Квадратичная математика

В.А. Тиморин

Предисловие

Книга содержит конспекты четырех лекций, прочитанных мной в Дубне летом 2005 года. Для чтения не требуется почти никаких предварительных знаний, по крайней мере, ничего выходящего за рамки школьной программы. Исключение составляет только последний раздел. Однако требуется хорошая математическая культура. Некоторые из обсуждаемых в книге вопросов отражают современное состояние предмета.

Название книги и курса не вполне честно. Оно намеренно звучит как название некоторой математической дисциплины. Однако, такой дисциплины не существует. По крайней мере, на сегодняшний день. А в книге речь пойдет о нескольких задачах из разных областей математики — алгебры, геометрии, теории чисел. Записанные на языке формул, эти задачи содержат квадратичные формы — выражения, являющиеся суммами квадратов или попарных произведений переменных с некоторыми коэффициентами. Само по себе, это обстоятельство еще не является поводом к тому, чтобы объединять эти задачи в одну науку. Однако у меня все же создается сильное впечатление, что такая наука существует — что все обсуждаемые вопросы являются разными гранями чего-то общего. Конечно, у читателя, может возникнуть и другое впечатление.

Несколько слов о задачах и упражнениях. Обычно упражнениями называют те задачи, которые легче остальных. В этой книге противоположная терминология. Задачи встречаются в тексте, и их решение существенно для понимания. Предполагается, что читатель либо знает решение задачи (многие из задач входят в школьную программу), либо может решить задачу за короткое время. Дополнительные упражнения приведены в конце каждого раздела. Это утверждения, дополняющие основной текст. Не все из этих утверждений просто доказать. Дополнительные упражнения образуют как бы второй слой книги, менее элементарный, чем основной текст.

Читателю, для которого обсуждаемые вопросы новы, при первом чтении рекомендуется пропустить дополнительные упражнения. А читателю, уже что-то знающему про все эти вопросы, можно рассматривать книгу как сборник задач повышенной сложности, состоящий из дополнительных упражнений. Основной текст тогда можно только бегло просмотреть. Я надеюсь, что книга будет интересна как школьникам старших классов, так и студентам.

1 Обзор основных сюжетов

Какие целые числа можно представить в виде суммы квадратов двух целых чисел? Это один из самых старых вопросов теории чисел, восходящий по крайней мере к Диофанту. Полный ответ на этот вопрос дал Ферма. Напишем первые несколько целых чисел, представимых в виде суммы квадратов:

$$0, 1, 2, 4, 5, 8, 9, 10, 13, 16, 17, 18, 20, 25, 26, 29, 32, 34, 36, 37, 40, 41, 45,$$

$$49, 50, 52, 53, 58, 61, 64, 65, 68, 72, 73, 74, 80, 81, 82, 85, 89, 90, 97, 98, 100$$

Можно сделать несколько экспериментальных выводов. Во-первых, не каждое число представимо в виде суммы двух квадратов. Например, 3, 6, 11, 12 не представляются в таком виде. Более того, можно заметить, что ни одно число вида $4k + 3$ не представляется в виде суммы двух квадратов (при целом k). Во-вторых, если каждое из двух чисел является суммой квадратов, то таково и их произведение. Конечно, можно сделать и другие, более глубокие, заключения.

Остановимся, однако, на втором заключении и попробуем его обосновать теоретически. Справедлива формула

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2) = (x_1y_1 - x_2y_2)^2 + (x_1y_2 + x_2y_1)^2. \tag{1}$$

Из этой формулы, в частности, вытекает, что если каждое из чисел a и b можно представить как сумму квадратов двух целых чисел, то произведение ab тоже представимо в таком виде.

Формула (1) тесно связана с законом умножения комплексных чисел. Напомним, что *комплексное число* — это выражение вида $u + vi$, где u и v — действительные числа. Символ i, по определению, не является действительным числом, однако он ведет себя как действительное число при арифметических операциях, например, $i + i = 2i$ и $xi = ix$ для всякого действительного числа x. Единственная разница состоит в том, что $i^2 = -1$, в то время квадрат никакого действительного числа не может быть отрицательным числом.

Рассмотрим комплексные числа $z = x_1 + x_2i$ и $w = y_1 + y_2i$. Их произведение равно

$$zw = (x_1y_1 - x_2y_2) + (x_1y_2 + x_2y_1)i.$$

Это можно проверить обычным раскрытием скобок. При этом нужно воспользоваться равенством $i^2 = -1$. Если ввести обозначение $|z|^2 = x_1^2+x_2^2$, то формулу (1) можно переписать в виде

$$|z|^2|w|^2 = |zw|^2$$

Модуль комплексного числа z — квадратный корень из $|z|^2$ — имеет простой геометрический смысл. Число $z = x_1 + x_2i$ удобно изображать точкой плоскости с координатами (x_1, x_2). Тогда модуль числа z — это расстояние от соответствующей точки до начала координат. Пользуясь введенной терминологией, можно переформулировать равенство (1) так: модуль произведения комплексных чисел равен произведению модулей.

Как мы уже видели, формула (1) важна для теории чисел. В следующих разделах мы обсудим ее теоретико-числовые приложения, а также другие аналогичные формулы, важные для теории чисел. Мы докажем, например, такую теорему, принадлежащую Ферма:

Теорема Ферма о суммах двух квадратов. *Положительное целое число x представляется в виде суммы квадратов двух целых чисел тогда и только тогда, когда все простые числа, отличные от 2 и входящие в разложение числа x в нечетной степени, имеют вид $4k+1$ для некоторых целых чисел k.*

А сейчас опишем геометрический смысл формулы (1). Оказывается, что с этой формулой связано некоторое замечательное отображение из трехмерной сферы в двумерную. Однако, прежде чем определять это отображение, следует понять, что такое трехмерная сфера. Согласно стандартному определению, это подмножество в четырехмерном пространстве, определенное уравнением, аналогичным уравнению двумерной сферы в трехмерном пространстве. В трехмерном пространстве с координатами (x_1, x_2, x_3), сфера радиуса 1 с центром в начале координат задается уравнением

$$x_1^2 + x_2^2 + x_3^2 = 1.$$

Аналогично, в четырехмерном пространстве с координатами (x_1, x_2, x_3, x_4), сфера радиуса 1 с центром в начале координат задается уравнением

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1.$$

Из этого описания может показаться, что наглядно представить себе трехмерную сферу трудно. Однако это не так. Представить себе трехмерную сферу не намного сложнее, чем двумерную.

Двумерную сферу можно изобразить на плоскости, спроецировав ее из любой точки. Точку, из которой производится проекция, можно условно назвать “северным полюсом”. Для того, чтобы изобразить точку сферы на плоскости, нужно соединить ее прямой линией с северным полюсом, после чего отметить точку пересечения этой прямой с плоскостью проекции. Лучше всего, чтобы плоскость проекции была параллельна касательной плоскости к сфере в северном полюсе. Тогда каждая точка сферы, кроме северного полюса, будет однозначно проецироваться на плоскость, и каждая точка плоскости будет проекцией некоторой (однозначно определенной) точки сферы. Описанная проекция называется *стереографической проекцией* (см. рис. 1).

Стереографическая проекция позволяет отождествить двумерную сферу с плоскостью, пополненной *“бесконечно удаленной точкой”*. При данном отождествлении, точка на сфере, отличная от северного полюса,

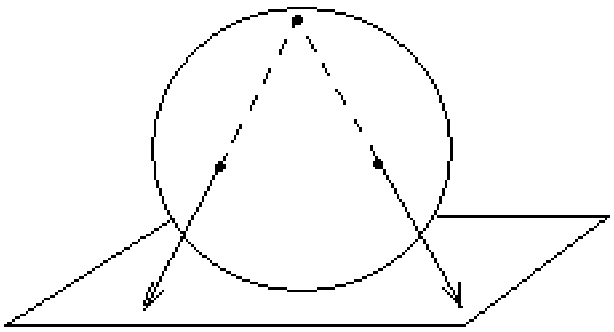

Рис. 1: стереографическая проекция.

отождествляется со своей стереографической проекцией. Однако проекции северного полюса не существует. Поэтому мы условно считаем, что северный полюс проецируется в бесконечно удаленную точку. Название "бесконечно удаленная точка" оправдывается следующим наблюдением. Если точка на сфере приближается к северному полюсу, то ее стереографическая проекция стремится к бесконечности.

Итак, двумерную сферу можно представить себе как плоскость, пополненную бесконечно удаленной точкой. Совершенно аналогично, трехмерную сферу удобно представлять как трехмерное пространство плюс точка на бесконечности. В таком пространстве, если идти по прямой достаточно долго, то можно прийти в исходную точку (через бесконечность).

Наконец, опишем отображение из трехмерной сферы в двумерную, связанное с формулой (1). Чтобы определить это отображение в координатах, нам придется воспользоваться исходным определением трехмерной сферы как подмножества четырехмерного пространства. Определим отображение четырехмерного пространства в трехмерное по следующей формуле:

$$f : (x_1, x_2, y_1, y_2) \mapsto (2(x_1 y_1 - x_2 y_2),\ 2(x_1 y_2 + x_2 y_1),\ x_1^2 + x_2^2 - y_1^2 - y_2^2). \quad (2)$$

Эта формула означает, что точка четырехмерного пространства с координатами, указанными в левой части, то есть (x_1, x_2, y_1, y_2), отображается в точку трехмерного пространства с координатами, указанными в правой части. Здесь x_1, x_2, y_1, y_2 могут быть любыми действительными числами. Мы выбрали нестандартные обозначения для координат — (x_1, x_2, y_1, y_2) вместо (x_1, x_2, x_3, x_4) — из соображений удобства. Отображение f, заданное формулой (2), называется *отображением Хопфа.*

Для точки X четырехмерного пространства с координатами (x_1, x_2, x_3, x_4) обозначим через $|X|^2$ число $x_1^2 + x_2^2 + x_3^2 + x_4^2$. Квадратный корень $|X|$ из этого числа называется *расстоянием от точки X до начала координат.* Аналогичным обозначением мы будем пользоваться и в трехмерном случае. Замечательное свойство отображения Хопфа, вытекающее из формулы (1), состоит в следующем:

$$|f(X)| = |X|^2 \tag{3}$$

Задача. Выведите эту формулу из формулы (1).

Заметим, что если $|X| = 1$, то $|f(X)| = 1$ по формуле (3). Это означает, что образ трехмерной сферы содержится в двумерной сфере. Нетрудно проверить, что на самом деле образ — вся двумерная сфера. Ограничение отображения Хопфа на трехмерную сферу тоже называется отображением Хопфа. Это отображение трехмерной сферы в двумерную.

Аналогичного отображения двумерной сферы в одномерную сферу (т.е. в окружность) не существует. Конечно, можно придумать много непрерывных отображений из двумерной сферы в окружность. Но все они устроены следующим образом. Сначала нужно каким-либо образом отобразить двумерную сферу в прямую — образ сферы будет отрезком — а потом намотать прямую на окружность. Такие отображения не очень интересны (с топологической точки зрения) по следующей причине. Их можно непрерывно продеформировать в отображение, переводящее всю сферу в одну точку. Такое отображение называется *тривиальным.* Отображение, которое можно непрерывно продеформировать в тривиальное отображение, называется *гомотопически тривиальным.*

Все отображения из двумерной сферы в прямую гомотопически тривиальны. Действительно, образом сферы при любом таком отображении будет отрезок, и этот отрезок можно непрерывно сжать в точку. Следовательно, все отображения из двумерной сферы в окружность гомотопически тривиальны. Конечно, чтобы доработать эти неформальные рассуждения до строгого доказательства, нужно хорошо потрудиться. Читателю, не знакомому с основами топологии, рекомендуется отложить это занятие до изучения соответствующего предмета. Тем не менее,

Задача. Посмотрите в какой-нибудь книжке определение непрерывного отображения и непрерывной деформации.

Отображение Хопфа не является гомотопически тривиальным. Это свойство важно для *топологии* — науки, изучающей геометрические объекты с точностью до непрерывных деформаций.

Перейдем теперь к чисто геометрическим свойствам отображения Хопфа. Сначала скажем несколько слов о геометрии трехмерной сферы. По аналогии с двумерной сферой, определим *большие окружности* на трехмерной сфере как пересечения трехмерной сферы с двумерными плоскостями, проходящими через начало координат. *Двумерная плоскость* в четырехмерном пространстве задается системой из двух линейных уравнений

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 = 0,$$

таких, что одно не получается из другого умножением на постоянное действительное число. Здесь $x_1, \ldots, x_4$ — переменные, а $a_1, \ldots, b_4$ — постоянные действительные числа. При стереографической проекции, большие окружности переходят в прямые или окружности в трехмерном пространстве. В определенном смысле, про большие окружности следует думать как про "прямые на сфере". Они реализуют кратчайшее расстояние (по поверхности сферы) от одной точки до другой.

Напомним, что окружности на двумерной сфере — это пересечения двумерной сферы с плоскостями, не обязательно проходящими через начало координат. Тем самым, кроме больших окружностей, мы рассматриваем много других окружностей — "малые окружности". Заметим, что согласно данному определению, отдельные точки тоже считаются окружностями. Это окружности нулевого радиуса. Все окружности на сфере переходят в прямые, окружности или точки на плоскости при стереографической проекции. Конечно, можно аналогичным образом определить произвольные окружности на трехмерной сфере.

Теорема об отображении Хопфа и окружностях.

- *Отображение Хопфа переводит все большие окружности на трехмерной сфере в окружности на двумерной сфере.*
- *Всякая окружность на двумерной сфере является образом одной и только одной большой окружности на трехмерной сфере, и наоборот.*

Вернемся к формулам с суммами квадратов. Теперь нас интересует такая алгебраическая задача: какие формулы с суммами квадратов можно написать для случая многих переменных? Сформулируем эту задачу более точно. Рассмотрим формулу вида

$$(x_1^2 + x_2^2 + \cdots + x_r^2)(y_1^2 + y_2^2 + \cdots + y_s^2) = z_1^2 + \cdots + z_n^2,$$

в которой $z_1, \ldots, z_n$ являются билинейными формами от $x_1, \ldots, x_r$ и $y_1, \ldots, y_s$. *Билинейные формы* — это выражения вида x_1y_1, x_3y_2, x_2y_1 и т.д., а также суммы таких выражений, взятых с произвольными действительными коэффициентами. Формулу, написанную выше, будем называть формулой типа (r, s, n). Мы докажем такую теорему:

Теорема. *Не существует формулы типа* $(2, 3, 3)$.

Отсюда, в частности, следует, что не существует и формулы типа $(3, 3, 3)$. Действительно, чтобы из формулы типа $(3, 3, 3)$ получить формулу типа $(2, 3, 3)$, достаточно положить $x_3 = 0$.

Однако оказывается, что формула типа $(4,4,4)$ существует! Это связано со знаменитой алгеброй *кватернионов*, построенной Гамильтоном. Комплексные числа удобно отождествлять с точками плоскости, поскольку они имеют две координаты — вещественную часть и мнимую часть. (Для комплексного числа $z = x + iy$, действительное число x называется *действительной частью*, а действительное число y — *мнимой частью* числа z). По аналогии с комплексными числами, Гамильтон долго пытался построить "трехмерные числа", то есть наделить точки трехмерного пространства естественными операциями сложения и умножения, удовлетворяющими некоторым естественным свойствам. Однако, ему это не удалось. Более того, в некотором естественном смысле, таких "хороших" операций не существует. Все же поиски не были бесполезны. В результате своих поисков, Гамильтон наткнулся на замечательную и естественную конструкцию "четырехмерных" чисел — кватернионов.

Кватернионом называется выражение вида

$$x_1 + x_2 i + x_3 j + x_4 k,$$

в котором i, j и k — формальные символы, не являющиеся действительными числами. Эти символы удовлетворяют следующим соотношениям:

$$i^2 = j^2 = k^2 = -1, \quad ij == -ji = k,\ jk = -kj = i,\ ki = -ik = j.$$

Первая серия соотношений говорит о том, что каждое из чисел i, j, k является *мнимой единицей* (быть мнимой единицей значит давать -1 при возведении в квадрат). Вторая серия соотношений содержит две вещи. Первая — мнимые единицы i, j и k антикоммутируют. Другими словами, если в произведении любых двух из этих мнимых единиц мы поменяем порядок сомножителей, то от этого произведение поменяет знак. Конечно, мы при этом предполагаем, что сомножители разные. Кроме этого, вторая серия соотношений выражает произведение любых двух мнимых единиц из трех указанных через эти же самые мнимые единицы. Запомнить эти соотношения проще всего так. Запишем символы i, j, k по кругу в указанном порядке по часовой стрелке. Таким образом, за i следует j, за j следует k, за k следует i, и так далее. Произведение любых двух соседних символов в том порядке, в котором они идут по часовой стрелке, равно третьему символу. Произведение любых двух соседних символов, идущих один за другим против часовой стрелки, равно третьему символу со знаком минус.

Как складывать и перемножать произвольные кватернионы? Для этого нужно воспользоваться правилами умножения мнимых единиц i, j и k, а также всеми обычными законами сложения (коммутативность — переместительный закон, ассоциативность — сочетательный закон, и т.д.) и обычным законом дистрибутивности (распределительный закон умножения). Например,

$$(i + j) + (j + k) = i + 2j + k.$$

$$(i+j)(j+k) = ij + ik + j^2 + jk = k + (-j) + (-1) + i = -1 + i - j + k.$$

Задача. Докажите, что умножение кватернионов ассоциативно, то есть для любых трех кватернионов q, q' и q'' выполнено равенство $(qq')q'' = q(q'q'')$. *Указание:* сначала докажите это свойство в том случае, когда каждый из трех кватернионов q, q' и q'' совпадает с одной из мнимых единиц i, j или k. Общий случай сводится к этому при помощи закона дистрибутивности.

Задача. Найдите все кватернионы q, являющиеся мнимыми единицами, то есть удовлетворяющие соотношению $q^2 = -1$.

Задача. *Модуль* $|q|$ кватерниона $q = x_1 + x_2 i + x_3 j + x_4 k$ определяется как неотрицательное значение корня из неотрицательного действительного числа

$$x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

Определим *сопряженный кватернион* $\overline{q}$ формулой

$$\overline{q} = x_1 - x_2 i - x_3 j - x_4 k.$$

Покажите, что $q\overline{q} = \overline{q}q = |q|^2$.

Задача. Докажите, что для любой пары кватернионов q, q' выполнено соотношение

$$\overline{qq'} = \overline{q'} \cdot \overline{q}.$$

Задача. Докажите, что для любой пары кватернионов q, q' выполнено соотношение

$$|q|^2 |q'|^2 = |qq'|^2.$$

Эту формулу можно проинтерпретировать как формулу типа $(4,4,4)$ для произведения сумм квадратов.

А. Гурвиц в 1898 году поставил следующий вопрос, который до сих пор остается открытым:

Вопрос Гурвица. Для каких целых чисел r, s и n существует формула типа (r,s,n) для произведения сумм квадратов?

Этот вопрос имеет несколько вариантов. В формуле типа (r,s,n) сумма квадратов r переменных, умноженная на сумму квадратов s других переменных, представляется в виде суммы квадратов n билинейных форм от этих двух групп переменных. Однако коэффициенты в этих билинейных формах могут быть целыми, вещественными, комплексными и т.д. Ни в одной из этих ситуаций, на вопрос Гурвица не найдено полного ответа. Кажется, что ответ не должен зависеть от выбора коэффициентов. Для всех известных примеров формул с комплексными коэффициентами, существуют формулы того же типа с вещественными, и даже с целыми (!) коэффициентами.

В некоторых частных случаях, однако, ответы известны. Например, Гурвиц доказал следующую теорему. Формула типа (n,n,n) существует только для следующих четырех значений n: 1, 2, 4 и 8. Формула для $n = 1$

очевидна — она соответствует умножению действительных чисел. Формула для $n = 2$ соответствует умножению комплексных чисел. Формула для $n = 4$ соответствует умножению кватернионов. Наконец, формула для $n = 8$ соответствует умножению октав, или чисел Кэли, которые мы определим позже.

Можно ответить на вопрос Гурвица в случае $s = n$. Это сделал сам Гурвиц в конце жизни, через двадцать лет после того, как поставил свой вопрос. Ответ оказывается связан с *представлениями алгебр Клиффорда*, которые мы кратко обсудим позже. Сейчас сформулируем сам ответ. Формула типа (r, n, n) существует тогда и только тогда, когда число r не превосходит числа ρ, зависящего от n следующим образом. Пусть σ — наибольшая степень двойки, на которую делится число n. Разделим число σ на 4 с остатком. Обозначим через a неполное частное, а через b остаток. Тогда

$$\sigma = 4a + b, \quad 0 \leq b < 4.$$

Число ρ равно $8a + 2^b$.

Дополнительные упражнения.

Следующие задачи посвящены отображениям Хопфа и, в частности, доказательству теоремы об отображении Хопфа и окружностях. При первом чтении, решение этих задач рекомендуется отложить.

1. Докажите, что на всякой двумерной плоскости в четырехмерном пространстве можно выбрать 2 координаты (параметра) (ξ, η) так, что любая точка плоскости однозначно задается соответствующей парой параметров (ξ, η) и так, что координаты точки плоскости в объемлющем четырехмерном пространстве выражаются через параметры (ξ, η) следующим образом:

$$x_1 = a_1\xi + b_1\eta, \quad x_2 = a_2\xi + b_2\eta, \quad x_3 = a_3\xi + b_3\eta, \quad x_4 = a_4\xi + b_4\eta.$$

Для решения этой задачи желательно некоторое знакомство с линейной алгеброй.

2. В обозначениях предыдущей задачи, параметры ξ и η на двумерной плоскости в четырехмерном пространстве можно ввести таким образом, чтобы расстояние от точки плоскости с координатами (ξ, η) до начала координат выражалось формулой $\sqrt{\xi^2 + \eta^2}$. Тогда, в частности, пересечение трехмерной сферы с центром в начале координат и радиусом 1 с рассматриваемой плоскостью задается уравнением

$$\xi^2 + \eta^2 = 1$$

или параметрически

$$\xi = \cos\phi, \quad \eta = \sin\phi. \tag{4}$$

3. Рассмотрим двумерную плоскость в четырехмерном пространстве вместе с параметрами (ξ, η), введенными в предыдущей задаче. Пусть h — *квадратичная форма* на четырехмерном пространстве, то есть функция от координат x_1, x_2, x_3, x_4, равная сумме выражений вида $x_k x_m$ с действительными коэффициентами, где индексы k и m независимо друг от друга пробегают значения от 1 до 4. Докажите, что ограничение квадратичной формы h на окружность (4) задается формулой

$$a\cos(2\phi) + b\sin(2\phi) + c$$

для некоторых постоянных действительных коэффициентов a, b и c. *Указание:* воспользуйтесь формулами двойного угла из тригонометрии.

4. Рассмотрим отображение f из четырехмерного пространства в трехмерное, заданное формулами

$$\begin{aligned} y_1 &= f_1(x_1, x_2, x_3, x_4), \\ y_2 &= f_2(x_1, x_2, x_3, x_4), \\ y_3 &= f_3(x_1, x_2, x_3, x_4). \end{aligned}$$

Допустим, что функции f_1, f_2, f_3 являются квадратичными формами. Такое отображение f называется *квадратичным отображением.* В частности, отображение Хопфа является квадратичным отображением.

Докажите, что образ всякой большой окружности на трехмерной сфере при квадратичном отображении содержится в некоторой двумерной плоскости, не обязательно проходящей через начало координат. Более того, этот образ является эллипсом.

5. Докажите теорему об отображении Хопфа и окружностях.

6. (Кватернионные расслоения Хопфа) Определим квадратичное отображение из семимерной сферы в четырехмерную сферу следующим образом. Семимерная сфера вложена в восьмимерное пространство. Каждая точка восьмимерного пространства имеет 8 координат, которые можно разбить на 2 группы по 4 координаты в каждой. Четыре координаты каждой группы можно отождествить с кватернионом. Таким образом, точка восьмимерного пространства представляется парой кватернионов (x, x'). Определим отображение f из восьмимерного пространства в пятимерное по следующей формуле:

$$f : (x, x') \mapsto (2xx', |x|^2 - |x'|^2).$$

В правой части этой формулы записана в координатах точка пятимерного пространства: кватернион $2xx'$ представляет первые четыре координаты, а действительное число $|x|^2 - |x'|^2$ представляет пятую координату. Отображение f называется левым *кватернионным расслоением Хопфа.* Есть еще правое кватернионное расслоение Хопфа, которое задается аналогичной формулой с той лишь разницей, что произведение кватернионов x и x' взято в другом порядке. Докажите, что оба кватернионных расслоения Хопфа переводят семимерную сферу в четырехмерную сферу. Кроме того, оба кватернионных расслоения Хопфа переводят все большие окружности на семимерной сфере в окружности на четырехмерной сфере.

2 Квадратичные формы и решетки

Вернемся к задаче о числах, представимых в виде суммы квадратов двух целых чисел. Пусть $a = x_1^2 + x_2^2$ для некоторых целых чисел x_1 и x_2. Тогда a — это квадрат расстояния от точки с координатами (x_1, x_2) до начала координат. Отметим на плоскости все точки, обе координаты которых являются целыми числами. Полученный геометрический объект называется *целочисленной решеткой.* Целочисленная решетка (в определенном масштабе) хорошо видна на любой тетрадке в клеточку — как множество вершин всех клеток.

Рассмотрим две точки целочисленной решетки — A и B, такие, что прямая, проходящая через A и B, не содержит начала координат. Обозначим начало координат через O. Выбор точек A и B определяет некоторую систему координат на плоскости, отличную от исходной. Пара координат точки X определяется следующим образом (см. рис. 2). Проведем через точку X прямую, параллельную прямой OB, до пересечения с прямой OA. Обозначим точку пересечения через X_1. Аналогично, проведем через точку X прямую, параллельную прямой OA, до пересечения с прямой OB. Обозначим точку пересечения через X_2. Теперь координаты точки X определяются как отношения длин, взятые со знаком:

$$\tilde{x}_1 = \pm\frac{|OX_1|}{|OA|}, \quad \tilde{x}_2 = \pm\frac{|OX_2|}{|OA|}$$

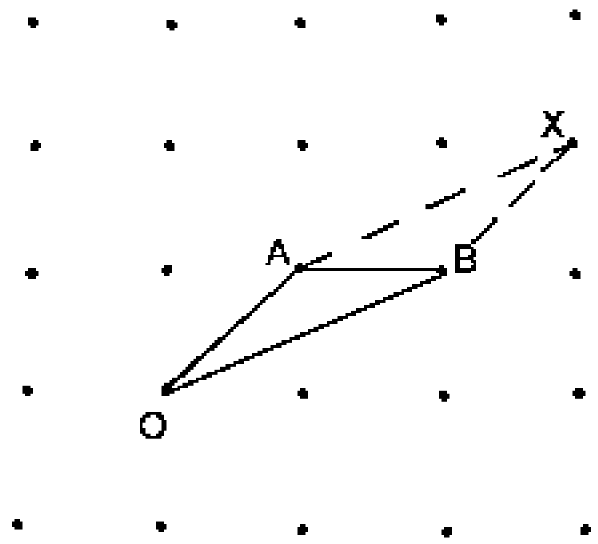

Рис. 2: целочисленная решетка.

Тильды над координатами поставлены для того, чтобы отличать новые координаты (с тильдами) от старых (без тильд). Знаки определяются следующим стандартным способом. Если точки X_1 и A находятся по одну сторону от точки O, то координата $\tilde{x}_1$ берется со знаком плюс, а если по разные стороны, то со знаком минус. То же самое для второй координаты.

Задача. Пусть (p_{11}, p_{21}) — координаты точки A относительно старой системы координат. Аналогично, обозначим через (p_{12}, p_{22}) старые координаты точки B. Поскольку точки A и B принадлежат целочисленной решетке, все числа p_{11}, p_{21}, p_{12} и p_{22} целые. Докажите, что старые координаты (x_1, x_2) любой точки плоскости выражаются через новые координаты $(\tilde{x}_1, \tilde{x}_2)$ следующим образом:

$$\begin{aligned} x_1 &= p_{11}\tilde{x}_1 + p_{12}\tilde{x}_2 \\ x_2 &= p_{21}\tilde{x}_1 + p_{22}\tilde{x}_2. \end{aligned}$$

Таблица

$$\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

называется *матрицей замены базиса.* Точки A и B однозначно определяются матрицей замены базиса.

Задача. Докажите, что новые координаты $(\tilde{x}_1, \tilde{x}_2)$ выражаются через старые координаты (x_1, x_2) следующим образом:

$$\begin{aligned} \tilde{x}_1 &= \tilde{p}_{11}x_1 + \tilde{p}_{12}x_2 \\ \tilde{x}_2 &= \tilde{p}_{21}x_1 + \tilde{p}_{22}x_2. \end{aligned}$$

Здесь коэффициенты $\tilde{p}_{11}$, $\tilde{p}_{12}$, $\tilde{p}_{21}$ и $\tilde{p}_{22}$ — рациональные числа.

Задача. Докажите, что площадь треугольника AOB равна

$$\left|\frac{p_{11}p_{22} - p_{12}p_{21}}{2}\right|.$$

Задача. Докажите, что произведение каждого из чисел $\tilde{p}_{11}$, $\tilde{p}_{12}$, $\tilde{p}_{21}$ и $\tilde{p}_{22}$ на

$$p_{11}p_{22} - p_{12}p_{21}$$

является целым числом.

Задача. Числа $\tilde{p}_{11}$, $\tilde{p}_{12}$, $\tilde{p}_{21}$ и $\tilde{p}_{22}$ целые тогда и только тогда, когда площадь треугольника AOB равна $1/2$.

Задача. Докажите, что в новой системе координат квадрат расстояния от точки X до начала координат выражается формулой

$$a\tilde{x}_1^2 + b\tilde{x}_1\tilde{x}_2 + c\tilde{x}_2^2, \tag{5}$$

в которой целые числа a, b и c не зависят от точки X (а зависят только от точек A и B, то есть от выбора новой системы координат). Более точно,

$$\begin{aligned} a &= p_{11}^2 + p_{21}^2 \\ b &= 2(p_{11}p_{12} + p_{21}p_{22}) \\ c &= p_{12}^2 + p_{22}^2. \end{aligned}$$

В частности, мы видим, что число b всегда четно.

Выражение (5) называется *квадратичной формой.* Если все три коэффициента a, b и c — целые числа, то это выражение называется *целочисленной квадратичной формой.* Пара точек A и B, для которой площадь треугольника AOB равна $1/2$, называется *базисом целочисленной решетки.* Начало координат мы всегда предполагаем фиксированным — мы меняли базис, но не меняли начало координат.

Задача. Докажите, что пара точек A и B образует базис целочисленной решетки тогда и только тогда, когда множество точек, имеющих целые координаты в новой системе координат (построенной по точкам A и B) совпадает с целочисленной решеткой.

Выбирая разные базисы целочисленной решетки, мы получим много разных целочисленных квадратичных форм. Однако все эти квадратичные формы будут иметь одно и то же множество значений. Под *значением* целочисленной квадратичной формы мы всегда понимаем результат подстановки в эту форму некоторых *целых* чисел $\tilde{x}_1$ и $\tilde{x}_2$. Другими словами, множество чисел, представимых в виде (5) для целых $\tilde{x}_1$ и $\tilde{x}_2$, будет одним и тем же, независимо от выбора точек A и B. Действительно, во всех случаях это множество составлено из квадратов расстояний от точек целочисленной решетки до начала координат. Поэтому это множество не может зависеть от выбора точек A и B.

Таким образом, мы приходим к следующему важному выводу. Изучать значения квадратичной формы $x_1^2 + x_2^2$ — это то же самое, что изучать значения любой другой квадратичной формы, полученной из этой изменением базиса целочисленной решетки. Например, с тем же успехом можно было бы рассматривать формы

$$2x_1^2 + 2x_1x_2 + x_2^2, \quad 5x_1^2 - 14x_1x_2 + 10x_2^2, \quad \dots$$

В частности, для всех этих форм верно следующее утверждение: произведение двух значений снова является значением.

Целочисленная квадратичная форма, получающаяся из формы $x_1^2 + x_2^2$ заменой базиса целочисленной решетки, считается *эквивалентной* форме $x_1^2 + x_2^2$. Как по коэффициентам a, b и c целочисленной квадратичной формы понять, является ли эта форма эквивалентной форме $x_1^2 + x_2^2$? Чтобы ответить на этот вопрос, нам понадобится понятие *дискриминанта* квадратичной формы.

Целочисленную форму удобно обозначать, просто перечисляя ее коэффициенты. Договоримся, что под формой (a, b, c) имеется в виду квадратичная форма с коэффициентами a, b и c, то есть форма, заданная выражением (5). *Дискриминантом* квадратичной формы (a, b, c) называется число $b^2 - 4ac$, хорошо известное по школьному учебнику.

Теорема об эквивалентности. *Квадратичная форма (a, b, c) эквивалентна форме $(1, 0, 1)$ тогда и только тогда, когда $a > 0$ и дискриминант $b^2 - 4ac$ равен -4.*

Прежде чем доказывать теорему об эквивалентности, опишем геометрический смысл положительно определенных целочисленных квадратичных форм. Квадратичная форма (a, b, c) *положительно* *определена*, если действительное число $ax_1^2+bx_1x_2+cx_2^2$ положительно при всех действительных значениях x_1 и x_2, не равных одновременно нулю.

Положительная определенность и дискриминант. *Квадратичная форма* (a,b,c) *положительно определена тогда и только тогда, когда*

- *дискриминант* b^2-4ac *отрицателен,*
- *коэффициент* a *положителен.*

Докажем эту теорему в одну сторону. Пусть квадратичная форма (a,b,c) положительно определена. Тогда при всех ненулевых значениях x_1 и x_2, имеем

$$ax_1^2+bx_1x_2+cx_2^2>0.$$

Положим $x_2=1$. Получим, что квадратичный трехчлен $ax_1^2+bx_1+c$ принимает только положительные значения. В частности, уравнение $ax_1^2+bx_1+c=0$ не имеет действительных корней. Это означает, что дискриминант b^2-4ac этого уравнения отрицателен.

Из определения положительной определенности формы (a,b,c), в частности, следует, что выражение ax_1^2 положительно при ненулевых действительных значениях x_1 (достаточно положить $x_2=0$). Следовательно, $a>0$. $\square$

Задача. Докажите эту теорему в другую сторону.

Напомним некоторые определения из геометрии векторов. Они все входят в школьную программу. Если Вы хорошо помните операции над векторами и их основные свойства, следующий абзац можно пропустить. *Вектором* называется ориентированный отрезок, рассматриваемый с точностью до параллельного переноса. Таким образом, любая пара точек A,B на плоскости определяет вектор $\vec{AB}$. Если $ABCD$ — параллелограмм, причем вершины написаны в порядке обхода вокруг границы, то векторы $\vec{AB}$ и $\vec{DC}$ считаются равными. Это правило позволяет откладывать любой вектор от любой точки плоскости. *Суммой* двух векторов $\vec{AB}$ и $\vec{BC}$ называется вектор $\vec{AC}$:

$$\vec{AB}+\vec{BC}=\vec{AC}$$

Чтобы воспользоваться этим определением, нужно второй вектор отложить от точки, в которой заканчивается первый вектор. *Произведением* вектора $\vec{AB}$ на действительное число α называется вектор $\vec{AC}$, для которого отношение длин $|AC|/|AB|$ равно α, причем если $\alpha>0$, то точки B и C находятся по одну сторону от точки A, а если $\alpha<0$ — то по разные стороны. Произведение вектора $\vec{v}$ на число α обозначается через $\alpha\vec{v}$. *Скалярным произведением* двух векторов $\vec{AB}$ и $\vec{AC}$ называется число

$$\langle\vec{AB},\vec{AC}\rangle=|AB|\cdot|AC|\cdot\cos(\phi),$$

где ϕ — это угол BAC. Заметим, что все приведенные выше определения имеют смысл не только на плоскости, но и в трехмерном пространстве (и даже в многомерном пространстве).

В следующих задачах содержатся наиболее важные свойства операций над векторами.

Задача. Докажите следующие основные свойства сложения векторов:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}, \quad (\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}).$$

Задача. Вектор, начинающийся и заканчивающийся в одной и той же точке, называется *нулевым вектором* и обозначается через $\vec{0}$. Докажите, что $\vec{a} + \vec{0} = \vec{a}$ для каждого вектора $\vec{a}$. Кроме того, существует единственный вектор $-\vec{a}$, для которого $\vec{a} + (-\vec{a}) = \vec{0}$.

Сумма $\vec{a} + (-\vec{b})$ обозначается через $\vec{a} - \vec{b}$.

Задача. Докажите, что если X — точка с координатами (x_1, x_2) (относительно декартовой прямоугольной системы координат), а Y — точка с координатами (y_1, y_2), то

$$\langle \vec{OX}, \vec{OY} \rangle = x_1 y_1 + x_2 y_2.$$

Здесь O — это начало координат.

Задача. Докажите следующие свойства скалярного произведения:

$$\langle \vec{a}, \vec{b} \rangle = \langle \vec{b}, \vec{a} \rangle, \quad \langle \vec{a} + \vec{b}, \vec{c} \rangle = \langle \vec{a}, \vec{c} \rangle + \langle \vec{b}, \vec{c} \rangle.$$

Следующая теорема описывает геометрический смысл положительно определенных форм.

Теорема. *Для каждой положительно определенной квадратичной формы (a, b, c) найдутся векторы $\vec{u}$ и $\vec{v}$, такие, что*

$$\langle \vec{u}, \vec{u} \rangle = a, \quad 2\langle \vec{u}, \vec{v} \rangle = b, \quad \langle \vec{v}, \vec{v} \rangle = c. \tag{6}$$

Если же, наоборот, коэффициенты квадратичной формы (a, b, c) заданы равенствами (6) для некоторых (произвольных) векторов $\vec{u}$ и $\vec{v}$, не пропорциональных друг другу, то форма (a, b, c) положительно определена.

Докажем первую часть этой теоремы. Рассмотрим произвольный вектор $\vec{u}$, длина которого равна $\sqrt{a}$. Если выбрать вектор $\vec{v}$ таким образом, чтобы его длина была равна $\sqrt{c}$, то первое и последнее равенства в (6) будут выполнены. Число c положительно, так как дискриминант $b^2 - 4ac$ отрицателен. Чтобы выполнялось среднее равенство в (6), следует специальным образом подобрать угол между векторами $\vec{u}$ и $\vec{v}$ (напомним, что длины этих векторов уже зафиксированы).

Поскольку квадратичная форма (a, b, c) положительно определена, ее дискриминант $b^2 - 4ac$ отрицателен. Следовательно,

$$-1 \le \frac{b}{2\sqrt{ac}} \le 1.$$

Значит, существует такой угол ϕ, что $\cos\phi = b/2\sqrt{ac}$. Положим угол между $\vec{u}$ и $\vec{v}$ равным ϕ. Тогда нетрудно проверить, что $\langle \vec{u}, \vec{v} \rangle = b/2$, то есть среднее равенство в (6) выполнено.

Для доказательства второй части теоремы заметим, что

$$ax_1^2 + bx_1x_2 + cx_2^2 = \langle x_1\vec{u} + x_2\vec{v}, x_1\vec{u} + x_2\vec{v}\rangle.$$

В правой части этого равенства стоит квадрат длины вектора $x_1\vec{u} + x_2\vec{v}$. Это положительное число при всех действительных x_1 и x_2, не равных одновременно нулю. □

Пользуясь обозначениями только что доказанной теоремы, отложим векторы $\vec{u}$ и $\vec{v}$ от одной и той же точки B. Обозначим через C конец вектора $\vec{u}$, а через A конец вектора $\vec{v}$. Полученный треугольник ABC назовем *фундаментальным треугольником* квадратичной формы (a, b, c). Доказанная выше теорема означает, что у каждой положительно определенной квадратичной формы есть фундаментальный треугольник, и что каждый треугольник на плоскости является фундаментальным треугольником некоторой квадратичной формы. Из приведенного доказательства также следует, что *фундаментальный треугольник положительно определенной квадратичной формы определен однозначно с точностью до евклидовых движений.* В частности, все углы и длины всех сторон этого треугольника определены однозначно.

Задача. Вычислите длины всех сторон фундаментального треугольника по коэффициентам соответствующей квадратичной формы.

Задача. Докажите, что целочисленной положительно определенной квадратичной форме соответствует фундаментальный треугольник, квадраты длин сторон которого являются целыми числами. Наоборот, пусть квадраты длин сторон некоторого треугольника на плоскости являются целыми числами. Тогда этот треугольник является фундаментальным треугольником некоторой целочисленной квадратичной формы.

Имеется замечательная связь между дискриминантом положительно определенной квадратичной формы и площадью ее фундаментального треугольника.

Теорема о дискриминанте и площади. *Пусть (a, b, c) — положительно определенная квадратичная форма, а ABC — ее фундаментальный треугольник. Тогда площадь S треугольника ABC выражается через дискриминант $\Delta = b^2 - 4ac$ формы (a, b, c) следующим образом:*

$$S^2 = -\Delta/16.$$

Задача. Докажите эту теорему.

Как представить себе геометрически множество значений положительно определенной целочисленной квадратичной формы? Пусть $\vec{u}$ и $\vec{v}$ — векторы, представленные сторонами $\vec{BC}$ и $\vec{BA}$ фундаментального треугольника ABC целочисленной квадратичной формы (a, b, c). Тогда, как мы видели, значение $ax_1^2 + bx_1x_2 + cx_2^2$ нашей квадратичной формы на паре целых чисел x_1 и x_2 равно квадрату длины вектора $x_1\vec{u} + x_2\vec{v}$.

Рассмотрим теперь векторы вида $x_1\vec{u} + x_2\vec{v}$ для всех возможных целых чисел x_1 и x_2. Множество концов этих векторов называется *решеткой*, соответствующей квадратичной форме (a, b, c). Как мы видели выше, квадратичной форме $(1, 0, 1)$ соответствует стандартная целочисленная решетка. Пара векторов $(\vec{u}, \vec{v})$ называется *базисом решетки*. Заметим, что одну и ту же решетку можно получить, исходя из разных базисов. Множество значений квадратичной формы определяется ее решеткой. Именно, значения формы — это в точности квадраты расстояний между точками решетки.

Вообще, *решетка на плоскости* определяется как множество всех векторов вида $x_1\vec{u} + x_2\vec{v}$, где коэффициенты x_1 и x_2 пробегают все целые числа, а векторы $\vec{u}$ и $\vec{v}$ фиксированы. Эти векторы предполагаются не параллельными. Решетка на плоскости не обязательно соответствует целочисленной квадратичной форме.

Задача. Решетка на плоскости соответствует целочисленной квадратичной форме тогда и только тогда, когда квадраты длин всех векторов решетки являются целыми числами. Мы будем называть такие решетки *целочисленно нормированными*.

Теперь мы можем ввести определение эквивалентности для произвольных целочисленных квадратичных форм. Рассмотрим некоторую целочисленно нормированную решетку на плоскости. Она допускает различные базисы. Выбор базиса $(\vec{u}, \vec{v})$ данной решетки однозначно определяет коэффициенты некоторой квадратичной формы (a, b, c) по формуле (6). Однако для разных базисов мы получаем разные квадратичные формы. Такие квадратичные формы (которые соответствуют разным базисам одной и той же решетки) называются *эквивалентными*. Поскольку множество значений формы зависит только от решетки, но не от конкретного базиса, эквивалентные формы имеют одно и то же множество значений. Ранее мы уже определяли, что значит эквивалентность данной квадратичной формы квадратичной форме $(1, 0, 1)$. Нетрудно понять, что это определение согласуется с приведенным только что более общим определением. Разница в том, что вместо стандартной целочисленной решетки, рассмотрением которой мы ранее ограничивались, мы сейчас рассматриваем и другие решетки.

Посмотрим, как связаны между собой фундаментальные треугольники эквивалентных квадратичных форм.

Прежде всего, заметим, что если поменять порядок вершин фундаментального треугольника, то полученный треугольник будет соответствовать квадратичной форме, которая эквивалентна исходной. Напомним, что при определении фундаментального треугольника порядок вершин был важен. Однако, легко видеть, что решетка, порожденная фундаментальным треугольником, не зависит от порядка вершин. Чтобы получить решетку, нужно просто достроить треугольник до параллелограмма, а потом замостить всю плоскость параллельными переносами этого параллелограмма. Геометрически очевидно, что результат не зависит от того, до какого параллелограмма достраивать.

Рассмотрим теперь параллелограмм $ABCD$ (вершины расположены в порядке обхода) и допустим, что треугольник ABC является фундаментальным треугольником некоторой целочисленной квадратичной формы.

Задача. Докажите, что треугольник BCD тоже является фундаментальным треугольником некоторой целочисленной квадратичной формы. Более того, квадратичные формы с фундаментальными треугольниками ABC и BCD эквиваленты.

Назовем треугольники ABC и BCD *соседними*. Таким образом, квадратичные формы с соседними фундаментальными треугольниками эквивалентны. Оказывается, что обратное утверждение тоже верно.

Теорема. *Фундаментальные треугольники любых двух эквивалентных целочисленных квадратичных форм можно включить в цепочку треугольников, такую, что два последовательных треугольника цепочки являются соседними в смысле приведенного выше определения.*

Доказать эту теорему сложно, но все же полезно попытаться. Мы не будем приводить доказательства, так как это утверждение не будет использоваться в дальнейшем.

Следующее утверждение вытекает из предыдущей теоремы, но мы его докажем из других соображений. Оно нам понадобится при доказательстве теоремы об эквивалентности.

Теорема о дискриминантах эквивалентных форм. *Дискриминанты эквивалентных квадратичных форм совпадают. Другими словами, фундаментальные треугольники эквивалентных форм имеют одну и ту же площадь.*

Доказательство. Мы будем доказывать это утверждение во второй, геометрической, формулировке. *Фундаментальным параллелограммом* решетки называется параллелограмм, составленный из двух фундаментальных треугольников. Конечно, можно составить много фундаментальных параллелограммов, отвечающих одной и той же решетке. Решетка получается из любого своего фундаментального параллелограмма следующим образом. Нужно рассмотреть параллельные переносы фундаментального параллелограмма, плотно примыкающие друг к другу. Из них, как из кирпичиков, выстраивается решетка.

Пусть K — круг достаточно большого радиуса. Число точек решетки внутри этого круга, как нетрудно видеть, примерно равно площади круга, делённой на площадь фундаментального параллелограмма. Дело в том, что к каждой точке решетки, лежащей внутри круга K, примыкает фундаментальный параллелограмм, причем почти все эти параллелограммы лежат внутри K, и объединение этих параллелограммов почти совпадает с K. Это все “почти” верно, а не просто верно, поскольку возникают проблемы около границы круга K. Однако число точек решетки, находящихся около границы круга K, растет примерно как радиус, а количество всех точек в круге растет примерно как квадрат радиуса. Стало быть, точками около границы можно пренебречь.

Получаем следующий результат: площадь фундаментального параллелограмма равна пределу отношения площади круга K к числу точек решетки, находящихся внутри этого круга, когда радиус стремится к бесконечности. Это описание никак не апеллирует к форме фундаментального параллелограмма. Оно использует только геометрические свойства самой решетки. Следовательно, все фундаментальные параллелограммы одной и той же решетки имеют одну и ту же площадь. Осталось только заметить, что площадь фундаментального треугольника вдвое меньше площади фундаментального параллелограмма. □

Теперь мы готовы к тому, чтобы доказать теорему об эквивалентности.

Доказательство теоремы об эквивалентности. Зафиксируем некоторую целочисленную квадратичную форму $f = (\tilde{a}, \tilde{b}, \tilde{c})$, удовлетворяющую условию теоремы об эквивалентности: коэффициент $\tilde{a} > 0$ и дискриминант $\tilde{b}^2 - 4\tilde{a}\tilde{c} = -4$. Поскольку коэффициент $\tilde{a}$ положителен, а дискриминант отрицателен, форма f положительно определена. Значит, имеет смысл говорить о ее фундаментальном треугольнике и о фундаментальных треугольниках всех форм, эквивалентных форме f.

Мы хотим доказать, что форма f эквивалентна форме $(1, 0, 1)$. Рассмотрим все целочисленные квадратичные формы, эквивалентные форме f. Среди них выберем форму, фундаментальный треугольник которой имеет наименьший периметр. Обозначим этот фундаментальный треугольник через ABC.

Достроим треугольник ABC до параллелограмма $ABCD$. Мы, как всегда, предполагаем, что вершины параллелограмма $ABCD$ написаны в порядке обхода его границы. Из двух диагоналей AC и BD параллелограмма $ABCD$, диагональ AC не должна быть большей. В самом деле, если длина отрезка AC превышает длину отрезка BD, то ровно на столько же периметр треугольника DBC превышает периметр треугольника ABC. Таким образом, $|AC| \leq |BD|$. Мы не исключаем равенства.

Напротив меньшей диагонали параллелограмма лежит острый угол, а если диагонали равны, то все углы параллелограмма прямые. Таким образом, в любом случае угол B треугольника ABC не превышает 90 градусов. Аналогичные рассуждения показывают, что любой угол треугольника ABC не превышает 90 градусов: нужно только рассмотреть другие способы достроить этот треугольник до параллелограмма. Таких способов три: треугольник можно отразить относительно середины любой из сторон, объединение полученного треугольника с исходным будет параллелограммом. Итак, ни один угол треугольника ABC не может быть тупым. Другими словами, наш треугольник не может быть тупоугольным.

Итак, у нас есть треугольник ABC, не являющийся тупоугольным. Кроме того, мы знаем, что это фундаментальный треугольник некоторой целочисленной положительно определенной квадратичной формы, и что дискриминант этой формы равен -4. Из целочисленности формы вытекает, что квадраты длин всех сторон треугольника ABC являются целыми числами. Из того, что дискриминант формы равен -4, вытекает, что площадь треугольника ABC равна $1/2$. Возникает задача: описать нетупоугольные треугольники площади $1/2$, такие что квадраты длин всех сторон являются целыми числами. Мы докажем, что это только прямоугольные равнобедренные треугольники, причем длина катетов равна 1.

Обозначим через a квадрат длины стороны BC, а через c — квадрат длины стороны AB. Скалярное произведение $\langle \vec{BA}, \vec{BC} \rangle$ обозначим через b. По теореме косинусов, квадрат длины стороны AC равен $a + c - b$. Числа a, b и c определены таким образом, что квадратичная форма (a, b, c) эквивалентна квадратичной форме $f = (\tilde{a}, \tilde{b}, \tilde{c})$. Без ограничения общности, можно считать, что квадраты длин сторон упорядочены следующим образом:

$$0 \ \leq \ a = |BC|^2 \ \leq \ c = |AB|^2 \ \leq \ a + c - b = |AC|^2.$$

Иначе достаточно просто переименовать вершины треугольника ABC так, чтобы эти неравенства были выполнены.

Поскольку треугольник ABC не является тупоугольным, число b неотрицательно. Заметим, что это число равно нулю тогда и только тогда, когда угол B прямой. Из написанных выше неравенств следует, что

$$0 \leq b \leq a \leq c.$$

Поскольку форма (a, b, c) эквивалентна форме f дискриминанта -4 (иначе, поскольку площадь треугольника ABC равна $1/2$), имеем $b^2 - 4ac = -4$. С другой стороны,

$$-4 = b^2 - 4ac \leq a^2 - 4a^2 = -3a^2.$$

Это неравенство получено заменой неотрицательного числа b на бо́льшее число a, с одновременной заменой числа c на меньшее число a. Доказанное неравенство означает, что $a^2 \leq 4/3$. Поскольку a является положительным целым числом, отсюда следует, что $a = 1$. Мы знаем, что $0 \leq b \leq a$, а значит, b равно либо 0, либо 1. С другой стороны, дискриминант равен $b^2 - 4c = -4$, откуда следует, что b четно. Таким образом, $b = 0$. Коэффициент c находится из уравнения на дискриминант: $c = 1$.

Мы видим, что $(a, b, c) = (1, 0, 1)$. С другой стороны, квадратичная форма (a, b, c) эквивалентна квадратичной форме f. Значит, форма f эквивалентна форме $(1, 0, 1)$, что и требовалось доказать. □

Предложенное доказательство теоремы об эквивалентности не является самым простым. Проще всего было бы доказывать это утверждение алгебраически, вообще не упоминая про фундаментальные треугольники. Однако геометрический подход с фундаментальными треугольниками представляется более естественным. До него проще додуматься.

Подчеркнем еще раз основную идею доказательства: чтобы привести данную квадратичную форму к стандартной, нужно выбрать форму, эквивалентную данной и имеющую фундаментальный треугольник наименьшего периметра. На самом деле, эта идея работает в более широком контексте: квадратичную форму произвольного дискриминанта можно привести к эквивалентной форме из конечного списка "стандартных" форм. Это основная идея *теории редукции* Лагранжа.

Дополнительные упражнения.

Следующие задачи призваны объяснить связь между композицией квадратичных форм и геометрией решеток на комплексной плоскости. Операция композиции была введена Гауссом. Несомненно, Гаусс знал геометрическую интерпретацию этой операции. Однако он старался не упоминать про комплексные числа без крайней необходимости (в те времена комплексные числа еще не пользовались популярностью), и поэтому ничего не написал про геометрическую интерпретацию. Про нее гораздо позже написал Ф. Клейн.

1. Будем отождествлять комплексные числа с точками плоскости по следующему правилу. Комплексному числу $z = x + yi$ соответствует точка плоскости с координатами (x, y). Докажите, что пара комплексных чисел

$$\sqrt{a}, \quad \frac{b+\sqrt{b^2-4ac}}{2\sqrt{a}}$$

образует базис решетки, соответствующей квадратичной форме (a, b, c).

3. Скажем, что целочисленно нормированная решетка L *минимальна*, если наибольший общий делитель квадратов длин всех векторов из L равен 1. Докажите, что, умножая все векторы произвольной целочисленно нормированной решетки на одно и то же действительное число, можно получить минимальную решетку.

2*. Рассмотрим две минимальные целочисленно нормированные решетки L и L', у которых площади фундаментальных параллелограммов совпадают. Определим множество LL', состоящее из всех комплексных чисел вида

$$\lambda_1 z_1 z_1' + \cdots + \lambda_n z_n z_n',$$

где комплексные числа $z_1, \ldots, z_n$ пробегают независимо друг от друга все точки решетки L, комплексные числа $z_1', \ldots, z_n'$ пробегают все точки решетки L', а коэффициенты $\lambda_1, \ldots, \lambda_n$ пробегают все целые числа. Докажите, что множество LL' тоже является минимальной целочисленно нормированной решеткой, причем площадь фундаментального параллелограмма этой решетки такая же, как и у решеток L и L'.

3. Рассмотрим две целочисленные квадратичные формы f и f' с одинаковым дискриминантом D. Предположим, что соответствующие решетки L и L' минимальны. Тогда *композицией* квадратичных форм f и f' называется форма f'', соответствующая решетке LL'. Дискриминант формы f'', очевидно, равен D. Докажите, что если a — некоторое значение формы f, а a' — некоторое значение формы f', то aa' является значением формы f''.

4. Пусть $f = (a, b, c)$, $f' = (a', b', c')$ и $f'' = (a'', b'', c'')$ — формы из предыдущей задачи. Тогда можно написать формулу

$$(ax_1^2 + bx_1x_2 + cx_2^2)(a'y_1^2 + b'y_1y_2 + c'y_2^2) = a''z_1^2 + b''z_1z_2 + c''z_2^2,$$

в которой z_1, z_2 являются билинейными формами от x_1, x_2 и y_1, y_2 с целыми коэффициентами. Выписанную выше формулу назовем *формулой композиции*. Эта формула является естественным обобщением формулы для произведения сумм квадратов.

5. Найдите композицию квадратичных форм $(2, 1, 2)$ и $(4, 1, 1)$. Выпишите соответствующую формулу композиции.

6. Целочисленная квадратичная форма называется *примитивной*, если соответствующая ей решетка минимальна. Докажите, что форма (a, b, c) примитивна тогда и только тогда, когда три целых числа a, b, c взаимно просты в совокупности, то есть наибольший общий делитель всех трех чисел равен 1.

7. Докажите, что композиция формы $(1, 0, d)$ с любой другой примитивной формой f дискриминанта $-4d$ эквивалентна f.

3 Суммы квадратов

В этом разделе мы опишем множество всех значений квадратичной формы $(1, 0, 1)$. Соответствующая теорема — теорема о суммах квадратов — была сформулирована в первом разделе. Приводимое доказательство является далеко не самым коротким. Однако оно поучительно по двум причинам. Во-первых, по ходу дела будут обсуждаться вещи, представляющие самостоятельный интерес или важные для других задач теории чисел. Во-вторых, многие рассуждения можно существенно обобщить. Например, они почти дословно годятся для описания значений некоторых других (но не всех!) целочисленных квадратичных форм.

Сначала обсудим одно свойство значений этой квадратичной формы, которое легко увидеть экспериментально:

Остатки от деления сумм квадратов на 4. *Остаток от деления числа вида $x^2 + y^2$ на 4 не может быть равен 3 (числа x и y предполагаются целыми).*

Для доказательства достаточно перебрать все возможные остатки чисел x и y от деления на 4.

Задача. Докажите следующие утверждения:

- квадрат любого четного числа делится на 4
- квадрат нечетного целого числа дает остаток 1 при делении на 4

Выведите теорему об остатках от деления сумм квадратов на 4 из этих утверждений.

Мы уже знаем, что произведение двух значений снова является значением. Поэтому естественно попытаться исследовать именно мультипликативные (то есть связанные с умножением) свойства чисел, представимых в виде сумм квадратов. Например, какие простые числа можно представить в виде суммы квадратов двух целых чисел?

При помощи теоремы об эквивалентности, мы можем доказать следующее утверждение:

Лемма. *Если x — целое число, а число x^2+1 делится на простое число p, то p представляется в виде суммы квадратов двух целых чисел.*

Доказательство. Действительно, $x^2 + 1 = py$ для некоторого целого числа y. Мы можем переписать это равенство в таком виде: $(2x)^2 - 4py = -4$. Это значит, что дискриминант целочисленной квадратичной формы $(p, 2x, y)$ равен -4. По теореме об эквивалентности, эта квадратичная форма эквивалентна форме $(1, 0, 1)$. В частности, множество значений формы $(p, 2x, y)$ совпадает с множеством значений формы $(1, 0, 1)$, т.е. с множеством чисел, представимых в виде суммы квадратов. Заметим, что p является значением формы $(p, 2x, y)$:

$$p = p \cdot 1^2 + (2x) \cdot 1 \cdot 0 + y \cdot 0^2$$

Следовательно, p является значением формы $(1, 0, 1)$, что и требовалось доказать. $\square$

Для дальнейшего нам потребуется дополнительная терминология. Скажем, что два числа a и b *сравнимы по модулю* m, если разность $a - b$ делится на m. Модуль m может быть любым ненулевым целым числом. Если числа a и b сравнимы по модулю m, то мы будем иногда писать

$$a \equiv b \pmod m$$

Следующие задачи описывают основные свойства сравнений.

Задача. Докажите, что сравнения по фиксированному модулю m, как и равенства, можно почленно складывать и умножать. Другими словами, из сравнений

$$a \equiv b \pmod m, \quad c \equiv d \pmod m$$

вытекают сравнения

$$a + c \equiv b + d \pmod m, \quad ac \equiv bd \pmod m.$$

Задача. Из сравнений

$$a \equiv b \pmod m, \quad b \equiv c \pmod m$$

вытекает сравнение

$$a \equiv c \pmod m.$$

Особенно удобно работать со сравнениями по простому модулю. Дело в том, что для таких сравнений есть аналог деления. Тот факт, что действительные числа можно делить друг на друга, выражается формально следующим образом: если действительное число $x \neq 0$, то существует действительное число y, такое что $xy = 1$. То же самое верно для сравнений по простому модулю p.

Свойство деления для сравнений по простому модулю. *Если целое число x не сравнимо с нулем по модулю p (то есть не делится на p), то существует целое число y, такое, что*

$$xy \equiv 1 \pmod{p}.$$

Это утверждение вытекает из алгоритма Евклида для нахождения наибольшего общего делителя чисел x и p, который, конечно, равен 1. Согласно алгоритму Евклида, наибольший общий делитель, в данном случае 1, представляется как $xy + pz$ для некоторых целых коэффициентов y и z. Из равенства $xy = 1 - pz$ и вытекает искомое сравнение.

То же самое соображение используется при доказательстве *основной теоремы арифметики*, утверждающей, что разложение целого числа на простые множители единственно с точностью до перестановки этих множителей. Поэтому, если Вы не знакомы с алгоритмом Евклида, обязательно изучите его.

Лемма. *Предположим, $x^2 + y^2$ делится на простое число p, но по крайней мере одно из целых чисел x или y не делится на p. Тогда найдется такое целое число z, что $z^2 + 1$ делится на p.*

Доказательство. Без ограничения общности, мы можем предполагать, что y не делится на p. Иначе просто поменяем x с y. Если y не делится на p, то, по свойству деления, найдется такое целое число t, что yt сравнимо с 1 по модулю p. Тогда $(xt)^2 + (yt)^2$ сравнимо с $(xt)^2 + 1$ по модулю p. В частности, $(xt)^2 + 1$ делится на p. Теперь достаточно положить $z = xt$. $\square$

Теперь мы можем свести описание множества всех значений квадратичной формы $(1, 0, 1)$ к описанию всех ее простых значений:

Лемма. *Целое положительное число a представляется как сумма квадратов двух целых чисел тогда и только тогда, когда всякое простое число, входящее в разложение числа a в нечетной степени, представляется в этом виде.*

Доказательство. Допустим сначала, что число a представляется как $x^2 + y^2$ для некоторых целых чисел x и y, а простое число p входит в разложение числа a в нечетной степени. Пусть p^k — максимальная степень p, на которую делятся оба числа x и y. Тогда обе части равенства $x^2 + y^2 = a$ делятся на p^{2k}. Положим $x' = x/p^k$, $y' = y/p^k$, $a' = a/p^{2k}$. Тогда ${x'}^2 + {y'}^2 = a'$, причем a' делится на p, а по крайней мере одно из чисел x' или y' не делится на p. Согласно ранее доказанной лемме, найдется целое число z, для которого $z^2 + 1$ делится на p. По другой лемме, отсюда следует, что p представляется как сумма двух квадратов.

Теперь, наоборот, предположим, что всякое простое число, входящее в разложение числа a в нечетной степени, представляется как сумма двух квадратов. Тогда и число a можно представить как сумму двух квадратов. Это вытекает из следующих двух фактов:

- произведение двух значений квадратичной формы $(1, 0, 1)$ снова является значением.
- всякий полный квадрат является значением формы $(1, 0, 1)$, так как $x^2 = x^2 + 0^2$.

□

Нам теперь остается только решить такую задачу: какие простые числа являются значениями формы $(1, 0, 1)$? В силу доказанных лемм, такая постановка вопроса эквивалентна следующей: какие простые числа делят числа вида $x^2 + 1$?

Закон взаимности. *Если простое число p делит число вида x^2+1, где x — целое число, то p сравнимо с 1 по модулю 4. Наоборот, всякое простое число p, сравнимое с 1 по модулю 4, делит некоторое число вида $x^2 + 1$.*

В одну сторону, это утверждение уже доказано. Именно, если простое число p делит число вида x^2+1, где x — целое число, или, что то же самое, если p представляется в виде суммы квадратов, то p сравнимо с 1 по модулю 4. Это частный случай теоремы об остатках от деления сумм квадратов на 4. Осталось доказать закон взаимности в обратную сторону.

Нам понадобится следующая знаменитая теорема:

Малая теорема Ферма. *Пусть p — простое число, а целое число a не делится на p. Тогда a^{p-1} сравнимо с 1 по модулю p.*

Доказательство. Приводимое ниже доказательство принадлежит Эйлеру. Утверждение малой теоремы Ферма эквивалентно следующему: все значения многочлена $f(x) = x^p - x$ при целых x делятся на p. На самом деле, достаточно доказать, что для всякого целого числа x, число $f(x+1) - f(x)$ делится на p. Тогда число

$$f(x) = [f(x) - f(x-1)] + [f(x-1) - f(x-2)] + \cdots + [f(1) - f(0)]$$

тоже будет делиться на p. В последнем равенстве мы воспользовались тем, что $f(0) = 0$.

Согласно формуле бинома Ньютона:

$$f(x+1) - f(x) = (x+1)^p - x^p - 1 = \frac{p!}{1!(p-1)!}x + \frac{p!}{2!(p-2)!}x^2 + \cdots + \frac{p!}{(p-1)!1!}x^p.$$

В правой части, коэффициент при x^k равен

$$\frac{p!}{k!(p-k)!}.$$

При $0 < k < p$, это число делится на p. В самом деле, числитель делится на p, а знаменатель нет. Итак, все коэффициенты многочлена $f(x+1) - f(x)$

делятся на p. Отсюда следует, что все значения этого многочлена при целых x тоже делятся на p. $\square$

Зафиксируем простое число p. Целое число a называется *квадратичным вычетом* по модулю p, если a сравнимо с x^2 по модулю p для некоторого целого числа x. В противном случае число a называется *квадратичным невычетом* по модулю p. Заметим, что закон взаимности, который мы пытаемся доказать, можно переформулировать следующим образом. Число -1 является квадратичным вычетом по простому модулю p тогда и только тогда, когда p дает 1 в остатке при делении на 4.

Лемма. *Пусть $p > 2$. Среди чисел*

$$1,\ 2,\ \ldots,\ p-1$$

квадратичных вычетов по модулю p ровно столько же, сколько квадратичных невычетов.

Доказательство. Рассмотрим отображение, сопоставляющее каждому числу x от 1 до $p-1$ остаток от деления числа x^2 на p. В качестве множества значений этого отображения, получаем в точности множество всех квадратичных вычетов по модулю p. Теперь достаточно доказать, что при рассматриваемом отображении, каждый квадратичный вычет имеет в точности два прообраза. Это означает, что квадратичных вычетов в точности в два раза меньше, чем всего чисел от 1 до $p-1$. Но это именно то, что мы хотим доказать.

Итак, рассмотрим квадратичный вычет a, получающийся как остаток от деления x^2 на p, где x — целое число от 1 до $p-1$. Докажем, что x^2 сравнимо с y^2 только в том случае, когда y сравнимо с x или с $p-x$. Это значит, что у числа a есть только два прообраза: x и $p-x$. Легко видеть, что эти числа всегда различны, так как p нечетно.

Если x^2 сравнимо с y^2 по модулю p, то $x^2 - y^2 = (x-y)(x+y)$ делится на p. Следовательно, $x-y$ делится на p или $x+y$ делится на p. В первом случае, x сравнимо с y по модулю p. Во втором случае, x сравнимо с $p-y$ по модулю p. $\square$

Теперь нам нужен критерий, позволяющий отличать квадратичные вычеты по модулю p от квадратичных невычетов.

Пусть f — многочлен с целыми коэффициентами. Рассмотрим сравнение

$$f(x) \equiv 0 \pmod{p}$$

с одной неизвестной x по простому модулю p. Допустим, задача состоит в том, чтобы найти все значения неизвестной x, при которых сравнение выполнено. Заметим, что эту задачу достаточно решить для x в пределах от 0 до $p-1$. Действительно, нетрудно видеть, что $f(x)$ делится на p тогда и только тогда, когда $f(x+p\,k)$ делится на p для некоторого (а тогда и для любого) целого k.

Задача. Докажите это утверждение.

Таким образом, если $a_1, \dots, a_m$ — все решения рассматриваемого сравнения в пределах от 0 до $p-1$, то любое решение имеет вид $a_l + p\,k$, где l — число от 1 до m, а k — произвольное целое число.

Теорема. *Пусть f — многочлен степени m с целыми коэффициентами, не все из которых делятся на p. Тогда сравнение*

$$f(x) \equiv 0 \pmod{p}$$

имеет не более m решений среди целых чисел от 0 до $p-1$.

Доказательство. Допустим, что a_1 является решением данного сравнения, то есть $f(a_1)$ делится на p, и что a_1 лежит в пределах от 0 до $p-1$. Многочлен f можно разделить на многочлен $x-a_1$ с остатком, то есть существует такой многочлен f_1 степени $m-1$ с целыми коэффициентами и такое целое число r_1, что

$$f(x) = f_1(x)(x-a_1) + r_1.$$

Для того, чтобы найти f_1 и r_1, достаточно сделать следующее:

- Подставить в многочлен f вместо переменной x выражение $y + a_1$, и раскрыть скобки.
- Свободный член полученного многочлена от y будет равен r_1.
- Если вычесть свободный член, то останется многочлен, делящийся на y. Разделим его на y. Получим многочлен на единицу меньшей степени, и в этот многочлен подставим $y = x - a_1$. Результат этой операции — многочлен f_1.

Очевидно, что если число $f(a_1)$ делится на p, то тогда остаток r_1 тоже делится на p. Далее, если a_2 — такое целое число от 0 до $p-1$, что $f_1(a_2)$ делится на p, то мы можем проделать ту же самую процедуру деления с остатком, и получить равенство

$$f_1(x) = f_2(x)(x-a_2) + r_2,$$

в котором многочлен f_2 с целыми коэффициентами имеет степень $m-2$, а остаток r_2 делится на p. Подставляя выражение для f_1 через f_2 в выражение для f через f_1, получаем:

$$f(x) = f_2(x)(x-a_1)(x-a_2) + r_2(x-a_1) + r_1.$$

Предположим, что $a_1, \dots, a_m$ — различные решения сравнения

$$f(x) \equiv 0 \pmod{p},$$

находящиеся в пределах от 0 до $p-1$. Если решений больше чем m, то мы возьмем только m из них (априори не все решения). Однако далее мы увидим, что никаких других решений существовать не может. Если же решений меньше чем m, то доказывать нечего.

Применяя m раз описанную выше процедуру, мы придем к многочлену f_m с целыми коэффициентами, степень которого будет равна 0. Это означает, что многочлен f_m — это просто целое число, не зависящее от x. Кроме того, мы получим последовательность остатков $r_1, \ldots, r_m$, все из которых делятся на p. Исходный многочлен f можно выразить через числа $a_1, \ldots, a_m$, число f_m и остатки $r_1, \ldots, r_m$ следующим образом:

$$f(x) = f_m(x - a_1) \cdots (x - a_m) + r_m(x - a_1) \cdots (x - a_{m-1}) + \cdots + r_1.$$

Поскольку все остатки $r_1, \ldots, r_m$ делятся на p, получаем такое сравнение:

$$f(x) \equiv f_m(x - a_1) \cdots (x - a_m) \pmod{p}.$$

Если число f_m делится на p, то все коэффициенты многочлена f делятся на p. Но это противоречит нашему предположению. Если f_m не делится на p, то любое решение сравнения

$$f(x) \equiv 0 \pmod{p}$$

сравнимо по модулю p с одним из чисел $a_1, \ldots, a_m$. Значит, $a_1, \ldots, a_m$ — это все решения данного сравнения, лежащие в пределах от 0 до $p - 1$. □

Применим доказанную теорему к описанию квадратичных вычетов по модулю p. Всякий квадратичный вычет a сравним с x^2 по модулю p, где x — некоторое целое число. Следовательно, если a не делится на p, то по малой теореме Ферма,

$$a^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod{p}.$$

Теперь рассмотрим уравнение

$$x^{\frac{p-1}{2}} \equiv 1 \pmod{p}. \tag{7}$$

Мы только что показали, что все квадратичные вычеты по модулю p, кроме 0, удовлетворяют этому уравнению. С другой стороны, чисел от 0 до $p - 1$, удовлетворяющих этому уравнению, не более чем $(p - 1)/2$. Но это в точности число ненулевых квадратичных вычетов. Таким образом, ненулевые квадратичные вычеты — это в точности числа от 1 до $p - 1$, удовлетворяющие уравнению (7).

Вернемся к закону взаимности. Мы хотели доказать, что $x^2 + 1$ делится на простое число p для некоторого целого x, если p сравнимо с 1 по модулю 4. Другими словами, -1 является квадратичным вычетом по модулю p, если $(p - 1)/2$ — четное целое число. Но для этого достаточно просто проверить условие (7)! Поскольку $(p - 1)/2$ четно, имеем

$$(-1)^{\frac{p-1}{2}} = 1.$$

Следовательно, -1 является квадратичным вычетом, что и требовалось доказать.

Теперь теорему Ферма о суммах квадратов можно считать доказанной. Напомним еще раз формулировку:

Целое число a представимо как значение квадратичной формы $(1, 0, 1)$ тогда и только тогда, когда всякое простое число, входящее в разложение числа a в нечетной степени, сравнимо с 1 по модулю 4.

Дополнительные упражнения.

Следующие упражнения имеют своей целью описание всех значений целочисленной квадратичной формы $(1, 0, 2)$. Это описание принадлежит Ферма. Оно следует тому же сценарию, по которому мы действовали при изучении значений формы $(1, 0, 1)$.

1. Выпишите первые несколько значений формы $(1, 0, 2)$. Какие закономерности можно увидеть?

2. Докажите, что произведение двух значений формы $(1, 0, 2)$ снова является значением. Более того, верна следующая формула

$$(x_1^2 + 2x_2^2)(y_1^2 + 2y_2^2) = (x_1y_1 - 2x_2y_2)^2 + 2(x_1y_2 + x_2y_1)^2.$$

Найдите интерпретацию этой формулы в терминах комплексных чисел.

3. Пусть a — значение квадратичной формы $(1, 0, 2)$. Остаток от деления числа a на 8 не может быть равен 5 или 7. Все остальные остатки встречаются.

4. Если простое число, отличное от 2, является значением квадратичной формы $(1, 0, 2)$, то оно дает остаток 1 или 3 при делении на 8.

5. Допустим, что простое число p входит в разложение числа $x^2 + 2y^2$ в нечетной степени. Тогда найдется такое целое число z, что $z^2 + 2$ делится на p.

6. Докажите теорему об эквивалентности для формы $(1, 0, 2)$: всякая целочисленная квадратичная форма дискриминанта -8 эквивалентна форме $(1, 0, 2)$.

7. Допустим, что $z^2 + 2$ делится на простое число p для некоторого целого z. Тогда число p является значением формы $(1, 0, 2)$.

8. Зафиксируем простое число p. Докажите, что произведение двух квадратичных невычетов по модулю p является квадратичным вычетом. *Указание:* воспользуйтесь тем, что квадратичных вычетов столько же, сколько квадратичных невычетов, и тем, что произведение квадратичного невычета и квадратичного вычета является квадратичным невычетом.

9. Пусть p — простое число, дающее 3 в остатке при делении на 8. Докажите, что -2 является квадратичным вычетом по модулю p. *Указание:* воспользуйтесь тем, что -1 является квадратичным невычетом по модулю p, и 2 тоже является квадратичным невычетом.

10. Пусть p — простое число, имеющее остаток 1 при делении на 8. Докажите, что сравнение

$$x^4 + 1 \equiv 0 \pmod p$$

разрешимо. *Указание:* воспользуйтесь тем, что многочлен $x^{p-1}-1$ делится на x^4+1.

11. Пусть p — простое число, имеющее остаток 1 при делении на 8. Докажите, что -2 является квадратичным вычетом по модулю p. *Указание:* воспользуйтесь формулой
$$x^4+1=(x^2-1)^2+2x^2.$$

12. На основании всех предыдущих задач, докажите следующую теорему:

Натуральное число a является значением квадратичной формы $(1,0,2)$ *тогда и только тогда, когда всякое простое число, входящее в разложение числа a в нечетной степени и отличное от 2, дает остаток 1 или 3 при делении на 8.*

4 Произведения сумм квадратов

Вернемся к обсуждению формулы (1):
$$(x_1^2+x_2^2)(y_1^2+y_2^2)=(x_1y_1-x_2y_2)^2+(x_1y_2+x_2y_1)^2.$$
Оказывается, что аналогичной формулы для трех квадратов не существует. Другими словами, произведение
$$(x_1^2+x_2^2+x_3^2)(y_1^2+y_2^2+y_3^2) \tag{8}$$
не представляется как сумма трех квадратов
$$z_1^2+z_2^2+z_3^2,$$
в которой z_1, z_2 и z_3 являются билинейными формами от x_1,x_2,x_3 и y_1,y_2,y_3. Согласно терминологии, введенной в первом разделе, формулы типа $(3,3,3)$ не существует. Ниже мы докажем это утверждение.

Верно даже и более сильное утверждение: произведение
$$(x_1^2+x_2^2)(y_1^2+y_2^2+y_3^2) \tag{9}$$
суммы двух квадратов на сумму трех квадратов не представляется как сумма трех квадратов билинейных форм от x_1,x_2 и y_1,y_2,y_3. Из второго утверждения вытекает первое, так как если бы произведение (8) представлялось в виде суммы трех квадратов, то полагая $x_3=0$, мы получили бы аналогичное представление для произведения (9).

Для доказательства сформулированных утверждений, нам нужно обсудить ортогональные линейные операторы. Рассмотрим отображение A из трехмерного пространства в себя. Наличие такого отображения означает, что у нас есть правило, по которому каждой точке $\mathbf{x}$ трехмерного пространства сопоставляется единственная точка $\mathbf{y}=A(\mathbf{x})$ того же пространства. В нашем случае мы даже можем предположить, что правило, о котором идет речь, задано явной формулой, выражающей координаты точки $\mathbf{y}$ через координаты точки $\mathbf{x}$.

Точку $\mathbf{x}$ трехмерного пространства мы будем всегда отождествлять с вектором, соединяющим начало координат с точкой $\mathbf{x}$. В частности, определено сложение точек, умножение точек на действительные числа, а также скалярное произведение точек. Если $\mathbf{x}$ — точка с координатами (x_1, x_2, x_3), а $\mathbf{y}$ — точка с координатами (y_1, y_2, y_3), то

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3),$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Кроме того, если α — действительное число, то

$$\alpha \mathbf{x} = (\alpha x_1, \alpha x_2, \alpha x_3).$$

Мы всегда придерживаемся соглашения, что набор координат в правой части равенства обозначает точку с такими координатами.

Отображение A называется *линейным оператором*, если для каждой пары точек $\mathbf{x}$ и $\mathbf{x}'$, выполнено равенство

$$A(\mathbf{x} + \mathbf{x}') = A(\mathbf{x}) + A(\mathbf{x}'),$$

и, кроме того, для каждой точки $\mathbf{x}$ и каждого действительного числа α, выполнено равенство

$$A(\alpha \mathbf{x}) = \alpha A(\mathbf{x}).$$

Задача. Докажите, что для каждого линейного оператора A, координаты точек $\mathbf{x}$ и $\mathbf{y} = A(\mathbf{x})$ связаны следующими линейными уравнениями:

$$\begin{aligned} y_1 &= a_{11}\, x_1 + a_{12}\, x_2 + a_{13}\, x_3, \\ y_2 &= a_{21}\, x_1 + a_{22}\, x_2 + a_{23}\, x_3, \\ y_3 &= a_{31}\, x_1 + a_{32}\, x_2 + a_{33}\, x_3, \end{aligned}$$

где a_{ij} — некоторые действительные коэффициенты, не зависящие ни от $\mathbf{x}$, ни от $\mathbf{y}$.

Таким образом, линейный оператор A определяется таблицей коэффициентов

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Эта таблица называется *матрицей* линейного оператора A. Если система координат фиксирована, то линейный оператор однозначно определяется своей матрицей. Однако важно иметь в виду, что матрица существенно зависит от выбора системы координат.

Линейный оператор A называется *ортогональным*, если для каждого вектора $\mathbf{x}$ имеем

$$\langle A(\mathbf{x}), A(\mathbf{x}) \rangle = \langle \mathbf{x}, \mathbf{x} \rangle.$$

Другими словами, линейный оператор ортогональный, если он сохраняет длины всех векторов. Отсюда, в частности, следует, что ортогональный оператор сохраняет расстояние между точками. Действительно, расстояние между двумя точками — это длина вектора, равного разности этих точек. Про линейный ортогональный оператор следует думать как про евклидово движение трехмерного пространства, сохраняющее начало координат.

Задача. Пусть A — ортогональный оператор, а $\mathbf{x}$ и $\mathbf{x}'$ — две точки. Тогда выполнено равенство

$$\langle A(\mathbf{x}), A(\mathbf{x}')\rangle = \langle \mathbf{x}, \mathbf{x}'\rangle.$$

В частности, если векторы $\mathbf{x}$ и $\mathbf{x}'$ перпендикулярны, то векторы $A(\mathbf{x})$ и $A(\mathbf{x}')$ тоже перпендикулярны. *Указание:* в соотношении

$$\langle A(\mathbf{x}+\lambda\mathbf{x}'), A(\mathbf{x}+\lambda\mathbf{x}')\rangle = \langle \mathbf{x}+\lambda\mathbf{x}', \mathbf{x}+\lambda\mathbf{x}'\rangle$$

раскройте скобки и приравняйте коэффициенты при λ в обеих частях полученного равенства. Равенство следует рассматривать как равенство многочленов от λ.

Рассмотрим векторы

$$\mathbf{e}_1 = (1,0,0), \quad \mathbf{e}_2 = (0,1,0), \quad \mathbf{e}_3 = (0,0,1).$$

Вектор $\mathbf{x}$ с координатами (x_1, x_2, x_3) выражается через эти векторы следующим образом:

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

Задача. Докажите, что оператор A ортогонален тогда и только тогда, когда выполнены следующие соотношения:

$$\langle A(\mathbf{e}_i), A(\mathbf{e}_i)\rangle = 1, \quad \langle A(\mathbf{e}_i), A(\mathbf{e}_j)\rangle = 0, \quad i,j = 1,2,3, \quad i \neq j.$$

Всякий ортогональный оператор A *взаимно однозначен*. Это означает, что каждая точка $\mathbf{y}$ трехмерного пространства является образом одной и только одной точки $\mathbf{x}$. Это утверждение интуитивно очевидно. Единственность точки $\mathbf{x}$ также следует из того, что оператор A сохраняет расстояния. Если бы две различные точки имели один и тот же образ, то тогда получалось бы, что две точки, находящиеся на ненулевом расстоянии, переходят в две точки на нулевом расстоянии. Противоречие. То, что каждая точка может быть получена как образ некоторой точки при отображении A, труднее доказать строго. Идея доказательства может состоять, например, в следующем. Подействуем оператором A на оси координат. Тем самым мы переведем один координатный репер (т.е. систему осей координат) в другой. Получим две системы координат — старую и новую. Теперь достаточно заметить, что всякая точка может быть однозначно задана своими координатами как относительно старой системы, так и относительно новой.

Поскольку оператор A взаимно однозначен, существует *обратный оператор* A^{-1}, который каждую точку $\mathbf{y}$ переводит в единственную точку $\mathbf{x}$, такую, что $\mathbf{y} = A(\mathbf{x})$.

Задача. Докажите, что A^{-1} является ортогональным линейным оператором.

Пусть A и B — два отображения из трехмерного пространства в себя (например, два линейных оператора). Тогда определена *композиция* AB. Это отображение, переводящее каждую точку x в точку $A(B(x))$. Заметим, что сначала применяется оператор B, а потом оператор A. Таким образом, запись AB следует читать справа налево. В частности, если A и B — линейные операторы, то их композиция AB — тоже линейный оператор. Если A и B — ортогональные операторы, то AB — тоже ортогональный оператор.

Задача. Докажите эти утверждения.

Тождественный оператор E определяется как оператор, который ничего не делает — он отправляет каждую точку $\mathbf{x}$ в нее саму, $A(\mathbf{x}) = \mathbf{x}$. По определению обратного оператора, выполнены соотношения

$$A^{-1}A = AA^{-1} = E.$$

Допустим, что существует формула типа $(2, 3, 3)$:

$$(x_1^2 + x_2^2)(y_1^2 + y_2^2 + y_3^2) = z_1^2 + z_2^2 + z_3^2,$$

где z_1, z_2 и z_3 — билинейные формы от x_1, x_2 и y_1, y_2, y_3. Зафиксируем двумерный вектор с координатами (x_1, x_2) и рассмотрим отображение A, переводящее точку с координатами (y_1, y_2, y_3) в точку с координатами (z_1, z_2, z_3). Нетрудно проверить, что это линейный оператор. Однако этот линейный оператор зависит от (x_1, x_2). Обозначим через A_1 оператор, соответствующий вектору $(1, 0)$, и через A_2 оператор, соответствующий вектору $(0, 1)$.

Задача. Докажите, что оба линейных оператора A_1 и A_2 являются ортогональными. Более того, если $x_1^2 + x_1^2 = 1$, то линейный оператор A, соответствующий вектору (x_1, x_2), является ортогональным.

Задача. Докажите, что для каждого трехмерного вектора $\mathbf{y}$, образы $A_1(\mathbf{y})$ и $A_2(\mathbf{y})$ перпендикулярны, то есть

$$\langle A_1(\mathbf{y}), A_2(\mathbf{y})\rangle = 0.$$

Указание: Верно более общее утверждение. Пусть операторы A и B соответствуют векторам $\mathbf{x} = (x_1, x_2)$ и $\mathbf{x}' = (x_1', x_2')$, соответственно. Тогда

$$\langle A(\mathbf{y}), B(\mathbf{y}))\rangle = \langle \mathbf{x}, \mathbf{x}'\rangle\langle \mathbf{y}, \mathbf{y}\rangle.$$

Рассмотрим теперь оператор $I = A_1^{-1}A_2$. Этот оператор ортогонален как композиция двух ортогональных операторов. Кроме того, оператор I *кососимметрический*, то есть для всякого трехмерного вектора $\mathbf{x}$ выполнено равенство

$$\langle \mathbf{x}, I(\mathbf{x})\rangle = 0.$$

(Заметьте, что мы сменили обозначения: теперь $\mathbf{x}$ обозначает трехмерный вектор, а не двумерный, как в предыдущей задаче). Действительно, имеем:

$$\langle \mathbf{x}, I(\mathbf{x})\rangle = \langle A_1(\mathbf{x}), A_1 I(\mathbf{x})\rangle = \langle A_1(\mathbf{x}), A_2(\mathbf{x})\rangle = 0.$$

Таким образом, оператор I как ортогональный, так и кососимметрический.

Задача. Докажите следующее утверждение. Если оператор I является кососимметрическим, то для каждой пары векторов $\mathbf{x}$ и $\mathbf{x}'$ выполнено соотношение

$$\langle \mathbf{x}, I(\mathbf{x}')\rangle = -\langle I(\mathbf{x}), \mathbf{x}'\rangle.$$

Мы докажем, что квадрат оператора I, то есть композиция этого оператора с самим собой, совпадает с $-E$, то есть с центральной симметрией относительно начала координат. Центральная симметрия $-E$ переводит всякий вектор $\mathbf{x}$ в антиподальный вектор $-\mathbf{x}$. Для того, чтобы доказать, что $I^2 = -E$, заметим, что для всякой пары векторов $\mathbf{x}$ и $\mathbf{x}'$ выполнено соотношение

$$\langle \mathbf{x}, I^2(\mathbf{x}')\rangle = -\langle I(\mathbf{x}), I(\mathbf{x}')\rangle = -\langle \mathbf{x}, \mathbf{x}'\rangle.$$

В первом равенстве была использована кососимметричность оператора I, а во втором равенстве — его ортогональность. Мы доказали, таким образом, что

$$\langle \mathbf{x}, I^2(\mathbf{x}')\rangle = \langle \mathbf{x}, -E(\mathbf{x}')\rangle.$$

Теперь результат $I^2 = -E$ вытекает из следующей задачи:

Задача. Предположим, что векторы $\mathbf{y}$ и $\mathbf{z}$ таковы, что для всякого вектора $\mathbf{x}$ выполнено равенство

$$\langle \mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{z}\rangle,$$

то $\mathbf{y} = \mathbf{z}$.

Мы собираемся построить противоречие, показывающее, что на самом деле, ортогонального оператора I, удовлетворяющего соотношению $I^2 = -E$, в трехмерном пространстве не существует. Таким образом, не существует и формулы типа $(2, 3, 3)$.

Для того, чтобы прийти к противоречию, нужно посмотреть, что происходит с *ориентацией* пространства при действии оператора I.

Неформально, *ориентация* трехмерного пространства — это выбор между правым винтом и левым винтом. Правый винт завинчивается вправо, а левый — влево. Даже если два винта очень похожи, а мы будем это предполагать, один винт нельзя перевести в другой собственным движением. Это означает, что один винт нельзя повернуть так, чтобы он стал идентичен другому винту. Однако зеркальное отражение левого винта является правым винтом и наоборот. Зеркальное отражение, хотя и не является собственным евклидовым движением, сохраняет все расстояния. Более того, зеркальное отражение осуществляется при помощи ортогонального линейного оператора. Ортогональные линейные операторы, меняющие ориентацию пространства, то есть переводящие правый винт в левый, и наоборот, называются *несобственными ортогональными операторами.*

Одно из замечательных свойств ориентации состоит в том, что преобразование (например, ортогональный оператор) либо сохраняет ориентацию, либо ее меняет на противоположную. Третьего не дано. Существует только два типа винтов — правые и левые. Из этого очевидного замечания, в частности, вытекает, что квадрат любого ортогонального оператора сохраняет ориентацию. Для того, чтобы в этом убедиться, достаточно рассмотреть два случая — когда оператор сохраняет ориентацию, и когда оператор меняет ориентацию. В обоих случаях, после двукратного применения оператора с ориентацией ничего не происходит.

Вернемся к нашему ортогональному оператору I, удовлетворяющему уравнению $I^2 = -E$. Оператор I^2 должен сохранять ориентацию. Однако, можно показать, что оператор $-E$ меняет ориентацию. Противоречие.

Задача. Убедитесь, что оператор $-E$ действительно меняет ориентацию.

Полученное противоречие показывает, что оператора I, удовлетворяющего описанным свойствам, не существует. Это означает, что не существует формулы типа $(2,3,3)$. Как мы видели, отсюда следует, что формулы типа $(3,3,3)$ тоже не существует.

Подчеркнем, что в отличие от размерности 3, в размерности 2 оператор центральной симметрии относительно начала координат сохраняет ориентацию. На плоскости, ориентация — это выбор направления кругового обхода: по часовой стрелке или против часовой стрелки. На плоскости, существует линейный ортогональный оператор I, удовлетворяющий соотношению $I^2 = -E$. Этот оператор удобнее всего задать, отождествив векторы на плоскости с комплексными числами. Тогда оператор I — это просто умножение на мнимую единицу. Другими словами, этот оператор переводит комплексное число z в комплексное число iz. Введем на плоскости комплексных чисел $z = x + yi$ координаты (x, y). Координатами комплексного числа будут его вещественная и мнимая части. В этой системе координат, оператор I выражается формулой

$$I(x, y) = (-y, x).$$

Иначе говоря, матрица оператора I такая:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Как мы видели ранее, кватернионы дают пример формулы типа $(4,4,4)$. Обозначим через I, J и K операторы умножения на i, j и k, соответственно. Тогда каждый из этих операторов в квадрате дает оператор $-E$. Более того, эти операторы антикоммутируют, то есть произведение любой пары различных операторов в одном порядке равно произведению в другом порядке со знаком минус.

Теперь обсудим очень кратко многомерную ситуацию. В общем случае, вопрос о том, при каких целых r, s и n существует формула типа (r, s, n), совершенно открыт (задача Гурвица). Однако в частном случае $s = n$ ответ известен, и получен самим Гурвицем (Радон, независимо от Гурвица и примерно в то же самое время, опубликовал решение этой задачи).

Как и в случае $n = 3$, если в формуле

$$(x_1^2 + \cdots + x_r^2)(y_1^2 + \cdots + y_n^2) = z_1^2 + \cdots + z_n^2$$

зафиксировать конкретные значения переменных $x_1, \ldots, x_r$, то получится линейный оператор A, действующий в n-мерном пространстве и зависящий от r-мерного вектора $(x_1, \ldots, x_r)$. Линейные операторы в n-мерном пространстве определяются точно так же, как и в трехмерном. Тоже имеет смысл говорить об ортогональных операторах, кососимметрических операторах и т.д. Мы здесь не будем вдаваться в подробности, однако читателю рекомендуется обратиться за более подробными обсуждениями к учебникам линейной алгебры.

Рассмотрим набор координат $(x_1, \ldots, x_r)$, в котором все координаты кроме одной, с номером k, равны нулю, а координата $x_k = 1$. Соответствующий линейный оператор A обозначим через A_k. Таким образом, у нас есть r линейных операторов $A_1, \ldots, A_r$. Нетрудно показать (это делается так же, как и в трехмерном случае), что эти операторы являются ортогональными. Кроме того, для каждого вектора $\mathbf{y}$ пространства размерности n, векторы

$$A_1(\mathbf{y}),\ A_2(\mathbf{y}),\ \ldots,\ A_r(\mathbf{y})$$

попарно *ортогональны*, то есть скалярное произведение любых двух различных векторов из этого списка равно нулю.

Аналогично тому, что мы делали в трехмерном случае, положим

$$I_1 = E,\ I_2 = A_1^{-1}A_2,\ \ldots,\ I_r = A_1^{-1}A_r.$$

Здесь E — это тождественный оператор, оставляющий все n-мерные векторы на месте. Это ортогональные операторы, которые к тому же кососимметричны. Отсюда вытекает, что эти операторы удовлетворяют соотношениям $I_k^2 = -E$. Кроме того, при $k \neq m$, мы имеем соотношение антикоммутативности $I_k I_m = -I_m I_k$.

Задача. Докажите соотношение антикоммутативности.

Набор операторов $I_1, \ldots, I_r$, удовлетворяющий указанным соотношениям, называется *представлением алгебры Клиффорда* с r образующими (операторы $I_1, \ldots, I_r$ называются образующими алгебры Клиффорда). Сама *алгебра Клиффорда* состоит из всех операторов I_k, их произведений друг с другом в любых количествах, и сумм таких произведений с действительными коэффициентами. При этом алгебра Клиффорда рассматривается вне всякой связи с тем, как операторы, ее образующие, действуют на n-мерном пространстве. Для описания алгебры Клиффорда нужно только знать, как складывать и перемножать операторы (перемножать операторы — это то же самое, что брать композиции операторов). А *представление* алгебры Клиффорда — это то, как она действует на n-мерном пространстве.

Резюмируем вышесказанное в виде следующей теоремы:

Теорема. *Предположим, что существует формула типа* (r, n, n). *Тогда можно определить набор* $I_1 = E, I_2, \ldots, I_r$ *из* r *ортогональных линейных операторов, действующих на* n*-мерном пространстве, такой, что* $I_k^2 = -E$ *для всех* k *и* $I_k I_m = -I_m I_k$ *для всех* $k \neq m$. *Другими словами, в* n*-мерном пространстве определено представление алгебры Клиффорда с* r *образующими.*

Дополнительные упражнения.

Следующие задачи связаны с вопросом Гурвица и с представлениями алгебр Клиффорда.

1. Докажите, что для каждого четного натурального числа n, существует формула типа $(2, n, n)$. Более того, в n-мерном пространстве определено представление алгебры Клиффорда с одной образующей. *Указание:* воспользуйтесь формулой для умножения комплексных чисел.

2. Докажите, что для каждого натурального числа n, делящегося на 4, существует формула типа $(4, n, n)$. Более того, в n-мерном пространстве определено представление алгебры Клиффорда с тремя образующими. *Указание:* воспользуйтесь формулой для умножения кватернионов.

3. (Формула типа $(8, 8, 8)$.) Точка восьмимерного пространства имеет 8 координат. Расщепим координаты на 2 группы по 4 координаты в каждой. Набор из четырех координат можно отождествить с кватернионом. Таким образом, точка восьмимерного пространства представляется парой кватернионов. Рассмотрим две такие пары (x, x') и (y, y'). Определим пару (z, z') следующим образом:

$$z = xy - y'\overline{x'}, \quad z' = yx' + \overline{x}y'.$$

Докажите следующую формулу:

$$(|x|^2 + |x'|^2)(|y|^2 + |y'|^2) = |z|^2 + |z'|^2.$$

Эта формула типа $(8, 8, 8)$. Указанное правило, по которому (z, z') получается из (x, x') и (y, y'), можно также проинтерпретировать как закон умножения точек в восьмимерном пространстве. Восьмимерное пространство, снабженное этим законом умножения и очевидным (по-координатным) законом сложения, называется *алгеброй октав*. Октавы придумал Дж. Грэйвз и, независимо от него, А. Кэли.

Читателю, знакомому с матричными обозначениями и с формулой для унитарных матриц два на два, возможно, будет легче запомнить формулу для умножения октав, если её записать в таком виде:

$$\begin{pmatrix} z \\ z' \end{pmatrix} = \begin{pmatrix} L_x & -R_{\overline{x'}} \\ R_{x'} & L_{\overline{x}} \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix}.$$

Здесь L_q обозначает оператор левого умножения (умножения слева) на кватернион q, а R_q — оператор правого умножения. Замечательное свойство матрицы два на два, фигурирующей в формуле, состоит в том, что все её элементы являются коммутирующими (перестановочными) друг с другом линейными операторами.

5 Прямые и окружности

В заключение, мне хотелось бы рассказать об одной задаче квадратичной математики, которой я занимался. Я перечислю только результаты, без доказательств. Этот раздел использует некоторые понятия, не входящие в школьную программу.

Задача такая: описать (достаточно хорошие, например, гладкие) преобразования n-мерного пространства, переводящие прямые в окружности. Эта формулировка нуждается в уточнении. Во-первых, преобразования не обязательно должны быть определены всюду — они могут быть определены только на открытом множестве n-мерного пространства. Предполагается, что преобразования — локальные диффеоморфизмы, и что они переводят отрезки прямых в дуги окружностей. Во-вторых, мы считаем, что прямые тоже являются окружностями (бесконечного радиуса).

Эта задача возникла из *номографии* — науки об изображении функций многих переменных плоскими схемами — номограммами. Задачу поставил Г.С. Хованский для $n = 2$.

Две похожие задачи были решены очень давно Мёбиусом, причем в любой размерности, а не только в размерности два. Первая: описать все преобразования, переводящие все прямые в прямые. Ответ: это *проективные преобразования*.

Проективные преобразования плоскости, вообще говоря, не всюду определены. Как правило, есть прямая, образ которой не определен. В определенном смысле, эта прямая уходит на бесконечность. Однако во всех остальных точках проективное преобразование определено. Исключения составляют *аффинные преобразования*, которые определены всюду.

Геометрически, проективные преобразования плоскости можно описать следующим образом. Вложим плоскость в трехмерное пространство. Затем возьмем другую плоскость в том же трехмерном пространстве, а затем спроецируем первую плоскость на вторую из некоторой точки. Получим отображение из плоскости в плоскость. Это частный случай проективного преобразования. Повторим описанную процедуру несколько раз. Все преобразования, которые можно получить таким образом, называются проективными преобразованиями. Аналогичное описание проективных преобразований можно дать в любой размерности.

Вторая задача такая: описать все преобразования, переводящие окружности в окружности. Ответ: это *преобразования Мёбиуса*. Ответ справедлив в любой размерности.

Преобразования Мёбиуса, как правило, тоже не всюду определены. В большинстве случаев есть точка, которая уходит на бесконечность.

Геометрическое описание преобразования Мёбиуса плоскости таково. Рассмотрим *инверсию* относительно окружности с центром в точке O и радиусом R. Инверсия определяется следующим образом. Каждая точка X плоскости, кроме точки O, переходит в единственную точку Y со следующими свойствами:

- точки O, X и Y лежат на одной прямой, более того, точки X и Y лежат по одну сторону от точки O;
- произведение длин отрезков OX и OY равно R^2.

Образ самой точки O не определен. В некотором смысле, точка O уходит на бесконечность.

Инверсия является частным случаем преобразования Мёбиуса. Общее преобразование Мёбиуса можно получить, проделав инверсии многократно (относительно разных окружностей — с разными центрами и разными радиусами). Преобразования Мёбиуса пространств большей размерности определяются аналогично. Нужно только заменить инверсии относительно окружностей инверсиями относительно сфер.

Оказывается, задача об описании отображений, переводящих все прямые в окружности, гораздо сложнее обеих задач, описанных выше. В размерности 2, А.Г. Хованский нашел полное решение этой задачи. Его ответ таков (мы сформулируем ответ не так, как он сформулирован у автора). Вложим плоскость в трехмерное пространство. Затем спроецируем плоскость на некоторую сферу из некоторой точки. Точка может быть бесконечно удаленной. Это означает, что мы включаем в рассмотрение *параллельные проекции* — проекции при помощи пучка параллельных прямых. Наконец, сферу отобразим на плоскость стереографически. Полученное отображение переводит все прямые в окружности. Действительно, при проекции на сферу, образ всякой прямой принадлежит некоторой плоскости (а именно, плоскости, натянутой на данную прямую и центр проекции). Одновременно с этим, образ лежит на сфере. Следовательно, образ всякой прямой принадлежит некоторой окружности на сфере. Стереографическая проекция переводит окружность на сфере в окружность на плоскости. Итак, описанные преобразования переводят прямые в окружности. В размерности 2, других преобразований, переводящих прямые в окружности, нет.

В размерности 3, описание преобразований, переводящих все прямые в окружности, совершенно аналогично. Это доказал Ф. Изади, непосредственно обобщив рассуждения Хованского. Однако в размерности 4 ответ совершенно другой, и он связан с кватернионными расслоениями Хопфа!

Напомним, что кватернионное расслоение Хопфа — это некоторое квадратичное отображение из семимерной сферы в четырехмерную. Это отображение переводит все большие окружности в окружности. На самом деле, есть два кватернионных расслоения Хопфа — правое и левое. Это обстоятельство связано с некоммутативностью умножения кватернионов. Однако, оба отображения обладают описанными свойствами.

Исходя из кватернионного расслоения Хопфа, можно определить очень много преобразований четырехмерного пространства, переводящих прямые в окружности. Именно, рассмотрим центральную проекцию четырехмерного пространства (вложенного в пятимерное пространство) на четырехмерную сферу. В качестве центра проекции возьмем центр сферы. При такой проекции все прямые в четырехмерном пространстве переходят в большие окружности на сфере. На следующем шаге, вложим четырехмерную сферу в семимерную в качестве большой сферы. Другими словами, мы предполагаем, что образ четырехмерной сферы в семимерной является пересечением семимерной сферы с некоторым пятимерным подпространством восьмимерного пространства, проходящим через начало координат. Таких вложений очень много. Наконец, просто ограничим кватернионное расслоение Хопфа (правое или левое) на образ четырехмерной сферы. Так как кватернионное расслоение Хопфа переводит большие окружности в окружности, полученное отображение — композиция всех перечисленных — переводит прямые в окружности.

Кроме отображений, приходящих из кватернионных расслоений Хопфа, есть еще отображения, определяемые так же, как в размерности 2 — это центральные проекции на четырехмерную сферу из некоторого четырехмерного подпространства пятимерного пространства. Однако следует заметить, что таких отображений гораздо меньше (словам "больше" и "меньше" можно придать чёткий смысл, рассматривая размерности различных многообразий отображений). Недавно я доказал, что никаких других преобразований четырехмерного пространства, переводящих все прямые в окружности (кроме перечисленных выше), не существует.

Задача об описании преобразований n-мерного пространства, переводящих прямые в окружности, открыта при всех $n > 4$.

Список литературы

[1] К.Ф. Гаусс: *Арифметические исследования*, Серия "Классики Науки", Издательство АН СССР, Москва 1959

[2] П.Г. Лежен Дирихле: *Лекции по теории чисел*, ОНТИ 1936

[3] В.А. Тиморин: *Отображения, переводящие прямые в окружности, в размерности 4.* http://www.arXiv.org/math.DG/abs/0309053, будет опубликовано в "Функциональный Анализ и его Приложения"

[4] Г.С. Хованский: *Основы номографии*, "Наука", Москва, 1976

[5] А.Г. Хованский: *Выпрямление окружностей*, Сиб. Мат. Ж., **21** (1980), 221–226

[6] V. Arnold: *Arithmetics of binary quadratic forms, symmetry of their continued fractions and geometry of their de Sitter world*, Bull. of Braz. Math. Soc., Vol. 34 No 1, 2003, p.1-41.

[7] A. Cayley: *On Jacobi's elliptic functions, in reply to the Rev. B. Bronwin; and on quaternions* (appendix only), in The Collected Mathematical Papers, Johnson Reprint Co., New York, 1963, p. 127

[8] É. Cartan: *Nombres complexes*, pp. 329-448 in J. Molk (red.): Encyclopédie des sciences mathématiques, Tome I, Vol. 1, Fasc. 4, art. 15 (1908). Reprinted in E. Cartan: *Œuvres complètes*, Partie II. Gauthier-Villars, Paris, 1953, pp. 107-246

[9] W. K. Clifford: *Applications of Grassmann's extensive algebra*, Amer. Jour. Math. **1** (1878), 350-358

[10] W.R. Hamilton: *Lectures on Quaternions.* Royal Irish Academy, 1853

[11] A. Hurwitz: *Über die Komposition der quadratischen Formen*, Math. Ann. 88 (1923) 1-25. Reprinted in Math. Werke II, 641-666

[12] F.A. Izadi: *Rectification of circles, spheres, and classical geometries*, PhD thesis, University of Toronto, (2001)

[13] J. Radon: *Lineare Scharen orthogonaler Matrizen*, Abh. Math. Sem. Univ. Hamburg 1 (1922), 1-14