# Содержание

# Лекция 1. Диофантовы уравнения в истории математики

## 1 Введение

Изучение диофантовых уравнений исторически было одним из древнейших подходов к математике. С этой точки зрения, эта область уникальна тем, что даёт представление о пути, по которому шли математики прошлого, и помогает сформировать представление о том, на чем строятся и как были получены современные результаты. И, хотя на первый взгляд может быть неясно, какой практический смысл имеет, например, описание всех решений диофантового уравнения $x^a - y^b = 1$ или результат, известный как Великая теорема Ферма, однако конструкции и методы, полученные при исследовании этих и многих других схожих вопросов (а Великой теоремой Ферма занимались практически 350 лет), во многом составляют фундамент современной математики, в особенности алгебры.

Теорема Пифагора, известна уже как минимум 2500 лет. Довольно естественно, что математики древности, старавшиеся, если возможно, не выходить за пределы целых положительных чисел, пришли к задаче об отыскании всех прямоугольных треугольников с целыми стронами. Примеры таких треугольников (в частности самый знаменитый — египетский треугольник с соотношением сторон 3:4:5) известны по крайней мере 4000 лет. Вопрос о том, были ли результаты этих исследований напрямую применены на практике (некоторые авторы полагают, что древние использовали пифагоровы тройки для построения прямого угла) по большей части остаётся спорным, как и вопрос о практической пользе другой задачи, уходящей корнями в древность — задаче о построенях циркулем и линейкой, однако, с точки зрения математики, в процессе рещения этих задач (одно из древнейших дошедших до нас решений задачи о пифагоровых тройках принадлежит Евклиду; теория построений циркулем и линейкой оконачательно сформировалась в начале XIX века) было получено много других результатов, ставших неотъемлемыми частями современной математики.

## 2 Задача о пифагоровых тройках

### 2.1 Задача о пифагоровых тройках. Классическое решение

Задача о пифагоровых тройках ставится как задача об отыскании всех натуральных решений уравнения $x^2 + y^2 = z^2$. Это задача интересна также тем, что среди ее решений моно найти применение разных классических техник из разных областей математики. Классический подход к решению задачи — рассмотрение остатков и применение основной теоремы арифметики.

Нахождение всех натуральных решений, очевидно, решает задачу и для $x, y, z \in \mathbb{Z}$, если не рассматривать решения, где одно из слагаемых равно нулю. Также разумно сузить круг рассматриваемых троек до взаимно простых, то есть таких, что $(x, y, z) = 1$, так как все остальные тройки могут быть получены из взаимно простых домножением на число.

**Утверждение 1.** Если $x, y, z \in \mathbb{N}$ — пифагорова тройка, $(x, y, z) = 1$, то $(x, y) = 1$.

*Доказательство.* Докажем от противного. Пусть $x$ и $y$ делятся на какое-то простое число $p$. Тогда $x = p \cdot x', y = p \cdot y'$, следовательно

$$x^2 + y^2 = p^2 \cdot \left(x'^2 + y'^2\right) = z^2 \Rightarrow p \mid z^2 \Rightarrow p \mid z.$$

Из этого следует, что $(x, y, z) \neq 1$, противоречие. □

В доказательстве утверждения стоит выделить неявно используемое утверждение, которое, по сути, является одной из формулировок основной теоремы арифметики для целых чисел.

**Утверждение 2.** Если число $p$ простое и $p \mid a \cdot b$, то либо $p \mid a$, либо $p \mid b$.

Следующий шаг заключается в рассмотрении чётности слагаемых. Никакие два числа не могут быть чётными, так как тройка взамно проста. Также все три не могут быть нечётными, так как сумма двух нечётных чисел есть чётное число. Таким образом, ровно одно из чисел чётное.

**Утверждение 3.** Если $(x, y, z) = 1$, то $z$ нечётно.

*Доказательство.* Легко видеть, что квадраты натуральных чисел дают только остатки 0 и 1 при делении на 4. Если $x$ и $y$ нечётные, то

$$x^2 \equiv y^2 \equiv 1 \pmod 4 \Rightarrow z^2 \equiv 2 \pmod 4,$$

противоречие. □

Теперь без потери общности $y$ — чётное число: $y = 2k$. Тогда

$$4k^2 = z^2 - x^2 = (z - x)(z + x) = \frac{z - x}{2} \cdot \frac{z + x}{2}.$$

В последней части равенства оба числа целые в силу нечётности $z$ и $x$.

**Утверждение 4.** $(\frac{z+x}{2}, \frac{z-x}{2}) = 1$.

*Доказательство.* Если какое-то число делит одновременно $\frac{z+x}{2}$ и $\frac{z-x}{2}$, то оно делит и их сумму, то есть $z$. Также оно должно делить и их разность, то есть $x$. Значит $(\frac{z+x}{2}, \frac{z-x}{2}) \mid (x, z) = 1$, откуда следует утверждение. □

Из предыдущих утверждений следует, что $k^2$ есть произведение двух взаимно простых чисел.

**Утверждение 5.** Числа $\frac{z-x}{2}$ и $\frac{z+x}{2}$ являются квадратами натуральных чисел.

*Доказательство.* Пусть $k = p_1^{\alpha_1} \cdot \ldots \cdot p_q^{\alpha_q}$. Тогда, если $p_i \mid \frac{z+x}{2}$, то $p_i^{2\alpha_i} \mid \frac{z+x}{2}$. В самом деле, если это не так, то $p_i \mid \frac{z-x}{2}$ и одновременно $p_i \mid \frac{z+x}{2}$, что входит в противоречие с условием $(\frac{z-x}{2}, \frac{z+x}{2}) = 1$. Отсюда $\frac{z-x}{2} = \prod\limits_{i \in I} p_i^{2\alpha_i}$, $I \subset \{1, \ldots, q\}$, то есть $\frac{z-x}{2}$ является полным квадратом. Аналогичные рассуждения проходят и для $\frac{z+x}{2}$. □

Таким образом,

$$\frac{z+x}{2} = m^2, \frac{z-x}{2} = n^2,$$
$$z = m^2 + n^2, x = m^2 - n^2, y = 2mn$$

**Упражнение 1.** Если $m$ и $n$ — взаимно простые разной чётности, $m > n$, то числа $m^2 + n^2$, $m^2 - n^2$, $2mn$ взаимно просты.

**Вывод.** *Все решения диофантова уравнения $x^2 + y^2 = z^2$ имеют вид*

$$x = m^2 - n^2, y = 2mn, z = m^2 + n^2,$$

*где $m, n$ —взаимно простые натуральные числа разной чётности, $m > n$; при этом каждой такой паре $n, m$ соответсвует решение.*

## 2.2 Задача о пифагоровых тройках. Алгебро-геометрическое решение

Ключевое наблюдение в этом решении таково: ненулевой (возможно тривиальной или не взаимно простой) пифагоровой тройке $x, y, z$ можно сопоставить рациональные числа $\alpha = \frac{x}{z}$ и $\beta = \frac{y}{z}$, причем $\alpha^2 + \beta^2 = 1$, то есть точка $(\alpha, \beta)$ лежит на единичной окружности.

Более того, верно и обратное. Если рациональная точка $(\frac{a_1}{b_1}, \frac{a_2}{b_2})$ лежит на единичной окружности, то $\forall L \in \mathbb{N}$ такого, что $L\frac{a_1}{b_1}$ и $L\frac{a_2}{b_2}$ — целые, тройка $(L\frac{a_1}{b_1}, L\frac{a_2}{b_2}, L)$ — пифагорова.

Таким образом задача поиска рациональных точек окружности оказывается эквивалентной задаче нахождения всех пифагоровых троек. Для описания всех рациональных точек окружности используется следующий классический прием.

Пусть $O$ — тривиальная рациональная точка, имеющая координаты $(0, -1)$. Пусть $(\alpha, \beta)$ — рациональная точка и $l$ — прямая, проходящая через нее и через точку $O$. Уравнение $l$ имеет вид:

$$y = k \cdot x - 1$$

**Утверждение 6.** $k \in \mathbb{Q} \Leftrightarrow \alpha, \beta \in \mathbb{Q}$.

*Замечание.* Двум особым случаям $\alpha = 0, \beta = 1$ и $\alpha = 0, \beta = -1$ соответствуют значения $k = \infty$ и $k = 0$.

*Доказательство.* $k = \frac{\beta+1}{\alpha} \in \mathbb{Q}$, если $\alpha, \beta \in \mathbb{Q}$. В обратную сторону: пересечение прямой $l$ и единичной окружности может быть найдено из уравнения $x^2 + (kx-1)^2 = 1$. Отсюда

$$x^2 + k^2x^2 - 2kx + 1 = 1$$
$$x^2 + k^2x^2 - 2kx = 0$$
$$x^2 \cdot (1 + k^2) = 2kx$$
$$x \cdot (1 + k^2) = 2k$$
$$x = \frac{2k}{1+k^2}$$
$$y = kx - 1 = \frac{2k^2}{k^2+1} - 1 = \frac{k^2-1}{k^2+1}$$

Очевидно, что $k \in \mathbb{Q} \Rightarrow x, y \in \mathbb{Q}$. Также очевидно, что $(-1, 0)$ — тривиальное решение, поэтому сокращение при решении уравнение было произведено правомерно. □

Если $k \in \mathbb{Q}$, то $k = \frac{m}{n}, m, n \in \mathbb{Z}, n \neq 0$. Тогда $\alpha = \frac{2mn}{m^2+n^2}, \beta = \frac{m^2-n^2}{m^2+n^2}$. При выборе числа $L = m^2 + n^2$, получаем пифагорову тройку $(2mn, m^2 - n^2, m^2 + n^2)$. Как видно, соотношения получаются те же, что и в предыдущем решении, однако получены они совершенно иными методами.

## 2.3 Задача о пифагоровых тройках. Решение в гауссовых целых числах

Здесь и далее все рассматриваемые кольца будут коммутативными, будут содержать единицу и не будет иметь делителей нуля.

**Определение 1.** Кольцом *гауссовых чисел* $\mathbb{Z}[i]$ называется подкольцо поля $\mathbb{C}$, состоящее из чисел с целой вещественной и мнимой частью.

**Определение 2.** Ненулевой элемент кольца $a$ называется *обратимым*, если $\exists a^{-1} : a \cdot a^{-1} = a^{-1} \cdot a = 1$

**Определение 3.** Ненулевой необратимый элемент кольца $a$ называется *неразложимым*, если из того, что $a = b \cdot c$ следует, что один из элементов $b, c$ обратим.

**Определение 4.** Кольцо называется *факториальным*, если для любой ненулевой необратимый элемент представляется в виде произведения неразложимых элементов, причем единственным образом*.

*С точностью до обратимых элементов (в $\mathbb{Z}[i]$ это $1, -1, i, -i$), которые составляют *мультипликативную группу* кольца. Элементы кольца, отличающиеся домножением на обратимый, называются *ассоциированным* (обозначение: $a \sim b$). В дальнейшем, все рассуждения будут проводиться с точностью до умножения на обратимые элементы. Ремарки, подобные этой, будут, как правило, опускаться.

Хорошо известно, что наличия в кольце нормы и деления с остатком достатоно для выполнения большинства свойств целых чисел. Кольцо, в котором введена целочисленная норма и определено деление с остатком, называется *евклидовым*. Кольцо гаусовых чисел *евклидово*, его норма унаследованна из поля $\mathbb{C}$. Любое евклидово кольцо факториально, значит каждый неразложимый элемент является *простым*, то есть удовлетворяет свойству из утверждения 2. Таким образом, делимость в кольце гауссовых чисел устроена привычным образом, то есть выполнена основная теорема арифметики.

Гауссовы числа удобно представлять как целочисленную решетку на комплексной плоскости. Сопряженные числа получаются друг из друга поворотом на 90 градусов, поэтому с точки зрения теории делимости можно ограничиться рассмотрением свойств чисел первой четверти.

Простые гауссовы числа без мнимой части будут простыми и в кольце $\mathbb{Z}$. Обратное неверно, в чем можно убедиться на примере $2 = (1+i)(1-i)$. Более того, так как $1+i$ и $1-i$ ассоциированы, то число 2 можно считать точным квадратом. Точно так же, число $5 = (2+i)(2-i)$ перестает быть простым в гауссовых числах, однако, точным квадратом оно уже не будет.

В гауссовых числах уравнение $x^2 + y^2 = z^2$ переписывается как $(x + yi)(x - yi) = z^2$.

**Упражнение 2.** $(1 + i) \mid (a + bi) \Leftrightarrow a$ и $b$ — одной чётности.

**Упражнение 3.** Если $x, y$ — взаимно просты в $\mathbb{Z}$, то они взаимно просты и в $\mathbb{Z}[i]$. Более того, для любых $x, y \in \mathbb{Z}$ верно, что $\gcd_{\mathbb{Z}}(x, y) = \gcd_{\mathbb{Z}[i]}(x, y)$.

**Утверждение 7.** Если $x, y \in \mathbb{Z}$ — взаимно простые разной чётности, то гауссовы числа $x + yi$ и $x - yi$ взаимно просты.

*Доказательство.* Пусть $x + yi$ и $x - yi$ имеют общий простой делитель $p$, то есть $p \mid (x + yi), p \mid (x - yi)$. Тогда $p \mid 2x, p \mid 2y$. Так как $x, y$ — разной чётности, то $p$ не может быть ассоциированным с $1 + i$ и не может делить число 2. Тогда $p \mid x, p \mid y$. Противоречие. □

Из утверждения 7 следует, что числа $x + yi$, $x - yi$ являются точными квадратами. Пусть $x+yi \sim (m+ni)^2$, тогда $x+yi \sim m^2-n^2+2mni$. Отсюда, с точностью до порядка $x, y$ и их знаков, вытекает $x = m^2-n^2, y = 2mn, z = m^2 + n^2$, как и в предыдущих решениях.

# 3 Экскурс в историю математики

## 3.1 Великая теорема Ферма

Методы, подобные описанным выше (конечно, не в современной математической формулировке, а в традиционной записи того времени), довольно глубоко развил древнегреческий математик Диофант Александрийский, живший примерно в III веке нашей эры. Из 13 томов его «Арифметики»

до нас дошли только первые 6. Его труды, будучи обнаружены в Ватиканской библиотеке Рафаэлем Бомбелли в XVI веке, были частично включены в его труд «Алгебра» и позже полностью изданы, оказав большое влияние на некоторых математиков того времени. Именно на полях «Арифметики», в главе, посвященной исследованию задачи о пифагоровых тройках, Пьер Ферма сделал знаменитое замечание о том, что никакую другую степень нельзя представить в виде суммы двух таких же степеней и он «нашел поистине чудесное доказательство этой теоремы, однако поля книги слишком узки, чтобы его привести». Впоследствии, Ферма опубликовал доказательство для уравнения $x^4 + y^4 = z^2$ (что, очевидно, представляет собой более общий случай), но доказательство общего случая так и не было найдено. Так родилась Великая теорема Ферма, на доказательство которой потребовалось более 350 лет.

Современные Ферма математики в основном были увлечены активно развивающейся дисциплиной математического анализа, однако, эта задача необычайно заинтересовала математиков следующих поколений. Случай $n = 3$ был разобран (с некоторыми оговорками, связанными с основной теоремой арифметики) Эйлером. Доказательство для $n = 5$ было найдено независимо Лежандром и Дирихле, позже альтернативные доказательства нашли Гаусс, Лебег, Ламе, что еще раз показывает внимание, уделяемое великой проблеме.

В 1847 году Габриэлем Ламе была предпринята попытка доказательства общего случая, основанная на идее рассмотрения кольца чисел, подобных гауссовым, но включающего комплексный примитивный корень $n$-й степени из единицы. Формально, рассматривалось минимальное кольцо, содержащее одновременно целые числа и комплексный корень $n$-й степени из единицы, обозначаемый, обычно $\xi_n$ или, когда из контекста ясен порядок, просто $\xi$. Существование такого кольца легко обосновать — так как поле $\mathbb{C}$ есть кольцо, содержащее целые числа и $\xi$, а пересечение двух колец является кольцом, то среди всех подходящих подколец $\mathbb{C}$ найдется наименьшее по включению. Такая конструкция часто используется в алгебре, а искомое кольцо обозначается как $\mathbb{Z}[\xi]$ (что согласуется с использованным ранее обозначением $\mathbb{Z}[i]$ для гауссовых чисел).

Полученное кольцо $\mathbb{Z}[\xi]$ гораздо сложнее описать, по сравнению с гауссовыми числами. Легко видеть, что кольцо чисел вида $a_0 + a_1\xi + a_2\xi^2 + \ldots + a_{n-2}\xi^{n-2}$ содержит искомое кольцо*. Однако точнее определить, какая наивысшая степень в действительности нужна, не так просто.

**Упражнение 4** (сложное)**.** Если $n$ — простое число, то $1, \xi, \ldots, \xi^{n-2}$ линейно независимы.

*Замечание.* Смысл этого утверждения в том, что для простых значений $n$ придется задействовать все степени $\xi$ вплоть до $n - 2$.

* В самом деле, если $\xi \neq 1$ и $\xi^n - 1 = 0$, то $\xi^{n-1} + \ldots + \xi + 1 = 0$. Поэтому при умножении чисел такого вида от степеней $n$ и выше можно избавиться, используя равенство $\xi^n = 1$, а $\xi^{n-1}$ выражается из предыдущего равенства.

Так или иначе, в этом кольце, выражение $x^n + y^n$ разложится на линейные множители: $(x + y)(x + y\xi) \dots (x + y\xi^{n-1})$. Кажется, что для решения задачи осталось только аккуратно доказать, что все эти числа взаимно просты и являются точными $n$-ми степенями. Однако, в этом месте появляется другое, гораздо более неприятное препятствие.

**Утверждение 8.** Кольцо $\mathbb{Z}[\xi_{23}]$ — не факториально.

Это утверждение, технически довольно сложное, в смысловом плане тем не менее одозначно сводит все предудыщие попытки доказательства на нет, так как основная теорема арифметики оказывается необходимым в рассуждениях инструментом. Ошибочное предположение факториальности построенного кольца, практически сразу же отмеченное Жозефом Лиувиллем, оказывается серьезным пробелом в доказательстве Ламе, пробелом, восполнить который так и не удастся.

Более того, даже для малых значений $n$, структура кольца оказывается очень и очень сложной.

**Утверждение 9.** Кольцо $\mathbb{Z}[\xi_5]$ — факториально и содержит бесконечное количество обратимых элементов.

*Замечание.* Удивительный факт также состоит в том, что описание всех обратимых элеметов $\mathbb{Z}[\xi_5]$ сводится к описанию всех целых решений уравнения Пелля: $x^2 - 5y^2 = 1$, что приводит к исследованию таких объектов, как цепные дроби, и изучению приближений иррациональных чисел рациональными.

Великая теорема в итоге, после огромного числа неудачных попыток, была доказана в 1994 году Эндрю Уайлсом. Примечательно, что первый вариант доказательства он опубликовал в 1993 году, но в нем был обнаружен серьезный пробел, который удалось, тем не менее, оперативно устранить. Однако, потребовалось достаточно много времени для того, чтобы математическое сообщество окончательно признало факт доказательства, что впрочем есть здоровый скептицизм, накопленный за 357 полных неудачных попыток лет. Доказательство построено на теории элиптических кривых, разделе алгебрической геометрии, не имея практически ничего общего с приведенными выше попытками. Так, за века, проведенные над решением задачи, не только были развиты и отточены известные еще древним идеи и методы, но и придуманы совершенно новые подходы и концепции, нашедшие применение во многих других областях математики.

## 3.2 Наследие иных методов

В итоге, из элементарных подходов, помогших тремя разными способами решить задачу о пифагоровых тройках, развились совершенно разные науки. Классическая арифметика и теория делимости, с которых все начиналось, в конечном итоге стали неиссякаемым источников задач и различных техник, породив в том числе и саму Великую теорему Ферма.

Алгебраическая геометрия была развита в основном в XX веке, развив теорию элептических кривых и много других красивых разделов, открыв в конечном итоге путь к решению задачи, стоявшей больше трех столетий. Еще одним примечательным для этой науки уравнением стало безобидное с виду $x^a - y^b = 1$. Удивительный результат касательно этой, окончательно решенной в 2002 году проблемы, таков.

**Утверждение 10.** Если $x, y, a, b$ — натуральные числа, $a, b > 1$, то $x = 3, a = 2, y = 2, b = 3$.

Гауссовы же числа, введенные в мнографии Карла Фридриха Гаусса «Теория биквадратичных вычетов» также стали необходимым инструментом в математическом арсенале. С их помощью могут быть решены различные задачи о суммах квадратов, например, какие числа могут быть представлены в виде суммы двух квадратов и сколькими способами. С этой задачей связана еще одна, описание которой начинается казалось бы совсем изделека.

Первое известное доказательство того, что простых числе бесконечно много предложил еще Евклид. Сейчас это рассуждение знакомо каждому школьнику: если простые числа исчерпываются каким-то набором $p_1, \ldots, p_n$, то наименьший простой делитель числа $p_1 \cdot \ldots \cdot p_n + 1$ не может входить в этот набор по очевидным причинам. Поэтому это «новое» простое число, не входящее в наш набор, что влечёт противоречие.

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, .... После недолгого созерцания первых нескольких простых чисел можно заметить, что нечётные простые, имеющие остаток 1 от деления на 4 встречаются не реже тех, что имеют остаток 3. Нет ни какой-либо видимой закономерности, ни очевидной причины для того, чтобы какие-то из этих чисел когда-нибудь закончились. Возникает вопрос: конечно ли количество простых вида $4k + 3$? Ответ отрицательный, и это не очень сложно доказать, используя знакомый прием.

**Утверждение 11.** Существует бесконечно много простых чисел вида $4k + 3$.

*Доказательство.* Пусть простые вида $4k + 3$ исчерпываются каким-то набором $p_1, \ldots, p_n$. Тогда число $4p_1 \cdot \ldots \cdot p_n - 1$, имея остаток 3 по модулю 4, не делится ни на одно из чисел $p_1, \ldots, p_n$. Значит, все его простые делители имеют вид $4k + 1$. Но тогда оно само имеет остаток 1 по модулю 4, так как два числа вида $4k + 1$ при умножении дают число того же вида. Противоречие. □

Легко видеть, что с простыми вида $4k+1$ этот метод не сработает и нужно придумывать что-то новое. Ответ на этот вопрос неожиданным образом возвращает к задаче о числах, представимых в виде суммы двух квадратов, что в свою очередь приводит нас к гауссовым числам.

Примечательно, что выполнена гораздо более общая (вместе с тем, гораздо более сложная) теорема.

**Теорема 1** (Дирихле)**.** *Если $a, b \in \mathbb{N}$ — взаимно простые, то существует бесконечно много простых чисел вида $ak + b$.*

## 3.3 Задачи о простых

Простые числа, естественно, стали источником огромного числа математических задач. В их числе древнейшая, однако, все еще не решенная, проблема простых-близнецов. Вопрос: конечно ли число соседних простых чисел, отличающихся ровно на 2. Из теоремы о законе распределения простых чисел, доказанная в конце XIX столетия Адамаром, следует, что $n$-е простое число асимптотически имеет порядок $n \cdot \ln n$. Промежуток же между соседними простыми числами растет асимптотически не медленнее, чем $\ln n$. Функция бесконечно растущая, и кажется, что это весомый аргумент в пользу того, что количество простых-близнецов все-таки конечно.

Однако, ошеломляющие, притом совсем недавние (2013 год), результаты Итана Чжана показывают, что для некоторого $N < 7 \cdot 10^7$ существует бесконечно много простых, отстоящих друг от друга ровно на $N$. К 2014 году совместными усилиями Теренса Тао и проекта Polymath константа $7 \cdot 10^7$ была последовательно улучшена до 4680, затем до 600 и, наконец до 246. Используя широко известные недоказанные, но предположительно верные, гипотезы, оценку улучшили до $N = 12$, а позже и до $N = 6$.

Методы, используемые в этих работах относятся к очень современному разделу математики — аддитивной комбинаторике, который, можно сказать, складывается прямо сейчас. Так, одна из древнейших математических проблем может в скорейшем времени оказать решенной, породив, как и все перечисленные необычайно короткие и простые в постановке задачи, большой и очень плодотворный раздел математики.

Следующая примечательная задача состоит в описании совершенных чисел.

**Определение 5.** Число $n$ называется совершенным, если $\sum\limits_{d|n} d = n$.

**Упражнение 5** (простое)**.** Если $2^n - 1$ — простое, то число $2^{n-1}(2^n - 1)$ — совершенное.

**Упражнение 6** (посложнее)**.** Все чётные совершенные числа представляются в виде $2^{n-1}(2^n - 1)$, где $2^n - 1$ —простое.

**Упражнение 7** (открытая проблема)**.** Доказать, что нечётных совершенных чисел не существует.

Более того, неизвестно даже, бесконечно ли количество чётных совершенных чисел, то есть неизвестно, конечно ли число простых вида $2^n - 1$ — простых чисел Мерсенна.

Если же полюбопытствовать и задать очень похожий вопрос: конечно ли число простых вида $2^n + 1$, то очень скоро выяснится, что $n$ в свою очередь тоже должно быть степенью двойки.

**Утверждение 12.** Если $n$ не является степенью двойки, то $2^n + 1$ — составное.

*Доказательство.* Если $n$ не является степенью двойки, то оно представляется в виде $n = rs$, где $r$ нечётно. Тогда $2^n + 1 = 2^{rs} + 1 = (2^s)^r + 1 = (2^s + 1)(2^{s(r-1)} - 2^{s(r-2)} + \ldots + 1)$. Оба множителя, очевидно, отличны от 1, значит $2^n + 1$ — составное. □

Так, сменой знака «+» на «−» получается другая все ещё открытая проблема — вопрос существовании бесконечного количества простых чисел Ферма. При том, что сам Ферма предполагал, что все числа такого вида просты, что было опровергнуто Эйлером (из числа $2^{32} + 1$ он выделил множитель 641), на сегодняшний день не известно никаких простых чисел Ферма, кроме первых четырёх.

Еще одна гипотеза о простых числах — гипотеза Гольдбаха-Виноградова: любое чётное число представляется в виде суммы двух простых чисел. В начале XX века совершенно неожиданным открытием стало доказательство того, что любое число представляется в виде суммы $k < 700$ простых чисел. Но вскоре после этого Виноградов доказал еще более потрясающее утверждение: любое нечётное число представляется в виде суммы 3-х простых чисел. Для чётных чисел тем самым, достаточно четырёх простых. Вопрос о том, достаточно ли двух, остаётся открытым.

## 3.4 Математическая революция XIX века

Наконец, еще один плодотворнейший раздел, в корне преобразивший в свое время алгебру, о котором уже упоминалось. К началу XIX века были хорошо известны четыре задачи, стоявшие со времен античности.

Имея единичный отрезок:

- Построить круг площади 1 (построить отерзок длины $\pi$).
- Построить куб объёма 2 (построить отрезок длины $\sqrt[3]{2}$).
- Разбить данных угол на три равных.
- Найти все правильные многоугольники, которые можно построить

К тому моменту сформировалась достаточно развитая алгебра, чтобы снять второй и третий вопросы (по сути, стало известно, какие длины отрезков можно построить, однако, долго было не понятно, является ли $\pi$ таким числом). С последней задачей дела обстояли сложнее.

Античным геометрам были известны способы построения 3-х и 5-и угольника (можно показать, что для решения задачи достаточно выяснить, какие простые $n$-угольники можно построить). В 1796 году Гаусс нашел способ построить 17-угольник (и завещал изобразить его на своей могиле). Более того, позже он сформулировал полный критерий построимости многоугольника цикрулем и линейкой, который вновь отсылает нас к задаче о простых числах Ферма.

Еще одна очень простая в постановке задача, долгое время волновавшая умы математиков состоит в решении уравнений в радикалах. Каждый знает формулу корней квадратного уравнения через его коэффициенты. Также весьма известна аналогичная формула для корней кубического уравнения — формула Кардано. Для уравнения четвёртой степени аналогичная формула носит имя Феррари. К XVII веку эти формулы были уже хорошо известны и велись активные поиски общей формулы для уравнения $n$-й или хотя бы 5-й степени. В начале XIX века произошёл ряд продвижений в этой задаче. Сперва было доказано, что общей формулы нет для всех $n$, но было неясно, можно ли найти какой-то конкретный многочлен, корни которого бы не выражались в радикалах. Построить такой многочлен удалось Галуа, он же развил полную теорию, получившую после его имя, описывающую полный критерий разрешимости уравнения в радикалах. На тот момент открытие было потрясающего масштаба. Снятие бывших открытыми с античности проблем, связанных с разрешимостью, ознаменовало настоящую математическую революцию и создание той алгебры, которую мы знаем сейчас.

# Лекция 2. Великая теорема Ферма для малых показателей

**Теорема 1** (Ферма)**.** *Для любого $n > 2$ не существует отличных от нуля натуральных чисел $x, y, z$ таких, что $x^n + y^n = z^n$.*

# 4 Великая теорема Ферма при $n = 4$

## 4.1 Доказательство при $n = 4$

Это самый простой случай теоремы, доказать его можно, не выходя за пределы натуральных чисел. Доказательство, которое в свое время придумал Пьер Ферма*, основано на изобретенном им методе «бесконечного спуска», в сущности, одном из видов индукции. В первую очередь, как это часто бывает, рассуждая по индукции, удобно перейти к более общему уравнению $x^4 + y^4 = z^2$. Шаг будет заключаться в следующем: пусть из всех решений данного уравнения некоторое (положительное, так как знаки переменных неважны) решение $(x, y, z)$ имеет наименьший $z$. Если по этому решению можно построить другое, с меньшим $z$, то теорема будет доказана.

**Утверждение 1.** Числа $x, y, z$ попарно взаимнопросты (пишут $\ll x, y, z \gg = 1$, где под записью $\ll a_1, \ldots, a_n \gg$ понимается максимальный из попарных наибольших общих делителей $a_i$ и $a_j, 1 \leqslant i \neq j \leqslant n$).

*Доказательство.* Пусть какие-либо два числа из $x, y, z$ делятся на простое

*Это одно их тех редких его доказательств, что были записаны и дошли до нас

число $p$. Тогда очевидно, что третье также делится на $p$. Тогда

$$x = p\overline{x}, y = p\overline{y}, z = p\overline{z},$$
$$p^4(\overline{x}^4 + \overline{y}^4) = p^2\overline{z}^2,$$
$$p^2(\overline{x}^4 + \overline{y}^4) = \overline{z}^2,$$
$$p \mid \overline{z}^2 \Rightarrow p \mid \overline{z} \Rightarrow \overline{z} = p\overline{\overline{z}},$$
$$\overline{x}^4 + \overline{y}^4 = \overline{\overline{z}}^2, \overline{\overline{z}} < z.$$

Противоречие с минимальностью $z$. □

Далее, $(x^2, y^2, z)$ — пифагорова тройка. Число $z$ обязательно нечётно, а среди $x$ и $y$ чётным без потери общности можно считать $x$. Тогда

$$x = 2x_1,$$
$$4x_1^2 = 2mn, y^2 = m^2 - n^2, z = m^2 + n^2,$$
$$y^2 \bmod 4 = 1 \Rightarrow 2 \nmid m, 2 \mid n \Rightarrow n = 2n_1,$$
$$x_1^2 = mn_1$$

Далее, если $p \mid \gcd(m, n)$, то $p \mid \gcd(x, y) = 1$, значит $m$ и $n$ взаимно просты, поэтому $m$ и $n_1$ тоже. Тогда $m = a^2, n_1 = b^2$ для каких-то натуральных $a, b$.

Из того, что $y^2 = m^2 - 4n_1^2 \Rightarrow y^2 + 4n_1^2 = m^2$ следует, что $(y, m, 2n_1)$ — пифагорова тройка. Тогда для каких-то взаимно простых натуральных $q$ и $r$ выполнено $m = q^2 + r^2, 2n_1 = 2qr$. Итак, $n_1 = qr$ и одновременно $n_1 = b^2$. Так как $q$ и $r$ взаинопросты, то они должны являться полными квадратами: $q = t^2, r = s^2$. Итого,

$$m = q^2 + r^2, m = a^2, q = t^2, r = s^2,$$
$$a^2 = t^4 + s^4.$$

Так как $a = \sqrt{m} \leqslant m \leqslant m^2 < z$, то тройка $(t, s, a)$ даёт необходимое противоречие.

## 4.2 Роль случая $n = 4$ в общей задаче

Стоит отдельно отметить, что в общем случае теоремы Ферма, если $n > 2$, то либо $4 \mid n$, либо $n$ имеет нечётный простой делитель. Если $n = 4k$, то соотношение $(x^k)^4 + (y^k)^4 = (z^k)^4$ противоречит только что проведённым рассуждениям. Аналогичное рассуждение полностью сводит задачу к рассмотрению только простых значений $n$.

Куммер в середине XIX века доказал теорему для широкого класса регулярных простых (предположительно их плотность в натуральном ряде не превосходит 40%, а в первой сотне нерегулярных простых только три: 37, 59, 67). Позже с помощью компьютера его доказательство было доработано для всех простых, не превосходящих 2521 (1954 г.), а позже 125,000 (1978 г.) и 4,000,000 (1993 г.). Однако полностью решить задачу используя машинные вычисления практически невозможно.

**Упражнение 1.** Если бы теорема Ферма была бы неверна и $x^n + y^n = z^n$, то $|x|, |y|, |z| > n$.

Это упражнение иллюстрирует, что компьютерный поиск контрпримера требовал бы проведения операций с числами порядка $n^n$, что для даже для $n$ порядка 125,000 представляет вычислительно сложную задачу.

# 5 Числа Эйзенштейна

## 5.1 Норма и обратимые элементы

Для решения задачи при $n = 3$ необходимо исследовать структуру кольца $\mathbb{Z}[\omega]$* — так называемых чисел Эйзенштейна.

**Упражнение 2.** Пусть $\xi \neq 1$ — любой нетривиальный корень из единицы степени $p$ ($p$ — простое), тогда

$$x^p + y^p = (x + y)(x + \xi y) \dots (x + \xi^{p-1} y).$$

Таким образом, в кольце $\mathbb{Z}[\omega]$ выражение $x^3 + y^3$ раскладывается на линейные множители, что значительно облегчает анализ.

Стоит напомнить, что в общем случае $\mathbb{Z}[\xi_p]$ не является факториальным кольцом и основная теорема арифметики в нем не выполнена (первый такой пример при $p = 23$), что в свое время помешало Ламе построить доказательство для общего случая. Для некоторых простых можно доказать теорему из других соображений, в частности, Софи Жермен сделала это в случае, если $p$ и $2p + 1$ одновременно простые† и $p \nmid xyz$. Надо заметить, что случай $p \mid xyz$ сильно сложнее для анализа даже при $p = 3$ (в доказательстве существенно используется то обстоятельство, что 3 — не простое число в $\mathbb{Z}[\omega]$).

**Упражнение 3** (сложное)**.** Найти разложения на простые множители 5 в $\mathbb{Z}[\xi_5]$ и $p$ в $\mathbb{Z}[\xi_p]$‡.

Итак, $\mathbb{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbb{Z}\}$, так как $\omega^2 = -1 - \omega$ (более строго, здесь сказано, что этот набор чисел является кольцом и что все числа такого вида лежат в $\mathbb{Z}[\omega]$, которое по определению есть минимальное кольцо, содержащее $\mathbb{Z}$ и $\omega$). Невероятно удобная и естественная визуализация чисел Эйзенштейна — изображение их на комплексной плоскости, где они формируют полное замощение правильными треугольниками.

**Определение 1.** Нормой числа $z = a + b\omega$ называется $N(z) = a^2 - ab + b^2$. Легко убедиться, что $N(z) = a^2 + ab(\omega^2 + \omega) + b^2\omega^3 = (a + b\omega)(a + b\omega^2) = (a + b\omega)(a + b\overline{\omega}) = z\overline{z}$, то есть $N(z)$ — это квадрат привычной комплексной нормы.

---

*$\omega = e^{\frac{2\pi}{3} i} = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$ — примитивный корень из единицы степени 3.

†Такие простые в честь неё названы простыми Софи Жермен.

‡В нефакториальных кольцах за определение простого берется свойство $p \mid ab \Rightarrow p \mid a$ или $p \mid b$.

Удобное свойство такой нормы — мультипликативность. В самом деле, очевидно, что $\forall a, b \in \mathbb{Z}[\omega] \to N(ab) = N(a)N(b)$. Стоит отметить, что это ни в коем случае не является аксимой нормы, и в других кольцах это свойство может не быть выполнено.

При исследовании евклидового кольца первоочередная задача заключается в описании его мультипликативной группы ведь основная теорема арифметики верна с точностью до умножения на обратимые элементы.

**Утверждение 2.** $a \in \mathbb{Z}[\omega]$ — обратим $\Leftrightarrow N(a) = 1$.

*Доказательство.* $N(a) = 1 \Rightarrow a\overline{a} = 1 \Rightarrow \overline{a}$ — обратный к $a$ элемент.

Если $ab = 1$, то $ab\overline{ab} = 1 \Rightarrow N(a)N(b) = 1$. Произведение двух целых положительных чисел равно 1 тогда и только тогда, когда каждой из них равно 1, то есть $N(a) = 1$. □

$$N(a + b\omega) = 1 \Leftrightarrow a^2 - ab + b^2 = 1,$$
$$4a^2 - 4ab + 4b^2 = 4,$$
$$(2a - b)^2 + 3b^2 = 4.$$

Далее очевидно, что задача сводится к перебору целых значений $b$ от $-1$ до 1. При фиксированном значении $b$ для $a$ существует не более двух возможных значений.

**Упражнение 4.** Перебрав варианты, показать, что обратимыми в кольце чисел Эйзенштейна являются элементы $1, -1, \omega, -\omega, 1 + \omega, -1 - \omega$.

Если записать эти элементы в другом виде, то можно увидеть, что все они представляют собой степени числа $1 + \omega$, которое является примитивным корнем степени шесть из единицы. Это также означает, что мультипликативная группа кольца чисел Эйзенштейна изоморфна $\mathbb{Z}_6$.

## 5.2 Число $\lambda$

Дальнейшее исследование коснется важного числа $\lambda = 1 - \omega$. Первое наблюдение состоит в том, что $N(\lambda) = 3$.

**Утверждение 3.** Если норма числа $p \in \mathbb{Z}[\omega]$ — простое число, то само число $p$ тоже простое.

*Доказательство.* Если $p$ равно произведению двух неразложимых необратимых элементов $a$ и $b$, то $N(p) = N(a)N(b)$, то есть какая-то из норм равна 1, а в этом случае или $a$ или $b$ — обратимый элемент, как было показано ранее*. □

* В этом доказательстве использована как основная теорема арифметики, так и мультипликативность нормы, поэтому в произвольном кольце оно не проходит. Более того, можно убедиться, что в произвольном кольце утверждение неверно.

Более того, тот факт, что норма числа $\lambda$ равна 3, автоматически означает, что $3 = \lambda\overline{\lambda} = (1-\omega)(2+\omega)$ в этом кольце не является простым числом. Это обстоятельство помогает разобрать важный случай $3 \mid xyz$, который в общем случае ($n \mid xyz$) представляет наибольшую трудность (в частности, как было сказано выше, Софи Жермен удалось доказать вариант теоремы для обширного класса простых, но только при $n \nmid xyz$). С точки зрения делимости, $\lambda$, как и любое другое простое число, имеет ровно 12 делителей — 6 обратимых и 6 ассоциированных.

Естественным образом, решение уравнения Ферма (доказательство отсутствия решений) в числах Эйзенштейна автоматически решает задачу и в целых числах. Используя внутренние симметрии кольца, можно заметить, что домножение $x$, $y$ или $z$ на любой корень из единицы третьей степени (и даже шестой, так от этого меняется только знак соответсвующего слагаемого) не меняет множество решений. Так, например, если одно из чисел $x+y$, $x+y\omega$, $x+y\omega^2$ делится на какое-то число $q$, то без потери общности можно считать, что это число $x+y$, так как в противном случае, домножая $x$ и $y$ на нужную степень $\omega$, можно получить тройку, все еще являющуюся решением уравнения Ферма и удовлетворяющую нужному свойству.

**Утверждение 4.** Пусть $x \in \mathbb{Z}[\omega]$, $\lambda \nmid x$. Тогда $x^3 \equiv \pm 1 \pmod 9$.

**Упражнение 5.**

(1) В $\mathbb{Z}[\omega]$ существует ровно 3 класса вычетов по модулю $\lambda$: $\{0, 1, -1\}$.

(2) В $\mathbb{Z}[\omega]$ существует ровно $N(p)$ классов вычетов по модулю $p$.

*Доказательство.* Если $x$ не кратен $\lambda$, то $x = r\lambda \pm 1$. Тогда

$$x^3 = r^3\lambda^3 \pm 3r^2\lambda^2 + 3r\lambda \pm 1.$$

Учитывая, что $3\lambda^2 \equiv 0 \pmod 9$, необходимо показать, что $r^3\lambda^3 + 3r\lambda$ делится на 9.

$$r^3\lambda^3 + 3r\lambda = 3r\lambda - 3r^3\lambda\omega = 3\lambda(r - r^3\omega).$$

В свою очередь $r = 0, 1$ или $-1 \pmod \lambda$. Если $\lambda \mid r$, то при вынесении $r$ выражение перед скобками делится на 9. Иначе $r = q\lambda \pm 1$, в этом случае $r^2 = q^2\lambda^2 \pm 2q\lambda + 1$. Тогда $\lambda \mid (r^2 - 1)$, то есть $r^2 = \lambda s + 1$. Итого,

$$3\lambda r(1 - r^2\omega) = 3\lambda r(1 - \omega - \lambda s\omega) = 3\lambda r(\lambda - \lambda s\omega) = 3\lambda^2(1 - s\omega) \equiv 0 \pmod 9$$

□

*Замечание.* Геометрический смысл утверждения заключается в том, что числа, кратные $\lambda$, но не кратные $\lambda^2$, в кубе не могут попасть на расстояние меньше 3 от чисел, кратных $\lambda^4$.

# 6 Великая теорема Ферма при $n = 3$

## 6.1 Основная теорема арифметики в $\mathbb{Z}[\omega]$

**Утверждение 5.** $\pm x^3 \pm y^3 \pm z^3 \neq 0$ при $\lambda \nmid xyz$.

*Доказательство.* Если ни одно из чисел $x, y, z$ не делится на $\lambda$, то их кубы дают остаток $\pm 1$ по модулю 9, а значит их сумма не сравнима с 0 по модулю 9, то есть не равна 0. □

Далее можно считать, что $\ll x, y, z \gg = 1$, значит ровно одно число делится на $\lambda$. Исследовать этот случай можно методом «бесконечного спуска», рассмотрев более общее уравнение $\varepsilon_1 x^3 + \varepsilon_2 y^3 + \varepsilon_3 z^3 = 0$, где $\varepsilon_1, \varepsilon_2, \varepsilon_3$ — произвольные обратимые, $xyz \neq 0$, $\ll x, y, z \gg = 1, \lambda \mid xyz$. Первое ключевое рассуждение заключается в следующем утверждении.

**Утверждение 6.** $\lambda^2 \mid xyz$.

*Доказательство.* Пусть без потери общности $z = -\lambda \overline{z}$. Тогда

$$\varepsilon_1 x^3 + \varepsilon_2 y^3 = \varepsilon_3 \lambda^3 \overline{z}^3, \lambda \nmid \overline{z}.$$

Тогда по модулю 9 левая часть есть $\pm\varepsilon_1 \pm \varepsilon_2$, а правая не равна 0 и делится на $\lambda^3$, но не на $\lambda^4$. По предыдущим утверждениям это противоречие. □

Прежде чем предпринять следующий шаг, необходимо внести ясность в вопрос об основной теореме арифметики в кольце чисел Эйзенштейна. Так как норма уже была введена, осталось только привести правило деления с остатком, согласующееся с нормой в смысле определения евклидового кольца.

**Утверждение 7.** Если $z_1, z_2 \in \mathbb{Z}[\omega]$, то $\exists \beta, \gamma \in \mathbb{Z}[\omega]\colon z_1 = z_2\beta + \gamma$, причём $N(\gamma) < N(z_2)$*.

*Доказательство.* $\frac{z_1}{z_2} = \frac{a+b\omega}{c+d\omega} = \alpha + \beta\omega, \alpha, \beta \in \mathbb{R}$. Пусть $r, s$ являются округлёнными до ближайшего целого числами $\alpha, \beta$. Тогда для числа $\gamma/z_2 = (\alpha - r) + (\beta - s)\omega$ верно, что $|\alpha - r|, |\beta - s| \leqslant 0.5$ (здесь число $\gamma$ неявно определено через числа $\alpha, \beta, r, s, z_2$)). В таком случае, сообразно с формулой для нормы, получается

$$N(\frac{\gamma}{z_2}) \leqslant |\alpha - r|^2 + |\alpha - r||\beta - s| + |\beta - r|^2 \leqslant \frac{3}{4} \Rightarrow N(\gamma) < N(z_2).$$

Итого $z_1 = z_2(r + s\omega) + \gamma$, $N(\gamma) < N(z_2)$. $\gamma$ будет числом Эйзенштейна, так как все остальные числа в равенстве лежат в $\mathbb{Z}[\omega]$. □

**Упражнение 6** (сложное)**.** Найти норму в кольце $\mathbb{Z}[\xi_5]$.

*Вообще говоря, норма нулевого элемента евклидового кольца не определена (для примера можно рассмотреть кольцо многочленов, которое тоже является евклидовым с нормой, равной степени многочлена). Но, работая с кольцами, аналогичными числам Эйзенштейна, как правило оставляют за нулевым элементом нулевую норму

## 6.2 Доказательство при $n = 3$

Теперь, автоматически получив основную теорему арифметики для кольца чисел Эйзенштейна, можно сформулировать второе ключевое утверждение, которое по сути является шагом метода «бесконечного спуска».

**Утверждение 8.** Если $\varepsilon_1 x^3 + \varepsilon_2 y^3 = \varepsilon_3 \lambda^{3k} z^3, k \geqslant 2$, причём $\lambda \nmid z$, то существуют числа $\overline{x}, \overline{y}, \overline{z}, \sigma_1, \sigma_2, \sigma_3$, являющиеся решенем уравнения $\overline{\sigma}_1 \overline{x}^3 + \overline{\sigma}_2 \overline{y}^3 = \overline{\sigma}^3 \lambda^{3k-3} \overline{z}^3$, причём $\lambda \nmid \overline{z}$.

*Доказательство.* По модулю 9 данное уравнение обращается в сравнение $\pm\varepsilon_1 \pm \varepsilon_2 \equiv 0 \pmod 9$, что конечно, заменяется обычным равенством, то есть $\varepsilon_1 = \pm\varepsilon_2$. Итак,

$$\pm\varepsilon_2 x^3 + \varepsilon_2 y^3 = \varepsilon_3 \lambda^{3k} z^3.$$

Далее, поделив на $\varepsilon_2$ и меняя при необходимости знак $x$, получаем

$$x^3 + y^3 = u\lambda^{3k} z^3,$$
$$(x+y)(x+y\omega)(x+y\omega^2) = u\lambda^{3k} z^3,$$
$$(x+y) - (x-y\omega) = y\lambda, (x+y) - (x+y\omega^2) = y\lambda(1+\omega)$$

Во-первых, невозможна ситуация, когда $\lambda^2 \mid (x+y, x+y\omega)$, так как в этом случае $\lambda^2 \mid y\lambda \Rightarrow \lambda \mid y$, но в то же время $\lambda \mid z$, противоречие. Аналогично, наибольший общий делитель любой другой пары скобок не делится на $\lambda^2$.

Во-вторых, в разложении $(x+y, x+y\omega)$ не существует никаких простых множителей, кроме $\lambda$. Если $p \mid (x+y, x+y\omega)$, то $p \mid y\lambda \Rightarrow p \mid y$. Однако, $p \mid (\omega(x+y) - (x+y\omega)) = -x\omega \Rightarrow p \mid x$, что невозможно, так как $(x, y) = 1$. Так как $x + y \equiv x + y\omega \equiv x + y\omega^2 \pmod \lambda$, то $\ll x, y, z \gg \in \{1, \lambda\}$.

Однако, так как правая часть уравнения делится на $\lambda$, то и левая тоже, значит каждая скобка делится на $\lambda$. Более того, в одну скобку $\lambda$ входит в степени $3k - 2 \geqslant 4$, а в остальные в степени 1. Без потери общности $\lambda^{3k-2} \mid (x+y)$.

$$x + y = \alpha\lambda^{3k-2}, x + y\omega = \beta\lambda, x + y\omega^2 = \gamma\lambda,$$
$$\ll \alpha, \beta, \gamma \gg = 1,$$
$$\alpha\beta\gamma = uz^3,$$

По основной теореме арифметики:

$$x + y = \lambda^{3k-2}\alpha = \sigma_1 \lambda^{3k-2} \overline{x}^3,$$
$$x + y\omega = \lambda\beta = \sigma_2 \lambda \overline{y}^3,$$
$$x + y\omega^2 = \lambda\gamma = \sigma_3 \lambda \overline{z}^3.$$

Сложив с коэффициентами $1, \omega, \omega^2$, получаем 0 в левой, части, то есть

$$0 = \sigma_1 \lambda^{3k-2} \overline{x}^3 + \sigma_2 \omega \lambda \overline{y}^3 + \sigma_3 \omega^2 \lambda \overline{z}^3,$$
$$\overline{\sigma}_1 \lambda^{3k-3} \overline{x}^3 = \overline{\sigma}_2 \overline{y}^3 + \overline{\sigma}_3 \overline{z}^3,$$
$$\ll \overline{x}, \overline{y}, \overline{z} \gg = 1.$$

Что в точности и нужно доказать. □

Таким образом, если существует решение уравнения Ферма со степенью вхождения $\lambda$, равной $k$, то по доказанному можно за $k-1$ шаг перейти к решению со степенью 1, а таких решений нет.

# 7 Великая теорема Ферма при $n = 5$

## 7.1 План доказательства при $n = 5$

Первый важный вопрос уже фигурировал раньше в виде упражнения 6. Далее, необходимо немного исследовать структуру кольца $\mathbb{Z}[\xi_5]$.

**Упражнение 7** (сложное)**.** Найти мультипликативную группу $\mathbb{Z}[\xi_5]$. Сперва может быть полезно найти обратимые из $\mathbb{Z}[\xi_5] \cap \mathbb{R}$.

**Упражнение 8.** Обобщить прием, использованный в утверждении 8, то есть найти способ скомбинировать уравнения

$$x + y = \alpha_1 \lambda^q,$$
$$x + y\xi = \alpha_2 \lambda,$$
$$\dots,$$

чтобы снизить степень делимости на $\lambda$. Найти, чему в этом случае равняется $\lambda$ и найти разложение числа 5 на простые множители.

**Упражнение 9.** Доказать теорему Ферма при $n = 5$.

# Лекция 3. Уравнение Пелля

# 8 Задачи, сводимые к уравнению Пелля

**Задача.** Имеется урна, внутри которой $n$ чёрных шаров и $m$ белых. Какими должны быть $n$ и $m$, чтобы вероятность наугад вытянуть два белых шара была равна $\frac{1}{2}$?

Вероятность нужного события, очевидно, равна $\frac{C_m^2}{C_{m+n}^2} = \frac{m(m-1)}{(m+n)(m+n-1)}$. То есть нужно решить следующее диофантово уравнение:

$$2m(m-1) = (m+n)(m+n-1),$$
$$2m^2 - 2m = m^2 + 2mn + n^2 - n - m,$$
$$m^2 - 2mn - n^2 - m + n = 0,$$
$$(m-n)^2 - 2n^2 - (m-n) = 0,$$
$$4(m-n)^2 - 8n^2 - 4(m-n) + 1 = 1(2(m-n)-1)^2 - 2(2n)^2 = 1.$$

Обозначая $2(m-n)-1 = x, 2n = y$, получаем уравнение, названное в честь Джона Пелля*:

$$x^2 - 2y^2 = 1.$$

В общем случае уравнением Пелля называется уравнение вида $x^2 - my^2 = 1$, где число $m$ не является полным квадратом.

**Задача.** Найти все прямоугольные треугольники с целыми сторонам, у которых катеты отличаются ровно на 1.

**Упражнение 1.** Свести задачу о прямоугольных треугольниках к уравнению Пелля.

Ещё удивительный один факт, мотивирующий задачу, состоит в том, что при решении уравнения Пелля анализируются так называемые «гиперболические повороты», которые играют важную роль в теории относительности Эйнштейна.

# 9 Уравнения Пелля для $m = 2$

## 9.1 Отображения $\psi$ и $\overline{\psi}$

Решения двух похожих уравнений

$$x^2 - 2y^2 = 1, \tag{1}$$
$$x^2 - 2y^2 = -1 \tag{2}$$

в вещественных числах представляют собой две гиперболы на плоскости. Очевидно, что общие амсимптоты, задаваемые уравнением $x^2 - 2y^2 = 0$, не содержат целых точек и имеют угловой коэффициент $\sqrt{2}$. Это придает решению уравнения Пелля ещё один смысл, заключающийся в поиске рациональных приближений числа $\sqrt{2}$ (так как с ростом $x$ и $y$, если конечно решений бесконечно много, отношение $\frac{y}{x}$ всё ближе приближается к $\sqrt{2}$).

*Леонард Эйлер ошибочно приписывал авторство общей постановки этой задачи Пеллю, который, однако, не имел к уравнению никакого отношения. Несмотря на то, что оно упоминается в трудах древнегреческих и древнеиндийских математиков, а также в современной постановке в трудах Пьера Ферма, название уравнения прочно закрепилось в литературе.

Несмотря на то, что для решения уравнений можно ограничиться рассмотрением первого квадранта, удобнее искать решения с $x + \sqrt{2}y > 0$. Несколько решений сразу угадываются, например, $(1, 1)$ для уравнения с $-1$ в правой части, и $(1, 0)$ для другого.

Пусть задано оторбражние $\psi\colon \mathbb{R}^2 \to \mathbb{R}, \psi(x, y) = x + y\sqrt{2}$, представляющее собой проекцию вдоль направления, задаваемого вектором $(\sqrt{2}, -1)$, на ось ординат. Легко видеть, что $\psi^2(1, 1) = (1 + \sqrt{2})^2 = 3 + 2\sqrt{2}$, а также, что пара $(3, 2)$ является решением уравнения Пелля. Двойственное отображение $\overline{\psi}(x, y) = x - y\sqrt{2}$ тоже обладает полезными свойствами.

**Утверждение 1.** Если $\psi^k(1, 1) = m + n\sqrt{2}$ (очевидно, что $m, n > 0$), то $\overline{\psi}^k(1, 1) = m - n\sqrt{2}$.

*Доказательство.* Рассмотрев бином Ньютона для $(1 + \sqrt{2})^k$ и $(1 - \sqrt{2})^k$, легко заметить, что они отличаются только знаком каждого из слагаемых, в которых $\sqrt{2}$ присутствует в нечётной степени, из чего утверждение немедленно следует. □

Из утверждения следует, что $(\psi \cdot \overline{\psi})^k(1, 1) = m^2 - 2n^2$ для некоторых $n, m \in \mathbb{N}$. Вместе с тем, $(\psi \cdot \overline{\psi})^k = (1 + \sqrt{2})^k(1 - \sqrt{2})^k = (-1)^k$. Итак, для чётных $k$ результат действия $\psi^k$ на $(1, 1)$ будет являться решением (1), для нечётных — решением (2).

Более того, предыдущие рассуждения можно провести не только для точки $(1, 1)$, но и для произвольного решения (2), получив тем самым бесконечную цепочку решений для обоих уравнений в чередующемся порядке.

**Упражнение 2.** Доказать, что если взять за начальную точку решение уравнения (1), то получится бесконечная цепочка только из решений (1).

## 9.2 Структура множества решений при $m = 2$

Естественным образом возникает вопрос: верно ли, что все решения (1) и (2) могут быть порождены каким-то одним решением?

**Теорема 1.** *Все решения уравнений (1) и (2), лежащие в полуплоскости $x > 0$, могут быть получены указанным способом из решения $(1, 1)$ при некотром $k \in \mathbb{Z}$.*

*Доказательство.* Пусть существует некоторое решение $(A, B)$ одного из уравнений, которое не может быть представлено в виде $(1 + \sqrt{2})^k$. Без потери общности, $A, B > 0$. Тогда $\frac{\psi(A,B)}{1+\sqrt{2}} = -(A + B\sqrt{2})(1 - \sqrt{2}) = (2B - A) + (A - B)\sqrt{2}$. Это снова решение одного из уравнений (1) или (2), что легко проверить напрямую или так, как это было проделано раньше. Более, того, можно заметить, что проделать это можно не только с $(A, B)$ и $(1, 1)$, но и с любой парой решений.

**Упражнение 3.** Если $(A, B), (C, D)$ — два решения указанных уравнений, то $(AC + 2BD, AD + BC)$ тоже является решением.

Эта операция обратима, а значит является групповой и вводит на точках двух частей гипербол, лежащих в полуплоскости $x + y\sqrt{2} > 0$, структуру группы Ли, что означает её непрерывность по всем аргументам. Более того, точки гиперболы, соотвествующей (1), образуют подгруппу описанной группы индекса 2, которая также является группой Ли.

Таким образом, полученное решение $(2B - A) + (A - B)\sqrt{2}$ тоже не может быть представлено в виде $(1 + \sqrt{2})^k$, так как иначе исходное решение обязательно будет представимо. Для завершения доказательства необходимо ввести норму на решениях уравнения, чтобы показать, что бесконечный спуск по этому правилу невозможен. Норму задаёт не что иное, как функция $\psi(A, B)$.

Норма $\psi(A, B)$ лежит между нормами каких-то двух последовательных решений, имеющих вид $(1 + \sqrt{2})^k$ и $(1 + \sqrt{2})^{k+1}$. Если делить $(A, B)$ $k$ раз на $1 + \sqrt{2}$ по описанному правилу, то получится решение $(X, Y), Y > 0$, не представимое в виде степени $(1 + \sqrt{2})$, лежащее по норме между решениями $(1, 1)$ и $(1, 0)$, что невозможно.

В терминах теории групп, сказанное выше можно пересказать так: группа Ли, соответствующая уравнению (1) содержит дискретную подгруппу, содержащую все целочисленные его решения. Эта подгруппа изоморфна $\mathbb{Z}$. □

# 10 Уравнение Пелля в общем случае

## 10.1 Распространение групповой операции на кольцо $\mathbb{Z}^2$

Неожиданно сложной задачей в общем случае оказывается отыскание хотя бы одного решения. В частности, при $m = 61$, $x$-координата наименьшего решения больше $10^9$. Если же хотя бы одно решение найдено, то дальнейший анализ строится на тех же идеях, что и в случае $m = 2$.

**Теорема 2.** *При любом $m$, для которого существует хотя бы одно решение, решения образуют группу с операцией $*\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$, $(a, b) * (c, d) = (ac + m \cdot bd, ac + bd)$, причем эта группа изоморфна $\mathbb{Z}$.*

Пусть $\Gamma_\eta$ — гипербола $x^2 - my^2 = \eta$. $\bigcup\limits_{\eta \in \mathbb{Z}} \Gamma_\eta$ покрывает все целые точки плоскости, так как $\mathbb{Z}^2 \ni (x, y) \in \Gamma_{x^2 - my^2}$.

**Упражнение 4.** Пусть $L_{a,b}\colon \mathbb{R}^2 \to \mathbb{R}^2, L_{a,b}(c, d) = (a, b) * (c, d)$. Тогда $L$ линейно и обратимо тогда и только тогда, когда $a^2 - mb^2 \neq 0$.

Из упражнения следует, что $(\mathbb{R}^2, +, *)$ является алгеброй, а её обратимые элементы лежат в объединении прямых $x + \sqrt{m}y = 0$ и $x - \sqrt{m}y = 0$.

Пусть снова $\psi\colon \mathbb{R}^2 \to \mathbb{R}, \psi(a, b) = a + b\sqrt{m}$.

**Утверждение 2.** $\psi$ является гомоморфизмом алгебр $(\mathbb{R}^2, +, *)$ и $(\mathbb{R}, +, \cdot)$.

*Доказательство.* Чтобы показать, что $\psi$ — гомоморфизм необходимо и достаточно проверить, что оно сохраняет сложение и умножение. Линейность $\psi$ очевидна: $\psi(a+b) = \psi(a)+\psi(b)$. Проверка второго условия также несложна:

$$\psi((a,b)*(c,d)) = \psi(ac+mbd, ad+bc) =$$
$$(ac+mbd) + (ad+bc)\sqrt{m} = (a+b\sqrt{m})(c+d\sqrt{m}) = \psi(a,b)\psi(c,d).$$

□

**Упражнение 5.** Показать, что $\overline{\psi}(a,b) = a - b\sqrt{m}$ также является гомоморфизмом.

**Определение 1.** Пусть *норма* $N(a,b)$ точки $(a,b)$ задана формулой $N(a,b) = \psi(a,b)\overline{\psi}(a,b) = a^2 - mb^2$. Легко видеть, что для целых $a, b$ норма целая и обладает свойством мультипликативности.

**Вывод.** *Если* $\xi = N(a,b), \eta = N(c,d)$, *то* $(a,b) \in \Gamma_\xi$, $(c,d) \in \Gamma_\eta$, $(a,b) * (c,d) \in \Gamma_{\xi\eta}$.

## 10.2 Групповая структура решений в общем случае

Если существует хотя бы одно целое решение уравнения Пелля $(x,y) \in \Gamma_1$, то все его целые степени относительно операции $*$ образуют дискретную изоморфную $\mathbb{Z}$ подгруппу в описанной алгебре. Все элементы этой подгруппы, естественно, также будут целочисленными решениями уравнения Пелля. Чтобы доказать, что никаких других решений нет, необходимо развить технику, использованную в случае $m = 2$.

Пусть для определённости $(x,y)$ таково, что $x, y > 0$ и $\nexists (x_1, y_1) \in \mathbb{Z}^2 \cap \Gamma_1 : x_1 < x$. Пусть существует какое-нибудь решение $(x', y') \in \mathbb{Z}^2 \cap \Gamma_1, x', y' > 0$, не являющееся степенью $(x,y)$. $(x,-y)$ тоже является целым решением уравнения Пелля, притом $\psi(x,-y) < 1$, так как $\psi((x,y)*(x,-y)) = \psi(x,y)\psi(x,-y) = 1, \psi(x,y) > 1$.

Тогда $\psi((x',y')*(x,-y)) < \psi(x',y')$, притом $(x',y')*(x,-y)$ является целым решением уравнения Пелля. Если $\psi(x,y)^k < \psi(x',y') < \psi(x,y)^{k+1}$ (такое $k$ существует, так как $(x',y')$ не является степенью $(x,y)$), то если $p = (x',y')*(x,-y)^k$, то $p$ снова будет решением уравнения Пелля, при этом $1 = \psi(1,0) < \psi(p) < \psi(x,y)$. Но это противоречит предположению о том, что решение $(x,y)$ обладает наименьшим $x$ (так как при движении по гиперболе в верхней полуплоскости $x$ и $y$ монотонно возрастают, то $(x,y)$ будет решением с наименьшим $\psi(x,y)$, при условии, что $\psi(x,y) > 1$).

**Упражнение 6.** Показать, что если существует хотя бы одно целое решение уравнения $x^2 - my^2 = -1$, то порождённая им группа включает в себя все решения этого уравнения и уравнения Пелля и изоморфна $\mathbb{Z} \oplus \mathbb{Z}_2$.

# 11 Существование решения уравнения Пелля

## 11.1 Методы геометри чисел

**Упражнение 7.** Доказать, что отображние $\psi$, ограниченное на $\mathbb{Q}^2$, вкладывается в $\mathbb{R}$ (взаимно-однозначно), а множество $\psi(\mathbb{Q}^2) = \{a + b\sqrt{m} \mid a, b \in \mathbb{Q}\}$ является двумерным линейным пространством над $\mathbb{Q}$ и более того, полем.

Итак, $(\mathbb{Z}^2, *, +)$ является алгеброй (а значит и кольцом), изоморфной кольцу $\{a + b\sqrt{m} \mid a, b \in \mathbb{Z}\} \subset \mathbb{R}$. Уравнения Пелля в такой интерпретации ставит вопрос о поиске обратимых элементов (мультипликативной группы) этого кольца, так как для поиска обратного к элементу $(a, b)$ нужно

$$\begin{aligned}(a + b\sqrt{m})(c + d\sqrt{m}) &= 1,\\ (a - b\sqrt{m})(c + d\sqrt{m}) &= 1,\\ (a^2 - mb^2)(c^2 - md^2) &= 1.\end{aligned}$$

В последнем равенстве все числа целые, откуда $a^2 - mb^2 = \pm 1$ и $c^2 - md^2 = \pm 1$. С другой стороны, если $a^2 - mb^2 = \pm 1$, то $(a+b\sqrt{m})(a-b\sqrt{m}) = \pm 1$ и оба числа обратимы. Итак, решение уравнения Пелля — это в сущности исследование обратимых элементов описанного кольца. Так как структура этого кольца уже выяснена ранее, остается только доказать существование хотя бы одного его элемента, для чего используются методы геометрии чисел.

**Лемма.** *Пусть задана (измеримая и ограниченная) фигура $\Phi \subset \mathbb{R}^2$ с площадью $S(\Phi) > 1$. Тогда $\exists x, y \in \Phi\colon (x - y)$ — целочисленный вектор.*

*Доказательство.* Поскольку фигура ограничена, то $\exists r \in \mathbb{N}\colon \forall p \in \Phi \to \max\{p_x, p_y\} < r$. Иными словами, фигура $\Phi$ вписывается в достаточно большой целочисленный квадрат.

Пусть теперь $\Phi_1, \ldots, \Phi_k$ — части фигуры, попавшие в полуоткрытые ячейки целочисленной решётки ($Q_{a,b} = \{x, y \mid a \leqslant x < a + 1, b \leqslant y < b + 1\}$). Пусть также $\Psi_i = \{(\{x\}, \{y\}) \mid (x, y) \in \Phi_i\}$, где $\{x\}$ обозначает дробную часть $x$. Иными словами, $\Psi_i$ есть параллельный перенос фигуры $\Phi_i$ на целочисленный вектор в единичный квадрат $0 \leqslant x < 1, 0 \leqslant y < 1$.

По свойству площадей (во всех операциях выше рассматриваемые объекты остаются измеримыми) $\sum_i S(\Psi_i) = \sum_i S(\Phi_i) = S(\Phi) > 1$. Если никакие две фигуры $\Psi_i, \Psi_j$ не пересекаются, то $S(\bigcup_i \Psi_i) = \sum_i S(\Psi_i) > 1$. С другой стороны, $\bigcup_i \Psi_i \in [0; 1) \times [0; 1)$, откуда $S(\bigcup_i \Psi_i) \leqslant 1$, противоречие.

Тогда существует точка $p \in [0; 1) \times [0; 1)$, принадлежащая одновременно двум фигурам $\Psi_i, \Psi_j$. Однако, из того, что $p \in \Psi_i$ следует, что $p$ отличается от какой-то точки $\Phi$ на целочисленный вектор. Аналогично, $p$ отличается от какой-то другой (отличной от предыдущей по построению) точки $\Phi$ на целочисленный вектор. Из этого немедленно следует, что в $\Phi$ есть две точки, отстоящие на целочисленный вектор. □

*Замечание.* Доказательство леммы легко обобщается на многомерный случай.

**Лемма** (Минковского о выпуклом теле)**.** *Пусть* $\Phi \subset \mathbb{R}^2$ *центрально-симметричная, выпуклая, измеримая фигура с площадью* $S(\Phi) > 4$*. Тогда* $\exists p \in \Phi \cap (\mathbb{Z}^2 \setminus (0,0))$.

*Доказательство.* Пусть $\Psi = \{(\frac{x}{2}, \frac{y}{2}) \mid (x,y) \in \Phi\}$. Очевидно, что $\Psi$ измерима и имеет площадь больше 1. Тогда по предыдущей лемме в $\Psi$ существуют две точки $p \neq q$, отстоящие на целочисленный вектор $v$. Остаётся доказать, что $v \in \Phi$. Это верно, так как $(-p) \in \Psi \Rightarrow \frac{(-p)+q}{2} \in \Psi$ (из-за выпуклости). Тогда $q - p = v \in \Phi$ по определению фигуры $\Psi$. Итого, ненулевая точка с целыми координатами вектора $v$ лежит в фигуре $\Phi$. □

## 11.2 Существование решений уравнения Пелля

Лемма Минковского оказывается невероятно полезной для отыскания целой точки на гиперболах $\Gamma_n$. Пусть $p_{\pm n}, q_{\pm n}$ — это точки пересечения гипербол $\Gamma_{\pm n}$ с координатными осями.

**Упражнение 8.** Найти наименьшее $n$, такое что площадь ромба, построенного на точках $p_{\pm n}, q_{\pm n}$ больше 4.

**Упражнение 9.** Пусть $A \in \Gamma_n$, параллелограмм $ABCD$ имеет все точки на гиперболах $\Gamma_{\pm n}$, причем стороны его параллельны асимптотам $y = \pm\sqrt{m}x$. Доказать, что его площадь не зависит от выбора точки $A$ и найти её.

Пусть теперь фигура $X = \bigcup\limits_{|x|<n, x\in\mathbb{R}} \Gamma_x$. Ключевое наблюдение заключается в том, что $X$ содержит бесконечно много целых точек. В самом деле, если их конечное число, то пусть точка $p$ лежит в первом квадранте и является самой близкой к асимптоте $y = \sqrt{m}x$. Важно, что точка $p$ не может лежать на асимптоте, так как наличие целой точки на асимптоте значило бы, что $m$ является полным квадратом. Тогда можно выбрать на гиперболе $\Gamma_n$ точку $A$, такую что расстояние от $A$ до асимптоты меньше, чем расстояние от $p$ до неё же. Тогда параллелограмм из упражнения 9 содержит целую точку, которая принадлежит $X$ и находится ближе к асимптоте, чем точка $p$, что противоречит предположению.

Итак, фигура $X$ содержит бесконечное число целых точек. Однако, все её целые точки покрываются конечным числом гипербол $\Gamma_{-n}, \ldots, \Gamma_n$. Это значит, что на какой-то гиперболе $\Gamma_l$ находится бесконечно много целых точек.

**Утверждение 3.** На гиперболе $\Gamma_l$ существуют точки $(a,b)$ и $(c,d)$, такие что $a \equiv c \pmod{l}, b \equiv d \pmod{l}$.

*Доказательство.* Поскольку на $\Gamma_l$ бесконечно много целых точек, то пусть $p_1, \ldots, p_{l^2+1}$ — проивзольные $l^2+1$ из них. Тогда по принципу Дирихле в один из $l^2$ классов сравнимости по модулю $l$ (обеих координат) попадёт хотя бы две точки. □

**Утверждение 4.** $(c,d) \mid (a,b)$ в кольце $(\mathbb{Z}^2, *)$.

*Доказательство.* Так как кольцо изоморфно $\mathbb{Z}[\sqrt{m}]$, то можно проверить делимость в нём.

$$\frac{a+b\sqrt{m}}{c+d\sqrt{m}} = \frac{(a+b\sqrt{m})(c-d\sqrt{m})}{l} = \frac{(ac-mbd)+\sqrt{m}(bc-ad)}{l}.$$

Так как $a \equiv c \pmod{l}, b \equiv d \pmod{l}$, то $ac-mbd \equiv a^2-mb^2 = l \equiv 0 \pmod{l}$. Аналогично $bc-ad \equiv ab-ab \equiv 0 \pmod{l}$. Значит $a+b\sqrt{m}$ делится нацело на $c+d\sqrt{m}$. □

Аналогично, $(a,b) \mid (c,d)$, что означает, что на самом деле эти числа ассоциированы в кольце $(\mathbb{Z}^2, *)$, то есть отличаются домножением на обратимый элемент. Так как решения уравнения Пелля и составляют обратимые элементы этого кольца, то частное $(a,b)$ и $(c,d)$ (в любом порядке) будет решением уравнения Пелля.

# 12 Уравнение Пелля и цепные дроби

## 12.1 Пример нахождения решений уравнения Пелля через цепные дроби

$$\sqrt{15} = 3 + (\sqrt{15}-3) = 3 + \cfrac{1}{\cfrac{\sqrt{15}-3}{6}} = 3 + \cfrac{1}{1+\cfrac{1}{\cfrac{6}{\sqrt{15}-3}}} = 3 + \cfrac{1}{1+\cfrac{1}{3+\sqrt{15}}}.$$

Таким образом,

$$\sqrt{15} = 3 + \cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{\ddots}}}}}$$

Можно заметить, что дроби

$$3 + \frac{1}{1} = \frac{4}{1}, 3 + \cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{1}}} = \frac{31}{8}$$

соответствуют решениям $(4,1),(31,8)$ уравнения Пелля $x^2 - 15y^2 = 1$. На самом деле этот факт не является простым совпадением. Разумная гипотеза, возникающая при анализе вышеописанных преобразований, состоит

в том, что для любого целого числа $m$, не являющегося полным квадратом выполнено

$$\sqrt{m} = a_0 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{2a_0 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{2a_0 + \cfrac{1}{\ddots}}}}}}}$$

Гипотеза подкрепляется следующей важной теоремой.

**Теорема 3** (Лагранжа)**.** *Пусть $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Тогда $\alpha$ обладает конечно-периодической цепной дробью тогда и только тогда, когда $\alpha$ удовлетворяет уравнению $ax^2 + bx + c = 0$ для некоторых целых $a, b, c$.*

## 12.2 Основные свойства цепных дробей

Для более тонкого понимания связи этих явлений нужно вспомнить основные свойства цепных дробей. В самом общем случае цепная дробь*

$$[a_0; a_1, \dots a_n] = a_0 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}$$

рассматривается в поле рациональных дробей $\mathbb{Z}(a_0, \dots, a_n)$ с $n+1$ независимой переменной $a_0, \dots, a_n$, которые следует воспринимать просто как формальные символы. Если сворачивать цепную дробь от конца к началу, то полученное выражение будет иметь вид $\frac{P_n}{Q_n}$, где $P_n, Q_n \in \mathbb{Z}[a_0, \dots, a_n]$.

Правило вычисления $P_n$ и $Q_n$ также выполнено в самом общем случае и выглядит следующим образом:

$$\begin{cases} P_n = a_n P_{n-1} + P_{n-2}, \\ Q_n = a_n Q_{n-1} + Q_{n-2}. \end{cases}$$

Основное свойство числителей и знаменателей цепных дробей $P_n Q_{n-1} - P_{n-1} Q_n = (-1)^n$ также выполнено в общем случае.

Более того, полезно заметить, что при вычислении цепных дробей не используется вычитание, то есть в некотором смысле вместо кольца $\mathbb{Z}$ можно поставить не являющееся кольцом множество целых положительных чисел $\mathbb{Z}_+$ с операциями сложения и умножения.

*Для более подробного и исчерпывающего описания можно обратиться к книге А. Я. Хинчина «Цепные дроби»

Если же рассмотреть более специальный случай, когда $a_0, \ldots, a_n \in \mathbb{R}_+$, то значение цепной дроби также будет лежать в $\mathbb{R}_+$. Поскольку наибольший интерес представляет исследование свойств бесконечных цепных дробей, то естественным образом возникает вопрос, когда ряд из значений подходящих цепных дробей сходится к какому-то числу.

**Теорема 4.** *Пусть числа $a_0, \ldots, a_n, \ldots \in \mathbb{R}_+$. Тогда бесконечная цепная дробь $[a_0; a_1, \ldots, a_n, \ldots]$ определена (сходится последовательность $\frac{P_n}{Q_n}$) тогда и только тогда, когда ряд $\sum a_i$ расходится.*

**Упражнение 10** (сложное)**.** Доказать теорему.

Однако, даже если ряд сходится, можно показать, что подходящие дроби с чётными номерами возрастают, а с нечётными убывают и всегда будут меньше, чем дроби на чётных местах. Если сумма $\sum a_i$ расходится, то между пределом последовательности чётных приближений и пределом последовательности нечётных будет какое-то расстояние, в противном случае последовательности стремятся к одному и тому же числу.

Из теоремы также вытекает то, что любая цепная дробь с целыми положительными коэффициентами заведомо сходится к какому-то вещественному числу.

## 12.3 Геометрическая интерпретация цепных дробей

Полезным инстрементом для работы с цепными дробями является их визуальное представление на координатной плоскости. Если положить $P_{-2} = 0, Q_{-2} = 1, P_{-1} = 1, Q_{-1} = 0$[†] и изображать дробь $\frac{x}{y}$ как точку $(x, y)$ плоскости (значением дроби в этом случае будет наклон прямой, проходящей через начало координат и её точку), то процесс нахождения подходящей дроби $\frac{P_n}{Q_n}$ числа $\alpha$ из двух предыдущих может быть описан как последовательное прибавление вектора $(P_{n-1}, Q_{n-1})$ к начальному вектору $(P_{n-2}, Q_{n-2})$ до тех пор, пока получающаяся точка не перейдет по другую сторону прямой $y = \alpha x$. Последняя точка перед переходом и будет соотвествовать подходящей дроби.

Операции сложения векторов в вещественных числах соотвествует операции взятия медианты двух дробей: $\frac{a}{b} * \frac{c}{d} = \frac{a+b}{c+d}$. Все дроби, получаемые в процессе построения (последовательного взятия медианты) носят название *промежуточных* и исследуются наравне с подходящими.

**Упражнение 11.** Доказать, что описанный процесс построения строит такую же цепную дробь, что и стандартный процесс построения, использующий операцию взятия целой части.

Основное свойство подходящих дробей на плоскости соответсвует тому факту, что параллелограмм, построенный на соседних подходящих дробях

[†]можно убедиться, что никакие из правил для подходящих дробей от этого не нарушаются

имеет единичную площадь. Это следует из того, что площадь такого параллелограмма может быть записана как определитель $\begin{Vmatrix} P_n & Q_n \\ P_{n-1} & Q_{n-1} \end{Vmatrix} = 1$. Таким образом число $\alpha$ помещается в цепочку все более узких параллелограммов единичной площади, что связано с исследованием группы невырожденных преобразований плоскости, сохраняющих площадь $SL(2,\mathbb{Z})$. Эта же группа отвечает за вид разных целочисленных квадратичных форм[‡].

## 12.4 Наилучшие приближения

Один из главных смыслов представления числа в виде цепной дроби — это получение его хороших рациональных приближений.

**Определение 2.** Число $\frac{p}{q} \in \mathbb{Q}$ является наилучшим приближением *первого рода* числа $\alpha \in \mathbb{R}$ , если $\forall\ \mathbb{Q} \ni \frac{a}{b} \neq \frac{p}{q}, |b| \leqslant |q| \to \left|\alpha - \frac{p}{q}\right| < \left|\alpha - \frac{a}{b}\right|$.

**Определение 3.** Число $\frac{p}{q} \in \mathbb{Q}$ является наилучшим приближением *второго рода* числа $\alpha \in \mathbb{R}$ , если $\forall\ \mathbb{Q} \ni \frac{a}{b} \neq \frac{p}{q}, |b| \leqslant |q| \to |q\alpha - p| < |b\alpha - a|$.

**Теорема 5.** *Любое приближение первого рода является промежуточной дробью, если в качестве промежуточных рассматривать в том числе и дроби, полученные при* $p = 1, q = 0$.

**Упражнение 12.** Доказать теорему и убедиться, что обратное утверждение верно не всегда.

**Теорема 6.** *Любое приближение втогоро рода является подходящей дробью. Если ещё* $\alpha - \frac{1}{2}$ *не является целым числом, то любая подходящая дробь является наилучшим приближением второго рода.*

**Упражнение 13.** Доказать теорему.

## 12.5 Формулы для решений уравнения Пелля

**Упражнение 14.** Для подходящей дроби $\frac{p_n}{q_n}$ выполнено

$$\frac{1}{q_n(q_n + q_{n+1})} < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}}.$$

**Упражнение 15.** Если для каких-то $p, q \in \mathbb{N}$ выполнено $\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q^2}$, то $\frac{p}{q}$ является подходящей дробью.

Остаётся только заметить, что любое решение уравнения Пелля удовлетворяет условию предыдущего упражнения, а значит искать решения следует исключительно среди подходящих дробей числа $\sqrt{m}$.

[‡]Об этом подробно написано в книге Джона Хортона Конвея «Квадратичные формы, данные нам в ощущениях» (John Horton Conway - The Sensual (Quadratic) Form).

**Упражнение 16.** Для любого $m \in \mathbb{N}$, не являющегося полным квадратом его цепная дробь имеет вид $\sqrt{m} = [a_0; a_1, \dots a_n, 2a_0, a_1, \dots, a_n, 2a_0, \dots]$, а решения уравнения Пелля имеют вид

$$\frac{x_1}{y_1} = [a_0; a_1 \dots a_n],$$
$$\frac{x_2}{y_2} = [a_0; a_1 \dots a_n, 2a_0, a_1, \dots, a_n],$$
$$\frac{x_3}{y_3} = [a_0; a_1, \dots, a_n, 2a_0, a_1, \dots, a_n, 2a_0, a_1, \dots, a_n],$$
$$\vdots$$

# Лекция 4. Построения циркулем и линейкой

# 13 Представление чисел в виде суммы двух квадратов

## 13.1 Одна из задач Эрдёша

Одно из классических диофантовых уравнений второй степени записывается как $x^2 + y^2 = m, m \in \mathbb{N}$ и ставит вопрос о количестве целых точек на окружности радиуса $\sqrt{m}$. Одним из интересных приложений, мотивирующих задачу, является открытая проблема, поставленная впервые Палом Эрдёшем.

**Задача.** Пусть $P_n$ — набор, состоящий из точек плоскости $p_1, \dots, p_n$, а $f(P_n)$ есть наибольшее количество одинаковых расстояний между какими-либо двумя точками. Какова асимптотическая скорость роста $f(P)$ при $n \to \infty$, если из всех конфигураций точек в качестве $P_n$ берется та, у которой наибольшее значение $f(P_n)$?

Перебрав некоторые простые конструкции, легко получить примеры линейного роста искомой величины. Содержательный же вопрос в том, можно ли построить серию конфигураций со сверхлинейным ростом. Наилучшую известную на сегодня конструкцию построил сам Эрдёш. Его утверждение заключалось в том, что на обычной квадратной сетке $\sqrt{n} \times \sqrt{n}$ можно найти расстояние $m$ (зависящее от $n$), которое будет встречаться асимптотически чаще, чем $c \cdot n$ раз для любого $c > 0$. В качестве такого значения $m$ берется как раз то, для которого на окружности радиуса $\sqrt{m}$ лежит много целых точек. Эрдёш показал, что при правильном выборе $m$ в зависимости от $n$, можно найти асимптотически $n \cdot \frac{\log n}{\log \log n}$ расстояний, равных $n$, сняв тем самым вопрос о возможности сверхлинейного роста. Остаётся, однако, открытым вопрос о том, можно ли получить рост быстрее, чем $n^\alpha$ для какого-либо $\alpha > 1$.

## 13.2 Решение задачи о сумме двух квадратов

**Утверждение 1.** Для простого числа $p > 2$ следующие три утверждения равносильны:

(1) $p$ имеет вид $4k + 1, k > 0$.

(2) $p$ не является простым элементом кольца гауссовых чисел $\mathbb{Z}[i]$.

(3) $p$ представляется в виде суммы двух квадратов натуральных чисел, притом единственным образом.

*Доказательство.* Легче всего установить равносильность (2) и (3). В самом деле, если $p = x^2 + y^2$, то в гауссовых числах можно записать $p = (x + yi)(x-yi)$. Поскольку $x, y \in \mathbb{N}$, то оба элемента из правой части необратимы, то есть $p$ разложим в гауссовых числах и не является простым элементом кольца.

Если $p = (a + bi)(c + di)$, то $p = (a - bi)(c - di)$ и $p^2 = (a^2 + b^2)(c^2 + d^2)$. Последнее равенство не выходит за пределы целых чисел, поэтому в силу простоты $p$ и необратимости множителей, каждый из них равен $p$. То есть $p = a^2+b^2 = c^2+d^2$. Каждое из чисел $a\pm bi, c\pm di$ является простым в $\mathbb{Z}[i]$, так как имеет норму $p$, поэтому из основной теоремы арифметики получается, что либо $(a+bi) \mid (a-bi), (c+di) \mid (c-di)$, либо $(a+bi) \mid (c-di), (c+di) \mid (a-bi)$.

**Упражнение 1.** $(a + bi) \sim (a - bi) \Leftrightarrow (a + bi) \sim (1 + i)$.

Первый случай влечёт $(a + bi) \sim (a - bi)$, то есть $p = 2$, что противоречит условию. Второй случай влечет $(a + bi) \sim (c + di)$, что означает, что разложения $a^2 + b^2$ и $c^2 + d^2$ совпадают. Аналогично, любое разложение $p$ в сумму двух квадратом совпадает с $a^2 + b^2$.

Так как $p$ вида $4k + 3$ не может быть суммой двух квадратов, то остается доказать, что любое число вида $4k + 1$ раскладывается в гауссовых числах. Для этого нужно доказать, что $-1$ является квадратичным вычетом по модулю $p$, что можно сделать, вычислив символ Лежандра или же воспользовавшись более общей теоремой.

**Теорема 1.** *Пусть $p - 1 = r \cdot l, r, l > 1$. Тогда $0 \neq a \in \mathbb{Z}_p$ является $r$-й степенью ($\exists x : x^r \equiv a \pmod p$) тогда и только тогда, когда $a^l \equiv 1 \pmod p$.*

Так как $p-1 = 2r$, то $-1$ является квадратичным вычетом тогда и только тогда, когда $(-1)^r \equiv 1 \pmod p$, то есть $r = 2k$ и $p = 4k + 1$. Итак, $\exists x : p \mid (x^2+1)$, что в гауссовых числах записывается как $p \mid (x+i)(x-i)$. Если бы $p$ было простым числом, то из этого следовало бы, что $p \mid (x+i)$ или $p \mid (x-i)$. Легко видеть, что это противоречие при $p > 2$, так как если $p \mid (a + bi)$, то $p \mid a$ и $p \mid b$. Итак, число вида $4k + 1$ простым в гауссовых числах быть не может, что завершает доказательство утверждения. □

**Упражнение 2.** Пусть $n = l^2 \cdot p_1 \cdot \ldots \cdot p_m$, где $p_i$ — различные простые. Тогда $n$ представимо в виде суммы двух квадратов, если все $p_i$ имеют вид $4k+1$, притом количество разложений равно $2^{m-1}$.

В связи с тем, что по доказанному простые пары $p$, $p+2$ не могут существовать в $\mathbb{Z}[i]$, можно поставить задачу о простых близнецах в гауссовых числах по-другому.

**Упражнение 3** (задача для исследования)**.** Выяснить, конечно ли число пар простых вида $(a \pm 1) + (b \pm 1)i$ в гауссовых числах.

# 14 Построения циркулем и линейкой

## 14.1 Стандартная постановка задач на построение

Задаваясь вопросом о построимости того или иного геометрического объекта, неоходимо предельно строго сформулировать задачу. К примеру, если ставить вопрос о построимости отрезка длиной $x^2$, если дан отрезок длины $x$, необходимо оговорить, дан ли единичный отрезок. Если, к примеру, дан отрезок длины 1, то важно оговорить, где лежат его конечные точки, потому что если один из его концов имеет координаты $(\sqrt[3]{2}, 0)$, то отрезок длины $\sqrt[3]{2}$ легко построим, однако построить его не удастся, если единичный отрезок дан на оси абсцисс с одним из концов в начале координат.

Устоявшаяся формулировка начальных условий и список разрешённых действий в задачах на построение подразумевает следующие условия:

- Даны координатные оси, перпендикулярные друг другу, и точка их пересечения.
- Слова «дан единичный отрезок» трактуются как «отмечена точка $(1, 0)$».
- Если в процессе построения используется произвольная точка (прямая), то координаты выбираемой точки (коэффициенты уравнения прямой) подразумеваются какими-либо рациональными числами. Это необходимо, так как в результате выбора произвольной точки может быть получена, к примеру, точка $(\pi, 0)$, построить которую в стандартных условиях невозможно.
- Алгоритм построения конечен.
- Одним шагом алгоритма считается одно из следующих действий:
  - Проведение прямой через две уже построенные точки.
  - Проведение окружности с одной из построенных точек в качетсве центра через другую построенную точку.
  - Взятие пересечения двух уже построенных прямых или окружностей или же взятие пересечения уже построенной прямой и уже построенной окружности.

– Взятие «произвольной» точки или прямой в смысле, оговоренном выше.

Только формализовав таким образом круг возможных действий, можно перейти к доказательству содержательных теорем о непостроимости. В их числе: непостроимость отрезка длины $\sqrt[3]{2}$, непостроимость углов величиной $\frac{\pi}{18}$ (что опровергает возможность трисекции угла в 30 градусов) и $\frac{2\pi}{7}$ (опровергает возможность построения правильного 7-угольника), непостроимость отрезка длины $\pi$.

*Замечание.* Стоит оговориться, что поскольку построение любого отрезка длины $x$ эквивалентно построению его на оси абсцисс, то можно вести речь попросту о «построимости числа».

Если рассматривать координатную плоскость как комплексную, то практически все задачи могут быть сформулированы как построение какого-то определенного числа $z \in \mathbb{C}$ или же набора чисел.

## 14.2 Поле построимых чисел

Уточнив постановку задачи, можно сформулировать несколько простых наблюдений. Первое из них состоит в том, что задача о построении некоторой точки на плоскости эквивалентна построению обеих её координат (проекций на оси или любых отрезков, равных проекциям по длине). Второе же заключается в том, что множество чисел, построимых, например, на оси абсцисс замкнуто относительно операций сложения, умножения и обратных к ним (для каждой операции можно поредъявить свое несложное геометрическое построеное), то есть представляет собой поле. Эти два наблюдения подитоживаются следующим предложением.

**Утверждение 2.** Множество построимых комплексных чисел $\widetilde{P}$ является полем и представимо как множество пар построимых вещественных чисел $\{(x, y) \mid x, y \in P\}$, которые также образуют поле.

*Замечание.* Аналогичное утверждение можно сформулировать в случае, если в задаче на уже даны какие-то числа, отрезки или углы. Множество построимых чисел по-прежнему останется полем, структура которого, как выяснится, устроена похожим образом.

## 14.3 Расширения полей. Квадратичные расширения

**Определение 1.** Ситуацию, когда поле $K_1$ является подполем поля $K_2$, называют *расширением* полей. Одно или несколько последовательных расширений $K_1 \subset \ldots \subset K_n$ называют *башней* расширений.

*Замечание.* Важное свойство расширения $K_1 \subset K_2$ состоит в том, что $K_2$ представляет собой линейное пространство над $K_1$. Размерность такого линейного пространства называется *степенью* расширения.

**Определение 2.** Расширение $K_1 \subset K_2$ называется квадратичным, если его степень равна 2 (пишут $[K_2 \colon K_1] = 2$).

**Утверждение 3.** Пусть есть некоторое поле $K$, являющееся для простоты подполем $\mathbb{C}$ и задано уравнение квадратное уравнение $x^2 = a$, которое не имеет решений в $K$. Пусть $\sqrt{a}$ — это какое-то из решений уравнения в $\mathbb{C}$. Тогда минимальное поле $\widetilde{K}$, содержащее $K$ и $\sqrt{a}$ (обозначается $K[\sqrt{a}]$), можно представить как $\widetilde{K} = \{x + y\sqrt{a} \mid x, y \in K\}$, причём все такие линейные комбинации различны.

*Доказательство.* Очевидно, что все такие линейные комбинации должны лежать в $K[\sqrt{a}]$ в силу того, что оно замкнуто и содержит $K$ и $\sqrt{a}$. Достаточно непосредственно проверить, что $\widetilde{K}$ является полем, тогда по стандартному рассуждению оно и будет минимальным.

Если же какие-то из линейных комбинаций $x + y\sqrt{a}$ совпадают, то это бы значило, что $\sqrt{a}$ выражается через элементы $K$, то есть лежит в $K$, что противоречит посылке. □

*Замечание.* Расширение $K \subset K[\sqrt{a}]$ квадратично.

*Замечание.* В доказательстве мы использовали то, что $\frac{x+y\sqrt{a}}{p+q\sqrt{a}} \in \widetilde{K}$. Приём, использующийся для доказательства этого простого утверждения, называется домножением на «сопряжённое».[§]

Следующее наблюдение состоит в том, что любое квадратичное расширение поля $K$ может быть получено добавлением квадратного корня некоторого числа $x \in K$. В самом деле, в качестве базиса в $L \supset K$ могут быть выбраны числа $1, x$, притом известно, что число $(1 + x)(1 - x) = 1 - x^2 \in L$, что означает, существует квадратное уравнение, имеющее $x$ своим корнем. Тогда $x$ лежит в $K[\sqrt{D}]$, где $D$ — дискриминант этого уравнения.

Имея число $a$, простым геометрическим построением можно получить число $\sqrt{a}$, поэтому любое число из любой башни квадратичных расширений сторится циркулем и линейкой.

С другой стороны, в соответствии с описанием алгоритма построения циркулем и линейкой, получение новых точек на каких-то шагах алгоритма происходит с помощью пересечения прямых и окружностей, параметры которых (угловые коэффициенты, центры и радиусы) лежат в некотором поле чисел, которые можно считать уже построенными (в комплексной семантике это минимальное поле, содержащее $\mathbb{C}$ и все точки, отмеченные в ходе алгоритма до рассматриваемого шага).

---

[§] В общем случае операция сопряжения в расширении $K_1 \subset K_2$ — это автоморфизм $K_2$, сохраняющий $K_1$.

Легко убедиться, что единственный нетривиальный автоморфизм в расширении $K \subset K[\sqrt{a}]$ переводит $x + y\sqrt{a}$ в $x - y\sqrt{a}$.

Как устанавливается в теории Галуа, для широкого класса расширений количество таких автоморфизмов совпадает со степенью расширения (если она конечна), а их группа, называемая группой Галуа, заключает в себе много информации о свойствах расширения. В частности со свойствами группы Галуа связана выразимость корней уравнений высших степеней в радикалах.

Самый простой случай — пресечение двух прямых. Легкая выкладка показывает, что точка пересечения двух прямых, заданных уравнениями с коэффициентами в $K$, лежит в $K$. Пересечение прямой и окружности же сводится к решению квадратного уравнения, корни которого лежат в $K[\sqrt{D}]$. Наконец, пересечение двух окружностей может быть сведено к пересечению прямой и окружности, так как разность уравенений вида $(x-a)^2+(y-b)^2 = c^2$ будет линейным уравенинем. Итого, точка, построенная на следующем шаге алгоритма либо лежит в $K$, либо в квадратичном расширении $K$.

Итого, сделанные наблюдения позволяют сформулировать следующее утверждение, характеризующее поле построимых чисел.

**Утверждение 4.** Поле построимых чисел $P$ состоит из всех чисел $\alpha$, для которых $\exists K_0 \subset K_1 \subset \ldots \subset K_n, K_0 = \mathbb{Q}, [K_{i+1} : K_i] = 2, \alpha \in K_n$.

Иными словами, построимы только элементы, лежащие в какой-либо башне квадратичных расширений.

Следующее важное наблюдение описывает строение башен квадратичных расишрений.

**Теорема 2.** *Пусть $K \subset L \subset T$ — двухэтажная башня конечных расширений полей, причем элементы $\{x_1, \ldots, x_{[L:K]}\}$ представляют собой базис $L$ над $K$, а $\{y_1, \ldots, y_{[T:L]}\}$ — базис $T$ над $L$. Тогда расширение $K \subset T$ конечно, имеет базис $\{x_i y_j \mid 1 \leqslant i \leqslant [L : K], 1 \leqslant j \leqslant [T : L]\}$ и степень $[T : K] = [T : L] \cdot [L : K]$.*

**Упражнение 4.** Доказать теорему.

*Замечание.* Простое, но очень важное следствие теоремы: если каждое расширение в башне $K_1 \subset \ldots \subset K_n$ конечно, то $K_1 \subset K_n$ тоже конечно.

Несмотря на простоту, теорема представляет собой мощный инструмент: она позволяет по-другому доказать то, что построимые числа образуют поле, а также, например, то, что полем являются все алгебриаические числа.

Другим немедленным следствием теоремы является такое утверждение:

**Утверждение 5.** Любое построимое число $\alpha \in P$ обладает свойством $[\mathbb{Q}[\alpha] : \mathbb{Q}] = 2^r$ для некоторого натурального $r$.

*Доказательство.* По характеристическому свойству построимых чисел получаем, что существует башня расширений $K_0 \subset \ldots K_n$, $[K_{i+1} : K_i] = 2$, $\alpha \in K_n, K_0 = \mathbb{Q}$. Из того, что $\alpha \in K_n$ следует, что $\mathbb{Q}[\alpha] \subset K_n$, то есть имеет место башня расширений $\mathbb{Q} \subset \mathbb{Q}[\alpha] \subset K_n$. По тоереме о степени расширения, $2^n = [K_n : \mathbb{Q}] = [\mathbb{Q}[\alpha] : \mathbb{Q}] \cdot [K_n : \mathbb{Q}[\alpha]]$, откуда немедленно следует, что $[\mathbb{Q}[\alpha] : \mathbb{Q}]$ является степенью двойки. □

*Замечание.* Утверждение можно использовать как инструмент для доказательства непостроимости каких-либо чисел. Так, если показать, что $[\mathbb{Q}[\alpha] : \mathbb{Q}]$ не является степенью двойки, то из характеристического свойства и утверждения выше немедленно получается, что $\alpha \notin P$.

Теперь задача о построимости числа $\alpha$ практически свелась к задаче о подсчёте степени расширения $\mathbb{Q}[\alpha]$. Мощным инструментом для поиска степени расширения оказывается следующая теорема:

**Теорема 3.** *Пусть $K$ — поле, $f(x) \in K[x]$ — неразложимый многочлен, $\deg f = l, \alpha$ — корень[¶] $f(x)$. Тогда $[K[\alpha] : K] = l$.*

*Доказательство.* Покажем, что $K[\alpha] = \{a_0 + \ldots + a_{l-1}\alpha^{l-1} \mid a_0, \ldots, a_{l-1} \in K\}$. Тогда теорема будет доказана, так как такие выражения $K[\alpha]$ содержать обязано. Осталось показать, что они образуют поле.

Проверка замкнутости относительно сложения и вычитания тривиальна. Для проверки умножения можно без потери общности считать, что старший коэффициент многочлена $f$ равен единице (он не может быть нулём, так как $\deg f = l$). Пользуясь тем, что $f(\alpha) = 0$, можно выразить $\alpha^l$ как линейную комбинацию $1, \alpha, \ldots, \alpha^{l-1}$. Поэтому в произведении двух линейных комбинаций вида $\sum a_i\alpha^i$ от всех степеней выше $l$ можно избавиться.

Осталось проверить наличие обратного по умножению элемента. Для этого как минимум нужно показать, что для никакая нетривиальная линейная комбинация $1, \alpha, \ldots, \alpha^{l-1}$ не равна нулю. Такое равенство повлекло бы существование многочлена степени меньше $l$, у которого $\alpha$ является корнем. Пусть $g(x)$ — многочлен минимальной степени среди всех многочленов, обнуляющих $\alpha, \deg g < l$. Пусть также $f$ даёт остаток $\sigma$ при делении на $g$: $f = gh + \sigma$. Но тогда, так как $f(\alpha) = 0, g(\alpha) = 0$, то $\sigma(\alpha) = 0$. Если $\sigma \not\equiv 0$, то $\deg\sigma < r$ по определению деления с остатком, что противоречит определению $g$. Значит $\sigma \equiv 0$, что в свою очередь влечёт противоречие с неразложимостью $f$.

Таким образом, все линейные комбинации в $K[x]$ различны. Осталось предъявить обратный по умножению элемент к $h(\alpha) = v_0 + \ldots + v_{l-1}\alpha^{l-1}$. Пусть $h(x) = v_0 + \ldots + v_{l-1}x^{l-1}$, тогда, очевидно, что $(f, h) = 1$, так как иначе $f$ разложим. Но в таком случае $\exists g_1, g_2 \in K[x] : f(x)g_1(x)+h(x)g_2(x) \equiv 1$. При подставлении $\alpha$ получается, что $h(\alpha)g_2(\alpha) \equiv 1$, что означает, что $g_2(\alpha)$ и будет обратным к $h(\alpha)$ (если $\deg g_2 \geqslant l$, то от членов с $\alpha$ в степени выше, чем $l - 1$ можно избавиться стандартным способом). □

## 14.4 Примеры непостроимых чисел

**Утверждение 6.** $x^3 - 2$ является неразложимым над $\mathbb{Q}$ многочленом.

*Доказательство.* Пусть это неверно, тогда $x^3 - 2 = (x - a)h(x)$, тогда $a \in \mathbb{Q}$ зануляет левую часть, то есть у $x^3 - 2$ есть рациональный корень, что невозможно. □

Итак, $[\mathbb{Q}[\sqrt[3]{2}] : \mathbb{Q}] = 3$ и число $\sqrt[3]{2}$ непостроимо.

[¶]В этом месте стоит оговориться, откуда берётся $\alpha$. Например, если $K$ является подполем $\mathbb{C}$, то $\alpha$ можно брать из $\mathbb{C}$. В общем же случае можно показать, что существует конструкция поля, являющаяся расширением $K$, в котором у $f$ есть один корень или даже все $l$. Здесь и далее неявно полагается, что $K \subset \mathbb{C}$, однако соответсвующие рассуждения можно провести и в общем случае

**Упражнение 5.** Записать минимальный многочлен для числа $\alpha = \sin \frac{\pi}{18}$ и доказать его неразложимость.

## 14.5 Построимость правильных многоугольников

### 14.5.1 Правильный 7-угольник

Вопрос о построимости правильного $n$-угольника равносилен вопросу о построимости комплексного числа $\xi = e^{\frac{2\pi}{n}}$ с помощью циркуля и линейки. Несложно заметить, что $2\cos\frac{2\pi}{n} = \xi + \xi^{-1}$.

Пусть $n$ нечётно и $\sigma_r = \xi^r + \xi^{-r} = 2\cos(\frac{2\pi r}{n})$ для $r = 1, \ldots, \frac{n-1}{2}$. Для решения вопроса о построимости $\xi$ или, что тоже самое, $\sigma_1$, нужно исследовать строение расширения $\mathbb{Q} \subset \mathbb{Q}[\sigma_1]$.

Сперва можно исследовать арифметические свойства чисел $\sigma_n$ для $n = 7$. Как несложно посчитать, $\sigma_1^2 = \sigma_2 + 2, \sigma_2^2 = \sigma_4 + 2, \sigma_4^2 = \sigma_1 + 2$. Вообще, таблица умножения в $\mathbb{Q}[\sigma_1]$ выглядит следующим образом.

| | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ |
|---|---|---|---|
| $\sigma_1$ | $\sigma_2 + 2$ | $\sigma_1 + \sigma_3$ | $\sigma_2 + \sigma_3$ |
| $\sigma_2$ | $\sigma_1 + \sigma_3$ | $\sigma_4 + 2$ | $\sigma_1 + \sigma_2$ |
| $\sigma_3$ | $\sigma_2 + \sigma_3$ | $\sigma_1 + \sigma_2$ | $\sigma_1 + 2$ |

Также $\sigma_1 + \sigma_2 + \sigma_3 = -1$.

**Упражнение 6.** Пользуясь полученной таблицей, показать, что $\sigma_1$ удовлетворяет уравнению $\sigma_1^3 + \sigma_1^2 - 2\sigma_1 - 1 = 0$.

Таким образом вопрос о построимости правильного семиугольника сведен к вопросу о разложимости многочлена $x^3 + x^2 - 2x - 1$. Однако рациональных корней у него нет (так как числитель рационального корня должен делить свободный член, а знаменатель — старший коэффициент), поэтому приводимым он быть не может и степень расширения $[\mathbb{Q}[\sigma_1] : \mathbb{Q}] = 3$ при $n = 7$. Итак, правильный семиугольник непостроим с помощью циркуля и линейки.

### 14.5.2 Правильный 17-угольник

Как было доказано Гауссом в своё время, для правильного 17-угольника алгоритм построения циркулем и линейкой существует. Поэтому целью здесь будет являться не просто нахождение степени расширения $\mathbb{Q}[\sigma_1]$, а построение конкретной башни квадратичных расширений, последнее из которых содержит $\sigma_1$. Для начала стоит снова немного изучить арифметические свойства чисел $\sigma_i$.

Пусть $\tau_1 = \sigma_1 + \sigma_2 + \sigma_4 + \sigma_8, \tau_2 = \sigma_3 + \sigma_5 + \sigma_6 + \sigma_7, \tau_1 + \tau_2 = -1$.

Тогда $\tau_1^2 = \tau_1 + 8 + 2(\sigma_1 + \sigma_3 + \sigma_3 + \sigma_5 + \sigma_7 + \sigma_8 + \sigma_2 + \sigma_6 + \sigma_6 + \sigma_7 + \sigma_4 + \sigma_5) = \tau_1 + 8 + 2(\tau_1 + 2\tau_2) = 8 + \tau_1 + 2\tau_1 + 4(-1 - \tau_1) = 4 - \tau_1$. Значит, числа $\tau_1$ и $\tau_2$ строятся циркулем и линейкой (лежат в $\mathbb{Q}[\sqrt{17}]$).

Далее нужно разбить $\tau_1$ и $\tau_2$ следующим образом:

$$\tau_1 = \underbrace{\sigma_1 + \sigma_4}_{\beta_1} + \underbrace{\sigma_2 + \sigma_8}_{\beta_2},$$
$$\tau_2 = \underbrace{\sigma_3 + \sigma_5}_{\beta_3} + \underbrace{\sigma_6 + \sigma_7}_{\beta_4}.$$

Число $\beta_1 + \beta_2$ уже построено, поэтому нужно понять, чему равно произведение $\beta_1\beta_2 = (\sigma_1+\sigma_4)(\sigma_2+\sigma_8) = \sigma_1+\sigma_3+\sigma_2+\sigma_6+\sigma_7+\sigma_8+\sigma_4+\sigma_5 = -1$. Тогда по теореме Виетта, числа $\beta_1, \beta_2$ также будут построимыми. Аналогично, $\beta_3\beta_4 = -1$, поэтому все $\beta_i$ будут построимы, притом добавить надо числа $\sqrt{\tau_1^2+4}$ и $\sqrt{\tau_2^2+4}$.

Осталось вычислить, что такое $\sigma_1\sigma_4 = \sigma_3+\sigma_5 = \beta_3$. Тогда сумма и произведение чисел $\sigma_1$ и $\sigma_4$ оказываются уже построены, то есть при расширении поля корнем $\sqrt{\beta_1^2 - 4\beta_3}$, получается поле, содержащее $\sigma_1$.

Итак, алгоритм построения правильного 17-угольника полностью описан.

# Лекция 5. Сложение точек на эллиптических кривых

## 15 Одна школьная задача

**Задача** (Задача 500 из сборника Шарыгина)**.** Существует ли неравнобедренный треугольник, такой, что точки пересечения биссектрисс его углов с противоположными сторонами образуют равнобедренный треугольник?

Оказывается, что такой треугольник на самом деле существует. Нетрудно понять, что если это так, то он будет тупоугольным. Явных же конструкций до некотрого времени придумано не было. Один из подходов к решению: записать соотношение между сторонами в виде уравнения и работать с ним. Такой подход приводит к ответу, однако никакой «наглядности» в таком подходе нет. Другое, совершенно неожиданное решение получается, если рассмотреть правильный 7-угольник. Однако, и в этом подходе ясно, что стороны треугольника иррациональны и являются громоздкими конструкциями из корней разных степеней (будь они квадратны, 7-угольник бы строился циркулем и линейкой). Вопрос, который можно поставить: а существует ли хоть один целочисленный треугольник с таким свойством.

Компьютерный перебор показывает, что среди целых чисел до $10^6$ искомой комбинации нет. С другой стороны, если посмотреть на условие задачи, совершенно ясно, что указанное свойство инвариантно относительно гомотетии, а значит, задача состоит в поиске рациональных точек на какой-то плоской кривой, задаваемой каким-то алгебраическим уравнением $\psi(a, b) = 0$. Может показаться странным, что на всей этой кривой нет рациональных точек и эти сомнения оказываются небеспочвенны.

Вычислив отрезки, на которые разбивают биссектриссы каждую сторону и записав теорему косинусов, можно получить следующее соотношение на стороны треугольника:

$$\cos\varphi = \frac{a^2+c^2-b^2}{2ac}, \cos\psi = \frac{b^2+c^2-a^2}{2bc}$$

$$\left(\frac{ac}{b+c}\right)^2 + \left(\frac{ac}{a+b}\right)^2 - 2\frac{ac}{b+c}\frac{ac}{a+b}\frac{a^2+c^2-b^2}{2ac} = \left(\frac{bc}{a+c}\right)^2 + \left(\frac{bc}{a+b}\right)^2 - 2\frac{bc}{a+c}\frac{bc}{a+b}\frac{b^2+c^2-a^2}{2bc}.$$

**Упражнение 1.** Упростить соотношение до вида:

$$\frac{a}{b+c} + \frac{b}{a+c} = \frac{c}{a+b}$$

У этого соотношения безусловно есть «лишние решения», соответствующие разным вырожденным случаям, равнобедренным треугольникам, а также тем $a, b, c$, которые вообще не задают никакой треугольник.

**Упражнение 2.** Получить другие формы того же соотношения:

- $(a+b+c)(a^2+b^2+c^2)+abc=0$
- $\frac{a(c-a)}{(b+c)^2} = \frac{b(c-b)}{(a+c)^2}$
- $c^3+c^2(a+b) = c(a^2+ab+b^2)+(a^3+a^2b+ab^2+b^3)$

Еще одна параметризация кривой, хорошая тем, что в ней можно не беспокоиться о соблюдении неравенства треугольника, связана с отрезками, на которые разбиваются стороны высотами, проведёнными из точки пересечения биссектрисс. Пусть эти отрезки отрезки равны $x, y, z$, тогда соотношение переписывается с учётом того, что $a = x+y, b = x+z, c = y+z$:

$$4x^3+9x^2y+9x^2z+6xyz+5xy^2+5xz^2-3y^2z-3yz^2=0.$$

## 15.1 Проективная плоскость и проективизация пространства

При изучении однородных уравнений произвольной степени естественным образом получается, что вместе с любым решением появляются все, отличающиеся от него в константное число раз. Путём отождествления всех решений, отличающихся в константное число раз, происходит *проективизация* пространства решений. Таким образом кривые, задаваемые однородными формами, образуют проективное пространство, так как получаются путем проективизации линейного пространства всех форм.

Однако проективные пространства, хоть и представляют собой более точную модель, обладают меньшей геометрической наглядностью. Поэтому часто рассматривают более простую вещь — отождествление лишь тех решений, которые лежат на одном луче, проходящем через точку 0. Полученный объект гораздо более интуитивен, так как представляет собой просто точки на сфере, являющиеся решением данного уравнения. В этом случае, конечно, всегда нужно помнить, что вместе с любым решением входит и диаметрально ему противоположное.

Непосредственно на сфере или в проективной записывать уравнения не слишком удобно. Поэтому используется метод афинных карт. В общем случае, выбирается афинное подпространство (гиперплоскость) размерности на 1 меньше, чем все пространство, не проходящая через 0, и каждой точке сопоставляется её центральная проекция на выбранное подпространство. Естественно, что точки, лежащие в плоскости, параллельной данной и проходящей через 0 не могут быть центрально спроектированы (проектируются на бесконечно удалённую прямую) и не представлены координатно на выбранной карте.

Однако, то, что некоторые решения выпадают из рассмотрения на самом деле не представляет собой большой проблемы. Однородное уравнение, получаемое в случае, если координаты решения лежат в параллельной карте плоскости (для простоты можно представить, что рассматривается карта $z = 1$, что соотвествует делению уравнения на $z$ в соответствующей степени и отдельному рассмотрению случая $z = 0$) имеет не более, чем конечное (ограниченное степенью формы) множество решений.

На этом наблюдении основан один трюк, который оказался довольно успешным при исследовании рациональных точек на произвольных кривых, заданных алгебраическими уравнениями. Стандартное исследование симметрий уравнения $\varphi(x, y) = 0$ подразумевает, например, исследование афинных преобразований, упрощающих вид уравнения. Однако, если добавить дополнительную переменную $z$ и дополнить умеющуюся форму до однородной, домножив каждое слагаемое на $z$ в нужной степени, то множество простых преобразований, с помощью которых можно упрощать уравнение сильно обогащается: в частности очень полезным является домножение на какую-либо матрицу $3 \times 3$. С другой стороны, из предыдущих рассуждений получается, что при таких преобразованиях «потеряно» может быть лишь конечное число решений. Сила проективных преобразовний состоит в том, что многие случаи, возникающие при рассмотрении других групп (евклидовых или афинных преобразований) перестают быть отдельными случаями в проективном случае.

В качестве примера можно привести поиск рациональных точек на кривой, заданной квадратичной формой $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$. Принципиально различны в проективном подходе следующие три случая: форма задаёт две совпадающие прямые, две пересекающиеся прямые или форма не раскладывается в произведение линейных сомножителей и задаёт эллипс. Случаи параллельных прямых и разных видов коник оказываются автоматически разобранными, эквивалентными уже перечисленным (так

как любые две прямых на проективной плоскости пересекаются, а вид коники определяется просто выбором карты).

Строгое утверждение состоит в том, что любое однородное уравнение $Ax^2 + Bxy + Cy^2 + Dxz + Eyz + Fz^2 = 0$, которое не раскладывается в произведение двух линейных форм, может быть переведено преобразованием

$$\begin{pmatrix} \widetilde{x} \\ \widetilde{y} \\ \widetilde{z} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

в уравнение вида $\widetilde{x}\widetilde{z} = \widetilde{y}^2$, которое обладает рациональной параметризацией, то есть отображение $\mathbb{PQ}^2 \to \mathbb{PQ}^3$, заданное формулой $\widetilde{x} = u^2, \widetilde{y} = uv, \widetilde{z} = v^2$, полностью описывает рациональные решения уравнения кроме, быть может, двух.

### 15.2 Получение одной рациональной точки на кривой

Возвращаясь к уравнению конкретной кривой, записанной в координатах $x, y, z$, можно отметить примечательную его особенность: будучи записанной в карте $y, z$, то есть в координатах $y, z$ при $x = 1$, оно принимает вид $4 + 9y + 9z + 6yz + 5y^2 + 5z^2 - 3y^2z - 3yz^2 = 0$. Можно заметить, что по каждой координате в отдельности оно является квадратным.

Тогда применим следующий метод: возьмём какую-либо нетривиальную целую точку, например, подойдет $(y, z) = (-3, 1)$. При $y = -3$ уравнение обращается в квадратное, притом известен его рациональный коренью. Это значит, что найденный второй корень также будет рациональным. Перейдя ко второму корню, можно получить третий, зафиксировав $z$ и найдя соотвествующее $y$. Итерируя этот процесс, можно получать новые рациональные точки на кривой до тех пор, пока не будет обнаружен цикл.

Таким способом на четвертой итерации получается точка $(y_4, z_4)$, причём обе её координаты положительны. Это значит, что она даёт решение исходной задачи. Если выразить его в исходный координатах, получится:

$$a = 1481089,\ b = 18800081,\ c = 19214131.$$

Однако такой метод анализа больше похож на некоторое ad-hoc наблюдение, которое к тому же не позволяет получить характеризацию всех точек на этой кривой, поэтому важно уделить внимание тому, как тот же результат ложится в рамки более общего подхода к анализу эллиптических кривых.

## 16 Эллиптические кривые

### 16.1 Понятие эллиптической кривой, особые точки и точки перегиба

При исследовании кривой, заданной однородной 3-формой $\Phi(a, b, c) = 0$ выделяется несколько случаев вырожденного поведения. Первый из них

заключается в том, что форма $\Phi$ факторизуется, то есть представляется в виде произведения двух форм меньшей степени. Этот случай менее интересен, так как кривая, задаваемая такой формой распадается в объединение прямой и коники или даже просто в три прямые.

Второй, более тонкий вид вырождения можно рассмотреть на примере кривой, задаваемой уравнением $y^2 = x^3$ (проективной формой $\Phi(a, b, c) = ac^2 - b^3$). На аналитическом языке можно сказать, что кривая имеет излом в точке $(0, 0)$, однако в алгебраической геометрии принято формулировать это иначе (это связано с желанием работать в более общем случае, нежели просто над полем $\mathbb{R}$ или $\mathbb{C}$). Можно заметить, что любая прямая, проходящая через $(0, 0, 0)$, является касательной к кривой, то есть $\Phi(a + \alpha t, b + \beta t, c + \gamma t)$ делится на $t^2$ для любых $(\alpha, \beta, \gamma)$. Точка иррегулярности такого рода называется *особой*, а кривые, содержащие особые точки носят название *особых* кривых.

*Замечание.* Поиск особых точек на кривой эквивалентен решению системы уравнений $\frac{\partial \Phi}{\partial a} = \frac{\partial \Phi}{\partial b} = \frac{\partial \Phi}{\partial c} = 0$.

**Упражнение 3.**

- Показать, что форма кривой из задачи про биссектральный треугольник неразложима.
- Показать, что эта кривая не является особой.

**Определение 1** (Основное определение теории элиптических кривых)**.** Кривая $\Phi(a, b, c) = 0$ третьей степени называется *эллиптической* над полем $K$ называется эллиптической, если она неразложима, неособа и существует хотя бы одна точка, все координаты которой лежат в $K$.

*Замечание.* Исследуемая кривая содержит пару рациональных точек $(1, 0, 1)$ и $(0, 1, 1)$, а значит, по упражнению **??** является эллиптической.

Основным инструментом получения точек на эллиптических кривых является следующая операция, называющаяся *сложением* точек: имея две точки на кривой, можно провести через них прямую. Точка её пересечения с эллиптической кривой[‖] будет рациональна, если исходные две точки были таковыми.

В частности, между известными точками $(1, 0, 1)$ и $(0, 1, 1)$ должна находиться одна несовпадающая с ними точка (так как уравнение кривой инвариантно при замене $a$ и $b$). Эта точка $(1, -1, 0)$. Стоит отметить, что эта точка инвариантна относительно замены $a$ и $b$, так как переходит в точку, получающуюся домножением на $-1$, то есть эквивалентна ей в проективном смысле.

[‖] Для эллиптической кривой такая точка будет единственна в силу простого варианта теоремы Гильберта о нулях: если форма степени три принимает значение 0 хотя бы в четырёх точках данной прямой, то она принимает нулевое значение в каждой точке прямой и её форма делится на уравнение прямой в алгебре многочленов, что исключено, так как формы эллиптических кривых неразложимы

Вторым способом получения рациональных точек на кривой является проведение касательной в какой-то известной рациональной точке.

*Замечание.* Точки перегиба кривой задаются уравнением

$$\det \begin{pmatrix} \frac{\partial^2 \Phi}{\partial a^2} & \frac{\partial^2 \Phi}{\partial a \partial b} & \frac{\partial^2 \Phi}{\partial a \partial c} \\ \frac{\partial^2 \Phi}{\partial b \partial a} & \frac{\partial^2 \Phi}{\partial b^2} & \frac{\partial^2 \Phi}{\partial b \partial c} \\ \frac{\partial^2 \Phi}{\partial c \partial a} & \frac{\partial^2 \Phi}{\partial c \partial b} & \frac{\partial^2 \Phi}{\partial c^2} \end{pmatrix} = 0.$$

Указанная матрица называется *гессианом* формы (для 3-формы её определитель также является 3-формой). В точке, где $\det H = 0$, можно найти также направление перегиба $v$, такое, что $Hv = 0$.

**Упражнение 4.** $(1, -1, 0)$ — точка перегиба для данной кривой с направлением перегиба $(1, 1, -4)$.

Если на кривой найдена точка перегиба, то бывает разумно рассмотреть связанный с ней базис: выбрать в качестве первого вектора радиус-вектор точки, направление перегиба взять за второй элемент базиса, а третий вектор достроить перпендикулярно первым двум. В этом базисе точка перегиба перейдет в бесконечно удалённую точку, а вектор перегиба будет сонаправлен с бесконечно удалённой прямой.

**Упражнение 5.** Переписать уравнение кривой в базисе для точки перегиба $(1, -1, 0)$.

Это упражнение на самом деле является частным случаем гораздо более общего факта. А именно, любую эллиптическую кривую можно привести к нормальной форме Вейерштрасса:

$$y^2 + axy + by = x^3 + \alpha x^2 + \beta x + \gamma.$$

В случае, если действие происходит в поле $\mathbb{Q}$, этот вид можно упростить ещё больше, получив выражение:

$$y^2 = x^3 + Ax + B,$$

притом $A$ и $B$ определяются однозначно.

## 16.2 Свойства операции сложения

Пусть операция сложения двух точек $A, B$ на эллиптической кривой обозначена как $A + B$, а точка перегиба обозначена нулём**. Тогда несложно

**Замечание, которое нужно сделать в общем случае для анализа произвольной эллиптической кривой: точки перегиба кривой не обязаны быть рациональными (вещественная точка перегиба у кривой третьего порядка есть всегда). В этом случае операцию сложения, определенную ниже тем не менее можно определить относительно произвольной рациональной точки на кривой, однако, самым простым для анализа случаем является наличие рациональной точки перегиба. Общие результаты, описанные ниже, тем не менее, выполнены и в общем случае

проверить, что $0 + 0 = 0$ и $A + 0 = \widetilde{A}$, где $\widetilde{A}$ — точка, симметричная к $A$ относительно точки перегиба. Следуя естественному желанию превратить точку 0 действительно в 0 какой-то абелевой группы относитльно операции сложния, разумно переопределить сумму двух точек следующим образом:

**Определение 2.** *Суммой* точек $A$ и $B$ на эллиптической кривых называется точка $C$, полученная следующим построением: построить третью точку $D$ на прямой $AB$, и взять в качестве $C$ третью точку на прямой $0D$, где 0 — точка перегиба.

*Замечание.* При таком определении можно в самом деле убедиться, что $0 + 0 = 0, A + 0 = A$, а также, что $\exists \overline{A} : A + \overline{A} = 0$ и $A + B = B + A$. Для группы осталось только проверить ассоциативность сложения.

Две великие по силе и красоте теоремы, полученные около 100 лет назад, дают практически исчерпывающую характеризацию рациональных точек на эллиптических кривых.

**Теорема 1** (Пуанкаре)**.** *Сложение точек на любой эллиптической кривой ассоциативно.*

**Теорема 2** (Морделла)**.** *Абелева группа рациональных точек на любой эллиптической кривой конечнопорождена.*

## 16.3 Примерный анализ группы исследуемой кривой и гипотезы

Пользуясь этими мощными инструментами, можно исследовать данную в задаче эллиптическую кривую. Так, на ней можно найти элемент кручения $D$ : $D + D = 0$ (касательная, проведённая в такой точке должна пересекать кривую в точке её перегиба) и порождающий элемент бесконечного порядка, что даёт право предположить, что группа рациональных точек кривой изоморфна $\mathbb{Z} \oplus (\mathbb{Z}/2\mathbb{Z})$.

Найденные образующие: $A = (1, 0, -1), D = (1, 1, -1)$, порядок первой бесконечен, вторая образует кручение степени 2. Параллель с поиском рациональной точки в других координатах путём отражения с фиксированной одной координатой по-видимому состоит в том, что одна результат $n$-й итерации есть $(2n + 1)A + D$, то есть точка, полученная на 4-й итерации того алгоритма должна быть равна $9A + D$, что верно и проверяется непосредственно.

Другие наблюдения о данной кривой, которые можно развивать: число $6A$ очень близко к $D$. Если к тому же $(12n + 6)A \to D$ при $n \to \infty$, то это может стать отправной точкой для доказательства бесконечности числа целочисленных треугольников, подходящих под условие задачи.

# 17 Доказательство теоремы Пуанкаре

Пусть даны точки $A, B, E$ на эллиптической кривой. Пусть также

- $R$ — третья точка на прямой $A, B$, то есть $A + B + R = 0$
- $L_1$ — прямая, содержащая $A, B, R$ (обозначение $L_1 = ABR$)
- $L_2 = R0\overline{R}$, где $R + \overline{R} = 0$ (ясно, что $\overline{R} = A + B$)
- $L_3 = E\overline{R}S$
- $L_4 = S0\overline{S}$, $\overline{S} = (A + B) + E$
- $M_1 = BEQ$
- $M_2 = Q0\overline{Q}$
- $M_3 = A\overline{Q}T$
- $M_4 = T0\overline{T}, \overline{T} = A + (B + E)$

Таким образом, необходимо показать, что $S = T$ или, что тоже самое $\overline{S} = \overline{T}$ (то есть в доказательстве использовать прямые $L_4$ и $M_4$ не придётся). Итого, получилось 10 точек на кривой: $A, B, E, O, R, \overline{R}, Q, \overline{Q}, S, T$.

**Упражнение 6.** Если $A, B, E, 0$ — различны, то никакие две точки из наборов $\{A, B, \ldots, \overline{Q}\} \cup \{S\}$, $\{A, B, \ldots, \overline{Q}\} \cup \{T\}$ не совпадают.

**Утверждение 1.** Пусть форма $\Phi(x, y, z)$ степени $d$ задаёт кривую $\Phi(x, y, z) = 0$. Пусть также нашлась такая коника, на которой можно выбрать $2d + 1$ точек, лежащих на кривой. Тогда $\Phi(x, y, z)$ делится на уравнение коники как многочлен.

*Доказательство.* Сначала можно показать, что $\Phi(x, y, z) = 0$ на любой точке коники. В самом деле, как было выяснено, в проективных координатах коника приводится к виду $xz = y$ и параметризуется с помощью $u, v$ вектором $(u^2, uv, v^2)$. Тогда, при подстановке в уравнение кривой получается, что однородный многочлен степени $2d$ равен нулю в $2d + 1$ точке, откуда следует, что он есть тождественный ноль.

Тогда многочлен $\Phi$ представляется в виде $\Phi_1 \cdot (zx - y^2) + \Phi_2(x, z) + y\Phi_3(x, z)$ и в силу наложенных ограничений $\Phi_2 \equiv \Phi_3 \equiv 0$ (второе слагаемое содержит только чётные степени $u, v$, третье только нечётные). □

*Замечание.* Совершенно аналогично, если прямая, заданная линейной формой $L = 0$, содержит $d + 1$ точку, на которой $\Phi(x, y, z) = 0$, то $\Phi$ делится на $L$ как многочлен.

Пусть теперь даны $k$ точек $P_1, \ldots, P_k$ и $C^2_{d+2}$-мерное пространство всех форм степени $d$. Тогда подпространство, задаваемое линейными уравнениями на коэффициенты формы $\Phi(P_1) = 0, \ldots, \Phi(P_k) = 0$, имеет размерность $\dim V \geqslant C^2_{d+2} - k$.

**Утверждение 2.** Если точки $P_1, \ldots, P_5$ таковы, что никакие 4 из них не лежат на одной прямой, то пространство всех 2-форм, зануляющихся на этих точках одномерно.

*Доказательство.* Первый случай: если никакие три точки не лежат на одной прямой. Тогда если $\dim V > 1$, то существовали бы две различные коники, обнуляющиеся на этих точках. Тогда так как при $d = 2$ выполнено $2d + 1 = 5$, то их уравнения делятся друг на друга в силу утверждения **??**, что ведёт к противоречию. □

**Упражнение 7.** Разобрать случай, когда три точки лежат на одной прямой.

**Утверждение 3.** Если точки $P_1, \ldots, P_8$ таковы, что никакие 7 из них не лежат на невырожденной конике и никакие 4 из них не лежат на одной прямой. Тогда размерность пространства 3-форм, проходящих через эти точки равна 2.

*Доказательство.* Первый случай: пусть никакие 6 не лежат на невырожденной конике и никакие 3 не лежат на одной прямой. Тогда через первые 5 точек проведем конику (она будет невырождена, так как никакие 3 не лежат на одной прямой). Тогда если размерность пространства форм больше 2, то при наложении ещё два линейных ограничений $\Phi(Q) = 0$, для каких-то двух точек $Q_1, Q_2$ на конике получится пространство размерности не меньше 1, то есть существует кривая степени 3, проходящая через $P_1, \ldots, P_8, Q_1, Q_2$. Тогда так как $2 \cdot 3 + 1 = 7$, то $\Phi = Q \cdot L$, где $Q$ — уравнение коники. Но тогда $L$ принимает нулевое значение на оставшихся 3 из 10 точек, то есть они лежат на одной прямой, противоречие.

Пусть теперь существуют 6 точек, лежащих на невырожденной конике. Тогда если к ним добавить еще одну и доказать, что полученное пространство одномерно, то из этого будет следовать утверждение. Тогда существует кривая второго порядка, проходящая через все эти 9 точек, в том числе через 7 точек коники, значит $\Phi = QL$, где $Q$ —уравнение коники. Оставшаяся часть $L$ проходит через две остальные точки, то есть задана однозначно, значит существует единственная с точностью до пропорциональности форма, проходящая через эти 8 точек, что и доказывает одномерность этого пространства, то есть исходное двумерно.

Последний случай: есть 3 точки на одной прямой. Тогда можно дополнить их четвертой точкой и форма будет делиться на уравнение полученной прямой $L$, то есть $\Phi = QL$, причём форма $Q$ содержит оставшиеся 5 точек, то есть строится однозначно. Аналогично предыдущему пункту, пространство 2-форм, проходящих через исходные 8 точек двумерно. □

Пусть теперь $C$ — исходная эллиптическая кривая, $D_1 = L_1M_2L_3$ (как произведение форм), $D_2 = M_1L_2M_3$. Несложно заметить, что на $D_1$ лежат 9 из точек $A, B, E, O, R, \overline{R}, Q, \overline{Q}, S, T$, притом в их числе точка $S$, на $D_2$ аналогично лежат 9 точек, в том числе $T$. Итак, $C$ и $D_1$ имеют 9 общих точек и, поскольку $C$ невырожденна, больше точек пересечения нет и более того, никакие 7 не лежат на одной конике и никакие 4 не лежат на одной прямой. Значит по утверждению, $C, D_1$ образуют базис в пространстве 3-форм, проходящих через первые 8 точек. Так как $D_2$ тоже проходит через

эти 8 точек, то $D_2 = \lambda C + \mu D_1$. Но точка $S$ лежит на $C$ и на $D_1$, значит и на $D_2$. Итак, 10 точек лежат на трёх прямых $M_1, L_2, M_3$ и значит $S = T$. Доказательство завершено.