1879–81*

`http://en.wikipedia.org/wiki/Tide-predicting_machine`

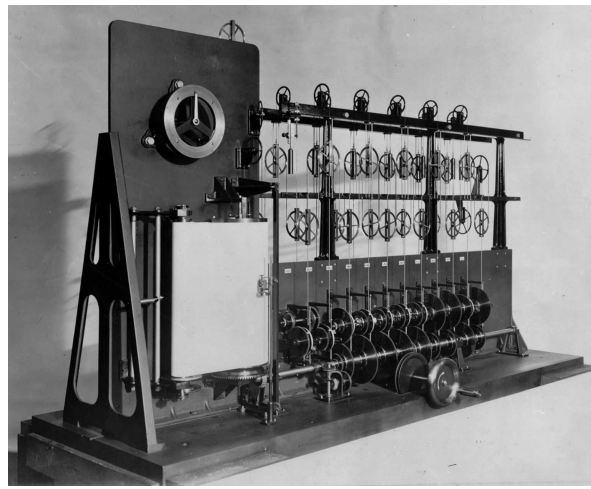


*Рис. 3. Машина для предсказания приливов*

`http://tidesandcurrents.noaa.gov/predma3.html`

позволяет игнорировать горизонтальные перемещения штырьков, а вертикальные преобразовывать в строго вертикальные перемещения гладких (без зубьев) колёс, расположенных над шестерёнками. Для суммирования служит нить, огибающая все гладкие колеса. Правый конец нити жёстко закреплён, а на левом висит перемещающийся груз, положение которого определяется суммой (16) и тем самым предсказывает высоту прилива.

Таким образом, идея конструкции «аппарата для вычисления дзета-функции» не была новой, но оригинальным было применение подобного устройства не для практической цели — вычисления приливов, а для сугубо теоретических целей. В архиве Тьюринга [24] сохранились «синьки», показывающие, что конструкция «аппарата» не была слепым подражанием машине для расчёта приливов.

Тьюринг получил грант (запрошенные 40 фунтов стерлингов) и приступил к работе. Ему помогал студент инженерного факультета Дональд Макфейл (Donald C. MacPhail). Удивительно, что они взялись сделать этот достаточно сложный механизм вдвоём. Для требуемого интервала $1450 < t < 6000$ (см. рис. 1) значение $m$ в (16) равно 30, то есть одних зубчатых колёс требовалось изготовить более 60 штук.

Тьюрингу требовались в качестве $\omega_k$ в (16) логарифмы целых чисел, то есть числа иррациональные, и он использовал цепные дроби для приближения этих логарифмов рациональными числами, числители и знаменатели которых задавали бы количество зубьев на шестерёнках. Некоторое количество этих шестерёнок было изготовлено, но начавшаяся Вторая мировая война прервала работу, и работа над «аппаратом» Тьюринга для вычисления дзета-функции никогда не была завершена.

Понятно, что замена логарифмов натуральных чисел на их рациональные приближения приводила к дополнительной погрешности в вычислении значения дзета-функции (помимо ошибки, связанной с использованием конечной суммы вместо бесконечной), и в заявке на грант Тьюринг указал, что в некоторых случаях потребуется дополнить счёт на «аппарате» традиционным вычислением. Этому была посвящена работа Тьюринга [18], поданная в печать в том же мае 1939 года, что и заявка на грант для механического вычисления дзета-функции.

Техника из [18] стала новым инструментом для вычисления дзета-функции, но впоследствии эта работа Тьюринга была превзойдена. В частности, это сделали Эндрю Михаель Одлызко (Andrew Michael Odlyzko) и Арнольд Шёнхаге (Arnold Schönhage), которые предложили метод, особенно эффективный, когда надо найти значение $\zeta(s)$ не для одного, а для многих близких значений аргумента. Вот что они написали в [12].

> Only one other method, besides the Euler – Maclaurin and Riemann – Siegel ones, seems to have been proposed for computing

> $\zeta(s)$ to moderate accuracy at large heights, namely the one due to Turing. It was designed to provide higher accuracy than was guaranteed by the crude bounds on the remainder term in the Riemann – Siegel formula that were available at that time, and at the same time be more efficient than the Euler – Maclaurin formula. However, very good estimates for remainder terms in the Riemann – Siegel formula are now available, which seem to make Turing's method unnecessary.[1)]

Здесь под «методом Тьюринга» авторы имеют в виду предложенный им способ вычисления значений дзета-функции Римана. То, что в теории чисел традиционно называется «методом Тьюринга» — это средство для проверки того, что все нули дзета-функции с мнимыми частями в заданном интервале удовлетворяют гипотезе Римана.

Этот метод был введён Аланом Тьюрингом в работе [19], опубликованной в 1953 году. Вот как оценил её Эндрю Букер (Andrew R. Booker) в [3]:

> Reading Turing's paper on the subject, which was one of his last, one marvels at what he accomplished with the limited computational resources of the day. His method was truly ahead of its time.[2)]

По форме статья [19] выглядит как отчёт об одном конкретном вычислении, проведённом Тьюрингом в 1950 году на ЭВМ «Mark 1» в Манчестерском университете, но во Введении Тьюринг написал пророческие слова о значимости вводимого им метода:

> This paper is divided into two parts. The first part is devoted to the analysis connected with the problem. All results obtained in this part are likely to be applicable to any further calculations to the same end, whether carried on the Manchester Computer or by any other means.[3)]

---

[1)] Похоже, что только один метод, помимо методов Эйлера – Маклорена и Римана – Зигеля, был предложен для вычисления $\zeta(s)$ с умеренной точностью для больших высот, а именно, метод, принадлежащий Тьюрингу. Он был предназначен давать бóльшую точность, чем та, которую гарантировали грубые оценки остаточного члена в формуле Римана – Зигеля, имевшиеся в то время, и в то же время быть более эффективным, чем формула Эйлера – Маклорена. Теперь, однако, имеются очень хорошие оценки остаточного члена в формуле Римана – Зигеля, которые делают метод Тьюринга излишним.

[2)] Читая работу Тьюринга по этому предмету, которая была одной из его последних, поражаешься тому, что он совершил с ограниченными вычислительными ресурсами тех дней. Его метод поистине опередил своё время.

[3)] Эта статья состоит из двух частей. Первая посвящена анализу рассматриваемой проблемы. Все результаты, полученные в этой части, вероятно, будут применимы ко всем последующим вычислениям, проводимым с той же целью, независимо от того, будут ли они они выполняться на Манчестерском Компьютере или другими средствами.

Интересно отметить, что Тьюринг не верил в справедливость гипотезы Римана:

> The calculations were done in an optimistic hope that a zero would be found off the critical line, and the calculations were directed more towards finding such zeros than proving that none existed.[4)]

При этом Тьюринг надеялся найти нуль $s = \sigma + \mathrm{i}t = \sigma + 2\pi\mathrm{i}\tau$, опровергающий гипотезу Римана, уже при сравнительно малой величине $t$:

> The principal investigation concerned the range $63^2 < \tau < 64^2$ ... The result of this investigation, so far as it can be relied on, was that ... all zeros of $\zeta(s)$ in the region $2\pi 63^2 < t < 2\pi 64^2$ are simple zeros on the critical line.[5)]

Понятно, что наличие нуля вне критической прямой можно установить, вычислив этот нуль с точностью, достаточной для вывода о том, что его вещественная часть не равна $\frac{1}{2}$, но каким образом проведя конечное вычисление с ограниченной точностью можно заключить, что рассматриваемые нули лежат ровно на критической прямой? Тьюринг придавал математической строгости большое значение:

> There is no reason in principle why computation should not be carried through with the rigour usual in mathematical analysis.[6)]

> The procedure was such that if it had been accurately followed, and if the machine made no errors in the period, then one could be sure that there were no zeros off the critical line in the interval in question.[7)]

> Even with the automatic computer this rigour can be rather tiresome to achieve, but in connexion with such a subject as the analytical theory of numbers, where rigour is the essence, it seems worth while.[8)]

---

[4)] Вычисления были проведены в оптимистической надежде найти ноль вне критической прямой и вычисления были направлены скорее на нахождение таких нулей, чем на доказательство их отсутствия.

[5)] Основное исследование относилось к области $63^2 < \tau < 64^2$ ... Результат этого исследования, насколько можно на него полагаться, состоит в том, что ... все нули $\zeta(s)$ в области $2\pi 63^2 < t < 2\pi 64^2$ лежат на критической прямой.

[6)] Нет причины, по которой нельзя было бы выполнить вычисление со строгостью, обычной в математическом анализе.

[7)] Процедура была такова, что если следовать ей аккуратно, и если машина не сделает в это время ошибки, то можно быть уверенным, что в пределах рассматриваемого интервала нет нулей вне критической прямой.

[8)] Даже с автоматическим компьютером достижение такой строгости является весьма утомительным, но в связи с таким предметом, как аналитическая теория чисел, для которой строгость является квинтэссенцией, это представляется заслуживающим усилий.

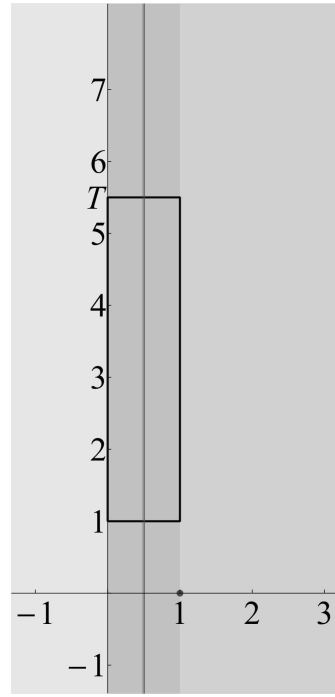



*Рис. 4. Полуплоскость сходимости, критические полоса и прямая, контур интегрирования*

Здесь интересно обратить внимание на словосочетание «автоматический компьютер» — чуть ранее Тьюринг написал следующее.

> The computer will probably have his own ideas as to how certain steps should be done. When certain steps may be omitted without serious loss of accuracy he will wish to do so.[9)]

Понятно, что под просто «компьютером» Тьюринг имеет в виду человека-вычислителя.

Предшественники Тьюринга в вычислении нулей дзета-функции Римана, проводя свои вычисления вручную, также получали математически строгое доказательство того, что найденные ими нули удовлетворяют гипотезе Римана. Грам использовал для этого *теорему Руше*, а последующие исследователи — *теорему Коши*.

[9)] Компьютер, вероятно, будет иметь свои собственные идеи по поводу того, каким образом следует выполнять определённые шаги. Когда некоторые шаги могут быть опущены без серьёзной потери точности, он пожелает это сделать.

Пусть $N(T)$ обозначает количество нулей дзета-функции, лежащих в прямоугольнике

$$0 \leqslant \mathrm{Re}(s) \leqslant 1,\ 1 \leqslant \mathrm{Im}(s) \leqslant T. \tag{17}$$

В предположении, что на его границе нет нулей дзета-функции, теорема Коши говорит, что

$$N(T) = \frac{1}{2\pi \mathrm{i}} \oint \frac{\zeta'(s)}{\zeta(s)}\, \mathrm{d}s, \tag{18}$$

где контурный интеграл берётся по границе прямоугольника (17). Поскольку по определению $N(T)$ — число целое, для его нахождения достаточно вычислить правую часть (18) с ошибкой, не превосходящей, скажем, $1/3$, а затем провести округление до ближайшего целого числа. Такое вычисление и строгую оценку его погрешности можно сделать средствами численного анализа.

Теорема Коши — это общее свойство всех аналитических функций. В нашем конкретном случае дзета-функции есть ещё одна специфическая техника для подсчёта количества нулей.

Рассмотрим функцию

$$Z(t) = e^{\mathrm{i}\theta(t)} \zeta\big(\tfrac{1}{2} + \mathrm{i}t\big), \tag{19}$$

где

$$\theta(t) = \mathrm{Im} \ln \left( \Gamma\left( \frac{\mathrm{i}t}{2} + \frac{1}{4} \right) \right) - \frac{t}{2} \ln(\pi). \tag{20}$$

Экспоненциальный множитель в (19) нулей, очевидно, не имеет. Когда $t$ принимает вещественные значения, аргумент $\frac{1}{2} + \mathrm{i}t$ у дзета-функции в (19) принимает значения на критической прямой, а значения функции $Z(t)$ оказываются вещественными.

Представим теперь, что мы нашли $N(T) + 1$ вещественное число $f_0$, …, $f_{N(T)}$ (белые точки на оси абсцисс на рис. 5) такое, что

$$1 \leqslant f_0 < \cdots < f_k < f_{k+1} < \cdots < f_{N(T)} \leqslant T \tag{21}$$

и

$$Z(f_{k-1}) Z(f_k) < 0 \tag{22}$$

при $k = 1, \ldots, N(T)$. Очевидно, что в таком случае функция $Z(t)$ имеет, по крайней мере, один нуль в каждом из открытых интервалов $(f_0, f_1)$, …, $(f_{N(T)-1}, f_{N(T)})$ и, следовательно, функция $Z(t)$, а, значит, и функция $\zeta(\frac{1}{2} + \mathrm{i}t)$, имеют по крайней мере $N(T)$ нулей при $1 \leqslant t \leqslant T$. Поскольку $N(T)$ — количество нулей во всём прямоугольнике, то никаких других нулей дзета-функции в нём быть не может.

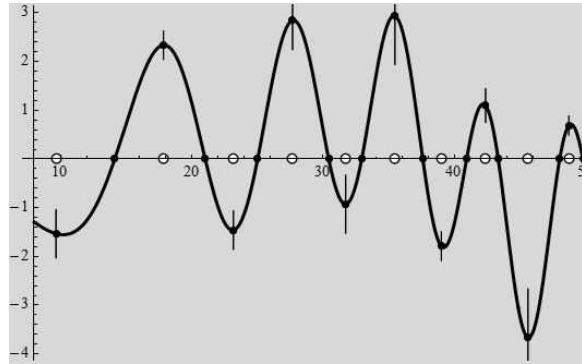



*Рис. 5. Функция $Z(t)$*

Отметим, что для проверки неравенств (22) нам не требуется вычислять точные значения функции $Z(t)$, мы можем делать это с некоторой погрешностью и оценивать её. Если оценка погрешности окажется по абсолютной величине меньше вычисленного приближенного значения $Z(t)$, то мы будем достоверно знать знак истинного значения $Z(t)$.

При практической реализации этого плана возникает вопрос — как найти числа $f_0, \ldots, f_{N(T)}$ с требуемыми свойствами? Конечно, можно вычислять значение $Z(1+kh)$ при $k = 1, \ldots, (T-1)/h$ и все уменьшающимися значения $h$ и ждать появления $N(T)$ перемен знака, но это очень трудоёмкий процесс.

Грам указал эвристический метод выбора значений $f_0, \ldots, f_{N(T)}$. Он рассмотрел мнимую часть экспоненциального сомножителя в (19)

$$\mathrm{Im}(e^{\mathrm{i}\theta(t)}) = \sin(\theta(t)). \tag{23}$$

Функция $\theta(t)$ имеет асимптотическое разложение

$$\theta(t) = \frac{1}{2}\left(\ln\left(\frac{t}{2\pi}\right) - 1\right)t - \frac{\pi}{8} + O(t^{-1}), \tag{24}$$

из которого видно, что $\theta(t)$ растёт почти как линейная функция с логарифмически медленно увеличивающимся коэффициентом при $t$. Соответственно, график функции $\sin(\theta(t))$ выглядит как синусоида с возрастающей частотой (см. рис. 6).

Из рис. 7, на котором изображены функции $Z(t)$ и $\sin(\theta(t))$, видно, что они ведут себя подобно синусу и косинусу — нули каждой из этих функций соответствуют экстремумам другой. Положительные нули функции $\sin(\theta(t))$ получили название *точек Грама*, по традиции их обозначают $g_m$ и нумеруют так, что

$$\theta(g_m) = m\pi, \tag{25}$$

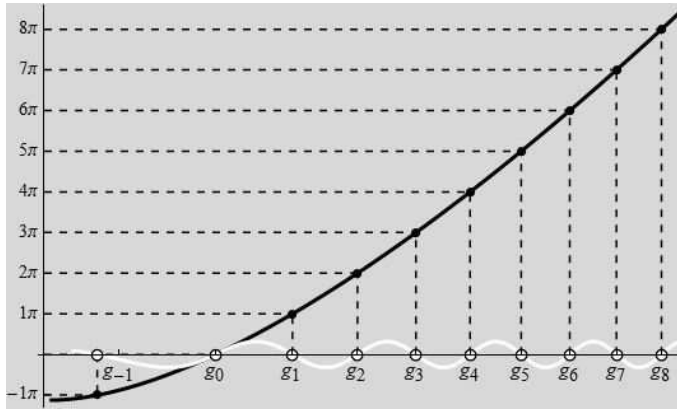



*Рис. 6. Функции $\theta(t)$ и $\sin(\theta(t))$*

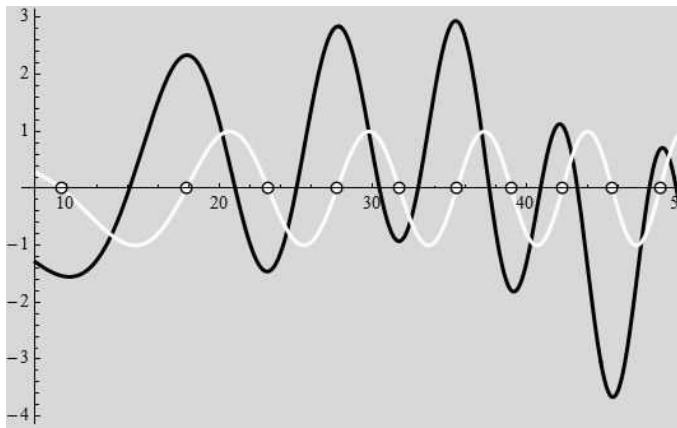



*Рис. 7. Функции $Z(t)$ и $\sin(\theta(t))$*

то есть первый нуль имеет индекс $-1$. Грам, вычислив начальные нули дзета-функции, обнаружил, что они чередуются с числами $g_m$ и при этом

$$(-1)^m Z(g_m) > 0. \tag{26}$$

Джон Ирвин Хатчинсон (John Irwin Hutchinson) назвал неравенство (26) *законом Грама*, и, по иронии судьбы, вычислив начальные 138 нулей, он же обнаружил первые исключения из этого «закона»

$$\begin{aligned} Z(g_{133}) &= -0.7763\ldots, \\ Z(g_{134}) &= -0.0169\ldots, \\ Z(g_{135}) &= -3.4698\ldots. \end{aligned}$$

Сейчас мы знаем, что «закон Грама» нарушается бесконечно часто, и можем различать *хорошие точки Грама*, удовлетворяющие неравенству (26), и *плохие*, ему неудовлетворяющие. Таким образом, при большой величине $T$ мы не можем брать точки Грама непосредственно на роль чисел $f_0, \ldots, f_{N(T)}$ в (21) и (22). Мы можем, однако, попытаться немного сдвинуть плохие точки, то есть найти такие (маленькие) числа $h_0, \ldots, h_{N(T)}$, для которых неравенства (21) и (22) выполняются при выборе $f_k = g_k + h_k$.

В итоге мы приходим к следующему классическому методу (использованному предшественниками Тьюринга) для проверки того, что все нули дзета-функции Римана, имеющие мнимую часть, не превосходящую некоторого числа $T$, лежат на критической прямой.

1. Вычислить с погрешностью не более $1/3$ интеграл
$$N(T) = \frac{1}{2\pi \mathrm{i}} \oint \frac{\zeta'(s)}{\zeta(s)}\, \mathrm{d}s,$$
округлить полученное значение до ближайшего целого числа, найдя тем самым $N(T)$ — количество нулей в прямоугольнике (17).
2. Найти $N(T)+1$ чисел $h_{-1}, h_0, \ldots, h_{N(T)-1}$ таких, что
$$1 < g_{-1} + h_{-1} < \ldots < g_m + h_m < \ldots < g_{N(T)-1} + h_{N(T)-1} < T,$$
и проверить, что при $k = 0, \ldots, N(T)-1$
$$Z(g_{k-1} + h_{k-1}) Z(g_k + h_k) < 0.$$

Если нам удастся это сделать, то мы докажем, что все нули функции $\zeta(s)$ в прямоугольнике (17) лежат ровно на критической прямой.

Что же, кроме возможной ошибочности гипотезы Римана, может помешать нам в осуществлении этого плана? Количество нулей можно подсчитывать двумя способами — с учётом кратности и без учёта кратности. Интеграл Коши (18) учитывает кратность, но когда мы оцениваем количество нулей через количество перемен знаков, каждый нуль нечётной кратности считается за один нуль, а нули чётных кратностей вообще не могут быть обнаружены по перемене знака функции. По этому поводу Тьюринг написал:

> We know no way of dealing with multiple zeros, and simply hope that none are present.[10)]

Надежда Тьюринга до сих пор оправдывалась — кратные нули дзета-функции Римана пока не обнаружены.

[10)] Мы не знаем никакого способа для работы с кратными нулями и просто надеемся, что их нет.

Усовершенствование, которое Тьюринг внёс в описанный выше метод проверки гипотезы Римана, относится к первой части — вычислению $N(t)$. Метод Тьюринга основан на одной теореме, опубликованной Литлвудом в 1914 году. Она производит сравнение $N(T)$ с количеством точек Грама, не превосходящих $T$, которое легко определить. Действительно, неравенство $g_m < T$ в силу монотонности функции $\theta$ эквивалентно неравенству $\theta(g_m) < \theta(T)$ и, согласно (25), неравенству $\pi m < \theta(T)$, то есть имеется $\lfloor \theta(T)/\pi \rfloor + 2$ точки Грама, не превосходящие $T$ (слагаемое 2 вызвано нумерацией точек начиная с $-1$). Определив функцию $S(T)$ через равенство

$$N(T) = \frac{\theta(T)}{\pi} + 1 + S(T), \tag{27}$$

мы можем ожидать, что она будет принимать не слишком большие значения. Теорема, доказанная Литлвудом, утверждает, что это верно в среднем.

Теорема [10].

$$\int_0^T S(t)\,\mathrm{d}t = O(\ln(T)). \tag{28}$$

Сам Литлвуд не указал, как эта теорема может быть использована для вычисления $N(T)$. Не увидел этого и Титчмарш, проверивший в 1936 году гипотезу Римана для первых 1041 нулей, и только Тьюринг нашёл применение результата Литлвуда для упрощения вычислений.

Теорема Литлвуда сформулирована с использованием символа $O$ и по этой причине она сама по себе бесполезна для каких-либо фактических вычислений, ибо не утверждает ничего ни про одно конкретное значение $T$. В связи с этим Тьюринг пишет:

> In analysis it is customary to use the notation $O\{f(x)\}$ to indicate 'some function whose ratio to $f(x)$ is bounded'. In the theory of a computation one needs a similar notation, but one is interested in the value of the bound concerned. We therefore use the notation $\Theta(\alpha)$ to indicate 'some number not greater in modulus than $\alpha$'. The symbol $\Theta$ has been chosen partly because of a typographical similarity to $O$, partly because of the relation with the use $\vartheta$ to indicate 'a number less than 1'.[11)]

[11)] В анализе принято использовать запись $O\{f(x)\}$ для обозначения «некоторой функции, отношение которой к $f(x)$ ограниченно». В теории вычислений необходимо подобное обозначение, но при этом интерес представляет значение такой границы. Соответственно, мы будем использовать запись $\Theta(\alpha)$ для обозначения «некоторого числа, не превосходящего по модулю числа $\alpha$». Символ $\Theta$ выбран отчасти из-за типографской схожести с $O$, отчасти из-за использования $\vartheta$ для обозначения «некоторого числа, меньшего 1».

Используя введённое обозначение, Тьюринг уточнил теорему Литлвуда, доказав, что если $168\pi < T_1 < T_2$, то

$$\int_{T_1}^{T_2} S(t)\,\mathrm{d}t = \Theta\left(2.30 + 0.128 \ln\left(\tfrac{T_2}{2\pi}\right)\right). \tag{29}$$

Доказательство занимает в [19] пять страниц традиционных теоретико-числовых оценок.

Дальнейший ход мыслей Тьюринга можно описать так. Наша конечная цель — установить положение нулей дзета-функции Римана с мнимыми частями, не превосходящими по абсолютной величине некоторого числа $T$. Вместо этого мы можем попытаться проделать это для некоторого $T_0$, которое больше $T$, поскольку может оказаться, что вычислить $N(T_0)$ легче, чем вычислить $N(T)$. И действительно, оказалось, что этого можно достичь, если вместо того, чтобы вычислять $S(T_0)$, исходя из значения $T_0$, поступить наоборот — потребовать, чтобы

$$S(T_0) = 0, \tag{30}$$

и искать такое значение $T_0$.

Из (30) и (27) следует, что значение $\theta(T_0)$ кратно $\pi$ и по определению $T_0$ должно быть некоторой точкой Грама $g_m$. Наоборот, если $T_0 = g_m$, то $\theta(T_0)$ также кратно $\pi$ и согласно (27) $S(T_0)$ — число целое. Это означает, что для установления равенства (30) достаточно показать, что выполнены неравенства

$$-1 < S(T_0) < 1. \tag{31}$$

Более того, нетрудно проверить, что если в качестве $T_0$ мы возьмём не произвольную, а хорошую точку Грама, то $S(T_0)$ обязательно будет чётным числом, и вместо неравенств (31) достаточно получить более слабые неравенства

$$-2 < S(T_0) < 2. \tag{32}$$

Итак, пусть $g_m$ — хорошая точка Грама. Несколько последующих точек могут оказаться плохими, и нам потребуется найти сдвиги $h_{m+1}, \ldots, h_{m+k-1}$ такие, что

$$\begin{aligned} (-1)^{m+1} Z(g_{m+1} + h_{m+1}) &> 0, \\ &\vdots \\ (-1)^{m+k-1} Z(g_{m+k-1} + h_{m+k-1}) &> 0. \end{aligned} \tag{33}$$

Рано или поздно (а обычно весьма скоро) мы найдём следующую хорошую точку Грама $g_{m+k}$, для которой $(-1)^{m+k} Z(g_{m+k}) > 0$. Используя

свою версию (29) теоремы Литлвуда, Тьюринг доказал, что

$$S(g_m) \leqslant 1 + \frac{2.30 + 0.128 \ln \left( \frac{g_{m+k}}{2\pi} \right) + \sum_{j=1}^{k-1} h_{m+j}}{g_{m+k} - g_m}. \tag{34}$$

Аналогично можно получить двойственную оценку снизу:

$$S(g_m) \geqslant -1 - \frac{2.30 + 0.128 \ln \left( \frac{g_m}{2\pi} \right) - \sum_{j=1}^{k-1} h_{m-j}}{g_m - g_{m-k}}. \tag{35}$$

Если в (34) и (35) слагаемые под знаками сумм достаточно малы, то мы получаем требуемые неравенства (32). Таким образом, главное преимущество метода Тьюринга состоит в замене трудоёмкого вычисления контурного интеграла в (18) на вычисление значений функции $Z(t)$ в небольшом количестве дополнительных точек.

Вторая часть статьи [19], представляющая ныне главным образом исторический интерес, посвящена конкретному счёту, проведённому Тьюрингом. Это было одно из первых применений ЭВМ для получения нетривиальных математических результатов, и Тьюринг приводит много деталей, которые сегодня уже не принято публиковать в математических журналах. В частности, он описывает устройство машины:

> The storage of the machine is of two kinds, known as 'electronic' and 'magnetic' storage. The electronic storage consisted of four 'pages' each of thirty-two lines of forty binary digits. The magnetic storage consisted of a certain number of tracks each of two pages of similar capacity. Only about eight of these tracks were available for the zeta-function calculations.[12)]

Тьюринг далее приводит таблицу распределения памяти — см. рис. 8.

Результаты счёта выдавались на перфоленту с пятью рядами отверстий, и потому использовалась система счисления с основанием 32. Содержимое ленты можно было потом распечатать. Тьюринг счёл нужным привести таблицу соответствия цифр и символов на распечатке — см. рис. 9.

Про использование десятичной записи Тьюринг написал следующее:

> More conventionally the scale of 10 can be used, but this would require the storage of a conversion routine, and the writer was entirely content to see the results in the scale of 32, with which he is sufficiently familiar.[13)]

[12)] Память машины была двух типов, известных как «электронная» и «магнитная» памяти. Электронная память состояла из четырёх «страниц», каждая из которых имела по тридцать две строки из сорока двоичных разрядов. Магнитная память имела некоторое количество треков по две страницы такого же размера. Только примерно восемь из этих треков были доступны при вычислениях дзета-функции.

[13)] Более традиционно можно использовать основание 10, но это потребовало бы

The storage available was distributed as follows:

*Magnetic store*

| | |
|---|---|
| Logarithms routine (for $\kappa$) | 1 page |
| Table of logarithms and reciprocal square roots | 4 pages |
| Routine for calculating the terms $n^{-\frac{1}{4}}\cos 2\pi(\tau \log n - \kappa)$ and table of cosines | 2 pages |
| Remainder of routine for calculating the function $Z(\tau)$ | 2 pages |
| Input routine | 2 pages |
| Output routine | 2 pages |

*Electronic store*, as occupied during the greater part of the time

| | |
|---|---|
| Instructions and cosines | 2 pages |
| Logarithms and reciprocal square roots | 1 page |
| Miscellaneous data and working space | 1 page |

*Рис. 8. Распределение памяти*

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| / | E | @ | A | : | S | I | U | ‡ | D | R | J | N | F | C | K | T | Z | L | W |

| 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| H | Y | P | Q | O | B | G | ‖ | M | X | V | £ |

*Рис. 9. Кодировка цифр системы счисления с основанием 32*

Большой проблемой была работоспособность машины. Тьюринг пишет:

> If it had not been for the fact that the computer remained in serviceable condition for an unusually long period from 3 p.m. one afternoon to 8 a.m. the following morning it is probable that the calculations would never have been done at all.[14)]

Помимо интервала $2\pi 63^2 < t < 2\pi 64^2$, Тьюринг хотел продолжить вычисления Титчмарша:

> Another investigation was also started with a view to extending the range of relatively small values of $t$ for which the Riemann hypothesis holds. Titchmarsh has already proved that it holds up to $t = 1468$, i.e. to about $\tau = 231\ldots$ It was intended to continue the work up to about $\tau = 500\ldots$[15)]

---

памяти для процедуры преобразования, а автор [Тьюринг] был вполне удовлетворён увидеть результаты с основанием 32, к которому он достаточно привычен.

[14)] Если бы не тот факт, что компьютер однажды оставался в рабочем состоянии необычно долгий период от 3 часов пополудни до 8 часов утра следующего дня, то вычисления вообще никогда не были бы проведены.

[15)] Было начато другое исследование с целью расширить область сравнительно маленьких значений $t$, для которых гипотеза Римана справедлива. Титчмарш уже доказал,

*Табл. 3. Проверка гипотезы Римана, начиная с А. Тьюринга*

| Год | Количество нулей | Автор |
|---|---|---|
| 1953 | 1104 | A. M. Turing |
| 1956 | 25000 | D. H. Lehmer |
| 1958 | 35337 | N. A. Meller |
| 1966 | 250000 | R. S. Lehman |
| 1968 | 3500000 | J. B. Rosser, J. M. Yohe, L. Schoenfeld |
| 1977 | 40000000 | R. P. Brent |
| 1979 | 81000001 | R. P. Brent |
| 1982 | 200000001 | R. P. Brent, J. van de Lune, H. J. J. te Riele, D. T. Winter |
| 1983 | 300000001 | J. van de Lune, H. J. J. te Riele |
| 1986 | 1500000001 | J. van de Lune, H. J. J. te Riele, D. T. Winter |
| 2004 | 900000000000 | S. Wedeniwski |
| 2004 | 10000000000000 | X. Gourdon |

> The interval $1414 < t < 1608$ was investigated and checked, but unfortunately at this point the machine broke down and no further work was done. Furthermore this interval was subsequently found to have been run with a wrong error value, and the most that can consequently be asserted with certainty is that the zeros lie on the critical line up to $t = 1540, \ldots$ a negligible advance.[16)]

Действительно, если сравнить количество нулей, проверенных Тьюрингом, с количеством нулей, проверенных до него (таблица 2) и после него (таблица 3), то достижения Тьюринга кажутся незначительными. На самом деле вклад Тьюринга огромен. Он не только был первым, кто применил компьютер для проверки гипотезы Римана (рано или поздно это сделал бы кто-либо иной), но, что гораздо важнее, как и предвидел

что она верна до $t = 1468$, то есть примерно до $\tau = 231\ldots$ Предполагалось продолжить работу до примерно $\tau = 500\ldots$

[16)] Интервал $1414 < t < 1608$ был исследован и проверен, но, к сожалению, в этот момент машина сломалась и никакой дальнейшей работы проведено не было. Впоследствии было обнаружено, что счёт для этого интервала был проведён с неправильной величиной ошибки, и самое большое, что можно утверждать с уверенностью, — это то, что нули лежат на критической прямой вплоть до $t = 1540, \ldots$ несущественное продвижение.

Тьюринг, все последующие вычисления, вплоть до наших дней, проводились по его методу.

Традиционно в описаниях исследований Алана Тьюринга по теории чисел говорится, что его научное наследие состоит из нескольких писем к другим математикам, неопубликованных рукописей, и только двух печатных работ — [18, 19]. Это, однако, не совсем так. Есть ещё одна опубликованная Тьюрингом работа, где он изучает гипотезу Римана. По какой-то причине эта работа практически не цитируется специалистами по теории чисел. Возможно, они не ценят полученный там результат, а, возможно, они просто не знают, что в работе [17], названной *Systems of logic based on ordinals*[17)] Тьюринг изучает, в частности, гипотезу Римана (эта статья — изложение диссертации Тьюринга).

Третий раздел этой работы назван «Теоретико-числовые теоремы». Тьюринг начинает его с того, что даёт формальное определение:

> By a number-theoretic theorem we shall mean a theorem of the form "$\theta(x)$ vanishes for infinitely many natural numbers $x$", where $\theta(x)$ is a primitive recursive function.[18)]

Тьюринг тут же делает подстрочное примечание:

> I believe that there is no generally accepted meaning for this term, but it should be noticed that we are using it in a rather restricted sense.[19)]

Затем Тьюринг даёт второе, эквивалентное определение «теоретико-числовых теорем» в его смысле:

> An alternative form for number-theoretic theorems is "for each natural number $x$ there exists a natural number $y$ such that $f(x, y)$ vanishes", where $f(x, y)$ is primitive recursive.[20)]

Оба этих определения, по-видимому, чужды специалистам по теории чисел, поскольку в них используется понятие примитивно рекурсивной функции из теории вычислимости. Можно, однако, дать третье определение, более близкое по духу к теории чисел.

Обозначим через $\Pi_0^0$ и через $\Sigma_0^0$ класс всех арифметических формул, в которых все кванторы ограничены. После выбора значений свободных

---

[17)] *Системы логики, основанные на ординалах.*

[18)] Под теоретико-числовой теоремой мы будем понимать теорему вида «$\theta(x)$ обращается в ноль для бесконечно многих натуральных $x$», где $\theta(x)$ — примитивно рекурсивная функция.

[19)] Я полагаю, что нет общепринятого понимания этого термина, но следует отметить, что мы используем его в очень узком смысле.

[20)] Альтернативной формой теоретико-числовых теорем является «для каждого натурального числа $x$ существует натуральное число $y$, при котором $f(x, y)$ обращается в ноль», где $f(x, y)$ — примитивно рекурсивная функция.

переменных в таких формулах мы, очевидно, можем установить их истинность или ложность. Далее, пусть $\Pi^0_n$ обозначает при $n > 0$ класс всех формул вида

$$\forall x_1 \ldots x_m \varphi \tag{36}$$

где $\varphi$ — формула из класса $\Sigma^0_{n-1}$, и аналогично $\Sigma^0_n$ — класс формул вида

$$\exists x_1 \ldots x_m \psi \tag{37}$$

где $\psi$ — формула из класса $\Pi^0_{n-1}$. Мы получаем *арифметическую иерархию*, в которой каждый класс содержит любой другой класс, расположенный левее него на этой схеме:

$$\begin{array}{ccccccc} & & \Pi^0_1 & \Pi^0_2 & \Pi^0_3 & \ldots \\ \Pi^0_0 = \Sigma^0_0 & & & & & & \\ & & \Sigma^0_1 & \Sigma^0_2 & \Sigma^0_3 & \ldots \end{array} \tag{38}$$

Теоретико-числовые теоремы в смысле Тьюринга — это в точности (истинные) формулы из класса $\Pi^0_2$.

Чтобы мотивировать введённое название, Тьюринг приводит пример — *Великую теорему Ферма*. В качестве менее очевидного примера Тьюринг указывает гипотезу Римана.

Сначала вообще непонятно, почему где-то в арифметической иерархии лежит гипотеза Римана, поскольку исходно она была сформулирована как утверждение про комплексные числа. Тем не менее, их можно представить как пары вещественных чисел, которые, в свою очередь, могут быть представлены как пределы (сходящихся) последовательностей рациональных чисел. Мощная техника *арифметизации*, развитая Куртом Гёделем, позволяет в итоге записать гипотезу Римана как арифметическую формулу с бо́льши́м числом кванторов по натуральным числам.

Таким образом, в арифметической иерархии (38) есть формулы, эквивалентные гипотезе Римана, причём они могут встречаться в разных классах. Возникает естественный вопрос — сколь малым может быть такой класс?

Поставленный в такой форме, этот вопрос является бессодержательным или тривиальным, поскольку ответ очевиден — конечно, это самый маленький класс $\Pi^0_0 = \Sigma^0_0$. Действительно, гипотеза Римана — это конкретное утверждение, не содержащее параметров, и оно либо истинно, либо ложно. В первом случае гипотеза Римана эквивалентна формуле $0 = 0$, во втором — формуле $0 = 1$, и обе эти формулы принадлежат классу $\Pi^0_0 = \Sigma^0_0$.

Более корректной является следующая постановка вопроса: где в арифметической иерархии (38) мы можем указать, *при нашем нынешнем уровне знаний*, формулу, эквивалентную гипотезе Римана?

Алан Тьюринг нашёл такую формулу в классе $\Pi_2^0$, то есть в его классе теоретико-числовых теорем. Несложное доказательство, занимающее чуть больше одной страницы, не вызвало, по-видимому, большого интереса у специалистов по теории чисел.

В 1958 году математический логик Георг Крайзель (Georg Kreisel) усилил результат Тьюринга, построив формулу из класса $\Pi_1^0$, эквивалентную гипотезе Римана. Идею его конструкции можно пояснить следующим образом.

Известно, что все невещественные нули дзета-функции Римана расположены симметрично относительно прямых $\mathrm{Re}(s) = 1/2$ и $\mathrm{Im}(s) = 0$, разбивающих всю комплексную плоскость на четыре квадранта, поэтому достаточно проверить, что нет нулей, скажем, в открытом квадранте $\mathrm{Re}(s) > 1/2 \;\&\; \mathrm{Im}(s) > 0$.

Мы можем представить этот квадрант как счётное объединение содержащихся в нём прямоугольников и сформулировать гипотезу Римана как утверждение, что ни один из этих прямоугольников не содержит нулей дзета-функции. К сожалению, непосредственно записать это утверждение как неравенство с соответствующим контурным интегралом мы не можем, поскольку a priori ноль может лежать на границе использованного нами прямоугольника, но Крайзель нашёл способ обойти эту трудность.

Следующим шагом в усилении результатов Тьюринга и Крайзеля стало бы помещение гипотезы Римана в класс $\Pi_0^0 = \Sigma_0^0$, то есть её доказательство или опровержение, но это сейчас находится вне пределов наших возможностей.

Тем не менее, это направление исследований получило некоторое развитие. В 1970 году была доказана так называемая DPRM-теорема[21)], устанавливающая, что в определении класса $\Pi_1^0$ в роли формулы $\varphi$ всегда можно брать формулу вида $P(x_1 \dots x_m) \not\equiv 0$, где $P$ — многочлен с целыми коэффициентами. Совместно с результатом Крайзеля это дало такое следствие: *можно построить конкретный многочлен $R(x_1 \dots x_m)$ с целыми коэффициентами такой, что гипотеза Римана эквивалентна утверждению об отсутствии решения у диофантова уравнения $R(x_1 \dots x_m) = 0$* (конкретный способ построения такого уравнения описан, например, в [1]).

Гипотеза Римана была включена Давидом Гильбертом в его 8-ю проблему, а в 10-й проблеме он предлагал найти алгоритм для распознавания разрешимости произвольного диофантова уравнения. Таким образом, начатое Тьюрингом выяснение вопроса о положении гипотезы Римана в

[21)] Аббревиатура в названии теоремы образована из первых букв фамилий математиков Мартина Дэвиса (Martin Davis), Хилари Патнема (Hilary Putnam), Джулии Робинсон (Julia Robinson) и Юрия Матиясевича. — *Прим. ред.*


*Рис. 10. Памятник Алану Тьюрингу в Сэквиль-парке, Манчестер*

арифметической иерархии привело к установлению неожиданной для специалистов по теории чисел связи между 8-й и 10-й проблемами Гильберта — оказалось, что гипотезу Римана можно переформулировать как очень частный случай 10-й проблемы Гильберта.

Без тезиса Чёрча – Тьюринга мы не могли бы даже сформулировать, что означает неразрешимость 10-й проблемы Гильберта, установление которой было стимулом для доказательства DPRM-теоремы. Сейчас мы имеем формальное доказательство того, что никакая машина Тьюринга не может распознавать, имеет ли произвольное диофантово уравнение решение или нет. Сведéние же гипотезы Римана к конкретному диофантову уравнению, первый шаг на пути к которому сделал Алан Тьюринг, даёт «психологическое» объяснение трудности диофантовых уравнений — было бы удивительно ожидать существование требуемого Гильбертом алгоритма для диофантовых уравнений, ибо тогда можно было бы, по крайней мере в принципе, «механически» установить справедливость или ложность такой трудной гипотезы, как гипотеза Римана.

## Список литературы

[1] Матиясевич Ю. В. *Десятая проблема Гильберта*. М.: Физматлит, 1993. `http://logic.pdmi.ras.ru/~yumat/H10Pbook`,

[2] Booker A. R. *Turing and the Riemann Hypothesis* // Notices Amer. Math. Soc. Vol. 53, no. 10. 2006 P. 1208–1211.

[3] Booker A. R. *Artin's Conjecture, Turing's Method, and the Riemann Hypothesis* // Experiment. Math. Vol. 15, issue 4. 2006. P. 385–408.

[4] *Alan Turing—His Work and Impact*. Под ред. S. B. Cooper, J. van Leeuwen. Elsevier, 2013.

[5] *The Undecidable. Basic Papers on Undecidable Propositions, Unsolvable Problems and Computable Functions*. Под ред. M. Davis. Hewlett, NY: Raven Press, 1965. Переиздано Dover Publications, 2004.

[6] Hejhal D. A. *A few comments about Turing's method* // [4].

[7] Hejhal D. A., Odlyzko A. M. *Alan Turing and the Riemann zeta function* // [4].

[8] Ingham A. E. *The Distribution of Prime Numbers*. Cambridge University Press, 1932; Перепечатано Stechert-Hafner, Inc., New York, 1964; Cambridge University Press, Cambridge, 1990. Переводы на русский язык: Главная редакция общетехнической литературы и номографии, Москва–Ленинград, 1936; Едиториал УРСС, 2005.

[9] Ingham A. E. *A note on the distribution of primes* // Acta Arithmetica. Vol. 1. 1936. P. 201–211.

[10] Littlewood J. E. *Sur la distribution des nombres premiers* // Comptes rendus de l'Academie des sciences. Vol. 158. 1914. P. 1869–1872. Перепечатано в *Collected Papers of J. E. Littlewood*. The Clarendon Press, Oxford University Press, New York, 1982.

[11] Kreisel G. *Mathematical significance of consistency proofs* // Journal of Symbolic Logic. Vol. 23(2). 1958. P. 155–182.

[12] Odlyzko A. M., Schönhage A. *Fast Algorithms for Multiple Evaluations of the Riemann Zeta Function* // Transactions of the American Mathematical Society. Vol. 309. No 2. 1988. P. 797–809.

[13] Riemann B. *Über die Anzhal der Primzahlen unter einer gegebenen Grösse*. Monatsberichter der Berliner Akademie, 1859. // Riemann B. *Gesammelte Werke*. Teubner, Leipzig, 1892; reprinted by Dover Books, New York, 1953. `http://www.claymath.org/millennium/Riemann_Hypothesis/1859_manuscript/zeta.pdf` , English translation `http://www.maths.tcd.ie/pub/HistMath/People/Riemann/Zeta/EZeta.pdf` .

[14] Skewes S. *On the Difference* $\pi(x) - \mathrm{Li}(x)$ (I) // Journal of the London Mathematical Society. Vol. 8. 1933. P. 227–283.

[15] Skewes S. *On the Difference* $\pi(x)-\mathrm{Li}(x)$ (II) // Proceedings of the London Mathematical Society. Vol. 5. 1955. P. 48–70.

[16] Turing A. M. *On computable numbers, with an application to the Entscheidungsproblem* // Proc. London Math. Soc. Ser. 2. Vol. 42. 1937. P. 230–265. Correction, ibid. Vol. 43. 1938. P. 544–546. Перепечатано в [4, 5, 20]

[17] Turing A. M. *Systems of logic based on ordinals* // Proceedings of the London Mathematical Society. Ser. 2. Vol. 45. 1939. P. 161–228. Перепечатано в [4, 5, 20]

[18] Turing A. M. *A method for the calculation of the zeta-function* // Proceedings of the London Mathematical Society. Ser. 2. Vol. 48. 1943. P. 180–197. Перепечатано в [4, 21].

[19] Turing A. M. *Some calculations of the Riemann zeta-function* // Proceedings of the London Mathematical Society. Ser. 3. Vol. 3. 1953. P. 99–117. Перепечатано в [4, 21].

[20] Turing A. M. *Collected Works of A. M. Turing: Mathematical Logic*. (Под ред. R. O. Gandy, C. E. M. Yates.) Amsterdam: Elsevier Science Publishers. 2001.

[21] Turing A. M. *Collected Works of A. M. Turing: Pure Mathematics*. (Под ред. J. L. Britton.) Amsterdam: North-Holland. 1992.

[22] Welles D. *The Penguin Dictionary of Curious and Interesting Numbers*. London: Penguin Books. 1997.

[23] Clay Mathematics Institute Millenium Problems `http://www.claymath.org/millennium`

[24] the Turing Digital Archive. `http://www.turingarchive.org`

Ю. В. Матиясевич, академик РАН, Санкт-Петербургское отделение Математического института им. В. А. Стеклова. 191023, г. Санкт-Петербург, наб. реки Фонтанки, д. 27.
Email: `yumat@pdmi.ras.ru`

# Заметки о математическом образовании во Франции

С. Тищенко

*От редколлегии*. С. Тищенко — выпускник Второй школы, который затем учился и преподавал во Франции. Нам представляется, что его заметки о системе французского математического образования будут интересны многим читателям.

Сравнение образовательных систем — сложная и богатая тема, для неё необходимо серьёзное многотомное научное исследование, которое обязательно должно проводиться в нашей стране. Без крепкой фундаментальной базы опрометчиво браться за реформы в России, от которых будет зависеть многое в нашем будущем. Я уверен, что серьёзное изучение французской системы образования уже проводилось и существуют более подробные научные труды на эту тему. Тем не менее, опишу те моменты, которые мне видятся определяющими и острыми. Ограничусь рассмотрением только той стороны французской системы образования, которая касается математики и её преподавания.

Математика во Франции есть любого уровня, в любом количестве и на любой вкус. Нужно только знать «правильный адрес» и иметь «пропуск», чтобы увидеть те «кузницы», в которых французская математика обретает свои формы, те «фабрики», в которых математическое образование реализует себя. Однако, французский мир математики, как науки, так и образования, не имеет строгой иерархии. Математика во Франции не выстроена вдоль какой-то вертикали, а существует отделёнными микросообществами, живущими независимо. Даже французу, находящемуся внутри французской системы образования, часто бывает мало что понятно о ней в целом. Поэтому нашему соотечественнику, столкнувшемуся с французскими реалиями, в первые годы тоже не всё понятно и многое вызывает удивление или даже возмущение.

## Высшее образование

Первое, и самое важное, что надо учитывать, — это разделение французского высшего образования на две параллельные системы. Университетскую и Высшие школы. Это разделение приводит к тому, что каждый

год лучшие 3 тысячи студентов-математиков идут в систему Высших школ, и только иногда на пятом — последнем — году обучения появляются в университете. Во французских университетах, а их около 80, на первых курсах учатся только «условно способные» к математике студенты, либо большие романтики, бóльшую часть времени в университете откровенно скучающие.

Высшие школы во Франции крайне разнообразны. Большинство из них выпускает инженеров, функционеров, администраторов и финансистов. Эти школы благодаря более привлекательным перспективам последипломного трудоустройства каждый год уводят «из-под носа» университетов около 3 тысяч наиболее способных к математике абитуриентов, а также «львиную» долю государственного финансирования образовательной системы. Это, казалось бы, должно наносить удар по французской науке и математике в частности. Но в реальности это не так. Около 200 лучших студентов второго года в системе Высших школ по итогам общенациональных конкурсов каждый год становятся студентами Высших Нормальных школ и специализируются в математике. Из их числа 50 студентов становятся элитой французской математики, поступив в Высшую Нормальную школу «Ульм». Иностранцев среди них практически нет.

Система отбора в элиту французской математики работает во Франции практически идеально. Среди французских лауреатов Филдсовской премии 10 из 11 закончили Высшую Нормальную школу «Ульм», это больше, чем число всех советских и российских лауреатов вместе взятых. А ведь в эту школу каждый год поступает на отделение математики только 50 человек. Стоит сказать и то, что Школа «Ульм» значительно старше самой Филдсовской премии и среди её выпускников — плеяды ярчайших математиков.

Время реванша для университетов наступает на 5 курсе и в аспирантуре. Высшие школы, за редким исключением, занимаются научной деятельностью значительно меньше, чем университеты. Французские университеты — это, в первую очередь, научные центры. И неудивительно, например, что на 5 курсе таких университетов как Пьер и Мари Кюри или Дени Дидро все или почти все студенты из Высших Нормальных школ. Практически все аспиранты парижских университетов либо иностранцы, либо выпускники Высших Нормальных школ. В провинциальных университетах эта тенденция менее подчёркнута, но местами повторяется с тем лишь отличием, что студенты приходят из менее престижных Высших школ, расположенных в том же городе, что и провинциальный университет.

Куда же пропадают студенты, учившиеся в университете до 5 курса? Многих из них не переводят на следующий курс и отпускают из

университета с дипломом о неполном высшем образовании. Некоторые переводятся в университеты «послабее», некоторые уходят на специализированные курсы и становятся преподавателями средних школ и лицеев и, наконец, многие покидают университетскую среду и начинают свою профессиональную жизнь уже вдали от математики.

То, с чем может столкнуться преподаватель младших курсов французского университета, удивит даже человека, видавшего многое. Случается, что на 2 курсе парижского университета по специальности «фундаментальная математика» только половина студентов может посчитать сумму геометрической прогрессии, на 3 курсе — только малая часть студентов сосчитает векторное произведение. Но, как следует из написанного ранее, это нисколько не характеризует французскую математику или саму систему образования. Во Франции никто не ставит цель производить каждый год десять тысяч высококлассных математиков. На деле получается, что преподаватели французских университетов каждый год и на каждом курсе сталкиваются со слабо подготовленными студентами и начинают свой курс с повторения элементарных основ математики. Преподаватели оказываются «прижатыми спиной к стене» и вынуждены приспосабливаться и «урезать» программу курса. И можно понять возникающее в такой ситуации удивление. Такую ситуацию отчасти можно объяснить высокой безработицей среди молодёжи, и университет в такой ситуации оказывается едва ли не единственной возможностью найти себе в дальнейшем работу, не связанную с тяжёлым физическим трудом.

Сам сотрудник французского университета — в первую очередь учёный, преподающий на младших курсах ради возможности заниматься наукой, обладая статусом государственного функционера. Статус этот даётся один раз и на всю жизнь. Постоянные профессора и доценты работают во французских университетах по контрактам, действующим до выхода на пенсию. При этом сам контракт отличается от того, который подписывает обычный сотрудник обычного предприятия. Функционеры во Франции обладают множеством положенных по статусу прав и привилегий.

Таким образом, каждый год во Франции почти все лучшие в математике студенты учатся в Высших школах и Подготовительных классах и становятся менеджерами, инженерами, финансистами, политиками, военными и чиновниками. Полсотни лучших из них поступают в Высшую Нормальную школу и становятся учёными и, в большинстве своём, профессорами французских университетов. В университетах же на младших курсах встретить человека, интересующегося математикой крайне сложно. Но, бесспорно, есть и яркие исключения.

## Среднее образование

Уровень среднего образования во Франции последние 40 лет немного снижается, понижается и уровень требований. Со слов старшего поколения преподавателей, немного ухудшилась дисциплина и отношение к учёбе в школах. Появились так называемые «Зоны приоритетного образования», т. е. кварталы с неспокойной социальной обстановкой, большим числом безработных и не полностью адаптировавшихся иммигрантов. Математика в таких школах заботит всех в последнюю очередь.

Большинство учителей во французских школах и лицеях — бывшие студенты младших курсов французских университетов, прошедшие один или два года интенсивной подготовки к преподавательской деятельности. Сама подготовка не педагогическая, а скорее научная. Всё внимание уделяется совершенствованию владения предметом преподавания и строгости и ясности его изложения.

Учителя во французских школах, как и преподаватели в университетах, обладают статусом государственных служащих, и контракт с ними заключается до выхода на пенсию. Распределение учителей происходит по результатам ежегодного общенационального конкурса согласно предпочтениям кандидатов, сами школы выбирать возможности не имеют. В результате в самых отстающих средних школах в самых проблемных кварталах оказываются наименее подготовленные учителя.

Во Франции государственное образование и обучение в университетах полностью бесплатное. Но практически нет и стипендий. Существуют очень редкие государственные, региональные и социальные стипендии. Раньше существовали «учительские» стипендии, которые выдавались лучшим ученикам отдельных школ для получения высшего образования с обязательством вернуться в свою школу и стать учителем. Отмена этой практики, постоянный спрос на финансистов и программистов и возникновение всё новых Высших школ по этим специальностям приводит к снижению общего уровня преподавательского состава средних школ во Франции.

Тем не менее, математика во французских школах преподаётся на высоком среднем уровне, причём нет большой разницы между столичными школами и школами из «глубинки». Учителя во Франции располагают великолепно проработанной методической литературой и методологической подготовкой. Учителя редко выходят за рамки общего курса, хотя и располагают полной свободой в выборе учебной программы, учебника по которому изучается предмет. Сама программа, с учётом дополнительного года для французских школьников, не уступает программе российской. Здесь нужна небольшая оговорка: аттестат о среднем образовании вручается после прохождения французского единого государственного экзамена — Baccalauréat, сокращённо — Бак. Аттестат этот даётся по результатам

общей суммы баллов, набранных по всем дисциплинам с соответствующими им коэффициентами. Это значит, что можно окончить школу и получить диплом, не набрав на выпускном экзамене по математике практически ни одного балла, если по остальным предметам у вас достаточно хорошие оценки.

Бак существенно отличается от ЕГЭ. Это не тестирование, а настоящая письменная работа. Задание составлено таким образом, что для школьников набрать абсолютный балл за отведённое время практически невозможно. В случае если становится известно о каких-то нарушениях в процедуре проведения экзамена Бак, все результаты по области аннулируются и тысячи учащихся сдают предмет заново. Если оказывается, что ученик каким-то образом нарушил правила проведения экзамена, то его лишают права на обучение во всех вузах страны на срок 5 лет. То же наказание предусмотрено в теории и за списывание на любых школьных и университетских экзаменах. Сама система экзамена Бак хорошо защищена, и о нарушениях можно слышать крайне редко.

## Лицеи — подготовительные курсы

Ещё не сданы все выпускные экзамены, а будущие студенты уже начинают устраивать свою взрослую жизнь. Высшие школы, в отличие от университетов набирают студентов с 3 курса. Поэтому все стремятся попасть на так называемые Подготовительные курсы к Высшим школам. Подготовительные курсы являются первыми двумя годами высшего образования. Учебные заведения, готовящие по программе Подготовительных курсов во Франции, — лицеи. В стенах самих лицеев учат в первую очередь учеников трёх старших классов средней школы. При лучших лицеях открываются Подготовительные классы для студентов младших курсов. Подготовительные курсы представляют собой 2 года занятий очень высокой интенсивности. Занятия ведутся 6 дней в неделю, начинаются каждый будний день в 8 утра и заканчиваются в 8 вечера с перерывом только на обед. Занятия проходят «парами» по два астрономических часа. Студенты сверх того регулярно получают очень тяжёлые домашние задания. В результате практически всё время на обед и все воскресения уходят на выполнение этих заданий. Большинство студентов Подготовительных классов работают над предметами до двух-трёх часов ночи, сильно ограничивая время своего сна. Пропускать занятия невозможно, за каждый прогул следует немедленный выговор.

Во Франции уверены, что только человек, способный выдержать такие нагрузки, может профессионально заниматься наукой. Математиков при этом в обязательном порядке учат и другим дисциплинам: физике, инженерии, информатике, химии и т. д. Два года физики, например, на

направлении «математика» в лучших Подготовительных классах во Франции эквивалентны по объёму материала и лабораторных работ объёму общей физики в лучших российских физических вузах, таких как, например, Физтех или Физфак МГУ. Надо при этом заметить, что программа Подготовительных классов не выходит за рамки общей физики и многие аспекты затрагиваются на лекциях лишь поверхностно. Основная часть времени уделяется теоретической физике. Преподаются также инженерные науки. Таким образом сохраняются старые традиции французской школы, согласно которым французские математики всегда хорошо владели предметом физики.

## Поступление в Высшую школу

По прошествии двух лет студенты сдают около дюжины параллельных экзаменов в разные Высшие школы. Эти экзамены в обязательном порядке состоят из письменных и устных испытаний по всем предметам. Экзамены проходят во всех уголках Франции, в каждую школу в новом городе, и для студентов это часто бывает уникальной возможностью за один месяц объехать всю свою страну по всему её периметру. По результатам этих экзаменов студентов классифицируют. У всех Высших школ появляется своя классификация. И лучший в каждом списке получает почётное право первым выбрать ту Высшую школу, в которой хочет учиться, за ним выбирает второй и так далее. Заканчивается это обычно тем, что первые студенты из каждого списка поступают в Высшую Нормальную школу «Ульм» или Политехническую школу, в которой, в условиях иногда близких к военным, готовят большую часть высших функционеров Франции. Остальные же выбирают то, что им останется. Так в один день люди становятся метеорологами, банкирами, военными офицерами или чиновниками, часто выбирая по остаточному принципу. Университетам же остаётся принимать всех тех, кто не выживает в этой системе или идёт против неё:

- тех, кто по каким-то причинам не выдерживает интенсивного ритма французских лицеев;
- тех, кто хочет заниматься наукой и не хочет становиться банкиром (во Франции есть и такие), но не поступил в Высшую Нормальную школу;
- тех, кто хочет стать преподавателем средних школ и доучиться до третьего или четвёртого курса университета;
- тех, кто получил на Баке высокий балл, но кого не взяли в Подготовительные классы. В основном это школьники, набравшие высокий балл по всем предметам кроме математики;

– ну и, наконец, это студенты, приезжающие учиться во Францию из других стран.

Поступить по общему конкурсу в лучшие Высшие школы для иностранца, не владеющего в необходимом совершенстве французским языком, невозможно. Один только письменный экзамен по французской философии остановит любого на полпути. Поблажек не делают никому. Иностранцам поступать в Высшую Нормальную школу надо по университетскому обмену из лучшего вуза своей страны, или приезжать заблаговременно и лет пять учиться во французской школе. Или же поступать по упрощённому конкурсу «для иностранцев». При этом вы откажетесь от многих привилегий, положенных студентам Высшей школы, и будете её выпускником с обязательными оговорками. Что в реальности может затруднить вашу дальнейшую интеграцию в систему французских закрытых сообществ, образующих профессиональную элиту.

Таким образом, в отличие от Высших школ, публика в университетах очень разношёрстная. Особенно, на младших курсах. Профессора же в университетах очень высокого уровня. Подобная ситуация создаёт ту общую атмосферу возмущения, которую можно часто наблюдать в разговорах с коллегами из французских университетов. В основном, в группах только несколько человек интересуются математикой и из них часто большая часть — иностранцы.

---

С. Тищенко, экономический факультет МГУ

# Алгебра геометрических построений

## Построения циркулем и линейкой

А. Г. Хованский

Древние греки решили много красивых задач на построение циркулем и линейкой и обнаружили несколько проблем, которые стали чрезвычайно знаменитыми из-за многочисленных безуспешных попыток их решения, предпринимавшихся на протяжении многих веков. Древние греки построили правильные $n$-угольники для $n = 2^k \cdot 3$, $2^k \cdot 4$, $2^k \cdot 5$, $2^k \cdot 15$, где $k$ — любое неотрицательное число. Построить правильный $n$-угольник для какого-либо другого $n$ никому не удавалось до тех пор, пока Гаусс не построил правильный 17-угольник и не описал полностью числа $n$, для которых задача построения разрешима. Гаусс получил этот замечательный результат еще до возникновения теории Галуа. Его удивительное открытие оказало огромное влияние на развитие ряда областей математики. Здесь мы рассказываем об этой задаче и о других задачах на построение.

Задачи на построение — самые старые задачи о «разрешимости в явном виде». Мы различаем три класса построений. В первом, самом простом, классе допускается лишь построение прямой, проходящей через две заданные точки, окружности данного радиуса с данным центром и точки пересечения данных прямых и окружностей.

В третьем, самом сложном, классе кроме перечисленных построений допускается выбор произвольных точек, но требуется, чтобы результат построения не зависел от сделанных произвольных выборов.

Во втором, промежуточном, классе произвольный выбор не допускается, но разрешаются два построения, осуществляемые при помощи этой операции: построение центра данной окружности и построение перпендикуляра к данной прямой, проходящего через данную точку, не лежащую на прямой.

Логически построения третьего класса отличаются от построений остальных классов: промежуточные объекты, получающиеся в процессе построения, могут зависеть от произвольного выбора. В этом случае они не считаются построенными циркулем и линейкой. Мы стараемся обойтись, когда это возможно, без операции выбора произвольной точки. Мы доказываем, что использование этой операции в большинстве задач на построение не добавляет ничего нового (все, что строится при помощи этой операции, можно построить и без нее). Часть классических задач на построение вполне удовлетворительно формулируется и решается внутри первого класса построений.

Задача о трисекции угла нуждается в операции выбора произвольной точки: по двум данным прямым, проходящим через данную точку, остальными операциями вообще нельзя построить ничего нового. Мы показываем, что если к двум данным прямым добавить произвольно выбранную точку на одной из них, то по этим новым данным операциями первого класса можно построить все, что строится операциями третьего класса по двум данным прямым.

Второй класс построений нужен, чтобы расширить исходное множество начальных данных. Например, операции первого класса построений не позволяют построить ничего нового, если множество начальных данных — несколько непересекающихся окружностей. Но если к окружностям добавить их центры, то по этим новым данным операциями первого класса можно построить все, что строится операциями третьего класса по исходным данным (если среди окружностей есть две с разными центрами).

В первом параграфе мы обсуждаем задачу о разрешимости алгебраических уравнений при помощи квадратных корней, нужную для задач на построение. Мы делаем это даже в большей общности, чем необходимо (мы не предполагаем, что основное поле совершенно и что его характеристика не равна двум). Эта задача интересна сама по себе, а лишняя общность не добавляет больших хлопот. Второй параграф посвящен задачам на построение.

## §1. Разрешимость уравнений в квадратных корнях

В этом параграфе мы обсуждаем следующий вопрос о разрешимости уравнений в конечном виде: когда неприводимое алгебраическое уравнение над полем $K$ решается при помощи арифметических операций и извлечения корней степени два? Теория Галуа отвечает на этот вопрос, если поле $K$ совершенно и его характеристика не равна двум. Простые дополнительные рассуждения позволяют ответить на этот вопрос, не делая никаких предположений о поле $K$.

О расположении материала. В п. 1.1 собран нужный вспомогательный материал. В п. 1.2 приводится необходимое и достаточное условие разрешимости уравнения при помощи квадратных корней, если характеристика поля $K$ не равна двум. В п. 1.3 тот же вопрос решается для полей характеристики два. В пп. 1.4–1.5 приводятся результаты Гаусса о корнях степени $n$ из единицы (нужные для задачи о правильном $n$-угольнике).

## 1.1. Вспомогательный материал

Напомним несколько элементарных утверждений. Если поле $K$ содержится в поле $F$, то $F$ является векторным пространством над $K$. Если $\dim_K F < \infty$, то поле $F$ называется *конечным расширением* поля $K$, размерность $\dim_K F$ называется *степенью* расширения и обозначается символом $[F : K]$.

Теорема 1. *Если $K \subset F$ и $F \subset M$ — конечные расширения, то*:
1) *$K \subset M$ — конечное расширение*;
2) $[M : K] = [M : F][F : K]$.

Доказательство. Пусть $u_1, \dots, u_n$ — базис $F$ над $K$ и $v_1, \dots, v_m$ — базис $M$ над $F$. Легко видеть, что элементы $u_i v_j$, где $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, образуют базис $M$ над $K$. Откуда вытекают оба утверждения теоремы.

Утверждение 2. 1) *Если $[F : K] = n$ и $a \in F$, то существует полином $Q$ над полем $K$ степени, не большей чем $n$, такой, что $Q(a) = 0$.*

2) *Если $Q$ — неприводимый над $K$ полином и $Q(a) = 0$, то* $[K(a) : K] = \deg Q$.

Доказательство. 1) Так как $\dim_K F = n$, то элементы $1$, $a$, $\dots$, $a^n$ зависимы, т. е. найдутся $\lambda_i \in K$, такие, что $\lambda_n a^n + \lambda_{n-1} a^{n-1} + \dots + \lambda_0 = 0$, причем для некоторого $i > 0$ коэффициент $\lambda_i$ не равен нулю.

2) Поле $K(a)$ изоморфно полю $K[x]/I$, где $I$ — идеал, порожденный полиномом $Q$ степени $n$.

Утверждение 3. *Если степень неприводимого полинома над полем характеристики $p > 0$ не делится на $p$, то все его корни некратны.*

Доказательство. Кратный корень полинома $Q$ является корнем его производной $Q'$. Неприводимый полином $Q$ не может иметь общего корня с ненулевым полиномом меньшей степени. Поэтому если $Q$ имеет кратный корень, то $Q' \equiv 0$, т. е. $Q(x) = R(x^p)$, где $R$ некоторый полином. В этом случае $\deg Q = p \cdot \deg R$ и, следовательно, $\deg Q$ делится на $p$.

### 1.2. 2-радикальные расширения.

Вернемся к вопросу о разрешимости уравнений при помощи квадратных корней.

Определение. Расширение $K \subset F$ называется *2-радикальным*, если существует башня полей $K = F_0 \subset F_1 \subset \cdots \subset F_n$, такая, что $F \subset F_n$ и $F_i = F_{i-1}(a_i)$ при $1 \leqslant i \leqslant n$, где $a_i^2 \in F_{i-1}$ и $a_i \notin F_{i-1}$.

Теорема 4. *Если $K \subset F$ — 2-радикальное расширение, то $[F : K] = 2^k$.*

Доказательство. Для $K = F_0 \subset F_1 \subset \cdots \subset F_n$ имеем $[F_n : K] = [F_n : F_{n-1}] \cdot \ldots \cdot [F_1 : F_0] = 2^n$. Если $K \subset F \subset F_n$, то $[F : K] \cdot [F_n : F] = 2^n$. Поэтому $[F : K]$ — степень двойки.

Следствие 5. *Если неприводимый над $K$ полином $P$ имеет корень в некотором 2-радикальном расширении поля $K$, то $\deg P = 2^k$.*

Доказательство. Если $P(a) = 0$, то $[K(a) : K] = \deg P$.

Следствие 6. *Пусть характеристика поля $K$ не равна двум. Кубическое уравнение $P = 0$ над полем $K$ решается в квадратных корнях, если и только если один из корней уравнения содержится в поле $K$.*

Доказательство. Если $a \in K$ и $P(a) = 0$, то $P = (x - a)Q$, где $Q \in K[x]$. Квадратное уравнение $Q = 0$ решается в квадратных корнях, так как характеристика поля $K$ не равна двум. Если кубический полином не имеет корня в $K$, то он неприводим и можно воспользоваться следствием 5.

Замечание. Для $K = \mathbb{Q}$ следствие 6 принимает явный вид: для полинома $P \in \mathbb{Q}[x]$ можно явно найти все его рациональные корни (в частности, можно явно проверить, что таких корней нет). Если корень $a$ найден, то квадратное уравнение $0 = Q(x) = P/(x - a)$ решается явно.

Обозначим через $E_P$ поле разложения полинома $P$ над полем $K$.

Следствие 7. *Если неприводимый над $K$ полином $P$ имеет корень в некотором 2-радикальном расширении поля $K$, то $[E_P : K] = 2^m$.*

Доказательство. Расширение $K \subset E_P$ в условиях следствия 2-радикально.

Следствие 7 допускает следующее частичное обращение.

Теорема 8. *Если для неприводимого полинома $P$ над полем $K$, характеристика которого не равна двум, выполняется равенство $[E_P : K] = 2^m$, то расширение $K \subset E_P$ является 2-радикальным.*

Доказательство. Степень расширения $[E_P : K]$ делится на степень полинома $P$, поэтому $\deg P = 2^k$. Следовательно, по утверждению 3, уравнение $P = 0$ сепарабельно и к нему применима теория Галуа. Порядок группы Галуа $G$ поля $E_P$ над полем $K$ равен числу $[E_P : K] = 2^m$. Так как порядок группы $G$ является степенью двойки, то существует нормальная башня подгрупп $G = G_0 \supset G_1 \supset \cdots \supset G_m = e$, такая, что $G_i/G_{i-1} = \mathbb{Z}_2$ при $1 \leqslant i \leqslant k$. Для башни полей $K = K_0 \subset K_1 \subset \cdots \subset K_m = E_P$, соответствующей этой башне подгрупп, имеем $[K_i : K_{i-1}] = 2$. Так как характеристика поля $K$ (а, значит, и поля $K_i$) не равна двум, то поле $K_i$ получается из поля $K_{i-1}$ присоединением квадратного корня.

### 1.3. 2-радикальные расширения полей характеристики два

В этом пункте мы будем обозначать символом $K$ некоторое поле характеристики два и символом $\overline{K}$ — его алгебраическое замыкание. Нас интересуют лишь алгебраические элементы над полем $K$ и алгебраические расширения поля $K$. Не ограничивая общности, можно считать, что эти элементы и расширения содержатся в поле $\overline{K}$.

Лемма 9. *Множество элементов $y$, таких, что $y^2 \in K$, является полем.*

Доказательство. Лемма вытекает из равенства $(a + b)^2 = a^2 + b^2$, справедливого в полях характеристики два.

Определим цепочку подполей $K = K_0 \subset K_1 \subset \cdots \subset K_n \subset \ldots$ поля $\overline{K}$ при помощи соотношения $y \in K_{i+1}$, если и только если $y^2 \in K_i$. Поле $\widetilde{K} = \bigcup K_i$ будем называть *совершенным замыканием* поля $K$. Легко видеть, что поле $\widetilde{K}$ является минимальным совершенным полем, содержащим поле $K$.

Теорема 10. *Конечное расширение $K \subset M$ поля $K$ является 2-радикальным, если и только если $M \subset \widetilde{K}$.*

Доказательство. Если для башни полей $K = F_0 \subset F_1 \subset \cdots \subset F_n$ имеем $F_i = F_{i-1}(a_i)$, где $a_i^2 \in F_{i-1}$, то $F_i \subset K_i$.

Полином $P \in K[x]$ называется *минимальным полиномом* алгебраического элемента $a$ над $K$, если $P(a) = 0$, полином $P$ неприводим и унимодален (т. е. старший коэффициент полинома $P$ равен единице).

Теорема 11. *Полином $P$ является минимальным полиномом некоторого элемента $a \in K_n \setminus K_{n-1}$, если и только если $P(x) = x^{2^n} - b$, где $b \in K$ и $b \neq c^2$ для всякого $c \in K$.*

Доказательство. Элемент $a \in K_n$ является единственным (кратным) корнем полинома $x^{2^n} - b$, где $b = a^{2^n} \in K$, поэтому минимальный

полином $P$ элемента $a$ имеет единственный корень $a$. Числа $m$, такие, что $a^m \in K$, образуют аддитивную подгруппу в $\mathbb{Z}$. Если $a \in K_n \setminus K_{n-1}$, то $a^m \in K$, только если $m$ делится на $2^n$. Откуда видно, что $P(x) = x^{2^n} - b$.

Следствие 12. *Если $P \in K[x]$ — унимодальный и неприводимый над $K$, то уравнение $P(x) = 0$ разрешимо в квадратных корнях, если и только если $P(x) = x^{2^n} - b$, где $b \in K$ и $b \neq c^2$ для всякого $c \in K$. В частности, всякое неприводимое уравнение степени большей единицы над совершенным полем $K$ не решается при помощи квадратных корней.*

## 1.4. Корни из единицы

Здесь мы напоминаем классические результаты Гаусса, открытые еще до возникновения теории Галуа.

Пусть $\Omega_n \subset \mathbb{C}$ — множество чисел $x$, таких, что $x^n = 1$, и $\Omega_n^*$ — множество всех *примитивных корней из единицы* степени $n$, т. е. множество чисел $a \in \Omega_n$, таких, что $a^m \neq 1$ при $0 < m < n$. Если $\omega \in \Omega_n^*$, то: 1) $a \in \Omega_n$, если и только если $a = \omega^k$ для некоторого целого $k$; 2) $a \in \Omega^*$, если и только если $a = \omega^k$, где $k$ взаимно просто с $n$, т. е. вычет $k$ по модулю $n$ лежит в мультипликативной группе $U(n)$ обратимых элементов кольца $\mathbb{Z}/n\mathbb{Z}$. *Циклотомическим полиномом* степени $n$ называется полином $\Phi_n(x) = \prod_{a \in \Omega_n^*}(x - a)$.

Лемма 13. *Справедливо равенство $x^n - 1 = \prod_{d|n} \Phi_d(x)$, где произведение берется по всем делителям $d$ числа $n$.*

Доказательство. Вытекает из соотношений

$$\Omega_n = \bigcup_{d|n} \Omega_d^* \quad \text{и} \quad \Omega_{d_1}^* \bigcap \Omega_{d_2}^* = \varnothing \text{ при } d_1 \neq d_2.$$

Следствие 14. *Полином $\Phi_n$ унимодален и имеет целые коэффициенты.*

Доказательство. Если $P$, $Q \in \mathbb{Z}[x]$ — унимодальные полиномы, то: 1) полином $PQ$ унимодален и $PQ \in \mathbb{Z}[x]$; 2) если $T = P/Q$ — полином, то полином $T$ унимодален и $T \in \mathbb{Z}[x]$. Мы используем очевидные факты 1)–2) для индукционного доказательства следствия. Для $n = 1$ следствие верно, так как $\Phi_1(x) = x - 1$. Положим $\Psi_n = \prod \Phi_{d'}$, где произведение берется по делителям $d'$ числа $n$, меньшим чем $n$. Если следствие верно для $d' < n$, то, согласно 1), полином $\Psi_n$ унимодален и $\Psi_n \in \mathbb{Z}[x]$. По лемме 13 $\Phi_n(x) = (x^n - 1)/\Psi_n(x)$. Согласно 2), следствие верно для $d = n$.

Полином $f \in \mathbb{Z}[x]$ называется *примитивным*, если его коэффициенты не имеют общего делителя. *Произведение примитивных полиномов является примитивным полиномом.* Из этого факта и теоремы Гаусса

автоматически вытекает следующее *свойство целочисленности*: если для $f_1 = f_2 f_3$, где $f_1, f_2$ — унимодальные полиномы, $f_1 \in \mathbb{Z}[x]$ и $f_2, f_3 \in \mathbb{Q}[x]$, то полиномы $f_2, f_3$ имеют целые коэффициенты и полином $f_3$ унимодален.

Напомним также, что если $p$ взаимно просто с $n$, то полином $x^n - 1 \in \mathbb{Z}_p[x]$ не имеет кратных корней (так как $n \neq 0 \mod p$, то производная $nx^{n-1}$ полинома $x^n - 1$ не имеет ненулевых корней).

Теорема 15 (Гаусс). *Полином $\Phi_n$ неприводим над $\mathbb{Z}$.*

Лемма 16 (Гаусс). *Пусть $\omega \in \Omega_n^*$, $f$ — минимальный полином числа $\omega$ и $p$ — простое число, взаимно простое с $n$. Тогда $f(\omega^p) = 0$.*

Доказательство. Число $\omega$ — корень полинома $\Phi_n$. Поэтому $\Phi_n = fg$, где $g \in \mathbb{Q}[x]$. Согласно свойству целочисленности, $g \in \mathbb{Z}[x]$ и полином $g$ унимодален. Допустим, что $f(\omega^p) \neq 0$. Тогда $g(\omega^p) = 0$, так как $0 = \Phi_n(\omega^p) = f(\omega^p)g(\omega^p)$. В этом случае $\omega$ — корень полинома $g(x^p)$, поэтому $g(x^p) = fh$, где $h \in \mathbb{Q}[x]$. Согласно свойству целочисленности, $h \in \mathbb{Z}[x]$ и полином $h$ унимодален. Пусть $\pi\colon \mathbb{Z}[x] \to \mathbb{Z}_p[x]$ — гомоморфизм, продолжающий на кольцо полиномов естественное отображение $\mathbb{Z} \to \mathbb{Z}_p$. Имеем $(\pi g)(x^p) = (\pi f)(\pi h)$. В кольце $\mathbb{Z}_p[x]$ для всякого полинома $\varphi$ справедливо тождество $\varphi(x^p) = \varphi^p(x)$ (оно вытекает из тождества $a^p = a$ в $\mathbb{Z}_p$ и из тождества $(\varphi + \psi)^p = \varphi^p + \psi^p$ в $\mathbb{Z}_p[x]$), поэтому $(\pi g)^p = (\pi f)(\pi h)$. Следовательно, полиномы $\pi(g)$ и $\pi(f)$ имеют общий множитель, поэтому полином $\pi(f)\pi(g)$ имеет кратный корень. Но полином $\pi(\Phi_n) = \pi(f)\pi(g)$ является делителем полинома $\pi(x^n - 1) = (x^n - 1) \in \mathbb{Z}_p[x]$, который не имеет кратных корней. Противоречие доказывает лемму Гаусса.

Доказательство теоремы Гаусса. Пусть $(\omega,\ f,\ p)$ такие, как в лемме Гаусса. Для тройки $(\omega_1,\ f,\ p_2)$, где $\omega_1 = \omega^{p_1}$, $p_1 = p$ и $p_2$ любое простое число, взаимно простое с $n$, применима лемма Гаусса. Действительно, $\omega_1 \in \Omega_n^*$ и $f(\omega_1) = 0$. Аналогично $f(\omega^{p_1 \cdot \ldots \cdot p_m}) = 0$ для любой последовательности простых чисел $p_1, \ldots, p_m$, взаимно простых с $n$. Каждый элемент $\alpha \in \Omega_n^*$ представим в виде $\alpha = \omega^m$, где $m$ — произведение простых чисел, взаимно простых с $n$. Унимодальный полином $\Phi_n$ имеет те же корни, что и унимодальный полином $f$. Поэтому $\Phi_n = f$ и полином $\Phi_n$ неприводим.

Следствие 17. *Группа Галуа $G$ поля $E_n$ разложения полинома $x^n - 1$ над полем $\mathbb{Q}$ изоморфна мультипликативной группе $U(n)$ кольца $\mathbb{Z}/n\mathbb{Z}$.*

Доказательство. Легко видеть, что группа $G$ является подгруппой группы $U(n)$. Корни неприводимого полинома $\Phi_n$ лежат в поле разложения полинома $x^n - 1 = 0$, поэтому $\#G \geqslant \deg \Phi_n$. Но $\deg \Phi_n = \#U(n)$. Поэтому группа $G$ совпадает с группой $U(n)$.

Замечание. Пусть $\mathbb{Q} \subset E$ расширение Галуа и пусть $E \subset E_n$. Тогда группа Галуа $G$ расширения $\mathbb{Q} \subset E$ коммутативна, так как $G$ — факторгруппа группы $U(n)$. Согласно знаменитой теореме Кронекера – Вебера, верно и обратное утверждение: *если группа Галуа расширения $\mathbb{Q} \subset E$ коммутативна, то $E$ содержится в поле $E_n$ при некотором $n$.*

### 1.5. Разрешимость уравнения $x^n - 1 = 0$

Здесь описываются числа $n$, для которых 2-радикально расширение $\mathbb{Q} \subset E_n$.

Утверждение 18. *Пусть $n = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m}$ — разложение $n$ на простые множители. Тогда $U(n) = U(p_1)^{k_1} \times \cdots \times U(p_m)^{k_m}$ и $\#U(n) = \prod(p_i^k - p^{k_i - 1})$.*

Доказательство. Вытекает из китайской теоремы об остатках и из того, что в кольце $\mathbb{Z}/p^k\mathbb{Z}$ есть ровно $p^{k-1}$ необратимых элементов.

Простое число $p$ называется *простым числом Ферма*, если $p = 2^n + 1$. Для нечетного $m$ число $2^{km} + 1$ делится на $2^k + 1$ и является составным. Поэтому простые числа Ферма представимы в виде $p = 2^{2^q} + 1$. Числа 3, 5, 17, 257, 65537 доставляют примеры простых чисел Ферма. Неизвестно, является ли множество простых чисел Ферма бесконечным. Целое число $n$ будем называть *числом Гаусса*, если $n = 2^k p_1 \cdot \ldots \cdot p_m$, где $k \geqslant 0$, а $p_1, \ldots, p_m$ — различные простые числа Ферма.

Теорема 19 (Гаусс). *Расширение $\mathbb{Q} \subset E_n$ 2-радикально, если и только если $n$ — число Гаусса.*

Доказательство. Действительно, $\deg \Phi(n) = \#U(n)$. Из утверждения 18 видно, что $\#U(n) = 2^k$, если и только если число $n$ — число Гаусса.

Пример. Решим уравнение $\Phi_5(x) = 0$. Имеем

$$\Phi_5(x) = (x^5 - 1)/(x - 1) = x^4 + x^3 + x^2 + x + 1 = 0.$$

Далее,

$$x^{-2}\Phi_5(x) = x^2 + x + 1 + x^{-1} + x^{-2} = u^2 + u - 1,$$

где $u = x + x^{-1}$. Чтобы найти $x$, достаточно сначала решить квадратное уравнение $u^2 + u - 1 = 0$ и затем решить квадратное уравнение $xu = x^2 + 1$.

Явное решение уравнения $\Phi_{17}(x) = 0$ было найдено Гауссом. Оно послужило отправной точкой и для других его замечательных открытий. И сейчас, владея теорией Галуа, самостоятельно решить это уравнение далеко не просто. Моим студентам Ю. Бурда и Л. Кадец это удалось (см. [1]).

## §2. Что можно построить циркулем и линейкой?

Этот параграф посвящен вопросам разрешимости и неразрешимости задач на построение. Несколько слов о расположении материала.

В пп. 2.1–2.2 описан класс точек, прямых и окружностей, которые могут быть построены при помощи операций из первого класса построений по начальным данным, являющимся конечным множеством точек: в п. 2.1 даны необходимые условия принадлежности этому классу, в п. 2.2 достаточные. В п. 2.3 мы обсуждаем несколько классических задач на построение (включая задачу о построении правильного $n$-угольника), которые укладываются в картину, разобранную в пп. 2.1–2.2.

В п. 2.4 выделены два построения, использующие выбор произвольных точек, которые позже рассматриваются как две новые операции. В п. 2.6 описано, что можно построить по любым (кроме нескольких исключительных типов) начальным данным с использованием операции выбора произвольных точек. Оказывается, что все, что можно построить с ее использованием, можно построить и без нее, пользуясь двумя новыми операциями и построениями из пп. 2.1–2.2.

В п. 2.7 мы описываем, что можно построить по множеству начальных данных одного исключительного типа, связанного с задачей о трисекции угла, и подробно обсуждаем вопрос о разрешимости этой задачи.

В п. 2.8 доказана одна теорема из вещественной аффинной геометрии, связанная с выполнимостью арифметических операций над вещественными числами при помощи геометрических построений.

### 2.1. Неразрешимость некоторых задач на построение

Прежде чем доказывать невозможность того или иного построения, нужно точно определить, что это такое. Пусть $M$ — множество всех точек, прямых и окружностей на плоскости (только такие объекты можно строить циркулем и линейкой). Можно задать некоторый *допустимый класс* $\mathcal{M} \subset M$ и сказать, что точка, прямая или окружность могут быть построены, если они принадлежат этому классу. Класс $\mathcal{M}$ можно определить, задав начальные данные и допустимые операции.

Список допустимых операций (начало списка)

1. *Операция построения прямой* сопоставляет паре различных точек проходящую через них прямую.

2. *Операция построения окружности* сопоставляет точкам $P$, $Q$, $O$, где $P \neq Q$, окружность с центром $O$ и радиусом, равным $[P, Q]$.

3. *Операция пересечения* сопоставляет паре несовпадающих пересекающихся кривых $\Gamma_1$, $\Gamma_2$, где $\Gamma_i$ — либо прямая, либо окружность,

их точки пересечения (пересечение может содержать одну или две точки).

Определение. Класс $\mathcal{M}(\mathcal{D}) \subset M$ точек, прямых и окружностей, которые *строятся по начальным данным* $\mathcal{D} \subset M$, — это минимальный класс, содержащий $\mathcal{D}$ и замкнутый относительно допустимых операций 1)–3).

Замкнутость класса $\mathcal{M}(\mathcal{D})$ относительно пересечения означает, что если $\Gamma_1$, $\Gamma_2 \in \mathcal{M}(\mathcal{D})$, кривые $\Gamma_1$, $\Gamma_2$ не совпадают и $P \in \Gamma_1 \cap \Gamma_2$, то $P \in \mathcal{M}(\mathcal{D})$. Аналогично определяется замкнутость относительно других операций.

Теоремы о невозможности тех или иных построений основаны на формулируемой ниже простой алгебраической теореме 20.

Определение. *Класс* $\mathcal{M}_T$ — это класс всех точек, прямых и окружностей на координатной плоскости, определенных *над некоторым вещественным полем* $T$ (точка определена над $T$, если обе ее координаты лежат в $T$, прямая или окружность определены над $T$, если их можно задать уравнениями $ax + by + c = 0$ или $(x-a)^2 + (y-b)^2 + c = 0$, где $a, b, c \in T$).

Если *вещественное поле* $T \subset \mathbb{R}$ *замкнуто относительно извлечения квадратных корней*, т. е. если $a \in \mathbb{R}$ и $a^2 \in T$, то $a \in T$, то класс $\mathcal{M}_T$ вместе с каждой окружностью содержит ее центр и расстояние между точками $P, Q \in \mathcal{M}_T$ лежит в поле $T$.

Теорема 20. *Если вещественное поле* $T \subset \mathbb{R}$ *замкнуто относительно извлечения квадратных корней, то класс* $\mathcal{M}_T$ *замкнут относительно допустимых операций* 1)–3).

Доказательство. В координатах плоскости $\mathbb{R}^2$ операции 1)–3) сводятся к нахождению вещественных решений линейных и квадратных уравнений. Решение таких уравнений не выводит из поля $T$, так как оно замкнуто относительно извлечения вещественных квадратных корней.

Пусть $\mathcal{D}_0$ — некоторое множество точек на плоскости, содержащее не менее двух точек. Евклидовы движения и гомотетии переводят прямую в прямую, окружность с отмеченным центром в окружность с отмеченным центром. Такие движения согласованы с задачами построения.

Определение. *Полем, соответствующим* $\mathcal{D}_0$, назовем наименьшее вещественное поле $T(\mathcal{D}_0)$, замкнутое относительно извлечения квадратных корней и содержащее отношения длин отрезков, концы которых лежат в $\mathcal{D}_0$.

Выберем две разные точки $O$, $E \in \mathcal{D}_0$ и нормируем расстояние так, чтобы длина отрезка $[O, E]$ равнялась единице. Скажем, что ортонормированная система координат *согласована с* $\mathcal{D}_0$, если $O = (0, 0)$ и $E = (0, 1)$.

Теорема 21. *В согласованной с $\mathcal{D}_0$ системе координат справедливо включение $\mathcal{M}(\mathcal{D}_0) \subset \mathcal{M}_T$, где $T = T(\mathcal{D}_0)$.*

Другими словами, если точка, прямая или окружность не определены над полем $T$, то их нельзя построить при помощи операций 1)–3) по точкам из множества $\mathcal{D}_0$.

Доказательство. В условиях теоремы координаты точек множества $\mathcal{D}_0$ лежат в поле $T(\mathcal{D}_0)$. Теперь теорема вытекает из теоремы 20.

## 2.2. Несколько явных построений

Для выполнения тех или иных построений в качестве «строительных кирпичиков» нужны решения нескольких школьных задач на построения. Напомним их.

1) *По точкам $A, B$ построить точку, лежащую вне прямой $AB$, и середину $P$ отрезка $[A, B]$.* Пусть $Q$, $R$ — точки пересечения окружностей с центрами $A$ и $B$ и радиусами, равными $[A, B]$. Каждая из точек $Q, R$ лежит вне $AB$, и $P$ — точка пересечения $AB$ и $QR$.

2) *Восстановить перпендикуляр к прямой $l$ из точки $P \in L$.* В нашей ситуации требуется по точкам $A, P \in l$ построить точки $Q$, $R$, такие, что прямые $AP$ и $QR$ перпендикулярны и $P \in AQ$. Пусть $B \neq A$ — точка пересечения прямой $AP$ с окружностью с центром $P$ и с радиусом, равным $[P, A]$. В качестве искомых точек можно взять точки $Q, R$ из предыдущего построения.

3) *На прямую $l$ опустить перпендикуляр из точки $P \notin l$.* В нашей ситуации требуется по трем точкам $A$, $B$, $P$, не лежащим на одной прямой, построить точку $Q$, такую, что прямые $AB$ и $PQ$ перпендикулярны. В качестве $Q$ можно взять точку пересечения $Q \neq P$ окружностей с центрами $A$ и $B$, проходящими через $P$.

4) *построить прямую $l_1$, параллельную прямой $l$ и проходящую через точку $P \notin l$.* Достаточно из точки $P$ опустить перпендикуляр $l_2$ к прямой $l$ и затем восстановить перпендикуляр $l_1$ к прямой $l_2$.

Пусть $O \neq E$ — две точки. Рассмотрим систему координат на плоскости, согласованную с множеством $\mathcal{D}_0 = \{O, E\}$. Первая координата на плоскости задает координату на прямой $l_0 = OE$. Отождествим точку на $l_0$ с числом, равным ее координате. При этом $O$ и $E$ отождествятся с 0 и 1.

Лемма 22. *Пусть $a, b \in l_0 \cap \mathcal{M}(\mathcal{D}_0)$. Тогда*: 1) *$-a$, $a^{-1}$, $a + b$, $ab \in l_0 \cap \mathcal{M}(M_0)$;* 2) *если $ab > 0$, то $(ab)^{1/2} \in l_0 \cap \mathcal{M}(\mathcal{D}_0)$.*

Ограничимся картинками, поясняющими доказательство, см. рис. 1.

Замечание. При построении прямой, параллельной данной и проходящей через данную точку, мы пользовались циркулем и линейкой. Это построение можно рассматривать как единую операцию. Такой операции

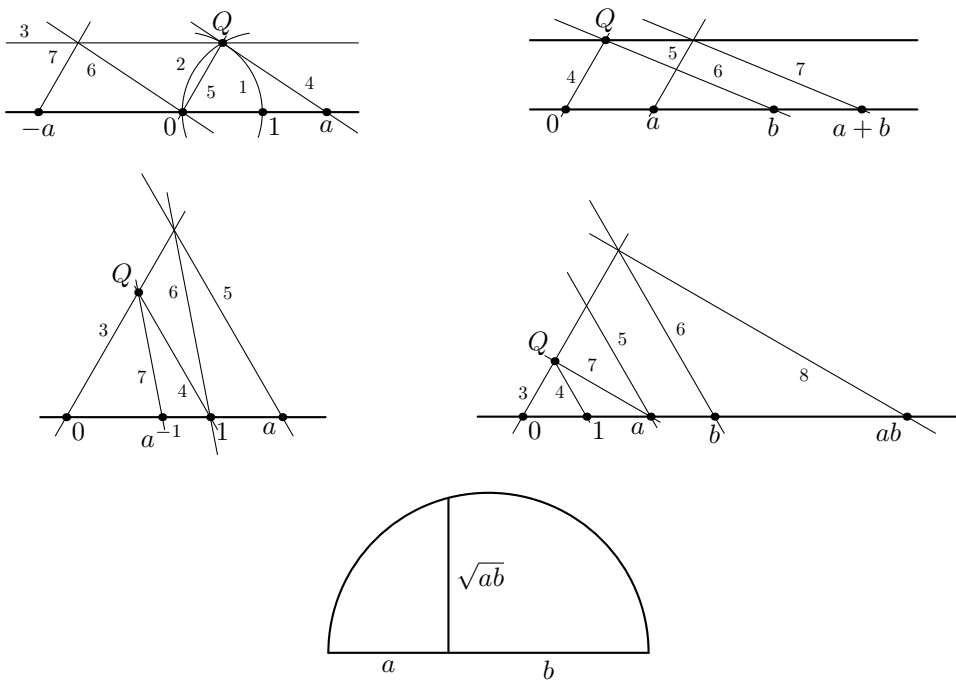



*Рис. 1. Линии занумерованы в порядке построения*

достаточно, чтобы, имея точки 0 и 1, построить точки $-a$, $a^{-1}$, $a+b$, $ab$ по точкам $a$, $b$ на числовой прямой (смотри рис. 1, точка $Q$ в этом случае выбирается произвольно). Этот факт имеет красивое применение в аффинной геометрии (см. п. 2.8).

Теорема 23. *В условиях теоремы* 21 *справедливо равенство* $\mathcal{M}(\mathcal{D}_0) = \mathcal{M}_T$.

Доказательство. Достаточно показать, что $\mathcal{M}(\mathcal{D}_0) \supset \mathcal{M}_T$, т. е. что можно построить любой элемент из $\mathcal{M}_T$. Если $P$, $Q \in \mathcal{D}_0$, то класс $\mathcal{M}(\mathcal{D}_0)$ содержит точку $\rho \in l_0$, где $\rho$ — отношение длин отрезков $[P,Q]$ и $[O,E]$: это одна из точек пересечения прямой $l_0$ с окружностью с центром $O$ и радиусом, равным $[P,Q]$. Согласно лемме 22, каждая точка $(a,0)$, где $a \in T$, лежит в $\mathcal{M}(\mathcal{D}_0)$. Точка $(0,b)$, где $b \in T$, тоже лежит в $\mathcal{M}(\mathcal{D}_0)$: ее можно построить, пересекая ось $y$ с окружностью с центром $O$, проходящую через точку $(b,0)$. Строя прямые, перпендикулярные к осям, убеждаемся, что при $a, b \in T$ точка $(a,b)$ лежит в $\mathcal{M}(\mathcal{D}_0)$. Прямая $l$, определенная над $T$, содержит пару точек, определенных над $T$, поэтому $l \in \mathcal{M}(D_0)$. Окружность $S$, определенная над $T$, содержит точку, определенную над $T$. Ее центр тоже определен над $T$, поэтому $S \in \mathcal{M}(D_0)$.

### 2.3. Классические задачи на построение

В нескольких классических задачах на построение начальным данным является отрезок, или, что то же самое, пара его концов $O$, $E$. В этом пункте символом $T$ мы будем обозначать *поле конструируемых чисел*, соответствующее множеству $\mathcal{D}_0 = \{O, E\}$. По теореме 23 в системе координат, согласованной с $\mathcal{D}_0$, справедливо равенство $\mathcal{M}(\mathcal{D}_0) = \mathcal{M}_T$.

Квадратура круга. *По точкам $O$, $E$ построить отрезок $I$ такой, что площадь круга радиуса $[OE]$ равна площади квадрата со стороной $I$.*

Теорема 24. *Для любых точек $P, Q \in \mathcal{M}_T$ отрезок $[PQ]$ не равен отрезку $I$. Другими словами, в классе $\mathcal{M}_T$ квадратура круга не осуществима.*

Доказательство. Расстояние между точками $P, Q \in \mathcal{M}_T$ — конструируемое число, а длина отрезка $I$ — трансцендентное число $\pi^{1/2}$.

Удвоение куба. *По точкам $O$, $E$ построить отрезок $J$ такой, что объем куба со стороной $J$ в два раза больше объема куба со стороной $I$.*

Теорема 25. *Для любых точек $P, Q \in \mathcal{M}_T$ отрезок $[PQ]$ не равен отрезку $J$. Другими словами, в классе $\mathcal{M}_T$ удвоение куба не осуществимо.*

Доказательство. Расстояние между точками $P, Q \in \mathcal{M}_T$ — конструируемое число, а длина отрезка $J$ равна $2^{1/3}$. Уравнение $x^3 - 2 = 0$ неприводимо над $\mathbb{Q}$ и не решается при помощи квадратных корней.

Задача о правильном $n$-угольнике. *Построить правильный $n$-угольник с данной стороной $[OE]$.*

Теорема 26 (Гаусс). *Правильный $n$-угольник можно построить* (т. е. его вершины лежат в классе $\mathcal{M}_T$), *если и только если $n$ — число Гаусса.*

Доказательство. Несложно видеть, что задача эквивалентна следующей: построить вершины правильного $n$-угольника с центром $O$, одна из вершин которого — точка $E$. При отождествлении плоскости с комплексной прямой вершины этого $n$-угольника — корни уравнения $z^n = 1$. Это уравнение решается при помощи квадратных корней (в поле комплексных чисел), если и только если $n$ — число Гаусса. Осталось заметить, что комплексное число выражается над полем $\mathbb{Q}$ при помощи квадратных корней, если и только если его вещественная и мнимая части — конструируемые числа.

### 2.4. Два специальных построения

Выделим два простых построения, использующие выбор произвольной точки из континуального множества. Эти построения невозможно

осуществить при помощи операций 1)–3), и их можно рассматривать как новые операции (что позже мы и будем делать).

Задача 1. *По прямой $l$ и точке $P \notin l$ найти точку $E \in l$, такую, что прямые $l$ и $EP$ перпендикулярны.*

Класс $\mathcal{M}(\mathcal{D})$, где множество начальных данных $\mathcal{D}$ состоит из прямой $l$ и точки $P \notin l$, совпадает с множеством $\mathcal{D}$: применение допустимых операций не увеличивает множества $\mathcal{D}$. Однако если на прямой $l$ *произвольным образом* выбрать две различные точки $A$, $B$, то такое построение легко выполнить (см. п. 2.2). Его результатом являются перпендикуляр $l_P$ и точка $E = l \cap l_P$, которые *не зависят от сделанного произвольного выбора*.

По точкам $O = P$ и $E$ можно построить все объекты класса $\mathcal{M}_T$. Результат каждого из этих построений не зависит от выбранных произвольно точек $A$, $B \in l$, $A \neq B$, которые использовались в построении.

Задача 2. *Построить центр данной окружности $S$.*

Класс $\mathcal{M}(\mathcal{D})$, где $\mathcal{D} = \{S\}$, совпадает с множеством $\mathcal{D}$. Однако если выбрать две разные точки $A$, $B \in S$, то перпендикуляр к прямой $AB$, делящий отрезок $[A, B]$ пополам, проходит через центр окружности. Находя середину построенного диаметра, получим центр окружности $O$.

Чтобы включить такие конструкции в нашу схему, нужно допустить выбор *произвольных точек*, лежащих в одном из множеств (стратов), на которые разбивается плоскость уже построенными точками, прямыми и окружностями. Но при этом считать построенными лишь объекты, которые не зависят от сделанных произвольных выборов. Ниже мы покажем, что такая расширенная интерпретация процесса построения циркулем и линейкой не меняет уже полученных результатов и позволяет рассматривать и другие задачи, в частности, задачу о трисекции угла.

Начнем с рассмотрения стратификации плоскости, связанной с конечным подмножеством множества $M$ всех точек, прямых и окружностей.

## 2.5. Стратификация плоскости

Пусть $V \subset M$ — конечное подмножество. С $V$ связана *стратификация* $\Sigma_V$ плоскости, т. е. ее разбиение на *страты* $S_\alpha \in \Sigma_V$ разных размерностей. (Если точки, прямые и окружности из $V$ нарисованы, то стратификация $\Sigma_V$ видна на картинке.)

*Нульмерный страт* в $\Sigma_V$ — любая точка множества $V_0$ всех точек пересечения различных прямых и окружностей из $V$ и любое из одноточечных множеств, содержащихся в $V$.

*Одномерный страт* в $\Sigma_V$ — любая компонента связности множества $\Gamma_i \setminus (\Gamma_i \cap V_0)$, где $\Gamma_i$ — любая прямая или окружность из $V$.

*Двумерный страт* в $\Sigma_V$ — любая компонента связности дополнения плоскости к объединению всех точек, прямых и окружностей из $V$.

Пусть $T$ — вещественное поле, замкнутое относительно извлечения квадратных корней. Из теоремы 20 вытекает следующее следствие.

Следствие 27. *Если $V \subset \mathcal{M}_T$ и $P$ — нульмерный страт в $\Sigma_V$, то $P \subset \mathcal{M}_T$.*

Утверждение 28. *Точки, определенные над $T$, плотны на*: 1) *плоскости*; 2) *прямой, определенной над $T$*; 3) *окружности, определенной над $T$.*

Доказательство. Пункты 1) и 2) очевидны. Прямые вида $y = c$, где $c \in T$, плотны в $\mathbb{R}^2$. Они пересекают окружность, определенную над $T$, в точках, определенных над $T$. Множество таких точек всюду плотно на окружности.

Следствие 29. *Если $V \subset \mathcal{M}_T$, то точки, определенные над $T$, плотны в каждом страте положительной размерности стратификации $\Sigma_V$.*

### 2.6. Классы построений, допускающие произвольный выбор

Результат применения операции пересечения зависит от выбора одной из точек пересечения двух кривых. Определим операцию 4), зависящую не только от дискретного, но и от континуального выбора. При ее помощи можно выполнить два простых построения (см. п. 2.4), которые можно принять за новые операции 5) и 6).

Продолжение списка допустимых операций

4. *Операция выбора точки* применима к конечному множеству $V \subset M$ и заключается в выборе страта $S_\alpha \in \Sigma_V$ положительной размерности и точки $P$ из этого страта.

5. *Операция построения основания перпендикуляра* сопоставляет прямой $l$ и точке $P \notin l$ точку $E \in l$ такую, что прямые $EP$ и $l$ перпендикулярны.

6. *Операция восстановления центра* сопоставляет окружности ее центр.

Определим класс $\mathcal{M}_G(\mathcal{D})$ элементов, которые можно *в обобщенном смысле построить* по конечному множеству $\mathcal{D}$. Скажем, что $v \in \mathcal{M}_G(\mathcal{D})$, если существует конечный алгоритм (т. е. правило, описывающее все дискретные выборы), $k$-й шаг которого — переход от одного конечного множества $V_k \subset M$ к следующему $V_{k+1} \subset M$. При этом: 1) $V_1 = \mathcal{D}$; 2) $V_{k+1} = V_k \cup \{a\}$, где $a$ или получается применением к некоторым элементам множества $V_k$ одной из операций 1)–3) (см. п. 2.1), или $a = P$ и точка $P$

получена при помощи операции выбора из множества $V_k$; 3) элемент $v$ содержится в некотором из множеств $V_N$ вне зависимости от континуальных выборов, которые встречались на предыдущих шагах.

Теорема 30. *Пусть $T$ — вещественное поле, замкнутое относительно извлечения квадратных корней и $\mathcal{D} \subset \mathcal{M}_T$ — конечное множество. Тогда $\mathcal{M}_G(\mathcal{D}) \subset \mathcal{M}_T$.*

Доказательство. Если $v \in \mathcal{M}_G(\mathcal{D})$, то континуальным выбором, встречающимся в процессе построения $v$, можно распорядиться по своему усмотрению. По условию $V_1 = \mathcal{D} \subset \mathcal{M}_T$ и $V_2 = V_1 \bigcup \{a\}$. Если первый шаг состоит в присоединении точки $a$ из страта положительной размерности в стратификации $\Sigma_{V_1}$, то выберем точку $a$, определенную над $T$. По следствию 29 это можно сделать. При таком выборе $V_2 \subset \mathcal{M}_T$. Если точка, прямая или окружность $a$ получаются применением к некоторым элементам множества $V_1$ одной из операций 1)–3), то $V_2 \subset \mathcal{M}_T$ по теореме 20. Будем последовательно на каждом шаге построения, состоящем в присоединении произвольно выбранной точки, выбирать точку, определенную над $T$. При применении это правила выбора $V_k \subset \mathcal{M}_T$ для любого $k > 0$.

Следствие 31. *Если $\mathcal{D}_0$ — конечное множество точек, содержащее не менее двух точек, то $\mathcal{M}_G(\mathcal{D}_0) = \mathcal{M}(\mathcal{D}_0)$. В частности, операция 4) не помогает решить задачи о квадратуре круга и об удвоения куба. С ее помощью можно построить только те правильные $n$-угольники, которые строятся и без нее.*

Определение. Минимальный класс $\mathcal{M}_r(\mathcal{D})$, содержащий $\mathcal{D}$ и замкнутый относительно операций 1)–3) и 5), 6), будем называть классом объектов, которые *в расширенном смысле строятся по начальным данным $\mathcal{D}$*.

Скажем, что $D$ — *исключительное множество типа $R_i$*, если $D$ содержит:

для типа $R_1$ — *единственную точку*;
для типа $R_2$ — *единственную прямую*;
для типа $R_3$ — $k > 1$ *параллельных прямых*;
для типа $R_4$ — $k > 0$ *прямых, проходящих через точку $O$*;
для типа $R_5$ — $k > 0$ *прямых, проходящих через точку $O$, и точку $O$*;
для типа $R_6$ — $k > 0$ *окружностей с общим центром $O$*;
для типа $R_7$ — $k > 0$ *окружностей с общим центром $O$ и точку $O$*.

Утверждение 32. *Для конечного неисключительного множества $\mathcal{D}$ существует конечное множество $\mathcal{D}_0 \subset \mathcal{M}_r(\mathcal{D})$, содержащее только точки плоскости, и такое, что $\mathcal{D} \subset \mathcal{M}(\mathcal{D}_0)$* (более того, для данного $\mathcal{D}$ множества $\mathcal{D}_0$ можно предъявить явно).

Например, для $\mathcal{D} = \{S, l\}$, где $S$ — окружность с центром $O$, $l$ — прямая и $O \notin l$, достаточно положить $\mathcal{D}_0 = \{O, E, P\}$, где $E \in l$ — основание перпендикуляра, опущенного из $O$ на $l$, и $P \in S \cap l$. Для других неисключительных множеств $D$ множество $\mathcal{D}_0$ предъявляется столь же явно.

Следствие 33. *Для конечного неисключительного множества начальных данных $\mathcal{D} \subset M$ справедливы равенства $\mathcal{M}_G(\mathcal{D}) = \mathcal{M}_r(\mathcal{D}) = \mathcal{M}(\mathcal{D}_0) = \mathcal{M}_T$, где $T$ — поле, согласованное с $\mathcal{D}_0$.*

Доказательство. Согласно утверждению, $\mathcal{D} \subset \mathcal{M}(\mathcal{D}_0)$. Но $\mathcal{M}(\mathcal{D}_0) = \mathcal{M}_T$ (см. теорему 23) и $\mathcal{M}_G(\mathcal{D}) \subset \mathcal{M}_T$ (см. теорему 30). Справедливы включения $\mathcal{M}_G(\mathcal{D}) \supset \mathcal{M}_r(\mathcal{D}) \supset \mathcal{M}(\mathcal{D}_0)$. Следствие доказано.

Мы описали класс $\mathcal{M}_G(\mathcal{D})$ для неисключительных $\mathcal{D}$ и показали, что для построения его объектов операция 4) не требуется: $\mathcal{M}_G(\mathcal{D}) = \mathcal{M}_r(\mathcal{D})$.

## 2.7. Трисекция угла

Следующая классическая задача связана с классом $\mathcal{M}_G(\mathcal{D})$ для исключительного множества $\mathcal{D}$ типа $R_4$.

Задача о трисекции угла. *Разделить данный угол на три равных части.*

Опишем класс $\mathcal{M}_G(\mathcal{D})$ для множества $\mathcal{D}$ типа $R_4$ (классы $\mathcal{M}_G(\mathcal{D})$ для исключительных множеств $\mathcal{D}$ других типов описываются также). Итак, пусть $\mathcal{D}$ — это $k > 0$ прямых, проходящих через точку $O$. Фиксируем любую окружность $S$ с центром $O$. Будем пользоваться следующими обозначениями: $\mathcal{D}'$ — множество точек, равное $\bigcup_{l \in \mathcal{D}} (S \cap l)$; $T$ — поле, согласованное с $\mathcal{D}'$; $l_0 \in \mathcal{D}$ — фиксированная прямая.

Теорема 34. *Класс $\mathcal{M}_G(\mathcal{D})$ состоит из точки $O$ и из всех прямых $l$, проходящих через $O$ и таких, что $|\cos(l, l_0)| \in T$* (здесь $(l, l_0)$ — любой и из двух углов, образованных прямыми $l$ и $l_0$).

Доказательство. Выберем любую точку $E \in l_0 \setminus O$, пользуясь операцией 4). Построим окружность $S$ с центром $O$ и радиусом, равным $[O, E]$. Вместе с $S$ построим множество $\mathcal{D}'$. Класс $\mathcal{M}(\mathcal{D}') = \mathcal{M}_T$ содержит все прямые $l$, проходящие через $O$, такие, что $|\cos(l, l_0)| \in T$, и не содержит других прямых, проходящих через $O$. Класс $\mathcal{M}_G(\mathcal{D})$ принадлежит классу $\mathcal{M}(T)$ и тоже не содержит других прямых, проходящих через $O$.

Объекты множества $\mathcal{D}$ инвариантны относительно группы $G_O$ гомотетий с центром $O$, поэтому все объекты класса $\mathcal{M}_G(\mathcal{D})$ тоже инвариантны относительно $G_O$. Действительно, при гомотетии построение переходит в гомотетичное построение, если произвольные точки выбирать гомотетичными точкам в исходном построении. Но объекты класса $\mathcal{M}_G(\mathcal{D})$ не

зависят от произвольных выборов, сделанных при их построении, т. е. они инвариантны относительно группы $G_O$. Только прямые, проходящие через $O$, и точка $O$ инвариантны относительно этой группы.

Разрешимость задачи о трисекции угла существенно зависит от величины этого угла (см. следствия 35–38).

Следствие 35. *Если $\mathcal{D} = \{l_0, l_1\}$, где $l_0$, $l_1$ — прямые, проходящие через $O$, и $a = |\cos(l_0, l_1)|$, то класс $\mathcal{M}_G(\mathcal{D})$ состоит из $O$ и из прямых $l$, таких, что $O \in l$ и $|\cos(l, l_0)| \in T$, где $T$ — минимальное вещественное поле, содержащее $a$ и замкнутое относительно извлечения квадратных корней.*

Следствие 36. *В условиях следствия* 35 *можно построить прямые, делящие угол $(l_0, l_1)$ на $n$ равных частей, если и только если уравнение $P_n(x) = a$, где $P_n$ — полином Чебышёва степени $n$, разрешимо в 2-радикалах над $T$.*

Отметим, что *если $a$ — трансцендентное число, то уравнение $P_n(x) = a$ неприводимо над полем $\mathbb{Q}(a)$.* Действительно, поле $\mathbb{Q}(a)$ изоморфно полю рациональных функций $\mathbb{Q}(t)$ над $\mathbb{Q}$, а уравнение $P_n(x) = t$ неприводимо даже над полем $\mathbb{C}(t)$ (риманова поверхность алгебраической функции $x(t)$, определенной этим уравнением, — сфера Римана).

Следствие 37. *Если в условиях следствия* 36 *число $a$ трансцендентно, то угол можно поделить на $n$ равных частей циркулем и линейкой, если и только если $n = 2^k$. В частности, трисекция такого угла невозможна.*

Действительно, если неприводимое уравнение решается в 2-радикалах, то его степень равна $2^k$. С другой стороны, угол можно поделить на $2^k$ частей, последовательно строя биссектрисы.

Следствие 38. *Если в условиях следствия* 35 *число $a$ рационально, то трисекция угла возможна, если и только если уравнение $4x^3 - 3x = a$ имеет рациональный корень.*

Следствие 38 доставляет явно проверяемый критерий разрешимости задачи о трисекции угла, косинус которого рационален. В частности, легко видеть, что трисекция угла в $60^\circ$ невозможна.

## 2.8. Одна теорема из аффинной геометрии

Формулируемая ниже теорема показывает, что для построения вещественной аффинной геометрии на плоскости нужны лишь точки и прямые,

а понятие непрерывности не нужно. Ее доказательство основано на возможности выполнения арифметических операций с помощью параллельных прямых (см. п. 2.2).

Теорема 39. *Пусть* $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ *— взаимно однозначное отображение, переводящее каждую прямую в прямую. Тогда* $F$ *— аффинное преобразование.*

Лемма 40. *Если* $\varphi\colon \mathbb{R} \to \mathbb{R}$ *автоморфизм поля* $\mathbb{R}$, *то* $\varphi(x) = x$ *при* $x \in \mathbb{R}$.

Доказательство. Если $x \in \mathbb{Q}$, то очевидно $\varphi(x) = x$. Если $x \geqslant 0$, то $x = a^2$ и $\varphi(x) = \varphi^2(a) \geqslant 0$, т. е. $\varphi$ монотонно. Значит, $\varphi(x) = x$ при $x \in \mathbb{R}$.

Лемма 41. *Если в условиях теоремы* $F(O) = O$ *и* $F(E) = E$, *где* $O \neq E$, *то ограничение* $F$ *на прямую* $OE$ *— тождественное отображение.*

Доказательство. Введем координату на прямой $OE$, отождествляя $O$ с нулем, а $E$ с единицей. Отображение $F$ переводит непересекающиеся прямые в непересекающиеся прямые, т. е. оно сохраняет соотношение параллельности между прямыми. Используя параллельные прямые и точки $O = 0$ и $E = 1$, по точкам $a, b \in OE$ можно построить точки $-a, a^{-1}, a + b, ab$ (см. лемму 22). Поэтому ограничение $F$ на $OE$ задает автоморфизм числовой прямой. Осталось воспользоваться леммой 40.

Доказательство теоремы. Отображения, удовлетворяющие условию теоремы, образуют группу $G$, содержащую группу аффинных преобразований. Подгруппа $G_0$, фиксирующая точки $A$, $B$, $C$, не лежащие на одной и прямой, тривиальна. Действительно, если $\Psi \in G_0$, то по лемме 41 ограничение $\Psi$ на продолжения сторон треугольника $ABC$ является тождественным преобразованием (так как $\Psi$ фиксирует вершины треугольника). Через каждую точку $P$ можно провести прямую $l_P$, пересекающую стороны треугольника $ABC$ в двух разных точках (которые $\Psi$ фиксирует). Применяя лемму 41 к прямой $l_P$, получим $\Psi(P) = P$, т. е. группа $G_0$ тривиальна. Следовательно, группа $G$ содержит не более одного преобразования, переводящего $A$, $B$, $C$ в точки $A'$, $B'$, $C'$, не лежащие на одной прямой. Но среди таких преобразований есть аффинное преобразование. Теорема доказана.

## Список литературы

[1] Бурда Ю., Кадец Л. *Семнадцатиугольник и закон взаимности Гаусса* // Математическое просвещение. Сер. 3. Вып. 17. 2013. С. 61–67.

---

А. Г. Хованский, Univ. of Toronto

# Семнадцатиугольник и закон взаимности Гаусса

Бурда Ю. Кадец Л.

В этой заметке обсуждается как построить семнадцатиугольник при помощи циркуля и линейки и как данное построение связано с квадратичным законом взаимности Гаусса.

## 1. Введение

Теория Галуа позволяет дать полный ответ на вопрос о возможности построения правильного $n$-угольника с помощью циркуля и линейки:

Теорема 1. *Правильный $n$-угольник можно построить циркулем и линейкой если и только если $n$ имеет вид $n = 2^m \cdot p_1 \cdot \ldots \cdot p_k$, где $p_1, \ldots, p_k$ — различные простые числа вида $2^{2^s} + 1$.*

К сожалению теория Галуа позволяет лишь судить о возможности того или иного построения и подсказывает некоторые шаги на пути к нему, но явное построение приходится находить отдельно в каждом конкретном случае.

Тем более удивительно, что построение семнадцатиугольника с помощью циркуля и линейки было придумано *до* появления теории Галуа. Первым это построение нашёл Гаусс в конце восемнадцатого века, спустя много столетий после того как правильные 3-, 4-, 5-, 6-, 8-, 10-, 12-, 15- и 16-угольники были построены древними геометрами.

В этой заметке представлен способ построения правильного 17-угольника. Описание построения не использует никаких идей, которые не были известны в то время, когда Гаусс его придумал. Увы, некоторые действия в этом построении выглядят несколько загадочно и могут быть куда лучше поняты в свете более общей теории. Объяснения такого рода даны в замечаниях 1 и 2. В замечании 3 мы объясняем связь построения 17-угольника с другими работами Гаусса, а именно, с квадратичным законом взаимности.

Благодарности. Мы благодарим Аскольда Георгиевича Хованского, чьи замечательные лекции по теории Галуа вдохновили нас задуматься о построении 17-угольника. Мы также благодарим Бориса Кадеца, который

сообщил нам нескольких крайне красивых доказательств квадратичного закона взаимности.

## 2. План построения

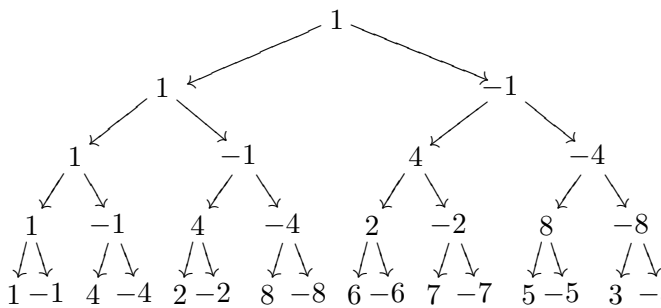
Если отождествить плоскость $\mathbb{R}^2$ с плоскостью комплексных чисел $\mathbb{C}$ (с отмеченными точками 0 и 1), то задача о построении правильного $n$-угольника переходит в задачу о построении примитивного корня из единицы $\xi$ степени $n$ (например $\xi = e^{2\pi i/n}$). Действительно, если это число построено, то его степени являются вершинами правильного $n$-угольника.

В дальнейшем мы будем часто использовать следующее построение: если числа $z_1$ и $z_2$ являются решениями квадратного уравнения $z^2+az+b = 0$, и точки $a$ и $b$ уже построены, то точки $z_1$ и $z_2$ можно построить циркулем и линейкой. Это построение основывается на формуле для решений квадратного уравнения, в которой присутствуют лишь операции сложения, вычитания, умножения, деления и извлечения квадратного корня. Все эти операции с уже построенными точками можно выполнить с помощью циркуля и линейки.

Наш план тогда сводится к нахождению явных квадратных уравнений со следующими свойствами:

- первое из этих уравнений имеет целочисленные коэффициенты;
- коэффициенты каждого из уравнений либо целые числа, либо (с точностью до знака) совпадают с корнями предыдущего уравнения;
- число $\xi$ является одним из решений последнего уравнения.

Эти уравнения будут продвигать нас вниз по следующей диаграмме:




Стрелки в этом дереве ведут от чисел к их квадратным корням по модулю 17.

С каждым узлом дерева мы свяжем число $\sum_i \xi^i$, где суммирование производится по набору тех чисел $i$ из самого нижнего ряда, до которых можно добраться по стрелкам из данного узла.

Ниже мы покажем, что числа, связанные с узлами в каждой строке, удовлетворяют квадратным уравнениям с коэффициентами, которые являются либо целыми числами, либо (с точностью до знака) числами, связанными с узлами предыдущей строки.

Замечание 1. Минимальным уравнением над полем $\mathbb{Q}$, которому удовлетворяет примитивный корень из единицы $\xi$ простой степени $p$, является уравнение $1 + x + \ldots + x^{p-1} = 0$. Его группа Галуа — циклическая группа $G = \mathbb{Z}_p^*$. Для каждого делителя $d$ числа $p - 1$ эта группа содержит единственную подгруппу порядка $d$. В частности если $p = 2^k + 1$, подгруппы группы $G$ образуют башню $G = G_k \supset G_{k-1} \supset \ldots \supset G_0 = \{1\}$, где $G_m$ имеет порядок $2^m$. Соответствие Галуа сопоставляет этой башне подгрупп цепочку квадратичных расширений полей $\mathbb{Q}(\xi) = L_0 \supset L_1 \supset \ldots \supset L_k = \mathbb{Q}$.

Вычеты из $m$-й строки дерева составляют группу $G_m$. Числа $\sum\limits_{i \in G_m} \xi^i$, связанные с узлами диаграммы, порождают расширение $L_m/\mathbb{Q}$. Коэффициенты квадратного уравнения, которое мы находим на $m$-м шаге, находятся в поле $L_{m-1}$, а его корни порождают расширение $L_m/L_{m-1}$.

## 3. Шаг номер 0

Пусть $\xi \neq 1$ — корень из единицы степени $p$. Тогда

$$\sum_{i=1}^{p-1} \xi^i = \frac{\xi^p - \xi}{\xi - 1} = -1.$$

Таким образом, с корнем дерева связано число $-1$.

## 4. Первый шаг — квадратичные вычеты

Теорема 2. *Пусть $Q$ — множество квадратичных вычетов по модулю нечётного простого числа $p$. Пусть $\xi$ — примитивный корень из единицы степени $p$ и пусть $x = \sum\limits_{i \in Q} \xi^i$.*

*Для $p$ вида $p = 4m + 1$ выполняется* $x^2 + x - \dfrac{p-1}{4} = 0$.

*Для $p$ вида $p = 4m - 1$ выполняется* $x^2 + x + \dfrac{p+1}{4} = 0$.

Замечание 2. Для любого нечётного простого $p$ группа $\mathbb{Z}_p^*$ содержит единственную подгруппу индекса 2, а именно группу $Q$ квадратичных вычетов по модулю $p$. Группа $\mathbb{Z}_p^*$ является группой Галуа расширения Галуа $\mathbb{Q}(\xi)/\mathbb{Q}$: элемент $m \in \mathbb{Z}_p^*$ действует автоморфизмом, переводящим $\xi$ в $\xi^m$.

Орбита числа $x = \sum\limits_{i \in Q} \xi^i$ при действии этой группы состоит из двух элементов. Таким образом, $x$ является корнем некоторого квадратного уравнения с рациональными коэффициентами. Теорема 2 предоставляет явный вид этого уравнения.

Доказательство теоремы 2. Чтобы найти $x$, мы вычислим $y = \sum\limits_{i=0}^{p-1} \xi^{i^2}$ и используем тождество $y = 2x - 1$. Комплексное сопряжение числа $y$ равно $\overline{y} = \sum\limits_{i=0}^{p-1} \xi^{-i^2}$, а значит,

$$y\overline{y} = \left(\sum_{i=0}^{p-1} \xi^{i^2}\right)\left(\sum_{i=0}^{p-1} \xi^{-j^2}\right) = \sum_{i,j=0}^{p-1} \xi^{i^2-j^2}.$$

Коэффициент при $\xi^k$ в этой формуле равен количеству решений $(i, j)$ сравнения $i^2 - j^2 \equiv k \mod p$. Обратимой заменой координат $(a, b) = (i - j, i + j)$ это сравнение переводится в сравнение $ab \equiv k \mod p$. У последнего сравнения $p - 1$ решений, если $k \neq 0$, и $2p - 1$ решений, если $k = 0$.

Таким образом,

$$y\overline{y} = 2p - 1 + (p - 1)\sum_{k=1}^{p-1} \xi^k = 2p - 1 - (p - 1) = p.$$

Если $p \equiv 1 \mod 4$, то $\overline{y} = y$. Действительно, так как $-1$ является квадратичным вычетом по модулю $p$, выражение $\sum_{i=0}^{p-1} \xi^{-i^2}$ не отличается от выражения $\sum_{i'=0}^{p-1} \xi^{i'^2}$.

Если $p \equiv -1 \mod 4$, то $\overline{y} = -y$. Действительно, так как $-1$ не является квадратичным вычетом по модулю $p$, в выражении $y + \overline{y}$ каждая степень $\xi$ встречается ровно дважды. А значит, $y + \overline{y} = 2\sum_{k=0}^{p-1} \xi^k = 0$.

В обоих случаях окончательный результат следует из тождества $y = 2x - 1$.

Из теоремы следует, что числа, связанные с узлами во второй строке, являются решениями уравнения $x^2 + x - \dfrac{17 - 1}{4} = 0$, то есть $x_1, x_2 = \dfrac{-1 \pm \sqrt{17}}{2}$.

Замечание 3. Интересно заметить, что вычисление из доказательства теоремы 2 можно использовать для простого и элегантного доказательства квадратичного закона взаимности Гаусса.

Чтобы выяснить, является ли $p$ квадратичным вычетом по модулю простого числа $q$, достаточно выяснить, принадлежит ли $\sqrt{p}$ полю $\mathbb{F}_q$ вычетов по модулю $q$.

Поле $\mathbb{F}_q$ состоит в точности из $q$ корней уравнения $y^q = y$, так что элемент $y$ расширения поля $\mathbb{F}_q$ находится в $\mathbb{F}_q$ в точности когда $y^q = y$.

Пусть, как и раньше, $y = \sum_{i=0}^{p-1} \xi^{i^2}$, где $\xi$ — примитивный корень из единицы степени $p$, лежащий в некотором расширении поля $\mathbb{F}_q$.

Вычисление из доказательства теоремы 2 показывает, что если $p \equiv 1 \bmod 4$, то $y = \sqrt{p}$ и находится в $\mathbb{F}_q$, если и только если $y^q = y$. Однако в расширениях поля $\mathbb{F}_q$ имеет место формула $(a+b)^q = a^q + b^q$, а значит, $y^q = \sum_{i=0}^{p-1} \xi^{i^2 q}$. Таким образом, $y^q = y$ в точности тогда, когда $q$ является квадратичным вычетом по модулю $p$.

Итак, если $p \equiv 1 \mod 4$ и $q \neq p$ — простые числа, то $p$ является квадратичным вычетом по модулю $q$ тогда и только тогда, когда $q$ является квадратичным вычетом по модулю $p$.

Если $p \equiv -1 \mod 4$, то $y = \sqrt{-p}$ и те же рассуждения показывают, что $-p$ является квадратичным вычетом по модулю $q$ тогда и только тогда, когда $q$ является квадратичным вычетом по модулю $p$.

## 5. Шаг второй, несколько тригонометрических тождеств

Лемма 1. *Для любого* $\xi \neq \pm 1$

$$(\xi + \xi^{-1})(\xi^2 + \xi^{-2})(\xi^4 + \xi^{-4}) \cdot \ldots \cdot (\xi^{2^n} + \xi^{-2^n}) = \frac{\xi^{2^{n+1}} - \xi^{-2^{n+1}}}{\xi - \xi^{-1}}\,.$$

Это тождество легко проверить, домножив обе части равенства на $\xi - \xi^{-1}$, а затем $n$ раз использовав тождество $(\xi^{2^k} - \xi^{2^k})(\xi^{2^k} + \xi^{2^k}) = (\xi^{2^{k+1}} - \xi^{2^{k+1}})$.

Следствие 1. *Если* $\xi \neq \pm 1$ *— корень из единицы степени* $p$, *где* $p = 2^{n+1} + 1$, *то*

$$(\xi + \xi^{-1})(\xi^2 + \xi^{-2})(\xi^4 + \xi^{-4}) \cdot \ldots \cdot (\xi^{2^n} + \xi^{-2^n}) = -1.$$

Пусть теперь $c_k = \xi^k + \xi^{-k}$.

Лемма 2. *Для любых* $s, t$

$$c_s \cdot c_t = c_{s+t} + c_{s-t}\,.$$

Замечание 4. Если $\xi$ равен $e^{i\alpha}$, то $c_k = 2\cos k\alpha$. В этом случае лемма 2 следует из формулы $2\cos(s\alpha)\cos(t\alpha) = \cos((s+t)\alpha) + \cos((s-t)\alpha)$, а

следствие 1 следует из тождества

$$2^n \cos\alpha \cos(2\alpha) \cos(4\alpha) \cdot \ldots \cdot \cos(2^n\alpha) = \frac{\sin(2^{n+1}\alpha)}{\sin\alpha}\,.$$

Найдём теперь квадратное уравнение с корнями $c_1 + c_4$ и $c_2 + c_8$.

Сумма корней этого уравнения — число $c_1 + c_2 + c_4 + c_8$, которое мы нашли на прошлом шаге: это корень $x_1$ уравнения $x^2 + x - 4 = 0$.

Чтобы найти произведение $(c_1 + c_4)(c_2 + c_8)$ будем рассуждать так: из следствия 1, применённого к корню из единицы $\xi$, мы находим $c_1c_2c_4c_8 = = -1$. То же следствие, применённое к корню из единицы $\xi^3$, приводит к тождеству $c_3c_6c_{12}c_{24} = -1$, то есть $c_3c_6c_5c_7 = -1$.

Подставляя тождества $c_3c_5 = c_2 + c_8$, $c_6c_7 = c_1 + c_4$ в $c_3c_6c_5c_7 = -1$, получаем $(c_1 + c_4)(c_2 + c_8) = -1$.

Таким образом, $c_1 + c_4$ и $c_2 + c_8$ являются решениями $y_1, y_2$ уравнения $y^2 - x_1y - 1 = 0$, где $x_1$ — решение уравнения $x^2 + x - 4 = 0$.

Так же находим, что $c_3 + c_5$ и $c_6 + c_7$ являются решениями $y_3, y_4$ уравнения $y^2 - x_2y - 1 = 0$, где $x_2$ — другой корень уравнения $x^2 + x - 4 = 0$.

## 6. Шаг третий

Сумма чисел $c_1$ и $c_4$ равна $c_1 + c_4 = y_1$, а их произведение равно $c_1c_4 = = c_3 + c_5 = y_3$. Следовательно, $c_1, c_4$ являются решениями уравнения $z^2 - - y_1z + y_3 = 0$.

## 7. Последний шаг

Поскольку $\xi + \xi^{-1} = c_1$ и $\xi \cdot \xi^{-1} = 1$, то $\xi$ и $\xi^{-1}$ являются корнями уравнения $w^2 - c_1w + 1 = 0$.

## 8. Подведение итогов

При решении четырёх квадратных уравнений, выписанных выше, нам приходится четыре раза выбирать один из двух корней. Таким образом, мы можем получить 16 различных ответов, являющихся различными примитивными корнями из единицы степени 17.

Если мы выберем $\xi = e^{2\pi i/17}$, то мы можем проследить, какие корни нам следует выбирать, чтобы найти $\xi$:

$$c_1 + c_2 + c_4 + c_8 = \frac{-1 + \sqrt{17}}{2}\,,$$

$$c_3 + c_5 + c_6 + c_7 = \frac{-1 - \sqrt{17}}{2}\,,$$

$$c_1 + c_4 = \frac{\frac{-1+\sqrt{17}}{2} + \sqrt{\frac{17-\sqrt{17}}{2}}}{2} = A,$$
$$c_3 + c_5 = \frac{\frac{-1-\sqrt{17}}{2} + \sqrt{\frac{17+\sqrt{17}}{2}}}{2} = B,$$
$$c_1 = \frac{A+\sqrt{A^2-4B}}{2} = C,$$
$$\xi = \frac{C+\sqrt{C^2-4}}{2}\,.$$

Юрий Бурда, Univ. of Toronto
Email: yburda@math.toronto.edu
Люся Кадец, Univ. of Toronto
Email: lucy.kadets@math.toronto.edu

# Оригами, или что можно получить с помощью складывания листа бумаги

Д. И. Грищенко

## 1. Введение

Хорошо известно, что множество координат точек, которые можно получить при построениях с помощью циркуля и линейки, есть поле, получающееся из $\mathbb{Q}$ итерированием операции извлечения квадратного корня. Естественно поставить вопрос о множествах точек, получающихся в результате других геометрических построений.

В этой работе мы изучим множество чисел оригами, то есть множество координат точек, которые можно получить при помощи складывания листа бумаги (при этом образуются прямые, которые, конечно, при пересечении дают нам точки). Оказывается, с помощью оригами выполнимы такие операции как удвоение куба, трисекция угла ... Более того, множество чисел оригами есть поле, получающееся из $\mathbb{Q}$ итерированием операций извлечения квадратного и кубического корней.

Основой для данной работы послужила статья [3], в которой формулируется результат об описании чисел оригами. Однако, многие существенные доказательства в [3] отсутствуют, в том числе и доказательство того, что множество чисел оригами не выходит за рамки того, что описано в теореме о характеризации чисел оригами. Кроме восполнения недостающих доказательств, мы получаем новые результаты, отсутствовавшие в [3]. Среди прочего, это теорема о характеризации фалесовых чисел и теорема о выделении полной независимой системы аксиом оригами.

Основываясь на [5], мы берем естественный набор из шести аксиом, описывающих процесс складывания из бумаги. Мы последовательно изучаем все увеличивающиеся множества точек, получающиеся добавлением к первым трем аксиомам четвертой, пятой, а потом рассматривая все шесть аксиом (числа Фалеса, Пифагора, Евклида и числа оригами). Основной результат описывает следующая теорема:

Теорема. *Множество точек, построимых с помощью складывания из бумаги при изначально данных точках* $0$ *и* $1$ *— это наименьшее подполе*

$\mathbb{C}$, *замкнутое относительно операций извлечения квадратного и кубического корня.*

В конце нами доказана теорема, выделяющая из шести аксиом независимую подсистему, состоящую из аксиом 2, 3 и 6.

## 2. Аксиомы складывания

В этой работе мы будем рассматривать только так называемое "One-Fold Origami", когда разрешается складывать только по одной линии и нельзя осуществлять одновременное складывание по двум и более линиям. Помимо этого есть "Two-Fold Origami" и "Multi-Fold Origami", когда разрешается складывать лист бумаги в двух (или трех и более) местах одновременно. Эти варианты и многое другое подробно описано в статье Р. Альперина и Р. Ленга [4].

Имеется шесть аксиом складывания листа бумаги, выполнение которых легко проверить. Мы не будем этого делать, как и обсуждать вопрос о том, почему эти аксиомы описывают все возможные складывания. Наша цель — описать множество построимых точек.

(A1) Если точки $A$ и $B$ построимы, то прямая $l$, проходящая через $A$ и $B$, построима.

(A2) Если прямые $k$ и $l$ построимы, то точка $k \cap l$ построима.

(A3) Можно построить серединный перпендикуляр к отрезку с концами в построимых точках.

(A4) Можно построить биссектрису любого построимого угла (угол считаем построимым, если мы можем построить его стороны).

(A5) Если дана построимая прямая $l$ и построимые точки $P$, $Q$, то построима такая прямая, проходящая через $Q$, что при симметрии относительно нее $P$ попадает на $l$ (при условии, что такая прямая существует).

(A6) Если даны прямые $k, l$ и точки $P, Q$, то мы можем построить такую прямую $m$ (если она существует), что при симметрии относительно $m$, точка $P$ переходит на $k$, а точка $Q$ переходит на $l$.

Эти аксиомы являются зависимыми и эквивалентная им независимая подсистема аксиом будет описана в седьмом разделе.

Конструкции, которые можно получить, используя только первые три аксиомы называются фалесовскими, четыре — пифагоровыми, пять — евклидовыми. До перехода к полю чисел оригами, опишем множества точек, получаемых с помощью каждой из этих конструкций.

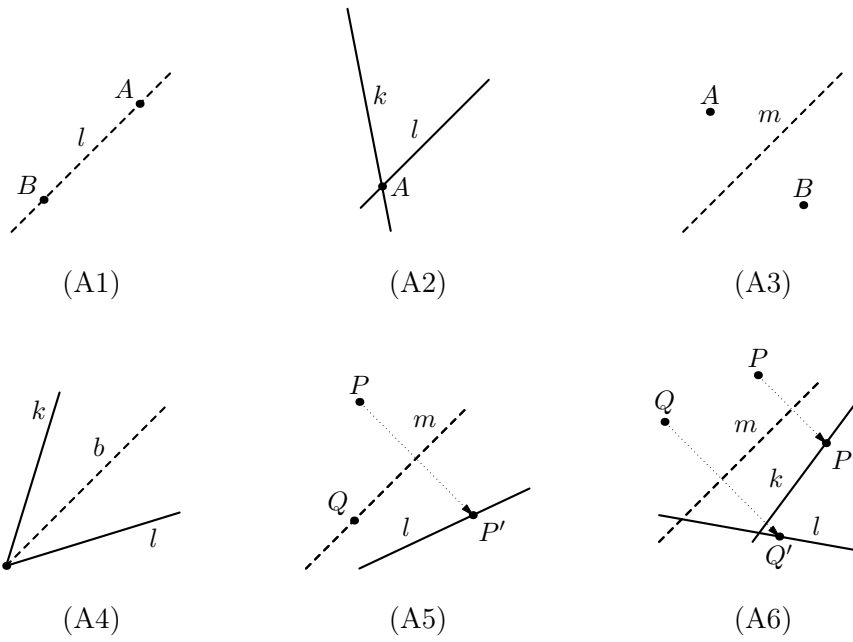



*Рис. 1. Аксиомы складывания*

## 3. Фалесовские конструкции

Для начала хотелось бы разобраться с тем, что мы можем получить при использовании только первых трех аксиом.

Хотелось бы предупредить читателя, что не все факты, описанные в этом разделе, является обязательными для доказательства основной теоремы. Если интересует только часть, связанная с ней, то достаточно прочесть только леммы 3.1–3.2, следствия 3.1–3.5, предложение 3.1.

Все конструкции мы рассматриваем над полем $\mathbb{C}$ (т. е. используем на плоскости комплексную координату).

Пусть даны три точки $A$, $B$, $C$, не лежащие на одной прямой. Определим множества точек $\Pi = \Pi(A, B, C)$ и прямых $\mathcal{L} = \mathcal{L}(A, B, C)$ как наименьшие множества, элементы которых — точки и прямые над $\mathbb{C}$ со следующими свойствами:

(0) $\Pi$ содержит точки $A$, $B$, $C$;

(1) прямая, проходящая через точки множества $\Pi$, принадлежит $\mathcal{L}$;

(2) точка пересечения двух прямых из $\mathcal{L}$ принадлежит $\Pi$;

(3) серединный перпендикуляр к отрезку с концами в точках из $\Pi$ принадлежит $\mathcal{L}$.

Будем называть точку построимой, если она принадлежит $\Pi$. Будем называть прямую построимой, если она принадлежит $\mathcal{L}$.

Заметим, что $\Pi$ — это множество, которые мы можем получить из точек $A, B, C$, складывая лист бумаги с использованием только первой, второй и третьей аксиом.

На протяжении этого параграфа мы будем работать именно с такими построениями.

Лемма 3.1. *Если имеется отрезок $AB$ и точка $P$, то мы можем построить отрезок с концом в этой точке, параллельный исходному и равный ему по длине.*

Доказательство. Пусть точка $P$ не лежит на отрезке $AB$ (см. рис. 2). Соединим точки $A$ и $B$ с точкой $P$. Отметим середины сторон получившегося треугольника $p$, $a$, $b$, лежащие против соответствующих вершин (эти точки построимы по третьей аксиоме).

Пересечем $Pp$ с $ba$, получим точку $O$. Пересечем $pa$ и $AO$, получим точку $T$. Заметим, что $PT \parallel AB$ и $|AB| = 2|PT|$.

Теперь пересечем прямые $Aa$ и $PT$ получим точку $V$. Тогда $PV$ — требуемый отрезок.

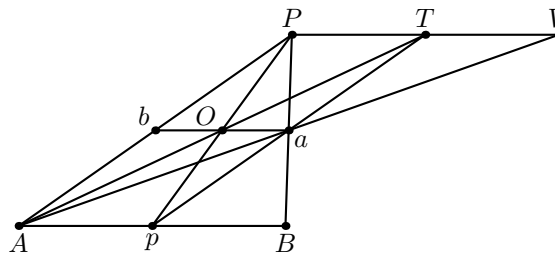



*Рис. 2.*

Пусть $P$ лежит на $AB$, тогда воспользуемся некоторой точкой $Q$, не лежащей на $AB$ (она существует, так как по условию точки $A$, $B$, $C$ не лежат на одной прямой), по первой части доказательства из нее можно отложить отрезок, затем отложим равный данному отрезок из точки $P$. □

Следствие 3.1. *Из данной точки можно опустить перпендикуляр на данную прямую.*

Доказательство. По лемме 3.1 мы можем провести через данную точку $P$ прямую $m'$ параллельную данной прямой $m$ (мы можем провести параллельную прямую, так как по лемме 3.1 мы можем провести отрезок равный и параллельный данному, а через него можно провести прямую). Теперь возьмем на $m'$ такой отрезок что, $P$ — его середина (см. рис. 3).

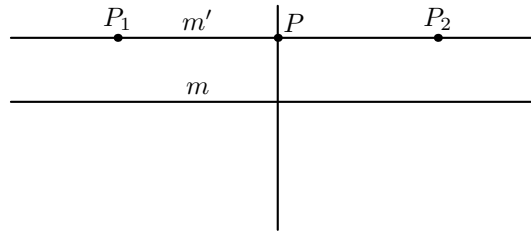



*Рис. 3.*

Проведем к нему серединный перпендикуляр (это возможно по аксиоме 3). Это и будет требуемый перпендикуляр. □

Следствие 3.2. *Если даны три точки $A$, $B$, $C$, лежащие на одной прямой, и точка $P$, не лежащая на ней, то мы можем построить такую точку $D$, что $\triangle ABD \sim \triangle ACP$ (см. рис. 4).*

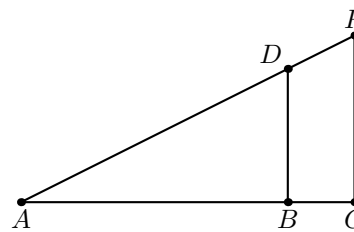



*Рис. 4.*

Доказательство. Через точку $B$ (рис. 4) проведем прямую, параллельную $PC$, это возможно по лемме 3.1. Она пересечет прямую $AP$ в некоторой точке, это и будет точка $D$ (очевидное подобие по 2 углам). □

Следствие 3.3. *Если даны точка $P$ и прямая $l$, то мы можем отразить $P$ относительно $l$. Если даны прямые $k$, $l$, то мы можем отразить $k$ относительно $l$.*

Доказательство. Если мы умеем отражать точку, то, очевидно, мы умеем отражать прямую (отразив две точки, лежащие на этой прямой).

Если $P$ лежит на прямой $l$, то ее образ есть она сама. Если нет, то проведем через точку $P$ прямую $l_1$, параллельную $l$ (см. рис. 5). Теперь мы возьмем на $l_1$ точку $A$, отличную от $P$. Из точек $A$ и $P$ опустим перпендикуляры на прямую $l$. Они пересекут $l$ в точках $B$ и $C$ соответственно. Теперь по лемме 3.1 мы можем построить отрезок из точки $C$, равный и параллельный $AB$, не совпадающий с $PC$ (это возможно, так как мы

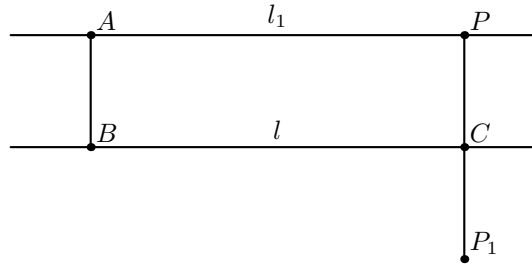



*Рис. 5.*

умеем строить два таких отрезка). Тогда его конец, отличный от $C$, и будет образом $P$ при отражении относительно прямой $l$. □

Следствие 3.4. *Множество* $\Pi$ *замкнуто относительно добавления отрезков в* $\mathbb{C}$, *т. е. если у нас есть два отрезка, то мы можем приписовывать их друг к другу и получать построимые точки. Если положить* $A = 0$, *то* $\Pi$ — *абелева группа.*

Доказательство. Возьмем $AQ$, $RT$, тогда мы можем из точки $Q$ отложить отрезок, равный и параллельный $RT$, получим отрезок $QO$. Тогда $A$ и $O$ построимы, а значит отрезок $AO$ построим.

Если мы положим, что $A = 0$, то мы получим абелеву группу, так как все свойства абелевой группы выполнены. □

Без ограничения общности полагаем, что $A = 0$ и $B = 1$. Тогда $\Pi = \Pi(0, 1, C)$ зависит только от комплексного числа $C$. Обозначим $C \equiv z$ и будем далее писать $\Pi(z)$ вместо $\Pi(0, 1, C)$.

Теперь мы можем построить координатные оси как подмножества $\mathbb{R}$ (мы можем провести прямую $AB$, назвав ее осью $\mathcal{X}$, и перпендикулярную к ней прямую через $A$, котоую мы назовем осью $\mathcal{Y}$). Обозначим через $\mathcal{X}(z)$ проекцию $\Pi(z)$ на ось $\mathcal{X}$, а через $\mathcal{Y}(z)$ проекцию $\Pi(z)$ на ось $\mathcal{Y}$. Считаем что $\mathcal{X}(z), \mathcal{Y}(z) \subset \mathbb{R}$. Тогда $\mathcal{X}(z)$ и $i\mathcal{Y}(z)$ это подмножества $\Pi$. Так как $\Pi$ абелева группа, то $\mathcal{X}$ и $\mathcal{Y}$ тоже абелевы группы. Очевидно, что $\Pi(z) = \mathcal{X}(z) \oplus i\mathcal{Y}(z)$ (по следствию 3.4 мы можем «складывать отрезки», а все элементы $\mathcal{X}$ и $\mathcal{Y}$ построимы, так как мы можем опустить перпендикуляр на каждую ось).

Следствие 3.5. *Для любого* $z \in \mathbb{C} \setminus \mathbb{R}$ *множество* $\Pi(z)$ — *векторное пространство над* $\mathbb{Q}$, *замкнутое относительно операции комплексного сопряжения и содержащее подпространства* $\mathcal{X}(z)$ *и* $i\mathcal{Y}(z)$.

Доказательство. Если нам дана точка $W$ и некоторое натуральное число $n$, то мы можем решить относительно $U$ (то есть найти точку $U$) уравнение $nU = W$. Действительно, возьмем некоторую точку $V$, не лежащую на прямой, проходящей через 0 и $W$. По лемме 3.1, мы можем

построить точку $nV$ (отрезок $0(nV)$ в $n$ раз длиннее отрезка $0V$ и они сонаправлены). По следствию 3.2, взяв за точку $P$ точку $W$, за точку $C$ — точку $nV$, за точку $B$ — точку $V$, мы можем построить точку $D$, которая как раз и будет искомой точкой $U$. Из этого факта следует что $\Pi(z)$ — векторное пространство над $\mathbb{Q}$. Замкнутость относительно комплексного сопряжения получаем из следствия 3.3. □

Лемма 3.2. *Если $t \neq 0$ и $t \in \mathcal{Y}$, то $1/t \in \mathcal{Y}$.*

Доказательство. Мы можем построить точку $P$ с координатами $(1, t)$ (выберем на оси $\mathcal{Y}$ точку $t$ (см. рис. 6) и отложим от нее отрезок единичной длины, параллельный оси $\mathcal{X}$). Теперь мы можем восставить

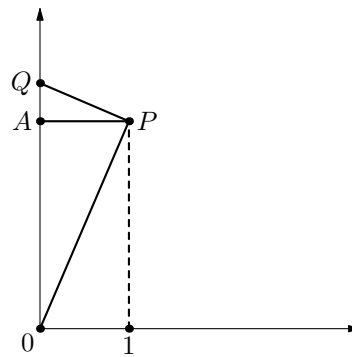



*Рис. 6.*

из $P$ перпендикуляр к отрезку $OP$, тогда он пересечет ось $\mathcal{Y}$ в точке $Q$. Из простых геометрических соображений (формула для высоты в прямоугольном треугольнике) $AQ = 1/t$, где $A = (0, t)$. Но тогда этот отрезок можно отложить из $(0, 0)$ и получить точку $1/t$. □

Предложение 3.1. *Пусть $x \in \mathcal{X}$ и $v, y \in \mathcal{Y}$, тогда выполнено следующее*:

1. *$vy \in \mathcal{X}$, $xy \in \mathcal{Y}$.*
2. *$x^2 \in \mathcal{X}$ и $\mathcal{X}$ есть $\mathbb{Q}$-алгебра.*
3. *Если $y \neq 0$, то $\mathcal{Y} = Xy$.*
4. *Если $x \neq 0$, то $1/x \in \mathcal{X}$ и $\mathcal{X}$ — поле.*
5. *Если $\mathcal{Y} \cap \mathcal{X} \neq \{0\}$, то $\mathcal{X} = \mathcal{Y}$ и $\Pi(z) = \mathcal{X}(i)$.*
6. *$\mu \in \mathbb{R}$ является угловым коэффициентом построимой прямой, тогда и только тогда, когда $\mu \in \mathcal{Y}$.*

Доказательство. 1. Пусть у нас есть две точки $u, x \in \mathcal{X}$ и точка $y \in \mathcal{Y}$. Тогда мы можем построить точку $(x, y)$, а значит и точку $(u, uy/x)$ тоже можно построить по следствию 3.3, следовательно $uy/x \in \mathcal{Y}$. Если $u = 1$, то $y/x \in \mathcal{Y}$.

Пусть $x \in \mathcal{X}$ и $v, y \in \mathcal{Y}$, тогда аналогичные рассуждения приводят к тому, что $vx/y \in \mathcal{X}$. Если $x = 1$, то $v/y \in \mathcal{X}$.

Выводы по поводу $vy$ легко сделать, пользуясь написанным выше, а именно $1/v \in \mathcal{Y}$, а значит $y/(1/x) \in \mathcal{X}$. Положим $1 = u$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, тогда по лемме 3.2 $1/y \in \mathcal{Y}$, а значит и $1/xy \in \mathcal{Y}$, но тогда по лемме 3.2 $xy \in \mathcal{Y}$.

2. Для начала построим точку $(x, y)$. Тогда прямая, проходящая через (0,0) и $(x, y)$, и прямая, параллельная оси $\mathcal{X}$, проведенная на высоте $xy$, пересекаются в точке $(x^2, xy)$, а значит $x^2 \in \mathcal{X}$. Так как $\mathcal{X}$ содержит $\mathbb{Q}$, то это векторное пространство над $\mathbb{Q}$. Замкнутость относительно умножения следует из равенства:

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

3. Из пункта 1 мы знаем, что $\mathcal{X}y \subset \mathcal{Y}$. Возьмем $v, y \in \mathcal{Y}$, $x \in \mathcal{X}$, тогда $xy \in \mathcal{Y}$, а значит $vx/(xy) \in \mathcal{X}$. Перенесем все, кроме $v$, в правую часть и получим, что $v \in \mathcal{X}y$, а значит $\mathcal{Y} \subset Xy$.

4. Пусть $x \neq 0$, $x \in \mathcal{X}$, $y \neq 0$, $y \in \mathcal{Y}$. Тогда из пункта 2 и леммы 3.2 следует, что $1/xy \in \mathcal{Y}$ и $y^2 \in \mathcal{X}$, но $\frac{1}{x} = \frac{1}{y}\frac{y}{x} \in \mathcal{X}$ по лемме 3.2. Тогда из этого и пункта 3 следует, что $\mathcal{X}$ — это поле.

5. Если $y \neq 0$ и $y \in \mathcal{X} \cap \mathcal{Y}$, то $1/y \in \mathcal{Y}$, и тогда $1 \in \mathcal{Y}$, а значит $\mathcal{X} = \mathcal{Y}$, следовательно $\Pi(z) = \mathcal{X}(i)$.

6. Возьмем прямую, проходящую через начало координат и имеющую тангенс наклона $\mu$. Она пересекает вертикальную прямую, проходящую через (1,0) в точке $\mu \in \mathcal{Y}$, а значит точка $(0, \mu)$ построима. Отсюда $\mathcal{Y} \cup \infty$ — это множество построимых $\mu$. □

Теорема 3.1. (i) *Для любого $z = a + bi$, $z \in \mathbb{C} \setminus \mathbb{R}$, множество $\Pi(z)$ — это поле, являющееся расширением $\mathbb{Q}$ и содержащее элемент $z$, замкнутое относительно операции компексного сопряжения и содержащееся в поле $\mathbb{Q}(a, b, i)$.*

(ii) $\Pi(z) = \mathbb{Q}(a, b^2) + \mathbb{Q}(a, b^2)bi$.

Доказательство. (i) Покажем, что $\Pi(z)$ — это поле. Сумма и разность построимых точек, очевидно, построимы. Из предложения 3.1 следует, что для любого построимого $u = x + yi$ точка $u^2 = x^2 - y^2 + 2xyi$ также построима. Если $w$ и $z$ построимы, то $wz$ построима, так как

$$2wz = (u + w)^2 - w^2 - z^2.$$

Сумма квадратов координат $r^2 = x^2+y^2$ построимой точки $z = x+yi = re^{it}$ построима и принадлежит $\mathcal{X}$ по предложению 3.1. По лемме 3.2 точка $1/r^2$ построима, значит $1/z = re^{-it}/r^2$ построима. Таким образом, $\Pi(z)$ — поле. Кроме того, $\overline{w} = r^2/w$, а значит построима, то есть наше поле замкнуто относительно операции комплексного сопряжения.

Если у нас есть точка $z = a + bi$, тогда серединные перпендикуляры к любым построимым отрезкам лежат в множестве, порожденном $a$, $b$ и $i$ над $\mathbb{Q}$. Это утверждение легко доказать индукцией по построению множества $\Pi$. Кроме того, любая построимая прямая имеет угловой коэффициент из $\mathbb{Q}(a, b)$. Тогда координаты любых построимых точек принадлежат $\mathbb{Q}(a, b)$. А значит, для любого $z \in \mathbb{C} \setminus \mathbb{R}$ поле $\Pi(z) \subseteq \mathbb{Q}(a, b, i)$.

(ii) Докажем индукцией по построению множества $\Pi$. База очевидно верна. Пусть имеется подмножество $\mathbb{Q}(a, b^2) + \mathbb{Q}(a, b^2)bi$. Покажем, что применив любую аксиому к точкам этого подмножества, мы не выйдем за пределы множества. Действительно, запишем наши три аксиомы в координатах.

(A1) Пусть у нас были точки $A(0, 0)$ и $B(c, d)$ (мы умеем сдвинуть первую точку в (0,0), какая бы она ни была). Тогда прямая $l$, проходящая через эти точки, будет иметь уравнение: $dx - cy = 0$.

(A2) Пусть у нас есть прямые $l\colon l_1x + l_2y = l_3$ и $m\colon m_1x + m_2y = m_3$, тогда точка их пересечения $P$ имеет координаты

$$\left(\frac{m_2l_3 - l_2m_3}{m_1l_2 - m_2l_1}, \frac{m_1l_3 - l_1m_3}{m_1l_2 - m_2l_1}\right).$$

(A3) Если концы отрезка — это точки $A(0, 0)$ и $B(c, d)$, то серединный перпендикуляр задается уравнением: $cx - dy = (c/2)^2 + (d/2)^2$.

Проверим что первая аксиома не выводит нас из множества $\mathbb{Q}(a, b^2) + +\mathbb{Q}(a, b^2)bi$. Мы знаем из предложения 3.1, что угловой коэффициент прямой всегда лежит в $\mathcal{Y}$. Но мы также знаем, что наши $c, d$ таковы, что $c \in \mathbb{Q}(a, b^2)b$, а $d \in \mathbb{Q}(a, b^2)$, а значит по предложению 3.1 $-\dfrac{d}{c} \in \mathbb{Q}(a, b^2)b$, то есть с аксиомой 1 мы разобрались.

Теперь проверим для второй аксиомы. Нетрудно, заметить что $x$-координата будет принадлежать $\mathbb{Q}(a, b^2)$, так как числитель и знаменатель дроби принадлежат $\mathbb{Q}(a, b^2)bi$. А $y$-координата принадлежит $\mathbb{Q}(a, b^2)bi$, так как числитель принадлежит $\mathbb{Q}(a, b^2)$, а знаменатель $\mathbb{Q}(a, b^2)bi$.

Осталось проверить условие для третьей аксиомы, это делается из соображений, использовавшихся для первой аксиомы. □

## 4. Пифагоровские числа и конструкции

В этом разделе мы будем работать уже с первыми четырьмя аксиомами. К использованным в предыдущем разделе аксиомам добавляется

(A4) прямая, делящая любой построимый угол пополам, построима.

Как и выше, мы предполагаем, что у нас есть точки 0 и 1. Будем называть это свойство аксиомой 0.

Предложение 4.1. *Если нам даны аксиомы* 0–3, *то следующие утверждения эквивалентны:*

1. *Аксиома* 4.
2. *Отрезок единичной длины может быть отложен в любом построимом направлении и из любой построимой точки.*
3. *Любой построимый отрезок может быть отложен в любом построимом направлении и из любой построимой точки.*

Доказательство. $(3 \Rightarrow 2)$ Очевидно.

$(2 \Rightarrow 1)$ Можно считать, что угол, образованный нашим направлением с вещественной осью, меньше $180°$. Отложим на каждой из сторон угла единичный отрезок. Затем из вторых концов отложенных отрезков $A, B$ (не являющихся вершиной угла) восставим перпендикуляры к соответствующим сторонам. Они пересекутся в некоторой точке $C$. Соединим вершину угла и $C$. Тогда образовались два прямоугольных треугольника, которые равны по катету и гипотенузе, а значит полученная прямая — биссектриса угла.

$(1 \Rightarrow 3)$ Мы предполагаем, что можно провести биссектрису у любого построимого угла. Предположим, что отрезок $AB$ построим и дано построимое направление $CL$, в котором требуется отложить отрезок с началом в точке $C$, равный отрезку $AB$. Для начала подвинем отрезок $AB$ так, чтобы он начинался в $C$, получим отрезок $CD$. Отразим его относительно биссектрисы угла $LCD$ и получим некоторый отрезок $CS$, лежащий на стороне $CL$. Это и есть требуемый отрезок, так как биссектриса является осью симметрии угла. □

Очевидно, что точка $i$ построима (мы можем провести биссектрису прямого угла, образованного координатными осями, и отразить относительно нее любую точку из $\mathcal{X}$), а значит, в обозначениях предыдущего раздела, $\mathcal{X} = \mathcal{Y}$. Легко видеть, что с добавлением новой аксиомы мы получаем возможность строить точки пересечения окружности любого построимого радиуса и с центром в любой построимой точке и прямой, проходящей через центр с любым построимым коэффициентом. Оказывается, множество $x$-координат построимых точек замкнуто относительно $\sqrt{a^2 + b^2}$ для

любой построимой точки $(a, b)$. Получившееся множество $x$-координат построимых точек $\mathcal{X}$ принято называть пифагоровскими числами, $\mathcal{P}$.

Теорема 4.1. *Пусть $\pi$ это множество чисел, построимых с помощью аксиом 0–4. Тогда $\pi = \mathcal{P} \oplus i\mathcal{P}$, где $\mathcal{P}$ — такое подполе действиельных чисел, что положительные элементы из $\mathcal{P}$ — это всевозможные длины построимых отрезков.*

Доказательство. Заметим, что мы можем построить любую точку вида $\sqrt{a^2 + b^2}$, где $a, b \in \Pi(z)$. Действительно, мы можем отложить отрезок, равный данному, в любом построимом направлении, а, так как направление перпендикулярное данному построимо, то мы можем построить прямоугольный треугольник с катетами $a$ и $b$. Тогда его гипотенуза будет искомым элементом.

Докажем индукцией по построению, что никаких других точек не добавится.

Исходно имеются элементы 0 и 1.

Предположим, что в результате $k$ применений аксиом 1–4 получаются лишь элементы из требуемого в условии теоремы множества. Применим к полученным точкам одну из аксиом 1–4 и посмотрим, что происходит с координатами точек.

($A1$) Пусть у нас были точки $A(0, 0)$ и $B(a, b)$ (мы умеем сдвинуть первую точку в (0,0), какая бы она ни была). Тогда прямая $l$, проходящая через эти точки будет иметь уравнение: $bx - ay = 0$

($A2$) Пусть у нас есть прямые $l\colon l_1x + l_2y = l_3$ и $m\colon m_1x + m_2y = m_3$, тогда точка их пересечения

$$P = \left(\frac{m_2l_3 - l_2m_3}{m_1l_2 - m_2l_1}, \frac{m_1l_3 - l_1m_3}{m_1l_2 - m_2l_1}\right).$$

($A3$) Если концы отрезка это точки $A(0, 0)$ и $B(a, b)$, то серединный перпендикуляр задается уравнением: $ax - by = (a/2)^2 + (b/2)^2$.

($A4$) Если стороны угла это прямые $l\colon l_1x + l_2y = 0$ и $m\colon m_1x + m_2y = 0$ (можно переместить точку пересечения в $(0, 0)$), то уравнение биссектрисы будет

$$\left(\frac{l_1}{\sqrt{l_1^2 + l_2^2}} + \frac{m_1}{\sqrt{m_1^2 + m_2^2}}\right)x + \left(\frac{l_2}{\sqrt{l_1^2 + l_2^2}} + \frac{m_2}{\sqrt{m_1^2 + m_2^2}}\right)y = 0.$$

Так как все коэффициенты в начальных прямых и координаты точек были из нашего поля, то и все коэффициенты в уравнениях конечных прямых и координаты конечных точек будут из нашего поля. □

## 5. Евклидовы конструкции и числа

Теперь пришло время разобраться, что получится, если у нас даны аксиомы 0–5. Напомним пятую аксиому:

(A5) Если дана построимая прямая $l$ и построимые точки $P$, $Q$, то построима такая прямая, проходящая через $Q$, что при симметрии относительно нее $P$ попадает на $l$ (при условии, что такая прямая существует).

Если нам даны точки 0 и 1, эти аксиомы позволяют получить так называемые евклидовы числа $\mathcal{E}$. На самом деле, $\mathcal{E}$ — это в точности то же самое поле, которое мы можем получить с помощью циркуля и линейки. А это, как известно [2], наименьшее поле, содержащее $\mathbb{Q}$, и замкнутое относительно операции извлечения квадратного корня.

Опишем, как мы можем извлечь квадратный корень из комплексного числа $z = re^{i\theta}$ с помощью аксиом 1–5. Этот процесс можно выполнить в два шага.

1. Извлечение квадратного корня из положительного действительного числа $r$.

2. Построение угла $\theta/2$ посредством проведения биссектрисы угла $\theta$.

Мы умеем извлекать корень из некоторых пифагоровских чисел, но только из тех, что имеют вид $a^2 + b^2$ для построимых $a$ и $b$. Пора использовать пятую аксиому: она как раз и добавляет замкнутость поля относительно операции извлечения квадратного корня. Рассмотрим параболу $\mathcal{T}$ с директрисой $l$ и фокусом $F$.[1)] Тогда пятая аксиома позволяет нам построить точки пересечения этой параболы с построимыми прямыми и касательные в них. Чтобы в этом убедиться, возьмем некую вспомогательную точку $Q$ и проведем через нее такую прямую $t$, что при симметрии относительноо нее $F$ попадает на $l$. Перпендикулярная прямая, $t$, проходящая через $F$, пересечет $t$ в построимой точке $R$. Теперь построим перпендикуляр к $l$ из точки $R$. Он пересечет прямую $t$ в точке $S$. Так как $t$ — это серединный перпендикуляр к $FR$, точка $S$ равноудалена от точек $F, R$, а значит она лежит на параболе $\mathcal{T}$. Прямая $t$ — касательная к параболе в точке $S$, так как она является биссектрисой $\angle FSR$.

С помощью точек параболы легко получать квадратные корни. Положим, $P = (0, 1)$, и в качестве директрисы $l$ рассмотрим прямую $y = -1$. Тогда нетрудно заметить, что соответствующая парабола задается уравнением $y = \frac{1}{4}x^2$. Касательная к этой параболе в точке $(x_0, \frac{1}{4}x_0^2)$ имеет

[1)] Фокус и директриса — это точка и прямая, связанные с параболой. Они характеризуются следующим свойством: парабола — это множество точек, равноудаленных от фокуса и директрисы.

угловой коэффициент $m = \frac{1}{2}x_0$, а значит уранение касательной имеет вид $y - \frac{1}{4}x_0^2 = \frac{1}{2}x_0(x - x_0)$. Пересечение этой прямой с прямой $x = 0$ назовем точкой $Q = (0, -\frac{1}{4}x_0^2)$. Тогда, выбрав в качестве точки $Q$ точку $(0, -\frac{1}{4}r)$, мы с помощью описанных выше конструкций можем получить на параболе построимую точку с абсциссой $\sqrt{r}$.

Теорема 5.1. *Построимые с использованием только первых пяти аксиом складывания из начальных точек* 0 *и* 1 *элементы поля* $\mathbb{C}$, *то есть поле евклидовых чисел* $\mathcal{E}$, *— наименьшее подполе* $\mathbb{C}$, *замкнутое относительно операции извлечения квадратного корня.*

Доказательство. То, что из всех построимых чисел можно извлечь квадратный корень, мы уже показали выше. Чтобы доказать, что ничего другого построить нельзя, воспользуемся тем фактом, что наименьшее подполе $\mathbb{C}$, замкнутое относительно операции извлечения квадратного корня — это поле построимых с помощью циркуля и линейки точек.[2)] Теперь для завершения доказательства теоремы достаточно показать, что действия описанные в любой из первых пяти аксиом складывания, можно осуществить с помощью циркуля и линейки. Заметим, что про первые 4 аксиомы это очевидно, так как построимое с помощью них поле было описано выше, и оно содержится в множестве построимых точек с помощью циркуля и линейки. Рассмотрим пятую аксиому и опишем соответствующее построение с помощью циркуля и линейки:

1. Строим окружность с центром в точке $Q$ и радиусом $QP$. Точка пересечения этой окружности с прямой $l$ есть точка $Q_1$.
2. Строим серединный перпендикуляр к отрезку $QQ_1$ — это и будет искомая прямая

Мы доказали вложение в обе стороны, а значит мы доказали утверждение теоремы. □

## 6. Конические конструкции и поле чисел оригами

Добавим шестую и последнюю аксиому складывания:

(A6) Если даны прямые $k, l$ и точки $P, Q$, то мы можем построить такую прямую $m$ (если она существует), что при симметрии относительно $m$, $P$ переходит на $k$, $Q$ переходит на $l$.

Чтобы построить такую прямую с помощью складывания бумаги, надо отразить $P$ в какую-то точку прямой $k$, а потом начать непрерывно двигать полученную точку до выполнения второго условия.

[2)] Другое доказательство теоремы 5.1 содержится в доказательстве теоремы 6.1.

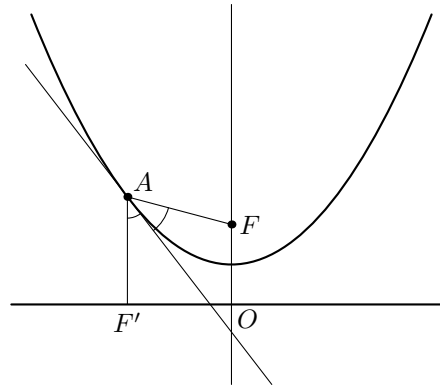



*Рис. 7.*

Предложение 6.1. *Аксиома* 6 *позволяет построить общую касательную к двум параболам, которые заданы своими фокусами и директрисами.*

Доказательство. Действительно, пусть две прямые, о которых идет речь в аксиоме шесть, — это директрисы, а точки — это фокусы парабол ($P$ и $l$ у одной, $Q$ и $k$ у другой). Тогда построенная в аксиоме прямая — это общая касательная этих двух парабол, так как у параболы есть следующее свойство: образ фокуса при отражении относительно касательной лежит на директрисе (см. рис. 7). □

Общая касательная существует не для всякой пары парабол, поэтому в условии аксиомы существование такой прямой оговаривается отдельно.

Аксиомы 1 – 6 называются аксиомами оригами. Множество чисел, которые можно получить с их использованием, — это множество чисел оригами $\mathcal{O}$. Эти конструкции позволяют нам построить вещественные решения кубических уравнений с вещественными коэффициентами из поля $\mathcal{O}$. Чтобы убедиться в этом, рассмотрим две параболы:

$$\left(y - \frac{1}{2}a\right)^2 = 2bx, \quad y = \frac{1}{2}x^2.$$

Эти параболы имеют фокусы и директрисы, построимые в поле, в котором построимы $a$ и $b$. Поэтому к ним можно провести общую касательную. Она пересечет их в точках $(x_0, y_0)$, $(x_1, y_1)$ соответственно. Из доказательства предложения 3.1 видно, что прямая построима только тогда, когда ее угловой коэффициент $\mu$ принадлежит полю или прямая вертикальна. Если продифференцировать уравнения парабол, то получим уравнения, определяющие касательные: $\dfrac{b}{y_0 - \frac{a}{2}} = \mu = x_1$ и $y_1 = \dfrac{1}{2}\mu^2$, $x_0 = \dfrac{(y_0 - \frac{a}{2})^2}{2b} = \dfrac{1}{2}\mu^2$.

Отсюда

$$\mu = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\frac{\mu^2}{2} - \frac{a}{2} - \frac{b}{\mu}}{\mu - \frac{b}{2}\mu^2}, \quad \text{т. е.}$$

$$\mu^3 + a\mu + b = 0.$$

Значит, мы можем найти все действительные корни любого кубического уравнения такого вида, если $a, b \in \mathcal{O} \cap \mathbb{R}$.

В частности, мы можем извлекать кубический корень из числа $b \in \mathcal{O} \cap \mathbb{R}$ с помощью уравнения $x^3 - b = 0$.

Кроме того, мы можем сделать трисекцию угла. Для этого достаточно воспользоваться уравнением Чебышёва: если $4x^3 - 3x = \cos(3\theta)$, то $x = \cos(\theta)$.

Отсюда следует, что мы можем извлекать кубические корни из построимых чисел (при извлечении кубического корня из комплексного числа угол уменьшается в три раза, а модуль меняется, как у вещественных чисел). Значит, корни кубических уравнений с коэффициентами из $\mathcal{O}$ также построимы, что видно из формулы для решения кубического уравнения (формула Кардано), которая включает в себя только извлечение квадратных и кубических корней из многочленов от коэффициентов уравнения. Корни уравнений четвертой степени также построимы, так как для них имеется формула Феррари, включающая в себя только извлечение квадратных и кубических корней (корень 4 степени — это квадратный корень, взятый от квадратного корня).

Теперь мы готовы доказать основную теорему:

Теорема 6.1. *Построимые из начальных точек* 0 *и* 1 *с помощью аксиом* 1–6 *точки в* $\mathbb{C}$, *то есть множество построимых чисел оригами* $\mathcal{O}$, *— это наименьшее подполе* $\mathbb{C}$, *замкнутое относительно операций извлечения квадратного и кубического корня.*

Доказательство. Заметим, что мы можем построить любую точку такого вида, как описано в условии теоремы (это было разобрано выше).

Докажем индукцией по построению, что других точек нет.

Исходно имеются элементы 0 и 1.

Предположим, что в результате $k$ применений аксиом 1–6 получаются лишь элементы из описанного в условии теоремы множества. Применим к полученным точкам одну из аксиом 1–6 и посмотрим, что происходит с координатами точек.

(A1) Пусть у нас были точки $A(0,0)$ и $B(a,b)$ (мы умеем сдвинуть первую точку в $(0,0)$, какая бы она ни была). Тогда прямая $l$, проходящая через эти точки, будет иметь уравнение: $bx - ay = 0$.

(A2) Пусть у нас есть прямые $l\colon l_1x + l_2y = l_3$ и $m\colon m_1x + m_2y = m_3$, тогда точка их пересечения

$$P = \left(\frac{m_2l_3 - l_2m_3}{m_1l_2 - m_2l_1}, \frac{m_1l_3 - l_1m_3}{m_1l_2 - m_2l_1}\right).$$

(A3) Если концы отрезка это точки $A(0,0)$ и $B(a,b)$, то серединный перпендикуляр задается уравнением: $ax - by = (a/2)^2 + (b/2)^2$.

(A4) Если стороны угла это прямые $l\colon l_1x + l_2y = 0$ и $m\colon m_1x + m_2y = 0$ (можно переместить точку пересечения в (0,0)), то уравнение биссектрисы будет:

$$\left(\frac{l_1}{\sqrt{l_1^2 + l_2^2}} + \frac{m_1}{\sqrt{m_1^2 + m_2^2}}\right)x + \left(\frac{l_2}{\sqrt{l_1^2 + l_2^2}} + \frac{m_2}{\sqrt{m_1^2 + m_2^2}}\right)y = 0.$$

(A5) Если даны точки $P = (a,b)$ и $Q = (0,0)$ (можем подвинуть точку) и прямая $l\colon x + y = -1$ (можем повернуть относительно $Q$ и сжать/растянуть), то прямая $m$, описываемая в условии аксиомы 5, будет задаваться уравнением: $(a+t)x - (b+t+1)y = 0$, где $t = \dfrac{-1 + \sqrt{2a^2 + 2b^2 - 1}}{2}$.

(A6) Проверим для шестой аксиомы. Точка в проективном пространстве — это набор из трех вещественных чисел $(x,y,z)$, не все из которых равны 0, причем для любого $c \neq 0$, точка $(x,y,z) \equiv (cx,cy,cz)$. Стандартная аффинная карта в таком пространстве — это карта $z = 1$ $(z \neq 0)$. Тогда бесконечно удаленная прямая — это множество точек, для которых $z = 0$. Однородное уравнение второй степени от трех переменных описывает на проективной плоскости точки коники. В матричном виде такое уравнение выглядит следующим образом:

$$F(x,y,z) = (x,y,z)A(x,y,z)^t = 0,$$

где $A$ — ненулевая симметрическая матрица $3 \times 3$.

Мы хотим доказать, что к двум параболам можно провести общую касательную, не выходя за пределы нашего поля. Для доказательства нам понадобится понятие проективной двойственности (более подробное изложение см в [1]). Для векторного пространства $V$ через $V^*$ обозначим двойственное ему векторное пространство. Проективные пространства $\mathbb{P}(V)$ и $\mathbb{P}(V^*)$ называются двойственными проективными пространствами. Геометрически, каждое из них — это пространство гиперплоскостей в другом, так как уравнение $\xi(v) = 0$, где $\xi \in V^*$, а $v \in V$ при фиксированном $\xi$ задает гиперплоскость в $\mathbb{P}(V)$, а при фиксированном $v$ — гиперплоскость в $\mathbb{P}(V^*)$. Двойственная кривая к кривой $\mathcal{F}$ — это

множество гиперплоскостей, касающихся $\mathcal{F}$. Построение общей касательной к двум кривым — это построение общей точки двойственных кривых. Двойственная кривая к конике $F(x, y, z) = 0$ задается уравнением $H(u, v, w) = (u, v, w)\mathrm{Adj}(A)(u, v, w)^t$, где $\mathrm{Adj}(A)$ — матрица, комплексно-сопряженная к $A$. Эта кривая тоже коника. Например, если $F(x, y, z) = \frac{1}{2}x^2 - yz = 0$, то уравнение двойственной коники — $H(u, v, w) = -uw + \frac{1}{2}v^2 = 0$. Заметим, что, если в матрице $A$ все элементы были из нашего поля, то и в матрице $\mathrm{Adj}(A)$ они тоже будут из нашего поля.

Для того, чтобы доказать, что построение общих касательных к кониках не выводит из нашего поля, мы можем перевести одну из двух прямых и соответствующую ей точку в прямую $y = -1$ и точку $(0, 1)$ с помощью поворота, гомотетии и параллельного переноса, которые не выводят за пределы поля (прямая $y = -1$ и точка $(0, 1)$ — это фокус и директриса параболы $y = \frac{1}{2}x^2$). Тогда вторая прямая и точка будут директрисой и фокусом некоторой параболы с коэффициентами из поля. Возьмем двойственные коники к этим двум параболам, у них тоже коэффициенты из поля, причем уравнение первой из них — $-uw + \frac{1}{2}v^2 = 0$, так что $u$ легко выразить через $v$ и $w$. Сделаем это и подставим в уравнение второй коники, получим уравнение четвертой степени с коэффициентами из $\mathcal{O}$, его решения тоже принадлежат $\mathcal{O}$. Но эти решения задают уравнения общих касательных к исходным параболам, а значит коэффициенты в уравнениях общих касательных из $\mathcal{O}$. □

## 7. О независимости аксиом

В этом разделе будут описаны некоторые факты про зависимость аксиом.

Предложение 7.1. *Аксиома* 4 *следует из аксиомы* 5.

Доказательство. Действительно, возьмем угол с вершиной $A$. Пусть точка $B$ лежит на стороне угла, а прямая $l$ — та его сторона, которая не содержит $B$. По пятой аксиоме мы можем через $A$ провести такую прямую $k$, что образ $B$ при симметрии относительно нее лежит на $l$. Но тогда несложно убедиться в том, что $k$ — биссектриса данного угла. □

Предложение 7.2. *Если даны аксиомы* 1, 2, 3 *и* 6, *то мы можем сделать построение, описанное в* 5.

Доказательство. Возьмем две точки $A$, $B$ и прямую $l$, о которых говорится в аксиоме 5. Проведем такую прямую $l'$, что $l'$ параллельна $l$ и $l'$ проходит через $B$ (это можно сделать, потому что у нас есть первые три аксиомы). Теперь по шестой аксиоме мы проведем такую прямую $m$, что образ $A$ при симметрии относительно нее попадет на $l$, а образ $B$ — на $l'$. Тогда имеется два случая. Либо $m$ перпендикулярна $l'$, то есть перпендикулярна $l$, но тогда образ $A$ никак не может попасть на $l$, если $A$ не лежит на $l$ (если лежит, то прямая, удовлетворяющая условию пятой аксиомы — это либо прямая, проходящая через $AB$, либо перпендикуляр к $l$, опущенный из $B$, но обе этих прямых мы можем построить). Либо $m$ проходит через $B$, тогда это и есть прямая, которая удовлетворяет условию пятой аксиомы. □

Предложение 7.3. *Если даны аксиомы* 2, 3 *и* 6, *то мы можем сделать построение, описанное в первой аксиоме.*

Доказательство. Опишем построение.

1. Построим серединный перпендикуляр $l$ (см рис. 8).

2. Воспользуемся шестой аксиомой, применив ее к точкам $P, Q$ и прямой $l$. Получим прямую $m$. Точка пересечения $l$ и $m$ — это точка $R$.

3. Проведем серединные перпендикуляры к отрезкам $PR$ и $QR$. Они пересекут прямую $m$ в точках $S$ и $T$ соответственно.

4. Проведем серединные перпендикуляры к отрезкам $SR$ и $TR$. Они пересекут прямую $l$ в точках $X$ и $Y$ соответственно.

5. Проведем серединный перпендикуляр к отрезку $XY$ — это и будет искомая прямая. □

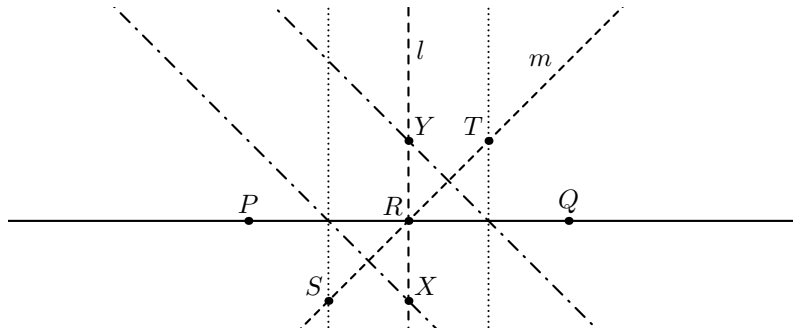



*Рис. 8.*

Теорема 7.1. (i) *Набор аксиом* 2, 3 *и* 6 *эквивалентен набору из всех шести аксиом. Кроме того это минимальный* (*нельзя выкинуть никакой аксиомы*) *набор, удовлетворяющий этому требованию.*

(ii) *Если у нас есть третья точка, не лежащая на прямой, проходящей через* 0 *и* 1, *и прямые, проходящие через любые две из этих трех точек, то минимальным будет набор аксиом* 2 *и* 6.

Доказательство. (i) Из предложений 7.1, 7.2 и 7.3 следует, что мы можем все построить. Проверим, что ни одну аксиому нельзя выкинуть.

(A2) Аксиома 2 — это единственная аксиома, которая описывает, как можно получить точку, а так как мы строим поле из координат точек, то она нам необходима.

(A3) Если убрать аксиому 3, то с помощью оставшихся трех аксиом мы не сможем выйти за пределы прямой, проходящей через 0 и 1.

(A6) Аксиома 6 расширяет поле фалесовских чисел, поэтому ее нельзя выкинуть.

(ii) Достаточно доказать, что можно построить серединный перпендикуляр к отрезку $AB$. Возьмем точку $A$ и две прямые: первая проходит через $A$ и $B$, а вторая проходит через $B$ и построимую точку, не лежащую на первой прямой (такая существует по условию). Тогда, по шестой аксиоме, мы можем провести такую прямую $l$, что точка $A$ при симметрии относительно $l$ попадет на первую прямую, а также попадет и на вторую прямую. Образ при симметрии единственен, значит образ — точка пересечения первой и второй прямых, то есть $l$ — серединный перпендикуляр к отрезку $AB$. □

Замечание. Если считать, что образы точек, упомянутых в аксиоме 6, построимы и есть третья построимая точка (кроме 0 и 1), не лежащая на прямой, проходящей через 0 и 1, и прямые, проходящие через любые две из этих трех точек, то можно обойтись только шестой аксиомой для построения поля оригами.

Действительно, возьмем две прямые, точку пересечения которых надо построить и произвольную построимую точку $P$. Применим шестую аксиому и получим, что образ $P$ — это точка пересечения двух данных прямых, то есть требуемая в условии точка, а она по условию построима. Значит, по теореме 7.1 достаточно только аксиомы 6.

Иногда к списку аксиом оригами добавляют еще одну.

(A7) Если дана точка $P$ и прямые $k$ и $l$, то мы можем провести такую прямую $m$, перпендикулярную $k$, что образ точки $P$ при симметрии относительно нее попадет на прямую $l$.

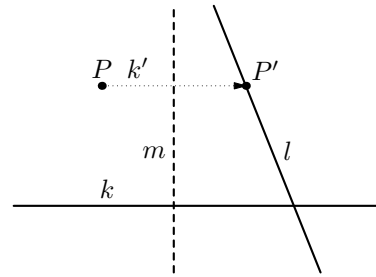



*Рис. 9. Избыточная аксиома 7*

На самом деле, добавление этой аксиомы никак не расширит поле построимых точек. Проверим это.

Предложение 7.4. *Аксиома* 7 *является следствием аксиом* 1–3.

Доказательство. Выполним с помощью первых трех аксиом действие, описанное в седьмой аксиоме.

1. Проведем через $P$ прямую $k'$, параллельную $k$ (см. рис. 9). Она пересечет $l$ в точке $Q$.
2. Построим серединный перпендикуляр $m$ к отрезку $PQ$ — это и будет искомая прямая.

Действительно, образ точки $P$ при симметрии относительно $m$ попадет в точку $Q$. Кроме того, так как прямая $k'$ параллельна $k$, то $m \perp k$. □

## Список литературы

[1] Городенцев А. Л. *Алгебра–1. Учебник для студентов-математиков первого курса*. М.: ВШЭ, 2011. 526 с.

[2] Кириченко В. А. *Построения циркулем и линейкой и теория Галуа*. 2005. `www.mccme.ru/∼valya/dubna05.pdf`

[3] Alperin R. C. *A mathematical theory of origami constructions and numbers* // New York Journal of Mathematics. Vol. 6. 2000. P. 119–133.

[4] Alperin R. C., Lang R. J. *One-, Two-, and Multi-Fold Origami Axioms* // Origami (Robert J. Lang, ed.). London: A K Peters Ltd. 2009. P. 371–394.

[5] Lang R. J. *Origami and Geometric Constructions*. 2003. `www.langorigami.com/science/math/hja/hja.php`

---

Д. И. Грищенко, факультет математики НИУ-ВШЭ

# Наш семинар: математические сюжеты

## Короткое опровержение гипотезы Борсука

А. Б. Скопенков*

Приводится простейшее из известных опровержений следующей гипотезы Борсука: *любое ограниченное подмножество $n$-мерного евклидова пространства, содержащее более $n$ точек, можно разбить на $n+1$ непустых частей меньшего диаметра.*

Доказательство принадлежит Н. Алону и является замечательным приложением комбинаторики и алгебры к геометрии.

Эта методическая заметка доступна студентам, старшеклассникам и учителям, интересующимся математикой.

Теорема 1 (Борсук). *Любое ограниченное подмножество плоскости, в котором более двух точек, можно разбить на три непустые части меньшего диаметра.*[1)]

*Диаметром* непустого подмножества плоскости называется наибольшее расстояние между его точками (точнее, супремум таких расстояний). Подмножество плоскости называется *ограниченным*, если его диаметр конечен.

*Точкой* $x = (x_1, \dots, x_n)$ $n$-мерного евклидова пространства называется упорядоченный набор $n$ чисел. *Расстояние* между точками $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$ определяется формулой

$$|x, y| := \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}\,.$$

*Поддержан грантом фонда Саймонса.

[1)] *Указание к доказательству.* Сначала, используя «соображения непрерывности», докажите, что любую плоскую фигуру диаметра 1 можно заключить в правильный шестиугольник, диаметр вписанной окружности которого равен 1. Затем докажите, что хотя диаметр полученного правильного шестиугольника больше 1, его можно разрезать на три части диаметра меньше 1. Ср. [11].

*Диаметр* и *ограниченность* подмножества $n$-мерного евклидова пространства определяются точно так же, как и в случае плоскости.

Борсук предложил следующее обобщение своего результата, которое долгие годы было одной из наиболее интригующих проблем комбинаторной геометрии.

Гипотеза 1 (Борсук). *Любое ограниченное подмножество $n$-мерного евклидова пространства, содержащее более $n$ точек, можно разбить на $n+1$ непустых частей меньшего диаметра.*

В 1993 г. Д. Кан и Дж. Калаи, следуя идеям Болтянского, Эрдёша и Лармана о применении комбинаторики для построения контрпримера, нашли контрпример к гипотезе Борсука [7, 10]. Подробно история вопроса описана в [3, 6].

Теорема 2. *Существует $n$ и ограниченное подмножество $n$-мерного евклидова пространства, содержащее более $n$ точек и которое невозможно разбить на $n+1$ часть меньшего диаметра.*

Мы приведем простейшее из известных доказательств, принадлежащее Н. Алону, ср. [1, 3, 5, 6, 8, 9]. (При этом другие доказательства дают более сильные результаты.) Это удивительный пример важного результата в современной математике, не требующего для полного понимания полугодового специального университетского курса (после двухгодового обязательного курса). Более простые применения аналогичных алгебраических соображений в комбинаторике можно найти в [2, 4].

Через $|X|$ обозначается число элементов в множестве $X$. *Скалярное произведение* векторов $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ определяется как $x \cdot y := x_1y_1 + \cdots + x_ny_n$. Векторы $x$ и $y$ называются *ортогональными*, если $x \cdot y = 0$.

Доказательство теоремы 2. Обозначим

$$M = \{(x_1, \ldots, x_n) \mid x_1 = 1,\ x_k \in \{1, -1\} \text{ и среди } x_2, \ldots, x_n \text{ число минус единиц четно}\}.$$

Вершина $n^2$-мерного куба — набор длины $n^2$ из плюс или минус единиц. Его удобно представлять себе как таблицу $n \times n$. Поставим в соответствие каждой точке $x = (x_1, \ldots, x_n) \in M$ таблицу $fx$, определенную формулой $(fx)_{ij} := x_ix_j$. Например,

$$f(1, -1, -1) = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Докажем, что контрпримером к гипотезе Борсука является $f$-образ множества $M$ для достаточно большого простого числа $p$ и $n = 4p$.

Пусть $x, y \in M$. Тогда $(x_ix_j - y_iy_j)^2 = (1 - x_iy_ix_jy_j)^2$. Обозначим через $a$ количество индексов $i$, для которых $x_i = y_i$. Тогда $x_iy_i = 1$ для $a$ индексов $i$ и $x_iy_i = -1$ для $n - a$ индексов $i$. Поэтому $|fx, fy|^2 = 4a(n - a)$. Это выражение максимально при $a = n/2$. Значит, условие $|fx, fy| = \operatorname{diam} fM$ равносильно условию $a = n/2$ и равносильно условию $x \cdot y = 0$.

Поэтому если множество $fM$ разбито на $k$ частей $Z_1, \dots, Z_k$ меньшего диаметра, то в каждом $f^{-1}Z_i$ никакие два вектора не ортогональны. Так как $x_1 = 1$ для любого $x \in M$, то $f$ инъективно. Значит, $|Z_i| = |f^{-1}Z_i|$. Теперь теорема вытекает из следующих леммы 1 об оценке и утверждения 1. □

Лемма 1 (оценка). *Пусть $p$ — простое (не обязательно большое), $n = 4p$, $A \subset M$ и никакие два вектора из $A$ не ортогональны. Тогда*

$$|A| \leqslant \alpha(n) := C_{n-1}^0 + C_{n-1}^1 + \dots + C_{n-1}^{n/4-1}.$$

Утверждение 1. *$\alpha(n)(n^2 + 1) < |M| = 2^{n-2}$ для достаточно больших $n$.*

Доказательство. Для достаточно больших $n$ и любых $1 \leqslant s, k \leqslant n/4$ имеем $\frac{5n}{4} > 5k - s - 2$, откуда $\frac{n-k-s+1}{\frac{n}{4}+k-s} > \frac{3}{2}$. Значит,

$$\frac{C_{n-1}^{n/4+k-1}}{C_{n-1}^{k-1}} = \frac{(n-k)(n-k-1)\cdot\ldots\cdot(n-k-\frac{n}{4}+1)}{(\frac{n}{4}+k-1)(\frac{n}{4}+k-2)\cdot\ldots\cdot k} > \left(\frac{3}{2}\right)^{n/4} > n^2.$$

Поэтому $\alpha(n)(n^2 + 1) < C_{n-1}^0 + C_{n-1}^1 + \dots + C_{n-1}^{n/2-1} = 2^{n-2}$. □

Осталось доказать лемму об оценке. При ее доказательстве можно забыть про конструкцию отображения $f$. Следующее утверждение очевидно.

Утверждение 2. *Для простого $p$ и целого $t$ число*

$$G(t) := (t-1)(t-2)\cdot\ldots\cdot(t-p+1)$$

*делится на $p$ тогда и только тогда, когда $t$ не делится на $p$.*

*Рациональной линейной комбинацией* многочленов $F_1, \dots, F_s$ называется любой многочлен $\lambda_1F_1 + \dots + \lambda_sF_s$ с рациональными $\lambda_1, \dots, \lambda_s$. Например, многочлен $x_2$ является рациональной линейной комбинацией многочленов $2x_1$, 1 и $x_1 + x_2$.

Многочлены называются *линейно независимыми*, если любая их рациональная линейная комбинация, в которой не все $\lambda_k$ нулевые, не равна нулю. Например, $n$ многочленов 1, $x_2$, $x_3$, …, $x_n$ являются линейно независимыми.

Многочлен с рациональными коэффициентами от $n - 1$ переменной $x_2, \dots, x_n$ называется *степени менее $n/4$ и свободным от квадратов*, если

он является рациональной линейной комбинацией многочленов

$$\begin{aligned} &x_{i_1} \cdot \ldots \cdot x_{i_s}, \quad \text{где} \quad s = 0, \ldots, p-1 \\ &\text{и} \quad i_1, \ldots, i_s \ \text{— различные числа от 2 до } n. \end{aligned} \tag{$*$}$$

Лемма об оценке вытекает из нижеследующих леммы 2 о линейной независимости и утверждения 3.

Лемма 2 (линейная независимость). *Пусть $p$ простое, $n = 4p$, $A \subset M$ и никакие два вектора из $A$ не ортогональны. Возьмем вектор $a \in A$. Раскроем скобки в произведении $G(a \cdot (1, x_2, \ldots, x_n))$. В каждом из полученных одночленов для каждого $i$ будем заменять $x_i^2$ на 1 пока это возможно. Полученный многочлен обозначим $F_a(x_2, \ldots, x_n)$. Тогда каждый многочлен $F_a(x_2, \ldots, x_n)$, $a \in A$, степени меньше $n/4$ и свободен от квадратов; эти многочлены линейно независимы.*

Утверждение 3. *Любое линейно независимое семейство многочленов от $x_2, \ldots, x_n$ степени менее $n/4$ и свободных от квадратов, содержит не более $\alpha(n)$ многочленов.*

Доказательство леммы о линейной независимости. Утверждения о степени и о свободе от квадратов очевидны. Докажем линейную независимость. Пусть, напротив, $\lambda_1 F_{a_1} + \cdots + \lambda_s F_{a_s} = 0$ для некоторых $a_1, \ldots, a_s \in A$ и рациональных $\lambda_1, \ldots, \lambda_s$, причем не все $\lambda_k$ нулевые. Здесь $a_1, \ldots, a_s$ — векторы, а не координаты. Можно считать, что $\lambda_1, \ldots, \lambda_s$ целые (иначе умножим это равенство на произведение их знаменателей). Можно также считать, что не все они делятся на $p$ (иначе поделим это равенство на их наибольший общий делитель). Не уменьшая общности считаем, что $\lambda_1$ не делится на $p$. Подставим в полученное равенство значения $x_2 = (a_1)_2, \ldots, x_n = (a_1)_n$.

Из $a_1 \cdot a_1 = n = 4p$ и утверждения 2 вытекает, что $\lambda_1 F_{a_1}$ не делится на $p$.

Так как $n$ делится на 4 и для любых $x, y \in M$ число минус единиц в $x$ и в $y$ нечетно, $x \cdot y$ делится на 4. Поэтому $x \cdot y \notin \{\pm p, \pm 2p, \pm 3p\}$. Так как $x \cdot y \neq 0$, то $x \cdot y$ не делится на $p$. Значит, по утверждению 2 $\lambda_k F_{a_k}$ делится на $p$ при любом $k > 1$. Противоречие. □

Набросок доказательства утверждения 3. Обозначим через $Q_1, \ldots, Q_{\alpha(n)}$ семейство многочленов $(*)$ и через $F_1, \ldots, F_k$ данное линейно независимое семейство. Возьмем таблицу $k \times \alpha(n)$ рациональных чисел $\lambda_{ij}$, для которых $F_i = \sum_j \lambda_{ij} Q_j$ при любом $i = 1, \ldots, k$. Семейство многочленов, полученное из семейства $F_1, \ldots, F_k$ заменой $F_i$ на $F_i + \lambda F_j$, $j \neq i$, линейно независимо. Такими заменами и перестановками многочленов $Q_1, \ldots, Q_{\alpha(n)}$ можно провести рассматриваемую таблицу $k \times \alpha(n)$ к «верхнетреугольному» виду. Так как в новой таблице нет нулевой строки, то $k \leqslant \alpha(n)$. □

Благодарности. Благодарю Н. П. Долбилина и А. М. Райгородского, от которых я узнал контрпримеры к гипотезе Борсука, учеников физ.-мат. школы им. А. Н. Колмогорова и школы №57 г. Москвы, которые узнали эти контрпримеры от меня, а также М. Б. Ахмедова, В. Н. Дубровского и А. Д. Руховича за полезные обсуждения.

## Список литературы

[1] Гервер М. Л. *О разбиении множеств на части меньшего диаметра: теоремы и контрпримеры* // Мат. Просвещение. Сер. 3. Вып. 3. 1999. С. 168–183.

[2] Ильинский Д., Купавский А., Райгородский А., Скопенков А. *Дискретный анализ для математиков и программистов (подборка задач)* // Мат. Просвещение. Сер. 3. Вып. 17. 2013. С. 162–181.

[3] Райгородский А. М. *Проблема Борсука*. М.: МЦНМО. 2006.

[4] Райгородский А. М. *Линейно-алгебраический метод в комбинаторике*. М.: МЦНМО, 2007.

[5] Скопенков А. *$N$-мерный куб, многочлены и решение проблемы Борсука* // Мат. Просвещение. Сер. 3. Вып. 3. 1999. С. 184–188. Эл. версия arXiv:0712.4009v1.

[6] Aigner M., Ziegler G. *Proofs from the Book*. NY, Berlin, Heidelberg: Springer. 2004. Рус. пер. Айгнер М., Циглер Г. *Доказательства из Книги*. М.: Мир. 2006.

[7] Kahn J., Kalai G. *A counterexample to Borsuk's conjecture* // Bull. AMS. Vol. 29(1). 1993. P. 60–62.

[8] Nilli A. *On Borsuk's problem* // Contemp. Math. Vol. 178. 1994. P. 209–210.

[9] Raigorodskii A. M. *The Borsuk partition problem: the seventieth anniversary* // Math. Intelligencer. Vol. 26 (3). 2004. P. 4–12.

[10] Skopenkov A. *The Borsuk problem* // Quantum. Vol. 7 (1). 1996. P. 16–21, 63.

[11] Yang D. *An elementary proof of Borsuk theorem*. arXiv:1010.1990. 2010.

---

А. Б. Скопенков, Московский физико-технический институт (государственный университет), Независимый московский университет и ГБОУ Центр педагогического мастерства
Email: skopenko@mccme.ru
Личная страница: www.mccme.ru/∼skopenko

# Полиномы Чебышёва и их обращения

А. Г. Хованский

Полином Чебышёва степени $n$ определяется следующей формулой:

$$T_n(x) = \cos n \arccos x.$$

Эти полиномы были открыты Чебышёвым в связи с задачей о наилучшем приближении заданной функции полиномами степени $\leqslant n$. Они играют большую роль в теории приближений. Удивительно, что эти полиномы оказались полезными и в алгебре: ведь задача, в связи с которой они возникли, от алгебры далека, а их исходное определение использует трансцендентные функции.

Тем не менее в ряде задач алгебры наряду с серией степенных полиномов $x^n$ встречается серия полиномов $T_n$. С «философской» точки зрения появление этих двух серий полиномов связано с существованием двух серий конечных групп проективных преобразований пространства $\mathbb{C}P^1$: циклических групп $C_n$ и групп диэдра $D_n$.

В комплексном анализе серия полиномов $x^n$ расширяется до семейства многозначных аналитических функций $x^\alpha$, $\alpha \in \mathbb{R}$, содержащего, наряду с полиномами $x^n$, их обращения $x^{1/n}$ и удовлетворяющего прежним композиционным соотношениям $(x^\alpha)^\beta = x^{\alpha\beta}$.

Аналогично мы расширяем серию полиномов Чебышёва $T_n$ до семейства многозначных аналитических функций $T_\alpha$, $\alpha \in \mathbb{R}$, содержащего, наряду с полиномами $T_n$, их обращения $T_{1/n}$ и удовлетворяющего прежним композиционным соотношениям $T_\beta \circ T_\alpha = T_{\alpha\beta}$.

Многозначную функцию можно определить без использования аналитического продолжения, описав множество ее значений в каждой точке. Это иногда дает возможность перенести определение функции на любое поле (над которым операция аналитического продолжения не определена). Например, при натуральном $n$ функция $x^{1/n}$ определена над любым полем $\Bbbk$: это многозначная функция, которая сопоставляет $x \in \Bbbk$ множество элементов $z$, лежащих в замыкании поля $\Bbbk$ и таких, что $z^n = x$.

Легче иметь дело с ростком однозначной функций, чем с многозначной функцией. Во многих вопросах этим можно ограничиться, если все значения многозначной функции получаются при аналитическом продолжении однозначного ростка.

В п. 1.1 определяется многозначная функция Чебышёва $T_\alpha$, $\alpha \in \mathbb{R}$, комплексного переменного $x$ при помощи описания множества ее значений. В п. 1.2 определяется ряд в точке $x = 1$, аналитическим продолжением которого она является (см. п. 1.3).

В п. 2.1 мы приводим алгебраическое определение полиномов Чебышёва и их обращений над любым полем, характеристика которого $\neq 2$. Если, дополнительно, характеристика поля $\neq 3$, то эти функции применимы для решения в радикалах уравнений степени три и степени четыре над этим полем (см. пп. 2.2–2.3).

В пп. 3.1–3.3 мы обсуждаем три классические задачи, в решении которых встретились серии полиномов $x^n$ и $T_n$. В п. 3.1 обсуждается решенная Риттом задача об описании всех комплексных полиномов, обращения которых представимы в радикалах. В п. 3.2 обсуждается решенная Фридом проблема Шура об описании всех полиномов $P \in \mathbb{Q}[x]$, для которых отображения $P\colon \mathbb{Z}_p \to \mathbb{Z}_p$ обратимы для бесконечного множества простых чисел $p$. В п. 3.3 мы формулируем результат Жулиа, Фату и Ритта об аффинной классификации *интегрируемых* (см. определение из этого пункта) полиномиальных отображений комплексной прямой в себя.

## §1. Функции Чебышёва над комплексными числами

### 1.1. Многозначные функции Чебышёва

*Функцией Чебышёва степени* $\alpha \in \mathbb{R}$ назовем многозначную функцию $T_\alpha$ комплексного переменного $x$, определенную соотношением:

$$T_\alpha(x) = \frac{u^\alpha(x) + u^{-\alpha}(x)}{2}\,, \tag{1}$$

где $u$ — двузначная функция, определенная соотношением

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

В формуле (1) имеется в виду, что каждое значение $f(x)$ многозначной функции $u^\alpha(x)$ складывается со значением $f^{-1}(x)$ функции $u^{-\alpha}(x)$ (а не с каким-либо другим ее значением). Согласно (2) функция $u(x)$ удовлетворяет уравнению $u^2(x) - 2xu(x) + 1 = 0$. Его корни $u_1(x)$, $u_2(x)$ связаны соотношением $u_1(x)u_2(x) = 1$, поэтому не важно, какой из двух корней использовать в формуле (1). (Отметим, что эти корни вычисляются явно: $u_{1,2}(x) = x \pm \sqrt{x^2 - 1}$.) Выбор другого корня лишь переставляет слагаемые $u^\alpha(x)$ и $u^{-\alpha}(x)$ и не меняет их суммы.

Теорема 1. *Функцию* $T_\alpha$ *можно определить соотношениями*:

$$x = \cos z(x), \quad T_\alpha(x) = \cos \alpha z(x).$$

Доказательство. Если $x = \cos z_0$, то $z(x) = \pm(z_0 + 2k\pi)$ и

$$\cos(\alpha z(x)) = \frac{\exp(i\alpha z(x)) + \exp(-i\alpha z(x))}{2}.$$

При этом $u_{1,2}(x) = \exp(\pm iz(x))$ и $u_{1,2}^{\pm\alpha}(x) = \exp i\alpha(\pm z(x))$. Откуда и вытекает теорема.

Утверждение 2. *Функция $T_n$ для натурального $n$ является полиномом степени $n$ с целыми коэффициентами. Справедлива формула*

$$T_n(x) = \sum_{0\leqslant k\leqslant[n/2]} \binom{n}{2k} x^{n-2k}(x^2-1)^k.$$

Доказательство. Соотношение $T_n(x) = \left(u^n(x) + u^{-n}(x)\right)/2$ с учетом равенств $u^n(x) = (x + \sqrt{x^2-1})^n$ и $u^{-n}(x) = (x - \sqrt{x^2-1})^n$ и бинома Ньютона превращается в формулу для $T_n(x)$.

Определение. Функция $T_n$ называется *полиномом Чебышёва степени $n$*.

Справедливо тождество $T_n(\cos z) = \cos nz$ (см. теорему 1). Полином Чебышёва можно определить, пользуясь этим тождеством (собственно, так и сделал сам Чебышёв). Полином $T_n$ является четной функцией при четном $n$ и нечетной функцией при нечетном $n$. Старший коэффициент полинома $T_n$ равен $2^n$. Ниже нам понадобится формула $T_3(x) = 4x^3 - 3x$.

Следствие 3. *Уравнение $T_n(x) = a$ явно решается в радикалах. Именно, его корни — значения в точке $a$ многозначной функции $T_{1/n}(a)$.*

Доказательство. Если $\cos z = a$ и $x = \cos\dfrac{z}{n}$, то $x = T_{1/n}(a)$. С другой стороны, в этом случае $T_n(x) = a$.

Эта «тригонометрическая» выкладка переносится в алгебру и позволяет решить уравнение $T_n(x) = a$, где $a$ — элемент поля, характеристика которого не равна двум (см. п. 1.4). Отметим, что $T_{1/n}$ — $n$-значная функция: выбор значения функции $u(a)$ не меняет значений $T_\alpha(a)$, а функция $u^{1/n}(a)$ принимает $n$ значений.

## 1.2. Ростки функций Чебышёва в единице

Многозначная функция $T_\alpha(x)$, так же как и степенная функция $x^\alpha$, имеет выделенный росток в точке $x = 1$, значение которого равно 1. С однозначными ростками легче иметь дело, чем с их многозначными аналитическими продолжениями. Ниже символом $x^\alpha$ мы обозначаем росток $\sum \dfrac{\alpha\cdot\ldots\cdot(\alpha-k+1)}{k!}(x-1)^k$.

Свойства ростков степенных функций в единице:

1) *свойство композиции*: если $f = x^\alpha$ и $g = x^\beta$, то $f \circ g = x^{\alpha\beta}$; другими словами, $(x^\beta)^\alpha = x^{\alpha\beta}$;

2) *свойство мультипликативности*: $x^\alpha x^\beta = x^{\alpha+\beta}$;

3) *свойство алгебраичности*: для $\alpha = 1/n$, где $n$ — натуральное число, росток $z = x^\alpha$ удовлетворяет алгебраическому уравнению $z^n = x$.

Аналитические ростки, инвариантные при инволюции.

Инволюция $\tau$ комплексной прямой $\tau(u) = u^{-1}$ переводит точку $u = 1$ в себя. Легко описать все ростки $f$ аналитических функций в этой точке, инвариантные относительно инволюции $\tau$, т. е. такие, что $f = f(\tau)$.

Утверждение 4. *Равенство $f = f(\tau)$ справедливо, если и только если $f(u) = \varphi(x)$, где $x = (u + u^{-1})/2$ и $\varphi$ — росток аналитической функции в точке $x = 1$.*

Доказательство. Если $f = f(\tau)$, то функция $\varphi(x) = f(u(x))$, где $u(x)$ — одна из двух ветвей функции, определенной уравнением $(u(x) + u^{-1}(x))/2 = x$, не зависит от выбора ветви и аналитична в проколотой окрестности точки $x = 1$. По теореме об устранимой особенности она аналитична и в этой точке тоже.

Ростки аналитических функции от $u$, не инвариантные относительно инволюции $\tau$, задают *двузначные ростки Пьюизо* от $x$.

*Ростком функции Чебышёва $T_\alpha$ в точке $x = 1$* мы будем называть росток аналитической функции от $x$, такой, что росток функции $\dfrac{u^\alpha + u^{-\alpha}}{2}$ (инвариантный при инволюции $\tau$) равен $T_\alpha(x(u))$, где $x(u) = (u + u^{-1})/2$. В этом пункте мы будем обозначать росток функции Чебышёва тем же символом $T_\alpha$, что и саму многозначную функцию. Ростки $T_\alpha$ наследуют свойства ростков степенных функций.

Свойства ростков функции Чебышёва в единице:

1) *свойство композиции*: $T_\alpha \circ T_\beta = T_{\beta\alpha}$;

2) *свойство мультипликативности*: $T_\alpha T_\beta = (T_{\alpha+\beta} + T_{\alpha-\beta})/2$;

3) *свойства алгебраичности*: для $\alpha = n$, где $n$ — натуральное число, росток $T_\alpha$ является ростком полинома Чебышёва $T_n$. Росток $T_{1/n}$ удовлетворяет алгебраическому уравнению $T_n(T_{1/n}(x)) = x$;

4) *тригонометрическое свойство*: $T_\alpha(\cos z) = \cos \alpha z$. Под этим равенством мы подразумеваем равенство ростков функций от $z$ в точке $z = 0$. Суперпозиция $T_\alpha(\cos z)$ определена, так как $\cos 0 = 1$.

Утверждение 5. *Семейство ростков функций Чебышёва в единице удовлетворяет свойствам* 1)–4).

Доказательство. 4) следует из теоремы 1. Это свойство полностью характеризует росток $T_\alpha$. Действительно, функция $\cos z$ четная. По теореме о неявной функции росток в нуле функции $z^2$ является аналитической функцией от ростка в единице функции $\cos z$. В свою очередь функция $\cos \alpha z$ — аналитическая функция от $z^2$. 1)–3) — это простые свойства функции cos: 1) если $\cos v = \cos \beta z = T_\beta(\cos z)$, то $\cos \alpha v = T_\alpha(\cos v)$ и $T_\alpha T_\beta \cos z = \cos \alpha\beta z$; 2) вытекает из тождества $\cos \alpha z \cos \beta z = [\cos((\alpha+\beta)z) + \cos((\alpha-\beta)z)]/2$; 3) для $\alpha = n$ доказано в утверждении 2, для $\alpha = \dfrac{1}{n}$ вытекает из свойства композиции.

### 1.3. Аналитическое продолжение ростков

В этом пункте мы покажем, что множество значений многозначной функции, порожденной ростком $T_\alpha$, согласуется с определением из п. 1.1.

Обращение ростка в нуле функции $\cos z$ — двузначный росток Пьюизо в точке $x = 1$, значения которого различаются знаком. Пусть $\pi^{-1}(x)$ — одно из двух различающихся знаком многозначных обращений функции $\cos z = x$, имеющих в точке $x = 1$ этот росток Пьюизо. Рассмотрим четную функцию $\Phi_\alpha(z) = \cos \alpha z$ переменной $z$. По определению $T_\alpha = \Phi_\alpha \circ \pi^{-1}$.

Функция $\cos z$ имеет некратные критические точки $z = k\pi$ и два критических значения $x = \pm 1$. Скажем, что кривая $x(t)$, идущая из точки 1 в точку $x_0$, т. е. $x(0) = 1$, $x(1) = x_0$, *допустима*, если $x(t) \neq \pm 1$ при $0 \leqslant t \leqslant 1$. Росток Пьюизо в точке $x = 1$ функции $\pi^{-1}$ в следующем смысле продолжается вдоль допустимой кривой $x(t)$, идущей из $x = 1$ в точку $x_0$: 1) любая из двух ветвей ростка аналитически продолжается вдоль $x(t)$ вплоть до $t = 1$, если $x_0 \neq \pm 1$, и вплоть до любого $t < 1$, если $x_0 = \pm 1$. В последнем случае продолжение до $t = 1$ — двузначный росток Пьюизо в точке $x_0 = \pm 1$ (ветви которого в $x_0$ совпадают).

Росток $T_\alpha = \Phi_\alpha \circ \pi^{-1}$ в этом же смысле продолжается вдоль любой допустимой кривой $x(t)$. Росток $T_\alpha$ регулярен и однозначен (а не двузначен, как $\pi^{-1}$), поэтому он имеет *единственное продолжение* вдоль допустимой кривой. Для некоторых допустимых кривых, идущих из точки $x = 1$ в точку $x = k\pi$, результат продолжения тоже может оказаться аналитическим ростком (а не двузначным ростком Пьюизо).

Покажем, что формулы (1), (2) описывают все значения многозначной функции, полученной продолжением ростка $T_\alpha$. Пусть $x_0$ и $a = T_\alpha(x_0)$ — любые числа, удовлетворяющие (1), (2).

Утверждение 6. *Существует допустимая кривая $x(t)$, идущая из точки $x=1$ в точку $x_0$, такая, что аналитический росток (или росток Пьюизо), полученный продолжением ростка $T_\alpha$ вдоль $x(t)$, принимает в точке $x_0$ значение $a$, определенное выше.*

Доказательство. Выберем $z_0$ так, чтобы $\exp iz_0 = u(x_0)$, $\exp(\alpha i z_0) = u^\alpha(x_0)$. Пусть $z(t)$ — кривая, такая, что $z(0)=0$, $z(1)=z_0$ и $z(t)$ не проходит через точки $z=k\pi$ при $0<t<1$. Тогда кривая $x(t)=\cos z(t)$ допустима, идет из точки $x=1$ в точку $x_0$ и аналитическое продолжение вдоль этой кривой ростка $T_\alpha = \cos\alpha(\cos^{-1})$ дает росток, принимающий в точке $x_0$ значение $a$.

Для нас особенно важны полиномы Чебышёва $T_n$ и функции $T_{1/n}$, обратные к ним. Благодаря утверждению 6, мы имеем *описание множества значений функции* $T_{1/n}$ в точке $a$. Пусть $u_1$, $u_2$ — корни уравнения $\dfrac{u+u^{-1}}{2}=a$ (достаточно взять один из этих корней). Пусть $\{v_{i,j}\}$ — корни уравнения $v^n=u_i$, где $i=1,2$; $1\leqslant j\leqslant n$. Множество $\{T_{1/n}(a)\}$ всех значений функции в точке $a$ равно множеству $\left\{\dfrac{v_{1,j}+v_{i,j}^{-1}}{2}\right\}$ и множеству $\left\{\dfrac{v_{2,j}+v_{2,j}^{-1}}{2}\right\}$.

## §2. Функции Чебышёва над полями

### 2.1. Алгебраическое определение

Полином Чебышёва $T_n\in\mathbb{Z}[x]$ определен над любым полем $\Bbbk$. Если характеристика поля равна нулю, то $\mathbb{Z}\subseteq\Bbbk$ и $T_n\in\Bbbk[x]$. Если поле имеет характеристику $p>0$, то $\mathbb{Z}_p\subseteq\Bbbk$ и полином, полученный из $T_n$ приведением его коэффициентов по модулю $p$ (который мы будем обозначать тем же символом $T_n$), принадлежит $\Bbbk[x]$. Если $p\neq 2$, то $\deg T_n = n$, так как старший коэффициент полинома $T_n$ равен $2^{n-1}$.

Утверждение 7. *Если характеристика поля $\Bbbk$ не равна двум, то в поле рациональных функций $\Bbbk(x)$ справедливо тождество*

$$T_n\left(\frac{x+x^{-1}}{2}\right)=\frac{x^n+x^{-n}}{2}. \tag{3}$$

Доказательство. Вытекает из формул (1), (2).

Следствие 8. *Если характеристика поля $\Bbbk$ не равна двум, то уравнение $T_n(x)=a$ над полем $\Bbbk$, где $a\in\Bbbk$, явно решается в радикалах.*

Доказательство. В тождество (3) подставим $x=(v+v^{-1})/2$. Получим $(v^n+v^{-n})/2=a$. Решим квадратное уравнение $u^2-2au+1=0$

для $u = v^n$. Пусть $u_1$, $u_2$ — его корни и $\{v_{1,j}\}$ — множество всех корней степени $n$ из $u_1$. Тогда элементы $v_{2,j} = v_{1,j}^{-1}$ образуют множество всех корней степени $n$ из $u_2$, так как $u_1 u_2 = 1$. Все корни уравнения $T_n(x) = a$ представимы в виде $x = (v_{1,j} + v_{1,j}^{-1})/2$, а также в виде $x = (v_{2,j} + v_{2,j}^{-1})/2$.

Доказательство следствия 8 показывает, что *уравнение* $T_n(x) = a$ над полем $\Bbbk$, характеристика которого не равна двум, *решается явно при помощи формулы* $x = T_{1/n}(a)$, которая имеет смысл и над полем $\Bbbk$.

## 2.2. Уравнения степени три

Пусть $F$ — полином степени $n$ над полем $\Bbbk$, характеристика которого или равна нулю, или больше чем $n$. Положим $Q(y) = aF(\lambda y + x_0)$, где $a \neq 0$, $\lambda \neq 0$ и $x_0$ — элементы поля $\Bbbk$ или его расширения. При сделанных предположениях о характеристике поля $\Bbbk$ имеем

$$Q(y) = \sum \frac{a\lambda^k F^{(k)}(x_0)}{k!} y^k.$$

Линейная функция $Q^{(n-1)}$ обращается в нуль в некоторой точке $q$. Положим $x_0 = q$, тогда коэффициент в $Q$ при $y^{n-1}$ обратится в нуль. Меняя $a$ и $\lambda$, можно добиться, чтобы два ненулевых коэффициента полинома $Q$ приняли заданные ненулевые значения.

Описанным преобразованием полином $F(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ можно привести либо к виду $y^3 + c$, либо к виду $4y^3 - 3y + c$. Полином $F''$ обращается в нуль в точке $x_0 = -a_2/3a_3$. Возможны два случая:

1) $F'(x_0) = 0$. В этом случае полином $F$ приводится к виду $y^3 + c$ преобразованием $aF(y + x_0)$, где $a = a_3^{-1}$. При этом $c = F(x_0)a$.

2) $F'(x_0) \neq 0$. В этом случае полином $F$ приводится к виду $4y^3 - 3y + c$ преобразованием $aF(\lambda y + x_0)$, где

$$\lambda = (-4F'(x_0)/3a_3)^{1/2}; \quad a = -3(\lambda F'(x_0))^{-1}.$$

При этом $c = F(x_0)a$. (Знак $\lambda$ можно выбрать любым: мы ищем одно преобразование, обладающее нужным свойством, а не описываем их все.)

Следствие 9. *Кубическое уравнение* $F(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ *над полем* $\Bbbk$, *характеристика которого не равна двум и трем, следующим образом решается в радикалах. Пусть* $x_0 = -a_2/3a_3$ — *корень полинома* $F''$. *Тогда*:

1) *если* $F'(x_0) = 0$, *то* $x = x_0 + (-F(x_0)/a_3)^{1/3}$;

2) *если* $F'(x_0) \neq 0$, *то* $x = x_0 + \lambda T_{1/3}(-c)$, *где* $\lambda$ *и* $c$ *определены выше.*

### 2.3. Уравнения степени четыре

Уравнение степени четыре можно свести к уравнению третьей степени (которое решается с помощью функции $T_{1/3}$), рассматривая пучок плоских квадрик [4].

Пусть $Q\colon V \to \Bbbk$ квадратичная форма и $\dim_{\mathbf{k}} V = n$. Квадратичную форму на плоскости и на прямой можно разложить на линейные множители (возможно, не над исходным полем $\Bbbk$, а над его квадратичным расширением $K$). Пусть $K$ — расширение поля $\Bbbk$, а $V_K$ и $Q_K$ — пространство и форма, соответствующие $V$ и $Q$ при расширении $\Bbbk \subset K$.

Лемма 10. *Если $Q_K$ раскладывается на множители, то $\dim_{\Bbbk} \ker Q \geqslant \geqslant n-2$. Если это неравенство выполнено, то можно явно найти разложение $Q_K = L_1 L_2$ над квадратичным расширением $K$ поля $\Bbbk$.*

Доказательство. Если $Q_K = L_1 L_2$, то $\ker Q_K \supset \bigcap_{i=1,2}\{L_i = 0\}$ и $\dim_K \ker Q_K \geqslant n-2$. Форма $Q$ определена над $k$, поэтому $\dim_{\Bbbk} \ker Q \geqslant \geqslant n-2$. Если неравенство выполнено, то $V$ представимо в виде $V = \ker Q \oplus W$, где $\dim_{\Bbbk} W \leqslant 2$. Пусть $\pi : V \to W$ — проекция вдоль $\ker Q$ и $\tilde{Q}$ — ограничение формы $Q$ на $W$. На $W$ есть разложение $\tilde{Q} = \tilde{L}_1 \tilde{L}_2$ и, следовательно, $Q = (\pi^* \tilde{L}_1)(\pi^* \tilde{L}_1)$.

Утверждение 10. *Координаты $x$, $y$ точек пересечения двух плоских квадрик $\mathcal{P} = 0$ и $\mathcal{R} = 0$, где $\mathcal{P}$ и $\mathcal{R}$ — полиномы второй степени, можно найти, решая одно кубическое и несколько квадратных и линейных уравнений.*

Доказательство. Все квадрики пучка $0 = \mathcal{Q}_\lambda = \mathcal{P} + \lambda\mathcal{R}$, где $\lambda$ — параметр, проходят через искомые точки. При некоторых $\lambda$ квадрика $\mathcal{Q}_\lambda = 0$ распадается на пару прямых, т. е. $\mathcal{Q}_\lambda = \mathcal{L}_1\mathcal{L}_2$, где $\mathcal{L}_1$, $\mathcal{L}_2$ — полиномы первой степени. Это $\lambda$ удовлетворяет кубическому уравнению $\det(Q_\lambda) = 0$, где $Q_\lambda = P + \lambda Q$ — $(3\times 3)$-матрица квадратичной формы, соответствующей уравнению квадрики в однородных координатах. Действительно, при этом $\lambda$ форма $Q_\lambda$ имеет нетривиальное ядро, поэтому $Q_\lambda = L_1 L_2$, причем $L_1$, $L_2$ можно найти, решая одно квадратное и несколько линейных уравнений. Возвращаясь к координатам $x, y$, из $L_1$, $L_2$ получим нужные полиномы $\mathcal{L}_1$, $\mathcal{L}_2$. Остается решить квадратные уравнения для нахождения точек пересечения квадрики $\mathcal{P} = 0$ и прямых $\mathcal{L}_1 = 0$ и $\mathcal{L}_2 = 0$.

Следствие 11. *Корни полинома $a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4$ можно найти, решая одно кубическое и несколько квадратных и линейных уравнений.*

Доказательство. Корни этого полинома — проекции на ось $x$ точек пересечения квадрик $y = x^2$ и $a_0y^2 + a_1xy + a_2y + a_3x + a_4 = 0$.

Полином $F$ называется *композиционно разложимым*, если он представим в виде $F = P(Q)$, где $P$ и $Q$ — полиномы степени, большей чем один.

Утверждение 12. *Полином $F$ степени четыре композиционно разложим, если и только если выполнено одно их следующих эквивалентных условий*:

1) *для некоторого $x_0$ справедливо тождество $F(x - x_0) \equiv F(x_0 - x)$;*

2) *$F'(x_0) = 0$, где $x_0$ — такая точка, что $F^{(3)}(x_0) = 0$.*

Доказательство. Если тождество справедливо, то $F$ — полином второй степени от $y^2$, где $y = x - x_0$. По формуле Тейлора это свойство эквивалентно равенствам $F'(x_0) = F^{(3)}(x_0) = 0$. Пусть $F = Q(P)$, тогда так как полином $P$ представим в виде $P = a(x - x_0)^2 + b$, имеем $F(x - x_0) \equiv F(x_0 - x)$.

## §3. О трех классических задачах

### 3.1. Обращение отображений в радикалах

Когда полиномиальное отображение $P : \mathbb{C} \to \mathbb{C}$ обратимо в радикалах? Начнем с примеров.

Пример 1. Если $P$ — степенной полином $x^n$, то обратное отображение $x = z^{1/n}$, по определению, представимо в радикалах. Если $n = km$ — составное число, то отображение $x^n$ раскладывается в композицию $x^n = (x^m)^k$. Для простого $n$ полином $x^n$ композиционно неразложим.

Пример 2. Если $P = T_n$ — полином Чебышёва, то обратное отображение $T_{1/n}$ представимо в радикалах. Если $n = km$ — составное число, то отображение $T_n$ раскладывается в композицию $T_n = T_k(T_m)$. Для простого $n$ полином $T_n$ композиционно неразложим.

Пример 3. Если $P$ — полином степени четыре, то обратное отображение представимо в радикалах (так как уравнения четвертой степени решаются в радикалах). Как правило, полиномы степени четыре композиционно неразложимы. Исключения описаны в утверждении 12.

Теорема 13. *Если $P = P_1 \circ \ldots \circ P_k$, где при $1 \leqslant i \leqslant k$ полином $P_i$ либо линеен, либо равен композиционно неразложимому полиному степени четыре, либо равен $x^n$, где $n$ — простое число, либо равен $T_n$, где $n > 2$ — простое число. Тогда отображение $P\colon \mathbb{C} \to \mathbb{C}$ обратимо в радикалах.*

Доказательство. Следует из рассмотренных примеров 1)–3).

Ритт [14] доказал обратную теорему (см. также [3, 5]).

Теорема 14 (J. Ritt). *Если отображение* $P\colon \mathbb{C} \to \mathbb{C}$ *обратимо в радикалах, то полином* $P$ *представим в виде, описанном в теореме* 13.

С теоремой 14 связан следующий интересный вопрос. Насколько единственно представление полинома в виде

$$P = P_1 \circ \ldots \circ P_k, \tag{4}$$

где при $1 \leqslant i \leqslant k$ полиномы $P_i$ композиционно неразложимы? Ритт дал полный ответ на этот вопрос ([15], см. также [17]). Есть ряд соотношений

$$A \circ B = C \circ D, \tag{5}$$

в которых $A$, $B$, $C$, $D$ — полиномы. Например, есть равенство $T_m \circ T_n = T_n \circ T_m$. Есть следующее обобщение равенства $(x^m)^n = (x^n)^m$: для всякого полинома $H$ равенство (5) выполнено для $A(x) = x^n$, $B(x) = x^m H(x^n)$, $C(x) = x^m H^n(x)$, $D(x) = x^n$. Ритт доказал, что по модулю выписанных равенств и композиционных соотношений с линейными функциями представление в виде (4) единственно.

Итак, Ритт полностью описал все полиномы, обратимые в радикалах. Семейства степенных полиномов и полиномов Чебышёва играют центральную роль в этом описании.

Замечание. В статье [7] полностью описаны все полиномы, обратимые в $k$-радикалах, т. е. обратимые при помощи радикалов и решения алгебраических уравнений степени не выше $k$ (где $k$ — любое заданное натуральное число). Это обобщение теоремы Ритта опирается на принадлежащую Мюллеру классификацию полиномов [13], обращение которых имеет примитивную группу монодромии.

Ритту также удалось полностью описать рациональные отображения $R\colon \mathbb{C} \to \mathbb{C}$ простой степени $p$, которые обратимы в радикалах [14]. В его описании фигурируют функции, связанные с делением аргумента эллиптической функции (подобно тому, как полином $T_n$ связан с делением аргумента функции cos). Подробнее о таких отображениях см. [5, 6].

### 3.2. Обратимость отображений конечных полей

Полином $P \in \mathbb{Q}[x]$ можно определить над $\mathbb{Z}_p$, если простое число $p$ не делит знаменатели его коэффициентов. Для каких $P$ отображение $P\colon \mathbb{Z}_p \to \mathbb{Z}_p$ обратимо (т. е. взаимно однозначно) для бесконечного множества простых чисел $p$? Этот вопрос был поставлен Шуром [16], который нашел гипотетический ответ и получил ряд результатов в этом направлении. Фрид доказал гипотезу Шура даже в большей общности [8] — вместо поля $\mathbb{Q}$ он рассматривал его конечное расширение $K$. Здесь мы ограничимся случаем $K = \mathbb{Q}$. Иногда нам понадобятся квадратичные расширения $\mathbb{k}$ полей $\mathbb{Z}_p$, содержащие $p^2$ элементов.

Пример 4. При $p > 2$ четный полином $P \in \mathbb{Z}[x]$ (например, $x^{2n}$ или $T_{2n}$) задает необратимое отображение $P\colon \mathbb{Z}_p \to \mathbb{Z}_p$, так как $P(x) = P(-x)$ и число значений полинома не больше чем $\dfrac{p-1}{2} + 1 < p$.

Пример 5. Для линейного полинома $P(x) = \dfrac{a_1}{b_1}x + \dfrac{a_2}{b_2}$ отображение $P\colon \mathbb{Z}_p \to \mathbb{Z}_p$ определено и обратимо, если $b_1 b_2$ не делится на $p$.

Пример 6. Отображение $P\colon K \to K$ для $P(x) = x^q$, где $q \neq 2$ — простое число и $K$ — конечное поле, обратимо, если $\#K \neq 1 \mod q$. Для $K = \mathbb{Z}_p$ условие $p \neq 1 \mod q$, в частности, выполнено для $p = 2 \mod q$. Для квадратичного расширения $\Bbbk$ поля $\mathbb{Z}_p$ условие $p \neq \pm 1 \mod q$ при $q > 3$, в частности, выполнено, если $p = 2 \mod q$.

Утверждение 15. *Пусть $q > 2$, $p > 2$ — простые числа и $p \neq \pm 1 \mod q$. Тогда отображение $T_q\colon \mathbb{Z}_p \to \mathbb{Z}_p$ обратимо.*

Доказательство. Докажем, что при любом $a \in \mathbb{Z}_p$ уравнение $T_q(x) = a$ имеет решение в $\mathbb{Z}_p$. Пусть $\Bbbk$ — расширение степени два поля $\mathbb{Z}_p$. Уравнение $v^2 - av + 1 = 0$ имеет решения $v_1, v_2 \in \Bbbk$. Так как $p \neq \pm 1 \mod q$, существует единственное решение $u_1 \in \Bbbk$ уравнения $u^q = v_1$, где $v_1$ — любое из решений $v_1, v_2$. Пусть $g$ — нетривиальный элемент группы Галуа поля $\Bbbk$ над $\mathbb{Z}_p$. Обозначим $g(u_1)$ через $u_2$. Так как $g(v_1) = v_2$, то $u_2^q = v_2$. Из равенства $(u_1 u_2)^q = v_1 v_2 = 1$ вытекает, что $u_1 u_2 = 1$. Откуда следует, что элемент $x = (u_1 + u_2)/2$ — решение уравнения $T_q(x) = a$. Так как $g(x) = x$, то $x \in \mathbb{Z}_p$. Мы доказали, что отображение $T_q\colon \mathbb{Z}_p \to \mathbb{Z}_p$ является отображением «на». Поскольку поле $\mathbb{Z}_p$ конечно, это отображение обратимо.

Замечание. Про $T_3$ в утверждении 15 говорится лишь, что отображение $T_3\colon \mathbb{Z}_3 \to \mathbb{Z}_3$ обратимо (что очевидно). Можно проверить, что отображение $T_3\colon \mathbb{Z}_p \to \mathbb{Z}_p$ необратимо при $p > 3$.

Теорема 16. *Пусть $P = P_1 \circ \ldots \circ P_k$, где при $1 \leqslant i \leqslant k$ полином $P_i \in \mathbb{Q}[x]$ либо линеен, либо равен $x^q$, где $q > 2$ — простое число, либо равен $T_q$, где $q > 3$ — простое число. Тогда отображение $P\colon \mathbb{Z}_p \to \mathbb{Z}_p$ обратимо для бесконечного множества простых чисел.*

Доказательство. Обозначим через $E$ конечное множество простых чисел $p$, для которых линейные полиномы, входящие в разложение полинома $P$, не определены над $\mathbb{Z}_p$. Пусть $M = \{q_i\}$ — множество различных степеней полиномов $T_{q_i}$ и $x^{q_i}$, входящих в разложение полинома $P$, и $m = \prod_{q_i \in M} q_i$. Пусть $S$ — множество натуральных чисел, равных двойке по модулю $m$. Если $a \in S$ и $q_i \in M$, то $a \mod q_i = 2$. По теореме Дирихле в арифметической последовательности $S$ есть бесконечно много простых чисел $p > 2$, не принадлежащих конечному множеству $E$. Для каждого из

таких простых чисел $p$ каждое из отображений $P_i\colon \mathbb{Z}_p \to \mathbb{Z}_p$ обратимо (см. примеры 5, 6 и утверждение 15). Теорема доказана.

Теорема 17 (Фрид). *Пусть для $P \in \mathbb{Q}[x]$ отображение $P\colon \mathbb{Z}_p \to \mathbb{Z}_p$ обратимо для бесконечного множества простых чисел $p$, тогда $P$ представим в виде $P = P_1 \circ \ldots \circ P_k$, где при $1 \leqslant i \leqslant k$ полином $P_i$ либо линеен, либо равен $x^q$, либо равен $T_q$.*

Статья Фрида [8] содержит красивые результаты о комплексных полиномах, близкие к теореме 14 Ритта. Она также использует связи между теорией чисел и алгебраической геометрией (в частности, некоторые результаты А. Вейля).

### 3.3. Интегрируемые отображения

Итерации полиномиального отображения $P\colon \mathbb{C} \to \mathbb{C}$ комплексной прямой в себя для полиномов $x^n$ и $T_n$ ведут себя очень необычно. Их динамика напоминает поведение вполне интегрируемых систем в гамильтоновой механике.

Пример 7. Итерации отображения $x \to x^n$ описываются явно: $k$-я итерация — это отображение $x \to x^{n^k}$. Если $k \to \infty$, то $x_0^{n^k} \to 0$ при $|x_0| < 1$ и $x_0^{n^k} \to \infty$ при $|x_0| > 1$. Проекция $x = \exp it$ прямой $\mathbb{R}$ на окружность $|x| = 1$ сопрягает растяжение $t \to nt$ с отображением $x \to x^n$. Отрезок $|t - t_0| \leqslant \varepsilon$ после $k$-й итерации растяжения переходит в отрезок $|t - n^k t_0| \leqslant \varepsilon n^k$. При $k \gg 0$ каждая точка окружности имеет порядка $\dfrac{\varepsilon}{\pi} n^k$ прообразов в этом отрезке. Точки $\exp 2\pi i n^{-k}$ после $k$-й итерации попадут в точку 1 и останутся в этой точке при следующих итерациях. Хотя итерации отображения описаны явно, его динамика хаотична на окружности $|x| = 1$.

Пример 8. Итерации отображения $x \to T_n(x)$ описываются явно: $k$-я итерация — это отображение $x \to T_{n^k}$. Если $k \to \infty$, то $T_{n^k}(x_0) \to \infty$ при $x_0 \notin I$, где $I \subset \mathbb{R}$ отрезок, определенный неравенством $|x| \leqslant 1$. Проекция $x = \dfrac{u + u^{-1}}{2}$ окружности $|u| = 1$ на отрезок $I$ сопрягает отображения $u \to u^n$ с отображением $x \to T_n(x)$. На отрезке $I$ динамика отображения $T_n$ столь же хаотична, как динамика отображения $u^n$ на окружности $|u| = 1$.

Определение. Полиномиальное отображение $P\colon \mathbb{C} \to \mathbb{C}$ *интегрируемо* (см. [1]), если существует полиномиальное отображение $G\colon \mathbb{C} \to \mathbb{C}$, такое, что $P \circ G = G \circ P$, причем: 1) $\deg P > 1$, $\deg G > 1$; 2) $k$-я итерация полинома $P$ не совпадает с $q$-й итерацией полинома $G$ для любых натуральных $k, q$.

Отображение $x \to x^n$ интегрируемо, так как оно коммутирует со всеми степенными отображениями $x \to x^m$. Если $m \neq n^{k/q}$, где $k, q \in \mathbb{Z}$, то

итерации этих отображений различны. Отображение $x \to T_n(x)$ интегрируемо, так как оно коммутирует со всеми отображениями $x \to T_m(x)$. Если $m \neq n^{k/q}$, где $k, q \in \mathbb{Z}$, то итерации этих отображений различны.

Полиномы $P$ и $G$ *эквивалентны*, если существует полином $H(x) = ax + b$, $a \neq 0$, такой, что $P \circ H = H \circ G$. Ритт, Жулиа и Фату описали все интегрируемые полиномиальные отображения с точностью до эквивалентности. Приведем их замечательный результат (см. [9, 10, 15]).

Теорема 17. *Отображение $P\colon \mathbb{C} \to \mathbb{C}$ интегрируемо, если и только если полином $P$ эквивалентен одному из полиномов $x^n$, $T_{2m}$, $T_{2m+1}$, $-T_{2m+1}$.*

Жулиа и Фату доказали эту теорему, используя методы динамики. Доказательство Ритта совершенно другое (ср. п. 3.1).

Ранее Латте привел примеры интегрируемых (в аналогичном смысле) рациональных отображений $\mathbb{C}P^1$ в себя [11, 12]. Ритт доказал, что нет интегрируемых рациональных отображений, кроме отображений Латте. Динамическими методами, восходящими к Жулиа и Фату, доказать эту теорему Ритта никто не мог, пока это не удалось Еременко [2].

Интересно, что все отображения Латте обратимы в радикалах. Ритт описал замечательный класс рациональных отображений, обратимых в радикалах (см. [5, 14]). Этот класс достаточно широк. Например, он содержит все отображения Латте и все обратимые в радикалах отображения простой степени.

Известны многомерные примеры интегрируемых полиномиальных и рациональных отображений (их можно найти в литературе, приведенной в обзоре Милнора [12]).

## Список литературы

[1] Веселов А. П. *Интегрируемые отображения* // УМН. Т. 45. Вып. 5 (281). 1991. С. 3–45.

[2] Еременко А. Э. *О некоторых функциональных уравнениях, связанных с итерацией рациональных функций* // Алгебра и анализ. Т. 1. Вып. 4. 1989. С. 102–116.

[3] Хованский А. Г. *Вариации на тему разрешимости в радикалах* // Тр. МИАН. Т. 259. 2007. С. 86–105.

[4] Berger M. *Geometry*. New York, Berlin, Heidelberg: Springer. 1987.

[5] Burda Y. *Around rational functions invertible in radicals*. arXiv:1005.4101. 2010.

[6] Burda Y., Khovanskii A. *Signature of Branch Coverings*. arXiv:1207.1211. 2012.

[7] Burda Y., Khovanskii A. *Polinomials invertible in k-radicals*. arXiv:1209.5137. 2012.

[8] Fried M. *On conjecture of Schur* // Michigan Math. J. Vol. 17. 1970. P. 41–55.

[9] Fatou P. *Sur lítertation analytique et les substitutions permutables* // J. math. pure appl. V. 23. 1924. P. 1–49.

[10] Julia G. *Memoire sur la permutabilite des fractions rationale* // Ann. sci. Ec. super. Vol. 39. 1922. P. 131–215.

[11] Lattès S. *Sur l'iteration des substitutions rationelles et les fonctions de Poincaré* // C.R. Acad. Sci. Paris. Vol. 166. 1918. P. 26–28.

[12] Milnor J. *On Lattès Maps*. Stony Brook IMS Preprint #2004/01.

[13] Muller P. *Primitive monodromy groups of polynomials* // Recent developments in the inverse Galois problem (Seattle,WA, 1993). Volume 186 of Contemp. Math. Providence, RI: AMS. 1995. P. 385–401.

[14] Ritt J. F. *On algebraic functions which can be expressed in terms of radicals* // Trans. Amer. Math. Soc. Vol. 24. 1922. P. 21–30.

[15] Ritt J. F. *Permutable rational functions* // Trans. Amer. Math. Soc. Vol. 25. No 4. 1923. P. 1–49.

[16] Shur I. *Über den Zusammenhang zwishen einnem Problem der Zahlentheorie polynomials* // Acta Arith. B. 12. 1966/1967. S. 289–299.

[17] Zieve M., Muller P. *On Ritt's polynomial decomposition theorems*. arXiv:0807.3578v1. 2008.

А. Г. Хованский, Univ. of Toronto

# Горизонтально-выпуклые полиамонды и их производящие функции

В. М. Журавлев

## 1. Введение

В литературе по комбинаторной геометрии хорошо известны фигуры, составленные из одинаковых квадратов — полимино, из одинаковых правильных треугольников — полиамонды и из одинаковых правильных шестиугольников — полигексы (см., например, [1]). Фигуры должны быть связными и каждый правильный многоугольник с длиной стороны 1 должен иметь общую сторону с каким-нибудь другим правильным многоугольником с длиной стороны 1. Более того, в упомянутой книге С. Голомб отмечает, что поставленная на любую клетку полимино ладья сможет за конечное число ходов перейти на любую другую клетку того же полимино. Таким образом, для определения полимино обычной геометрической связности составленной фигуры недостаточно. Термин «ход ладьи» не является общепринятым для случая фигур, составленных из правильных треугольников и шестиугольников. Чтобы выйти из этой ситуации, мы можем каждой такой фигуре сопоставить граф. Каждому правильному многоугольнику с длиной стороны 1 будет соответствовать вершина графа. Если два таких многоугольника имеют общую сторону, то соединим соответствующие им вершины ребром. Теперь требование связности составленной фигуры мы можем заменить требованием связности соответствующего этой фигуре графа.

Обычно эти фигуры рассматривают на соответствующих решётках: квадратной, треугольной и шестиугольной. Поэтому такие фигуры часто называют решётчатыми монстрами[1)]. Различные типы решётчатых монстров можно получить, если рассматривать различные группы движений плоскости. *Трансляционный* тип полимино (полиамондов, полигексов) получается, если мы считаем одинаковыми фигуры, совпадающие при параллельном переносе. Например, на рис. 1 показаны все трансляционные 1-амонды и 3-амонды. *Вращательный* (*ротационный*) тип возникает, если считать одинаковыми любые два монстра, эквивалентные относительно

[1)] Англ. “lattice animals”.

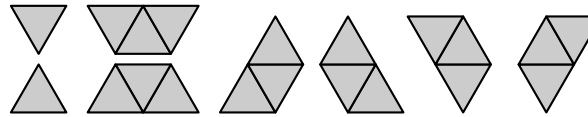


*Рис. 1.*

группы собственных движений плоскости — поворотов и параллельных переносов. *Изометрический* тип возникает, если считать одинаковыми (эквивалентными) любые два конгруэнтных монстра.

Трансляционное полимино называют *выпуклым по рядам* (соответственно *выпуклым по столбцам*), если пересечение любой горизонтальной (вертикальной) прямой с полимино, либо пусто, либо состоит только из одного отрезка. Иногда такие полимино называют горизонтально-выпуклыми (вертикально-выпуклыми) или «штабельными» полимино. *Выпуклыми* называют полимино, которые одновременно являются выпуклыми по рядам и выпуклыми по столбцам. Отметим, что выпуклое полимино не обязано быть выпуклой геометрической фигурой.

Пусть $s_n$ обозначает количество различных выпуклых по рядам $n$-мино (полимино состоящих из $n$ единичных квадратов). Нетрудно вычислить начальные члены этой последовательности $s_1 = 1$, $s_2 = 2$, $s_3 = 6$, $s_4 = 19$, $s_5 = 61$. Известно, что эта последовательность удовлетворяет рекуррентному соотношению третьего порядка

$$s_n = 5s_{n-1} - 7s_{n-2} + 4s_{n-3}, \quad \text{для } n \geqslant 5.$$

Производящей функцией для последовательности будет рациональная функция

$$S(x) = \sum_{n=1}^{\infty} s_n x^n = \frac{x(1-x)^3}{1 - 5x + 7x^2 - 4x^3}\,.$$

Эти результаты были получены Д. Кларнером в [8] с помощью комбинаторной интерпретации интеграла Фредгольма. Там же в [8] ему удалось вычислить производящую функцию для количества выпуклых по рядам полигексов

$$F(x) = \sum_{n=1}^{\infty} = \frac{x(1-x)^3}{1 - 6x + 10x^2 - 7x^3 + x^4}\,.$$

Как следствие, последовательность для количества выпуклых по рядам полигексов будет удовлетворять рекуррентному соотношению четвёртого порядка

$$f_n = 6f_{n-1} - 10f_{n-2} + 7f_{n-3} - f_{n-4}, \quad \text{для } n \geqslant 5.$$

Начальные члены этой последовательности таковы: $f_1 = 1$, $f_2 = 3$, $f_3 = 11$, $f_4 = 42$.

По каким-то причинам в упоминавшейся работе [8] Д. Кларнер не нашёл производящую функцию для полиамондов, хотя и получил оценку снизу для их количества.

Кроме производящих функций, соответствующих площади полимино, также рассматриваются другие виды производящих функций, в частности, производящие функции, соответствующие периметру полимино, или производящие функции нескольких переменных. Так М. Деле [5] нашёл рациональную производящую функцию двух переменных для количества выпуклых по рядам $n$-мино с $m$ столбцами и алгебраическую производящую функцию для количества выпуклых по рядам полимино с периметром $2n + 2$.

Для количества выпуклых $n$-мино получены асимптотические оценки см. [4,9] (начальные члены последовательности для их количества таковы: 1, 2, 6, 19, 59, ...). В то же время М. Деле и Ж. Вьенно [6] нашли точную формулу для количества выпуклых полимино $p_{2n}$ с периметром $2n$:

$$p_{2n+8} = (2n + 11) \cdot 4^n - 4(2n + 1)\binom{2n}{n}.$$

Асимптотические оценки для общего количества полигексов и полиамондов получены в работе [10].

По аналогии мы можем определить горизонтально-выпуклые полиамонды (полигексы). Поскольку для треугольной и шестиугольной решётки у нас имеется три основных направления 0 (горизонтальное), $\pi/3$ и $2\pi/3$, то для определения выпуклых полиамондов (полигексов) мы должны потребовать выполнение выпуклости по всем трём основным направлениям. В то же время, мы можем рассматривать полиамонды (полигексы) выпуклые по двум направлениям.

Публикации, связанные с изучением полиамондов, встречаются гораздо реже, чем публикации связанные с полимино. Мы хотим сократить этот разрыв, и поэтому в этой статье мы будем рассматривать горизонтально-выпуклые полиамонды. Мы не будем затрагивать проблем, связанных с выпуклыми полиамондами, поскольку это тема для отдельного исследования.

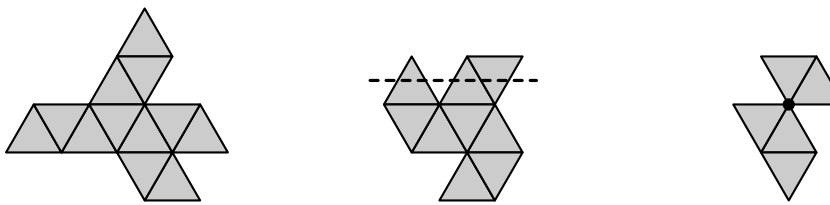
Что касается горизонтально-выпуклых полиамондов, то нам удалось вычислить несколько рациональных производящих функций для их количества, при этом мы использовали комбинаторные методы, описанные в [2]. В результате оказалось, что последовательность для количества горизонтально-выпуклых полиамондов удовлетворяет рекуррентному соотношению седьмого порядка. Кроме того, под впечатлением работы [7], мы нашли другое доказательство рекуррентного соотношения, использующее

элементарные выкладки. Это доказательство доступно для понимания подготовленным учащимся старших классов.

## 2. Определения и обозначения

Во множестве трансляционных полиамондов мы рассмотрим подмножество горизонтально-выпуклых полиамондов. В дальнейшем мы будем опускать упоминание трансляционного типа и говорить просто «полиамонды».

Горизонтально-выпуклыми мы называем такие полиамонды, что любая горизонтальная прямая либо не пересекает полиамонд, либо пересекает его по отрезку или точке. На рис. 2a) приведён пример горизонтально-выпуклого полиамонда, заметим, что этот полиамонд не является выпуклым по направлениям $\pi/3$ и $2\pi/3$. На рис. 2b) приведён пример полиамонда **не** являющегося горизонтально-выпуклым, поскольку существует горизонтальная прямая, которая пересекает его по двум отрезкам. На рис. 2c) приведён пример связной фигуры, которая не является полиамондом.



a) горизонтально-выпуклый полиамонд

b) не горизонтально-выпуклый полиамонд

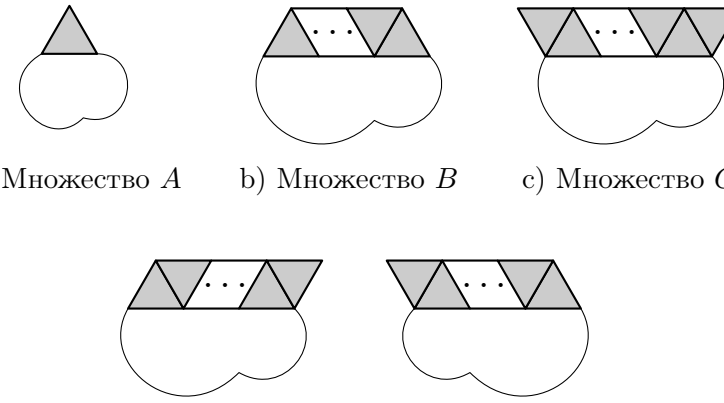
c) связная фигура, не являющаяся полиамондом

*Рис. 2.*

Мы отмечали, что имеется только два 1-амонда (рис. 1), они являются горизонтально-выпуклыми. Единичный правильный треугольник, расположенный остриём вверх, мы будем обозначать через $\Delta$, а единичный правильный треугольник, расположенный остриём вниз, мы будем обозначать через $\nabla$. Понятно, что любой полиамонд состоит из единичных треугольников $\nabla$ и $\Delta$.

Определим множества горизонтально-выпуклых полиамондов $A$, $B$, $C$, $D$ и $D^*$ в зависимости от формы их верхней строки (рис. 3):

a) если верхняя строка полиамонда состоит только из одного треугольника $\Delta$, то отнесём его к множеству $A$ (в частности к $A$ относится треугольник $\Delta$);



a) Множество $A$ b) Множество $B$ c) Множество $C$

d) Множества $D$ и $D^*$

*Рис. 3.*

b) если верхняя строка полиамонда является равнобедренной трапецией, нижнее основание которой длиннее верхнего, то отнесём его к множеству $B$;

c) если верхняя строка является равнобедренной трапецией, нижнее основание которой короче верхнего, то отнесём его к множеству $C$; для удобства дальнейших вычислений треугольник $\nabla$ также отнесём к множеству $C$;

d) если верхняя строка полиамонда является параллелограммом с крайним правым треугольником $\nabla$ либо треугольником $\Delta$, то отнесём его к множеству $D$ либо к множеству $D^*$ соответственно.

К множеству $H$ отнесём полиамонды, у которых в верхней строке находится не менее двух треугольников, т. е. множество $H$ является объединением множеств $B$, $C$, $D$ и $D^*$.

Множество $G$ будет состоять из всех горизонтально-выпуклых полиамондов, т. е. множество $G$ является объединением множеств $A$ и $H$.

Через $A_n$, $B_n$, $C_n$, $D_n$, $D_n^*$, $H_n$, $G_n$ обозначим подмножества множеств $A$, $B$, $C$, $D$, $D^*$, $H$, $G$, в которых полиамонды состоят из $n$ правильных треугольников с длиной стороны 1 ($n$-амондов). Количество элементов в соответствующих подмножествах обозначим через $a_n$, $b_n$, $c_n$, $d_n$, $d_n^*$, $h_n$, $g_n$.

Исходя из соображений симметрии, мы можем заключить, что полиамондов в множестве $D_n$ ровно столько же, сколько полиамондов в множестве $D_n^*$. Следовательно, для любого $n$ выполнено $d_n = d_n^*$.

Тогда из наших определений следует что

$$h_n = b_n + c_n + d_n + d_n^* = b_n + c_n + 2d_n, \tag{2.1}$$

$$g_n = a_n + b_n + c_n + d_n + d_n^* = a_n + h_n. \tag{2.2}$$

В дальнейшем мы хотим использовать и изучить детализацию полиамондов по количеству составляющих их треугольников $\Delta$ и $\nabla$. Более того, для нас будет важно, сколько треугольников вида $\nabla$ расположено в самой верхней строке полиамонда. Исходя из этих соображений, обозначим через $a(p,q,m)$, $b(p,q,m)$, $c(p,q,m)$, $d(p,q,m)$, $d^*(p,q,m)$, $h(p,q,m)$ и $g(p,q,m)$ количество полиамондов принадлежащих подмножествам $A$, $B$, $C$, $D$, $D^*$, $H$, $G$ соответственно, которые состоят из $p$ треугольников $\nabla$, $q$ треугольников $\Delta$ и верхняя строка которых содержит $m$ треугольников $\nabla$.

Поскольку у полиамондов из множества $A$ верхняя строка состоит только из одного треугольника $\Delta$, то $a(p,q,m) = 0$ при $m \geqslant 1$.

Применяя соображения симметрии, получаем $d(p,q,m) = d^*(p,q,m)$.

Мы можем записать равенства аналогичные полученным ранее

$$h(p,q,m) = b(p,q,m) + c(p,q,m) + 2d(p,q,m), \tag{2.3}$$

$$g(p,q,m) = a(p,q,m) + h(p,q,m). \tag{2.4}$$

Мы сделали хорошую подготовительную работу. И все же, прежде чем переходить к производящим функциям, мы хотим предложить вывод рекуррентного соотношения, использующий элементарные выкладки.

## 3. Рекуррентное соотношение

Структура предлагаемого доказательства представляет собой ряд лемм. Каждая лемма связывает рекуррентным соотношением последовательности количества полиамондов в определённых подмножествах. Фактически каждая лемма устанавливает взаимно однозначное соответствие между несколькими подмножествами полиамондов. В итоге мы получаем систему рекуррентных соотношений между последовательностями. Вывод итогового соотношения представляет собой техническое упражнение над полученными промежуточными соотношениями.

Теорема 1. *Для $n \geqslant 8$ выполняется рекуррентная формула*:

$$g_n = 3g_{n-1} - 4g_{n-2} + g_{n-4} + g_{n-5} + 3g_{n-6} - g_{n-7}\,. \tag{3.1}$$

Прежде чем приступать к доказательству теоремы 1 сформулируем и докажем небольшие леммы.

Лемма 1. *Для $n \geqslant 2$ выполняется соотношение*

$$c_n = g_{n-1}\,.$$

Доказательство. Если мы добавим треугольник $\nabla$ слева в верхнюю строку горизонтально-выпуклого $(n-1)$-амонда из множества $D_{n-1}$, то

получим горизонтально-выпуклый $n$-амонд из $C_n$. Эта процедура обратима, и из каждого горизонтально-выпуклого $n$-амонда из $C_n$, удаляя левый треугольник из верхней строки, мы получим горизонтально-выпуклый $(n-1)$-амонд из множества $D_{n-1}$ (рис. 4).

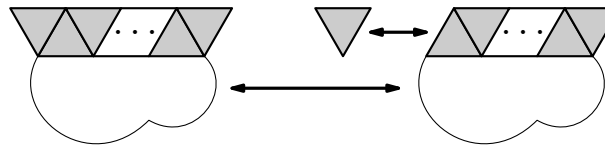


*Рис. 4.*

Мы получаем взаимно однозначное соответствие между такими множествами. Следовательно, для любого $n \geqslant 2$ выполнено $c_n = g_{n-1}$. □

Лемма 2. *Для $n \geqslant 2$ выполняется соотношение*

$$d_n = a_{n-1} + b_{n-1}\,.$$

Доказательство. Рассмотрим полиамонд из множества $D_n$. Если этот полиамонд имеет не менее четырёх треугольников в верхней строке, то мы можем удалить самый правый треугольник $\nabla$ из верхней строки и получить полиамонд из множества $B_{n-1}$. Если этот полиамонд имеет в точности два треугольника в верхней строке, то, так же удаляя самый правый треугольник $\nabla$ из верхней строки, мы получим полиамонд из множества $A_{n-1}$. Очевидно, что этот процесс обратим (рис. 5).

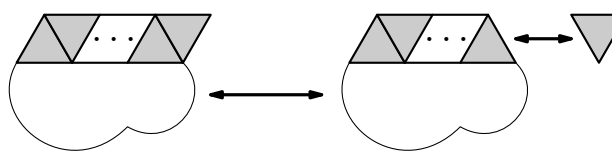
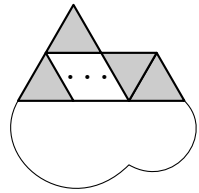


*Рис. 5.*

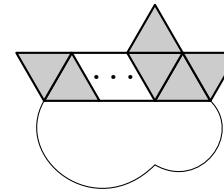
Значит для любого $n \geqslant 2$ мы получаем взаимно однозначное соответствие между множеством $D_n$ и объединением непересекающихся множеств $A_{n-1}$ и $B_{n-1}$. Следовательно, $d_n = a_{n-1} + b_{n-1}$. □

Для получения дополнительных соотношений введём ещё несколько подмножеств горизонтально-выпуклых полиамондов.

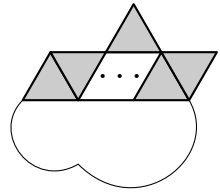
В множестве $A$ определим подмножества горизонтально-выпуклых полиамондов $I$, $J$, $E$ и $E^*$ в зависимости от формы их второй сверху строки (рис. 6):

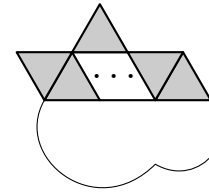


a) Множество $I$



b) Множество $J$





c) Множество $E$ и множество $E'$

*Рис. 6.*

a) если вторая сверху строка полиамонда является равнобедренной трапецией, нижнее основание которой длиннее верхнего, то отнесём его к множеству $I$;

b) если вторая сверху строка полиамонда является равнобедренной трапецией, нижнее основание которой короче верхнего, то отнесём его к множеству $J$;

c) если вторая сверху строка полиамонда является параллелограммом, крайний правый треугольник которого направлен остриём вниз или направлен остриём вверх, то отнесём его к множеству $E$ или множеству $E^*$ соответственно.

Через $I_n$, $J_n$, $E_n$ и $E_n^*$ обозначим множество горизонтально-выпуклых полиамондов из $I$, $J$, $E$ и $E^*$ соответственно, которые состоят из $n$ треугольников, а количество элементов в этих множествах $i_n$, $j_n$, $e_n$, $e_n^*$ соответственно.

Соображения симметрии дают равенство $e_n = e_n^*$.

Понятно, что множество $A$ является объединением непересекающихся подмножеств $I$, $J$, $E$ и $E^*$. Следовательно, для $n \geqslant 2$

$$a_n = i_n + j_n + e_n + e^* n = i_n + j_n + 2e_n \,.$$

Лемма 3. *Для $n \geqslant 2$ выполняются соотношения*

$$1) \qquad j_n = c_{n-1} + e_{n-1} \,,$$
$$2) \qquad e_n = d_{n-1} + i_{n-1} \,.$$

Доказательство. 1) Рассмотрим $n$-амонд из множества $J$. Если в таком $n$-амонде треугольник из верхней строки расположен в точности над самым правым треугольником второй строки, то, удалив треугольник из верхней строки, мы получим $(n-1)$-амонд из $C$ (рис. 7). Количество таких полиамондов будет $c_{n-1}$. Если же в таком $n$-амонде треугольник из верхней строки расположен не над самым правым треугольником второй строки, то удалим самый правый треугольник из второй строки и получим $(n-1)$-амонд из $E^*$. Количество таких полиамондов будет $e_{n-1}$.

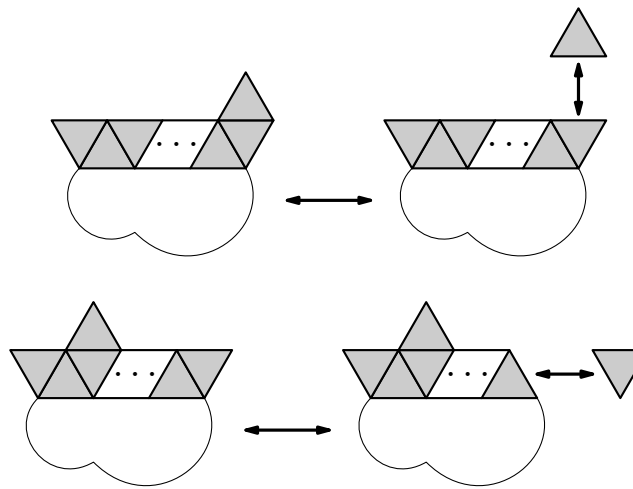


*Рис. 7.*

Эти операции обратимы, значит для любого $n \geqslant 2$ мы получаем взаимно однозначное соответствие между множеством $J_n$ и объединением непересекающихся множеств $C_{n-1}$ и $E^*_{n-1}$. Следовательно, $j_n = c_{n-1} + e_{n-1}$.

2) Будем рассуждать аналогично. Рассмотрим $n$-амонд из множества $E$. Если в таком $n$-амонде треугольник из верхней строки расположен в точности над самым правым треугольником второй строки, то, удалив этот треугольник из верхней строки, мы получим $(n-1)$-амонд из $D_{n-1}$. Количество таких полиамондов будет $d_{n-1}$. Если же в таком $n$-амонде треугольник из верхней строки расположен не над самым правым треугольником второй строки, то удалим самый правый треугольник из второй строки и получим $(n-1)$-амонд из $I$. Количество таких полиамондов будет $i_{n-1}$. Эти операции обратимы, значит для любого $n \geqslant 2$ мы получаем взаимно однозначное соответствие между множеством $E_n$ и объединением непересекающихся множеств $D_{n-1}$ и $I_{n-1}$. Следовательно, $e_n = d_{n-1} + i_{n-1}$. $\square$

Наконец нам понадобится ещё три подмножества горизонтально-выпуклых полиамондов.

К множеству $P$, отнесём полиамонды из множества $B$, у которых крайний правый треугольник $\Delta$ верхней строки расположен над самым левым треугольником $\nabla$ второй строки (самый левый треугольник $\nabla$ может быть вторым слева треугольником), см. рис. 8.

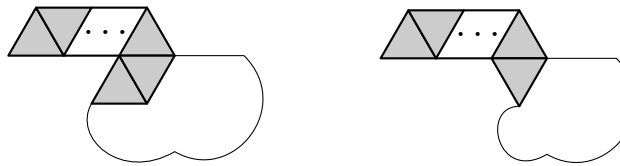


*Рис. 8. Полиамонды из множества $P$*

К множеству $Q$, отнесём те полиамонды из $I$, у которых треугольник из верхней строки не расположен над вторым справа треугольником второй строки (рис. 9).

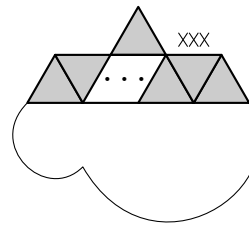



*Рис. 9. Множество $Q$*

К множеству $R$, отнесём те полиамонды из $Q$, у которых крайний правый треугольник $\Delta$ второй сверху строки расположен над самым левым треугольником $\nabla$ третьей сверху строки (самый левый треугольник $\nabla$ может быть вторым слева треугольником) (рис. 10). Таким образом, полиамонды из множества $R$ состоят из не менее чем трёх строк.

На рисунках 9 и 10 знаком «XXX» обозначена позиция, запрещённая для расположения треугольника $\Delta$.

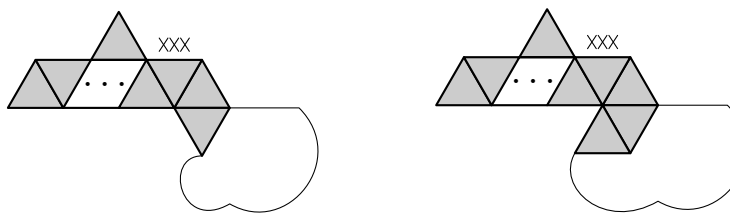



*Рис. 10. Множество $R$*

Через $P_n$, $Q_n$, $R_n$ обозначим множества горизонтально-выпуклых полиамондов типов $P$, $Q$, $R$ соответственно, которые состоят из $n$ треугольников, а через $p_n$, $q_n$, $r_n$ количество элементов в этих множествах.

Лемма 4. *Для $n \geqslant 2$ выполняется соотношение*

$$b_n = d_{n-1} + p_n\,.$$

Доказательство. Множество $B_n \setminus P_n$ состоит из горизонтально-выпуклых $n$-амондов, у которых самый правый треугольник $\Delta$ из верхней строки не расположен над самым левым треугольником $\nabla$ второй строки (самый левый треугольник $\nabla$ может оказаться вторым слева после $\Delta$). Это означает, что у верхней и второй сверху строки общими являются не менее двух сторон треугольников. Если мы удалим самый правый треугольник из верхней строки, то мы получим $(n-1)$-амонд из $D$ (рис. 11).

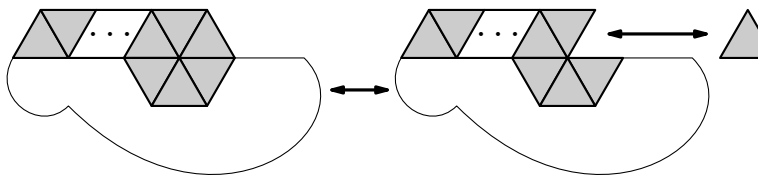


*Рис. 11.*

Понятно, что эту операцию можно обратить, поскольку к каждому $(n-1)$-амонду из $D$ мы можем добавить справа треугольник в верхнюю строку. Значит, количество горизонтально-выпуклых $n$-амондов в множестве $B_n \setminus P_n$ такое же, как и в множестве $D_{n-1}$. Следовательно, $b_n - p_n = d_{n-1}$. Осталось перенести слагаемые в нужные части равенства. □

Лемма 5. *Для $n \geqslant 2$ выполняется соотношение*

$$i_n = b_{n-1} + q_n\,.$$

Доказательство. Множество $I_n \setminus Q_n$ состоит из горизонтально-выпуклых $n$-амондов, у которых треугольник $\Delta$ из верхней строки расположен в точности над вторым справа треугольником второй строки. Если мы удалим треугольник из верхней строки, то получим $(n-1)$-амонд из $B$. Понятно, что эту операцию можно обратить, поскольку к каждому $(n-1)$-амонду из $B$ мы можем добавить треугольник в точности над вторым справа треугольником верхней строки (рис. 12).

Значит, количество горизонтально-выпуклых $n$-амондов в множестве $I_n \setminus Q_n$ такое же, как и в множестве $B_{n-1}$. Следовательно, $i_n - q_n = b_{n-1}$. Осталось перенести слагаемые в нужные части равенства. □

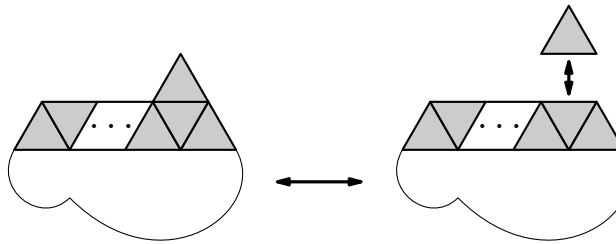


*Рис. 12.*

Лемма 6. *Для $n \geqslant 3$ выполняется соотношение*

$$q_n = r_n + i_{n-2}\,.$$

Доказательство. Множество $Q_n \setminus R_n$ состоит из горизонтально-выпуклых $n$-амондов, у которых треугольник из верхней строки не расположен над вторым справа треугольником второй строки, а самый правый треугольник второй строки не расположен над самым левым треугольником $\nabla$ третьей строки (самый левый треугольник $\nabla$ может оказаться вторым слева после $\Delta$). Это означает, что у второй и третьей сверху строки общими являются не менее двух сторон треугольников. Если мы удалим два крайних справа треугольника из второй сверху строки, то мы получим $(n-2)$-амонд из $I$ (рис. 13).

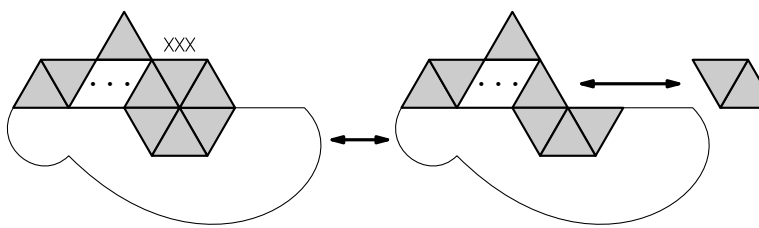



*Рис. 13.*

Понятно, что эту операцию можно обратить, поскольку к каждому $(n-2)$-амонду из $I$ мы можем добавить справа два треугольника во вторую сверху строчку. Значит, количество горизонтально-выпуклых $n$-амондов в множестве $Q_n \setminus R_n$ такое же, как и в множестве $I_{n-2}$. Следовательно, $q_n - r_n = i_{n-2}$. Осталось перенести слагаемые в нужные части равенства. □

Лемма 7. *Для $n \geqslant 4$ выполняется соотношение*

$$r_n = r_{n-2} + p_{n-3}\,.$$

Доказательство. Рассмотрим $n$-амонд из $R$. Если в таком $n$-амонде треугольник из верхней строки расположен в точности над вторым слева треугольником второй строки, то, удалив этот треугольник и два самых левых треугольника из второй сверху строки, мы получим $(n-3)$-амонд из $P$. Следовательно, число таких $n$-амондов будет $p_{n-3}$. Если же в таком $n$-амонде треугольник из верхней строки расположен не над вторым слева треугольником второй строки, то, удалив два самых левых треугольника из второй сверху строки, мы получим $(n-2)$-амонд из $R$ (рис. 14). Число таких $n$-амондов будет $r_{n-2}$.

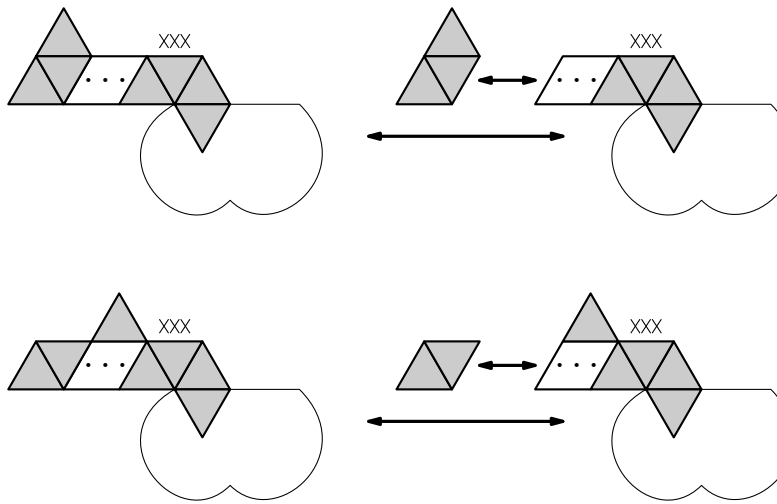



*Рис. 14.*

Эти операции обратимы. Следовательно, $r_n = r_{n-2} + p_{n-3}$. □

Лемма 8. *Для $n \geqslant 4$ выполняется соотношение*

$$p_n = p_{n-2} + h_{n-3}\,.$$

Доказательство. Рассмотрим $n$-амонд из $P$. Если в таком $n$-амонде верхняя строка состоит из 5 и более треугольников, то, удалив слева два треугольника из верхней строки, мы получим $(n-2)$-амонд из $P$ (рис. 15). Если в таком $n$-амонде верхняя строка состоит ровно из 3 треугольников, то, удалив всю эту строку из трёх треугольников, мы получим $(n-3)$-амонд из $H$.

Заметим, что все $n$-амонды из $P$ можно получить такими обратными операциями. Следовательно, $p_n = p_{n-2} + h_{n-3}$. □

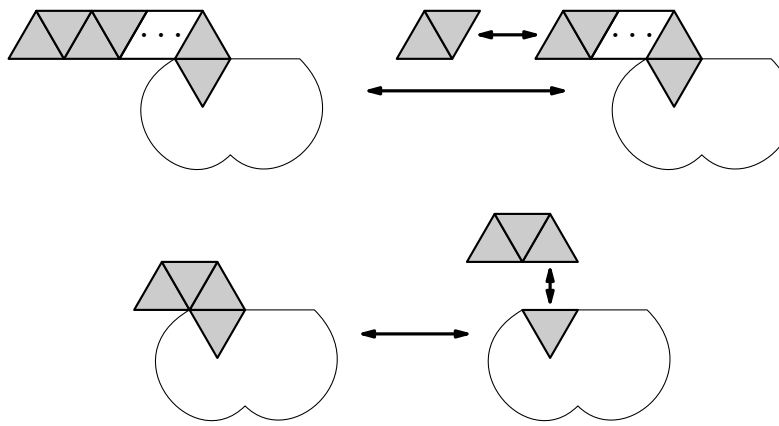


*Рис. 15.*

Сведём полученные рекуррентные соотношения вместе:

$$\begin{cases} g_n = a_n + h_n, \\ h_n = b_n + c_n + 2d_n, \\ c_n = d_{n-1}, \\ d_n = a_{n-1} + b_{n-1}, \\ a_n = i_n + j_n + 2e_n, \\ j_n = c_{n-1} + e_{n-1}, \\ e_n = d_{n-1} + i_{n-1}, \\ b_n = d_{n-1} + p_n, \\ i_n = b_{n-1} + q_n, \\ q_n = r_n + i_{n-2}, \\ r_n = r_{n-2} + p_{n-3}, \\ p_n = p_{n-2} + h_{n-3}. \end{cases}$$

Этих соотношений нам хватит, чтобы доказать теорему 1.

Дальнейшие выкладки будут иметь чисто технический характер с использованием отдельных рекуррентных соотношений системы для получения других рекуррентных соотношений. Чтобы не утомлять читателя, мы просто предложим ему в качестве упражнения промежуточные результаты.

Упражнение 1. Докажите соотношения:

a) $g_{n-1} = d_n + 2d_{n-1} + d_{n-2}$, для $n \geqslant 3$;

b) $g_n - g_{n-1} = 2d_n + d_{n-1} - d_{n-2} - 2d_{n-3} + b_n + b_{n-2} + r_n + r_{n-1}$, для $n \geqslant 4$;

c) $g_n - g_{n-1} + 2g_{n-2} = 2d_n + 3d_{n-1} + 3d_{n-2} + b_n + b_{n-2} + r_n + r_{n-1}$, для $n \geqslant 4$;

d) $g_n - 2g_{n-1} + 3g_{n-2} - 2g_{n-3} = 2d_n + 2d_{n-1} - d_{n-2} - 2d_{n-3} - d_{n-4} + p_n - p_{n-1} + p_{n-2}$, для $n \geqslant 5$;

e) $g_n - 2g_{n-1} + 3g_{n-2} - g_{n-3} = 2d_n + 2d_{n-1} + p_n - p_{n-1} + p_{n-2}$, для $n \geqslant 5$;

f) $p_n - p_{n-2} - p_{n-3} = 2(d_{n-3} + d_{n-4})$, для $n \geqslant 5$;

g) $g_n - 2g_{n-1} + 2g_{n-2} - g_{n-4} - 2g_{n-5} + g_{n-6} = 2d_n + 2d_{n-1} - 2d_{n-2} - 2d_{n-3} - 2d_{n-4} + 2d_{n-6}$, для $n \geqslant 7$;

h) $g_n - 2g_{n-1} + 2g_{n-2} - g_{n-4} - 4g_{n-5} + g_{n-6} = 2(d_n + d_{n-1}) - 2(d_{n-2} + d_{n-3}) - 4(d_{n-4} + d_{n-5})$, для $n \geqslant 7$.

Доказательство теоремы 1. Продемонстрируем технические выкладки для получения итогового соотношения.

Используя результат упражнения 1h) для $n \geqslant 8$ имеем

$$g_{n-1} - 2g_{n-2} + 2g_{n-3} - g_{n-5} - 4g_{n-6} + g_{n-7} = \\ = 2(d_{n-1} + d_{n-2}) - 2(d_{n-3} + d_{n-4}) - 4(d_{n-5} + d_{n-6}).$$

Тогда

$$\begin{gathered} g_n - 2g_{n-1} + 2g_{n-2} - g_{n-4} - 4g_{n-5} + g_{n-6} + \\ + (g_{n-1} - 2g_{n-2} + 2g_{n-3} - g_{n-5} - 4g_{n-6} + g_{n-7}) = \\ = 2(d_n + d_{n-1}) - 2(d_{n-2} + d_{n-3}) - 4(d_{n-4} + d_{n-5}) + \\ + 2(d_{n-1} + d_{n-2}) - 2(d_{n-3} + d_{n-4}) - 4(d_{n-5} + d_{n-6}) = \\ = 2(d_n + 2d_{n-1} + d_{n-2}) - 2(d_{n-2} + 2d_{n-3} + d_{n-4}) - 4(d_{n-4} + 2d_{n-5} + d_{n-6}) = \\ = 2g_{n-1} - 2g_{n-3} - 4g_{n-5}\,. \end{gathered}$$

Следовательно,

$$\begin{gathered} g_n - 2g_{n-1} + 2g_{n-2} - g_{n-4} - 4g_{n-5} + g_{n-6} + \\ + g_{n-1} - 2g_{n-2} + 2g_{n-3} - g_{n-5} - 4g_{n-6} + g_{n-7} = \\ = 2g_{n-1} - 2g_{n-3} - 4g_{n-5}\,. \end{gathered}$$

Перенесём все слагаемые кроме первого в правую часть равенства и приведём подобные члены. В итоге получаем требуемое соотношение

$$g_n = 3g_{n-1} - 4g_{n-2} + g_{n-4} + g_{n-5} + 3g_{n-6} - g_{n-7}\,.$$

Теорема 1 доказана. □

Приведём таблицу начальных значений последовательностей.

| $n$ | $a$ | $b$ | $c$ | $d$ | $i$ | $j$ | $e$ | $p$ | $q$ | $r$ | $h$ | $g$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| 2 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 3 |
| 3 | 2 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 4 | 6 |
| 4 | 5 | 2 | 1 | 3 | 1 | 2 | 1 | 1 | 0 | 0 | 9 | 14 |
| 5 | 12 | 5 | 3 | 7 | 2 | 2 | 4 | 2 | 0 | 0 | 22 | 34 |
| 6 | 31 | 12 | 7 | 17 | 6 | 7 | 9 | 5 | 1 | 0 | 53 | 84 |
| 7 | 77 | 28 | 17 | 43 | 15 | 16 | 23 | 11 | 3 | 1 | 131 | 208 |
| 8 | 192 | 70 | 43 | 105 | 36 | 40 | 58 | 27 | 8 | 2 | 323 | 515 |
| 9 | 474 | 169 | 105 | 262 | 91 | 101 | 141 | 64 | 21 | 6 | 798 | 1272 |

*Табл. 1.*

## 4. Производящие функции

Все ряды в этой статье мы рассматриваем как формальные степенные ряды, не вдаваясь в вопросы сходимости. Операции сложения, умножения, дифференцирования рядов мы рассматриваем как соответствующие операции над формальными степенными рядами.

Одной из наших целей будет нахождение производящей функции для горизонтально-выпуклых полиамондов

$$G(x) = \sum_{n=1}^{\infty} g_n x^n.$$

Попутно мы найдём производящую функцию двух переменных

$$\widetilde{G}(u,v) = \sum_{p,q} g_{p,q} u^p v^q,$$

где $g_{p,q}$ обозначает количество горизонтально-выпуклых полиамондов, состоящих из $p$ треугольников вида $\nabla$ и $q$ треугольников вида $\Delta$.

Заметим, что $G(x) = \widetilde{G}(x,x)$.

Для нахождения указанных функций рассмотрим производящие функции трёх переменных

$$G(u,v,y) = \sum_{p,q,m} g(p,q,m) u^p v^q y^m, \quad A(u,v,y) = \sum_{p,q,m} a(p,q,m) u^p v^q y^m,$$

$$B(u,v,y) = \sum_{p,q,m} b(p,q,m) u^p v^q y^m, \quad C(u,v,y) = \sum_{p,q,m} c(p,q,m) u^p v^q y^m,$$

$$D(u,v,y) = \sum_{p,q,m} d(p,q,m) u^p v^q y^m, \quad H(u,v,y) = \sum_{p,q,m} h(p,q,m) u^p v^q y^m.$$

Из наших построений видно, что $\widetilde{G}(u,v) = G(u,v,1)$, поэтому для наших целей не обязательно находить функцию $G(u,v,y)$, а достаточно найти $G(u,v,1)$.

Из (2.3) и (2.4) следует

$$\begin{aligned} H(u,v,y) &= B(u,v,y) + C(u,v,y) + 2D(u,v,y), \\ G(u,v,y) &= A(u,v,y) + H(u,v,y). \end{aligned} \tag{4.1}$$

Мы отмечали что $a(p,q,m) = 0$, если $m = 1$. Поэтому функция $A(u,v,y)$ фактически является функцией двух переменных и не зависит от $y$, в частности $A(u,v,y) = A(u,v,1)$.

Пусть $\chi(u,v) = \dfrac{\partial}{\partial y} H(u,v,y)\Big|_{y=1}$ обозначает функцию двух переменных, полученную при подстановке $y = 1$ в продифференцированный по переменной $y$ формальный ряд $H(u,v,y)$.

Проведём небольшое техническое вычисление, которое мы впоследствии используем.

При почленном дифференцировании формального ряда получим для всех целых $k \geqslant 0$

$$\frac{\partial}{\partial y}\big(y^k H(u,v,y)\big)\Big|_{y=1} = \frac{\partial}{\partial y}\left(\sum_{p,q,m} h(p,q,m) u^p v^q y^{m+k}\right)\Bigg|_{y=1} =$$

$$= \sum_{p,q,m} (m+k) h(p,q,m) u^p v^q = \sum_{p,q}\left(\sum_m (m+k) h(p,q,m)\right) u^p v^q.$$

С другой стороны, используя свойство дифференцирования произведения функций, получим

$$\frac{\partial}{\partial y}\big(y^k H(u,v,y)\big)\Big|_{y=1} = kH(u,v,1) + \chi(u,v).$$

Следовательно, для всех целых $k \geqslant 0$

$$kH(u,v,1) + \chi(u,v) = \sum_{p,q}\left(\sum_m (m+k) h(p,q,m)\right) u^p v^q. \tag{4.2}$$

В частности, при $k = 0$

$$\chi(u,v) = \sum_{p,q}\left(\sum_m m h(p,q,m)\right) u^p v^q. \tag{4.3}$$

Давайте получим рекуррентные соотношения на коэффициенты $a(p,q,m)$, $b(p,q,m)$, $c(p,q,m)$, $d(p,q,m)$ и затем составим соотношения для производящих функций.

Посмотрим, как получаются полиамонды из $A$. Возьмём треугольник $\Delta$ и начнём прикладывать его сверху к некоторому полиамонду (назовём его исходным) так, чтобы получился новый полиамонд. Мы не сможем проделать такую операцию, если исходный полиамонд был из $A$. Если исходный полиамонд из $H$ и в его верхней строке содержится $m$ треугольников $\nabla$, то такую операцию можно проделать $m$ способами. Заметим, что таким образом мы можем получить все полиамонды подмножества $A$ кроме собственно треугольника $\Delta$. Следовательно, для всех $p, q$ выполняются соотношения

$$a(p, q+1, 0) = \sum_m mh(p, q, m).$$

Для треугольника $\Delta$ мы имеем $a(0, 1, 0) = 1$.

Из этих соотношений следует соотношение для производящих функций

$$A(u, v, y) = v + v \cdot \frac{\partial}{\partial y} H(u, v, y)\Big|_{y=1} = v(1 + \chi(u, v)). \tag{4.4}$$

Действительно, вспоминая (4.3), получаем

$$v(1 + \chi(u, v)) = v + \sum_{p,q} \left( \sum_m mh(p, q, m) \right) u^p v^{q+1} =$$
$$= v + \sum_{p,q} a(p, q+1, 0) u^p v^{q+1} = A(u, v, y).$$

Посмотрим, как получаются полиамонды из множества $B$. Возьмём полиамонд из $B$, состоящий только из одной строки и составленный из $k$ треугольников $\nabla$, и $k+1$ треугольников $\Delta$, и начнём прикладывать его сверху к некоторому (исходному) полиамонду так, чтобы получился новый полиамонд. Мы не сможем проделать такую операцию, если исходным полиамондом был полиамонд из $A$. Если исходный полиамонд был из $H$ и в его верхнем слое содержится $m$ треугольников $\nabla$, то такую операцию можно проделать $m+k$ способами. Заметим, что таким образом мы можем получить все полиамонды из множества $B$ кроме тех, которые состоят из одной строки. Следовательно, для всех $p, q$ выполняются соотношения

$$b(p+k, q+k+1, k) = \sum_m (m+k)h(p, q, m).$$

Для полиамондов из множества $B$, состоящих только из одной строки, получаем

$$b(k, k+1, k) = 1.$$

Используя (4.2), выведем соотношение для производящих функций

$$B(u,v,y)=\sum_{p,q,m} b(p,q,m)u^p v^1 y^m =$$

$$=\sum_k b(k,k+1,k)u^k v^{k+1}y^k+\sum_{p,q,k} b(p+k,q+k+1,k)u^{p+k}v^{q+k+1}y^k =$$

$$=\sum_{k\geqslant 1} u^k v^{k+1}y^k+\sum_{k\geqslant 1} u^k v^{k+1}y^k\left[\sum_{p,q}\Big(\sum_m (m+k)h(p,q,m)\Big)u^p v^q\right]=$$

$$=\sum_{k\geqslant 1} u^k v^{k+1}y^k+\sum_{k\geqslant 1} u^k v^{k+1}y^k(kH(u,v,1)+\chi(u,v)).$$

Итак

$$B(u,v,y)=\sum_{k\geqslant 1} u^k v^{k+1}y^k+\sum_{k\geqslant 1} u^k v^{k+1}y^k(kH(u,v,1)+\chi(u,v)). \tag{4.5}$$

Проделав аналогичные рассуждения для полиамондов из множеств $C$ и $D$, мы получим ещё два соотношения для производящих функций

$$C(u,v,y)=uy+\sum_{k\geqslant 2} u^k v^{k-1}y^k+\sum_{k\geqslant 2} u^k v^{k-1}y^k((k-2)H(u,v,1)+\chi(u,v)), \tag{4.6}$$

$$D(u,v,y)=\sum_{k\geqslant 1} u^k v^k y^k+\sum_{k\geqslant 1} u^k v^k y^k((k-1)H(u,v,1)+\chi(u,v)). \tag{4.7}$$

Складываем равенства (4.5), (4.6) с удвоенным равенством (4.7) и, учитывая (4.1), получим уравнение на функцию $H(u,v,y)$

$$\begin{aligned} H(u,v,y)=&\left(\sum_{k\geqslant 1} u^k v^{k+1}y^k+uy+\sum_{k\geqslant 2} u^k v^{k-1}y^k+2\sum_{k\geqslant 1} u^k v^k y^k\right)+\\ &+\sum_{k\geqslant 1} u^k v^{k+1}y^k(kH(u,v,1)+\chi(u,v))+\\ &+\sum_{k\geqslant 2} u^k v^{k-1}y^k((k-2)H(u,v,1)+\chi(u,v))+\\ &+2\sum_{k\geqslant 1} u^k v^k y^k((k-1)H(u,v,1)+\chi(u,v)). \end{aligned} \tag{4.8}$$

Чтобы не утруждать читателя подробным выписыванием дальнейших громоздких преобразований, мы просто сообщим, какие шаги необходимо проделать, чтобы получить явный вид функции.

Часть громоздких выражений, поддаётся упрощениям.

Упражнение 2. Используя формулу суммы бесконечной геометрической прогрессии, докажите, что

$$\sum_{k\geqslant 1} u^k v^{k+1} y^k + uy + \sum_{k\geqslant 2} u^k v^{k-1} y^k + 2\sum_{k\geqslant 1} u^k v^k y^k = \\ = \frac{uv^2 y}{1-uvy} + \frac{u^2 vy2}{1-uvy} + 2\frac{uvy}{1-uvy} = \frac{uvy(v+2+uy)}{1-uvy}\,.$$

Упражнение 3. Используя интегрирование и дифференцирование формальных рядов, докажите, что

$$\sum_{k\geqslant 1} k u^k v^{k+1} y^k + uy + \sum_{k\geqslant 2} (k-2) u^k v^{k-1} y^k + 2\sum_{k\geqslant 1} (k-1) u^k v^k y^k = \\ = \frac{uv^2 y(1+uy)^2}{(1-uvy)^2}\,.$$

В итоге мы упростим (4.8) до следующего вида

$$H(u,v,y) = uy + \left(\frac{uvy(v+2+uy)}{1-uvy}\right)(1+\chi(u,v)) + H(u,v,1)\frac{uv^2 y(1+uy)^2}{(1-uvy)^2}\,. \tag{4.9}$$

Из уравнения (4.9) легко получить систему уравнений относительно $H(u,v,1)$ и $\chi(u,v)$. Первое уравнение системы мы получим, если в (4.9) сделаем подстановку $y=1$. Второе уравнение системы получается, если (4.9) продифференцировать по $y$ и затем сделать подстановку $y=1$. Имеем

$$\begin{cases} H(u,v,1) = u + \left(\dfrac{uv(v+2+u)}{1-uv}\right)(1+\chi(u,v)) + H(u,v,1)\dfrac{uv^2(1+u)^2}{(1-uv)^2}\,, \\ \chi(u,v) = u + (1+\chi(u,v))\dfrac{\partial}{\partial y}\left(\dfrac{uv(v+2+u)}{1-uv}\right)\Big|_{y=1} + \\ \qquad + H(u,v,1)\dfrac{\partial}{\partial y}\left(\dfrac{uv^2(1+u)^2}{(1-uv)^2}\right)\Big|_{y=1}\,. \end{cases}$$

Решив систему, найдём $H(u,v,1)$ и $\chi(u,v)$. Затем находим $A(u,v,y) = v(1+\chi(u,v)$ и $G(u,v,1) = A(u,v,1) + H(u,v,1)$.

Не будем утруждать читателя выписыванием всех функций двух переменных, для примера приведём две из них. Для простоты записи определим два многочлена от двух переменных

$$r(u,v) = 1 - 6uv - 2u^2 v - 2uv^2 + 8u^2 v^2 + 4u^3 v^2 + 4u^2 v^3 - 5u^3 v^3 - \\ - 6u^4 v^3 - 6u^3 v^4 - u^5 v^3 - u^3 v^5 - 5u^4 v^4 + u^5 v^5,$$

$$t(u,v) = u + v + 3uv - 2u^2 v - 2uv^2 - u^3 v - uv^3 - 7u^2 v^2 - u^3 v2 - u^2 v^3 + \\ + u^4 v^2 + u^2 v^4 + 5u^3 v^3 + 4u^4 v^3 + 4u^3 v^4 + 3u^4 v^4.$$

Тогда имеет место следующая теорема.

Теорема 2. *Производящими функциями двух переменных для множеств $H$ и $G$ горизонтально-выпуклых полиамондов будут функции*

$$H(u,v,1) = \frac{u(v+1)^2(1-uv)^2(1-uv-u^2v)}{r(u,v)}\,,$$

$$\widetilde{G}(u,v) = \sum_{p,q} g_{p,q}u^pv^q = G(u,v,1) = \frac{(1-uv)t(u,v)}{r(u,v)}\,.$$

Отметим, что рациональная функция $\widetilde{G}(u,v)$ является симметрической функцией относительно переменных $u, v$.

Сделав подстановку $u = v = x$, мы найдём производящие функции от одной переменной:

$$H(x,x,1) = \frac{x(x+1)(1-x)^2(1-x^2-x^3)}{1-3x+4x^3-x^4-x^5-3x^6+x^7}\,,$$

$$A(x,x,1) = \frac{x(1-x)(1-x)^2(1-x^2-x^3)}{1-3x+4x^3-x^4-x^5-3x^6+x^7}\,,$$

$$B(x,x,1) = \frac{x^3(1-2x^2+2x^5)}{(1+x)(1-3x+4x^3-x^4-x^5-3x^6+x^7)}\,,$$

$$C(x,x,1) = \frac{x(1-2x-2x^2+3x^3+x^4-3x^6-x^7)}{(1+x)(1-3x+4x^3-x^4-x^5-3x^6+x^7)}\,,$$

$$D(x,x,1) = \frac{x^2(1-x-2x^2+2x^3+x^4+x^5-x^6)}{(1+x)(1-3x+4x^3-x^4-x^5-3x^6+x^7)}\,.$$

и, конечно, получим следствие теоремы 2.

Следствие. *Производящей функцией для количества горизонтально-выпуклых полиамондов будет рациональная функция*

$$G(x) = \sum_{n=1}^{\infty} g_n x^n = \widetilde{G}(x,x) = G(x,x,1) = \frac{x(1-x)(2-x-4x^2+2x^4+3x^5)}{1-3x+4x^3-x^4-x^5-3x^6+x^7}\,. \tag{4.10}$$

Из (4.10) получаем

$$(1-3x+4x^3-x^4-x^5-3x^6+x^7)\sum_{n=1}^{\infty} g_n x^n = x(1-x)(2-x-4x^2+2x^4+3x^5).$$

Поскольку правая часть последнего равенства является многочленом, то и левая часть равенства должна быть многочленом. Следовательно, последовательность $g_n$ удовлетворяет рекуррентному соотношению, совпадающему с соотношением (3.1) из Теоремы 1

$$g_n = 3g_{n-1} - 4g_{n-2} + g_{n-4} + g_{n-5} + 3g_{n-6} - g_{n-7}\,.$$

Для последовательностей, заданных рекуррентными соотношениями, можно составить характеристическое уравнение (см., например, А. И. Маркушевич [3]). Для последовательности $g_n$ характеристическое уравнение будет следующим

$$x^7 - 3x^6 + 4x^4 - x^3 - x^2 - 3x + 1 = 0.$$

Это уравнение имеет действительные корни, наибольший из них

$$x_{\max} \approx 2.463536\,.$$

Поскольку $x_{\max}$ не является корнем числителя производящей функции, то имеем асимптотическую оценку $2.4635^n \leqslant g_n \leqslant 2.4636^n$. Что, естественно, лучше нижней оценки $2.13^n \leqslant g_n$, полученной Д. Кларнером в [8].

Остается отметить, что последовательности $a_n$, $h_n$ удовлетворяют рекуррентному соотношению седьмого порядка, аналогичному соотношению (3.1). В тоже время, как видно из производящих функций, последовательности $b_n$, $c_n$, $d_n$ будут удовлетворять рекуррентному соотношению восьмого порядка.

Чтобы найти производящую функцию двух переменных для количества горизонтально-выпуклых $n$-амондов с $m$ строками, как это сделал М. Деле [5] для полимино, нам следовало бы рассмотреть производящие функции от четырёх переменных. Повторив наши рассуждения, мы получили бы соотношение, аналогичное соотношению (4.9), но от четырёх переменных. Разрешив полученное соотношение, мы получим искомую рациональную функцию.

Для технических операций умножения многочленов и нахождения корней автором использовался портал `http://www.sagenb.org/` .

Автор благодарит К. А. Ванькова и П. И. Самовола за внимание к работе.

## Список литературы

[1] Голомб С. В. *Полимино*. М.: Мир. 1975.

[2] Гульден Я., Джексон Д. *Перечислительная комбинаторика*. М.: Наука. 1990.

[3] Маркушевич А. И. Возвратные последовательности. 3-е изд. М.: Наука, 1983. (серия «Популярные лекции по математике»).

[4] Bender E. *Convex n-ominoes* // Discrete Math. Vol. 8. 1974. P. 219–226.

[5] Delest M. P. *Generating Functions for Column-Convex Polyominoes* // Journal of Combinatorial Theory. Series A. Vol. 48. 1988. P. 12–31.

[6] Delest M. P., Viennot G. *Algebraic languages and polyominoes enumeration* // Theoret. Comput. Sci. Vol. 34. 1984. P. 169–206. [Русский перевод: М.-П. Делест, Ж. Вьенно. *Алгебраические языки и перечисление полимино* // Киб. сборник. Нов. сер. Вып. 26. М.: Мир. 1989. С. 113–156.]

[7] Hickerson D. *Counting Horizontally Convex Polyominoes* // Journal of Integer sequences. Vol. 2. 1999. Article 99.1.8.

[8] Klarner D. A. *Cell growth problems* // Canad. J. Math. Vol. 19. 1967. P. 851–863.

[9] Klarner D., Rivest R. *Asymptotic bounds for the number of convex n-ominoes* // Discrete Math. Vol. 8 1974. P. 31–40.

[10] Lunnon W. F. *Counting hexagonal and triangular polynominoes* // Graph Theory and Computing (R. Read, ed.). New York: Academic Press. 1972. P. 87–100.

В. М. Журавлев, 111141, Москва, ул. Кусковская, д. 17, кв. 32.
Email: ZhuravlevVM@mail.ru

# О некоторых свойствах производной, её характеризующих

Б. Кадец

Пусть $A$ и $B$ — две алгебры над одним полем, причём $B$ — подалгебра $A$. Линейный оператор $D\colon B \to A$ называется дифференцированием, если для любых $f, g \in A$ выполнено «правило Лейбница»: $D(fg) = f \cdot Dg + g \cdot Df$.

В математике подобные дифференцирования возникают при рассмотрении разных вопросов (см. например [3, 5, 6] и, для некоммутативного случая, [1]).

Известно, что любое дифференцирование алгебры $C^\infty(\mathbb{R})$ в себя — это, с точностью до нормировки, обычный оператор взятия производной [6, гл. 1, §2]. Однако аналог этого факта для других алгебр гладких функций не является широко известным. Мне не удалось найти в литературе описания дифференцирований из $C^1(\mathbb{R})$ в $C(\mathbb{R})$, а известные доказательства для $C^\infty$ на этот случай не переносятся. Возможно, что это описание можно получить из общих теорем теории дифференцирования банаховых алгебр (см. [1, 7]), однако мне такого общего утверждения найти не удалось.

В этой заметке приведено элементарное доказательство того, что любое дифференцирование из $C^1$ в $C$ после естественной нормировки совпадает с производной (нам даже не потребуется в полной мере условие линейности). Мы не знаем, верно ли аналогичное утверждение для дифференцирований из алгебры дифференцируемых функций в алгебру всех функций на $\mathbb{R}$.

Рассмотрим оператор $D\colon C^1(\mathbb{R}) \to C(\mathbb{R})$. Мы покажем, что, если он удовлетворяет одному из следующих наборов условий:

(1a) $D(fg) = f \cdot Dg + g \cdot Df$ («правило Лейбница»),

(2a) $D(a \cdot x + b \cdot \mathbf{1}) = a \cdot \mathbf{1}$,

или

(1b) $D(a \cdot x + b \cdot \mathbf{1}) = a \cdot \mathbf{1}$.

(2b) Существуют не сюръективная функция $\varphi \in C^1$ и точка $x_0$ такие, что $(D\varphi)(x_0) \neq 0$,

(3b) $D(f(g))(x) = (Df)(g(x)) \cdot (Dg)(x)$ («цепное правило»),

то для любой функции $f$ верно $Df = f'$. (Здесь и далее значком $\mathbf{1}$ обозначается функция, тождественно равная единице.)

Мы также покажем существование абстрактного дифференцирования на пространстве функций с локально ограниченной производной, не совпадающего с операцией взятия производной.

## 1. Две теоремы о совпадении абстрактного дифференцирования с обычным

Основным инструментом будет служить следующая простая лемма.

Лемма. *Пусть оператор* $D\colon C^1(\mathbb{R}) \to C(\mathbb{R})$ *удовлетворяет условиям*:
(1) *для любых* $a, b \in \mathbb{R}$ *выполнено* $D(ax + b\cdot\mathbf{1}) = a\cdot\mathbf{1}$;
(2) *для любого конечного интервала* $J = (s;t)$ *из* $f|_J = g|_J$ *следует, что* $(Df)|_J = (Dg)|_J$ (*локальность*).
*Тогда для любой функции* $f$ *верно* $Df = f'$.

Доказательство. Пусть $f \in C^1(\mathbb{R})$, $x_0 \in \mathbb{R}$, $I_1 = (x_0, +\infty)$, $I_2 = (-\infty, x_0)$, $g(x) = f(x_0) + f'(x_0)(x - x_0)$. Введём функцию $h$ такую, что $h|_{I_1} = f$, $h|_{I_2} = g$. Тогда $h$ непрерывна в точке $x_0$ и $\lim_{x\to x_0} h'(x) = f'(x_0)$. Поэтому $h \in C^1(\mathbb{R})$. Тогда, так как $I_1$ и $I_2$ есть (счётные) объединения интервалов, то $(Dh)|_{I_1} = (Df)|_{I_1}$, $(Dh)|_{I_2} = (Dg)|_{I_2} = f'(x_0)\cdot\mathbf{1}$. Но $Df$ и $Dh$ непрерывны, $x_0$ лежит на границе как $I_1$, так и $I_2$. Поэтому $(Df)(x_0) = f'(x_0)$. □

Теперь докажем следующую теорему.

Теорема 1. *Пусть оператор* $D\colon C^1(\mathbb{R}) \to C(\mathbb{R})$ *удовлетворяет условиям*:
(1) $D(fg) = f\cdot Dg + g\cdot Df$;
(2) $D(a\cdot x + b\cdot\mathbf{1}) = a\cdot\mathbf{1}$.
*Тогда для любой функции* $f$ *верно* $Df = f'$.

Доказательство. Докажем, что $D$ — «локальный» оператор. Пусть $I$ — интервал. Рассмотрим функцию $h \in C^1$ такую, что $h$ не равно нулю ни в одной точке интервала $I$, а $h|_{\mathbb{R}\setminus I} = 0$. Пусть $g|_I = f|_I$. Тогда $f\cdot h = g\cdot h$ на всей оси. Поэтому $f\cdot Dh + h\cdot Df = g\cdot Dh + h\cdot Dg$. Значит, $(Df)(x) = (Dg)(x)$ при всех $x \in I$. Таким образом, оператор $D$ удовлетворяет условиям леммы, поэтому $Df = f'$. □

Замечание. Без условия (2) теорема неверна. Действительно, отображение $(Df)(x) = L(f(x))$, где $L(x) = x\cdot\ln|x|$, $L(0) = 0$, удовлетворяет условию (1).

Попробуем теперь посмотреть, какие условия надо добавить к «цепному правилу» $(D(f(g))(x) = (Df)(g(x))\cdot(Dg)(x))$, чтобы полученные условия задавали оператор дифференцирования. Приведём сначала несколько примеров отображений, удовлетворяющих «цепному правилу».

Пример 1. $(Df)(x) = \dfrac{T(f(x))}{T(x)}$, где $T \in C^1$ — произвольная всюду положительная функция. Действительно,

$$D(f(g))(x) = \frac{T(f(g(x)))}{T(x)} = \frac{T(f(g(x)))}{T(g(x))} \cdot \frac{T(g(x))}{T(x)} = (Df)(g(x)) \cdot (Dg)(x).$$

Пример 2. Пусть $Df = f'$, если $f$ — биекция, и $Df$ тождественно равна 0 для остальных $f$ (для непрерывных функций из $\mathbb{R}$ в $\mathbb{R}$ композиция является биекцией только тогда, когда обе функции — биекции).

Последний пример объясняет естественность появления условия (2) в следующей теореме.

Теорема 2. *Пусть оператор $D\colon C^1(\mathbb{R}) \to C(\mathbb{R})$ удовлетворяет условиям*:

(1) $D(a \cdot x + b \cdot \mathbf{1}) = a \cdot \mathbf{1}$;

(2) *существует не сюръективная функция $\varphi \in C^1$ и точка $x_0$ такие, что $(D\varphi)(x_0) \neq 0$*;

(3) $D(f(g))(x) = (Df)(g(x)) \cdot (Dg)(x)$.

*Тогда $Df = f'$.*

Доказательство. Основой доказательства снова будет проверка локальности оператора. Будем считать, что образом $\varphi$ служит конечный отрезок, интервал или полуинтервал, внутренность которого мы обозначим $I$ (доказательство для случая открытого или замкнутого луча полностью аналогично).

Заметим, что из свойств (1) и (3) следует, что $D(f(ax + b\mathbf{1}))(x_0) = a \cdot D(f)(ax_0 + b)$ и $D(a \cdot f + b)(x) = aD(f)(x)$. Поэтому для любого интервала $J$ и любой точки $y \in J$ существует функция $f_0 \in C^1$ и точка $x_1$ такие, что $f_0(x_1) = y$, $(Df_0)(x_1) \neq 0$, $f_0(\mathbb{R}) \subset J$ ($f_0$ — это сдвинутая и растянутая функция $\varphi$).

Докажем теперь, что, если $f|_J = g|_J$, где $J$ — интервал, то $(Df)|_J = (Dg)|_J$. Зафиксируем произвольную точку $y \in J$. Выберем такую функцию $h$, что $h(x_1) = y$, $(Dh)(x_1) \neq 0$, $h(\mathbb{R}) \subset J$. Тогда $f(h(x)) = g(h(x))$ при всех $x \in \mathbb{R}$. Применяя оператор $D$ и подставляя $x = x_1$, получим $(Df)(y) \cdot (Dh)(x_1) = (Dg)(y) \cdot (Dh)(x_1)$, т. е. $(Df)(y) = (Dg)(y)$. Значит, оператор удовлетворяет условиям леммы, т. е. $Df = f'$. □

## 2. Фильтры и пример странного дифференцирования

Для доказательства следующей теоремы нам понадобится техника фильтров и ультрафильтров, хорошо известная специалистам в общей топологии, но редко упоминаемая в общих университетских курсах (подробнее см. [2, 4, 8]).

При рассмотрении функций, не имеющих предела (в обычном смысле), удобство предела по ультрафильтру заключается в таком единообразном выделении предельной точки, что выполняются арифметические свойства предела.

Напомним необходимые определения и факты.

Определение. Семейство подмножеств $\mathcal{F}$ множества $X$ называется *фильтром* на $X$, если

(1) $\mathcal{F}$ непусто;

(2) $\varnothing \notin \mathcal{F}$;

(3) если $A, B \in \mathcal{F}$, то $A \cap B \in \mathcal{F}$;

(4) если $A \in \mathcal{F}$, $A \subset B \subset X$, то $B \in \mathcal{F}$.

Определение. Семейство множеств $\mathcal{C}$ называется *центрированным*, если пересечение любого конечного набора элементов из $\mathcal{C}$ непусто.

Теорема. Каждое непустое центрированное семейство множеств содержится в некотором фильтре.

Определение. Пусть $X$ — множество, $\mathcal{F}$ — фильтр на $X$. Точка $y \in \mathbb{R}$ называется *пределом функции* $f\colon X \to \mathbb{R}$ по фильтру $\mathcal{F}$, если для любой окрестности $U$ точки $y$ существует $A \in F$ такое, что $f(A) \subset U$.

Определение. *Ультрафильтром* на $X$ называется максимальный по включению фильтр на $X$. То есть $\mathcal{F}$ — ультрафильтр, если не существует фильтра $\mathcal{U} \neq \mathcal{F}$ такого, что $\mathcal{F} \subset \mathcal{U}$.

Теорема. Любой фильтр содержится в некотором ультрафильтре.

Теорема (о существовании и единственности предела по ультрафильтру). Пусть $\mathcal{F}$ — ультрафильтр на $E$, $f : E \to \mathbb{R}$ функция, и образ некоторого элемента ультрафильтра относительно компактен (то есть лежит в каком-то отрезке). Тогда существует и единственен предел функции $f$ по $\mathcal{F}$.

Теорема (арифметические свойства предела по фильтру). Пусть $f, g$ — вещественнозначные функции на множестве $E$, $\mathcal{F}$ — фильтр на $E$, и существуют пределы функций $f$ и $g$ по фильтру $\mathcal{F}$, равные, соответственно, $a$ и $b$. Тогда существуют пределы по $\mathcal{F}$ функций $fg$ и $f + g$, равные, соответственно, $ab$ и $a + b$.

Вернёмся теперь к операторам, удовлетворяющим «условию Лейбница» $D(fg) = f \cdot Dg + g \cdot Df$. Обозначим через $DB_0(\mathbb{R})$ пространство дифференцируемых функций, имеющих ограниченную в окрестности нуля производную, а через $\mathbb{R}^{\mathbb{R}}$ — пространство всех функций из $\mathbb{R}$ в $\mathbb{R}$.

Теорема 3. *Существует оператор* $D\colon DB_0(\mathbb{R}) \to \mathbb{R}^{\mathbb{R}}$, *отличный от оператора дифференцирования и удовлетворяющий условиям* (1) *и* (2) *теоремы* 1.

Доказательство. Рассмотрим функцию $h(x) = x^2 \sin(1/x)$, и пусть $p(x) = h'(x)$, $b$ — предельная точка $p(x)$ при $x \to 0$, отличная от $p(0)$ (например, $b = \frac{1}{2}$). Рассмотрим семейство множеств вида

$$\{x : |p(x) - b| < \varepsilon\} \cap (-\delta, +\delta)$$

при различных $\varepsilon, \delta$. Оно центрировано, значит, существует содержащий это семейство ультрафильтр $\mathcal{F}$. Очевидно, что образ этого ультрафильтра под действием $p$ содержит базу окрестностей точки $b$, т. е. $\lim_{\mathcal{F}}(p) = b$. Теперь можно рассмотреть оператор $D$, такой что $(Df)(x) = f'(x)$ при $x \neq 0$ и $(Df)(0) = \lim_{\mathcal{F}}(f'(x))$. Так как любая функция из $DB_0(\mathbb{R})$ имеет ограниченную производную в окрестности нуля (а окрестности нуля — это элементы ультрафильтра), то по теореме о существовании предела по ультрафильтру $\lim_{\mathcal{F}}(f'(x))$ существует. Из арифметических свойств предела по фильтру следует, что равенство $D(fg)(x) = (f(Dg) + g(Df))(x)$ верно и при $x = 0$. Но $(Dh)(0) = b \neq 0 = h'(0)$. □

Мне не известен ответ на следующий вопрос.

Вопрос. Верно ли утверждение теоремы 3 для операторов из множества всех дифференцируемых функций в $\mathbb{R}^{\mathbb{R}}$?

Выражаю глубокую благодарность С. Л. Гефтеру за постановку задачи и ценное обсуждение.

## Список литературы

[1] Браттели У., Робинсон Д. *Операторные алгебры и квантовая статистическая механика*. М.: Мир. 1982.

[2] Бурбаки Н. *Общая топология. Основные структуры*. М.: Наука, Гл. ред. физ.-мат. лит. 1968.

[3] Винберг Э. Б., Онищик А. Л. *Семинар по группам Ли и алгебраическим группам*. М.: Наука. Гл. ред. физ.-мат. лит. 1988.

[4] Кадец В. М. *Курс функционального анализа*. Х.: ХНУ имени В.Н. Каразина. 2006. `http://page.mi.fu-berlin.de/werner99/kadetsbook/`

[5] Капланский И. *Введение в дифференциальную алгебру*. Библиотека сборника «Математика». М.: Из-во иностр. лит. 1959.

[6] Хелгасон С. *Дифференциальная геометрия и симметрические пространства*. М.: Мир. 1964.

[7] Dales H. G. *Automatic continuity: a survey* // Bull. London Math. Soc. Vol. 10. 1978. P. 129–183.

[8] Comfort W.W., Negrepontis S. *The theory of ultrafilters*. Berlin, New York: Springer-Verlag. 1974.

Б. Кадец, Харьковский национальный университет им. В.Н. Каразина
Email: borja.kadets@gmail.com

# О множествах с двумя расстояниями

А. В. Акопян* О. Р. Мусин†

В этой заметке пойдёт речь о множестве точек в пространстве или на сфере, расстояния между которыми принимают не более чем два значения. Обсуждается вопрос о том, как много точек может иметь такое множество, а также какие конфигурации образуют точки из экстремальных наборов.

## Введение

Сколько точек можно выбрать в $d$-мерном евклидовом пространстве так, что расстояния между любыми двумя из них будут равны? Легко видеть, что можно выбрать не более $d+1$ точек, а вершины $d$-мерного симплекса как раз и будут образовывать такое множество. Правильный симплекс вписан в $(d-1)$-мерную сферу, и поэтому на такой сфере тоже можно выбрать $(d+1)$ точку так, что расстояния между любыми двумя точками принимает фиксированное значение.

Отметим, что вопрос о том, какое максимальное количество точек на одинаковом друг от друга расстоянии гарантировано можно выбрать в $d$-мерном банаховом пространстве остаётся открытым (см. обзоры [16,17]).

Множество $\mathcal{S}$ в $\mathbb{R}^d$ или $\mathbb{S}^d$ (или любом другом метрическом пространстве) называется *множеством с $s$ расстояниями*[1)], если расстояния между его точками принимают не более $s$ значений.

В этой статье в основном обсуждаются множества с двумя расстояниями. Оказывается, что, несмотря на такое, в общем-то, достаточно простое, ограничение, эти множества обладают рядом довольно интересных и иногда неожиданных свойств.

Отметим, что Эйнхорн и Шонберг в [9, 10] показали, что существует лишь конечное число множеств (с точностью до подобия) с двумя расстояниями в $\mathbb{R}^d$, состоящими из более чем $d+2$ точек. Заметим, что подобный

*Работа выполнена при частичной поддержке фонда «Династия», грантов РФФИ 11-01-00735 и 12-01-31281 и гранта Правительства РФ №11.G34.31.0053.

†Работа выполнена при частичной поддержке гранта 11-01-00735 и гранта Правительства РФ №11.G34.31.0053.

[1)] По-английски *s-distance set.*

результат для множеств с $s > 2$ расстояниями был получен Х. Нозаки совсем недавно [19].

Имеется пример множества с двумя расстояниями в $\mathbb{R}^d$, состоящего из $C^2_{d+1} = d(d+1)/2$ точек. Мы будем обозначать это множество $S_d$. Рассмотрим правильный симплекс в $\mathbb{R}^d$, у которого длины всех рёбер равны 1. У этого симплекса всего $d(d+1)/2$ рёбер. Их середины будут образовывать множество с двумя расстояниями. Действительно, если два ребра имеют общую вершину, то расстояние между их серединами равно $1/2$ (поскольку соединяющий их отрезок будет средней линией треугольника, образованного вершинами этих рёбер). Если не имеют, то $1/\sqrt{2}$, поскольку в этом случае вершины этих рёбер являются вершинами правильного трёхмерного тетраэдра, а расстояние между серединами противоположных рёбер правильного тетраэдра именно таково.

Это множество можно описать также с помощью ортонормированного базиса $e_1, e_2, \ldots, e_{d+1}$ пространства $\mathbb{R}^{d+1}$. Рассмотрим точки вида

$$e_i + e_j \quad (1 \leqslant i < j \leqslant d+1).$$

Расстояние между такими точками может быть равно либо $\sqrt{2}$, либо 2, в зависимости от того, имеют ли они общую единицу в координатной записи или нет. Сумма координат получившихся $d(d+1)/2$ точек будет равна 2 и поэтому они будут лежать в гиперплоскости, задаваемой уравнением $x_1 + \ldots + x_{d+1} = 2$.

Заметим, что если $a$ и $b$ $(a < b)$ — два расстояния множества $S_d$, то $b^2/a^2 = 2$. Оказывается, что подобное свойство верно для всех достаточно больших множеств с двумя расстояниями.

Ларман, Роджерс и Зейдель в [14] доказали, что если множество с двумя расстояниями $a$ и $b$ $(a < b)$ в $\mathbb{R}^d$ состоит из более чем $2d+3$ точек, то

$$\frac{a^2}{b^2} = \frac{k-1}{k}\,, \text{ где } k \in \mathbb{N}, \quad \text{и} \quad 2 \leqslant k \leqslant \frac{1+\sqrt{2d}}{2}\,. \tag{1}$$

Недавно этот соотношение было обобщено Нозаки [19] на случай множеств с $s$ расстояниями.

Естественно, что нас будет интересовать максимальная мощность множеств с $s$ расстояниями. Часто экстремальные конфигурации, то есть множества с $s$ расстояниями, состоящие из максимального возможного числа точек, оказываются весьма любопытными и граф расстояний между точками имеет высокую степень симметрии. Это одна из причин, почему интересно исследовать вопросы о мощности и виде таких множеств.

Отметим, что верхние оценки на мощность множеств с $s$ расстояниями в $\mathbb{R}^d$ стали известны около 30 лет назад. В частности, Блокхаус доказал, что число точек у множества $\mathcal{S}$ с двумя расстояниями в $\mathbb{R}^d$ не превосходит

$(d+1)(d+2)/2$ (см. ниже теорему 1). Как показал Лисонек (см. раздел 3), эта оценка достигается в размерности 8.

В работе [8] были получены оценки для случая, когда точки множества $\mathcal{S}$ лежат на сфере в $\mathbb{R}^d$. (Мы будем называть такие множества *сферическими множествами с двумя расстояниями*.) В этом случае оценка будет $d(d+3)/2$ (см. теорему 2). Заметим, что эта оценка достигается для $d = 2$, 6 и 22.

В разделе 3 мы разбираем работу Лисонека по множествам с двумя расстояниями в $\mathbb{R}^d$ для $d \leqslant 8$. Кроме верхних оценок и работы Лисонека, основанной на компьютерном переборе, практически никаких результатов для максимальных евклидовых множеств с двумя расстояниями нет.

В отличии от евклидова, для сферического случая имеется значительный прогресс. Мы немного обсудим это в разделе 5. В работе [18] было показано, что для $6 < d < 22$ и $23 < d < 40$ количество точек не превосходит $d(d+1)/2$. Следовательно, в этих размерностях множество $S_d$ является максимальным. Недавно этот результат был расширен для $d = 23$ и $40 \leqslant d \leqslant 93$ ($d \neq 46, 78$) в работе А. Барга и В.-Ш. Ю [4].

## 1. Множества с двумя расстояниями для $d \leqslant 3$.

В случае, когда наше пространство — это прямая, легко видеть, что максимальная мощность множества с двумя расстояниями равна трём. Экстремальной конфигурацией тут служат концы отрезка и точка в его середине.

На плоскости такое множество может уже состоять из пяти точек — вершин правильного пятиугольника. Два расстояния — это длина стороны пятиугольника и длина его диагонали. В следующем разделе мы покажем, что эта конфигурация единственна.

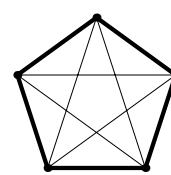


*Рис. 1.*

В трёхмерном пространстве максимальная мощность не сильно больше и равна шести. Это показал Крофт в [7]. Оказывается, что размерность три в данном случае является исключительной ситуацией, поскольку таких множеств из шести точек не одно, а целых шесть штук. Перечислим их.

Во-первых, существует два множества с отношением длин, равным $\sqrt{2}$. Первое — это вершины правильного октаэдра. Другой пример — это

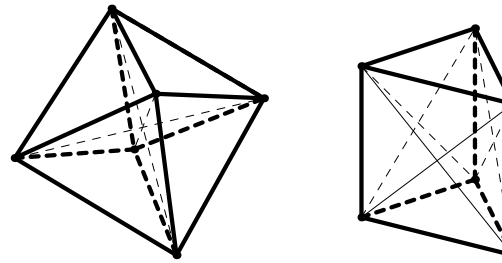


*Рис. 2.*

призма, основание которой — правильный треугольник, а боковые рёбра равны стороне это треугольника (рис. 2).

Есть ещё четыре довольно любопытных примера, которые получаются из правильного икосаэдра. В правильном икосаэдре 12 вершин (рис. 3). Для каждой вершины есть 5 соседних вершин (давайте считать, что расстояние до них равно 1), противоположная и 5 соседних с противоположной (расстояние до них равно $\sqrt{5/2}$). Давайте из каждой пары противоположных вершин выберем одну. Тогда мы получим 6 точек, между

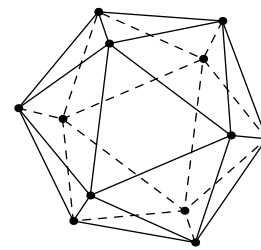


*Рис. 3.*

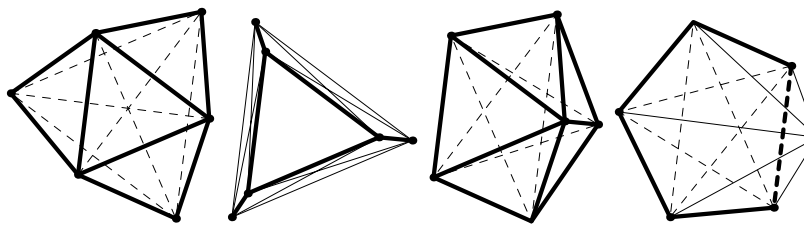


*Рис. 4. Четыре примера множеств с двумя расстояниями, получающимися из правильного икосаэдра. Жирные линии обозначают расстояния, равные* 1, *а тонкие* — $\sqrt{5}/2$

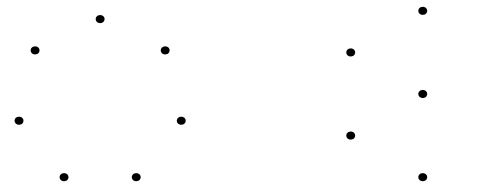


*Рис. 5. Множества из 7 точек на плоскости, между которыми только три вида расстояний*

которыми расстояние либо 1, либо $\sqrt{5/2}$. Таким образом, мы можем получить четыре новые конструкции, изображённые на рис. 4.

Интересно, что для плоскости и пространства также решена задача о множествах с тремя расстояниями. На плоскости оно может состоять максимум из семи точек, и существует две такие конфигурации: вершины правильного семиугольника и вершины правильного шестиугольника и его центр (рис. 5). Полностью все множества с тремя расстояниями на плоскости, состоящие из пяти и более точек классифицировал Шинохара [21] (их оказалось 24 штуки). Он же показал [20], что в трёхмерном пространстве единственное максимальное множество с тремя расстояниями — это икосаэдр (рис. 3).

Примеры множеств с двумя расстояниями в больших размерностях, обсуждаются в разделе 3.

## 2. Максимальное множество с двумя расстояниями на плоскости

Приведём здесь доказательство того, что пять точек на плоскости с двумя расстояниями, всегда образуют правильный пятиугольник. Первым это доказал Келли [12], решая похожую задачу Эрдёша из задачника American Mathematical Monthly. Задача Эрдёша состояла в следующем: сколько точек можно расположить на плоскости или в пространстве так, чтобы любой треугольник с вершинами в них был равнобедренным. Множество с двумя расстояниями, очевидно, является примером такого множества. Обратное неверно. В частности, ответом на вопрос Эрдёша для плоскости будет 6 — правильный пятиугольник и его центр. Келли показал, что эта конструкция из шести точек будет единственной. В трёхмерном пространстве задачу Эрдёша решил Крофт [7], показав, что в этом случае точек не более 8. Множество из 8 точек существует: правильный пятиугольник в горизонтальной плоскости, его центр и две дочки над и под центром на расстоянии равным радиусу описанной окружности пятиугольника. Однако доказать, что это единственная восьмиточечная

конфигурация удалось доказать лишь сравнительно недавно (Кидо [13]). Позже Ю. И. Ионину удалось получить решение этой задачи для пространств размерности не большей 8 [11]. Более подробно об этой задаче и похожих вопросах можно прочитать в его недавней статье [2].

Вернёмся к нашей задаче о множестве с двумя расстояниями на плоскости. Очевидно, что к вершинам правильного пятиугольника нельзя будет добавить ещё одну точку, так чтобы число различных расстояний не увеличилось. Поэтому, из единственности будет следовать, что максимальная мощность множества с двумя расстояниями на плоскости равна пяти.

Итак пусть дано множество $\mathcal{S}$ из 5 точек на плоскости, таких, что между ними только два расстояния $a$ и $b$ ($a < b$). Заметим, что среди этих точек нет трёх, образующих правильный треугольник. Действительно, пусть три точки $A$, $B$ и $C$ образуют правильный треугольник. Тогда про любую точку $X$ из двух оставшихся точек $\mathcal{S}$, можно сказать, что
а) из трёх расстояний до $A$, $B$ и $C$ два должны быть равны, поэтому $X$ лежит на одном из серединных перпендикулярах к сторонам;
б) все три расстояния до $A$, $B$ и $C$ не могут быть одинаковыми (случай, когда $X$ центр $\triangle ABC$ легко исключается, так как пятую точку к этому набору добавить нельзя), поэтому одно из них должно быть равно $|AB|$. Следовательно эта точка лежит на окружности радиуса $|AB|$ с центром в одной из вершин треугольника.

Таким образом, кандидатами в наше множество могут быть только 9 точек, отмеченных на рис. 6. Легко заметить, что в множестве с двумя расстояниями точки должны быть одного типа (т. е. одинаково отмечены на рис. 6). Но, проанализировав все три возможные в этом случае конструкции, мы видим, что они имеют три различных расстояния.

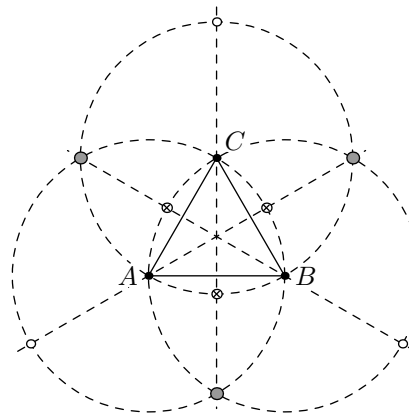



*Рис. 6.*

Заметим также, что для каждой точки $X$ из множества $\mathcal{S}$, существует ровно две точки из $\mathcal{S}$ на расстоянии $a$ и две на расстоянии $b$. Действительно, если три точки $A$, $B$ и $C$ находятся на расстоянии, скажем, $a$ от $X$, то либо между какими-то двумя из них расстояние $a$ и они образуют правильный треугольник вместе с точкой $X$, либо между ними всеми расстояние $b$ и $\triangle ABC$ правильный.

Раз так, то получаем, что если соединить отрезками вершины на расстоянии $b$, то они образуют пятизвенную ломаную, при этом каждое ребро этой ломаной является диаметром $\mathcal{S}$. Но любые два диаметра множества на плоскости должны либо пересекаться, либо иметь общий конец (если два диаметра не пересекаются, то из каждого можно выбрать по вершине так, что расстояние между выбранными вершинами будет больше). Следовательно, эта ломаная устроена звездой, как изображено на рис. 7. Осталось заметить, что все стороны образовавшегося пятиугольника должны быть равны $a$. Отсюда уже легко получить равенство его углов, а значит, этот пятиугольник правильный.

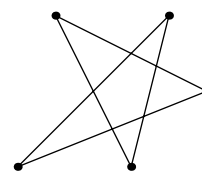


*Рис. 7.*

## 3. Множества с двумя расстояниями в пространствах размерности $d \leqslant 8$

В 1997 году Пётр Лисонек опубликовал интересную статью [15], в которой он с помощью компьютера показал единственность (с точностью до подобия) максимальных множеств с двумя расстояниями в размерностях 4, 5, 6 и 7 (сами конструкции были известны ещё до него), а в размерности 8 он обнаружил множество состоящее из 45 точек, т. е. максимально возможное согласно границе Блокхауса (см. теорему 1 в разделе 4).

Рассмотрим эту работу более подробно. Переберём все максимальные множества с двумя расстояниями в $\mathbb{R}^d$, где $d \leqslant 8$. Мы уже знакомы с такими множествами для $d = 2$ и $d = 3$.

**d = 4**. В этом случае число точек в множестве с двумя расстояниями не превосходит 10. И единственным множеством из 10 точек является описанное выше множество $S_5$, состоящее из середин рёбер правильного

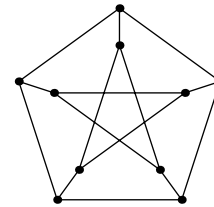


*Рис. 8.*

симплекса. Интересно, что граф рёбер большей длины в этой конфигурации образует знаменитый граф Петерсена (рис. 8).

**d = 5**. В этой размерности можно получить больше точек, чем даёт нам множество из середин рёбер симплекса — 16 вместо 15. Для этого надо «подкрутить» четыре точки, что позволит добавить ещё одну. Опишем эту конструкцию явно (она же является единственной максимальной в пятимерном пространстве). Пусть $e_1$, $e_2$, $e_3$, $e_4$ и $e_5$ — это ортонормированный базис. Наше множество будет состоять из пяти точек вида $-e_i + \sum_{i=j}^{5} e_j$ $(1 \leqslant i \leqslant 5)$, 10 точек вида $e_i + e_j$ $(1 \leqslant i < j \leqslant 5)$ и начала координат. Легко видеть, что между точками этого множества расстояние либо $\sqrt{2}$, либо 2. Кроме того, все они лежат на сфере радиуса $\sqrt{5}/2$ с центром в $\frac{1}{2}\sum_{i=1}^{5} e_i$.

**d = 6**. Оказывается, что в семимерном пространстве существует 28 прямых, проходящих через начало координат, таких, что угол между ними одинаковый и равен $\arccos 1/3$. Как построить такие прямые? Рассмотрим прямые в $\mathbb{R}^8$, выходящие из начала координат в направлениях вдоль векторов, у которых две координаты равны 3, а остальные шесть равны $-1$. Таких прямых $C_8^2 = 28$ и все они лежат на гиперплоскости задаваемой уравнением $\sum_{i=1}^{8} x_i = 0$. Все векторы, порождающие эти прямые, имеют длину $\sqrt{24}$, а скалярное произведение этих векторов равно либо 8, либо $-8$. Значит, угол между этими векторами либо $\arccos 1/3$, либо $-\arccos 1/3$.

Выберем единичный вектор $e$ на одной из прямой и 27 единичных векторов на оставшихся прямых, так что они образуют с $e$ тупой угол. Легко понять, что концы этих векторов лежат на шестимерной плоскости, перпендикулярной $e$, и расстояния между ними равны либо $\sqrt{4/3}$, либо $\sqrt{8/3}$. (И опять же, видно, что все они лежат на одной сфере.)

**d = 7**. В этой размерности Лисонек обнаружил, что максимальное множество единственно и состоит оно из 29 точек. Как и в размерности пять, можно «подкрутить» конструкцию, образованную серединами рёбер правильного симплекса и добавить ещё одну точку.

Мы опишем это множество с помощью ортонормированного базиса $e_1, e_2, \ldots, e_8$ пространства $\mathbb{R}^8$. Сумма координат получившихся точек будет равна 2 и поэтому они будут лежать в семимерной гиперплоскости, задаваемой уравнением $x_1 + \ldots + x_8 = 2$.

7 точек нашего множества задаются так:

$$-e_i + \frac{1}{2}\sum_{k=1}^{7} e_k - \frac{1}{2}e_8 \quad (1 \leqslant i \leqslant 7).$$

21 точка имеет вид:

$$e_i + e_j \quad (1 \leqslant i < j \leqslant 7),$$

и 29-я точка — это точка

$$\frac{1}{2}\sum_{k=1}^{7} e_k - \frac{3}{2}e_8.$$

Расстояния между точками опять же равны либо $\sqrt{2}$, либо 2, но в отличие от предыдущих случаев, на этот раз точки не будут лежать на одной сфере.

**d = 8**. Как и в предыдущем случае, зададим это множество в координатах пространства большей размерности. Пусть $e_1, e_2, \ldots, e_9$ ортонормированный базис в $\mathbb{R}^9$. Сумма координат построенных точек опять будет равна 2, то есть они будут лежать на гиперплоскости, задаваемой уравнением $x_1 + \ldots + x_9 = 2$.

9 точек задаются с помощью таких соотношений

$$-e_i + \frac{1}{3}\sum_{k=1}^{9} e_k \quad (1 \leqslant i \leqslant 9).$$

Остальные 36 образуют множество $S_8$ и имеют вид

$$e_i + e_j \quad (1 \leqslant i < j \leqslant 9).$$

Как и раньше расстояния между точками будут равны либо $\sqrt{2}$, либо 2. Этот пример является максимальным, поскольку на нём достигается граница Блокхауса (теорема 1). Но является ли эта конструкция единственной? Кроме того, отметим, что получившиеся точки будут лежать на двух концентрических сферах. Может быть, это верно и для других экстремальных конфигураций в бóльших размерностях (эта гипотеза была высказана несколько лет назад А. А. Глазыриным и О. Р. Мусиным)?

Метод перебора основан на соотношении (1), благодаря которому мы можем считать, что знаем отношения между расстояниями. После чего надо рассматривать различные графы и проверять, реализуются ли они в пространстве соответствующей размерности (для этого существуют классические методы, основанные на положительной определённости матрицы скалярных произведений).

В статье Лисонека также приводятся данные о времени, потраченном на вычисления на довольно мощном по меркам 1996 года компьютере (DEC Alpha 2000 4/275 с 256 MB RAM и суперкомпьютер IBM/SP2 с 40 узлами RS6000 по 256 MB RAM в каждой). Так на случай $d = 4$ ушло всего 0.2 секунды, на $d = 5 - 7$ секунд, на $d = 6 - 12$ минут, а на $d = 7$ понадобилось аж 40 дней.

## 4. Верхние оценки на мощность множества с двумя расстояниями в $\mathbb{R}^d$

Теорема 1. *Пусть $\mathcal{S}$ — это множество с двумя расстояниями в $\mathbb{R}^d$. Тогда количество точек в $\mathcal{S}$ не превосходит $(d+1)(d+2)/2$ (что равно $C^2_{d+2}$).*

Первоначально оценка $(d+1)(d+4)/2$ была получена Ларманом, Роджерсом и Зейделем в [14]. Потом Блокхаус [6] слегка дополнил это доказательство и получил оценку из теоремы 1. Мы сейчас приведём эти довольно простые рассуждения.

Доказательство теоремы 1. Для каждой точки $p \in \mathcal{S}$ заведём многочлен от $d$ переменных:

$$F_p(x) = (\|x-p\|^2 - a^2)(\|x-p\|^2 - b^2), \tag{2}$$

где $a$ и $b$ — это те самые два расстояния.

Легко видеть, что эти многочлены линейно независимы, поскольку, подставляя в качестве переменной $x$ точку $p$, мы получим, что все многочлены $F_{p'}$, $p' \in \mathcal{S}$, $p' \neq p$, обнуляются, а значение $F_p$ в этой точке равно $a^2b^2$. Заметим также, что они является линейными комбинациями следующих функций:

$$\|x\|^4,\ \ \|x\|^2 x_i,\ \ x_i x_j,\ \ x_i,\ \ 1. \tag{3}$$

Но размерность линейного пространства, задаваемого этими функциями, равна $1 + d + d(d+1)/2 + d + 1 = (d+1)(d+4)/2$.

Это рассуждение было приведено в работе [14]. Теперь приведём уточнение, придуманное Блокхаусом в [6].

Покажем, что набор функций $\{F_p, x_i, 1\}$ также является линейно независимым. Отсюда будет следовать оценка $(d+1)(d+4)/2 - d - 1 = (d+1)(d+2)/2 = C^2_{d+2}$.

Итак, пусть

$$\sum_{p\in\mathcal{S}} c_p F_p(x) + \sum_{i=1}^{d} c_i x_i + c = 0. \tag{4}$$

Заметим, что поскольку многочлен — тождественный нуль, коэффициент при $\|x\|^4$ должен равняться нулю, поэтому $\sum\limits_{p\in\mathcal{S}} c_p = 0$.

Подставляя $p$ в эти соотношения, получаем

$$a^2b^2c_p + \sum_{i=1}^{d} c_i p_i + c = 0, \tag{5}$$

где $p_i$ — это $i$-ая координата точки $p$. Умножая соотношение (5) на $c_p$ и суммируя по всем $p \in \mathcal{S}$, получаем

$$a^2b^2 \sum_{p\in\mathcal{S}} c_p^2 + \sum_{i=1}^{d} c_i \sum_{p\in\mathcal{S}} c_p p_i + c\sum_{p\in\mathcal{S}} c_p = 0. \tag{6}$$

Заметим, что последние два члена с левой стороны обнуляются, поэтому $\sum\limits_{p\in\mathcal{S}} c_p^2 = 0$. Откуда все $c_p = 0$, а значит и все $c_i$ и $c$ равны нулю, что и завершает доказательство.

Для множеств с $s$ расстояниями это оценка была обобщена в работах [3, 5]. Развивая этот метод, было показано, что $|\mathcal{S}| \leqslant C_{d+s}^{s}$.

В случае сферических множеств с двумя расстояниями оценку можно уменьшить ещё на 1, для этого достаточно показать, что набор построенных многочленов ещё линейно независим от многочлена $\|x\|^2 - 1$.

Теорема 2. *Пусть $\mathcal{S}$ — это множество с двумя расстояниями, располагающееся на сфере в $\mathbb{R}^d$. Тогда количество точек в $\mathcal{S}$ не превосходит $d(d+3)/2$.*

Доказательство. Без ограничения общности можно считать, что точки множества $\mathcal{S}$ лежат на единичной сфере. Как и в доказательстве теоремы 1 рассмотрим многочлены задаваемые в (2). Покажем, что функции из набора $\{F_p, x_i, 1, \|x\|^2 - 1\}$ линейно независимы.

Пусть это не так. Тогда

$$\sum_{p\in\mathcal{S}} c_p F_p(x) + \sum_{i=1}^{d} c_i x_i + c + c'(\|x\|^2 - 1) = 0. \tag{7}$$

Опять же, поскольку коэффициент при $\|x\|^4$ равен 0, имеем $\sum_{p\in\mathcal{S}} c_p = 0$. Поскольку $\|p\|^2-1=0$, как и раньше, подставляя в это соотношение $x=p$, $p\in\mathcal{S}$, мы получим соотношения (5).

Конец доказательства дословно совпадает с доказательством предыдущей теоремы. Мы получаем, что все $c_p$, все $c_i$ и $c$ равны нулю. А значит, и $c'=0$.

На самом деле, теорема 2 была доказана раньше теоремы 1 и сразу для множеств с несколькими расстояниями. Приведём здесь это доказательство.

Теорема 3 (Дельсарт, Гуталс и Зейдель, [8]). *Пусть $\mathcal{S}$ — это сферическое множество с $s$ расстояниями в $\mathbb{R}^d$. Тогда количество точек в $\mathcal{S}$ не превосходит $C^{d-1}_{d+s-1}+C^{d-1}_{d+s-2}$.*

Доказательство. Опять же, можно считать, что все точки располагаются на единичной сфере. И тогда расстоянию между двумя точками соответствует значение скалярного произведения их векторов. Поэтому можно считать, что у нас задано семейство единичных векторов, скалярные произведения которых принимают одно из $s$ значений $a_i$, $i=1,\dots,s$.

Каждой точке $p\in\mathcal{S}$ сопоставим многочлен степени $s$ следующим образом

$$F_p(x)=\prod_{i=1}^{s}(\langle x,p\rangle - a_i). \tag{8}$$

Заметим, что $F_p(x)$ равен нулю во всех точках множества $\mathcal{S}$, кроме $p$, в которой он равен произведению $a_i$ (умноженному на $-1$, если $s$ нечётно). Покажем, что эти многочлены линейно независимы, причём порождаемое ими линейное подпространство пересекается с пространством многочленов вида $(\|x\|^2-1)P(x)$, где $P(x)$ многочлены степени не больше $s-2$, только по 0 (это можно интерпретировать так, как будто мы рассматриваем многочлены, определённые только на сфере). А поскольку размерность пространства многочленов степени не больше $s$ равна $C^d_{d+s}$, а размерность подпространства многочленов вида $(\|x\|^2-1)P(x)$ равна размерности подпространства многочленов степени не большей $s-2$, то есть $C^d_{d+s-2}$, получаем, что размерность линейной оболочки $F_p(x)$ не больше, чем

$$\begin{aligned} C^d_{d+s}-C^d_{d+s-2} &= C^d_{d+s-1}+C^d_{d+s-1}-C^d_{d+s-2} = \\ &= C^d_{d+s-1}+C^d_{d+s-2}+C^{d-1}_{d+s-2}-C^d_{d+s-2} = C^d_{d+s-1}+C^{d-1}_{d+s-2}\,. \end{aligned} \tag{9}$$

Здесь мы пользуемся известным соотношением $C^n_m = C^n_{m-1}+C^{n-1}_{m-1}$.

Итак, пусть

$$\sum_{p\in\mathcal{S}} c_p F_p(x) + (\|x\|^2 - 1)P(x) = 0 \tag{10}$$

для некоторого ненулевого набора $c_p$ и многочлена $P(x)$. Тогда, подставляя $p \in \mathcal{S}$ в качестве $x$, мы обнулим все слагаемые в левой части этого равенства кроме $c_p F_p(p)$. Поскольку $F_p(p)$ не равно нулю, получаем, что $c_p = 0$. Поскольку это верно для всех $p$, получаем противоречие, что и доказывает теорему.

## 5. Недавние результаты по сферическим множествам с двумя расстояниями

Один из авторов этой статьи усилил верхнюю оценку для случая сферических множеств с двумя расстояниями, когда сумма этих расстояний не превосходит $\pi$.

Теорема 4 (О. Р. Мусин [18]). *Пусть $\mathcal{S}$ — это множество с двумя расстояниями, располагающееся на единичной сфере в $\mathbb{R}^d$, причём сумма этих (угловых) расстояний не превосходит $\pi$. Тогда количество точек в $\mathcal{S}$ не превосходит $d(d+1)/2$ (то есть $C^2_{d+1}$).*

Доказательство. Будем считать, что скалярные произведения, соответствующие указанным расстояниям, равны $a$ и $b$. Тогда условие, что сумма этих расстояний не превосходит $\pi$, эквивалентно тому, что $a+b \geqslant 0$.

Как и в предыдущих доказательствах, каждой точке $p \in \mathcal{S}$ сопоставим многочлен

$$F_p(x) = (\langle x, p\rangle - a)(\langle x, p\rangle - b). \tag{11}$$

Как мы знаем, эти многочлены линейно независимы и имеют степень 2. Покажем, что кроме этого они линейно независимы с однородными линейными функциями и многочленом $\|x\|^2 - 1$.

Пусть

$$\sum_{p\in\mathcal{S}} c_p F_p(x) + \langle v, x\rangle + c(\|x\|^2 - 1) = 0, \tag{12}$$

где без ограничения общности можно считать, что $\|v\| = 1$. Тогда, подставляя в качестве $x$ точки $p \in \mathcal{S}$, получим следующие соотношения:

$$\sum_{p\in\mathcal{S}} c_p F_p(p) + \langle v, p\rangle = 0, \text{ то есть } c_p = \frac{-\langle v, p\rangle}{(1-a)(1-b)}\,. \tag{13}$$

Теперь подставим в соотношение (11) в качестве $x$ точки $v$ и $-v$:

$$\sum_{p\in\mathcal{S}} \frac{-\langle v,p\rangle}{(1-a)(1-b)} F_p(v) + 1 = 0; \tag{14}$$

$$\sum_{p\in\mathcal{S}} \frac{-\langle v,p\rangle}{(1-a)(1-b)} F_p(-v) - 1 = 0. \tag{15}$$

Вычитая (15) из (14), получаем

$$0 = \sum_{p\in\mathcal{S}} \frac{-\langle v,p\rangle}{(1-a)(1-b)} (F_p(v) - F_p(-v)) + 2 =$$

$$= \sum_{p\in\mathcal{S}} \frac{-\langle v,p\rangle}{(1-a)(1-b)} (-2a-2b)\langle v,p\rangle + 2 = \sum_{p\in\mathcal{S}} \frac{2(a+b)\langle v,p\rangle^2}{(1-a)(1-b)} + 2, \tag{16}$$

чего не может быть, поскольку $a+b \geqslant 0$ и, значит, $\dfrac{2(a+b)\langle v,p\rangle^2}{(1-a)(1-b)}$ величина неотрицательная.

Теорема 2 даёт верхнюю оценку на мощность множества с двумя расстояниями на сфере в $\mathbb{R}^d$, равную $d(d+3)/2$. Мы уже видели, что эта оценка достигается при $d = 6$. Подобно случаю $d = 6$ строится пример для $d = 22$ [8]. Других примеров, когда достигается оценка $d(d+3)/2$, не известно.

С другой стороны, у нас есть универсальный пример сферического множества с двумя расстояниями $S_d$ — середины рёбер симплекса. В этом случае количество точек равно $d(d+1)/2$.

В работе [18] рассматривалась задача о максимальных множествах $\mathcal{S}$ с двумя расстояниями на сфере. По теореме 4, если сумма расстояний не превосходит $\pi$, то количество точек в $\mathcal{S}$ не превосходит $d(d+1)/2$. Таким образом, остаётся рассмотреть случай $a+b<0$.

Из теоремы Лармана – Роджерса – Зейделя следует, что $b = (ka-1)/(k-1)$, где $k = 2,\ldots,\lfloor(1+\sqrt{2d})/2\rfloor$. Далее в [18] применяется метод Дельсарта (см., например, [1]) для оценки мощности множества точек с такими расстояниями. Этот метод для $6 < d < 22$ и $23 < d < 40$ даёт оценку меньшую чем $d(d+1)/2$. Следовательно, в этих размерностях у максимального множества $a+b \geqslant 0$ и множество середин рёбер симплекса является максимальным.

Недавно в работе [4] этот результат был расширен на $d = 23$ и $40 \leqslant d \leqslant 93$ $(d \neq 46, 78)$.

Заметим, что 6, 22, 46, 78 являются числами вида $(2m+1)^2 - 3$, где $m = 1, 2, 3, 4$. Это приводит к следующей гипотезе:

*Мощность максимального сферического множества с двумя расстояниями в* $\mathbb{R}^d$, *где* $d > 6$ *и* $d \neq 4m^2+4m-2$, $m \in \mathbb{N}$, *равна* $d(d+1)/2$.

## Список литературы


[1] Акопян А. В., Кабатянский Г. А., Мусин О. Р. *Контактные числа, коды и сферические многочлены* // Математическое просвещение. Третья Серия. Вып. 16. 2012. С. 57–74.

[2] Ионин Ю. И. *Строго равнобедренные множества* // Математическое просвещение. Третья Серия. Вып. 15. 2011. С. 154–175.

[3] Bannai E., Bannai E., Stanton D. *An upper bound for the cardinality of an s-distance subset in real Euclidean space* II // Combinatorica. Vol. 3(2). 1983. P. 147–152.

[4] Barg A., Yu W.-H. *New bounds for spherical two-distance sets*. ArXiv: 1204.5268. 2012.

[5] Blokhuis A. *Few-Distance Sets*. CWI Tracts, vol. 7. Amsterdam: CWI. 1984.

[6] Blokhuis A. *A new upper bound for the cardinality of 2-distance sets in Euclidean space* // North-Holland Mathematics Studies. Vol. 87. 1984. P. 65–66.

[7] Croft H. T. 9-*point and* 7-*point configurations in* 3-*space* // Proc. London Math. Soc. Ser. 3. Vol. 12. 1962. P. 400–424.

[8] Delsarte P., Goethals J. M., Seidel J. J. *Spherical codes and designs* // Geometriae Dedicata. Vol. 6 (3). 1977. P. 363–388.

[9] Einhorn S. J., Schoenberg I. J. *On Euclidean sets having only two distances between points*. I // Nederl. Akad. Wetensch. Proc. Ser. A 69=Indag. Math. Vol. 28. 1966. P. 479–488.

[10] Einhorn S. J., Schoenberg I. J. *On Euclidean sets having only two distances between points*. II // Nederl. Akad. Wetensch. Proc. Ser. A 69=Indag. Math. Vol. 28. 1966. P. 489–504.

[11] Ionin Y. J. *Isosceles sets* // The Electronic Journal of Combinatorics. Vol. 16(R141):1. 2009.

[12] Kelly L. M. *Elementary Problems and Solutions. Isosceles n-points. Solutions: E735* // Amer. Math. Monthly. Vol. 54 (4). 1947. P. 227–229.

[13] Kido H. *Classification of isosceles eight-point sets in three-dimensional Euclidean space* // European Journal of Combinatorics. Vol. 27 (3). 2006. P. 329–341.

[14] Larman D. G., Rogers C. A., Seidel J. J. *On two-distance sets in Euclidean space* // Bulletin of the London Mathematical Society. Vol. 9(3). 1977. P. 261–267.

[15] Lisoněk P. *New maximal two-distance sets* // Journal of Combinatorial Theory. Series A. Vol. 77 (2). 1997. P. 318–338.

[16] Martini H., Swanepoel K. J. *The geometry of Minkowski spaces—a survey. Part* II // Expositiones mathematicae. Vol. 22(2). 2004. P. 93–144.

[17] Martini H., Swanepoel K. J., G. Weiß. *The geometry of Minkowski spaces—a survey. Part* I // Expositiones mathematicae. Vol. 19(2). 2001. P. 97–142.

[18] Musin O. R. *Spherical two-distance sets* // Journal of Combinatorial Theory. Series A. Vol. 116(4). 2009. P. 988–995.

[19] Nozaki H. *A generalization of Larman–Rogers–Seidel's theorem* // Discrete Mathematics. Vol. 311(10). 2011. P. 792–799.

[20] Shinohara M. *Uniqueness of maximum three-distance sets in the three-dimensional Euclidean space*. Preprint.

[21] Shinohara M. *Classification of three-distance sets in two dimensional euclidean space* // European Journal of Combinatorics. Vol. 25(7). 2004. P. 1039–1058.

А. В. Акопян, Институт проблем передачи информации им. А. А. Харкевича РАН и Ярославский государственный университет им. П. Г. Демидова
О. Р. Мусин, Институт проблем передачи информации им. А. А. Харкевича РАН и Ярославский государственный университет им. П. Г. Демидова

# Опять о многоугольниках Рейнхардта

С. Б. Гашков

В [9] Рейнхардт доказал изодиаметрическое неравенство для периметра $n$-угольника данного диаметра $d$. В [1] автором настоящей заметки было приведено более простое доказательство. Попутно было доказано аналогичное неравенство для ширины[1)]. А именно, доказана следующая теорема.

*Если $p$ периметр, $b$ ширина и $d$ диаметр выпуклого $n$-угольника, то справедливы неравенства* $2nb\operatorname{tg}\frac{\pi}{2n} \leqslant p \leqslant 2nd\sin\frac{\pi}{2n}$. *Каждое из неравенств точное, если* $n \neq 2^k$, $k \in \mathbb{N}$. *Достигаются они, в частности, на правильных $n$-угольниках при нечетном $n$, и на полуправильных $n$-угольниках при четном $n$, отличных от степени двойки*[2)].

Вопрос об экстремальных $n$-угольниках[3)] оказался непростым. Он был явно сформулирован автором в тексте, помещенном вслед за статьей (под названием «задачи для исследования»). В нем утверждалось, в частности, что для простого $p$ единственным экстремальным $p$-угольником является правильный, единственным экстремальным $2p$-угольником является полуправильный, среди девятиугольников имеется в точности два экстремальных (один правильный, другой полуправильный), а для всех остальных $n \neq 2^k$ существуют экстремальные $n$-угольники, не являющиеся ни правильными, ни полуправильными. Там же было приведено следующее неравенство Рейнхардта для площади $n$-угольника

$$s \leqslant \frac{n}{2}\cos\frac{\pi}{n}\operatorname{tg}\frac{\pi}{2n}d^2,$$

в котором равенство достигается только для правильных нечетноугольников.

Примерно в то же время были получены верхние и нижние оценки для числа различных экстремальных $n$-угольников, но опубликовать их в журнале «Квант» не предоставлялось возможным, а публикации в научных

[1)] Соответствующие определения можно найти в [1–3].

[2)] Полуправильными в [1] названы равносторонние $n$-угольники, вписанные в правильные $k$-угольники Рело, где $k$ — делитель $n$, но фактически там дано более явное определение.

[3)] Для которых обращается в равенство любое из двух неравенств, так как они обращаются в равенство одновременно.

журналах препятствовало отсутствие уверенности в новизне результатов, в частности изодиаметрического неравенства для ширины $2nb\,\mathrm{tg}\,\frac{\pi}{2n} \leqslant p$ (в [1] оно вместе с неравенством для периметра было названо неравенством Рейнхардта).

После возобновления ежегодника «Математическое просвещение» появилась возможность публикации в нем, что и было осуществлено[4)] в [2]. Там было кратко повторено доказательство из [1] (по не вполне понятной сейчас самому автору причине в [2] не было ссылки на [1]), доказаны утверждения, оставленные в [1] без доказательства, а также некоторые новые утверждения. Экстремальные $n$-угольники в [2] были названы многоугольниками Рейнхардта, им были сопоставлены двухцветные ожерелья и подмножества мощности $n$ в группе комплексных корней степени $2n$ из единицы, не содержащие противоположных корней, и с помощью этих элементарных соображений получены верхние и нижние оценки для числа $n$-угольников Рейнхардта с точностью до вращений[5)]. А именно, для него получены оценки

$$\frac{1}{2n}\sum_{2\nmid m\mid n} 2^{n/m}(\varphi(m)-\mu(m)) \leqslant R(n) < \frac{1}{2n}\sum_{2\nmid m\mid n} 2^{n/m}\varphi(m),$$

где $\mu(n)$ — функция Мёбиуса, $\varphi(n)$ — функция Эйлера; нижняя оценка — это в точности число «периодических» $n$-угольников[6)]. Для $n = 2^l p^k$, $k, l \in \mathbb{N}$, где $p > 2$ — простое число, в [2] было доказано с помощью круговых многочленов, что все $n$-угольники Рейнхардта периодические, откуда следует формула

$$\begin{aligned} R(n) &= \frac{1}{2n}\sum_{2\nmid m\mid n} 2^{n/m}(\varphi(m)-\mu(m)) = \\ &= \frac{1}{2n}(2^{n/p}p + 2^{n/p^2}p(p-1) + \ldots + 2^{n/p^k}p^{k-1}(p-1)) \sim 2^{n/p}p/2n. \end{aligned}$$

Недавно автор случайно увидел в интернете статью [7], где, в частности, были найдены формулы для числа периодических $n$-угольников

[4)] После того, как автор нашел свои старые записи и сделал из них TEX-файл, а также перевод на немецкий и послал в «Elemente der Mathematik», и получил оттуда отказ под предлогом, что подобного стиля статьи журнал не принимает.

[5)] Пользуясь возможностью, заметим здесь, что в [2] в двух формулах посреди с. 98 есть опечатки — в одной в последнем равенстве пропущен множитель $m/2n$, а другой между двумя суммами вставлен ненужный знак равенства.

[6)] Т. е. тех, которым соответствуют периодические ожерелья; в частности, полуправильные многоугольники являются периодическими.

Рейнхардта[7] с точностью до вращений и осевых симметрий[8], повторен результат из [2] о том, что при $n = 2^l p^k$, $k, l \in \mathbb{N}$ не бывает непериодических $n$-угольников Рейнхардта, с помощью компьютерного поиска найдены для некоторых $n$ непериодические (названные в [7] спорадическими) $n$-угольники, и было выдвинуто предположение о том, что при $n = pq$, где $p > q$ — простые, спорадических $n$-угольников не существует. Из [7] автор узнал также, что в [4] доказано неравенство для ширины $2nb \operatorname{tg} \frac{\pi}{2n} \leqslant p$ без ссылок на [1], где это было сделано намного ранее (и вероятно, проще). Значит, по-видимому, этот результат [1] являлся новым.

В [6], наряду с другими результатами, была доказана

Теорема 1. *Непериодических $n$-угольников Рейнхардта при $n$, равном произведению двух различных простых, не существует.*

Ниже мы приведем более простое доказательство этого факта. Полагаем, что и остальные результаты [6] могут быть доказаны более просто с использованием указанного ниже подхода.

Доказательство [6] основано на теореме Рейнхардта [9] о том, что экстремальным $n$-угольникам можно однозначно сопоставить *многочлены Рейнхардта $f(x)$ такие, что $f(0) = 1$, $\deg f < n$, $\Phi_{2n} \mid f$, где $\Phi_{2n}$ — круговой полином*[9] *порядка $2n$, а коэффициенты $f(x)$ равны $0, \pm 1$, причем число ненулевых коэффициентов нечетно.* Утверждение теоремы, т. е. отсутствие спорадических $n$-угольников, очевидно следует из следующей леммы [6]:

Лемма 1. *Если $n = pq$, где $p > q > 2$ — простые, то многочлен Рейнхардта $f(x)$ делится или на $\Phi_p(-x^q)$ или на $\Phi_q(-x^p)$.*

Действительно, так как $\Phi_p(-x^q) = x^{(p-1)q} - x^{(p-2)q} + \ldots \pm x^q \mp 1$, указанная делимость означает антипериодичность последовательности коэффициентов многочлена $f(x)$, т. е. выполнения равенств $f_i = -f_{i+p}$, где $0 \leqslant i < q(p-1)$, а значит и $q$-периодичность соответствующего $n$-угольника[10] (это не очевидно, так как здесь мы не приводим правила построения многочленов Рейнхардта, однако несложно доказывается в [6, 7], куда мы и отсылаем читателя).

---

[7] Термин $n$-угольники Рейнхардта, кажется, не использовался в [7] и появился только в [6].

[8] Почему-то мне не пришло в голову рассмотреть такую задачу.

[9] Определение круговых полиномов, или полиномов деления круга, можно найти в любом учебнике высшей алгебры.

[10] Многоугольник называется $t$-периодическим, если его группа вращений имеет порядок $t$.

Для доказательства этой леммы в [6] используется следующая лемма, фактически содержащаяся в [5, 8, 10][11)].

Лемма 2. *Если $g(x)$ — произвольный многочлен степени, меньшей $n = pq$, с целыми коэффициентами, делящийся на $\Phi_n$, то*

$$g(x) = a(x)\Phi_p(x^q)+b(x)\Phi_q(x^p),\ a(x), b(x) \in \mathbb{Z}[x],\ \deg a(x) < q,\ \deg b(x) < p.$$

Так как при нечетном $n$, как известно из алгебры, $\Phi_{2n}(x) = \Phi_n(-x)$, то из этой леммы (при подстановке $x = -y$) следует, что если $\Phi_{2n} \mid g(x)$, то $g(x) = a(x)\Phi_p(-x^q) + b(x)\Phi_q(-x^p)$, $\deg a(x) < q$, $\deg b(x) < p$.

В [1] экстремальным $n$-угольникам Рейнхардта другим способом однозначно сопоставлены многочлены $g(x)$, $\deg g < n$, у которых *все коэффициенты равны* $\pm 1$, такие, что $\Phi_{2n}$ делит $g$, причем, если коэффициенты $g_i$ многочлена удовлетворяют условию антипериодичности $g_i = -g_{i+t}$, где $n/t$ нечетно, то соответствующий $n$-угольник Рейнхардта будет $n/t$-периодическим.

Это позволяет упростить рассуждения [6] и вывести отсутствие спорадических $n$-угольников при $n = pq$, где $p > q > 2$ — простые, из следующей леммы.

Лемма 3. *Если сумма двух последовательностей, одна из которых антипериодична с периодом $p$, а другая антипериодична с периодом $q$, $(p, q) = 1$, есть последовательность, членами которой являются только два различных числа $a, b$, то эта последовательность тоже антипериодична либо с периодом $p$, либо с периодом $q$.*

Действительно, пусть $g(x)$, $\deg g < n$, многочлен с коэффициентами $\pm 1$, такой, что $\Phi_{2n} \mid g$, соответствующий произвольному $n$-угольнику Рейнхардта, $n = pq$, $p > q > 2$ — простые. Согласно лемме 2 имеем $g(x) = a(x)\Phi_p(-x^q) + b(x)\Phi_q(-x^p)$, $a(x), b(x) \in \mathbb{Z}[x]$, $\deg a(x) < q$, $\deg b(x) < p$. Так как последовательности коэффициентов многочленов $a(x)\Phi_p(-x^q)$, $b(x)\Phi_q(-x^p)$ антипериодичны с периодами $q, p$ соответственно (и обратно, любой многочлен степени $pq-1$ с антипериодичной с периодом $q$ последовательностью коэффициентов представим в виде $a(x)\Phi_p(-x^q)$, $\deg a(x) < q$), а последовательность коэффициентов многочлена $g(x)$ состоит только из двух чисел $\pm 1$, то из леммы 3 следует, что последовательность коэффициентов многочлена $g(x)$ антипериодична с периодом $p$ или $q$, значит соответствующий ему $n$-угольник периодичен (т. е. имеет группу вращений порядка $q$ или $p$).

Приведем доказательство леммы 3. Докажем вначале, что одна из двух суммируемых антипериодических последовательностей $\{a_i\}$, $\{b_i\}$ такова,

[11)] В [5, 8, 10] доказано более общее утверждение, но ограничений на степени многочленов там нет. В [8] дано более простое доказательство, чем в [5] и есть ссылка на [5]. В [10] ссылок на [5, 8] нет.

что все ее члены с четными номерами равны. Допустим, что это неверно, тогда найдутся четные числа $i_0$, $i_1$, $j_0$, $j_1$, такие, что $a_{i_0} < a_{i_1}$, $b_{j_0} < b_{j_1}$, $i_0, i_1 < 2q$, $j_0, j_1 < 2p$. Согласно китайской теореме об остатках для любых $\alpha, \beta \in \{0, 1\}$ найдется число $N_{\alpha,\beta} < 2pq$, такое что $N_{\alpha,\beta} = i_\alpha \bmod 2q$, $N_{\alpha,\beta} = j_\beta \bmod 2p$ (китайская теорема применяется к числам $i_\alpha/2$, $j_\beta/2$ и модулям $q$, $p$, а потом все умножается на 2). Тогда при $N = N_{\alpha,\beta}$ имеем в силу периодичности $c_N = a_N + b_N = a_{i_\alpha} + b_{j_\beta}$, значит $c_{N_{0,0}} = a_{i_0} + b_{j_0} < a_{i_0} + b_{j_1} = c_{N_{0,1}} < a_{i_1} + b_{j_1} = c_{N_{1,1}}$, т. е. сумма последовательностей состоит не из двух, а хотя бы из трех разных чисел, что противоречит условию. Поэтому можно считать, что, например, в последовательности $\{a_i\}$ все члены $a_{2i}$, $i = 0, 1, 2, \ldots$, равны некоторому числу $a$. В силу $q$-антипериодичности $a_{2i+1} = -a_{q+2i+1} = -a$, $i = 0, 1, 2, \ldots$, поэтому $\{a_i\} = a, -a, a, -a, \ldots$, т. е. $\{a_i\}$ на самом деле антипериодична с периодом 1, а потому и антипериодична с любым нечетным периодом, например $p$. Но тогда антипериодична с тем же периодом $p$ и сумма $\{c_i = a_i + b_i\}$, что доказывает лемму.

Замечание. Очевидно, таким же способом доказывается и следующая лемма (которая далее не понадобится).

*Если сумма двух последовательностей, одна из которых периодична с периодом $p$, а другая периодична с периодом $q$, $(p, q) = 1$, есть последовательность, членами которой являются только два различных числа $a, b$, то эта последовательность тоже периодична либо с периодом $p$, либо с периодом $q$.*

Для полноты приведем доказательство леммы 2 из [6]. Так как $\Phi_p, \Phi_q$ при простых $p > q$ взаимно просты, то из известного свойства алгоритма Евклида следует существование многочленов $a(x), b(x) \in \mathbb{Z}[x]$, $\deg a(x) < q$, $\deg b(x) < p$, таких, что $a(x)\Phi_p(x) + b(x)\Phi_q(x) = 1$. Если $g(x)$ — многочлен делящийся на $\Phi_n$, то $g(x) = \Phi_n(x)h(x)$, откуда

$$g(x) = h(x)a(x)\Phi_n(x)\Phi_p(x) + h(x)b(x)\Phi_n(x)\Phi_q(x).$$

Деля $h(x)a(x)$ на $\Phi_q(x)$, имеем $h(x)a(x) = c(x)\Phi_q(x) + f_1(x)$, $\deg f_1(x) < q$, и предыдущее равенство перепишем в виде

$$\begin{aligned} g(x) &= f_1(x)\Phi_n(x)\Phi_p(x) + (c(x)\Phi_p(x) + h(x)b(x))\Phi_n(x)\Phi_q(x) = \\ &= f_1(x)\Phi_n(x)\Phi_p(x) + f_2(x)\Phi_n(x)\Phi_q(x), \\ f_2(x) &= c(x)\Phi_p(x) + h(x)b(x). \end{aligned}$$

Так как $\Phi_{pq}(x) = \dfrac{(x^{pq}-1)(x-1)}{(x^p-1)(x^q-1)}$, то $\Phi_n(x)\Phi_p(x) = \dfrac{x^{pq}-1}{x^q-1} = \Phi_p(x^q)$, значит $\Phi_n(x)\Phi_p(x) = \Phi_p(x^q)$, $\Phi_n(x)\Phi_q(x) = \Phi_q(x^p)$, поэтому предыдущее равенство можно переписать в виде $g(x) = f_1(x)\Phi_p(x^q) + f_2(x)\Phi_q(x^p)$. Так как $\deg f_1 < q$, $\deg g < n$, то $\deg f_1(x)\Phi_p(x^q) < q + (p-1)q = pq = n$, значит

$\deg(f_2(x)\Phi_q(x^p)) = \deg(g(x) - f_1(x)\Phi_p(x^q)) < n$, поэтому $\deg f_2(x) < n - (q-1)p = p$, что и доказывает лемму.

Также для полноты приведем кратко доказательство из [1, 2] неравенства Рейнхардта и условий его обращения в равенство. Пусть $M^*$ — симметризация Минковского $n$-угольника $M$, т. е.

$$M^* = \frac{1}{2}(M + (-M)) = \{(x-y)/2 : x, y \in M\}.$$

Тогда множество $M^*$ есть центрально-симметричный выпуклый $2m$-угольник, где $m \leqslant n$, с теми же периметром, диаметром и шириной что и у $M$. Из центральной симметричности $M^*$ следует, что его внутренний радиус $r^* = b/2$ и внешний радиус $R^* = d/2$ (доказательства всех используемых фактов можно найти в [1,3]). Последовательность $n \operatorname{tg} \frac{\pi}{n}$ строго монотонно убывает, а $n \sin \frac{\pi}{n}$ строго монотонно возрастает. Поэтому

$$2nb \operatorname{tg} \frac{\pi}{n} \leqslant$$
$$\leqslant 2mb \operatorname{tg} \frac{\pi}{2m} = 4mr^* \operatorname{tg} \frac{\pi}{2m} \leqslant p \leqslant 4mR^* \sin \frac{\pi}{2m} = 2md \sin \frac{\pi}{2m} \leqslant$$
$$\leqslant 2nd \sin \frac{\pi}{2n},$$

причем равенства $p = 2nb \operatorname{tg} \frac{\pi}{2n}$, $p = 2nd \sin \frac{\pi}{2n}$ лишь тогда справедливы, когда $M^*$ есть правильный $2n$-угольник.

Для правильного $n$-угольника при нечетном $n$ справедливы равенства $p = 2nb \operatorname{tg} \frac{\pi}{2n} = 2nd \sin \frac{\pi}{2n}$, а при четном $n$ ни одно из них не верно. Однако при четном $n$, не равном степени двойки, указанные равенства могут достигаться. Примеры $n$-угольников, для которых они достигаются, можно построить следующим образом[12)]. Представим $n$ произвольным образом в виде $m(2k+1)$. Возьмем правильный $2k+1$-угольник с наибольшей диагональю $d$ и с центром в каждой его вершине построим круг радиуса $d$. Выпуклая фигура, являющаяся пересечением построенных кругов — это правильный $2k+1$-угольник Рело. Ее диаметр равен $d$, а периметр $\pi d$. Ширина ее во всех направлениях одинакова и равна $d$ (это фигура постоянной ширины). Впишем в нее $n$-угольник с равными сторонами так, чтобы среди его вершин содержались все вершины рассматриваемого $2k+1$-угольника. Каждая из $2k+1$ упомянутых дуг при этом будет разбиваться на $m$ равных дужек, хорды которых будут сторонами рассматриваемого $n$-угольника. Диаметр его $d$, а ширина равна $b = d\cos(\pi/2n)$ — высоте равнобедренного треугольника с ребром $d$ и углом при вершине $\pi/n$. Тогда его периметр равен $2nd \sin(\pi/2n) = 2b \operatorname{tg}(\pi/2n)$.

[12)] В [1] они назывались полуправильными $n$-угольниками

Укажем теперь вытекающее из приведенного доказательства полное описание $n$-угольников Рейнхардта (взятое из [2]). Пусть $M$ такой $n$-угольник, тогда $M^*$ — правильный $2n$-угольник (в противном случае равенство не достигается). Ориентируем все стороны многоугольника $M^*$ в одном направлении и сопоставим им вектора, изображающие комплексные корни $2n$-й степени из единицы $1, \varepsilon, \varepsilon^2, \dots \varepsilon^{2n-1}$, где $\varepsilon = \exp\frac{i\pi}{n}$. Тогда сторонам $n$-угольника $M$ однозначно сопоставляется $n$-элементное множество, также обозначаемое $M$, лежащее в группе корней из единицы $W_{2n} = \{1, \varepsilon, \dots, \varepsilon^{2n-1}\}$. При подходящем выборе обозначений, можно считать $1 \in M$. Если расположить вектора этого множества в определенном порядке так, чтобы начало очередного вектора вектора совпало с концом предыдущего, то они будут ограничивать $n$-угольник, подобный $n$-угольнику $M$. Естественно, центрально-симметричному многоугольнику $(-M)$ соответствует множество $(-M)$ такое, что $M \cap (-M) = \varnothing$. Ясно, что условия $M \cap (-M) = \varnothing$ и $M \cup (-M) = W_{2n}$ равносильны. Они равносильны также тому, что в каждой паре противоположных векторов $\{\varepsilon^i, \varepsilon^{n+i} = -\varepsilon^i\}$ один принадлежит $M$, а другой принадлежит $-M$. Также очевидно, что $\sum\limits_{\varepsilon^i \in M} \varepsilon^i = 0$, так как сумма векторов, в которые превращаются ориентированные в одном направлении стороны многоугольника, равна нулю. Расположение векторов в любом другом порядке, хотя и имеет всегда сумму равную нулю, приводит к невыпуклому или самопересекающемуся многоугольнику. Рассмотрим многочлен $f_M(x) = \sum\limits_{\varepsilon^i \in M} x^i$, степени, меньшей $2n$, состоящий из $n-1$ одночленов и единичного свободного члена. Его корнем будет $x = \epsilon$. Заменим в этом многочлене каждый член $x^{n+i}$ на $-x^i$, $0 \leqslant i < n$. Полученный многочлен обозначим $g_M(x)$. Он состоит из $n-1$ различных одночленов с коэффициентами $\pm 1$ (в силу условия $M \cap (-M) = \varnothing$) и единичного свободного члена, имеет степень меньше $n$, значит он не содержит нулевых коэффициентов. Его корнем также будет $x = \epsilon$. Согласно известным свойствам круговых многочленов (см., например, [5, 8, 10] и любой курс алгебры) последнее условие равносильно делимости $g_M(x)$ на круговой многочлен $\Phi_{2n}(x)$. По любому многочлену $g_M(x)$, удовлетворяющему указанным условиям, однозначно восстанавливается многочлен $f_M(x)$, множество $M$, соответствующее ему двухцветное ожерелье (см. [2]) и сам $n$-угольник Рейнхардта $M$. Поэтому задача описания всех таких $n$-угольников является чисто алгебраической (что и было показано в [2]).

Легко видеть, что условие антипериодичности коэффициентов многочлена $g_M$ равносильно условию периодичности коэффициентов многочлена $f_M$, множества $M$ и соответствующего ему двухцветного ожерелья. Из условия антипериодичности с периодом $t$ (тогда обязательно $n/t$ нечетно)

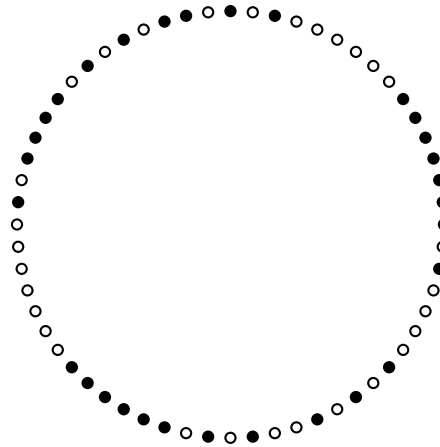


*Рис. 1. Непериодическое ожерелье для 30-угольника Рейнхардта*

коэффициентов $g_M(x)$ следует равенство $g_M(x) = h(x)(1 - x^t + x^{2t} - \ldots - x^{n-t})$, $\deg h < t$, из которого очевидно следует равенство $g_M(\epsilon) = 0$ (а значит, и условие $\Phi_{2n} \mid g_M$).[13)]

В качестве примера различия между многочленами Рейнхардта, использованными в [6, 7], и нашими многочленами, рассмотрим спорадический 30-угольник, найденный в [7] (там было доказано, что при $n < 30$ спорадических $n$-угольников не существует). Опишем его вначале, используя введенные выше множества, ожерелья и многочлены. Возьмем при $\varepsilon = e^{i\pi/30}$, $\varepsilon^{60} = 1$ непериодическое множество

$$\{1, \varepsilon, \varepsilon^2, \varepsilon^3, \varepsilon^4, \varepsilon^5, \varepsilon^6, -\varepsilon^7, -\varepsilon^8, -\varepsilon^9, -\varepsilon^{10}, -\varepsilon^{11}, -\varepsilon^{12}, \varepsilon^{13}, -\varepsilon^{14}, \varepsilon^{15}, -\varepsilon^{16}, \varepsilon^{17}, \varepsilon^{18}, -\varepsilon^{19}, \varepsilon^{20}, -\varepsilon^{21}, \varepsilon^{22}, -\varepsilon^{23}, \varepsilon^{24}, \varepsilon^{25}, \varepsilon^{26}, \varepsilon^{27}, -\varepsilon^{28}, \varepsilon^{29}\}$$

(в этой записи использовалось тождество $\varepsilon^i = -\varepsilon^{30+i}$, $i = 0, \ldots, 29$). Соответствующее ему ожерелье см. на рис. 1.

Проверим, что соответствующий многочлен $g(x)$ имеет корень $\varepsilon$. Для этого нужно установить, что

$$1 + \varepsilon + \varepsilon^2 + \varepsilon^3 + \varepsilon^4 + \varepsilon^5 + \varepsilon^6 - \varepsilon^7 - \varepsilon^8 - \varepsilon^9 - \varepsilon^{10} - \varepsilon^{11} - \varepsilon^{12} + \varepsilon^{13} - \varepsilon^{14} + \varepsilon^{15} - \varepsilon^{16} + \varepsilon^{17} + \varepsilon^{18} - \varepsilon^{19} + \varepsilon^{20} - \varepsilon^{21} + \varepsilon^{22} - \varepsilon^{23} + \varepsilon^{24} + \varepsilon^{25} + \varepsilon^{26} + \varepsilon^{27} - \varepsilon^{28} + \varepsilon^{29} = 0.$$

Непосредственно это сделать затруднительно, но можно непосредственно проверить, что

[13)] Геометрическое доказательство см. в [2].

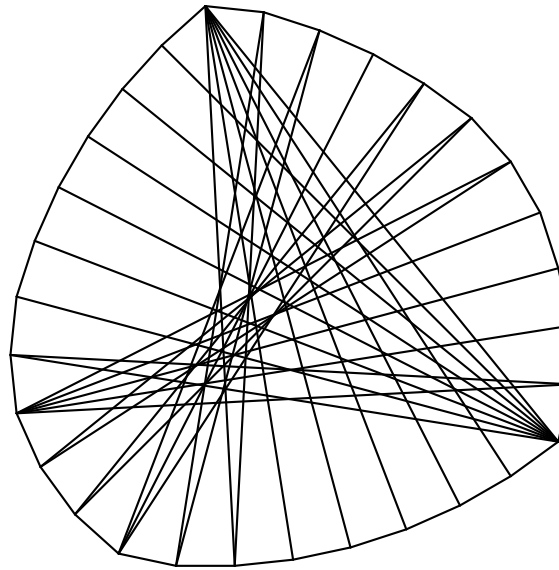


*Рис. 2. Непериодический* 30*-угольник Рейнхардта, соответствующий ожерелью рис.* 1

$$(x^{13}-x^{12}+2x^{10}+x^9-x^8-x^7+x^6+x^5+x^4+x+1)\cdot$$
$$\cdot(1+x^2-x^6-x^8-x^{10}+x^{14}+x^{16})=1+x+x^2+x^3+x^4+x^5+x^6-$$
$$-x^7-x^8-x^9-x^{10}-x^{11}-x^{12}+x^{13}-x^{14}+x^{15}-x^{16}+x^{17}+x^{18}-$$
$$-x^{19}+x^{20}-x^{21}+x^{22}-x^{23}+x^{24}+x^{25}+x^{26}+x^{27}-x^{28}+x^{29},$$

и заметить, что $\Phi_{60}(\varepsilon)=-1+\varepsilon^2-\varepsilon^6-\varepsilon^8-\varepsilon^{10}+\varepsilon^{14}+\varepsilon^{16}=0$, так как

$$0=(\varepsilon^{30}+1)(\varepsilon^2+1)=(\varepsilon^6+1)(\varepsilon^{10}+1)(1+\varepsilon^2-\varepsilon^6-\varepsilon^8-\varepsilon^{10}+\varepsilon^{14}+\varepsilon^{16}),$$
$$\varepsilon^6+1\neq 0,\quad \varepsilon^{10}+1\neq 0.$$

Непериодический 30-угольник Рейнхардта, соответствующий ожерелью рис. 1, изображен на рис. 2. Кроме сторон, на нем показаны также его диаметры.

В [7] для построения этого же 30-угольника использовался многочлен Рейнхардта

$$r(x)=$$
$$=1-x^7+x^{13}-x^{14}+x^{15}-x^{16}+x^{17}-x^{19}+x^{20}-x^{21}+x^{22}-x^{23}+x^{24}-x^{28}+x^{29},$$

найденный компьютерным поиском. Тот факт, что он делится на $\Phi_{60}$ (и

имеет корнем $\varepsilon$) следует из тождества

$$(x^{13}-x^{12}-x^{11}+x^{10}+x^9-x^7+x^4-x^2+1)(x^{16}+x^{14}-x^{10}-x^8-x^6+x^2+1) = r(x).$$

Многочлен $r(x)$ факторизован программой Maple, но проверку уже найденного тождества можно при желании выполнить и вручную.

Работа выполнена при финансовой поддержке РФФИ, проект 11–01–00508а.

## Список литературы

[1] Гашков С. Б. *Неравенства для площади и периметра выпуклого многоугольника* // Квант, №10. 1985. С. 15–19.

[2] Гашков С. Б. *Неравенства для выпуклых многоугольников и многоугольники Рейнхардта* // Математическое просвещение. Сер. 3. Вып. 11. 2007. С. 91–103.

[3] Яглом И. М., Болтянский В. Г. *Выпуклые фигуры*. М.: ГИТТЛ. 1951.

[4] Audet C., Hansen P., Messine F. *Isoperimetric polygons of maximum width* // Discrete Comput. Geom. Vol. 41, no 1. 2009. P. 45–60.

[5] de Bruijn N. G. *On the factorization of cyclic groups* // Nederl. Akad. Wetensch. Proc. Ser. A. Vol. 15. 1953. P. 370–377.

[6] Hare K. G., Mossinghoff M. J. *Sporadic Reinhardt polygons*. arXiv:1203.4107v2. 2012.

[7] Mossinghoff M. J. *Enumerating isodiametric and isoperimetric polygons* // Journal of Combinatorial theory. Ser. A. Vol. 118, no 6. 2011. P. 1801–1815.

[8] Rédei L. *Über das Kreisteilungspolynom* // Acta Math. Acad. Sci. Hungar. Vol. 5. 1954. P. 27–28. MR 0062760 (16,13h).

[9] Reinhardt K. *Extremal Polygone gegebenen Durchmessers* // Jber. Deutsch. Math.-Vereinig. Bd. 31. 1922. S. 251–270.

[10] Schoenberg I. J. *A note on the cyclotomic polynomial* // Mathematika. Vol. 11. 1964. P. 131–136.

С. Б. Гашков, механико-математический факультет МГУ им. М. В. Ломоносова
Email: sbgashkov@gmail.com

# Дискретный анализ для математиков и программистов (подборка задач)

Д. Ильинский А. Купавский А. Райгородский*
А. Скопенков†

## Введение

Мы приводим подборки задач по комбинаторике и теории графов (в том числе случайных). Эти задачи будут полезны руководителям и участникам кружков для старшеклассников и младшекурсников, ориентированных на олимпиады. Некоторые приводимые красивые задачи и важные темы малоизвестны в парадигме кружков по математике, но полезны как для математического образования, так и для подготовки к олимпиадам.

Решение этих задач будет полезно также всем, кто хочет стать математиком, специалистом по computer science или программистом-разработчиком. Именно таких специалистов мы готовим на факультете инноваций и высоких технологий Московского Физико-Технического Института. Приведённые задачи используются при изучении курса дискретного анализа на этом факультете, который читает с 2009 проф. А. М. Райгородский. Мы благодарим студентов за каверзные вопросы и указания на неточности.

Формулировки большинства задач доступны старшеклассникам, интересующимся математикой; мы приводим все определения, не так часто изучаемые на кружках. Однако многие задачи трудны, для их решения

*Поддержан грантом РФФИ 12-01-00683 и грантом поддержки ведущих научных школ НШ-2519.2012.1.

†Частично поддержан грантом фонда Саймонса.

нужно предварительно прорешать другие приведённые задачи на данную тему или знать лекционный материал.

## Общие соглашения и определения

Если условие задачи является утверждением, то в задаче требуется это утверждение доказать. Если некоторая задача не получается, то читайте дальше — соседние задачи могут оказаться подсказками.

Если не оговорено противное, асимптотики и пределы рассматриваются при $n \to \infty$.

*Графом без петель и кратных рёбер* $G = (V, E)$ называется конечное множество $V = V(G)$, некоторые двухэлементные подмножества (т. е. неупорядоченные пары несовпадающих элементов) которого выделены. Граф без петель и кратных рёбер мы коротко называем графом. Элементы данного множества $V$ называются *вершинами*. Выделенные пары вершин называются *рёбрами*. Если вершина принадлежит ребру, то говорится, что ребро называется *проходящим* через эту вершину или *инцидентным* этой вершине. Множество выделенных подмножеств обозначается $E = E(G)$. Если не оговорено противное, то через $n$ и $e$ обозначаются соответственно количества вершин и рёбер рассматриваемого графа.

Приведённое определение не годится для *графов с петлями и кратными рёбрами* (мультиграфов). С ними мы почти не работаем и определяем там, где они используются.

Граф можно представлять себе как набор точек (например, на плоскости), некоторые пары которых соединены ломаными, причём одна пара вершин не может быть соединена более чем одной линией. Точки называются *вершинами* графа, а ломаные — *рёбрами*. Ломаные могут пересекаться, но точки пересечения «не считаются», то есть не являются вершинами.

*Степенью* вершины графа называется число выходящих из неё рёбер.

*Путём* в графе называется последовательность вершин, в которой любые две соседние вершины соединены ребром.

Граф называется *связным*, если любые две его вершины можно соединить путём.

Ясно, что «соединённость некоторым путём» является отношением эквивалентности на множестве вершин графа. *Связной компонентой* графа называется класс эквивалентности.

Граф, у которого проведены все возможные ребра между вершинами называется *полным* и обозначается $K_n$. Если вершины графа можно разделить на две части так, что не существует рёбер, соединяющих вершины из одной и той же части, то граф называется *двудольным*, а части называются *долями*. Через $K_{m,n}$ обозначается двудольный граф с долями из

$m$ и из $n$ вершин, в котором имеются все ребра между вершинами разных долей.

*Изолированной вершиной* называется вершина, из которой не выходит ни одного ребра.

Раскраска вершин графа в несколько цветов называется *правильной*, если концы любого ребра разноцветны.

## 1. Комбинаторика и биномиальные коэффициенты

1.1.

(a) $$\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}.$$

(b) $$\left\{{n+1 \atop k+1}\right\} = (k+1)\left\{{n \atop k+1}\right\} + \left\{{n \atop k}\right\}.$$

(c) $$\left|{n+1 \atop k+1}\right| = \left|{n \atop k+1}\right| + 2^{n-k}\left|{n \atop k}\right|.$$

Здесь $\binom{n}{k}$, $\left\{{n \atop k}\right\}$, $\left|{n \atop k}\right|$ — соответственно количества

- $k$-элементных подмножеств $n$-элементного множества (другое обозначение: $C_n^k$);
- разбиений $n$-элементного множества на $k$ частей (т. е. непустых подмножеств), разбиения считаются неупорядоченными, т. е. $\{1,2\} \cup \{3\} = \{3\} \cup \{1,2\}$;
- $k$-мерных линейных подпространств линейного пространства $\mathbb{Z}_2^n$.

1.2. Найдите «явную» формулу для

$$\text{(a)} \sum_{k\geqslant 0}\binom{n}{2k}; \quad \text{(b)} \sum_{k\geqslant 0}\binom{n}{4k}; \quad \text{(c)} \sum_{k\geqslant 0}\binom{n}{3k}.$$

В ответе используйте только целочисленные функции целочисленного аргумента.

Будут полезны *тригонометрическая форма комплексного числа* $a + bi = r(\cos\varphi + i\sin\varphi)$, где $r = \sqrt{a^2+b^2}$ и $\varphi = \operatorname{arctg}(b/a)$ или $\varphi = \operatorname{arctg}(b/a) + \pi$, *формула Муавра* $(\cos\varphi + i\sin\varphi)^n = \cos n\varphi + i\sin n\varphi$.

1.3. (a) Во скольких подмножествах множества $\{1, 2, 3, \ldots, 11\}$ не найдётся двух подряд идущих чисел?

(b) То же для *трёх* подряд идущих чисел.

1.4. (a) При фиксированном $n$ число $\binom{n}{k}$ максимально при $k = [n/2]$.

(b) *Best in their own ways*. В математической олимпиаде участвовало $k$ школьников. Выяснилось, что для любых двух школьников $A$ и $B$ нашлась

задача, которую решил $A$ и не решил $B$, и задача, которую решил $B$, но не решил $A$. Какое наименьшее возможное количество $n$ задач могло быть при этом условии?

Иными словами, найдите наименьшее возможное $n$, для которого найдётся такое семейство из $k$ подмножеств $n$-элементного множества, что ни одно из подмножеств семейства не содержится (собственно) в другом.

1.5. (a) Найдите $\left|{2 \atop k}\right|$ для $k = 0, 1, 2$. (b) Найдите $\left|{3 \atop k}\right|$ для $k = 0, 1, 2, 3$. (c) $\left|{n \atop 0}\right| = \left|{n \atop n}\right| = 1$, $\left|{n \atop 1}\right| = \left|{n \atop n-1}\right| = 2^n - 1$. (d) $\left|{n \atop k}\right| = \left|{n \atop n-k}\right|$. (e) Найдите $\left|{n \atop 2}\right|$. (f) Найдите $\left|{n \atop k}\right|$.

## 2. Плоские графы

*Плоским графом* называется изображение графа на плоскости, для которого любые два ребра пересекаются только по их общим вершинам (в частности, если таких вершин нет, то не пересекаются).

Иногда такое изображение называют просто графом, но это неточно, поскольку планарный граф можно изобразить (без самопересечений) на плоскости разными способами.

*Рис. 1. Различные изображения графа на плоскости*

Плоский граф делит плоскость на части, называемые *гранями* графа.

2.1. Формула Эйлера. (a) Для связного плоского графа с $V$ вершинами, $E$ рёбрами и $F$ гранями имеем $V - E + F = 2$. (Доказательство этой теоремы см., например, в [1].)

(b) Найдите аналог этого результата для плоского графа с $k$ компонентами связности.

2.2. *Применения формулы Эйлера.* (a) Ни один из графов $K_5$ и $K_{3,3}$ (см. рис. 2) невозможно без самопересечений нарисовать на плоскости.

(b) На плоскости отмечено $n$ точек. Разрешается соединять некоторые две из них ломаной, не проходящей через другие точки. Два игрока по очереди соединяют ломаной какие-то две ещё не соединённые точки. При этом требуется, чтобы эти ломаные не самопересекались и не пересекались нигде, кроме отмеченных точек. Проигрывает тот, кто не может сделать ход. Кто выигрывает при правильной игре (в зависимости от $n$)?


*Рис. 2. Графы Куратовского*

(c) Опишите (с точностью до изоморфизма) плоские графы, у которых степени всех вершин равны и «степени» всех граней равны (т. е. в границе разных граней одинаковое количество рёбер).

Указание к (a), (b) и (c): если не получается, то решайте следующие пункты. Определение изоморфизма см. в задаче 3.5.

(d)* Выпуклых *правильных* многогранников (все грани — правильные многоугольники с одинаковым числом сторон, степени всех вершин равны) ровно 5 (с точностью до изоморфизма их графов). Конструкцию соответствующих многогранников нужно привести, она не предполагается известной.

(e) Для любого плоского связного графа без петель и кратных рёбер, имеющего более двух вершин, $2E \geqslant 3F$ и $E \leqslant 3V - 6$.

(f) В любом плоском графе есть вершина степени не более 5.

(g) Если каждая вершина плоского связного графа имеет степень $d$, а в границе каждой грани ровно $k \geqslant 3$ рёбер, то $\frac{1}{d} + \frac{1}{k} = \frac{1}{2} + \frac{1}{E}$.

2.3. Вершины любого плоского графа можно правильно раскрасить в (a) 6 цветов; (b) 5 цветов; (c)** [Все равно не докажете].

Тор и лист Мёбиуса изображены на рис. 3. Эти фигуры предполагаются *прозрачными*. Т. е. точка (или подмножество), «лежащая на одной стороне поверхности», «лежит и на другой стороне». Это аналогично тому, что при изучении геометрии мы говорим, например, о треугольнике на плоскости, а не о треугольнике на верхней (или нижней) стороне плоскости.

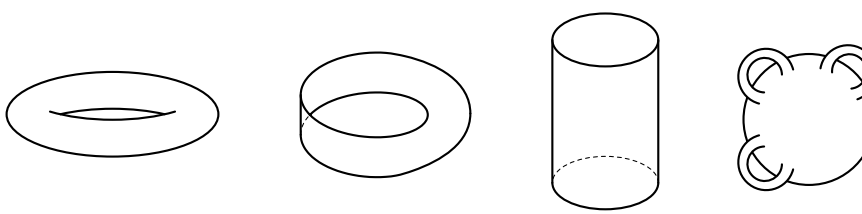


*Рис. 3. Тор, лист Мёбиуса, цилиндр и сфера с ручками*

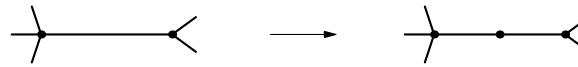


*Рис. 4. Подразделение ребра*

2.4. Нарисуйте без самопересечений (a) $K_5$ на торе. (b) $K_{3,3}$ на листе Мёбиуса.

Грубо говоря, *подграф* данного графа — это его часть. Формально, граф $G$ называется *подграфом* графа $H$, если множество вершин графа $G$ содержится в множестве вершин графа $H$ и каждое ребро графа $G$ является ребром графа $H$. При этом две вершины графа $G$, соединённые ребром в графе $H$, не обязательно соединены ребром в графе $G$.

Ясно, что любой подграф планарного графа планарен.

Операция *подразделения ребра* графа показана на рисунке 4.

Два графа называются *гомеоморфными*, если от одного можно перейти к другому при помощи операций подразделения ребра и обратных к ним. Или, эквивалентно, если существует граф $G$, полученный из обоих данных графов операциями подразделения ребра.

Ясно, что гомеоморфные графы являются или не являются планарными одновременно.

Теорема Куратовского. Граф является планарным тогда и только тогда, когда он не содержит подграфа, гомеоморфного графу $K_5$ или $K_{3,3}$ (рис. 2). (Доказательство этой теоремы см., например, в [6].)

2.5. Придумайте алгоритм

(a) распознавания планарности графа (здесь можно использовать без доказательства теорему Куратовского);

(b)* рисования без самопересечений заведомо планарного графа на плоскости.

Найдите асимптотику сложности вашего алгоритма в зависимости от числа $n$ рёбер графа, т. е. асимптотику максимума по графам с $n$ рёбрами от числа шагов в алгоритме, применённому к данному графу. (Известны линейные алгоритмы, но у вас вряд ли получится даже полиномиальный.)

См. «определение» нахождения асимптотики в задаче 8.1.a.

## 3. Перечисление графов

Заметим, что графы $(\{1,2,3\},\{\{1,2\}\})$ и $(\{1,2,3\},\{\{1,3\}\})$ различны. Графом называется именно граф, а не класс изоморфизма графов (определение изоморфизма напомнено в задаче 3.5, где оно используется). Или, говоря неформально, вершины графов считаются занумерованными. Поэтому граф иногда называют помеченным графом.

*Мультиграфом* (или графом с петлями и кратными рёбрами) называется квадратная таблица из целых неотрицательных чисел, симметричная относительно главной диагонали.

*Ориентированным графом (без петель и кратных рёбер)* $G = (V, E)$ называется конечное множество $V = V(G)$, некоторые *упорядоченные* пары несовпадающих элементов которого выделены.

*Ориентированным мультиграфом* (или ориентированным графом с петлями и кратными рёбрами) называется квадратная таблица из целых неотрицательных чисел.

Читатель легко сообразит, как изображать (с самопересечениями) ориентированные или неориентированные графы и мультиграфы на плоскости.

Ребра $ab$ и $ba$ ориентированного графа не считаются кратными.

3.1. Сколько всего графов с $n$ вершинами
(a) ориентированных без кратных рёбер, но, возможно, с петлями?
(b) неориентированных без петель, но, возможно, с кратными рёбрами?

3.2. Сколько всего графов с $n$ вершинами и $k$ рёбрами
(a) неориентированных без петель и кратных рёбер?
(b) неориентированных графов, у которых допускаются кратные ребра и петли?

3.3. Каких графов с $n$ вершинами больше:
(a) имеющих изолированную вершину (т. е., вершину, из которой не выходит ни одного ребра) или нет?
(b) связных или несвязных?

*Циклом* называется путь, в котором первая и последняя вершины совпадают. Путь (цикл) называется *несамопересекающимся*, если он проходит по каждой своей вершине только один раз. Граф называется *деревом*, если он связен и не содержит несамопересекающихся циклов. Граф называется *лесом*, если он не содержит несамопересекающихся циклов. Граф называется *унициклическим*, если он становится лесом после удаления некоторого ребра. Или, эквивалентно, если он имеет ровно один несамопересекающийся цикл.

3.4. Каких графов больше, деревьев со 100 вершинами или связных унициклических графов с 98 вершинами?

3.5. Грубо говоря, графы изоморфны, если они одинаковы (при этом их изображения на плоскости могут быть разными). Формально, графы $G_1$ и $G_2$ (без петель и кратных рёбер) называются *изоморфными*, если существует взаимно однозначное отображение $f : V_1 \to V_2$ множества $V_1$ вершин графа $G_1$ на множество $V_2$ вершин графа $G_2$, удовлетворяющее

условию: *вершины* $A, B \in V_1$ *соединены ребром в том и только в том случае, если вершины* $f(A), f(B) \in V_2$ *соединены ребром*.

Количество классов изоморфизма деревьев с $n$ вершинами (т. е. количество различных деревьев с $n$ незанумерованными вершинами) меньше $4^n$.

*Код Прюфера* сопоставляет дереву с вершинами $1, 2, \ldots, n$ последовательность чисел от 1 до $n$ по следующему алгоритму.

Сначала код Прюфера — пустое слово. Пока количество вершин больше двух

1. Выбирается *лист* (вершина степени 1) $v$ с минимальным номером.
2. В код Прюфера добавляется номер вершины, смежной с $v$.
3. Вершина $v$ (и инцидентное ей ребро) удаляются из дерева.

Когда осталось две вершины, алгоритм завершает работу.

3.6. (a) Найдите код Прюфера дерева с вершинами $1, 2, \ldots, 10$ и рёбрами $(8, 9)$, $(8, 4)$, $(4, 10)$, $(10, 3)$, $(3, 5)$, $(10, 6)$, $(10, 1)$, $(1, 7)$, $(1, 2)$.

(b) Восстановите дерево по коду Прюфера 1,1,2,5,4,2,7.

(c) *Формула Кэли*. Число деревьев с $n$ вершинами равно $n^{n-2}$.

3.7. (a) Последовательность из $n$ положительных целых чисел является последовательностью степеней вершин некоторого дерева тогда и только тогда, когда сумма её членов равна $2n - 2$.

(b) Если сумма целых положительных чисел $d_1, \ldots, d_n$, равна $2n - 2$, то количество деревьев с $n$ вершинами, у которых $i$-я вершина имеет степень $d_i$, равно $\dfrac{(n-2)!}{(d_1-1)! \cdot \ldots \cdot (d_n-1)!}$. Иными словами, $(x_1 + \ldots + x_n)^{n-2} = \sum_T x_1^{\deg_T(1)-1} \cdot \ldots \cdot x_n^{\deg_T(n)-1}$, где сумма — по всем деревьям $T$ с вершинами $1, 2, \ldots, n$, и через $\deg_T(k)$ обозначена степень вершины $k$ дерева $T$.

(c) Найдите число связных унициклических графов с $n$ вершинами.

3.8*. Пусть $T_1, \ldots, T_r$ — деревья, множества вершин которых не пересекаются. Сколько есть деревьев, множество вершин которых есть объединение множества вершин этих $r$ деревьев, и которые содержат $T_1, \ldots, T_r$?

## 4. Эйлеровы пути и циклы в графах

4.1. Вершины графа можно раскрасить в 2 цвета так, что каждое ребро будет иметь разноцветные концы, тогда и только тогда, когда граф содержит циклы только чётной длины.

*Эйлеров цикл* (путь) — цикл (путь), проходящий по всем рёбрам графа ровно по одному разу.

4.2. (a) В связном графе есть эйлеров цикл тогда и только тогда, когда степень каждой его вершины четна.

(b) При каком условии в графе существует эйлеров путь?

(c) При каком условии в *ориентированном* графе существует ориентированный эйлеров цикл?

4.3. (a) При каких $n$ граф $K_n$ имеет эйлеров цикл? (b) То же для графа $K_{m,n}$.

4.4. На рёбрах графа, у которого степень каждой вершины четна, можно поставить стрелки так, что у каждой вершины входящая степень будет совпадать с исходящей.

4.5. Если количество вершин нечётной степени в связном графе равно $2k$, то множество его рёбер можно представить в виде объединения $k$ путей, никакой из которых не проходит ни по какому ребру дважды и никакие два из которых не имеют общих рёбер.

4.6. Грани эйлерова плоского графа можно правильно раскрасить в 2 цвета.

4.7. (a) Математик забыл трёхзначный код своего замка́. Замок открывается, если три цифры кода набраны подряд (даже если перед этим были набраны другие цифры). Математик набирает одну цифру в секунду; набранная цифра добавляется в конец. Докажите, что математик сможет открыть замок за 1002 секунды, если в коде могут быть использованы только десять цифр.

(b) Сформулируйте и докажите правило «$0 < 1 < 2 < \cdots < 8 < 9$» открытия замка за 1002 секунды.

4.8. Пусть дан ориентированный граф $G$, у которого на каждом ребре $e$ написан вес $v(e)$. (Этот вес можно понимать как работу, которую нужно затратить для того, чтобы пройти по ребру от начала до конца.) Функция $p\colon V(G) \to \mathbb{R}$ («потенциал») такая, что $v(x,y) = p(x) - p(y)$ для любого ребра $e = (x,y)$ существует тогда и только тогда, когда суммарный вес ребер вдоль любого цикла равен нулю (при прохождении ребра по циклу в направлении, противоположном ориентации, вес в сумму берется с отрицательным знаком).

4.9* Дан связный ориентированный граф с вершинами $1, \dots, n$, у которого входящая степень $d_k$ каждой вершины $k$ равна исходящей.

(a) Существует дерево, содержащее все вершины ориентированного графа, у которого все ребра направлены в сторону вершины 1.

(b) Фиксируем дерево $T$ из (a). Будем обходить ориентированный граф (по стрелкам), проходя по каждому ребру не более одного раза. Сначала выйдем из вершины 1 в произвольном направлении. Далее, пусть мы пришли в некоторую вершину $v$. Выходим из нее по любому ребру, не принадлежащему $T$, если это возможно. А если невозможно, то выходим

из нее по ребру, принадлежащему $T$ (такое ребро единственно). Докажите, что обход закончится в вершине 1 и что в результате обхода получится ориентированный эйлеров цикл.

(c) Число ориентированных эйлеровых циклов в ориентированном графе кратно числу $(d_1-1)!\cdot\ldots\cdot(d_n-1)!$.

## 5. Теорема Турана. Дистанционные графы

*Треугольником* в графе называется цикл длины 3.

5.1. (a) Если граф не содержит треугольников, то $e \leqslant n^2/4$.

(b) Если $e = [n^2/4]+1$, то в графе есть по крайней мере $[n/2]$ треугольников.

(c) Если $n = km$ и граф не содержит полного подграфа с $k+1$ вершинами, то $2e \leqslant k(k-1)m^2$. (Переходя к дополнительному графу, получаем, что если $n = km$ и среди любых $k+1$ вершин некоторые две соединены ребром, то $2e \geqslant km(m-1)$.)

(d) Если среди любых $k+1$ вершин некоторые две соединены ребром, $m := [n/k]$ и $r := k\{n/k\}$, то $2e \geqslant km(m-1)+2mr$.

5.2. (a) Если граф двудолен и не содержит цикла длины 4, то $e \leqslant n^{3/2}/2$.

(b) В любом графе есть двудольный подграф, содержащий не менее половины ребер графа.

(c) Если граф не содержит цикла длины 4, то $e \leqslant n^{3/2}$.

(d) Если граф не содержит подграфа $K_{2,3}$, то $e \leqslant 2n^{3/2}$.

(e) Если граф не содержит подграфа $K_{3,3}$, то $e \leqslant 2n^{5/3}$.

(f)* Для любых целых $s, t$, $0 < s \leqslant t$, если граф с не содержит подграфа $K_{s,t}$, то $e \leqslant (s+t)v^{2-1/s}$.

5.3. Для любых $n$ точек на плоскости существует не более $n$ *диаметров*, т. е. (неупорядоченных) пар точек, расстояние между которыми равно максимуму из всех возможных расстояний между парами из этих $n$ точек.

5.4. Для любых $n$ точек $A_1, \ldots, A_n$ в $\mathbb{R}^d$ обозначим через $D(A_1, \ldots, A_n)$ число (неупорядоченных) пар точек, расстояние между которыми равно 1. Обозначим

$$E_n(d) = \max\{D(A_1,\ldots,A_n) \ : \ A_1,\ldots,A_n \in \mathbb{R}^d\}.$$

(a) $E_n(2) > n[\log_2 n]/4$. (b) $E_n(2) \leqslant 2n^{3/2}$. (c) $E_n(3) \leqslant 2n^{5/3}$.

(d) $\dfrac{(n-1)^2}{4} \leqslant E_n(4) \leqslant \dfrac{2(n+4)^2}{5}$.

5.5. (a) Пусть $V$ — $11^k$-элементное подмножество пространства $\mathbb{R}^k$, любое $10^k$-элементное подмножество которого содержит две точки $x, y$

на расстоянии 1: $|x-y|=1$. Докажите, что для достаточно большого $k$ количество единичных расстояний между точками множества $V$ больше, чем $12^k/2$:

$$\frac{1}{2}\,|\{(x,y)\in V\times V\colon\ |x-y|=1\}|>\frac{12^k}{2}\,.$$

(b) То же больше, чем $12.1^k$.

## 6. Гамильтоновы циклы и пути

*Гамильтонов путь* (*цикл*) в графе — путь (цикл), проходящий через каждую вершину ровно по одному разу.

6.1. *Реберным графом* графа $G$ называется граф, вершины которого — ребра графа $G$; две вершины реберного графа соединены ребром, если соответствующие ребра графа $G$ имеют общую вершину. Найдите в терминах графа $G$ необходимое и достаточное условие наличия гамильтонова цикла в его реберном графе.

6.2. (a) Граф, сумма степеней любых двух несмежных вершин которого не меньше $n-1$, имеет гамильтонов путь.

(b) Граф, сумма степеней любых двух несмежных вершин которого не меньше $n$, имеет гамильтонов цикл.

(c) Если граф связен и $2e\geqslant n^2-3n+6$, то в нем есть гамильтонов цикл.

(d) Если среди любых $k+1$ вершин графа есть ребро и после удаления любой $k-1$ вершины граф остается связным, то в нем есть гамильтонов путь.

6.3. [Метод Дирака] Если $a_1,\dots,a_k$ — максимальный путь среди несамопересекающихся по вершинам путей в графе, $k\geqslant 3$ и $\deg a_1+\deg a_k\geqslant k$, то в этом графе есть несамопересекающийся цикл длины $k$.

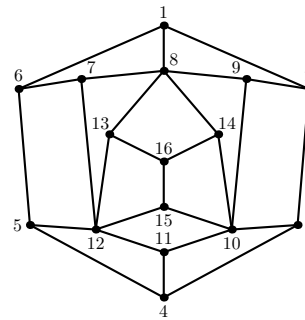



*Рис. 5.*

6.4. В графе на рис. 5 нет гамильтонова пути.

6.5. Максимальное число попарно непересекающихся по ребрам гамильтоновых циклов в полном графе $K_n$ равно $\left[\frac{n-1}{2}\right]$.

## 7. Хроматический многочлен и многочлен Татта

Обозначим через $\chi(G)$ минимальное количество цветов, в которые можно правильно покрасить вершины графа $G$.

7.1. Если для некоторого $k$ в графе с $n$ вершинами среди любых $k+1$ вершин есть ребро, то вершины невозможно правильно покрасить менее, чем в $n/k$ цветов.

Через $G-e$ обозначим граф, получаемый удалением из $G$ ребра $e$. Через $G/e$ обозначим граф, получаемый склеиванием концов ребра $e$ (при этом могут образоваться кратные ребра). *Мостом* называется ребро, при удалении которого количество связных компонент увеличивается. *Остовом графа* называется любое поддерево, содержащее все вершины графа.

7.2. (a) На какое число может измениться хроматическое число графа $G$, если добавить к графу одно ребро? Или, формально, найдите все целые $k$, для которых существует граф $G$ и его ребро $e$ такие, что $\chi(G) - \chi(G-e) = k$.

(b) $\chi(V, E_1 \cup E_2) \leqslant \chi(V, E_1)\chi(V, E_2)$.

(c) Для каждого $r > 0$ постройте такие $V, E_1, E_2$, что $\chi(V, E_1 \cup E_2) = \chi(V, E_1)\chi(V, E_2)$ и $\chi(V, E_1) = r$.

7.3. (a) Если максимальная степень вершины графа $G$ равна $\Delta(G)$, то $\chi(G) \leqslant \Delta(G) + 1$.

(b) Если дан граф $G$ с $n$ вершинами и $e$ ребрами, причём при удалении любой вершины хроматическое число уменьшается, то $2e \geqslant n(\chi(G) - 1)$.

7.4. (a) Для любого ребра $e$ графа $G$ выполнено $\chi(G) = \chi(G-e) - \chi(G/e)$. (Можно считать, что кратные ребра в графе $G/e$ заменены на некратные.)

(b) *Теорема Биркгофа – Уитни*. Число $\chi_G(t)$ правильных раскрасок графа $G$ в $t$ цветов есть многочлен от $t$.

(c) Его степень равна количеству вершин, старший коэффициент 1, второй коэффициент равен числу рёбер в графе, взятому со знаком минус, коэффициенты знакопеременны (т. е. коэффициенты при $t^{n-2k}$ неотрицательны и коэффициенты при $t^{n-2k+1}$ неположительны для любого целого $k$).

7.5. Вычислите хроматический многочлен для (a) полного графа; (b) пустого графа; (c) цепи; (d) цикла; (e) данного дерева с $n$ вершинами.

7.6. (a) Если хроматический многочлен графа есть $t(t-1)^{n-1}$, то граф — дерево.

(b) Не существует графа с хроматическим многочленом $t^4-3t^3+3t^2$.

7.7. Для графов с петлями и кратными рёбрами $T(G)=T(G-e)+$ $+T(G/e)$, где $e$ — ребро графа $G$, не являющееся ни петлей, ни мостом, и $T(G)$ —

(1,1) число остовных лесов (т. е. объединений остовных деревьев его компонент).

(1,2) число таких наборов рёбер, что для любой компоненты связности графа лежащие в ней ребра из набора образуют связный подграф.

(2,1) число ациклических наборов рёбер.

7.8. *Теорема Тамма*. В мультимножестве (т. е. в неупорядоченном наборе с кратностями) графов разрешается для любого графа $G$ и его ребра $e$, не являющегося ни петлей, ни мостом, заменять один граф $G$ на два графа $G-e$, $G/e$. Эта замена применяется до тех пор, пока не останутся только графы, в которых все ребра являются петлями или мостами. Тогда для любого исходного графа полученное мультимножество корректно определено, т. е. не зависит от порядка графов и рёбер, к которым мы применяем замены.

*Многочленом Тамма* называется многочлен $T(x,y)$ от двух переменных, определённый рекуррентной формулой $T_G=T_{G-e}+T_{G/e}$, если $e$ не петля и не мост, и $T_G(x,y)=x^iy^j$, если $G$ имеет $i$ мостов, $j$ петель и не имеет других рёбер.

7.9. (a) Выразите хроматический многочлен через многочлен Татта.

(b) Почему так странно занумерованы пункты в задаче 7.7?

## 8. Асимптотики

Будет полезна *формула Стирлинга* $n!\sim n^ne^{-n}\sqrt{2\pi n}$, т. е.

$$\lim_{n\to\infty}\frac{n!}{n^ne^{-n}\sqrt{2\pi n}}=1.$$

8.1. (a) Найдите асимптотику для количества $A_n$ подмножеств множества $\{1,2,\dots,n\}$, не содержащих двух подряд идущих чисел.

(Иными словами, найдите «явную» функцию $f(n)$, для которой $\lim\limits_{n\to\infty}\dfrac{A_n}{f(n)}=1$.)

(b*) То же для *трёх* подряд идущих чисел.

В ответе можно использовать функцию $x_P(a,b)$, которая по числам $a,b$ и многочлену $P$, имеющему единственный корень на отрезке $[a,b]$, выдаёт этот корень.

8.2. (a) $\dfrac{2^n}{n+1} < \binom{n}{[n/2]} < 2^n$. (b) $\binom{n}{[n/2]} \sim \dfrac{2^n}{\sqrt{\pi n/2}}$.

8.3. (a) $\binom{n}{[an]} = (a^{-a}(1-a)^{a-1} + o(1))^n$.

(b) $\dfrac{n!}{[a_1 n]! \cdot \ldots \cdot [a_s n]!} = (e^{-a_1 \ln a_1 - \cdots - a_s \ln a_s} + o(1))^n$, где $a_1 + \cdots + a_s = 1$.

8.4. (a) $n^2(2 + \frac{1}{n})^n = (2 + o(1))^n$. (b) $3^{\sqrt{n}}(2 + \frac{1}{n})^n = (2 + o(1))^n$.

В пунктах (b,c,d) следующей задачи достаточно интуитивного понимания того, что такое вероятность.

8.5. (a) $\ln \dfrac{(n-1)(n-2) \cdot \ldots \cdot (n-k)}{n^k} = -\dfrac{k(k+1)}{2n}\left(1 + O\left(\dfrac{k}{n}\right)\right)$ для $k = k_n$.

(b) $\binom{n}{k} \sim \dfrac{n^k}{k!}$ для $k = k_n = o(\sqrt{n})$. Значит, для $k \ll \sqrt{n}$ вероятность выпадения ровно $k$ орлов при $n$ подбрасываниях монеты приближённо равна $\dfrac{n^k}{k!} 2^{-n}$.

(c) Для $k \ll \sqrt{n}$ вероятность получения ровно $k$ успехов в серии из $n$ опытов, с вероятностью $\lambda/n$ успеха в одном опыте, приближённо равна $\dfrac{\lambda^k}{k!} e^{-\lambda}$ (*распределение Пуассона*).

(d) $\binom{2n}{n-k} \Big/ \binom{2n}{n} = e^{-\frac{k^2}{n}(1+o(1))}$ для $k = k_n = o(n)$. Значит, для $k \ll n$ вероятность $P_k$ выпадения ровно $n-k$ орлов при $2n$ подбрасываниях монеты приближённо равна $P_0 e^{-k^2/n}$ (*нормальное распределение*).

8.6. Верно ли, что записи $e^{o(n)}$ и $o\,(e^n)$ «равнозначны»? Т. е. верно ли, что для любой функции $f\colon \mathbb{Z} \to (0, +\infty)$ условия $\lim\limits_{n\to\infty} \dfrac{\ln f(n)}{n} = 0$ и $\lim\limits_{n\to\infty} f(n)e^{-n} = 0$ равносильны?

8.7. Какая функция растёт быстрее: $x^{(x^x)}$ или $(x!)^{(2^x)}$? Т.е. найдите $\lim\limits_{x\to\infty} x^{(x^x)}(x!)^{(-2^x)}$.

8.8. Найдите асимптотику для

(a) $\ln \binom{n^2}{n}$. (b)* $\binom{n}{[n^\alpha]}$, $\alpha \in (0,1)$. (c) $(2n-1)!!$. (d) $\sum\limits_{k=0}^{n} \binom{n}{k}^2$.

(e)* $\sum\limits_{k=0}^{n} \binom{n}{k}^4$.

8.9. Существует ли функция $\varphi(n) = o(1)$, для которой

$$(2 + \varphi(n))^n 2^{-n} e^{-\sqrt{n}} \to \infty?$$

(Как в любой математической задаче, нужно обосновать ответ: привести пример такой функции или доказать её существование или доказать, что такой функции не существует.)

8.10. Найдите асимптотику функции $s = s(n)$, заданной как (a) $s^s = n$; (b) $s^{s^3} = n$.

8.11. (a) Найдите асимптотику для величины из 1.4.b.

(b)* Найдите асимптотику для количества линейных подпространств в $\mathbb{Z}_2^n$. В ответе можно использовать константу, заданную в виде суммы ряда.

(c)* Найдите асимптотику числа связных унициклических графов с $n$ вершинами.

8.12. В этой задаче, в отличие от остальных, нельзя пользоваться формулой Стирлинга.

(a) $n^n e^{-n+1} < n! < n^{n+1} e^{-n+1}$; (b) $n! < n^n e^{-n+1}\sqrt{n}$.

## 9. Простейший вероятностный метод, или подсчёт

9.1. Существует такая раскраска рёбер графа $K_{m,n}$ в два цвета, что число одноцветных подграфов $K_{a,b}$ не больше $\binom{m}{a}\binom{n}{b}2^{1-ab}$.

9.2. *$k$-однородный гиперграф* $H = (V, E)$ — это множество $V$ вершин и система $E$ из $k$-элементных подмножеств множества вершин. Эти подмножества называют рёбрами (математики других специальностей называют их симплексами).

(a) Если $|E| \leqslant 2^{k-1}$, то вершины гиперграфа можно раскрасить в два цвета правильно (т. е. так, чтобы не нашлось ребра, все вершины которого одноцветны).

(b) Если для некоторого чётного $n$

$$\left(1 - 2\frac{\binom{n/2}{k}}{\binom{n}{k}}\right)^e 2^n < 1,$$

то существует $k$-однородный гиперграф $H = (V, E)$ с $e$ рёбрами, вершины которого нельзя правильно раскрасить в два цвета.

(c) Существует такое $c > 0$, что для любого $k$ существует $k$-однородный гиперграф, который нельзя правильно раскрасить в два цвета и который имеет не более $ck^2 2^k$ рёбер.

9.3. Для любых $n$ векторов $v_1, \ldots, v_n \in \mathbb{R}^n$ длины 1 существует такой набор $\varepsilon_1, \ldots, \varepsilon_n = \pm 1$, что (a) $|\sum_{k=1}^n \varepsilon_k v_k| \leqslant \sqrt{n}$; (b) $|\sum_{k=1}^n \varepsilon_k v_k| \geqslant \sqrt{n}$.

9.4. Имеется несколько цветов. Каждой вершине двудольного графа с $n$ вершинами сопоставлено не менее, чем $\log_2 n + 1$ из них. Тогда существует правильная раскраска графа, приписывающая каждой вершине некоторый сопоставленный ей цвет.

9.5. В любом множестве из $n$ различных натуральных чисел найдётся подмножество из более, чем $n/3$ чисел, в котором нет трёх чисел, сумма двух из которых равна третьему.

9.6. Если в графе $G = (V, E)$ с $n$ вершинами минимальная степень вершины равна $\delta$, то существует такое множество вершин

(a) $A \subset V$, что имеется не более $np + n(1-p)^{\delta+1}$ вершин, лежащих в $A$ или не соединённых ни с какой вершиной из $A$. Здесь $p \in (0,1)$ — произвольное наперёд заданное число.

(b) $D \subset V$, что любая вершина из $V \backslash D$ соединена ребром с некоторой вершиной из $D$, и $|D| \leqslant n\dfrac{1+\ln(\delta+1)}{\delta+1}$.

## 10. Случайные графы

Назовём *вероятностью* графа (в модели, или в вероятностном пространстве, $G(n,p)$) с $n$ вершинами $\{1,2,\ldots,n\}$ и $e$ рёбрами число $\mathrm{P}_p(G) := p^e(1-p)^{n(n-1)/2-e}$. *Вероятностью* произвольного семейства (или, что то же самое, свойства) графов с множеством вершин $\{1,2,\ldots,n\}$ называется сумма вероятностей входящих в него графов.

*Случайной величиной* называется функция $Y$, определённая на множестве графов.

Пусть случайная величина $Y$ принимает $k$ различных значений $y_1, \ldots, y_k$. Тогда *математическим ожиданием* (матожиданием) случайной величины $Y$ называется её «взвешенное среднее» $\mathrm{E}Y = \sum\limits_{s=1}^{k} y_i \mathrm{P}(Y^{-1}(y_s))$, где $Y^{-1}(y_s)$ — множество всех графов $G$, для которых $Y(G) = y_s$. Последнюю вероятность можно обозначать $\mathrm{P}(Y = y_s)$.

10.1. Для данных $n$ и $p$ найдите матожидание количества (a) гамильтоновых циклов; (b) несамопересекающихся циклов.

10.2. (a) Вероятность наличия $k$ вершин, между которыми нет рёбер, меньше $e^{k\ln n - pk(k-1)/2}$.

(b) Для любых целых $l, q > 0$ существует граф, не содержащий несамопересекающихся циклов длины $\leqslant l$, который невозможно правильно раскрасить в $q$ цветов.

10.3. Для данного $p$ вычислите асимптотику (при $n \to \infty$) для $\mathrm{E}^{(k)}(Y) := \mathrm{E}(Y(Y-1)\dots(Y-k+1))$ (т. е. для *$k$-го факториального момента*), если $Y$ — число изолированных вершин.

Событие $A_n$ *происходит асимптотически почти наверное* (или с асимптотической вероятностью 1), если $\mathrm{P}_{p(n)}(A_n) \to 1$.

10.4. При $p = p(n) = 1/(2n)$

(a) для некоторой последовательности $d_n \to 0$ асимптотически почти наверное имеется $(1+d_n)n/2$ изолированных вершин (специалисты говорят: имеется $(1+o(1))n/2$ изолированных вершин).

(b) для некоторого $C > 0$ асимптотически почти наверное каждая компонента связности имеет менее $C \ln n$ вершин (специалисты говорят: менее $O(\ln n)$ вершин).

(c) асимптотически почти наверное каждая компонента связности является деревом или унициклическим графом.

(d) для некоторого $C > 0$ асимптотически почти наверное имеется менее $C$ унициклических компонент.

10.5. (a) При $p = p(n) = o\left(n^{-3/2}\right)$ асимптотически почти наверное ребра попарно не пересекаются.

(b) При $p = p(n) = o\left(n^{-3/2}\right)$ и $pn^2 \to \infty$ существует такая функция $M(n)$, что $2M(n) \sim pn^2$ и асимптотически почти наверное $2M(n)$ степеней вершин равны 1, а остальные степени равны нулю.

10.6. (a) Найдите хотя бы одну такую функцию $p^*(n)$, что

– при $p(n)/p^*(n) \to 0$ асимптотически почти наверное граф не содержит подграфа, изоморфного $K_4$, и

– при $p(n)/p^*(n) \to +\infty$ асимптотически почти наверное граф содержит подграф, изоморфный $K_4$.

(b)* То же с заменой $K_4$ на заданный граф с $v$ вершинами и $e$ рёбрами.

## 11. Пересечения подмножеств

В этом разделе через $\mathcal{F}$ обозначается произвольное семейство $k$-элементных подмножеств $n$-элементного множества.

11.1. (a) Если $k \geqslant l$ и каждое $l$-элементное подмножество $n$-элементного множества содержится в некотором подмножестве из $\mathcal{F}$, то $|\mathcal{F}| \geqslant \binom{n}{l} \Big/ \binom{k}{l}$.

(b) Количество $(k-1)$-элементных подмножеств $n$-элементного множества, целиком содержащихся хотя бы в одном из подмножеств семейства $\mathcal{F}$, не менее $\dfrac{k|\mathcal{F}|}{n-k+1}$.

11.2. В любом семействе попарно пересекающихся подмножеств $n$-элементного множества не более $2^{n-1}$ подмножеств.

11.3. *Теорема Эрдёша – Ко – Радо.* (a) Если $2k \leqslant n$ и любые два подмножества из $\mathcal{F}$ пересекаются, то $|\mathcal{F}| \leqslant \binom{n-1}{k-1}$.

(b) Если $2k \geqslant n$ и объединение никаких двух подмножеств из $\mathcal{F}$ не есть все $n$-элементное множество, то $|\mathcal{F}| \leqslant \binom{n-1}{k}$.

11.4. Пусть $n-2 \geqslant t \geqslant 2$.

(a) Постройте семейство из $2^{n-t}$ подмножеств $n$-элементного множества, любые два из которых пересекаются не менее, чем по $t$ элементам.

(b) Существует ли такое семейство из $2^{n-t}+1$ подмножеств?

*Подсолнухом с $m$ лепестками и ядром $Y$* называют набор множеств $\{F_1, \ldots, F_m\}$ такой, что $|F_i \cap F_j| = Y$ при $i \neq j$ и все множества $F_i \setminus Y$ непусты.

Например, попарно непересекающиеся множества образуют подсолнух с пустым ядром.

11.5. (a) Если $|\mathcal{F}| > k!(m-1)^k$, то в $\mathcal{F}$ найдётся подсолнух с $m$ лепестками.

(b) Найдутся $(m-1)^k$ подмножеств $n$-элементного множества, в каждом из которых $k$ элементов и среди которых нельзя выбрать подсолнух с $m$ лепестками.

## 12. Линейно-алгебраический метод

12.1. Дано семейство $\mathcal{F}$ подмножеств множества $\{1, \ldots, n\}$.

(a) Если в каждом подмножестве из $\mathcal{F}$ нечётное число элементов, а в пересечении любых двух подмножеств из $\mathcal{F}$ чётное число элементов, то $|\mathcal{F}| \leqslant n$.

(b) Постройте пример, когда эта оценка достигается.

(c) Если в пересечении любых двух подмножеств из $\mathcal{F}$ ровно $k$ элементов и в каждом подмножестве из $\mathcal{F}$ более $k$ элементов, то $|\mathcal{F}| \leqslant n$.

(d) Если $k > 0$ и в пересечении любых двух подмножеств из $\mathcal{F}$ ровно $k$ элементов, то $|\mathcal{F}| \leqslant n$.

12.2. (a) Приведите пример $2^k$ подмножеств $2k$-элементного множества, в каждом из которых чётное число элементов, и в пересечении любых двух из которых чётное число элементов.

(b) Больше, чем $2^k$ подмножеств в условиях пункта (a) быть не может.

12.3. (a) Среди любых 327 попарно пересекающихся девятиэлементных подмножеств 25-элементного множества найдутся два подмножества, в пересечении которых ровно 3 или ровно 6 элементов.

(b) Для $k \in \mathbb{Z}$ обозначим $V_{n,k} := \{(x_1, \dots, x_n) \in \{0,1\}^n \ : \ \sum_s x_s = k\}$. Среди любых 327 точек в $V_{25,9}$ есть две, скалярное произведение которых лежит в $\{0, 3, 6\}$.

(c) Для любого $\vec{a} \in V_{25,9}$ раскроем скобки в произведении

$$(\vec{a} \cdot (x_1, x_2, \dots, x_{25}) - 1)(\vec{a} \cdot (x_1, x_2, \dots, x_{25}) - 2),$$

где $x_1, x_2, \dots, x_{25}$ — переменные. В каждом из полученных одночленов для каждого $i$ будем заменять $x_i^2$ на 1, пока это возможно. Полученный многочлен обозначим $F_{\vec{a}}(x_1, \dots, x_{25})$.

Докажите, что если скалярное произведение никаких двух векторов среди $\vec{a}_1, \dots, \vec{a}_s \in V_{25,9}$ не делится на 3, то многочлены $F_{\vec{a}_1}, \dots, F_{\vec{a}_s}$ линейно независимы над $\mathbb{Q}$.

12.4. (a) Среди любых 107 пятиэлементных подмножеств 14-элементного множества найдутся два подмножества, в пересечении которых ровно 2 элемента.

(b) То же для 93 подмножеств.

(c) То же для 92 подмножеств.

(d) Невозможно раскрасить в 21 цвет все пятиэлементные подмножества 14-элементного множества так, чтобы любые два пятиэлементные подмножества, пересекающиеся ровно по двум элементам, были разноцветны.

12.5. (a) Наибольшее число точек в $\mathbb{R}^n$ с равными попарными расстояниями равно $n + 1$.

(b) Постройте $n(n-1)/2$ точек в $\mathbb{R}^n$, попарные расстояния между которыми принимают только два различных значения.

(c) Если попарные расстояния между $m$ точками в $\mathbb{R}^n$ принимают только два различных значения, то $m \leqslant (n+1)(n+4)/2$.

12.6. (a) Для простого $p$ и целого $t$ число

$$G(t) := (t-1)(t-2) \cdot \ldots \cdot (t-p+1)$$

делится на $p$ тогда и только тогда, когда $t$ не делится на $p$.

(b) Пусть $p$ простое и $n = 4p$. Обозначим

$$M = \{(1, y_2, y_3, \dots, y_n) \mid y_1 = 1,\ y_k \in \{1, -1\} \text{ и среди } y_2, \dots, y_n \text{ число минус единиц чётно}\}.$$

Для любого $\vec{a} \in M$ раскроем скобки в произведении $G(\vec{a} \cdot (1, x_2, \dots, x_n))$, где $x_1, x_2, \dots, x_n$ — переменные. В каждом из полученных одночленов для

каждого $i$ будем заменять $x_i^2$ на 1, пока это возможно. Полученный многочлен обозначим $F_{\vec{a}}(x_2, \ldots, x_n)$.

Докажите, что если скалярное произведение никаких векторов среди $\vec{a}_1, \ldots, \vec{a}_s \in M$ не равно нулю, то многочлены $F_{\vec{a}_1}, \ldots, F_{\vec{a}_s}$ линейно независимы над $\mathbb{Q}$.

(c)* Существует $n$ и ограниченное подмножество в $\mathbb{R}^n$, которое невозможно разбить на $n+1$ непустых частей меньшего диаметра.

12.7. (a) Если множество рёбер графа $K_n$ является объединением множеств рёбер $m$ полных двудольных графов, не пересекающихся по рёбрам, то $m \geqslant n-1$.

(b) Постройте набор двудольных графов, на котором эта оценка достигается.

Задачи этого пункта заимствованы из [2, 3, 5, 7], где читатель найдёт решения и обобщения.

## Список литературы

[1] Прасолов В. В. *Элементы комбинаторной и дифференциальной топологии*. М.: МЦНМО. 2004. `http://www.mccme.ru/prasolov`

[2] Райгородский А. М. *Проблема Борсука*. М.: МЦНМО, 2006.

[3] Райгородский А. М. *Линейно-алгебраический метод в комбинаторике*. М.: МЦНМО, 2007.

[4] Райгородский А. М. *Вероятность и алгебра в комбинаторике*. М.: МЦНМО, 2010 (второе издание).

[5] Babai L., Frankl P. *Linear algebra methods in combinatorics. Part* 1. Department of Computer Science. The University of Chicago. Preliminary version 2. September 1992.

[6] Skopenkov A. *On the Kuratowski graph planarity criterion*. arXiv:0802.3820v3. 2012.

[7] Skopenkov A. *A two-page disproof of the Borsuk partition conjecture*. arXiv:0712.4009v2. 2012.

---

Д. Ильинский, А. Купавский, А. Райгородский, А. Скопенков — Московский Физико-Технический Институт
Личные страницы: `http://dm.fizteh.ru/staff`

# По мотивам задачника «Математического просвещения»

## Семь этюдов об одном несовпадении

Н. Н. Осипов

*— Видите, Балаганов, что можно сделать из простой швейной машины Зингера? Небольшое приспособление — и получилась прелестная колхозная сноповязалка.*

И. Ильф, Е. Петров. *Золотой телёнок*

На одной недавней математической олимпиаде[1)] её участникам была предложена следующая

Задача. *Докажите, что числа* $\operatorname{arctg}(4/3)$ *и* $\pi$ *несоизмеримы.*

Иными словами, требовалось показать, что число

$$\frac{\operatorname{arctg}(4/3)}{\pi}$$

иррационально. Если привлечь комплексные числа, то это утверждение можно сформулировать так: *число*

$$\frac{3+4\sqrt{-1}}{5}$$

[1)] *Региональная студенческая олимпиада по математике*, состоявшаяся в апреле 2010 года. Регулярно проводится Институтом математики Сибирского федерального университета (Красноярск) и обычно собирает команды университетов Абакана, Кемерово, Новосибирска, Улан-Удэ и самого Красноярска.

*не является корнем из единицы.* Таким образом, решение задачи сводится к доказательству такого факта: *при любом* $n = 1, 2, \ldots$ *равенство*

$$(3 + 4\sqrt{-1})^n = 5^n \qquad (*)$$

*невозможно.* Кому-то это может показаться очевидным: в самом деле, ведь не может быть, чтобы мнимое комплексное число $(3 + 4\sqrt{-1})^n$ оказалось равным вещественному числу $5^n$. Но почему, собственно, первое число является мнимым? Как это можно доказать? И вообще, какие есть способы аккуратно опровергнуть равенство $(*)$?

В дальнейшем изложении мы будем использовать некоторые стандартные понятия и факты из теории колец и теории многочленов. Все необходимые для понимания предварительные сведения читатель при желании сможет найти, например, в книгах [1] и [6].

## Этюд I. Канонический эпиморфизм

В *факторкольце* $R_1 = \mathbb{Z}[\sqrt{-1}]/I_1$ *кольца целых гауссовых чисел*

$$\mathbb{Z}[\sqrt{-1}] = \{a + b\sqrt{-1} : (a, b) \in \mathbb{Z}^2\}$$

по *идеалу* $I_1 = (5)$ имеем $5 = 0$ и $(3 + 4\sqrt{-1})^2 = 3 + 4\sqrt{-1}$. Но тогда

$$5^n = 0, \quad (3 + 4\sqrt{-1})^n = 3 + 4\sqrt{-1}$$

при любом натуральном $n$, и равенство $(*)$ не может быть верным.

А в факторкольце $R_2 = \mathbb{Z}[\sqrt{-1}]/I_2$, где $I_2 = (1+2\sqrt{-1})$, ещё нагляднее:

$$5^n = 0, \quad (3 + 4\sqrt{-1})^n = 1,$$

поскольку по-прежнему $5 = 0$, но теперь $3 + 4\sqrt{-1} = 1$.

Отметим кстати, что кольцо $R_1$ состоит из $25 = 5^2$ элементов и не является полем (ибо, например, содержит *делители нуля* $1 \pm 2\sqrt{-1}$), а $R_2$ — поле из 5 элементов (в отличие от $I_1 \subset I_2$, идеал $I_2$ уже *максимален*).

## Этюд II. Факториальное кольцо

Поскольку $3 + 4\sqrt{-1} = (2 + \sqrt{-1})^2$ и $5 = (2 + \sqrt{-1})(2 - \sqrt{-1})$, равенство $(*)$ можно переписать в виде

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

Но это равенство невозможно, поскольку в *евклидовом* (а значит, и *факториальном*) кольце $\mathbb{Z}[\sqrt{-1}]$ оба числа $2 \pm \sqrt{-1}$ являются *простыми*, причём *неассоциированными*.

Противоречивость равенства $(*)$ также станет очевидной, если заметить, что число $1+2\sqrt{-1}$ — простой делитель 5, который не делит $3+4\sqrt{-1}$ ввиду равенства $3 + 4\sqrt{-1} = 2(1 + 2\sqrt{-1}) + 1$.

## Этюд III. Циклотомические многочлены

Хорошо известно, что *циклотомический многочлен* порядка $n$

$$\Phi_n(x) = \prod_{\text{НОД}\,(a,n)=1} (x - \zeta_n^a), \quad \zeta_n = \exp\,(2\pi\sqrt{-1}/n),$$

имеет целочисленные коэффициенты и неприводим над полем рациональных чисел $\mathbb{Q}$. Как следствие, любой многочлен $f(x) \in \mathbb{Q}[x]$, имеющий общие корни с $\Phi_n(x)$, обязан делиться на $\Phi_n(x)$. В частности, справедливо неравенство

$$\deg f(x) \geqslant \deg \Phi_n(x) = \varphi(n), \qquad (\dagger)$$

где $\varphi(n)$ — *функция Эйлера*.

Пусть теперь $n$ — наименьшее натуральное число, для которого равенство $(*)$ справедливо. Тогда

$$\frac{3 + 4\sqrt{-1}}{5} = \zeta_n^a$$

для некоторого $a$, взаимно простого с $n$. Поскольку число слева — *квадратичная иррациональность*, имеем $\varphi(n) \leqslant 2$, откуда $n \in \{1, 3, 4, 6\}$. Но, как показывает непосредственная проверка, для этих значений $n$ равенство $(*)$ не выполняется.

Фактически из $(\dagger)$ вытекает следующая более общая

Теорема. *Если* $\varphi/\pi \in \mathbb{Q}$ *и* $\cos\varphi \in \mathbb{Q}$, *то* $\cos\varphi \in \{0, \pm 1/2, \pm 1\}$.

Действительно, если $\varphi = 2\pi a/n$, где $0 \leqslant a < n$ и $\text{НОД}\,(a, n) = 1$, то

$$2\cos\varphi = \zeta_n^a + \zeta_n^{-a},$$

откуда $\zeta_n^a$ — корень $f(x) = x^2 - (2\cos\varphi)x + 1 \in \mathbb{Q}[x]$.

Другие доказательства этой теоремы будут даны в этюдах V и VI.

## Этюд IV. Диадический показатель

Из равенства $(*)$ вытекает, что мнимая часть числа

$$(3 + 4\sqrt{-1})^n = (2 + \sqrt{-1})^{2n} = X_n + Y_n\sqrt{-1}$$

должна быть нулевой. Применив *биномиальную формулу*, найдём

$$Y_n = 2(-1)^{n-1} \sum_{k=0}^{n-1} (-1)^k C_{2n}^{2k+1} 2^{2k}.$$

Если равенство $Y_n = 0$ записать в виде

$$C_{2n}^1 = -\sum_{k=1}^{n-1} (-1)^k C_{2n}^{2k+1} 2^{2k},$$

станет понятно, что оно невозможно. Действительно,

$$C_{2n}^{2k+1} = \frac{C_{2n}^1 C_{2n-1}^{2k}}{2k+1},$$

поэтому при $k \geqslant 1$ имеем

$$\nu_2(C_{2n}^{2k+1} 2^{2k}) = \nu_2\left(\frac{C_{2n}^1 C_{2n-1}^{2k} 2^{2k}}{2k+1}\right) \geqslant \nu_2(C_{2n}^1) + 2k > \nu_2(C_{2n}^1),$$

где $\nu_2(m)$ — 2-*адический показатель* натурального числа $m$, т. е. такое целое число $l \geqslant 0$, что $m$ делится на $2^l$, но не делится на $2^{l+1}$.

## Этюд V. Многочлены Чебышёва и целые алгебраические числа

При любом $n = 1, 2, \ldots$ двучлен $z^n + 1/z^n$ можно представить в виде многочлена $p_n(y)$ от $y = z + 1/z$, при этом $p_n(y)$ нормирован и имеет целые коэффициенты. В самом деле, имеем

$$p_0(y) = 2, \quad p_1(y) = y, \quad p_{n+1}(y) = y p_n(y) - p_{n-1}(y).$$

Если $z = \cos\varphi + \sqrt{-1}\sin\varphi$, то $y = 2\cos\varphi$ и из *формулы Муавра* получим

$$p_n(y) = z^n + 1/z^n = 2\cos(n\varphi) = 2T_n(\cos\varphi) = 2T_n(y/2),$$

где $T_n(x)$ — *многочлен Чебышёва* 1-го рода.

Теперь запишем равенство $(*)$ в виде $z^n = 1$, где

$$z = \frac{3 + 4\sqrt{-1}}{5}.$$

Тогда $y = 6/5$, а также

$$p_n(y) = p_n(6/5) = z^n + 1/z^n = 2.$$

Итак, $6/5$ оказывается рациональным, но не целым корнем нормированного многочлена с целыми коэффициентами $p_n(y) - 2$, что невозможно.

Это рассуждение — ещё один способ доказательства теоремы из этюда III. Его можно сделать совсем кратким, если заметить, что число

$$y = 2\cos\varphi = \zeta_n^a + \zeta_n^{-a} = \zeta_n^a + \zeta_n^{n-a}$$

является *целым алгебраическим* (поскольку таково $\zeta_n$, а целые алгебраические числа образуют кольцо). Но в таком случае условие $y \in \mathbb{Q}$ равносильно условию $y \in \mathbb{Z}$, а значит, $\cos\varphi \in \{0, \pm 1/2, \pm 1\}$.

## Этюд VI. Динамические системы

Рассмотрим итерационный процесс, заданный формулой

$$x_{k+1} = 2x_k^2 - 1 \quad (k = 0, 1, 2, \ldots).$$

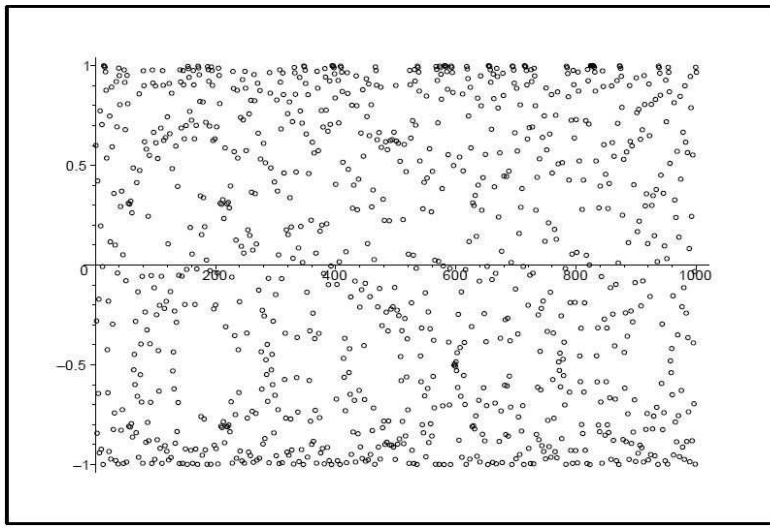



*Рис. 1. Первая тысяча членов последовательности* $\{x_k\}$ *при* $x_0 = 3/5$

Введём обозначение: $I_{\mathbb{Q}} = \{r \in \mathbb{Q} : -1 \leqslant r \leqslant 1\}$.

Предложение. *При* $x_0 \in I_{\mathbb{Q}} \setminus \{0, \pm 1/2, \pm 1\}$ *все члены последовательности* $\{x_k\}$ *попарно различны.*

В описанной ситуации поведение последовательности $\{x_k\}$ выглядит хаотическим (см. рис. 1). Для доказательства предложения нам понадобится

Лемма. *Пусть* $f(x) = 2x^2 - 1$. *Для* $k = 1, 2, \ldots$ *положим*

$$f_k(x) = \underbrace{f(\ldots(f}_{k}(x)\ldots).$$

*Тогда рациональные корни уравнения* $f_k(x) = x$ *суть* $1$ *и* $-1/2$.

Доказательство. При $k > 1$ справедливо представление

$$f_k(x) = 2^{2^k-1}x^{2^k} + \ldots + 1.$$

Поэтому любой рациональный корень $x_0$ многочлена $f_k(x) - x$ должен иметь вид $\pm 1/2^l$, где $l$ — некоторое целое неотрицательное число. Имеем

$$f_k(x_0) = f_2(y_0) = 8y_0^4 - 8y_0^2 + 1 = x_0,$$

где $y_0 = f_{k-2}(x_0)$ — рациональное число (считаем $f_0(x) = x$). Но, как нетрудно убедиться, уравнение вида

$$8y^4 - 8y^2 + 1 = \pm 1/2^l$$

имеет рациональные корни только тогда, когда его правая часть равна 1 или $-1/2$.

Переходя к доказательству предложения, заметим, что если $x_{N+k} = x_N$, то $f_k(x_N) = x_N$ и по лемме $x_N \in \{0, \pm 1/2, \pm 1\}$. Но в таком случае, как нетрудно проверить, $x_{N-1} \in \{0, \pm 1/2, \pm 1\}$ и т. д. — противоречие.

Расскажем ещё о двух других способах доказательства предложения, в каждом из которых эксплуатируется представление рациональных чисел $x_k$ в виде несократимых дробей:

$$x_k = \frac{a_k}{b_k}\,, \quad \text{НОД}\,(a_k, b_k) = 1.$$

I. Первый способ совсем короткий. Имеем

$$x_{k+1} = \frac{a_{k+1}}{b_{k+1}} = \frac{2a_k^2 - b_k^2}{b_k^2}\,. \qquad (\ddagger)$$

Так как $d_k = \text{НОД}\,(2a_k^2 - b_k^2, b_k^2) \in \{1, 2\}$, то

$$b_{k+1} = \frac{b_k^2}{d_k} \geqslant \frac{b_k^2}{2} > b_k\,,$$

как только $b_k > 2$. Осталось заметить, что $b_0 > 2$.

Подобные элементарные рассуждения составляют основу интересного сюжета о том, как из одной известной рациональной точки на данной *эллиптической кривой* можно изготовить бесконечную последовательность рациональных точек этой кривой.

Приведём один пример. Пусть $S$ — положительное рациональное число, $E_S$ — эллиптическая кривая, определяемая уравнением

$$Sy^2 = x^3 - x.$$

Теорема. *Если кривая $E_S$ содержит хотя бы одну рациональную точку $P_0 = (x_0, y_0)$ с $y_0 \neq 0$, то таких точек на ней бесконечно много.*

Доказательство. Проведём в точке $P_0$ касательную к $E_S$. Она пересечёт $E_S$ в некоторой другой точке $P^*$. Обозначим через $P_1$ точку, симметричную $P^*$ относительно прямой $y = 0$. Говорят, что точка $P_1$ получена *удвоением* точки $P_0$ (такая терминология связана с некоторой естественной *операцией сложения* точек эллиптической кривой).

Пусть $P_1 = (x_1, y_1)$. Простые, но несколько громоздкие вычисления дают

$$x_1 = \frac{(x_0^2 + 1)^2}{4x_0(x_0^2 - 1)}\,, \quad y_1 = \frac{x_0^6 - 5x_0^4 - 5x_0^2 + 1}{8Sx_0y_0(x_0^2 - 1)}\,.$$

Поскольку $y_1 \neq 0$, ибо многочлен

$$x^6 - 5x^4 - 5x^2 + 1 = (x^2 + 1)(x^2 - 2x - 1)(x^2 + 2x - 1)$$

не имеет рациональных корней, из точки $P_1$ удвоением можно получить точку $P_2 = (x_2, y_2)$, из точки $P_2$ аналогичным образом получить точку $P_3 = (x_3, y_3)$ и т. д. Покажем, что все точки $P_k$ будут попарно различны.

Пусть $x_k = a_k/b_k$ — запись в виде несократимой дроби. Тогда, как легко видеть,

$$b_1 = \frac{4|a_0(a_0^2 - b_0^2)|}{d_0}\, b_0, \quad d_0 = \text{НОД}\,((a_0^2 + b_0^2)^2, 4a_0b_0(a_0^2 - b_0^2)).$$

Нетрудно проверить, что $d_0 \in \{1, 4\}$. Поскольку $|a_0(a_0^2 - b_0^2)| > 1$, имеем $b_1 > b_0$. Аналогично $b_2 > b_1$ и т. д. Таким образом, знаменатели $b_k$ рациональных чисел $x_k$ монотонно возрастают, а значит, все $x_k$ и тем более $P_k$ попарно различны.

Кривая $E_S$ примечательна тем, что её рациональным точкам, отличным от $(0, 0)$ и $(\pm 1, 0)$, соответствуют прямоугольные треугольники площади $S$ с рациональными длинами сторон. Доказанную теорему можно сформулировать так: *если существует хотя бы один такой треугольник, то их существует бесконечно много* (утверждение, которое впервые высказал П. Ферма). Однако вопрос существования для данного $S$ хотя бы одного треугольника с требуемым свойством оказывается весьма нетривиальным (подробнее об этом см., например, в брошюре [5]).

II. Второй способ основан на следующем соображении. Пусть $p$ — простой делитель $b_0$. Тогда для $k = 0, 1, \ldots$ имеем

$$\nu_p(b_k) = \begin{cases} 2^k \nu_p(b_0), & \text{если } p > 2, \\ 2^k(\nu_2(b_0) - 1) + 1, & \text{если } p = 2, \end{cases}$$

где $\nu_p(m)$ — *$p$-адический показатель*, определяемый аналогично 2-адическому показателю (см. этюд IV). Доказательство проводится по индукции с использованием рекуррентного соотношения (‡). В частном случае это рассуждение можно найти, например, в опубликованном решении следующей фольклорной задачи: *доказать иррациональность градусной меры угла $\varphi$ при условии* $\cos\varphi = 1/3$ (см. [3]).

Теперь можно дать ещё одно доказательство теоремы из этюда III.

Пусть $x_0 = \cos\varphi \in I_{\mathbb{Q}} \setminus \{0, \pm 1/2, \pm 1\}$. Тогда

$$x_k = \cos\,(2^k\varphi), \quad k = 0, 1, 2, \ldots,$$

и по предложению в последовательности $\{x_k\}$ не должно быть совпадений. Но так не бывает, если $\varphi/\pi \in \mathbb{Q}$.

Некоторые элементарные свойства последовательности $\{x_k\}$ рассмотрены в статье [2].

## Этюд VII. Трансцендентные числа и мера иррациональности

Утверждение об иррациональности числа

$$\beta_0 = \frac{\operatorname{arctg}(4/3)}{\pi}$$

означает, что это число не может быть корнем многочлена 1-й степени с целыми коэффициентами. Но на самом деле оно не является корнем вообще никакого многочлена с целыми коэффициентами. Иными словами, это число — *трансцендентное*. Этот факт вытекает из следующей теоремы.

Теорема Гельфонда (1934). Если $\alpha$, $\beta$ — алгебраические числа, причём $\alpha \notin \{0, 1\}$, $\beta \notin \mathbb{Q}$, то число $\alpha^\beta$ трансцендентно.

Доказательство этой весьма нетривиальной теоремы, решающей 7-*ю проблему Гильберта*, читатель может найти в книге [7]. В нашем случае имеем

$$(-1)^{\beta_0} = \exp(\beta_0 \log(-1)) = \exp(\beta_0 \pi \sqrt{-1}) = \frac{3 + 4\sqrt{-1}}{5},$$

а последнее число, очевидно, алгебраическое.

Иррациональное число $\beta_0$ является вещественным, а для таких чисел представляет интерес вопрос о том, насколько хорошо они приближаются рациональными дробями.

*Мерой иррациональности* $\mu(\beta)$ числа $\beta \in \mathbb{R} \setminus \mathbb{Q}$ называют нижнюю грань таких чисел $\mu > 0$, что неравенство

$$\left| \beta - \frac{p}{q} \right| < \frac{1}{q^\mu} \qquad (\S)$$

имеет лишь конечное множество решений $(p, q) \in \mathbb{Z}^2$. Если чисел $\mu$ с указанным свойством нет, то полагают $\mu(\beta) = \infty$ (такие числа $\beta$ существуют и называются *числами Лиувилля*). Опираясь на *принцип Дирихле*, легко показать, что при $\mu = 2$ неравенство (§) имеет бесконечно много решений, поэтому всегда $\mu(\beta) \geqslant 2$.

Теорема Рота (1955). Если $\beta \in \mathbb{R} \setminus \mathbb{Q}$ — алгебраическое число, то для любого $\mu > 2$ неравенству (§) удовлетворяет конечное множество пар $(p, q) \in \mathbb{Z}^2$.

Итак, $\mu(\beta) = 2$ для любого алгебраического числа $\beta \in \mathbb{R} \setminus \mathbb{Q}$. Доказательство теоремы Рота также очень сложно и к тому же неэффективно, поскольку не позволяет указать явной оценки $q \leqslant q_0 = q_0(\beta, \mu)$ для возможных решений $(p, q)$ неравенства (§) (см., например, [4]).

Вместе с тем для любой вещественной квадратичной иррациональности $\beta$ такую оценку выписать вполне можно. Это связано с тем, что можно

предъявить в явном виде такую константу $c = c(\beta) > 0$, что неравенство

$$\left|\beta - \frac{p}{q}\right| \geqslant \frac{c(\beta)}{q^2}$$

будет верно для любых $(p, q) \in \mathbb{Z}^2$. Так, например, для $\beta = \sqrt{2}$ годится $c = 1/4$.

Для меры иррациональности $\mu(\beta)$ трансцендентных чисел $\beta \in \mathbb{R}$, не являющихся числами Лиувилля, обычно известны только верхние оценки (некоторые конкретные результаты читатель сможет найти по ссылке [10]).

Про число $\beta_0$ известно, что оно не число Лиувилля, при этом можно указать явную оценку $\mu(\beta_0) \leqslant a_0$ его меры иррациональности. Такого рода факты вытекают из степенной оценки для *линейной формы от логарифмов алгебраических чисел*, которую впервые получил Фельдман в 1968 году (см. [7]).

В частности, для числа

$$\beta_0 = \frac{\log(\alpha_1)}{\log(\alpha_2)}, \quad \alpha_1 = \frac{3 + 4\sqrt{-1}}{5}, \quad \alpha_2 = -1$$

речь идёт о линейной форме $\Lambda = q_1 \log(\alpha_1) + q_2 \log(\alpha_2)$, для которой справедлива оценка типа

$$|\Lambda| > L^{-b_0},$$

где $L = \max\{|q_1|, |q_2|, 2\}$, при этом константа $b_0$, зависящая только от $\alpha_1$, $\alpha_2$, а также от выбранных значений $\log(\alpha_1)$, $\log(\alpha_2)$, может быть явно вычислена. Для получения этого результата Фельдман усовершенствовал метод оценки линейной формы от логарифмов алгебраических чисел, предложенный в 1966 году Бейкером (см. [8]).

Однако константа $b_0$ (и, следовательно, константа $a_0$) оказывается очень большой, поэтому для практических приложений выгоднее применять оценку типа

$$|\Lambda| > \exp(-c_0 \log^2 L),$$

худшую по порядку, но с относительно небольшой константой $c_0$, также зависящей только от $\alpha_1$, $\alpha_2$, $\log(\alpha_1)$, $\log(\alpha_2)$. Пример такой оценки читатель сможет найти в работе [9].

## Список литературы

[1] Винберг Э. Б. *Курс алгебры*. М.: Факториал Пресс. 2001.

[2] Иванов О.А. *Современная математика в школьных задачах* // Соросовский образовательный журнал. № 6. 2000. С. 110–116.

[3] Канель-Белов А. Я. *Решение задачи* 12.1 // Математическое просвещение. Сер. 3. Вып. 15. 2011. С. 236–237.

[4] Касселс Дж. *Введение в теорию диофантовых приближений*. М.: Мир. 1961.

[5] Острик В. В., Цфасман М. А. *Алгебраическая геометрия и теория чисел*: *рациональные и эллиптические кривые*. М.: МЦНМО. 2001.

[6] Прасолов В. В. *Многочлены*. М.: МЦНМО. 2001.

[7] Фельдман Н.И. *Седьмая проблема Гильберта*. М.: МГУ. 1982.

[8] Baker A. *Transcendental Number Theory*. Cambridge: Cambridge Univ. Press. 1975.

[9] Laurent M., Mignotte M., Nesterenko Y. *Formes linéaires en deux logarithmes et déterminants d'interpolation* // J. Number Theory. Vol. 55. 1995. P. 285–321.

[10] `http://mathworld.wolfram.com/IrrationalityMeasure.html`

Н. Н. Осипов
Email: `nnosipov@rambler.ru`

# Об одном функциональном уравнении

Н. Николов Б. Станков

В статье обсуждается решение задачи 16.6 из задачника «Математического просвещения»

В этой статье мы исследуем функциональное уравнение

$$f(f(x)) = g(x), \tag{1}$$

на функцию $f\colon \mathbb{N} \to \mathbb{N}$,[1)] где $g\colon \mathbb{N} \to \mathbb{N}$ — данная инъективная функция.[2)]

Функции $g$ удобно сопоставить ориентированный граф $\Gamma(g)$, вершинами которого служат элементы $\mathbb{N}$, а ребра ведут от $x$ к $g(x)$ для всех $x \in \mathbb{N}$. Инъективность $g$ влечет следующую структуру графа: все вершины $\Gamma(g)$ разбиваются на непересекающиеся множества $K_i$, где каждое $K_i$ имеет один из следующих трех типов:

i) бесконечная *цепочка*: $x \to g(x) \to g^2(x) \to \ldots$[3)], в которой $g^{-1}(x) = \varnothing$, и $g^t(x) \neq g^s(x)$ при $s \neq t$;

ii) *цикл длины* $\infty$: $\ldots \to g^{-2}(x) \to g^{-1}(x) \to x \to g(x) \to g^2(x) \to \ldots$, где $g^t(x) \neq g^s(x)$ при $s \neq t$;

iii) *цикл длины* $k$ ($k \in \{1, 2, 3, \ldots\}$): $x \to g(x) \to g^2(x) \to \ldots \to g^k(x) = x$, где все элементы $x, g(x), g^2(x), \ldots, g^{k-1}(x)$ попарно различны.

Будем говорить, что $g$ — функция *типа* $(a; b; c_1, c_2, \ldots)$, если $\Gamma(g)$ разбивается на $a$ цепочек, $b$ циклов длины $\infty$ и $c_k$ циклов длины $k$, $k = 1, 2, \ldots$ (здесь $a, b, c_k$ могут независимо друг от друга принимать значения $0, 1, 2, \ldots, \infty$). Справедлива следующая теорема, в которой дается ответ (в терминах типа функции $g$) на вопрос о разрешимости уравнения (1), причем из доказательства ясно описание множества решений.[4)]

---

[1)] $\mathbb{N}$ можно заменить на произвольное счетное множество, так как арифметические операции здесь не используются.

[2)] Напомним, что $g$ называется инъективной, если из $g(x) = g(y)$ вытекает $x = y$.

[3)] Для натурального $k$ мы полагаем $g^k(x) = g(g^{k-1}(x))$. Кроме того, $g^0(x) = x$, $g^{-1}(x) = \{y \mid g(y) = x\}$, $g^{-k}(x) = g^{-1}(g^{-k+1}(x))$.

[4)] Метод доказательства применим для описания решений уравнения $f^k = g$, где $k$ — натуральное, а $g$ — данная инъективная функция.

Теорема. *Пусть* $g\colon \mathbb{N} \to \mathbb{N}$ — *функция типа* $(a; b; c_1, c_2, \ldots)$.

*а) Уравнение* (1) *неразрешимо тогда и только тогда, когда в множестве* $\{a, b, c_{2k} \mid k = 1, 2, \ldots\}$ *есть хотя бы одно нечетное число.*

*б) Уравнение* (1) *разрешимо и имеет конечное множество решений тогда и только тогда, когда выполнены следующие условия*: *в множестве* $\{a, c_{2k} \mid k = 1, 2, \ldots\}$ *все числа четны,* $b = 0$, *и сумма*

$$\sum_{k=1}^{\infty}(c_{2k-1}(c_{2k-1} - 1) + c_{2k})$$

*конечна.*

Отметим, что частный случай нашей задачи для функции $g(x) = x + 1987$, предлагался на Международной математической олимпиаде 1987 г. Как видим, эта функция имеет тип $(1987; 0; 0, 0, \ldots)$, и, согласно нашей теореме, для такой функции $g$ уравнение (1) не имеет решений.

## Доказательство теоремы

Пусть $N$ — это множество всех вершин $\Gamma(g)$, принадлежащих циклам (конечным и бесконечным), а $M$ — это множество всех вершин $\Gamma(g)$, принадлежащих бесконечным цепочкам. (Формально $N = \bigcap\limits_{i=0}^{\infty} g^i(\mathbb{N})$, $M = \mathbb{N} \setminus N$.) Ясно, что $g(N) = N$ и $g(M) \subset M$. Обозначим $M_i = g^i(M)$ $(i = 0, 1, 2, \ldots)$ и $S_i = M_{i+1} \setminus M_i$, так что $S_0$ — множество начальных элементов всех бесконечных цепочек, $S_1 = g(S_0)$, $S_2 = g(S_1)$, $\ldots$ .

Пусть функция $f$ удовлетворяет (1). Так как $g$ инъективна, то и $f$ инъективна. Отметим также, что

$$f(g(n)) = g(f(n)) \text{ для всех } n \in \mathbb{N}, \tag{2}$$

так как левая и правая части (2) равны $f(f(f(n)))$.

Покажем, что $f(M) \subset M$ и $f(N) \subset N$. Достаточно понять, что множества $K = \{k \in M \mid f(k) \in N\}$ и $L = \{l \in N \mid f(l) \in M\}$ пустые. Предположим $K$ непусто. Найдем минимальный номер $i$ такой что $K \cap S_i$ непусто, и возьмем $k \in K \cap S_i$. Так как $f(k) \in N$, то $f(k) = g(x)$ для некоторого $x \in N$. Из равенства $f(k) = f(f(x))$ и инъективности $f$ получаем $k = f(x)$. Для элемента $x$ найдем $y \in N$ такой что $g(y) = x$. Пусть $z = f(x)$. Имеем $k = f(x) = f(g(y)) = g(f(y))$. Если $i = 0$, то есть $k \in S_0$, получили противоречие, так как $g^{-1}(k) = \varnothing$. Иначе $f(y) \in S_{i-1}$. Кроме того $f(f(y)) = g(y) = x \in N$, поэтому $f(y) \in K \cap S_{i-1}$ в противоречие с выбором $i$.

Предположим, что $L$ непусто. Возьмем $z \in L$. По определению $f(z) \in M$. С другой стороны $f(f(z)) = g(z) \in N$. Отсюда $f(z) \in K$, и все сводится к предыдущему рассмотрению.

Итак, мы разделили нашу задачу на две независимых подзадачи для множеств $M$ и $N$.

1. Решим задачу для $M$ (можно считать, что $N = \varnothing$). Положим

$$A = \{a \in S_0 \mid f(a) \in S_0\}, \quad B = \{b \in S_0 \mid f(b) \in M \setminus S_0\}.$$

Очевидно $A \cup B = S_0$ и $A \cap B = \varnothing$.

Заметим, что если $x \in A$, то $f(x) \in B$. Действительно, предположим противное: $f(x) \in A$. Тогда $f(f(x)) \in S_0$ по определению $A$. Но с другой стороны $f(f(x)) = g(x) \in S_1$ — противоречие.

Покажем, что $f(B) \cap M_2 = \varnothing$. Предположим, что, напротив, существует $y \in B$ такой что $f(y) = g(g(x))$ для некоторого $x \in M$. Из (2) и (1) следует, что $g(y) = f(f(y)) = f(g(g(x))) = g(g(f(x)))$. Так как $g$ инъективна, то $y = g(f(x))$, что противоречит включению $y \in S_0$. Итак, для каждого $y \in B$ имеем $f(y) \in S_1$, то есть $f(y) = g(x)$ для некоторого $x \in S_0$. С другой стороны, $g(x) = f(f(x))$, и в силу инъективности $f$ получаем $y = f(x)$. Так как $y \in S_0$, то $x \in A$. Получаем, что $f(A) = B$, в частности, если $a = |S_0|$ конечно ($|S_0|$ равно количеству бесконечных цепочек), то оно четно.

С другой стороны, при любом разбиении множества $S_0$ на две части $A$ и $B$ такие, что $|A| = |B|$ любая биекция $h\colon A \to B$ продолжается единственным образом до искомой функции $f\colon M \to M$ следующим образом. Пусть $x \in S_i$, тогда

$$\begin{cases} f(x) = g^i(h(g^{-i}(x))), & \text{если } g^{-i}(x) \in A, \\ f(x) = g^{i+1}(h^{-1}(g^{-i}(x))), & \text{если } g^{-i}(x) \in B. \end{cases}$$

В случае четного $a$ это дает $\binom{a}{a/2} \cdot (a/2)!$ искомых функций.

2. Теперь решим задачу для $N$.[5)] Для каждого $x \in N$ пусть $G(x) = \{g^i(x) \mid i \in \mathbb{Z}\}$ — цикл (орбита) элемента $x$ (циклы двух элементов либо не пересекаются, либо совпадают).

Заметим, что для всякого $x \in N$ верно $f(G(x)) = G(f(x))$. Это следует из того, что $f(g^j(x)) = g^j(f(x))$ для всех $j \in \mathbb{Z}$. Отсюда получаем, что $f(G(f(x))) = f(f(G(x))) = g(G(x)) = G(x)$. Значит, для каждого цикла имеется две возможности: либо $f(G) = G$, либо найдется другой цикл $\overline{G}$ такой, что $f(G) = \overline{G}$ и $f(\overline{G}) = G$. Зафиксируем $l \in \{1, 2, \ldots, \infty\}$ и рассмотрим цикл $G$ длины $|G| = l$.

а) Рассмотрим случай $f(G) = G$. Возьмем $x \in G$. Так как $f(x) \in G$, имеем $f(x) = g^j(x)$ для некоторого $j \in \mathbb{Z}$. Для каждого $y \in G$, $y = g^k(x)$

[5)] Фактически это задача решения уравнения $x^2 = g$ в группе $S_\infty$ биекций счетного множества.

имеем $f(y) = f(g^k(x)) = g^k(f(x)) = g^k(g^j(x)) = g^j(g^k(x)) = g^j(y)$. Имеем $g(x) = f(f(x)) = g^j(g^j(x)) = g^{2j}(x)$ или $g^{2j-1}(x) = x$. Это означает, что $l$ конечно и $2j-1$ делится на $l$. Отсюда $l$ нечетно: $l = 2k-1$. Поэтому $j \equiv k \pmod{l}$, и значит сужение $f$ на $G$ однозначно определяется равенством

$$f(x) = g^k(x). \tag{3}$$

б) Рассмотрим второй случай: $f(G) = \overline{G}$ и $f(\overline{G}) = G$, $G \neq \overline{G}$. В этом случае сужение $f$ на $G$ — это биекция между $G$ и $\overline{G}$ ($f(G) = \overline{G}$ и $f$ инъективна). Значит, $|G| = |\overline{G}| = l$. Зафиксируем $x \in G$. Пусть $f(x) = \overline{x} \in \overline{G}$. Тогда однозначно

$$\begin{cases} f(y) = g^k(\overline{x}), & \text{если } y = g^k(x) \in G, \\ f(y) = g^{k+1}(x), & \text{если } y = g^k(\overline{x}) \in \overline{G}. \end{cases} \tag{4}$$

Наоборот, для произвольного $\overline{x} \in \overline{G}$ сужение на $G \cup \overline{G}$ такой функции $f$, что $f(x) = \overline{x}$, однозначно определяется формулами (4).

Итак, если $l$ бесконечно или четно, случай а) невозможен, поэтому количество циклов длины $l$ либо бесконечно, либо четно. При этом каждое разбиение циклов длины $l$ на пары дает $l$ способов определить $f$ на $G \bigcup \overline{G}$ по формулам (4) (итого $\dfrac{c_l! \cdot l}{(c_l/2)! \cdot 2^{c_l/2}}$ способов в том случае, когда $l$ и $c_l$ конечны).

Если $l$ нечетно, множество циклов длины $l$ разбиваются на множества $S$ и $T$, ($|T|$ четно или бесконечно). Тогда $f$ определяется на каждом цикле из $S$ единственным способом как показано в (3). Циклы множества $T$ произвольно разбиваются на пары, и для каждой пары $f$ определяется одним из $l$ способов как показано в (4).

Объединяя полученные результаты, получим утверждение теоремы.

Авторы благодарят П. Кожевникова за помощь в подготовке текста этой заметки на русском языке.

---

Н. Николов, Институт математики и информатики, Болгарская академия наук, ул. «Акад. Г. Бончев», бл. 8, 113 София
email: nik@math.bas.bg
Б.Станков, Lyceé Louis Le Grand, 123 rue Saint Jacques, 75005 Paris
email: thrall.warlord@gmail.com

# Задачный раздел

В этом разделе вниманию читателей предлагается подборка задач разной степени сложности, в основном трудных. Составителям этой подборки кажется, что предлагаемые ниже задачи окажутся интересными как для сильных школьников, интересующихся математикой, так и для студентов-математиков.

Мы обращаемся с просьбой ко всем читателям, имеющим свои собственные подборки таких задач, присылать их в редакцию. И, разумеется, мы с удовольствием будем публиковать свежие авторские задачи.

В скобках после условия задачи приводится фамилия автора (уточнения со стороны читателей приветствуются).

1. Можно ли в куб достаточно большой размерности с ребром 1 см вложить здание МГУ? (*Ф. Ивлев*)

2. а) Найти 300-ю цифру после запятой числа $\sqrt[3]{0.\underbrace{99\ldots9}_{100\text{ штук}}}$.

   б) С помощью калькулятора найти первую цифру числа $2^{10^6}$.

   (*А. Я. Белов*)

3. На плоскости дано множество $M$, площадь которого меньше 1, и $n$ точек. Доказать, что множество $M$ можно сдвинуть на вектор, длина которого меньше $\sqrt{n/\pi}$, где $\pi = 3,14159\ldots$, так, что множество, полученное в результате сдвига, не будет покрывать ни одной из данных $n$ точек. (*В. А. Сендеров*)

   б) (Задача на исследование) Постарайтесь получить оценки для $n$-мерного пространства.

4. $\mathcal{A}$ — отображение плоскости в себя, сохраняющее расстояние (т. е. $|XY| = |\mathcal{A}(X)\mathcal{A}(Y)|$ для любых точек $X$, $Y$ плоскости). Доказать, что $\mathcal{A}$ — отображение плоскости на себя (т. е. каждая точка имеет прообраз при этом отображении).

5. На плоскости нарисованы две а) пересекающиеся б) непересекающиеся окружности. Можно ли одной линейкой построить их центры?

6. Если целые $m$ и $n$ взаимно просты, а числа $x^n + x^{-n}, x^m + x^{-m}$ — целые, то $x + 1/x$ — тоже целое число ($x \in \mathbb{C}$).

7. На каждом ребре правильного многогранника $M$ с единичными ребрами взяли по точке $A_i$. Найти объем геометрического места центров масс таких наборов. Рассмотреть все 5 возможностей.

(*А. Я. Канель*)

8. Слова $u$ и $v$ *циклически сопряжены*, если $u = s_1 s_2, v = s_2 s_1$ для некоторых слов $s_1, s_2$. Слово $u$ называется *правильным*, если оно больше любого своего лексикографически сопряженного. а) Докажите, что в любом правильном слове $u$ можно так однозначно расставить лиевы скобки $[\cdot,\cdot]$, что при их раскрытии ($[st]$ раскрывается как $st - ts$) слово $u$ будет старшим членом получившегося (некоммутативного) многочлена.

   б) Докажите, что достаточно длинное слово содержит подслово вида $UXU$, где $U$, $X$ — правильные слова.

(*D. Bakelin, В. А. Уфнаровский*)

9. Имеется $2^n - 1$ коробок. В коробке первой величины содержатся две коробки второй величины. В каждой из $2^{k-1}$ коробок $k$-ой величины содержатся по две коробки $(k+1)$-ой величины. В коробках последней $n$-ой величины лежит по одной монете. За один ход разрешается в одной из коробок любой величины перевернуть все монеты. Доказать, что за $[n/2] + 1$ ходов можно уравнять число монет, лежащих орлом вверх и орлом вниз. Можно ли улучшить эту оценку?

(*А. Я. Белов*)

10. Дано векторное пространство $W$, $\dim(W) = m$, два его подпространства $U$ и $V$, такие что $U \cap V = 0$ ($\dim(u) = n_1$, $\dim(v) = n_2$) и обратимый оператор $A\colon W \to W$. Докажите, что $A^n(U) \cap V = 0$ при некотором $n \leqslant \min(\binom{m}{n_1}, \binom{m}{n_2})$.

11. Существует ли граф с хроматическим числом, большим 2013, все циклы которого имеют длину больше 2013? (*Хроматическое число графа* есть минимальное число цветов, в которые его можно правильно раскрасить.)

12. (Задача на исследование). а) Дан многочлен $P(x, y)$ степени $n$ такой, что $P(x, y) \geqslant 0$ при всех $x, y$. При этом $P(x, y) = 0$ только если $x = y = 0$. Верно ли, что для некоторой константы $C > 0$ выполняется неравенство $P(x, y) > C \cdot (|x| + |y|)^n$? б) Для каких натуральных $m$ можно утверждать что для некоторой константы $C > 0$ выполняется неравенство $P(x, y) > C \cdot (|x| + |y|)^m$ (при всех $x, y \in [-1, 1]$)? (*И. И. Богданов, Г. Р. Челноков*)

# Решения задач из предыдущих выпусков

6.11. Условие. Рассматриваются слова из букв русского алфавита. Слова вида $SUT$ и $SUUT$ имеют одинаковый смысл (здесь $S$, $U$, $T$ — произвольные слова, возможно, пустые). Докажите, что количество различных смыслов конечно.

Решение. Назовём слова $A$ и $B$ *эквивалентными* (обозначение: $A \sim B$), если одно можно получить из другого цепочкой замен вида $U \leftrightarrow UU$. Обозначим через $|A|$ длину слова $A$.

Докажем сначала следующие две леммы.

Лемма 1. *Пусть $A$ — произвольное слово, содержащее все буквы алфавита $X$, а $B$ — произвольное слово в этом алфавите. Тогда $A \sim ABC$ для некоторого слова $C$.*

Доказательство. Индукция по длине $|B|$ слова $B$. Если $|B| = 0$, утверждение тривиально. Пусть теперь $|B| > 0$, и $b$ — первая буква слова $B$, то есть $B = bB'$.

Буква $b$ встречается в $A$, то есть $A = UbV$ для некоторых слов $U$, $V$; тогда $A \sim UbVbV = AbV$. Применяя предположение индукции к словам $A' = Ab$ и $B'$, находим слово $C'$ такое, что $A' \sim A'B'C' = ABC'$; тогда $A \sim A'V \sim ABC'V$, и можно положить $C = C'V$. Лемма доказана. □

Лемма 2. *Существует такое число $f(n)$, что любое слово в $n$-буквенном алфавите длины, не меньшей $f(n)$, эквивалентно слову меньшей длины.*

Доказательство. Индукция по $n$. База индукции при $n = 1$ тривиальна: можно положить $f(1) = 2$. Для шага индукции докажем, что можно положить $f(n) = (n^{f(n-1)} + 1)f(n-1)$. Действительно, рассмотрим слово $W$ длины, не меньшей $f(n)$, и выделим в нём $n^{f(n-1)}+1$ непересекающихся подслов длины $f(n-1)$. По принципу Дирихле, два из этих подслов будут одинаковыми, то есть $W = UABAV$, где $A$ — слово длины $f(n-1)$. Если слово $A$ содержит не все $n$ букв нашего алфавита, то по предположению индукции $A \sim A'$, где $|A'| < f(n-1)$; тогда $W \sim UA'BA'V$, и утверждение доказано.

Пусть теперь $A$ содержит все буквы алфавита. Тогда, применяя лемму, получаем, что $A \sim ABC$ для некоторого слова $C$. Отсюда имеем

$ABA \sim ABABC \sim ABC \sim A$, то есть $W = UABAV \sim UAV$, и $|UAV| < |W|$, что и требовалось доказать. □

Теперь легко доказать утверждение задачи. Рассмотрим произвольное слово в $n$-буквенном алфавите; будем применять к нему лемму 2, пока его длина не меньше $f(n)$. В итоге получим, что наше слово эквивалентно слову длины, меньшей $f(n)$, то есть одному из конечного набора слов.

(*И. И. Богданов*)

11.3. Условие. $A_1, \ldots, A_n$ и $B_1, \ldots, B_n$ — два разбиения единичного квадрата на непересекающиеся измеримые множества. $S_{ij}$ — пересечение множеств $A_i$ и $B_j$, $|G|$ — площадь множества $G$. Докажите неравенство:

$$\sum_{ij} |S_{ij}| \cdot \ln(|S_{ij}|) \geqslant \sum_{i} |A_i| \cdot \ln(|A_i|) + \sum_{j} |B_j| \cdot \ln(|B_j|).$$

Решение. Неравенство задачи — это (с точностью до знака) стандартный факт из теории информации: энтропия совместного распределения не превосходит суммы энтропий. Приведём стандартное доказательство этого факта.

Обозначим $|S_{ij}| = s_{ij}$, $a_i = |A_i| = \sum_j s_{ij}$, $b_j = |B_j| = \sum_i s_{ij}$. Запишем разность

$$\sum_{ij} |S_{ij}| \cdot \ln(|S_{ij}|) - \sum_{i} |A_i| \cdot \ln(|A_i|)$$

в виде

$$\sum_{i,j} s_{ij} \ln s_{ij} - \sum_{i,j} s_{ij} \ln a_i = \sum_{i,j} a_i \frac{s_{ij}}{a_i} \ln \frac{s_{ij}}{a_i}\,. \qquad (*)$$

Заметим, что функция $f(x) = x \ln x$ выпукла при $x > 0$ (вторая производная положительна). Из неравенства Йенсена получаем при каждом $j$ следующее неравенство

$$\sum_{i} a_i \frac{s_{ij}}{a_i} \ln \frac{s_{ij}}{a_i} \geqslant \left( \sum_{i} a_i \frac{s_{ij}}{a_i} \right) \ln \left( \sum_{i} a_i \frac{s_{ij}}{a_i} \right) = b_j \ln b_j\,.$$

Отсюда оцениваем правую часть $(*)$ снизу как

$$\sum_{j} b_j \ln b_j\,,$$

откуда и следует неравенство задачи. (*М. Н. Вялый*)

11.10. Условие. Пусть $A, B$ — целочисленные матрицы. Известно, что $\det(A) = 1$, $\det(B) \neq 0$. Докажите, что существует $n \in \mathbb{N}$ такое, что $B^{-1}A^nB$ — целочисленная матрица.

Решение. Поскольку $\det(A) = 1$, матрица $A$ обратима по любому целому модулю.

Обозначим $\Delta = \det B$. Тогда $B^{-1} = \Delta^{-1}B'$, где $B$ — целочисленная матрица. С другой стороны, при некотором $n$ выполняется $A^n = 1 \bmod \Delta$. Это означает, что $A^n = I + \Delta A_n$, где $I$ — единичная, а $A_n$ — целочисленная матрица. Но тогда матрица

$$B^{-1}A^nB = \Delta^{-1}B'(I + \Delta A_n)B = I + B'A_nB$$

также целочисленная. (*М. Н. Вялый*)

12.3. Условие. Вершины $A$ и $B$ графа $G$ назовём *эквивалентными*, если существует такая последовательность вершин $A = A_0, A_1, \ldots, A_n = B$, что любые две соседние вершины $A_i$ и $A_{i+1}$ можно соединить $k$ путями без общих промежуточных вершин. Докажите, что любые две эквивалентные вершины можно соединить $k$ путями без общих рёбер.

Решение. Утверждение задачи вытекает из следующего факта.

*Вершины $A$ и $B$ графа назовём* эквивалентными, *если существует такая последовательность вершин $A = A_0, A_1, \ldots, A_n = B$, что любые две соседние вершины $A_i$ и $A_{i+1}$ можно соединить $k$ путями без общих рёбер. Любые две эквивалентные вершины можно соединить $k$ путями без общих рёбер.*

*Доказательство факта.* Если удалить любые $k-1$ рёбер, то для любого $i$ вершины $A_i$ и $A_{i+1}$ окажутся в одной компоненте связности. Значит, вершины $A_0$ и $A_n$ также окажутся в одной компоненте связности. Остается применить рёберную теорему Менгера.

Примечание. Это решение изоморфно первому решению, приведённому в прошлом выпуске сборника «Математическое просвещение». Действительно, теорема Форда – Фалкерсона для единичных пропускных способностей есть просто рёберная теорема Менгера. (*Б. Мохар*)

15.10. Условие. На грани правильного тетраэдра отмечена точка. Докажите, что тетраэдр можно разрезать на четыре равных выпуклых многогранника так, чтобы эта точка была вершиной одного из них.

Решение. Пусть $A_1A_2A_3A_4$ — исходный тетраэдр, $O$ — его центр, а $X_1$ — данная точка на грани $A_2A_3A_4$. Обозначим через $\varphi_{12}$ отражение относительно прямой $\ell_{12}$, проходящей через середины рёбер $A_1A_2$ и $A_3A_4$, через $\varphi_{13}$ и $\varphi_{14}$ — два аналогичных отражения относительно прямых $\ell_{13}$ и $\ell_{14}$. Тогда $\varphi_{1i} \circ \varphi_{1j} = \varphi_{1k}$, если $\{i, j, k\} = \{1, 2, 3\}$. Таким образом, отражения $\varphi_{12}$, $\varphi_{13}$, $\varphi_{14}$ вместе с тождественным отображением образуют

подгруппу в группе $S_4$ всех самосовмещений тетраэдра (эта подгруппа называется *группой Клейна*).

Положим $X_2 = \varphi_{12}(X_1)$, $X_3 = \varphi_{13}(X_1)$, $X_4 = \varphi_{14}(X_1)$. Тогда точка $X_i$ лежит на грани $A_{i+1}A_{i+2}A_{i+3}$ (все индексы берутся по модулю 4, то есть, например, $A_5 = A_1$). Заметим, что четвёрка точек $(X_1, X_2, X_3, X_4)$ переходит в себя при любом отражении $\varphi_{1i}$. Поэтому и сумма векторов $\vec{s} = \overrightarrow{OX_1} + \overrightarrow{OX_2} + \overrightarrow{OX_3} + \overrightarrow{OX_4}$ переходит в себя при действии нашими отражениями; значит, $\vec{s} = 0$. Таким образом, либо точка $O$ лежит внутри тетраэдра $X_1X_2X_3X_4$, либо точки $X_1$, $X_2$, $X_3$, $X_4$ компланарны.

Если тетраэдр $X_1X_2X_3X_4$ невырожден, то всё пространство разбивается на четыре трёхгранных угла $C_1 = OX_2X_3X_4$, $C_2 = OX_1X_3X_4$, $C_3 = OX_1X_2X_4$, $C_4 = OX_1X_2X_3$. Наши отражения переставляют эти трёхгранные углы и переводят тетраэдр в себя. Значит, если мы обозначим через $T_i$ пересечение нашего тетраэдра с $C_i$, то $\varphi_{1i}(T_1) = T_i$, поэтому все многогранники $T_i$ равны. Кроме того, $T_i$ выпуклы как пересечения выпуклых множеств, и точка $X_1$ является вершиной $T_1$. Итого, наше разбиение построено.

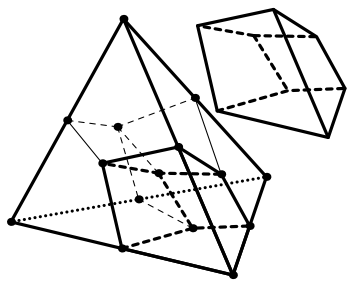


*Рис. 1*

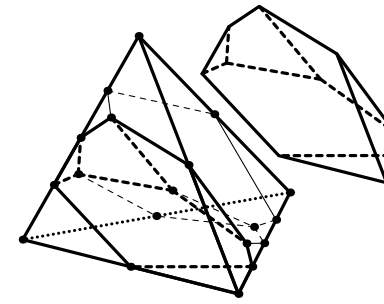


*Рис. 2*

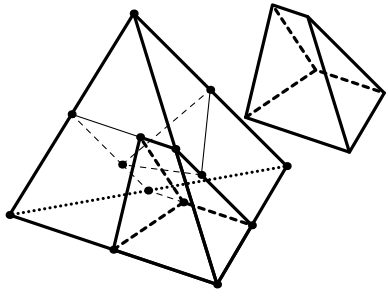


*Рис. 3*

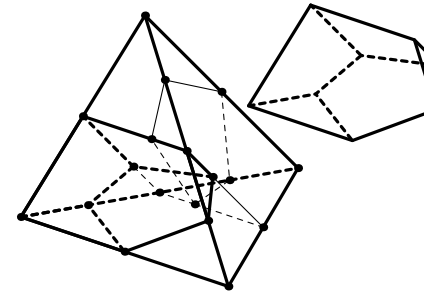


*Рис. 4*

Пусть, наконец, точки $X_1$, $X_2$, $X_3$, $X_4$ компланарны. Тогда четырёхугольник $X_1X_2X_3X_4$ имеет (в пространстве) три попарно перпендикулярных оси симметрии; значит, это — прямоугольник (возможно, вырожденный), его стороны параллельны двум из прямых $\ell_{1i}$ (скажем, $\ell_{12}$ и $\ell_{13}$) и пересекают их. Таким образом, эти четыре точки лежат в «серединной» плоскости, равноудалённой от рёбер $A_1A_4$ и $A_2A_3$. В этом случае построение может быть получено, например, предельным переходом из невырожденного случая.

На рис. 1 и 2 показаны два возможных невырожденных разрезания, полученных описанным способом (справа от каждого тетраэдра показан один из многогранников разбиения). На рис. 3 и 4 — «вырожденные» разрезания, полученные из них предельным переходом.

Замечание. Вместо группы Клейна можно бы было рассмотреть движение пространства, переставляющее все вершины тетраэдра циклически, и группу, порождённую этим движением. Аналогичное построение, естественно, тоже работает. На рис. 5 и 6 показаны два разрезания, полученные таким образом.

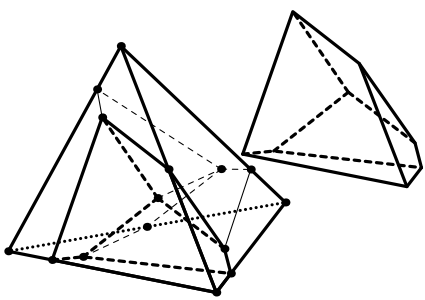


*Рис. 5*

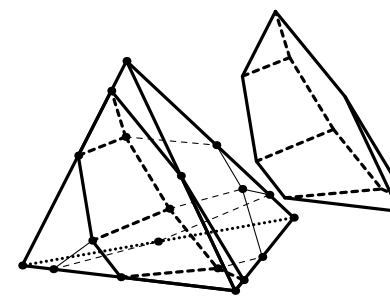


*Рис. 6*

Нетрудно видеть, что эта вариация допускает также обобщение на $n$-мерный случай.

(*И. И. Богданов*)

16.2. Условие. В Черноморске во время обсуждения вопроса о том, когда же наконец Черноморск объявят вольным городом, сложилась занятная ситуация. Все черноморцы разбились на партии, а все партии на фракции так, что: 1) существует партия, в которой объединились все $n$ жителей города; 2) каждая партия состояла ровно из двух непересекающихся фракций; 3) каждая фракция численностью более одного человека считала себя партией. Каждый житель города платит членский взнос (1 руб.) в каждой партии, членом которой является. Как им

надо было организоваться, чтобы сумма взносов была: а) максимальной; б) минимальной?

Решение. Пусть $\mathrm{Mx}(n)$ — максимально возможная сумма членских взносов, $\mathrm{Mn}(x)$ — минимально возможная. Пусть $k$ — оптимальное число изначального деления черноморцев на две партии, так чтобы сумма членских взносов была бы, скажем, минимальна. Тогда $\mathrm{Mn}(n) = \mathrm{Mn}(k) + \mathrm{Mn}(n-k) + n$, ибо внутри фракций черноморцы делятся оптимальным образом. Таким образом, имеет место равенство:

$$\mathrm{Mn}(n) = \min_k(\mathrm{Mn}(k) + \mathrm{Mn}(n-k)) + n. \tag{1}$$

Аналогичным образом, для задачи о максимума членских взносов имеет место равенство

$$\mathrm{Mn}(n) = \max_k(\mathrm{Mn}(k) + \mathrm{Mn}(n-k)) + n. \tag{2}$$

Из формулы (2) непосредственной индукцией легко проверить, что $\mathrm{Mx}(n) = n(n+1)/2 - 1$.

Равенство для минимума членских взносов несколько сложнее. Пусть $n = 2^k + l$; $0 \leqslant l \leqslant 2^k$. Тогда $\mathrm{Mn}(n) = kn + 2l$. Это равенство также с помощью индукции выводится из формулы (1). Соответственно, $\mathrm{Mn}(2047) = 11 \cdot 2047$.

Замечание. Можно рассматривать двоичное дерево, связанное с разбиением на партии/фракции и заметить, что если есть две висячие вершины, чьи уровни отличаются не менее чем на 2, то одну из них можно перенести к другой так, чтобы общая сумма членских взносов уменьшилась.

(*А. Я. Канель-Белов*)

16.11. Условие. При каких натуральных $n$ число $\dfrac{3^n - 1}{2}$ есть квадрат целого числа?

Ответ: $n \in \{1, 2, 5\}$.

Решение. Требуется решить уравнение

$$2x^2 + 1 = 3^n \tag{1}$$

в натуральных числах $n$ и $x$.

1-й способ. Этот способ решения опирается на теорию уравнений вида

$$X^2 - AY^2 = B, \tag{2}$$

где $A > 0$, $B \neq 0$ — целые числа и $\sqrt{A}$ иррационален. Частным случаем является *уравнение Пелля* с $B = 1$. Все решения уравнения (2) в целых числах $X$, $Y$ могут быть найдены из формул

$$X + Y\sqrt{A} = \pm(X_j + Y_j\sqrt{A})(X_0 + Y_0\sqrt{A})^k, \quad k \in \mathbb{Z},$$

где $(X_j, Y_j)$ — некоторые *базисные* решения, а $(X_0, Y_0)$ — *минимальное* решение ассоциированного уравнения Пелля в натуральных числах (подробности читатель может найти, например, в статье [2] или книге [3]).

Пусть сначала $n$ чётно, $n = 2m$. Положим $y = 3^m$. Уравнение Пелля

$$y^2 - 2x^2 = 1$$

в натуральных числах имеет единственную серию решений:

$$y + x\sqrt{2} = (3 + 2\sqrt{2})^k, \quad (k = 1, 2, \dots).$$

Если $m \geqslant 2$, то $y \equiv 0 \pmod 9$, а это имеет место тогда и только тогда, когда $k \equiv 3 \pmod 6$.

Действительно, последовательность чисел $(3 + 2\sqrt{2})^k$ является периодической по любому модулю (подумайте, почему). Заметим, что

$$(3 + 2\sqrt{2})^6 \equiv -1 \pmod 9.$$

Поэтому если $k = 6q + r$, где $r \in \{0, 1, \dots, 5\}$, то

$$(3 + 2\sqrt{2})^k \equiv (-1)^q (3 + 2\sqrt{2})^r \pmod 9.$$

Проверка показывает, что только при $r = 3$ рациональная часть числа $(3 + 2\sqrt{2})^r$ будет $\equiv 0 \pmod 9$.

Но для $k = 6q + 3$ имеем $y \equiv 0 \pmod{11}$, что невозможно, если $y = 3^m$.

В самом деле, поскольку $(3 + 2\sqrt{2})^6 \equiv -1 \pmod{11}$, имеем

$$(3 + 2\sqrt{2})^k \equiv (-1)^q (3 + 2\sqrt{2})^3 \equiv (-1)^q 4\sqrt{2} \pmod{11},$$

откуда и следует утверждение.

Итак, $m = 1$ — единственное возможное значение, откуда $(n, x) = (2, 2)$.

В случае нечётного $n = 2m + 1$ рассуждения аналогичны. Положим $z = 2x$, $y = 3^m$. Уравнение

$$z^2 - 6y^2 = -2$$

в натуральных числах также имеет единственную серию решений:

$$z + y\sqrt{6} = (2 + \sqrt{6})(5 + 2\sqrt{6})^k \quad (k = 0, 1, 2, \dots).$$

Далее можно проверить, что сравнение $y \equiv 0 \pmod{27}$ равносильно сравнению $k \equiv 4 \pmod 9$, однако для таких $k$ имеем $y \equiv 0 \pmod{17}$ — противоречие. Таким образом, $m \leqslant 2$, что даёт $(n, x) \in \{(1, 1), (5, 11)\}$.

2-й способ. Для чётных $n = 2m$ можно рассуждать совсем элементарно, предварительно переписав уравнение в виде

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

Положим $x = 2x_1$, тогда

$$\frac{3^m - 1}{2} \cdot \frac{3^m + 1}{2} = 2x_1^2,$$

при этом числа $(3^m \pm 1)/2$ взаимно просты. Возможны следующие два случая: либо

$$\frac{3^m-1}{2} = 2a^2, \quad \frac{3^m+1}{2} = b^2,$$

либо

$$\frac{3^m-1}{2} = a^2, \quad \frac{3^m+1}{2} = 2b^2$$

($a$, $b$ — некоторые натуральные числа). Но в первом случае равенство

$$3^m = 4a^2 + 1$$

невозможно, поскольку $4a^2 + 1$ не делится на 3. Во втором случае имеем

$$3^m = (2b-1)(2b+1),$$

откуда $2b - 1 = 1$ и $2b + 1 = 3^m$, т. е. $b = 1$, $m = 1$ и $x = 2$.

В случае нечётных $n = 2m + 1$ рассуждение менее элементарно. Мы можем воспользоваться тем, что кольцо

$$\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} : a,\, b \in \mathbb{Z}\}$$

*факториально*, т. е. для него справедлив аналог основной теоремы арифметики об однозначном разложении в произведение простых чисел. Проще всего это доказать, заметив, что кольцо $\mathbb{Z}[\sqrt{-2}]$ *евклидово*, т. е. в нём возможно *деление с остатком* (точные формулировки соответствующих определений и фактов см., например, в учебнике [1, с. 190–197]; проверка евклидовости кольца $\mathbb{Z}[\sqrt{-2}]$ относительно нормы $N(a+b\sqrt{-2}) = a^2 + 2b^2$ оставляется читателю в качестве несложного упражнения).

Перепишем уравнение (1) в виде

$$(1 + x\sqrt{-2})(1 - x\sqrt{-2}) = (1 + \sqrt{-2})^{2m+1}(1 - \sqrt{-2})^{2m+1}. \tag{3}$$

Заметим, что числа $1 \pm x\sqrt{-2}$ взаимно просты в кольце $\mathbb{Z}[\sqrt{-2}]$, а числа $1 \pm \sqrt{-2}$ являются простыми в этом кольце. Поэтому из равенства (3) и свойства факториальности следует, что

$$1 + x\sqrt{-2} = \pm(1 \pm \sqrt{-2})^{2m+1}$$

при некотором выборе знаков. Раскрыв бином и приравняв вещественные части, получим равенство

$$1 - C^2_{2m+1}2 + C^4_{2m+1}2^2 - \ldots + (-1)^m C^{2m}_{2m+1}2^m = \pm 1. \tag{4}$$

При знаке «минус» в случае $m \geqslant 2$ это равенство можно переписать в виде

$$1 - C^2_{2m+1} + C^4_{2m+1}2 - \ldots + (-1)^m C^{2m}_{2m+1}2^{m-1} = 0.$$

Но такое равенство невозможно, поскольку при любом $m$ число

$$1 - C^2_{2m+1} + C^4_{2m+1}2 = \frac{(2m-1)(2m^3 - m^2 - 4m - 3)}{3}$$

не делится на 4.

Пусть теперь в равенстве (4) взят знак «плюс». Тогда при $m \geqslant 5$ оно примет вид

$$C_{2m+1}^2 - C_{2m+1}^4 2 + C_{2m+1}^6 2^2 - C_{2m+1}^8 2^3 + \ldots - (-1)^m C_{2m+1}^{2m} 2^{m-1} = 0. \quad (5)$$

Ясно, что $m$ должно быть чётным. Положим

$$A = C_{2m+1}^2 - C_{2m+1}^4 2 + C_{2m+1}^6 2^2 - C_{2m+1}^8 2^3 = \\ = -\frac{m(m-2)(2m+1)(8m^5 - 68m^4 + 158m^3 - 139m^2 + 254m + 102)}{315}.$$

Для любого целого числа $P \neq 0$ пусть $\nu_2(P)$ обозначает такое целое неотрицательное число $\alpha$, что $P$ делится на $2^\alpha$, но не делится на $2^{\alpha+1}$. Для любого рационального числа $P/Q$ положим $\nu_2(P/Q) = \nu_2(P) - \nu_2(Q)$. Как нетрудно видеть,

$$\nu_2(A) = \nu_2(m(m-2)) + 1.$$

Для доказательства невозможности равенства (4) нам достаточно убедиться в справедливости неравенств $\nu_2(B_l) > \nu_2(A)$, где

$$B_l = C_{2m+1}^{2l} 2^{l-1}, \quad 5 \leqslant l \leqslant m.$$

Имеем $B_l = C_{2m-9}^{2l-10} B 2^{l-1}$, где

$$B = \frac{m(m-1)(m-2)(m-3)(m-4)(2m+1)(2m-1)(2m-3)(2m-5)(2m-7)}{l(l-1)(l-2)(l-3)(l-4)(2l-1)(2l-3)(2l-5)(2l-7)(2l-9)}.$$

Поэтому

$$\nu_2(B_l) \geqslant \nu_2(B 2^{l-1}) = \nu_2(m(m-2)(m-4)) + a \geqslant \\ \geqslant \nu_2(m(m-2)) + 1 + a = \nu_2(A) + a,$$

где

$$a = \nu_2\left(\frac{2^{l-1}}{l(l-1)(l-2)(l-3)(l-4)}\right) > 0$$

при любом $l \geqslant 5$.

Комментарии. I. Уравнения

$$x^2 + 2 = 3^n, \quad x^2 + 4 = 5^n$$

также могут быть решены двумя способами (первое из них было представлено на VIII Кубке памяти Колмогорова). Уравнение

$$7x^2 - 4 = 3^n$$

решается первым способом, а уравнение

$$x^2 + 7 = 2^n$$

(уравнение Рамануджана – Нагеля, см. [4], а также [5, p. 205–206]) — вторым, но технически более сложно. Этим же способом, но технически проще

решается уравнение

$$x^2 + 1 = y^n$$

при любом целом $y > 1$ (см. материалы X Кубка памяти Колмогорова).

Есть многочисленные примеры подобных уравнений, которые можно решить школьными методами, при этом используются только основная теорема арифметики и метод остатков.

Один из неэлементарных подходов связан с эллиптическими кривыми. Решение уравнения (1) сводится к отысканию всех целых точек на кривых

$$2x^2 + 1 = 3^r y^3, \quad r \in \{0, 1, 2\}.$$

При $r = 0$ эта задача может быть решена элементарно (снова можно воспользоваться факториальностью кольца $\mathbb{Z}[\sqrt{-2}]$).

II. Решение уравнения

$$2x^2 + 1 = 3^n$$

первым способом есть на `www.artofproblemsolving.com` (см. [6]).

## Список литературы

[1] Кострикин А. И. *Введение в алгебру. Часть* I. *Основы алгебры*. М.: Физико-математическая литература. 2000.

[2] Спивак А. В. *Уравнения Пелля* // Квант. №4. 2002. С. 5–11.

[3] Barbeau E. J. *Pell's equation*. New-York: Springer-Verlag. 2003.

[4] Nagell T. *The Diophantine equation* $x^2 + 7 = 2^n$ // Ark. Mat. Vol. 4. 1961. P. 185–187.

[5] Mordell L. J. *Diophantine equations*. London: Academic Press Inc. 1969.

[6] `http://.../Forum/viewtopic.php?f=56&t=346906`

(*Н. Н. Осипов*)

## Опечатки, замеченные в №16

| Страница, | Строка | Напечатано | Следует читать |
|---|---|---|---|
| 189, | 8 снизу | $2\pi$ | $2/\pi$ |
| 197, | 7 снизу | $2^{-N}$ | $1-2^{-N}$ |
| 201, | 18 сверху | $\pi/2$ | $\pi/8$ |
| 235, | 7 сверху | 12.1 | 14.2 |

Подготовка оригинал-макета: LaTeX2$\varepsilon$,
METAPOST, М. Н. Вялый

Издательство Московского Центра
непрерывного математического
образования
119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241 74 83

Отпечатано в ППП «Типография „Наука“».
121099, Москва, Шубинский пер., д. 6.
Подписано в печать 17.02.13. Формат 70×100/16. Бумага офсетная. Печать офсетная. Печ. л. 13,0. Тираж 1000 экз. Заказ №