Библиотека
«Математическое просвещение»
Выпуск 13

В. О. Бугаенко

# Уравнения Пелля

Издательство Московского центра
непрерывного математического образования
Москва • 2001

УДК 511.512
ББК 22.13
Б90

## Аннотация

Уравнения Пелля представляют собой класс диофантовых уравнений второй степени. Они связаны со многими важными задачами теории чисел. Решение уравнений Пелля — задача непростая, хотя и выполнимая методами элементарной математики. Ключевую роль в исследовании этих уравнений играет геометрическая лемма Минковского о выпуклом теле. Эта лемма неожиданно возникает во многих задачах теории чисел и является одним из ярких примеров связи алгебры и геометрии.

Основной результат, которому посвящена брошюра, — полное описание решений уравнений Пелля.

Текст брошюры представляет собой обработанную и расширенную запись двух лекций, прочитанных автором 19 февраля и 15 апреля 2000 года на малом мехмате МГУ для школьников 9–11 классов.

Брошюра рассчитана на широкий круг читателей, интересующихся математикой: школьников старших классов, студентов младших курсов, учителей...

ISBN 5–900916–96–0

*Бугаенко Вадим Олегович.*

Уравнения Пелля.

(Серия: «Библиотека „Математическое просвещение“»).
М.: МЦНМО, 2001. — 32 с.: ил.

Главный редактор серии *В. М. Тихомиров.*

Заведующая редакцией *В. Л. Браккер.*

Редактор *Р. М. Кузнец.* Техн. редактор *М. Ю. Панов.*

---

Лицензия ИД № 01335 от 24/III 2000 года. Подписано к печати 14/VI 2001 года. Формат бумаги $60 \times 88\ ^1/_{16}$. Офсетная бумага № 1. Офсетная печать. Физ. печ. л. 2,00. Усл. печ. л. 1,96. Уч.-изд. л. 1,86. Тираж 3000 экз. Заказ 6078.

---

Издательство Московского центра непрерывного математического образования.
121002, Москва, Г-2, Бол. Власьевский пер., 11. Тел. 241–05–00.

---

Отпечатано в Производственно-издательском комбинате ВИНИТИ.
140010, г. Люберцы Московской обл., Октябрьский пр-т, 403. Тел. 554–21–86.

## Диофантовы уравнения

В разных областях математики, а в особенности в алгебраической геометрии и теории чисел, встречаются уравнения, обе части которых представляют собой многочлены, а решения требуется найти в целых числах. Такие уравнения называются *диофантовыми*. Среди диофантовых уравнений встречаются как простые, легко решаемые элементарными методами, так и те, решения которых требуют применения современных математических теорий. Примером последних служат знаменитые уравнения Ферма

$$x^n + y^n = z^n, \quad n > 2,$$

которые человечество пыталось решить на протяжении более трёх веков, и лишь несколько лет назад английскому математику Эндрю Уайлсу удалось доказать, что они не имеют решений в целых положительных числах.

## Что такое уравнения Пелля?

Уравнения Пелля — это уравнения вида

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

где $m$ — целое положительное число, не являющееся точным квадратом. Они представляют собой класс диофантовых уравнений второй степени и связаны со многими важными задачами теории чисел. Решение уравнений Пелля — задача непростая, хотя и выполнимая методами элементарной математики. Полное описание решений этих уравнений и является нашей конечной целью. Попутно нам встретятся некоторые математические понятия и теоремы, которые на первый взгляд могут показаться не связанными с основной темой. Большинство из них представляют интерес не только как инструмент для исследования уравнений Пелля, но и сами по себе.

Сначала мы разберём, как решаются линейные диофантовы уравнения. Формально, эти результаты не будут использованы при решении основной задачи. Но некоторая тренировка в более простой ситуации поможет проще ориентироваться в последующих, достаточно сложных математических рассуждениях. При решении линейных диофантовых уравнений ключевую роль играет алгоритм Евклида, главное назначение которого — нахождение наибольшего общего делителя (НОД) целых чисел. Затем мы узнаем, что арифметические операции могут быть определены не только на числовых множествах, но и на более экзотических математических объектах, например на множестве точек плоскости. Попутно нам понадобится лемма Минковского о выпуклом теле — красивый геометрический факт, неожиданным образом возникающий при решении многих теоретико-числовых задач. Наконец, мы познакомимся с основами теории цепных дробей, которые помогут нам отыскивать решения уравнений Пелля.

Таким образом, двигаясь к нашей основной цели, мы коснёмся нескольких математических сюжетов. Мы не будем останавливаться на них подробно, но снабдим некоторые из них задачами для самостоятельного решения. В большинстве эти задачи не связаны с основной темой книги, однако в нескольких случаях их результаты используются в дальнейших доказательствах. В конце брошюры ко всем задачам приведены решения и указания.

Прежде всего сделаем два замечания. Во-первых, при любом $m$ уравнение (1) имеет по крайней мере два решения: $x = \pm 1$, $y = 0$. Эти решения мы назовём *тривиальными*. Во-вторых, поскольку при изменении знака у $x$ или $y$ левая часть уравнения (1) не изменится, достаточно ограничиться нахождением только неотрицательных решений (т. е. решений с неотрицательными $x$ и $y$).

Решая уравнение Пелля, мы будем отвечать на три вопроса.

1) Существует ли хотя бы одно нетривиальное решение?
2) Если да, то как его найти?
3) Как описать все решения?

Отвечать на эти вопросы нам будет удобнее в другом порядке. Начнём мы с последнего: в предположении, что одно решение уже найдено, мы покажем, как найти все. Их окажется бесконечно много. Затем перейдём к первому вопросу, а именно докажем, что уравнение Пелля всегда имеет нетривиальное решение. И наконец, мы покажем, как можно найти это решение.

Заметим, что ограничение на параметр $m$ является естественным. Если $m$ — точный квадрат, то уравнение (1) не имеет нетривиальных решений. Действительно, разность двух точных квадратов в левой части может равняться единице, только если первый из них равен единице, а второй — нулю.

## Пример: уравнение $x^2 - 2y^2 = 1$

Решим для начала уравнение Пелля при $m = 2$.

Несложная выкладка показывает, что если пара $(x, y)$ является решением рассматриваемого уравнения, то пара $(3x + 4y,\ 2x + 3y)$ тоже его решение. Действительно,

$$(3x+4y)^2-2(2x+3y)^2=(9x^2+24xy+16y^2)-\\ -2(4x^2+12xy+9y^2)=x^2-2y^2.$$

Поэтому если $x^2-2y^2=1$, то и $(3x+4y)^2-2(2x+3y)^2=1$.

Значит, исходя из тривиального решения $x_0=1$, $y_0=0$, мы можем получить бесконечную последовательность (нетривиальных) решений $(x_i,y_i)$ с помощью рекуррентной формулы $(x_i,y_i)=f(x_{i-1},y_{i-1})$, где $f(x,y)=(3x+4y,\ 2x+3y)$. Вот несколько первых её членов:

$$(3,2),\quad (17,12),\quad (99,70),\quad (577,408).$$

Докажем теперь, что этой последовательностью $(x_i,y_i)$ исчерпываются все неотрицательные решения уравнения. Тогда описание всех его решений можно будет считать завершённым.

Неотрицательные решения уравнения Пелля можно естественным образом упорядочить. Для этого рассмотрим на координатной плоскости множество точек, удовлетворяющих соотношению $x^2-2y^2=1$, лежащих в первой координатной четверти. Это — график функции $y=\sqrt{\frac{x^2-1}{2}}$, определённой при $x\geqslant 1$ (рис. 1). Будем говорить, что точка на этом графике тем больше, чем дальше она находится от точки $(1,0)$. Поскольку функция монотонна, большей из двух точек графика будет та, у которой больше как абсцисса, так и ордината. Неотрицательные решения уравнения Пелля суть целые точки на графике. Поэтому неравенство $(x',y')<(x'',y'')$ для двух различных неотрицательных решений $(x',y')$ и $(x'',y'')$ означает, что $x'<x''$ (и, тем самым, $y'<y''$).

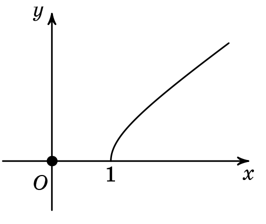


Рис. 1

Отображение $f$ является монотонным относительно введённого упорядочения. Действительно, для неотрицательных $x'$, $y'$, $x''$ и $y''$ из неравенств $x'<x''$ и $y'<y''$ очевидным образом следуют неравенства $3x'+4y'<3x''+4y''$ и $2x'+3y'<2x''+3y''$. Любое монотонное отображение имеет обратное, также являющееся монотонным. Обратным к $f$ является отображение $g(x,y)=(3x-4y,\ 3y-2x)$. Оно также переводит любое решение уравнения в решение.

Предположим, что существует решение $(x',y')$ уравнения $x^2-2y^2=1$, не совпадающее ни с одним из членов построенной последовательности $(x_i,y_i)$. Поскольку $x_i$ и $y_i$ неограниченно возрастают, решение $(x', y')$ лежит между какими-то двумя решениями из последовательности: $(x_i, y_i) < (x', y') < (x_{i+1}, y_{i+1})$. Применяя к этому двойному неравенству монотонное отображение $g$ последовательно $i$ раз, получим $(x_0, y_0) < g^i(x', y') < (x_1, y_1)$, где $g^i(x', y')$ тоже решение уравнения. Однако нетрудно убедиться, что между решениями $(1, 0)$ и $(3, 2)$ других решений рассматриваемое уравнение не имеет. Полученное противоречие доказывает, что любое неотрицательное решение принадлежит последовательности $(x_i, y_i)$.

Аналогичным образом можно описать решения и других уравнений Пелля. Для этого достаточно найти аналог отображения $f$ для произвольного параметра $m$. Это отображение должно переводить любое неотрицательное решение уравнения Пелля в другое неотрицательное решение и быть монотонным на множестве неотрицательных решений уравнения.

Прежде чем приступить к решению уравнений Пелля в общем случае, разберёмся, как решаются более простые диофантовы уравнения, линейные.

## Линейные диофантовы уравнения

Линейные диофантовы уравнения — это диофантовы уравнения вида

$$ax + by = c, \tag{2}$$

где $a$, $b$ и $c$ — некоторые целые числа, причём $a$ и $b$ не равны нулю одновременно.

Ответим сначала на вопрос, имеет ли линейное уравнение хотя бы одно решение. Обозначим через $d$ наибольший общий делитель чисел $a$ и $b$. Если число $c$ не делится на $d$, то решений нет, поскольку при любых $x$ и $y$ левая часть (2) делится на $d$, а правая — не делится.

Пусть теперь $c = kd$. В этом случае решение существует. Чтобы это доказать, достаточно показать, что имеет решение уравнение

$$ax + by = d. \tag{3}$$

Действительно, умножив решение (т. е. каждое из чисел $x$ и $y$) уравнения (3) на $k$, получим решение уравнения (2). Один из методов нахождения решения уравнения (3) основан на алгоритме Евклида.

## Алгоритм Евклида

Алгоритм Евклида служит для нахождения наибольшего общего делителя двух целых положительных чисел. Он основан на следующем простом наблюдении. Если $a = bq + r$ (где $q$ — частное, а $r$ — остаток от деления $a$ на $b$), то НОД$(a, b)$ = НОД$(b, r)$. Действительно, из формулы деления с остатком следует, что любой общий делитель чисел $b$ и $r$ является также делителем числа $a$, а любой общий делитель чисел $a$ и $b$ является также делителем числа $r$. Поэтому множества общих делителей пар чисел $(a, b)$ и $(b, r)$ совпадают, а значит, совпадают и их наибольшие общие делители.

Применение алгоритма Евклида заключается в последовательном делении с остатком. Сначала мы делим большее из двух чисел на меньшее. На каждом следующем шаге мы делим число, которое на предыдущем шаге было делителем, на число, которое было остатком. Так поступаем до тех пор, пока не получим нулевой остаток. Это обязательно произойдёт через конечное число шагов, поскольку остатки всё время уменьшаются. Последний ненулевой остаток и будет наибольшим общим делителем исходных чисел.

Алгоритм Евклида, применённый к паре чисел $(a, b)$, где $a > b$, может быть записан в виде цепочки равенств

$$\begin{aligned} a &= bq_1 + r_1, \\ b &= r_1q_2 + r_2, \\ r_1 &= r_2q_3 + r_3, \\ &\dots\dots\dots\dots\dots \\ r_{n-3} &= r_{n-2}q_{n-1} + r_{n-1}, \\ r_{n-2} &= r_{n-1}q_n + r_n, \\ r_{n-1} &= r_nq_{n+1}. \end{aligned} \tag{4}$$

Тогда $d$ = НОД$(a, b)$ = НОД$(b, r_1)$ = ... = НОД$(r_{n-1}, r_n) = r_n$.

**1.** Докажите, что если $m$ и $n$ — различные неотрицательные целые числа, то числа $2^{2^m} + 1$ и $2^{2^n} + 1$ являются взаимно простыми.

Покажем теперь, как найти решение уравнения (3) с помощью цепочки равенств (4). Выразим $d = r_n$ из предпоследнего равенства. В полученное выражение подставим значение $r_{n-1}$ из предыдущего равенства и т. д. Продолжая этот процесс, в конце концов получим выражение $d$ в виде левой части уравнения (3).

Для примера рассмотрим уравнение $355x + 78y = 1$. Сначала найдём наибольший общий делитель чисел 355 и 78 с помощью алгоритма

---

Двумя чертами слева выделены тексты задач для самостоятельного решения. Задачи повышенной сложности отмечены знаком *.

Евклида:

$$\begin{aligned} 355 &= 78 \cdot 4 + 43, \\ 78 &= 43 \cdot 1 + 35, \\ 43 &= 35 \cdot 1 + 8, \\ 35 &= 8 \cdot 4 + 3, \\ 8 &= 3 \cdot 2 + 2, \\ 3 &= 2 \cdot 1 + 1, \\ 2 &= 1 \cdot 2. \end{aligned}$$

Итак, $\text{НОД}(355, 78) = 1$. Перепишем полученные равенства в виде

$$\begin{aligned} 43 &= 355 - 78 \cdot 4, \\ 35 &= 78 - 43 \cdot 1, \\ 8 &= 43 - 35 \cdot 1, \\ 3 &= 35 - 8 \cdot 4, \\ 2 &= 8 - 3 \cdot 2, \\ 1 &= 3 - 2 \cdot 1. \end{aligned}$$

Теперь пойдём по цепочке равенств снизу вверх:

$$\begin{aligned} 1 &= 3 - 2 \cdot 1 = 3 - (8 - 3 \cdot 2) \cdot 1 = 3 \cdot 3 - 8 = (35 - 8 \cdot 4) \cdot 3 - 8 = \\ &= 35 \cdot 3 - 8 \cdot 13 = 35 \cdot 3 - (43 - 35 \cdot 1) \cdot 13 = 35 \cdot 16 - 43 \cdot 13 = \\ &= (78 - 43 \cdot 1) \cdot 16 - 43 \cdot 13 = 78 \cdot 16 - 43 \cdot 29 = \\ &= 78 \cdot 16 - (355 - 78 \cdot 4) \cdot 29 = 78 \cdot 132 - 355 \cdot 29. \end{aligned}$$

Найдено решение: $x = -29$, $y = 132$.

Нам осталось ответить на вопрос: как, зная одно решение линейного диофантова уравнения, найти все? Чтобы ответить на него, разберём сначала один пример.

### Пример: уравнение $3x + 5y = 22$

Подбором легко найти одно из решений этого уравнения, например, $x = 4$, $y = 2$. (Применение алгоритма Евклида приводит к другому решению: $x = 44$, $y = -22$.)

Нарисуем на координатной плоскости *график уравнения* — множество точек $(x, y)$, таких что $x$ и $y$ удовлетворяют этому уравнению. Этот график представляет собой прямую (рис. 2), которую мы обозначим через $l$. На ней нужно найти все точки с целыми координатами (которые мы для простоты будем называть *целыми точками*). Из рисунка видно, что точки $(4, 2)$ и $(-1, 5)$ лежат на прямой $l$ и на отрезке между этими точками других целых точек нет. Конечно, ссылка на график не является доказательством. Однако доказать это несложно: непосредственно проверяем, что указанные пары чисел удовлетворяют уравнению, а при $y = 3$ и $y = 4$ значения $x$, получаемые из уравнения, оказываются нецелыми.

Заметим, что если пара $(x_0, y_0)$ — решение уравнения, то пара $(x_0 - 5,\ y_0 + 3)$ тоже решение:

$$3(x_0 - 5) + 5(y_0 + 3) = 3x_0 - 15 + 5y_0 + 15 = 3x_0 + 5y_0 = 22.$$

Преобразование

$$(x, y) \mapsto (x - 5,\ y + 3), \tag{5}$$

во-первых, сохраняет прямую $l$ (оно является параллельным переносом вдоль неё), во-вторых, переводит целые точки в целые, и поэтому любое решение уравнения переводит в решение.

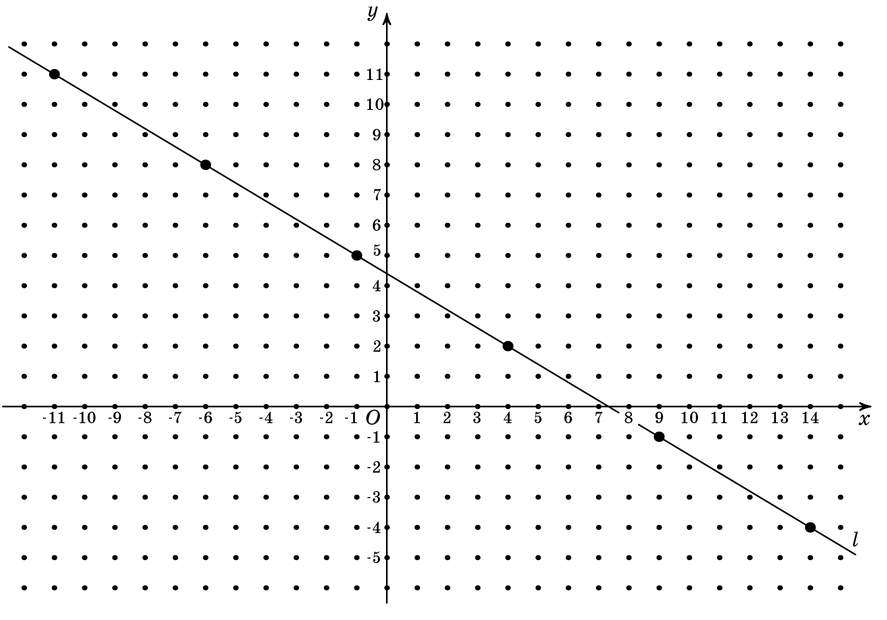



Рис. 2

Применяя к уже найденному решению преобразование (5), т. е. прибавляя к $x$ число $(-5)$, а к $y$ число 3, мы получим ещё одно решение, потом ещё одно и т. д. Точки, соответствующие этим решениям, располагаются на прямой $l$ через равные расстояния (см. рис. 2). Ясно, что можно двигаться и в обратную сторону.

Таким образом, мы нашли бесконечную серию решений

$$x = 4 - 5t, \quad y = 2 + 3t,$$

где $t$ — любое целое число. Докажем, что любое решение имеет такой вид. Рассуждаем от противного: пусть на прямой $l$ между точками $(4-5t,\ 2+3t)$ и $(4-5(t+1),\ 2+3(t+1))$ найдётся целая точка. Применяя несколько раз параллельный перенос (5), если $t < 0$, или обратный ему перенос, если $t > 0$, мы получим, что на интервале между точками $(4, 2)$ и $(-1, 5)$ тоже есть целая точка. Противоречие.

Итак, мы нашли все решения уравнения и доказали, что других нет.

## Общее решение линейного диофантова уравнения

Перейдём теперь к рассмотрению общего случая. Рассмотрим графики уравнений (2) при фиксированных $a$ и $b$ и пробегающем всевозможные целые значения параметре $c$. Это — бесконечное множество параллельных прямых $l_c$, причём каждая целая точка принадлежит ровно одной такой прямой (рис. 3).

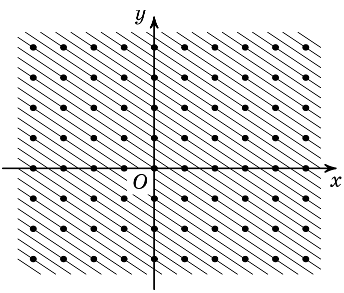


Рис. 3

Введём на координатной плоскости «сложение» точек:

$$(x_1,\ y_1) + (x_2,\ y_2) = (x_1 + x_2,\ y_1 + y_2).$$

Введённая операция сложения точек определена также на подмножестве целых точек. Действительно, сумма двух целых точек является целой точкой. Эта операция имеет естественную геометрическую интерпретацию: она соответствует сложению векторов с началами в начале координат и концами в данных точках. Как и обычное сложение, сложение точек имеет обратную операцию — вычитание.

Мы отметим следующее легко проверяемое свойство сложения: если две точки лежат на прямых $l_m$ и $l_n$ соответственно, то их сумма лежит на прямой $l_{m+n}$, а разность — на прямой $l_{m-n}$. Из этого следует, что если мы к одной целой точке, лежащей на прямой $l_c$ прибавим поочерёдно все целые точки, лежащие на прямой $l_0$, то получим все целые точки на прямой $l_c$. Иными словами, чтобы найти общее решение уравнения (2), нужно к его частному решению добавить общее решение уравнения

$$ax + by = 0. \qquad (6)$$

Осталось решить это уравнение. Снова введём $d = \text{НОД}(a, b)$. Тогда $a = a'd$, $b = b'd$, где $\text{НОД}(a', b') = 1$. Уравнение (6) перепишется в виде

$$a'x = -b'y,$$

а так как $a'x$ делится на $b'$ и $a'$ взаимно просто с $b'$, то $x$ делится на $b'$. Значит, $x = b't$, откуда $y = -a't$.

Теперь можем резюмировать. Если $c$ не делится на $\text{НОД}(a, b)$, то уравнение (2) решений не имеет. Если же $c$ делится на $\text{НОД}(a, b)$, то уравнение (2) имеет бесконечно много решений, получаемых по формулам

$$x = x_0 + \frac{b}{d}t, \qquad y = y_0 - \frac{a}{d}t,$$

где $t$ — произвольное целое число, $d = \text{НОД}(a, b)$, а $(x_0, y_0)$ — частное решение, которое может быть найдено с помощью алгоритма Евклида.

## График уравнения Пелля

Приступая к рассмотрению уравнения Пелля, прежде всего нарисуем его график. График уравнения $x^2 - my^2 = 1$ — это гипербола (рис. 4), асимптотами которой являются прямые $y = \pm\frac{x}{\sqrt{m}}$. Чтобы убедиться в этом, разложим левую часть уравнения на множители:

$$(x - \sqrt{m}\,y)(x + \sqrt{m}\,y) = 1,$$

и введём новую (косоугольную) систему координат, направив ось $Ox'$ вдоль прямой $y = -\frac{x}{\sqrt{m}}$, а ось $Oy'$ — вдоль прямой $y = \frac{x}{\sqrt{m}}$. В системе координат $Ox'y'$ уравнение нашей кривой запишется в привычном виде: $x'y' = \text{const}$.

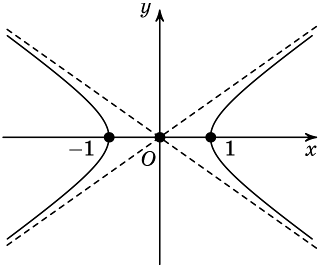


Рис. 4

При любом $m$ гипербола проходит через точку $(1, 0)$ и симметрична относительно обеих координатных осей.

Вместе с гиперболой $x^2 - my^2 = 1$ рассмотрим серию кривых $l_n$, задаваемых уравнениями

$$x^2 - my^2 = n, \tag{7}$$

где $n$ — всевозможные целые числа (рис. 5). Кривые $l_n$ при $n \neq 0$ представляют собой гиперболы, а $l_0$ — это пара прямых $y = \pm\frac{x}{\sqrt{m}}$, являющихся общими асимптотами этого семейства гипербол.

Поскольку для любой целой точки величина $x^2 - my^2$ является целым числом, каждая целая точка попадает на один из графиков $l_n$. Так как $m$ не является точным квадратом, на $l_0$ (паре асимптот) лежит лишь начало координат. Все остальные целые точки лежат на гиперболах.

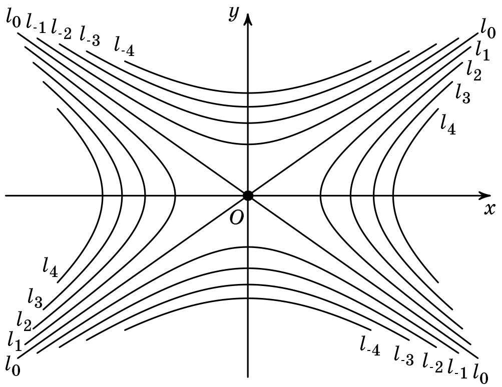



Рис. 5

С каждой гиперболой $l_n$ связана *сопряжённая* ей гипербола $l_{-n}$. Если мы выберем на одной из гипербол пару центрально-симметричных относительно начала координат точек, то на сопряжённой гиперболе можно выбрать такую пару центрально-симметричных точек, чтобы все четыре точки были вершинами параллелограмма со сторонами, параллельными асимптотам. Такие пары точек будем называть *сопряжёнными* друг другу. Действительно, если в системе координат, оси которой идут вдоль асимптот, пара симметричных точек имеет координаты $(x', y')$ и $(-x', -y')$, то сопряжённой ей в той же системе координат является пара симметричных точек $(x', -y')$ и $(-x', y')$.

## Умножение точек

Когда мы рассматривали линейные диофантовы уравнения, мы ввели «сложение» целых точек плоскости, и эта операция помогла нам разобраться, как устроены решения таких уравнений.

Сейчас мы введём новую операцию на множестве целых точек плоскости. Эту операцию мы назовём «умножением», и она нам поможет решить уравнения Пелля. Точнее, мы введём не одну, а бесконечно много операций, параметризованных целым положительным параметром $m$, который, как и в уравнениях Пелля, не должен являться точным квадратом. Вот определение умножения точек:

$$(x_1, y_1)\cdot(x_2, y_2) = (x_1 x_2 + m y_1 y_2,\ x_1 y_2 + x_2 y_1). \tag{8}$$

Введённая операция обладает теми же свойствами, что и обычное умножение чисел: коммутативностью, ассоциативностью и дистрибутивностью по отношению к сложению, введённому нами раньше. Эти свойства можно доказать непосредственной проверкой, однако мы не будем торопиться выписывать длинные цепочки тождеств. Доказательства безо всяких выкладок будут получены из алгебраической интерпретации нашего умножения, которую мы рассмотрим чуть позже.

Если точка $(x_1, y_1)$ лежит на кривой $l_n$, а $(x_2, y_2)$ — на кривой $l_k$, то их произведение $(x_1, y_1)\cdot(x_2, y_2)$ лежит на кривой $l_{n\cdot k}$. Действительно,

$$(x_1x_2 + my_1y_2)^2 - m(x_1y_2 + y_1x_2)^2 = x_1^2x_2^2 + 2mx_1x_2y_1y_2 + m^2y_1^2y_2^2 -$$
$$- mx_1^2y_2^2 - 2mx_1x_2y_1y_2 - my_1^2x_2^2 = (x_1^2 - my_1^2)(x_2^2 - my_2^2) = n\cdot k.$$

В частности, перемножая точки, лежащие на $l_1$, мы будем получать точки на $l_1$. Иными словами, произведение решений уравнения Пелля также является его решением. Умножение на неотрицательное тривиальное решение новых решений не даст, поскольку точка $(1, 0)$ играет для нашего умножения роль единицы: умножая на неё любую точку, мы получим её саму.

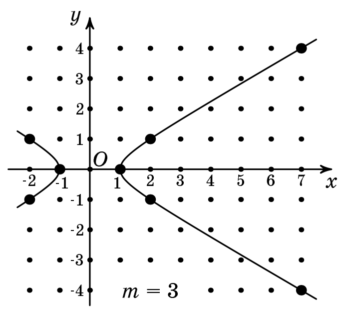


Рис. 6

### Общее решение уравнения Пелля

Если уравнение Пелля имеет хотя бы одно нетривиальное решение, то, умножая его многократно на себя, можно найти бесконечно много решений. При этом все решения можно найти аналогично тому, как мы действовали в частном случае $m = 2$. Двигаясь по графику уравнения (рис. 6) из точки $(1, 0)$ в направлении положительных значений $y$, находим первое нетривиальное решение. Это решение назовём *основным*.

**Теорема 1.** Все нетривиальные положительные решения получаются многократным умножением основного решения на себя.

Д о к а з а т е л ь с т в о. Рассмотрим последовательность $(x_1, y_1)$, $(x_2, y_2)$, ..., $(x_n, y_n)$, ... решений, получаемых из основного решения $(x_1, y_1)$ последовательным умножением на него. Предположим, что на графике уравнения между двумя её членами $(x_n, y_n)$ и $(x_{n+1}, y_{n+1})$ имеется некоторое решение. Умножив его на $(x_1, -y_1)$, получим новое решение, лежащее между $(x_{n-1}, y_{n-1})$ и $(x_n, y_n)$. Действительно, умножение на $(x_1, -y_1)$ является обратной операцией к умножению на $(x_1, y_1)$. Проделав такую операцию $n$ раз, получим решение, лежащее между $(1, 0)$ и $(x_1, y_1)$. Это противоречит тому, что $(x_1, y_1)$ — основное решение. □

Теперь перейдём к доказательству существования нетривиального решения любого уравнения Пелля. Для этого нам понадобится ввести деление точек — операцию, обратную умножению. Но прежде рассмотрим две интерпретации операции умножения точек, геометрическую и алгебраическую.

### Гиперболический поворот

Пусть $(x_0, y_0)$ — некоторое фиксированное нетривиальное решение уравнения Пелля. Рассмотрим преобразование, которое произвольную точку $(x, y)$ переводит в точку $(x_0 x + m y_0 y,\ x_0 y + y_0 x)$. Это преобразование — умножение на нетривиальное решение уравнения Пелля — любую точку на гиперболе $l_n$ переводит в точку на ней же, т. е. каждую гиперболу семейства сдвигает вдоль себя (и поэтому называется *гиперболическим поворотом*). При этом целые точки переходят в целые точки.

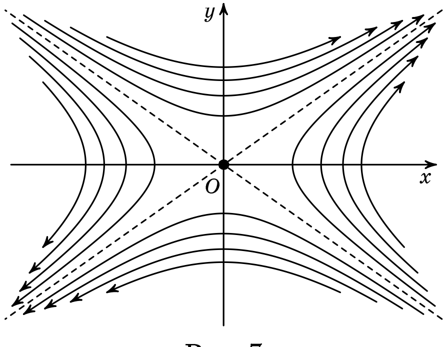


Рис. 7

Направления, в которых сдвигаются гиперболы, указаны на рис. 7 (предполагается, что начальное решение $(x_0, y_0)$ положительно).

При рассмотрении линейных диофантовых уравнений аналогичную роль играли параллельные переносы на целочисленные векторы вдоль прямой — графика уравнения.

**2.** К о н т р о л ь н ы й в о п р о с. Как указанное преобразование действует на асимптотах?

Теперь предложим для самостоятельного решения серию диофантовых уравнений второй степени, родственных уравнениям Пелля.

**3.** Докажите, что целые неотрицательные числа $x$, $y$ удовлетворяют уравнению $x^2-mxy+y^2=1$ (где $m$ — данное целое число) тогда и только тогда, когда $x$ и $y$ — соседние числа последовательности $\varphi_0=0$, $\varphi_1=1$, $\varphi_2=m$, $\varphi_3=m^2-1$, $\varphi_4=m^3-2m$, $\varphi_5=m^4-3m^2+1$, ..., в которой $\varphi_{k+1}=m\varphi_k-\varphi_{k-1}$ для всех $k\geqslant 1$.

### Квадратичные иррациональности

Рассмотрим множество $\mathbb{Z}[\sqrt{m}]$ чисел, представимых в виде $x+y\sqrt{m}$, где $x$ и $y$ — целые числа. Нетрудно проверить, что сумма, разность и произведение чисел из $\mathbb{Z}[\sqrt{m}]$ также лежат в $\mathbb{Z}[\sqrt{m}]$:

$$(x_1+y_1\sqrt{m})\pm(x_2+y_2\sqrt{m})=(x_1\pm x_2)+(y_1\pm y_2)\sqrt{m},$$
$$(x_1+y_1\sqrt{m})\cdot(x_2+y_2\sqrt{m})=(x_1x_2+my_1y_2)+(x_1y_2+x_2y_1)\sqrt{m}.$$

Мы видим, что если поставить в соответствие числу $x+y\sqrt{m}$ точку с координатами $(x,y)$, то операции сложения и умножения над числами будут соответствовать введённым нами операциям сложения и умножения целых точек на плоскости. Поскольку умножение чисел коммутативно, ассоциативно и дистрибутивно относительно сложения, этими же свойствами обладает и умножение точек.

Заметим, что указанное соответствие между числами из $\mathbb{Z}[\sqrt{m}]$ и целыми точками на плоскости является *взаимно однозначным*: предположение, что двум различным точкам $(x_1,y_1)$ и $(x_2,y_2)$ соответствует одно и то же число, приводит к равенству $x_1+y_1\sqrt{m}=x_2+y_2\sqrt{m}$, откуда $\sqrt{m}=\dfrac{x_1-x_2}{y_2-y_1}$. Это невозможно, поскольку число $\sqrt{m}$ является иррациональным.

Для каждого числа $x+y\sqrt{m}\in\mathbb{Z}[\sqrt{m}]$ можно определить *сопряжённое число* $\overline{x+y\sqrt{m}}=x-y\sqrt{m}$, а также *норму* $N(x+y\sqrt{m})=x^2-my^2$, равную произведению числа на его сопряжённое. Заметим, что норма числа — это номер гиперболы, на которую попадает соответствующая ему целая точка. То, что при перемножении точек номера́ гипербол, на которых они лежат, перемножаются, теперь можно сформулировать так: норма произведения двух чисел равна произведению их норм.

В этих терминах решить уравнение Пелля значит найти все числа из $\mathbb{Z}[\sqrt{m}]$ с нормой единица.

Результат деления двух чисел $x_1+y_1\sqrt{m}$ и $x_2+y_2\sqrt{m}$ из $\mathbb{Z}[\sqrt{m}]$ не всегда лежит в $\mathbb{Z}[\sqrt{m}]$. Это происходит только в случае, если $x_1x_2-my_1y_2$ и $x_2y_1-x_1y_2$ делятся на $N(x_2+y_2\sqrt{m})$. Действительно,

$$\frac{x_1+y_1\sqrt{m}}{x_2+y_2\sqrt{m}}=\frac{(x_1+y_1\sqrt{m})(x_2-y_2\sqrt{m})}{(x_2+y_2\sqrt{m})(x_2-y_2\sqrt{m})}=\frac{x_1x_2-my_1y_2}{N(x_2+y_2\sqrt{m})}+\frac{x_2y_1-x_1y_2}{N(x_2+y_2\sqrt{m})}\sqrt{m}.$$

В частности, если норма числа равна $\pm1$, то на него делятся все числа. Нам потребуется следующее утверждение.

**Лемма 1.** Пусть $x_1+y_1\sqrt{m}$ и $x_2+y_2\sqrt{m}$ принадлежат $\mathbb{Z}[\sqrt{m}]$, $n=|N(x_2+y_2\sqrt{m})|$. Тогда если $x_1\equiv x_2 \pmod{n}$ и $y_1\equiv y_2 \pmod{n}$, то $x_1+y_1\sqrt{m}$ делится на $x_2+y_2\sqrt{m}$.

Доказательство. По определению нормы, $|x_2^2-my_2^2|=n$, поэтому $x_2^2-my_2^2\equiv 0 \pmod{n}$. По условию, $x_1\equiv x_2 \pmod{n}$, поэтому $x_1x_2\equiv x_2^2 \pmod{n}$. Аналогично, из того что $y_1\equiv y_2 \pmod{n}$, следует $my_1y_2\equiv my_2^2 \pmod{n}$. Вычитая, получаем:

$$x_1x_2-my_1y_2\equiv x_2^2-my_2^2\equiv 0 \pmod{n}. \tag{9}$$

Кроме того, перемножив два данных в условии сравнения, предварительно поменяв местами правую и левую части одного из них, получаем $x_2y_1\equiv x_1y_2 \pmod{n}$, или

$$x_2y_1-x_1y_2\equiv 0 \pmod{n}. \tag{10}$$

Сравнения (9) и (10) означают, что $x_1+y_1\sqrt{m}$ делится на $x_2+y_2\sqrt{m}$. □

## Деление точек

Вернёмся к целым точкам, разместившимся на семействе гипербол. Всё, что говорилось о делимости чисел из $\mathbb{Z}[\sqrt{m}]$, справедливо также и для делимости точек ввиду установленного выше взаимно однозначного соответствия. Докажем, что на какой-нибудь гиперболе $l_n$ найдутся такие две целые положительные точки, что одна из них делится на другую. Для этого достаточно показать, что на этой гиперболе найдётся не менее $n^2+1$ различных целых положительных точек. Действительно, имеется всего $n^2$ различных пар остатков, которые могут давать при делении на $n$ абсциссы и ординаты этих точек. По принципу Дирихле, в любом множестве, содержащем $n^2+1$ целых точек, найдутся две такие, у которых как абсциссы, так и ординаты сравнимы между собой по модулю $n$. Согласно лемме 1, эти точки делятся друг на друга. Результат деления обязан лежать на гиперболе $l_1$, т. е. быть решением уравнения Пелля. Поскольку выбранные точки различны и положительны, полученное решение нетривиально.

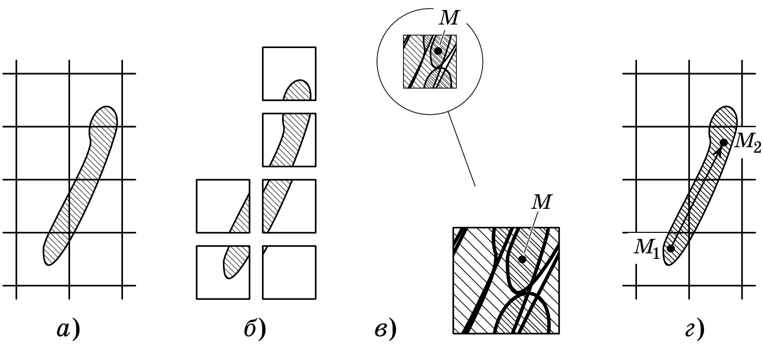
После небольшого геометрического отступления мы докажем, что на одной из гипербол лежит бесконечно много точек. Этим доказательство существования решения будет завершено.

### Лемма Минковского о выпуклом теле

Мы докажем две замечательных геометрических леммы. Вторая, называемая леммой Минковского, неожиданно появится при доказательстве существования решения уравнения Пелля. Первая лемма является вспомогательной для доказательства леммы Минковского.

**Лемма 2.** Пусть фигура Ф на плоскости имеет площадь больше 1. Тогда найдутся две точки $A$ и $B$, принадлежащие фигуре Ф, такие что вектор $\overrightarrow{AB}$ целочисленный. (Иными словами, фигуру Ф можно сдвинуть на целочисленный вектор так, чтобы она пересеклась со своим образом.)

Д о к а з а т е л ь с т в о. Разрежем координатную плоскость на единичные квадратики прямыми, параллельными осям (рис. 8, *а*, *б*). Перенесём параллельно все квадратики, которые имеют хотя бы одну общую точку с фигурой Ф, так, чтобы они все наложились друг




*а*) *б*) *в*) *г*)

Рис. 8

на друга. Для этого мы один из квадратиков оставляем на месте, а остальные накладываем на него. После такого переноса части, на которые была разрезана фигура Ф, оказались целиком лежащими внутри одного квадрата. Так как площадь фигуры больше площади квадрата, то какие-то части пересекаются. Пусть $M$ — какая-нибудь общая точка двух различных частей (рис. 8, *в*). Переместим эти части на их исходные места (обратными параллельными переносами). Точка $M$ при этих переносах перейдёт в точки, лежащие внутри фигуры Ф и являющиеся концами целочисленного вектора (рис. 8, *г*). □

**Лемма Минковского.** Пусть на координатной плоскости имеется центрально-симметричная (относительно начала координат) выпуклая фигура Ф площади больше 4. Тогда она содержит целую точку, отличную от начала координат.

Доказательство. Рассмотрим гомотетию с коэффициентом $\frac{1}{2}$ и центром в начале координат $O$. Она переводит рассматриваемую фигуру Ф в фигуру Ф$'$ площади больше 1, которая также является выпуклой и центрально-симметричной (рис. 9). Согласно лемме 2, найдутся точки $A, B \in$ Ф$'$, такие что вектор $\overrightarrow{AB}$ целочисленный. Пусть точка $A'$ симметрична точке $A$ относительно $O$. Из симметричности фигуры относительно начала координат следует, что $A' \in$ Ф$'$. Далее, пусть $M$ — середина отрезка $A'B$. Из выпуклости следует, что $M \in$ Ф$'$. С другой стороны,

$$\overrightarrow{OM} = \frac{1}{2}(\overrightarrow{OA'} + \overrightarrow{OB}) = \frac{1}{2}(\overrightarrow{AO} + \overrightarrow{OB}) = \frac{1}{2}\overrightarrow{AB}.$$

Рассмотрим точку $K$, такую что $\overrightarrow{OK} = 2\overrightarrow{OM}$. Поскольку $\overrightarrow{OK} = \overrightarrow{AB}$ — целочисленный вектор, точка $K$ — целая, а так как $M \in$ Ф$'$, то $K \in$ Ф. Итак, точка $K$ — искомая. □

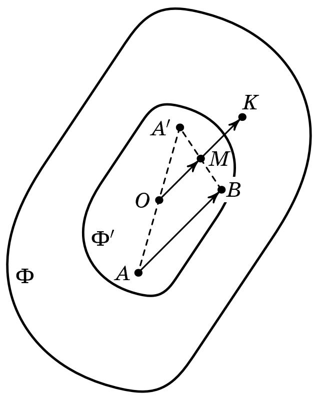


Рис. 9

**4.** В круглом саду радиуса 1 км деревья рассажены в вершинах квадратной сетки со стороной квадратов 1 м. В вершинах сетки, лежащих на границе сада, деревья тоже посажены. Если расстояние от вершины сетки до границы сада меньше радиуса дерева, то дерево вылезает за границу сада. Единственная вершина сетки, в которой дерево не посажено, — центр сада. Радиус каждого дерева равен 1 мм. Докажите, что вид из центра сада полностью заслонён (т. е. любой луч, выходящий из него, пересекает какой-нибудь ствол).

**5*.** Пусть $a$, $b$ и $c$ — целые числа, при этом $a > 0$, $ac - b^2 = 1$. Докажите, что уравнение $ax^2 + 2bxy + cy^2 = 1$ имеет целочисленное решение.

## Завершение доказательства существования нетривиального решения уравнения Пелля

Нам осталось доказать, что на одной из гипербол семейства разместилось бесконечно много целых точек. Предположим противное: каждая гипербола содержит лишь конечное число точек. Выберем гиперболу $l_N$ с достаточно большим номером (чему должно равняться $N$, мы уточним позже) и сопряжённую ей гиперболу $l_{-N}$. Внутри области, ограниченной этими гиперболами, расположено лишь конечное число целых точек. Действительно, все они лежат на конечном числе гипербол с номерами между $-N$ и $N$, а на каждой из них по предположению лежит лишь конечное число точек (в данном контексте вырожденная гипербола $l_0$ тоже учитывается; на ней лежит только начало координат). Каждая из этих точек, кроме начала координат, лежит в одном из четырёх углов, на которые делят плоскость асимптоты гипербол. Проведём лучи, параллельные и сонаправленные сторонам соответствующего угла, с началом в каждой из этих целых точек (рис. 10). Получим конечное множество углов, стороны каждого из которых пересекают одну из четырёх ветвей рассматриваемой пары гипербол, а значит, каждый угол высекает на гиперболах ограниченное множество. Выберем пару симметричных точек на одной из гипербол и сопряжённую ей пару точек на сопряжённой гиперболе (см. с. 12) так, чтобы ни одна из этих четырёх точек не покрывалась ни одним из нарисованных углов. Это можно сделать, выбрав точки достаточно далеко, поскольку углы покрывают лишь ограниченную область на каждой ветви. Параллелограмм с вершинами в этих четырёх точках (рис. 11) не содержит внутри себя ни одной целой точки, кроме начала координат. В противном случае угол с вершиной в этой точке покрыл бы одну из вершин параллелограмма. Но тогда мы получаем противоречие с леммой Минковского. Параллелограмм —

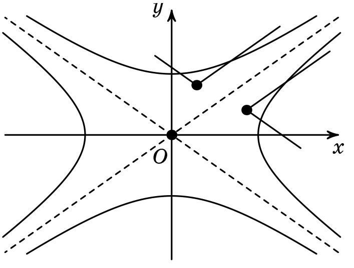



Рис. 10

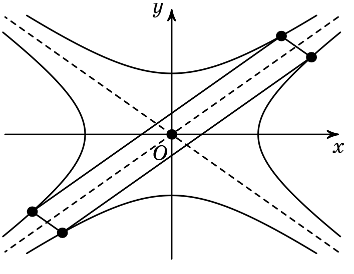



Рис. 11

центрально-симметричная выпуклая фигура, а выбрав параметр $N$ достаточно большим, можно добиться, чтобы его площадь была больше 4.

Осталось показать, каким нужно выбрать $N$. Заметим, что площадь параллелограмма зависит только от номера $N$ гиперболы, но не от выбора точек на ней. Действительно, площадь параллелограмма пропорциональна произведению сторон (коэффициент пропорциональности равен синусу его угла и зависит только от $m$). Но в системе координат, оси которой направлены вдоль асимптот, стороны параллелограмма равны удвоенным абсциссе и ординате одной из его вершин. А гипербола (в этой же системе координат) задаётся следующим уравнением: произведение координат равно константе. Значит, постоянной будет и площадь параллелограмма при произвольном выборе вершин на гиперболах.

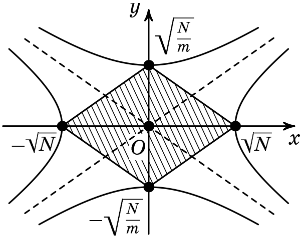


Рис. 12

Вернёмся к исходной системе координат. Для подсчёта площади удобнее всего выбрать параллелограмм с вершинами в точках пересечения гипербол с осями координат (рис. 12). Это точки $(\pm\sqrt{N}, 0)$ и $\left(0, \pm\sqrt{\frac{N}{m}}\right)$. Этот параллелограмм является ромбом, и его площадь легко найти: $S = \frac{2N}{\sqrt{m}}$. Из условия $S > 4$ получаем $N > 2\sqrt{m}$. Итак, мы доказали следующую теорему существования решения.

**Теорема 2.** Любое уравнение Пелля имеет нетривиальное решение.

Замечание. Теорема существования решения уравнения Пелля является частным случаем теоремы Дирихле о строении групп обратимых элементов колец целых алгебраических чисел. Эта теорема (не только доказательство, но и формулировка которой неэлементарны) является одним из замечательных результатов теории чисел. Формулировку теоремы Дирихле и её доказательство, фактически являющееся обобщением приведённого здесь, можно найти в книге [1].

**6.** Контрольный вопрос. Где в доказательстве теоремы 2 использовалось, что число $m$ не является точным квадратом?

## Как найти решение уравнения Пелля?

Приведённое доказательство существования нетривиального решения уравнения Пелля обладает существенным недостатком: оно неконструктивно. Иными словами, доказательство не даёт никаких путей к нахождению решения. Зададимся теперь вопросом, как найти частное (желательно основное) решение уравнения Пелля. Один из способов напрашивается сам собой: будем перебирать последовательно все целые положительные значения $y$ до тех пор, пока число $my^2+1$ не окажется точным квадратом. Такой алгоритм рано или поздно приведёт нас к основному решению. Однако же у нас нет никаких оценок того, насколько долго он будет работать. И действительно, этот алгоритм не всегда эффективен. Для некоторых небольших значений $m$ значения $x$ и $y$, представляющие основное решение, достаточно велики. Например, при $m = 109$ десятичная запись числа $x$ состоит из 15 цифр, а числа $y$ — из 14. Поэтому осуществление такого алгоритма за разумное время нереально даже с использованием компьютера.

Аналогично тому как алгоритм Евклида помог нам эффективно найти решение линейного диофантова уравнения, для нахождения решения уравнения Пелля нам понадобится понятие цепной дроби.

## Цепные дроби

Любое нецелое число $\alpha$ можно представить в виде $\alpha = a_0 + \dfrac{1}{\alpha_1}$, где $a_0$ — целое число, а $\alpha_1 > 1$. Действительно, в качестве $a_0$ нужно взять целую часть числа $\alpha$, а в качестве $\alpha_1$ — обратное число к его дробной части. Такое представление единственно, поскольку из условия $\alpha_1 > 1$ следует $0 < \dfrac{1}{\alpha_1} < 1$, поэтому $\dfrac{1}{\alpha_1}$ — дробная часть $\alpha$, а значит, $a_0$ — целая часть $\alpha$. Если число $\alpha_1$ оказалось нецелым, то его в свою очередь можно представить в виде $\alpha_1 = a_1 + \dfrac{1}{\alpha_2}$ и т. д. В результате мы получим представление числа $\alpha$ в виде

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{\alpha_n}}} \qquad (11)$$

(здесь $a_0$ — целое число, $a_1, a_2, \ldots, a_{n-1}$ — целые положительные числа, а $\alpha_n$ — число, не обязательно целое, большее единицы), причём такое представление (при фиксированном $n$) единственно. Может оказаться, что число $\alpha_n$ для какого-то $n$ целое, и тогда процесс прекратится. В этом случае правая часть выражения (11) называется *конечной цепной дробью*. Если же все $\alpha_i$ не являются целыми, то запишем формальное равенство

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}. \tag{12}$$

Выражение, стоящее в его правой части, называется *бесконечной цепной дробью*. Коэффициенты $a_0$, $a_1$, $a_2$, ... называются её элементами.

**7.** Цепная дробь является конечной тогда и только тогда, когда число $\alpha$ рационально.

Если разрешить последнему элементу конечной цепной дроби равняться единице, то единственность представления для рациональных чисел будет нарушена. Получаемые выражения будем называть *обобщёнными конечными цепными дробями*. Любое рациональное число представляется в виде обобщённой конечной цепной дроби ровно двумя способами, причём количества этажей в них различаются на единицу. Первый способ — это представление в виде конечной цепной дроби, а второй получается из первого заменой последнего элемента $a_n$ на $(a_n - 1) + \frac{1}{1}$.

**8.** Если последовательность элементов бесконечной цепной дроби периодическая, то число $\alpha$ является корнем квадратного уравнения.

З а м е ч а н и е. Верно и обратное утверждение: если $\alpha$ — иррациональный корень квадратного уравнения, то последовательность элементов цепной дроби для числа $\alpha$ является периодической. Однако его доказательство значительно сложнее.

Чтобы придать смысл равенству (12), следует корректно определить его правую часть. Для этого «обрубим хвост» бесконечной цепной дроби и получим обобщённую конечную цепную дробь

$$r_n = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}}, \tag{13}$$

которая представляет собой некоторое рациональное число $r_n = \frac{p_n}{q_n}$.

Дробь $\frac{p_n}{q_n}$ будем считать несократимой. Она называется *n-й подходя-*

*щей дробью* цепной дроби (а также числа $\alpha$, для которого мы строили цепную дробь). Подходящая дробь определена при всех $n \geqslant 0$. Мы докажем, что последовательность подходящих дробей сходится. Тогда естественно правую часть равенства (12) определить как предел последовательности соответствующих подходящих дробей. Для доказательства существования этого предела вначале сформулируем несколько вспомогательных утверждений. Два первых приведём в виде задач. Они могут быть легко доказаны методом математической индукции.

**9.** Докажите, что последовательности $p_n$ и $q_n$ удовлетворяют рекуррентным соотношениям

$$\begin{cases} p_{n+1} = p_n a_{n+1} + p_{n-1}, \\ q_{n+1} = q_n a_{n+1} + q_{n-1}. \end{cases} \tag{14}$$

Как следствие отметим, что последовательности $p_n$ и $q_n$ монотонно возрастают по абсолютной величине (при $n \geqslant 1$) и стремятся к бесконечности.

**10.** Докажите, что при всех $n \geqslant 1$

$$p_{n-1}q_n - p_n q_{n-1} = (-1)^n. \tag{15}$$

**Лемма 3.** Если $n$ чётно и $m > n$, то $r_n < \alpha$ и $r_n < r_m$. Если $n$ нечётно и $m > n$, то $r_n > \alpha$ и $r_n > r_m$.

Доказательство. Рассмотрим функцию от переменной $x$

$$f_n(x) = a_0 + \cfrac{1}{a_1 + \ddots + \cfrac{1}{a_{n-1} + \cfrac{1}{x}}}. \tag{16}$$

Она является монотонно возрастающей при чётном $n$ и убывающей при нечётном $n$. Действительно, её можно представить как $n$-кратную композицию двух функций: взятия обратной величины и прибавления константы. При этом взятие обратной величины каждый раз меняет характер монотонности функции с возрастания на убывание и наоборот, а прибавление константы сохраняет его.

Обозначим

$$a_{n,m} = a_n + \cfrac{1}{a_{n+1} + \ddots + \cfrac{1}{a_m}}.$$

Тогда $a_n < \alpha_n$, $a_n < a_{n,m}$ и $f_n(a_n) = r_n$, $f_n(\alpha_n) = \alpha$, $f_n(a_{n,m}) = r_m$. Из монотонности функции $f_n(x)$ следует, что $r_n < \alpha$, $r_n < r_m$ при чётном $n$ и $r_n > \alpha$, $r_n > r_m$ при нечётном $n$. □

**Лемма 4.** Разность $r_n - r_{n+1}$ стремится к нулю.

Доказательство. Из формулы (15) следует, что

$$|r_n - r_{n+1}| = \left|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}\right| = \left|\frac{p_n q_{n+1} - p_{n+1} q_n}{q_n q_{n+1}}\right| = \frac{1}{q_n q_{n+1}} \leqslant \frac{1}{q_n^2}. \qquad (17)$$

Теперь утверждение леммы вытекает из очевидного следствия формул (14), а именно из того что $q_n$ неограниченно возрастает. □

**Теорема 3.** Последовательность $r_n$ сходится к числу $\alpha$.

Доказательство. По лемме 3 подпоследовательность членов с чётными номерами возрастает и ограничена сверху (например числом $\alpha$ или любым членом последовательности с нечётным номером). Согласно теореме Больцано—Вейерштрасса, она сходится к числу, не превосходящему $\alpha$. Аналогично, подпоследовательность членов с нечётными номерами монотонно убывает и ограничена снизу числом $\alpha$, поэтому сходится к числу, не меньшему $\alpha$. По лемме 4 разность этих подпоследовательностей неограниченно убывает, поэтому оба предела равны $\alpha$, а значит, и предел всей последовательности $r_n$ равен $\alpha$. □

Замечание. Мы строили цепную дробь, исходя из заданного числа $\alpha$. Однако бесконечную цепную дробь можно построить, взяв в качестве её элементов любую последовательность положительных (кроме, возможно, $a_0$) целых чисел. Значением цепной дроби будет предел значений её подходящих дробей.

## Подходящие дроби как приближения действительных чисел рациональными

Поскольку $\alpha$ лежит между $r_n$ и $r_{n+1}$, из формулы (17) следует, что

$$\left|\frac{p_n}{q_n} - \alpha\right| < \frac{1}{q_n^2}.$$

Иными словами, подходящие дроби числа $\alpha$ являются хорошими приближениями к нему. Следующая теорема показывает, что в некотором смысле верно и обратное утверждение: если рациональное число хорошо приближает число $\alpha$, то оно является некоторой его подходящей дробью. Правда, нужно уточнить, что значит «хорошо приближать»: мера «хорошести» приближения в прямом и обратном утверждениях различаются в два раза.

**Теорема 4.** Если несократимая дробь $\frac{p}{q}$ такова, что

$$\left|\frac{p}{q}-\alpha\right|<\frac{1}{2q^2},$$

то $\frac{p}{q}$ **является подходящей дробью числа** $\alpha$.

Доказательство. Для начала рассмотрим функцию, задаваемую формулой (16), и выясним, как она будет выглядеть после приведения к обыкновенной дроби. Нам будет удобнее заменить индекс $n$ на $n+1$. Учитывая формулы (14), имеем

$$f_{n+1}(a_{n+1})=\frac{p_{n+1}}{q_{n+1}}=\frac{p_n a_{n+1}+p_{n-1}}{q_n a_{n+1}+q_{n-1}}.$$

Так как коэффициенты $p_n$, $p_{n-1}$, $q_n$ и $q_{n-1}$ не зависят от $a_{n+1}$, то $a_{n+1}$ можно считать переменной, и значит, для любого $x$

$$f_{n+1}(x)=\frac{p_n x+p_{n-1}}{q_n x+q_{n-1}}. \tag{18}$$

Нетрудно найти функцию $g_{n+1}(x)$, обратную к $f_{n+1}(x)$:

$$g_{n+1}(x)=\frac{p_{n-1}-q_{n-1}x}{q_n x-p_n}. \tag{19}$$

Теперь рассмотрим разложение числа $\frac{p}{q}$ в обобщённую конечную цепную дробь

$$\frac{p}{q}=a_0+\cfrac{1}{a_1+\cfrac{1}{a_2+\ddots+\cfrac{1}{a_n}}}.$$

Если $\alpha>\frac{p}{q}$, то из двух возможных выбираем представление с чётным $n$, а если $\alpha<\frac{p}{q}$, то с нечётным. Пусть $\omega=g_{n+1}(\alpha)$. Тогда $\alpha=f_{n+1}(\omega)$, кроме того,

$$\omega+\frac{q_{n-1}}{q_n}=\frac{p_{n-1}-\alpha q_{n-1}}{\alpha q_n-p_n}+\frac{q_{n-1}}{q_n}=\frac{(-1)^n}{q_n^2\left(\alpha-\frac{p_n}{q_n}\right)}=\frac{1}{q^2\left|\alpha-\frac{p}{q}\right|}>2.$$

Мы воспользовались формулой (15) и равенством $\frac{p_n}{q_n}=\frac{p}{q}$. Отсюда получаем:

$$\omega>2-\frac{q_{n-1}}{q_n}\geqslant 1.$$

Наконец, рассмотрим число

$$\alpha' = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n + \cfrac{1}{\omega}}}}.$$

Так как $\omega > 1$, то $a_0$, $a_1$, ..., $a_n$ — начальные элементы цепной дроби для числа $\alpha'$ (ведь, как мы уже отмечали, представление любого числа в виде (11) единственно). Значит, $\alpha' = f_{n+1}(\omega)$. С другой стороны, мы выбирали $\omega$ так, чтобы $\alpha = f_{n+1}(\omega)$. Поэтому $\alpha = \alpha'$. А отсюда следует, что $\frac{p}{q}$ — подходящая дробь $\alpha$. □

**11.** Контрольный вопрос. В приведённом доказательстве имеется пробел: если $\alpha$ не принадлежит области определения функции $g_{n+1}$, то число $\omega$ не определено. Восполните этот пробел.

Дальнейшие сведения из теории цепных дробей вы можете найти в книгах [2] и [4].

## Решения уравнения Пелля — числители и знаменатели подходящих дробей

**Теорема 5.** Пусть $(x, y)$ — положительное решение уравнения Пелля. Тогда $\frac{x}{y}$ является подходящей дробью $\sqrt{m}$.

Доказательство. Так как $x > y > 0$ и $\sqrt{m} > 1$, то $x + y\sqrt{m} > 2y$. Значит,

$$1 = x^2 - my^2 = (x - y\sqrt{m})(x + y\sqrt{m}) > (x - y\sqrt{m}) \cdot 2y.$$

Разделим полученное неравенство на $2y^2$:

$$\frac{x}{y} - \sqrt{m} < \frac{1}{2y^2}.$$

Поскольку $(x, y)$ — положительное решение уравнения Пелля, левая часть этого неравенства положительна и дробь $\frac{x}{y}$ несократима. Поэтому, согласно теореме 4, она является подходящей дробью числа $\sqrt{m}$. □

Итак, положительные решения уравнений Пелля следует искать только среди пар, составленных из числителя и знаменателя какой-нибудь подходящей дроби числа $\sqrt{m}$. Возникает вопрос, какие именно подходящие дроби соответствуют решениям уравнения Пелля. Ответ на него даёт следующая теорема, которую мы приводим без доказательства. В её формулировке используется, что элементы цепной дроби для числа $\sqrt{m}$ образуют периодическую последовательность (см. замечание к задаче 8).

**Теорема 6.** Пусть $n$ — длина периода последовательности элементов цепной дроби для числа $\sqrt{m}$. Тогда числитель и знаменатель подходящей дроби числа $\sqrt{m}$ являются решением уравнения Пелля тогда и только тогда, когда её номер имеет вид $kn-1$ (т. е. даёт при делении на $n$ остаток $n-1$) и нечётен.

Доказательство этой теоремы, а также другие интересные факты из теории чисел вы найдёте в книге [2]. Об уравнениях Пелля и других диофантовых уравнениях можно прочитать также в книгах [3] и [5].

Метод цепных дробей является вполне эффективным для отыскания решений уравнений Пелля. Например, в самом сложном среди всех значений $m < 100$ случае $m = 61$ решение находится всего на 21-м шаге. Такой подсчёт было под силу осуществить вручную даже древним математикам. А в настоящее время, при возможности использовать компьютер, это является простым упражнением. Получаем $x = 1\,766\,319\,049$, $y = 226\,153\,980$.

### Исторический комментарий

Упоминания об уравнениях, которые сейчас принято называть уравнениями Пелля, были найдены уже в работах математиков Древней Греции и древней Индии. В работах индийского математика XII века Бхаскары был предложен общий способ решения этих уравнений (так называемый «циклический метод»). В частности, при помощи этого метода им было найдено решение в случае $m = 61$, приведённое нами выше. Однако в те времена ещё не ставился вопрос о доказательстве того, что приведённый всегда приводит к решению.

В общем виде эту задачу сформулировал в середине XVII века великий французский математик Пьер Ферма. Вот какова была формулировка Ферма в письме, обращённом к математикам, его современникам.

> *Если дано произвольное число, которое не является квадратом, то найдётся также и бесконечное количество таких квадратов, что если этот квадрат умножить на данное число и к произведению прибавить единицу, то результат будет квадратом.*

Ферма утверждал, что умеет это доказывать. Однако, как и большинство его трудов, этот результат не был опубликован. Поэтому до нас не дошло доказательство Ферма и даже не известно, было ли оно верным.

Английские математики Джон Валлис и Уильям Броункер нашли способ решения этого уравнения, отличный от циклического метода. Однако и они не смогли доказать, что этот способ всегда приводит к успеху. Или даже не осознавали необходимости такого доказательства! И лишь в конце XVIII века французский математик Жозеф Луи Лагранж дал доказательство гипотезы, сформулированной в письме Ферма.

Леонард Эйлер ошибочно приписал авторство этих уравнений Джону Пеллю. Благодаря этому, уравнения получили имя математика, который не имел к ним практически никакого отношения. Впрочем, подобные казусы в истории математики нередки. Иногда уравнения Пелля называют неопределёнными уравнениями Ферма, хотя название «уравнения Пелля» в настоящее время более распространено.

В конце XIX века выдающийся немецкий математик и физик Герман Минковский разработал так называемую геометрию чисел, в которой методы геометрии используются для решения задач теории чисел. Основным объектом её изучения являются пространственные решётки. С помощью них Минковскому удалось получить много новых результатов в теории чисел и найти новые доказательства многих известных теорем. В частности, приведённое в брошюре доказательство теоремы существования решения уравнения Пелля основано на геометрических идеях Минковского.

Подробнее об истории уравнений Пелля и о различных алгоритмах нахождения его решений можно узнать в книге [5].

## Литература


[1] З. И. Боревич, И. Р. Шафаревич. Теория чисел. — М.: Наука, 1985.

[2] А. А. Бухштаб. Теория чисел. — М.: Учпедгиз, 1960.

[3] А. О. Гельфонд. Решение уравнений в целых числах. — (Серия «Популярные лекции по математике». Вып. 8). — М.: Наука, 1983.

[4] А. Я. Хинчин. Цепные дроби. — М.: Наука, 1978.

[5] Г. Эдвардс. Последняя теорема Ферма. Генетическое введение в алгебраическую теорию чисел. — М.: Мир, 1980.

# Ответы, указания, решения

**1.** Пусть $m > n$. Из равенства

$$2^{2^m}+1=(2^{2^n}-1)(2^{2^n}+1)(2^{2^{n+1}}+1)\cdot\ldots\cdot(2^{2^{m-1}}+1)+2$$

следует, что число $2^{2^m}+1$ при делении на $2^{2^n}+1$ даёт остаток 2. Поэтому НОД$(2^{2^m}+1,$ $2^{2^n}+1)=$ НОД$(2^{2^n}+1,\ 2)=1$.

**2.** Асимптота $x=y\sqrt{m}$ подвергается гомотетии с коэффициентом $x_0+y_0\sqrt{m}$ (все её точки удаляются от начала координат), а асимптота $x=-y\sqrt{m}$ подвергается гомотетии с коэффициентом $x_0-y_0\sqrt{m}$ (все её точки приближаются к началу координат). Начало координат (точка пересечения асимптот) является единственной неподвижной точкой этого преобразования.

**3.** Графиком рассматриваемого уравнения является либо эллипс (при $m=1$), либо пара параллельных прямых (при $m=2$), либо гипербола (при $m\geqslant 3$). Для случаев $m=1$, 2 и 3 графики с отмеченными на них целыми точками изображены на рис. 13.

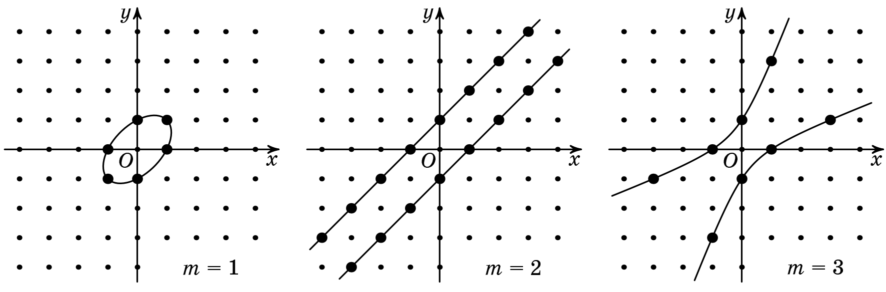


Рис. 13

Рассмотрим два преобразования плоскости, задаваемые формулами $f(x,y)=$ $=(y,\ my-x)$ и $g(x,y)=(mx-y,\ x)$. Непосредственная проверка показывает, что они переводят любое решение рассматриваемого уравнения в другое его решение и являются взаимно обратными. В частности, это означает, что пары $(\varphi_k,\varphi_{k+1})$ являются решениями уравнения. Докажем, что эти пары исчерпывают множество решений, удовлетворяющих условию $0\leqslant x<y$. Предположим противное: имеется решение, лежащее на графике между двумя решениями $(\varphi_{k-1},\varphi_k)$ и $(\varphi_k,\varphi_{k+1})$. Применив к нему преобразование $g^k$, получим решение, лежащее между $(-1,0)$ и $(0,1)$. Противоречие, поскольку такого решения не существует.

**4.** Выберем за единицу длины один метр — расстояние между соседними деревьями. Радиусы сада и дерева обозначим через $R=1000$ и $r=0{,}001$ соответственно.

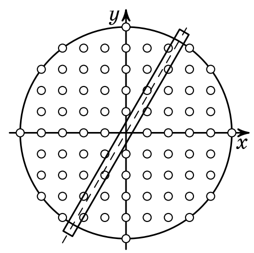


Рис. 14

Рассмотрим некоторое направление и построим прямоугольник с центром в центре сада, одна сторона которого параллельна выбранному направлению и имеет длину $2R'$ (где $R'$ чуть больше $R$), а другая равна $2r$ (рис. 14). Площадь такого прямоугольника равна $4R'r>4Rr=4$. Согласно лемме Минковского, внутри его имеется узел сетки, а дерево с центром в этом узле будет заслонять вид в выбранном направлении.

Осталось доказать, что этот узел не может находиться вне сада. Для этого рассмотрим часть прямоугольника, лежащую вне круга радиуса $R$ (на рис. 15 изображена одна из двух её компонент). Квадрат расстояния от любой точки этой области до центра сада больше чем $R^2$ и не превосходит $R'^2 + r^2$ (на максимальном расстоянии находится вершина прямоугольника). Выбрав $R'$ достаточно мало отличающимся от $R$, можно добиться, чтобы $R'^2 + r^2 < R^2 + 1$. Поэтому квадрат расстояния от любой точки рассматриваемой области до центра является дробным числом. Но квадрат расстояния между любыми двумя целыми точками — число целое. Следовательно, рассматриваемая область не содержит целых точек.

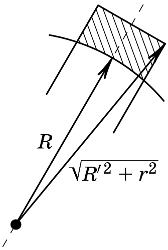



Рис. 15

**5.** Рассмотрим кривую, задаваемую уравнением $ax^2 + 2bxy + cy^2 = \lambda^2$, где $\frac{4}{\pi} < \lambda^2 < 2$. Это эллипс, полуоси которого равны

$$d_1 = \frac{\lambda}{\sqrt{\dfrac{a+c-\sqrt{(a+c)^2-4}}{2}}} \quad \text{и} \quad d_2 = \frac{\lambda}{\sqrt{\dfrac{a+c+\sqrt{(a+c)^2-4}}{2}}}.$$

Площадь ограниченной им области равна $\pi d_1 d_2 = \pi\lambda^2 > 4$. Согласно лемме Минковского, внутри эллипса содержится отличная от начала координат целая точка $(x_0, y_0)$. Величина $ax_0^2 + 2bx_0y_0 + cy_0^2$ должна быть целым числом, большим нуля и меньшим $\lambda^2$. Значит, она равна 1.

**6.** Если $m$ является точным квадратом, то имеются отличные от начала координат целые точки, лежащие на асимптотах гипербол. А в завершающей части доказательства мы пользовались тем, что все целые точки кроме начала координат лежат строго внутри одного из четырёх углов, на которые асимптоты делят плоскость.

**7.** Если цепная дробь конечна, то её можно преобразовать в обыкновенную дробь, т. е. рациональное число.

Обратно, пусть рациональное число $\alpha$ представлено в виде $\alpha = a_0 + \frac{1}{\alpha_1}$, где $a_0$ — целое число, а $\alpha_1 > 1$. Тогда если $\alpha = \frac{p}{q}$, то $a_0$ — частное от деления $p$ на $q$, а $\alpha_1 = \frac{q}{r}$, где $r$ — остаток от деления $p$ на $q$. Значит, последовательность $\alpha_i$ состоит из рациональных чисел с убывающими знаменателями. Следовательно, она конечна.

**8.** Если последовательность $a_n$ периодическая, то $\alpha_k = \alpha_l$ для некоторых $k$ и $l$. Следовательно, $g_k(\alpha) = g_l(\alpha)$, где $g_k$ и $g_l$ — дробно-линейные функции (задаваемые формулами (19)). Полученное уравнение относительно $\alpha$ простыми преобразованиями приводится к квадратному.

**9.** База индукции сводится к непосредственной проверке тождеств (14) для $n = 1$.

Шаг индукции. По предположению индукции, при $n \geqslant 2$

$$\frac{p_n}{q_n} = \frac{p_{n-1}a_n + p_{n-2}}{q_{n-1}a_n + q_{n-2}}.$$

Если в правую часть этого равенства вместо $a_n$ подставить $a_n + \frac{1}{a_{n+1}}$, получим $\frac{p_{n+1}}{q_{n+1}}$. Значит,

$$\frac{p_{n+1}}{q_{n+1}} = \frac{p_{n-1}\left(a_n + \dfrac{1}{a_{n+1}}\right) + p_{n-2}}{q_{n-1}\left(a_n + \dfrac{1}{a_{n+1}}\right) + q_{n-2}} = \frac{p_{n-1}a_n + p_{n-2} + \dfrac{p_{n-1}}{a_{n+1}}}{q_{n-1}a_n + q_{n-2} + \dfrac{q_{n-1}}{a_{n+1}}} = \frac{p_n a_{n+1} + p_{n-1}}{q_n a_{n+1} + q_{n-1}}.$$

Мы доказали, что получаемые с помощью рекуррентных соотношений (14) дроби $\frac{p_n}{q_n}$ равны подходящим дробям. Осталось доказать, что полученные дроби несократимы. Это следует из приведённого ниже решения следующей задачи 10 (им пользоваться можно, поскольку оно не использует взаимной простоты чисел $p_n$ и $q_n$). Если предположить, что числа $p_n$ и $q_n$ имеют общий делитель, то этот же делитель должно иметь число $p_{n-1}q_n - p_n q_{n-1} = (-1)^n$. Таким делителем может быть только единица.

**10.** База индукции ($n = 1$) проверяется непосредственной подстановкой $p_0 = a_0$, $q_0 = 1$, $p_1 = a_0 a_1 + 1$, $q_1 = a_1$.

Умножим первую из формул (14) на $q_n$, а вторую — на $p_n$ и вычтем одну из другой. Получим

$$p_n q_{n+1} - p_{n+1} q_n = -(p_{n-1} q_n - p_n q_{n-1}).$$

Отсюда следует шаг индукции.

**11.** Функция $g_{n+1}(x)$ не определена лишь при $x = \frac{p_n}{q_n}$. Но если $\alpha = \frac{p}{q}$, то утверждение теоремы, очевидно, верно.

## Оглавление