В. С. Губа, С. М. Львовский

«Парадокс» Банаха–Тарского

МЦНМО

Летняя школа «Современная математика»
Дубна, июль 2011

В. С. Губа, С. М. Львовский

«Парадокс» Банаха—Тарского

Электронное издание

Москва
Издательство МЦНМО
2014

УДК 512.54+510.222
ББК 22.144+22.12
Г93

Губа В. С., Львовский С. М.
«Парадокс» Банаха—Тарского
Электронное издание
М.: МЦНМО, 2014
44 с.
ISBN 978-5-4439-2015-3

В 1924 году выдающиеся польские математики Стефан Банах и Альфред Тарский доказали, что шар в пространстве можно разрезать на конечное число частей, из которых можно сложить шар другого объема. В брошюре мы расскажем, почему эта теорема, производящая впечатление нелепости, не противоречит возможности измерять объемы тел, и познакомим читателя с красивой математикой, стоящей за этим уже классическим результатом.

Для школьников старших классов и студентов младших курсов.

Подготовлено на основе книги: *В. С. Губа, С. М. Львовский*. «Парадокс» Банаха—Тарского. — М.: МЦНМО, 2012.

Издательство Московского центра
непрерывного математического образования
119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241-74-83
http://www.mccme.ru

ISBN 978-5-4439-2015-3

Предисловие

Эта брошюра представляет собой расширенную версию мини-курса, прочитанного вторым автором в июле 2011 года на летней школе «Современная математика» в Дубне. В курсе существенно использовались идеи, разработанные первым автором. В книгу мы добавили кое-что из того, на что не хватило времени на занятиях.

Парадоксом Банаха—Тарского называется следующая теорема, доказанная в 1924 году Стефаном Банахом и Альфредом Тарским, опиравшимися, в свою очередь, на теорему, опубликованную в 1914 году Феликсом Хаусдорфом.

Пусть B_1 и B_2 — два шара разных радиусов. Тогда шар B_1 можно представить в виде объединения конечного числа попарно непересекающихся множеств $B_1 = X_1 \cup X_2 \cup \ldots \cup X_n$, а шар B_2 — в виде объединения конечного числа попарно непересекающихся множеств $B_1' = X_1' \cup X_2' \cup \ldots \ldots \cup X_n'$ таким образом, что множество X_1 переводится некоторым движением пространства в X_1', X_2 переводится некоторым движением пространства в X_2', …, X_n переводится некоторым движением пространства в X_n'. Иными словами, шар B_1 можно разрезать на конечное число частей, из которых можно сложить B_2.

То же верно и для любых двух многогранников в пространстве и вообще для более-менее любых двух тел (точную формулировку мы приведем ниже).

Разумеется, это утверждение вопиющим образом противоречит интуиции: как же такое возможно, если у шаров разного радиуса объемы разные?! Почему факт наличия объемов у тел теорему Банаха—Тарского не опровергает и каков ее «философский» смысл, мы обсудим в разделе 4 после того как эту теорему докажем, а пока что скажем одно: за шокирующей формулировкой теоремы Банаха—Тарского стоит красивая и важная математика, и именно она является главным предметом нашей книжки.

Для выполнимости теоремы Банаха—Тарского очень существенно, что действие происходит именно в пространстве: на плоскости тот же номер уже не проходит, и невозможно разрезать многоугольник на конечное число частей, из которых складывается многоугольник другой площади. В заключительной части книжки мы постараемся объяснить, почему так выходит и чем в этом смысле пространство «хуже»[1] плоскости.

[1] Или лучше?

Для чтения основной части брошюры знать сверх школьной программы почти ничего не требуется: надо только быть знакомым с понятием множества и несколькими типичными примерами счетных и несчетных множеств. Кроме того, желательно знать, как устроены движения плоскости.

Странным образом оказывается, что доказательство противоречащей интуиции теоремы Банаха—Тарского проще, чем доказательство полностью соответствующего интуиции результата, согласно которому аналог теоремы Банаха—Тарского для плоскости места не имеет. Поэтому в посвященном этому заключительном разделе 5 требования к предварительным знаниям читателя существенно выше, чем в остальной части брошюры, и даже читателю, этими знаниями обладающему, многое придется принять на веру. По крайней мере мы старались, чтобы в этом трудном разделе сложность нарастала постепенно и чтобы читатель-школьник не «утратил нить» максимально долго.

Авторы благодарны Ф. В. Петрову за ценные обсуждения.

1. Равносоставленность в наивном и точном смысле

Два многоугольника на плоскости называются равносоставленными, если один из них можно разрезать на конечное число частей, из которых без зазоров и наложений можно сложить другой. Например, из квадрата со стороной 1 можно сделать прямоугольный треугольник с углом 45° и катетом $\sqrt{2}$ (рис. 1).

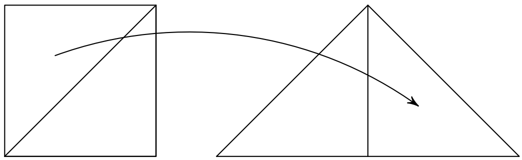


Рис. 1. Квадрат равносоставлен с равнобедренным прямоугольным треугольником

Из самогó определения очевидно, что у равносоставленных прямоугольников площади равны; в XIX веке было доказано и обратное: как гласит так называемая теорема Бойяи—Гервина, если площади двух прямоугольников равны, то они равносоставлены. Доказательство этой теоремы можно найти, например, в книге С. Л. Табачникова и Д. Б. Фукса «Математический дивертисмент» (М.: МЦНМО, 2011), лекция 22; мы не будем его воспроизводить.

Давайте теперь вглядимся в формулировку теоремы Бойяи—Гервина взглядом педанта (точнее говоря, человека, знакомого с понятием множества). Что в ней, собственно говоря, подразумевается под разрезанием и складыванием? Например, катеты прямоугольного треугольника на рис. 1 получаются из двух экземпляров диагонали квадрата, но ведь каждая точка диагонали может попасть либо на один, либо на другой катет, но не на оба одновременно! И вообще, если считать, что многоугольник как подмножество плоскости включает свою границу, то не получается разрезания на непересекающиеся подмножества, а если считать, что не включает, то многоугольник не будет совпадать с объединением тех частей, на которые он разрезан.

Эти шероховатости в формулировке можно устранять разными способами. Например, можно раз и навсегда условиться, что многоугольник как подмножество плоскости включает свою границу, а при определении разрезания многоугольника на меньшие — сказать, что многоугольникам разрешается иметь непустое пересечение, но при этом оно обязано содержаться в границе обоих многоугольников. Это вполне осмысленное уточнение, и при переходе от плоскости к пространству оно также ведет к интересным и неожиданным результатам[1], но мы пойдем по другому пути, потребовав, чтоб части, на которые разбиваются многоугольники, действительно не пересекались.

Перейдем к точным определениям

Мы будем обозначать множество точек плоскости через $\mathbb{R}^2$, а множество точек пространства через $\mathbb{R}^3$ (это стандартные обозначения).

Как известно, отображение $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ (соответственно $f\colon \mathbb{R}^3 \to \mathbb{R}^3$) называется *движением*, если оно сохраняет расстояния между точками: если p и q — точки в $\mathbb{R}^2$ (соответственно $\mathbb{R}^3$), то $|f(p), f(q)| = |p, q|$ (через $|AB|$ или $|A, B|$ мы будем обозначать расстояние между точками A и B). Иногда вместо «движение» говорят «перемещение».

Задача 1. Пусть $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ — отображение, сохраняющее расстояния между точками. Докажите, что f *инъективно*, то есть «не склеивает разные точки» (если $P \neq Q$, то $f(P) \neq f(Q)$), и *сюръективно*, то есть $f(\mathbb{R}^2) = \mathbb{R}^2$ (каждая точка плоскости является чьим-то образом). Аналогичное утверждение верно и для отображений $f\colon \mathbb{R}^3 \to \mathbb{R}^3$.

Взаимно однозначные отображения $f\colon X \to Y$, отображающие X на Y, обычно называют биекциями.

Определение 1. Пусть X и Y — два подмножества плоскости (или пространства). Мы будем говорить, что X и Y *конгруэнтны*, если су-

[1] Ключевые слова здесь — «третья проблема Гильберта» и «инвариант Дена»; см. цитированную выше книгу Табачникова и Фукса.

ществует движение $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ (соответственно $f\colon \mathbb{R}^3 \to \mathbb{R}^3$), для которого $f(X) = Y$. Если X и Y конгруэнтны, мы будем это записывать так: $X \cong Y$.

В современных школьных курсах конгруэнтные множества обычно называются равными; мы будем пользоваться длинным словом «конгруэнтность», чтобы исключить путаницу с равенством множеств. (Впрочем, авторы книжки застали то время, когда термин «конгруэнтность» использовался и в школьном курсе.)

Для дальнейшего нам понадобится одно теоретико-множественное обозначение. Если множество X является объединением множеств $X_1, \ldots, X_n$ и при этом множества $X_1, \ldots, X_n$ попарно не пересекаются, то мы будем записывать это обстоятельство в виде $X = X_1 \sqcup \ldots \sqcup X_n$ (а не просто $X = X_1 \cup \ldots \cup X_n$). Подчеркнем, что $\sqcup$, в отличие от $\cup$ — не символ операции над множествами: если у множеств A и B непустое пересечение, то образовать множество $A \sqcup B$ мы не можем. В ситуации, когда $X = X_1 \sqcup \ldots \sqcup X_n$, мы будем иногда говорить, что задано разбиение множества X на подмножества $X_1, \ldots, X_n$ или что X является дизъюнктным объединением множеств $X_1, \ldots, X_n$.

Определение 2. Пусть X и Y — подмножества плоскости (или пространства). Будем говорить, что X и Y *равносоставлены*, если существуют такие разбиения $X = X_1 \sqcup \ldots \sqcup X_n$, $Y = Y_1 \sqcup \ldots \sqcup Y_n$, что $X_1 \cong Y_1$, $X_2 \cong Y_2$, …, $X_n \cong Y_n$.

Если множества X и Y равносоставлены, мы будем записывать это так: $X \approx Y$.

Неформально говоря, равносоставленность означает, что X можно разбить на части, из которых можно сложить Y (в слове «который» из этого описания спряталось утверждение о конгруэнтности соответствующих частей). В отличие от «школьного» определения мы требуем, чтобы части, на которые разбиваются X и Y, не пересекались (а не просто «пересекались по сторонам, а не по внутренностям»); с другой стороны, сами эти части не обязательно являются многоугольниками или многогранниками — они могут быть сколь угодно плохими, настолько плохими, что их и на рисунке-то изобразить затруднительно. На количество частей, на которые разбиваются X и Y, никаких ограничений не накладывается: оно может быть любым, лишь бы оно было конечно.

Сразу скажем, что при этом «аккуратном» определении равносоставленности теорема Бойяи—Гервина остается верной, но для ее доказательства потребуются некоторые дополнительные усилия; мы расскажем об этом в приложении на с. 10.

Предложение 1. *Если $X \approx Y$ и $Y \approx Z$, то $X \approx Z$.*

Доказательство. Пусть $Y = Y_1 \sqcup \ldots \sqcup Y_n$ — разбиение Y на части, из которых можно сложить X, а $Y = Y'_1 \sqcup \ldots \sqcup Y'_m$ — разбиение Y на части, из которых можно сложить Z. Разобьем теперь Y на mn частей вида $Y_i \cap Y'_j$, где $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$. Из этих частей, очевидно, можно сложить и X, и Z.

Среди множеств $Y_i \cap Y'_j$ могли оказаться и пустые. Формально говоря, это ничему не мешает (в определении 2 нигде не сказано, что множества X_i или Y_j непусты), но если вам это неприятно, прибавьте к доказательству еще одну фразу: те из $Y_i \cap Y'_j$, которые являются пустыми множествами, мы из рассмотрения исключим. □

Если X и Y — два подмножества плоскости (соответственно пространства) и X равносоставлено с некоторым подмножеством в Y, будем обозначать это так: $X \leqslant Y$.

Следующее предложение доказывается точно так же, как предложение 1.

Предложение 2. *Если $X \leqslant Y$ и $Y \leqslant Z$, то $X \leqslant Z$.*

Теперь докажем теорему (одну из версий «теоремы Кантора—Бернштейна»), с помощью которой очень удобно устанавливать равносоставленность множеств.

Теорема 3. *Пусть X и Y — два подмножества плоскости (соответственно пространства). Если $X \leqslant Y$ и $Y \leqslant X$, то $X \approx Y$.*

Доказательство. Будем говорить, что биекция (взаимно однозначное отображение) $f: A \to B$ *задает равносоставленность*, если существует такое разбиение $A = A_1 \sqcup A_2 \sqcup \ldots \sqcup A_n$, что ограничение f на каждое из множеств A_j, где $1 \leqslant j \leqslant n$, сохраняет расстояния.[1] (Если $f: A \to B$ — отображение и $A_1 \subseteq A$, то ограничение f на A_1, обозначаемое $f|_{A_1}$, получится, если мы «забудем» о том, что f определено не только на элементах A_1, и будем рассматривать его исключительно как отображение из A_1 в B.) Ясно, что множества A и B равносоставлены тогда и только тогда, когда между ними есть биекция, задающая равносоставленность.

Пусть теперь X равносоставлено с подмножеством $Y_1 \subseteq Y$, а Y равносоставлено с подмножеством $X_1 \subseteq X$. Тогда существуют биекции $f: X \to Y_1$ и $g: Y \to X_1$, задающие равносоставленность. Мы будем рассматривать f и g как отображения из X в Y и из Y в X соответственно; эти отображения *инъективны* (это означает, что они не склеивают разные точки).

Теперь сформулируем такую лемму.

[1] Всякую биекцию $\varphi: A \to B$, сохраняющую расстояния, можно, как известно, продолжить до движения пространства, переводящего A в B.

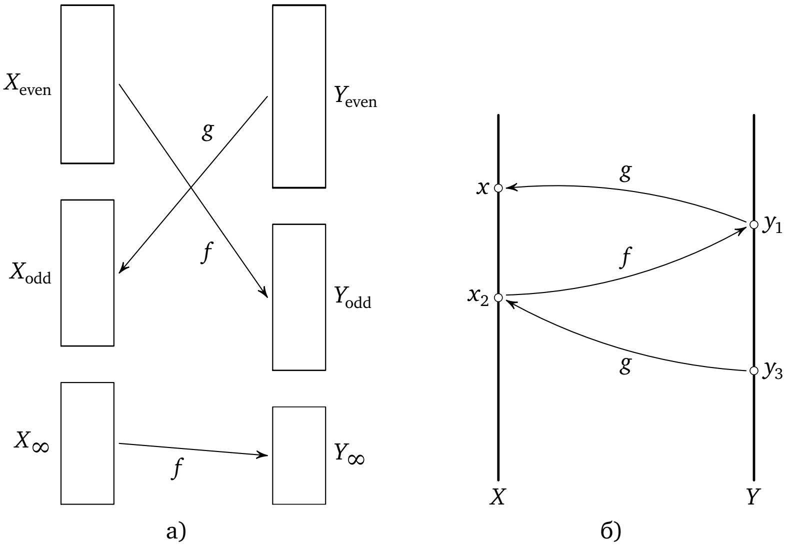


Рис. 2. Лемма 4: а) формулировка; б) доказательство

Лемма 4. *Пусть X и Y — два множества, и пусть $f: X \to Y$ и $g: Y \to X$ — инъективные отображения. Тогда существуют такие разбиения $X = X_{\mathrm{even}} \sqcup X_{\mathrm{odd}} \sqcup X_\infty$ и $Y = Y_{\mathrm{even}} \sqcup Y_{\mathrm{odd}} \sqcup Y_\infty$, что*:

1) ограничение $f|_{X_{\mathrm{even}}}$ задает биекцию между X_{even} и Y_{odd};

2) ограничение $g|_{Y_{\mathrm{even}}}$ задает биекцию между Y_{even} и X_{odd};

3) ограничение $f|_{X_\infty}$ задает биекцию между X_∞ и Y_∞.

См. рис. 2а.

Из этой леммы наша теорема сразу следует. В самом деле, определим отображение $\varphi: X \to Y$ так:

$$\varphi(x) = \begin{cases} f(x), & \text{если } x \in X_{\mathrm{even}}, \\ g^{-1}(x), & \text{если } x \in X_{\mathrm{odd}}, \\ f(x), & \text{если } x \in X_\infty; \end{cases}$$

тогда ясно, что φ — биективное отображение, задающее равносоставленность.

Осталось построить разбиения множеств X и Y, существование которых утверждается леммой. Для этого мы поступим следующим образом. Для каждого элемента $x \in X$ его предшественником назовем такой элемент $y \in Y$, что $g(y) = x$. Предшественника может вообще не быть (его нет, если x не лежит в X_1), но уж если он есть, то он единственен, так как отображение g является инъективным. Аналогично, предшественником элемента $y \in Y$ называется такой $x \in X$, что $f(x) = y$. Если он существует, то он опять-таки единственен.

Теперь по каждому элементу $x \in X$ построим цепочку его предшественников: $y_1 \in Y$ — предшественник элемента x (если предшественник есть), $x_2 \in X$ — предшественник элемента y_1 (опять-таки если таковой есть), и так далее. В конце концов либо цепочка оборвется (у одного из элементов предшественника не окажется), либо окажется, что цепочка бесконечна. Если цепочка конечна, то общее количество предшественников в ней назовем ее длиной: если предшественника у элемента x вообще нет, длина цепочки равна нулю, если у элемента x есть предшественник y_1, а у элемента y_1 предшественника уже нет, то длина цепочки равна 1, и т. д. Аналогично определим цепочки предшественников для элементов $y \in Y$. Заметим, что если длина цепочки предшественников для элемента $x \in X$ равна n, то длина цепочки предшественников для элемента $f(x) \in Y$ равна $n+1$ (предшественники элемента $f(x)$ — это сам x плюс все предшественники элемента x); если же цепочка предшественников для x бесконечна, то бесконечна она и для $f(x)$.

Теперь через X_{even} обозначим множество элементов X, у которых длина цепочки предшественников четна, через X_{odd} — множество элементов X, у которых длина цепочки предшественников нечетна, а через X_∞ — множество элементов X, у которых цепочка предшественников бесконечна (английские слова even и odd как раз и означают «четный» и «нечетный»). Аналогично определим Y_{even}, Y_{odd} и Y_∞. Из сказанного выше ясно, что $f(X_{\text{even}}) = Y_{\text{odd}}$, $g(Y_{\text{even}}) = X_{\text{odd}}$ и $f(X_\infty) = Y_\infty$. Так как f и g — инъективные отображения, лемма доказана, а с ней и теорема. □

Лемма 4 хорошо известна в теории множеств. Из нее точно таким же рассуждением, что и у нас, выводится следующая теорема: если множество X равномощно подмножеству в Y, а множество Y равномощно подмножеству в X, то X и Y равномощны. Эту теорему обычно и называют теоремой Кантора—Бернштейна (без всяких «версий»).

Теперь мы готовы сформулировать теорему Банаха—Тарского.

Будем говорить, что подмножество пространства *ограничено*, если оно содержится в каком-нибудь шаре. Будем говорить, что подмножество пространства *имеет непустую внутренность*, если оно, напротив, содержит какой-нибудь шар.

Теорема 5 (С. Банах, А. Тарский). *Пусть каждое из множеств X и Y является ограниченным подмножеством пространства $\mathbb{R}^3$, имеющим непустую внутренность. Тогда X равносоставлено с Y.*

Прежде чем доказывать теорему 5, сформулируем следующее более слабое утверждение.

Лемма 6. *Пусть B и B_1 — два непересекающихся шара (в пространстве) одинакового радиуса. Тогда $B_1 \cup B \leqslant B$.*

Опираясь на эту лемму, будем рассуждать так. Пусть множества X и Y удовлетворяют условиям теоремы Банаха—Тарского. В силу теоремы Кантора—Бернштейна нам достаточно установить, что $X \leqslant Y$ и $Y \leqslant X$; так как X и Y входят в условие теоремы Банаха—Тарского симметрично, достаточно установить только, что $X \leqslant Y$.

Так как множество Y имеет непустую внутренность, оно содержит некоторый шар B; обозначим его радиус через r. Так как множество X ограничено, его можно разбить на конечное число попарно непересекающихся подмножеств $X = X_1 \sqcup X_2 \sqcup \ldots \sqcup X_n$, каждое из которых помещается в шар радиуса r (например, можно поместить X в некоторый параллелепипед, разбить этот параллелепипед в объединение непересекающихся параллелепипедов с диагональю, меньшей $2r$, и взять в качестве подмножеств X_j пересечения X с этими параллелепипедами). Выберем в пространстве n попарно непересекающихся шаров радиуса r; обозначим их $B_1, \ldots, B_n$. Из всего сказанного выше вытекает, что $X \leqslant B_1 \sqcup \ldots \sqcup B_n$. Далее, из леммы 6 вытекает, что $B_1 \cup B_2 \leqslant B_2$, откуда

$$B_1 \sqcup B_2 \sqcup \ldots \sqcup B_n \leqslant B_2 \sqcup \ldots \sqcup B_n;$$

далее, из той же леммы 6 вытекает, что $B_2 \sqcup B_3 \leqslant B_3$, откуда

$$B_2 \sqcup B_3 \sqcup \ldots \sqcup B_n \leqslant B_3 \sqcup \ldots \sqcup B_n;$$

продолжая в том же духе и пользуясь предложением 2, получаем, что $X \leqslant B_n$; так как $B_n \cong B$, а $B \subseteq Y$, получаем, что $X \leqslant Y$.

Итак, чтобы завершить доказательство теоремы Банаха—Тарского, нам осталось доказать лемму 6; этому мы посвятим два следующих раздела.

Приложение: уточненная теорема Бойяи—Гервина

Мы начали эту книжку с того, что сформулировали теорему о том, что всякие два плоских многоугольника равной площади равносоставлены «в наивном смысле слова». Сейчас мы покажем, что такие многоугольники равносоставлены и в точном смысле (в смысле определения 2). Основную роль в доказательстве будет играть следующая лемма.

Лемма 7. *Пусть X — подмножество плоскости с непустой внутренностью*[1]*, и пусть Y — конечное объединение отрезков. Если $X \cap Y = \varnothing$, то X равносоставлено с $X \cup Y$.*

Пусть X содержит круг радиусом r. Разобьем Y на конечное число частей, каждая из которых представляет собой либо точку, либо полуинтервал достаточно малой длины (не больше r). Далее достаточно доказать, что если к X добавить точку или полуинтервал (дизъюнктно), то полученное объединение будет равносоставлено с X.

Сначала покажем, как присоединить к X полуинтервал, длина которого равна $d \leqslant r$. Возьмем круг радиусом d с центром O, содержащийся в X. На окружности, ограничивающей круг, выберем произвольную точку A_0, а также зададим угол φ, градусная мера которого иррациональна. Образ точки A_0 при повороте вокруг O на угол φ обозначим через A_1, образ точки A_1 при том же повороте — через A_2, и так далее. Это дает бесконечную последовательность точек A_n $(n \geq 0)$, среди которых нет совпадающих ввиду иррациональности φ.

Фигура Φ_0, полученная как объединение радиусов круга с центром O, проведенных из центра ко всем точкам вида A_n, $n \geqslant 0$ (центр круга в эти радиусы не включается), при повороте на угол φ перейдет в фигуру Φ_1, отличающуюся от Φ_0 тем, что полуинтервал $(O, A_0]$ в нее уже не входит. Но он имеет длину d и потому конгруэнтен тому подмножеству в Y, которое мы хотим добавить.

Это доказывает равносоставленность: X разбивается на три части, а именно $(0, A_0]$, Φ_1 и $X \setminus \Phi_0$. Первая из них конгруэнтна добавляемой части, а объединение второй и третьей равносоставлено с X.

Операция добавления точки (вместо полуинтервала) еще проще. Нужно проделать все то же самое, но радиусов не проводить, а фигуру X разбить на три части: $\{A_0\}$, $\{A_1, A_2, \ldots\}$ и дополнение к их объединению. Из второй и третьей части очевидным образом составляется X, а одноточечная фигура $\{A_0\}$, которая нами «высвобождается», есть то, что мы хотели присоединить.

Итак, лемма доказана. Теперь сформулируем и докажем уточненную теорему Бойяи—Гервина.

Теорема 8. *Пусть X и Y — два плоских многоугольника равной площади* (*мы подразумеваем, что граница входит в многоугольник*). *Тогда $X \approx Y$.*

Для доказательства теоремы заметим, что по теореме Бойяи—Гервина многоугольники X и Y можно так разбить конечным числом

[1] По аналогии с пространственным случаем это означает, что X содержит некоторый круг.

отрезков на многоугольники $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_n$ соответственно, что $X_1 \cong Y_1, \ldots, X_n \cong Y_n$. Обозначим через X_0 и Y_0 множества, получающиеся из X и Y после удаления границ всех многоугольников $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_n$. Ясно, что $X_0 \approx Y_0$ (вспомним, что при переходе от «наивного» определения равносоставленности к строгому проблемы возникали только с границами многоугольников). С другой стороны, по доказанной нами лемме имеем $X \approx X_0$, $Y \approx Y_0$. Следовательно, $X \approx Y$.

Задача 2. Постройте в явном виде биекцию между отрезком $[0; 1]$ и интервалом $[0; 1)$.

Задача 3. Покажите, что всякий плоский многоугольник с границей равносоставлен (в смысле определения 2) множеству, получаемому из него с помощью удаления произвольного подмножества границы.

2. Удвоение абстрактного яблока

Чтобы двигаться дальше, полезно посмотреть на понятия конгруэнтности и равносоставленности с немного более абстрактной точки зрения. Следующее определение относится к числу основных определений математики.

Определение 3. Пусть E — некоторое множество и G — множество, состоящее из биекций множества E в себя (не обязательно всех биекций). Множество G называется *группой преобразований* множества E, если выполняются следующие условия:

1) если $g_1 \in G$ и $g_2 \in G$, то и их композиция $g_1 \circ g_2$ также принадлежит G (напомним, что, согласно определению композиции отображений, запись $g_1 \circ g_2 = g_1 g_2$ означает, что сначала выполняется преобразование g_2, а потом g_1);

2) если $g \in G$, то и обратное преобразование g^{-1} принадлежит G;

3) тождественное отображение принадлежит G.

(Для педантов: условие 3 вытекает из условий 1 и 2, если заранее считать, что G непусто.)

Далее мы будем обозначать тождественное отображение символом 1; путаницы с числом один при этом не возникнет.

По мере того как вы будете изучать математику, вам будут встречаться самые разные группы преобразований; мы приведем только несколько примеров.

Примеры групп преобразований. 1. Пусть E — произвольное множество и G — множество *всех* биекций из E в E. Это, очевидно, группа преобразований. Этот пример очень интересен, если E конечно, и менее интересен, если E бесконечно; нам он не встретится.

2. Пусть E — пространство и G — множество всевозможных движений пространства. Тогда G — группа преобразований множества E.

3. Пусть E — плоскость и G — множество всевозможных движений плоскости. Тогда G — группа преобразований множества E.

$3'$. В предыдущем примере заменим слова «всевозможных движений» на «всевозможных движений, сохраняющих ориентацию»[1]. Это также группа преобразований множества E.

$3''$. Множество всевозможных параллельных переносов плоскости (или пространства) также является группой; множество поворотов плоскости относительно фиксированной точки — тоже группа. А вот множество поворотов плоскости относительно всевозможных точек группой уже не является: композицией двух поворотов относительно разных точек может, как известно, быть (неединичный) параллельный перенос.

Если задано множество E и его группа преобразований G, то два подмножества $X, Y \subseteq E$ будем называть конгруэнтными, если существует такое $g \in G$, что $g(X) = Y$. (В нашем определении 1 роль группы G играла группа движений плоскости или пространства.) Определение равносоставленности в этой общей ситуации остается дословно тем же (см. определение 2).

Наш первый шаг к доказательству леммы 6 (и тем самым теоремы Банаха—Тарского) — следующая конструкция.

В этом разделе мы построим такие множество E и его группу преобразований G, что существует разбиение $E = X \sqcup Y \sqcup Z$, для которого $X \cong Y \cong Z$ и при этом $X \cong Y \cup Z$.

Эта конструкция будет для нас модельным примером, но и не только: в следующем разделе мы существенно используем ее в доказательстве.

Сначала мы определим множество E.

Рассмотрим бесконечный граф[2], устроенный следующим образом: из вершины исходят три ребра, из второго конца каждого из этих ребер исходят еще по два ребра, из второго конца этих ребер — снова по два новых ребра, и так до бесконечности (рис. 3). В результате получится бесконечный граф, у которого из каждой вершины исходит по три ребра и в котором из каждой вершины в каждую можно единственным образом дойти по ребрам, не проходя по од-

[1] Иногда их еще называют «собственными движениями» или «движениями первого рода».

[2] Напомним, что граф — это множество *вершин*, в котором какие-то пары вершин соединены *ребром*. Когда графы изображают на рисунке, то вершины изображают точками, а ребра — отрезками или дужками кривых, соединяющих эти точки.

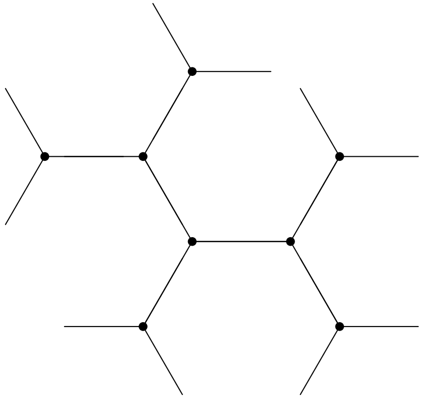
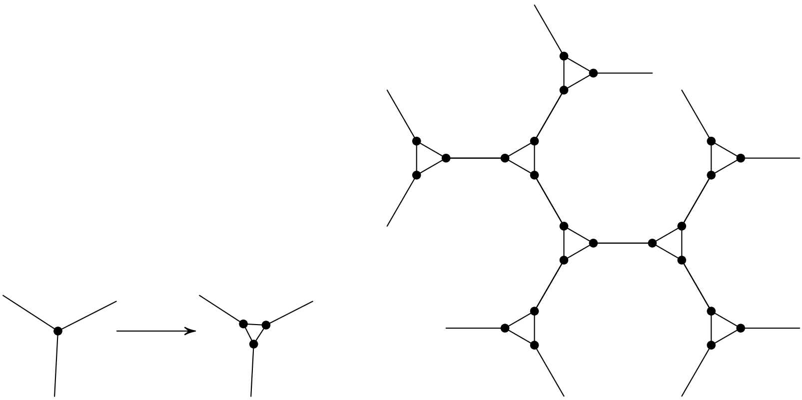


Рис. 3. Бесконечное троичное дерево

ному ребру дважды. Такой граф называется бесконечным троичным деревом.

Теперь «усложним» бесконечное троичное дерево, заменив каждую вершину на треугольник, как на рис. 4а. Граф, который в итоге получится, изображен на рис. 4б; им мы и будем заниматься в этом разделе. Мы назовем его ab-графом (кто такие a и b, скоро узнаете).

Множество E, на котором будет действовать группа — это множество вершин ab-графа.



а) б)

Рис. 4. а) От вершины к треугольнику; б) ab-граф в целом

Теперь определим группу, действующую на множестве E. Чтобы было удобнее ее описывать, условимся различать у ab-графа ребра двух сортов: «короткие» — стороны треугольников, добавленных как на рис. 4б, и «длинные» — ребра, соединяющие разные треугольники. Далее, на каждом из треугольников ab-графа зафиксируем порядок обхода вершин и назовем этот порядок «направлением по часовой стрелке» (если изобразить ab-граф на плоскости — что в принципе возможно, если не гнаться за тем, чтобы все треугольники были одинакового размера, — то можно и в самом деле считать, что направление обхода всех треугольников совпадает с обычным направлением «по часовой стрелке»).

Определим биекции $a\colon E \to E$ и $b\colon E \to E$ следующим образом. Преобразование a переводит каждую вершину ab-графа в другую вершину, лежащую на том же «длинном» ребре (равносильное описание: преобразование a меняет местами два конца у каждого «длинного» ребра). Преобразование b переводит каждую вершину $P \in E$ в вершину, лежащую в том же треугольнике и следующую за P по часовой стрелке (равносильное описание: преобразование b переставляет по кругу вершины каждого треугольника — на один шаг по часовой стрелке). См. рис. 5.

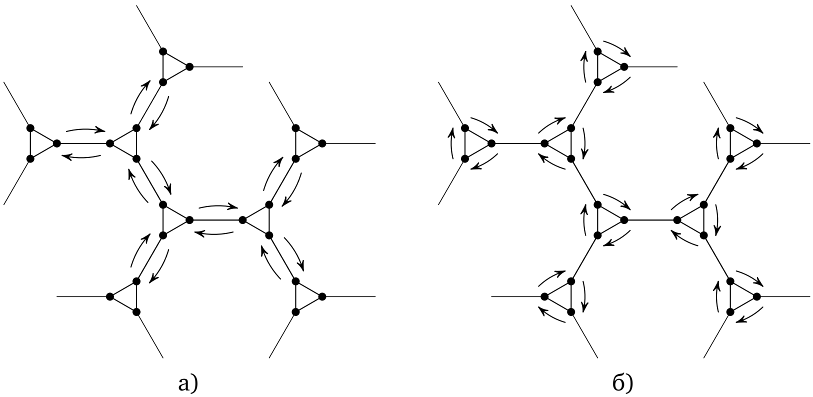


Рис. 5. а) Действие преобразования a; б) действие преобразования b

Из определений очевидно, что композиция преобразования a с самим собой есть тождественное преобразование ($a \circ a = 1$; мы будем это записывать короче: $a^2 = 1$) и что тройная композиция преобразования b с собой есть также тождественное преобразование ($b \circ b \circ b = 1$, или $b^3 = 1$; далее мы будем применять эту «степенную» запись и для других композиций преобразования с собой).

Теперь мы готовы определить и группу G — группу преобразований множества E.

Группа G состоит из тождественного преобразования и всех преобразований, полученных композициями из преобразований a и b в любом порядке и в любом количестве.

(В таких случаях говорят, что G — группа, порожденная преобразованиями a и b.)

Задача 4. Убедитесь, что мы имеем право называть G группой. (Выполнимость условий 1 и 3 из определения 3, конечно, очевидна, но надо еще объяснить, почему из того, что $g \in G$, вытекает, что и $g^{-1} \in G$.)

Посмотрим повнимательнее, из чего состоит группа G. Любая композиция преобразований a и b имеет вид: несколько раз a, затем несколько раз b, затем снова несколько раз a, и т. д. (или наоборот: сначала несколько раз b, затем несколько раз a...). Если писать a^n вместо $a \circ a \circ \ldots \circ a$ (n раз) и a^0 вместо 1, то всякий элемент группы G можно записать в виде

$$a^{m_1} b^{n_1} a^{m_2} b^{n_2} \ldots a^{m_k} b^{n_k}, \tag{2.1}$$

где все $m_1, \ldots, m_k$ и $n_1, \ldots, n_k$ — натуральные числа, кроме, возможно, m_1 и n_k, которым разрешается быть нулями.

Запись (2.1) можно упростить. В самом деле, коль скоро $a^2 = 1$, всякое a^m равно a, если m нечетно, и равно 1, если m четно. Аналогично, всякое b^n равно 1, если n делится на 3, равно b, если n имеет вид $3k+1$, где k целое, и равно b^2, если n имеет вид $3k+2$. Сделаем эти замены; если после этого два сомножителя вида a^m (или два сомножителя вида b^n) окажутся рядом, опять проведем соответствующее упрощение, и т. д. Вот пример:

$$b^2 a^8 b^7 a^7 b a^3 b^9 a^5 b^4 = b^2 \cdot 1 \cdot baba \cdot 1 \cdot a \cdot b = \\ = b^2 babaab = b^3 aba^2 b = 1 \cdot ab \cdot 1 \cdot b = ab^2.$$

То, что получается после всех этих упрощений, называется *приведенным словом*. Точнее говоря, приведенное слово — это выражение вида

$$a^m b^{n_1} a b^{n_2} a b^{n_3} \ldots a b^{n_k}, \tag{2.2}$$

где m равно 0 или 1, числа $n_1, \ldots, n_{k-1}$ равны 1 или 2, а число n_k равно 1, 2 или 0.

Оказывается, что дальше, чем до приведенного слова, упростить выражение (2.1) уже невозможно.

Предложение 9. *Разным приведенным словам соответствуют разные элементы группы G.*

Докажем это предложение. Пусть w_1 и w_2 — два разных приведенных слова. Будем сравнивать их справа налево; если обозначить через v максимальный общий «хвост» слов w_1 и w_2, то $w_1 = u_1 v$, $w_2 = u_2 v$, причем приведенные слова u_1 и u_2 отличаются уже в последнем символе. Выберем теперь произвольную точку $P \in E$ (т. е. произвольную вершину ab-графа) и покажем, что преобразования $w_1 = u_1 v$ и $w_2 = u_2 v$ переведут ее в две разные точки — этим и будет доказано, что $w_1 \neq w_2$.

Положим $Q = v(P)$; тогда нам надо только убедиться, что $u_1(Q) \neq$ $\neq u_2(Q)$. Если одно из слов u_1 и u_2 оканчивается на символ a, а другое оканчивается на b или b^2, то вершины $u_1(Q)$ и $u_2(Q)$ заведомо окажутся в разных треугольниках: одна из них уйдет в соседний треугольник по «длинному» ребру, исходящему из Q, а другая либо останется в том же треугольнике (если все слово имеет вид b или b^2), либо выйдет из этого треугольника по «длинному» ребру, исходящему из другой вершины; точки, вышедшие из одного треугольника по разным «длинным» ребрам, больше уже не встретятся, так как невозможно ни вернуться обратно по уже пройденному ребру, ни обойти целиком по кругу треугольник (в этом месте и используется, что наши слова — приведенные). Если одно из слов u_1 и u_2 оканчивается на b, а другое — на b^2, то рассуждения аналогичны: либо вершина Q попадет в разные вершины одного треугольника, либо выйдет из него по двум разным маршрутам. См. рис. 6.

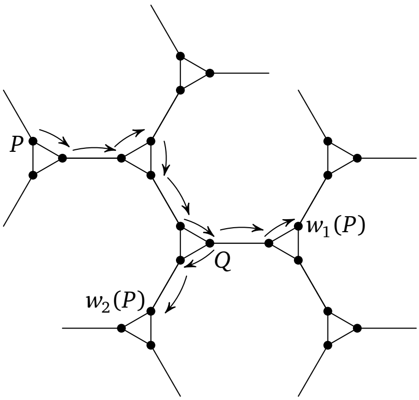


Рис. 6. К доказательству предложения 9:
$w_1 = babab^2ab$, $w_2 = ab^2ab^2ab$, $v = bab^2ab$

Итак, мы установили взаимно однозначное соответствие между элементами группы G и приведенными словами. Формально это соответствие описывается так: $G=\langle a, b \mid a^2=1, b^3=1\rangle$ (словами: «G — группа с образующими a и b и определяющими соотношениями $a^2=1$ и $b^3=1$»).

Покажем еще, что элементов множества E «столько же», сколько элементов группы G. Разумеется, как E, так и G — счетное множество, но утверждение про «столько же» верно и в гораздо более сильном смысле. Именно, докажем такое предложение.

Предложение 10. *Для любых двух вершин $P, Q \in E$ существует и единственно преобразование $g \in G$, для которого $Q=g(P)$.*

В самом деле, если записать g в виде приведенного слова (что, как мы знаем, можно сделать единственным образом), то равенство $Q=g(P)$ задает путь из P в Q по ребрам графа, удовлетворяющий следующим условиям:

1) ни одно ребро не проходится дважды;

2) «короткие» ребра (стороны треугольников) проходятся по часовой стрелке;

3) проходя треугольник, запрещено делать полный круг.

Ясно, что такой путь из P в Q существует и, более того, единственен: последовательность прохождения треугольников единственна, так как в дереве единственен путь из вершины в вершину, не проходящий дважды по одному ребру (условие 1), а переход от одного «длинного» ребра к другому, осуществляемый внутри треугольников, единственен ввиду условий 2 и 3. Этим предложение 10 и доказано.

Если теперь зафиксировать какую-нибудь вершину $P \in E$, то соответствие $g \mapsto g(P)$ задает, в силу только что доказанного предложения, биекцию между G и E. Таким образом, у нас имеется разумное взаимно однозначное соответствие между элементами группы G и элементами множества E.[1]

Теперь мы готовы построить искомое разбиение множества E. Для этого мы пометим каждую вершину ab-графа буквой x, y или z, руководствуясь следующим алгоритмом. Выберем сначала произвольно одну вершину и пометим ее буквой x. Затем будем последовательно помечать остальные вершины, руководствуясь следующими тремя правилами:

1) если одна из вершин треугольника помечена, то буквы на двух остальных ставятся таким образом, чтобы три вершины были поме-

[1] Правда, это не единственное разумное соответствие: если выбрать другую вершину P, соответствие будет уже другим.

Если G — группа преобразований множества E, для которой выполнено предложение 10, то говорят, что E является «главным однородным пространством над G».

чены тремя разными буквами, причем алфавитный порядок вершин x, y, z совпадал с порядком «по часовой стрелке»;

2) если одна вершина «длинного» ребра помечена буквой x, то другую его вершину помечают буквой y;

3) если одна вершина «длинного» ребра помечена буквой y или z, то другую его вершину помечают буквой x.

Эта расстановка меток проиллюстрирована на рис. 7.

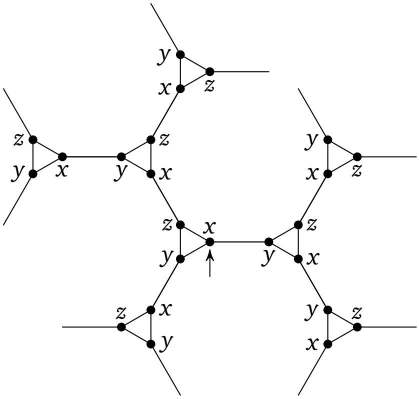


Рис. 7. Расстановка букв x, y и z. Вершина, с которой расстановка начинается (первая «x-вершина»), указана стрелкой

Обозначим теперь через X, Y и Z множество вершин, помеченных буквами x, y и z соответственно. Из правила 1 выше вытекает, что $b(X)=Y$, $b(Y)=Z$ и $b(Z)=X$; из правил 2 и 3 вытекает, что $a(X)=Y\cup Z$.

Стало быть, мы построили искомое разбиение множества E:

Предложение 11. *Существует разбиение $E=X\sqcup Y\sqcup Z$, для которого $X\cong Y\cong Z$ и $X\cong Y\sqcup Z$. При этом конгруэнтность множеств X, Y и Z задается преобразованием b, а конгруэнтность множеств X и $Y\sqcup Z$ задается преобразованием a.*

Конгруэнтность в этом предложении понимается, естественно, относительно группы преобразований G.

3. Основная конструкция

Нам пора возвращаться от абстрактных графов к геометрии. В этом разделе мы докажем, наконец, лемму 6 (рекомендуем читателю освежить в памяти ее формулировку) и тем самым завершим доказательство теоремы Банаха—Тарского.

Раз и навсегда заметим, что лемму 6 достаточно, очевидно, доказать для двух шаров радиуса 1, так что далее мы будем доказывать ее именно в этих предположениях.

Пусть теперь S — сфера радиуса 1 с центром в точке O. Начнем с того, что построим группу движений этой сферы, устроенную так же, как группа G из предыдущего раздела.[1]

Выберем прямые ℓ_A и ℓ_B, проходящие через точку O — центр сферы; угол между ними обозначим через φ (каков должен быть этот угол, мы уточним в дальнейшем). Обозначим через A поворот на 180° относительно прямой ℓ_A, а через B — поворот на 120° относительно прямой ℓ_B (в одном из двух возможных направлений; мы раз и навсегда выберем это направление и зафиксируем его до конца доказательства). Обозначим через Γ группу преобразований пространства (а также сферы S), состоящую из тождественного преобразования и всевозможных композиций A и B, в любом количестве и любом порядке. Ясно, что $A^2=1$ и $B^3=1$ (так что в этом отношении преобразования A и B аналогичны преобразованиям a и b из предыдущего раздела). Поэтому так же, как раньше, устанавливаем, что всякий элемент группы Γ можно записать в виде приведенного слова относительно A и B (под приведенным словом мы понимаем то же, что в выражении (2.2)).

У группы G было и еще одно замечательное свойство: разным приведенным словам соответствовали разные преобразования. Верно ли это для группы Γ? Если выбирать прямые ℓ_A и ℓ_B как попало, то ответ отрицателен. Например, нетрудно убедиться, что если ℓ_A и ℓ_B совпадают, то приведенные слова $ABAB$ и B^2 задают одно и то же преобразование (поворот на -120°). Тем не менее мы сейчас покажем, что можно выбрать прямые так, чтобы подобных безобразий не происходило.

Предложение 12. *Угол φ между прямыми ℓ_A и ℓ_B можно выбрать таким образом, что различным приведенным словам от A и B будут соответствовать различные движения пространства.*

Объясним, как можно доказать это предложение. Пусть

$$u=A^{m_1}B^{n_1}\dots A^{m_k}B^{n_k}, \qquad v=A^{m'_1}B^{n'_1}\dots A^{m'_r}B^{n'_r}$$

— два приведенных слова от A и B, задающие одно и то же преобразование пространства. Запишем

$$uv^{-1}=A^{m_1}B^{n_1}\dots A^{m_k}B^{n_k}B^{-n'_r}A^{-m'_r}\dots B^{-n'_1}A^{-m'_1}$$

и упростим правую часть, как в предыдущем разделе, чтобы получилось приведенное слово. Если правая часть упростилась до 1, то слова

[1] По-ученому говоря, «изоморфную группе G».

u и v совпадают; в противном случае в правой части останется нетривиальное приведенное слово от A и B, которому соответствует тождественное преобразование пространства.

Итак, нам достаточно доказать вот что: *при подходящем выборе угла φ между осями поворотов A и B всякое нетривиальное (отличное от $1 = A^0B^0$) приведенное слово от A и B задает нетождественное преобразование пространства*. Чтобы это сделать, надо повнимательнее разобраться с композициями поворотов.

Условимся обозначать через R_ℓ^α поворот пространства на угол α относительно прямой ℓ, а через S_π — симметрию пространства относительно плоскости π. Как известно, $R_\ell^\alpha = S_{\pi_1} \circ S_{\pi_2}$, где π_1 и π_2 — плоскости, проходящие через ℓ и образующие угол $\alpha/2$; при этом в качестве π_1 (или π_2) можно выбрать любую плоскость, проходящую через ℓ, и тогда вторая плоскость будет определена однозначно.

Рассмотрим теперь композицию поворотов $R_{\ell_1}^{\alpha_1} \circ R_{\ell_2}^{\alpha_2}$, где прямые ℓ_1 и ℓ_2 проходят через точку O. Обозначим через π плоскость, проходящую через ℓ_1 и ℓ_2; тогда $R_{\ell_1}^{\alpha_1} = S_{\pi_1} \circ S_\pi$, $R_{\ell_2}^{\alpha_2} = S_\pi \circ S_{\pi_2}$, где плоскость π_1 образует с плоскостью π угол $\alpha_1/2$, а плоскость π_2 образует с ней же угол $\alpha_2/2$. Следовательно,

$$R_{\ell_1}^{\alpha_1} \circ R_{\ell_2}^{\alpha_2} = S_{\pi_1} \circ S_\pi \circ S_\pi \circ S_{\pi_2} = S_{\pi_1} \circ S_{\pi_2} = R_{\ell'}^{2\beta},$$

где ℓ' — прямая, по которой пересекаются плоскости π_1 и π_2 (эта прямая проходит через O, так как и π_1, и π_2 содержит O), а β — угол между плоскостями π_1 и π_2 (см. рис. 8).

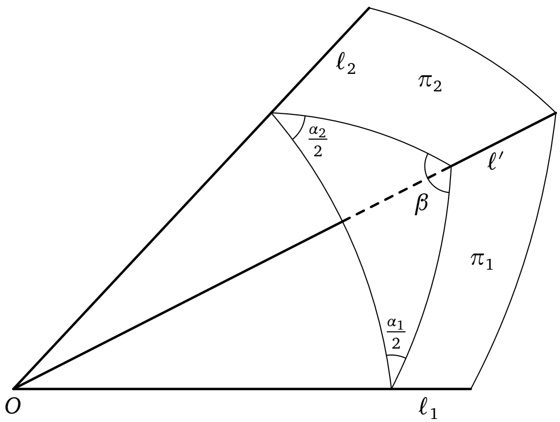


Рис. 8. Композиция поворотов

Мы видим, что композиция двух поворотов относительно прямых, проходящих через O — тоже поворот относительно прямой, проходящей через O. В частности, множество таких поворотов является группой преобразований пространства.[1] Кроме того, ясно, что существует формула, выражающая угол β через углы α, β и φ (раздел стереометрии, в котором изучаются эти формулы, называется сферической тригонометрией; они аналогичны планиметрическим формулам для нахождения, скажем, всех сторон и углов треугольника, если даны две стороны и угол).

Отсюда следует, что для любого (нетривиального) приведенного слова от $A = R_{\ell_A}^{\pi}$ и $B = R_{\ell_B}^{2\pi/3}$ условие «данное слово задает тождественное преобразование пространства» задает некоторое тригонометрическое уравнение на угол φ. Можно показать (мы приведем соответствующие вычисления в приложении к этому разделу), что оно имеет только конечное число решений (с точностью до кратных 360°, разумеется).[2] Следовательно, для всякого приведенного слова от A и B множество углов φ, для которых это слово задает тождественное преобразование, конечно. Однако множество всех приведенных слов, очевидно, счетно; поэтому множество углов φ, при которых какое-нибудь приведенное слово задает тождественное отображение, также счетно. Так как множество всех возможных углов несчетно, существуют углы φ, для которых *никакое* нетривиальное приведенное слово от A и B не задает тождественное отображение, что и требовалось.

Начиная с этого момента мы будем предполагать, что угол φ между осями поворотов A и B выбран таким образом, что разным приведенным словам от A и B соответствуют разные преобразования пространства. Стало быть, группа G, состоящая из тождественного преобразования и всевозможных композиций A и B, устроена так же, как группа Γ из предыдущего раздела: она состоит из тождественного преобразования и всевозможных приведенных слов от A и B (приведенных в смысле формулы (2.2), см. с. 16), причем $A^2 = B^3 = 1$ и разным приведенным словам соответствуют разные элементы группы.

Так как всякий элемент группы G — композиция поворотов относительно прямых, проходящих через O, всякий элемент G также являет-

[1] Эта группа обозначается $SO(3, \mathbb{R})$ (не будем обсуждать, почему); мы с ней еще встретимся.

[2] Это утверждение не самоочевидно: в принципе возможен случай, когда соответствующее уравнение после упрощений примет вид $0 = 0$. Например, если A — поворот плоскости на 180°, а B — поворот плоскости на 120° относительно другой точки, то приведенное слово $(AB)^6$ есть тождественное преобразование. То, что в пространстве этих патологий можно избежать, нуждается в обосновании.

ся поворотом относительно некоторой прямой, проходящей через O. Поскольку группа G счетна (так как счетно множество приведенных слов), множество осей поворотов, являющихся элементами группы G, также счетно. Обозначим через Σ множество точек пересечения этих осей со сферой S (напомним, что это сфера единичного радиуса с центром в точке O); так как прямая, проходящая через центр, пересекается со сферой в двух точках, множество Σ также счетно.

Множество Σ можно описать и по-другому. Именно, рассмотрим поворот R_ℓ^α (на угол α относительно оси ℓ, проходящей через точку O); точки прямой ℓ — единственные точки пространства, переходящие при этом повороте в себя («неподвижные точки» поворота).[1] Если рассматривать R_ℓ^α как преобразование сферы S, то его неподвижные точки — точки пересечения прямой ℓ со сферой. Тем самым *множество Σ есть множество неподвижных точек нетождественных преобразований $g \in G$*. Нам понадобится еще одно свойство множества Σ.

Лемма 13. *Если $p \in S \setminus \Sigma$ и $g \in G$, то $g(p) \in S \setminus \Sigma$.*

В самом деле, будем рассуждать от противного. Если $g(p) \in \Sigma$, то, как мы выяснили, существует такое нетождественное преобразование $h \in G$, что $h(g(p)) = g(p)$. Подействуем на обе части этого равенства преобразованием g^{-1}:

$$g^{-1}(h(g(p))) = g^{-1}(g(p)) = p.$$

Стало быть, точка p является неподвижной точкой преобразования $g^{-1}hg \in G$; так как $p \notin \Sigma$, отсюда следует, что $g^{-1}hg = 1$; умножим обе части этого равенства на g слева, получим $hg = g$; умножив обе части полученного равенства на g^{-1} справа, получим, что $h = 1$, а это противоречит тому, что h — нетождественное преобразование, и лемма доказана.

Операция перехода от h к $g^{-1}hg$ называется сопряжением; она часто встречается при работе с группами.

Рассмотрим теперь произвольную точку $p \in S \setminus \Sigma$; обозначим через E_p множество всевозможных точек вида $g(p)$, где $g \in G$ (это множество называется *орбитой* точки p). Докажем следующую лемму.

Лемма 14. *Для любых двух точек $p_1, p_2 \in E_p$ существует и единственен элемент $g \in G$, для которого $g(p_1) = g_2$.*

В самом деле, пусть $p_1 = g_1(p)$, $p_2 = g_2(p)$ и мы хотим найти такой $g \in G$, что $g(g_1(p)) = g_2(p)$. Действуя на обе части этого равенства преобразованием g_2^{-1}, получаем, что $g_2^{-1}gg_1(p) = p$, то есть p — неподвиж-

[1] Если, конечно, угол α не кратен 360° — тогда наш «поворот» является тождественным преобразованием и всякая точка пространства является для него неподвижной.

ная точка преобразования $g_2^{-1}gg_1$; поскольку $p \notin \Sigma$, отсюда вытекает, что

$$g_2^{-1}gg_1 = 1 \Rightarrow gg_1 = g_2 \Rightarrow g = g_1 g_2^{-1},$$

так что если искомое g существует, то оно единственно. Однако же найденное нами g не является «посторонним корнем»: $g_2 g_1^{-1}(g_1(p)) = g_2(g_1^{-1}(g(p))) = g_2(p)$, так что и существование, и единственность доказаны.

Давайте теперь сопоставим лемму 14 и предложение 10: оказывается, что не только группа G отождествляется с Γ при соответствии $A \mapsto a$ и $B \mapsto b$, но и множество E_p, группой преобразований которого является G, устроено в точности так же, как множество E (множество вершин ab-графа). Мы изобразили это отождествление E_p и E на рис. 9 (не забудьте, что A соответствует a и B соответствует b). Стало быть, множество E_p обладает теми же свойствами, что и множество E. В частности, можно применить предложение 11: существует такое разбиение $E_p = X_p \sqcup Y_p \sqcup Z_p$, что $B(X_p) = Y_p$, $B(Y_p) = Z_p$, $B(Z_p) = X_p$ и при этом $A(X_p) = Y_p \sqcup Z_p$.

Заметим теперь, что любые две орбиты точек из $S \setminus \Sigma$ либо вообще не пересекаются, либо совпадают: если орбита точки p пересекается с орбитой точки q, т. е. если $g(p) = h(q)$ для каких-то $g, h \in G$, то

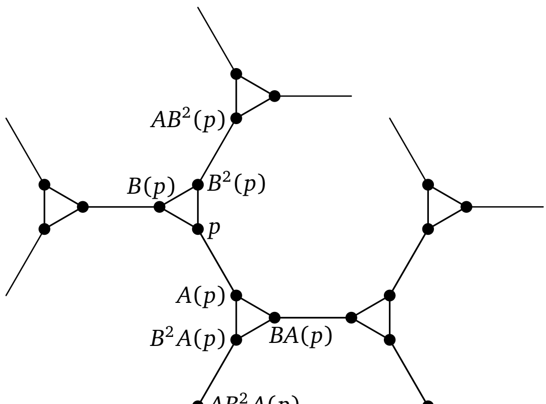


Рис. 9. Вершины ab-графа как элементы орбиты E_p

$q = h^{-1}g(p)$ и орбиты точек q и p совпадают: всякая точка вида $w(q)$, где $w \in G$, имеет также вид $wh^{-1}g(p)$. Итак, множество $X \setminus \Sigma$ разбивается на непересекающиеся орбиты; каждую такую орбиту E_q разобьем на X_q, Y_q и Z_q, как выше. Обозначим теперь через X объединение всех X_q, через Y — объединение всех Y_q и через Z — объединение всех Z_q. Тогда $B(X) = Y$, $B(Y) = Z$, $B(Z) = X$ и при этом $A(X) = Y \cup Z$. Мы доказали следующее утверждение.

Предложение 15. *Существует разбиение сферы $S = X \sqcup Y \sqcup Z \sqcup \Sigma$, в котором Σ — счетное множество, $X \cong Y \cong Z$ и $X \cong Y \sqcup Z$.*

Конгруэнтность здесь уже понимается в самом обычном смысле (определение 1).

Именно это предложение было доказано Хаусдорфом в 1914 году в его книге «Grundzüge der Mengenlehre» («Основы теории множеств»), и именно на него опирались Банах и Тарский в своей статье 1924 года.

Основное для доказательства леммы 6 мы уже сделали; то, что остается, — дело техники. Ей мы сейчас и займемся.

Главное — разобраться с «путающимся под ногами» счетным множеством Σ.

Лемма 16. *Существует такой поворот R сферы S (относительно прямой, проходящей через O), что $R(\Sigma) \cap \Sigma = \varnothing$.*

В доказательстве этой леммы специфика множества Σ никак использована не будет: нам будет достаточно того, что множество Σ счетно.

Пусть вообще m и n — две произвольные точки сферы и поворот $R = R_\ell^\alpha$, где $\ell \ni O$, таков, что $R(m) = n$. Тогда, очевидно, точки пересечения оси ℓ со сферой лежат на большом круге, проходящем через середину дуги mn и перпендикулярном ей (планиметрический аналог: если поворот плоскости относительно точки L переводит M в N, то L лежит на серединном перпендикуляре к отрезку MN); будем называть этот большой круг также «серединным перпендикуляром» точек m и n. Для каждой пары точек $m, n \in \Sigma$ рассмотрим большой круг — их «серединный перпендикуляр». Если пересечение оси ℓ со сферой S не лежит ни на одном из этих серединных перпендикуляров, то, очевидно, всякий (нетождественный) поворот относительно ℓ будет искомым. Значит, нам достаточно доказать, что существует точка сферы, не лежащая ни на одном из описанных «серединных перпендикуляров».

Поскольку множество Σ счетно, множество всевозможных пар его точек также счетно. Тем самым все будет доказано, как только мы докажем следующую лемму.

Лемма 17. *Пусть на сфере S задано счетное семейство больших кругов. Тогда существует точка $q \in S$, не лежащая ни на одном из них.*

Эта лемма является частным случаем так называемой «теоремы Бэра» из анализа. Мы дадим ей два доказательства: одно «научное», пригодное с небольшими изменениями и для доказательства общей теоремы Бэра, а другое — короткое и простое, использующее специфику нашей ситуации.

«Научное» доказательство. Пусть $K_1, K_2, \ldots, K_r, \ldots$ — счетное семейство больших кругов. Ясно, что существует сферический сегмент $B_1 \subset S$ радиусом, скажем, 1°, не пересекающийся с K_1 (здесь и далее мы считаем, что сферические сегменты включают в себя границу). Далее, даже если B_1 пересекается с K_2, заведомо существует сферический сегмент B_2 радиуса $(1/3)^\circ$, содержащийся в B_1 и не пересекающийся с K_2.

Аналогично, в B_2 можно найти сферический сегмент B_3 радиуса $(1/3^2)^\circ$, не пересекающийся с K_3, и так далее. Согласно одному аналогу известного «принципа стягивающихся отрезков», существует точка $q \in S$, лежащая во всех этих сегментах. Тогда ясно, что q не лежит ни на одном из K_j: она не лежит на K_1, поскольку $q \in B_1$, она не лежит на K_2, поскольку $q \in B_2$ и т. д.

Простое доказательство. Это доказательство было предложено Леней Янушевичем на лекции. Пусть опять $K_1, K_2, \ldots, K_r, \ldots$ — счетное семейство больших кругов. Поскольку множество всех больших кругов на сфере несчетно, существует большой круг K, не совпадающий ни с одним из K_j. Пересечение K с каждым из K_j состоит ровно из двух точек, так что множество точек на K, лежащих также хоть на каком-нибудь из K_j, является счетным. Поскольку множество всех точек на большом круге несчетно, на K найдется точка, не лежащая ни на одном из K_j.

Итак, лемма 16 доказана. Давайте ею воспользуемся. Пусть R — поворот, существование которого утверждается леммой. Тогда $R(\Sigma) \subset X \sqcup Y \sqcup Z$.

Положим

$$X_1 = \{p \in \Sigma \colon R(p) \in X\},$$
$$Y_1 = \{p \in \Sigma \colon R(p) \in Y\},$$
$$Z_1 = \{p \in \Sigma \colon R(p) \in Z\};$$

тогда $\Sigma = X_1 \sqcup Y_1 \sqcup Z_1$, причем $X_1 \leqslant X$, $Y_1 \leqslant Y$, $Z_1 \leqslant Z$. Так как $Y \cong X$ и $Z \cong X$, можно сказать и чуть больше: каждое из множеств X_1, Y_1 и Z_1 равносоставлено с некоторым подмножеством в X.

Обозначим теперь через B шар радиуса 1 с центром в O. По каждому из множеств X, Y, Z, X_1, Y_1, Z_1 построим подмножество в B, обозначаемое $\mathscr{X}, \mathscr{Y}, \mathscr{Z}, \mathscr{X}_1, \mathscr{Y}_1, \mathscr{Z}_1$ соответственно, следующим образом: каждую точку $p \in X$ соединим отрезком с центром шара и включим в множество $\mathscr{X}$ все точки этого отрезка, кроме самого центра. Тогда получаем, что

$$B = \mathscr{X} \sqcup \mathscr{Y} \sqcup \mathscr{Z} \sqcup \mathscr{X}_1 \sqcup \mathscr{Y}_1 \sqcup \mathscr{Z}_1 \sqcup \{O\},$$

причем каждое из семи подмножеств в правой части конгруэнтно некоторому подмножеству в $\mathscr{X}$: про первые шесть это очевидно из построения и из уже сказанного о подмножествах сферы, а одноточечное множество $\{O\}$, очевидно, конгруэнтно подмножеству всякого непустого множества. Пусть теперь B_1 — шар радиуса 1, не пересекающийся с B. У него имеется аналогичное разбиение на семь подмножеств $\mathscr{X}'$, $\mathscr{Y}'$,..., $\{O_1\}$, каждое из которых конгруэнтно подмножеству в $\mathscr{X}$. Стало быть, $B_1 \sqcup B$ разбивается в дизъюнктное объединение 14 подмножеств

$$B_1 \sqcup B = C_1 \sqcup C_2 \sqcup \ldots \sqcup C_{14},$$

причем их можно занумеровать так, что $C_1 = \mathscr{X}$, и при этом $C_j \leqslant \mathscr{X}$ (а тем самым $C_j \leqslant \mathscr{Y}$, $C_j \leqslant \mathscr{Z}$ — ведь множества $\mathscr{X}$, $\mathscr{Y}$ и $\mathscr{Z}$ конгруэнтны!) для всякого $j \leqslant 14$. Далее, если $i \leqslant 14$, $j \leqslant 14$, то $C_i \sqcup C_j \leqslant \mathscr{X}$; в самом деле, так как $C_i \leqslant \mathscr{Y}$ и $C_j \leqslant \mathscr{Z}$, имеем

$$C_i \sqcup C_j \leqslant \mathscr{Y} \sqcup \mathscr{Z} \cong \mathscr{X}.$$

Теперь запишем цепочку «неравенств»:

$$B_1 \sqcup B = (C_1 \sqcup C_2) \sqcup C_3 \sqcup \ldots \sqcup C_{14} \leqslant (\mathscr{X} \sqcup C_3) \sqcup C_4 \sqcup \ldots \sqcup C_{14} \leqslant \\ \leqslant \mathscr{X} \sqcup C_4 \sqcup \ldots \sqcup C_{14} \leqslant \ldots \leqslant \mathscr{X}.$$

Поскольку $\mathscr{X} \subset B$, получаем, что $B_1 \sqcup B \leqslant B$. Значит, лемма 6 доказана, а с ней и теорема Банаха—Тарского.

Приложение: об уравнениях на угол φ

Напомним основных действующих лиц. Через A обозначен поворот на $180°$ вокруг прямой ℓ_A, через B — поворот на $120°$ вокруг прямой ℓ_B; обе эти прямые проходят через точку O и образуют угол φ.

Будем называть приведенные слова от A и B, отличные от $1 = A^0 B^0$, *нетривиальными*. Нам надо доказать следующее.

Предложение 18. *Для всякого нетривиального приведенного слова от A и B существует лишь конечное число углов φ, при которых это слово задает тождественное отображение.*

Мы воспроизведем, с несущественными изменениями, доказательство самого Хаусдорфа. Начнем с простой леммы.

Лемма 19. *Если некоторое нетривиальное приведенное слово от A и B равно* 1 (*тождественному отображению*), *то существует нетривиальное приведенное слово от A и B, равное* 1 *и при этом начинаю-*

щееся на B в ненулевой степени, а оканчивающееся на A в ненулевой степени.[1]

В самом деле, пусть α — приведенное слово от A и B. Если оно не имеет указанного вида, то для него имеются три возможности:

1) α начинается на A и оканчивается на A;

2) α начинается на B и оканчивается на B;

3) α начинается на A и оканчивается на B.

Положим $\beta = B\alpha B^2\alpha$ в первом случае, $\beta = \alpha A\alpha A$ во втором и $\beta = A\alpha A$ — в третьем. В первом и втором случаях слово β уже приведенное, причем оно начинается на B и оканчивается на A; в третьем случае слово β становится приведенным после сокращения $A^2 = 1$ в начале, причем полученное приведенное слово также начинается на B и оканчивается на A. Если при этом $\alpha = 1$, то во всех трех случаях $\beta = 1$, так что лемма доказана.

Вернемся к доказательству предложения 18. Ввиду леммы 19 нам достаточно рассматривать приведенные слова, начинающиеся на B и оканчивающиеся на A. Заметим, что всякое приведенное слово такого вида есть произведение преобразований вида BA или B^2A.

Введем в пространстве прямоугольную систему координат, в которой точка O — начало координат, прямая ℓ_B совпадает с осью Oz, а прямая ℓ_A лежит в плоскости Oxz. Следуя Хаусдорфу, положим $\lambda = \cos 120^\circ = -1/2$, $\mu = \sin 120^\circ = \sqrt{3}/2$. Запишем в координатах, как действуют преобразования A и B, а также BA и B^2A. В формулах ниже подразумевается, что точка с координатами (x, y, z) переходит в точку с координатами (x', y', z').

$$A\colon\ \begin{aligned} x' &= \cos 2\varphi\cdot x + \sin 2\varphi\cdot z,\\ y' &= -y,\\ z' &= -\sin 2\varphi\cdot x + \cos 2\varphi\cdot z;\end{aligned} \qquad B\colon\ \begin{aligned} x' &= \lambda x - \mu y,\\ y' &= \mu x + \lambda y,\\ z' &= z;\end{aligned}$$

$$BA\colon\ \begin{aligned} x' &= \lambda x + \mu y,\\ y' &= -\mu x + \lambda y,\\ z' &= z;\end{aligned} \qquad B^2\colon\ \begin{aligned} x' &= \lambda\cos 2\varphi\cdot x + \mu y + \lambda\sin 2\varphi\cdot z,\\ y' &= \mu\cos 2\varphi\cdot x - \lambda y + \mu\sin 2\varphi\cdot z,\\ z' &= -\sin 2\varphi\cdot x + \cos 2\varphi\cdot z;\end{aligned} \tag{3.1}$$

$$B^2A\colon\ \begin{aligned} x' &= \lambda\cos 2\varphi\cdot x - \mu y + \lambda\sin 2\varphi\cdot z,\\ y' &= -\mu\cos 2\varphi\cdot x - \lambda y - \mu\sin 2\varphi\cdot z,\\ z' &= -\sin 2\varphi\cdot x + \cos 2\varphi\cdot z.\end{aligned}$$

Проследим за судьбой точки $p = (0; 0; 1)$, на которую действуют композиции преобразований BA и B^2A.

[1] Далее этих оговорок про ненулевую степень мы делать не будем.

Лемма 20. *Пусть преобразование u — композиция n преобразований вида BA или B^2A. Тогда координаты точки $u(p)$ имеют вид*

$$\begin{aligned} x &= \sin 2\varphi(a_{n-1}\cos^{n-1}2\varphi + a_{n-2}\cos^{n-2}2\varphi + \ldots + a_0),\\ y &= \sin 2\varphi(b_{n-1}\cos^{n-1}2\varphi + b_{n-2}\cos^{n-2}2\varphi + \ldots + b_0),\\ z &= c_n\cos^n 2\varphi + c_{n-1}\cos^{n-1}2\varphi + c_{n-2}\cos^{n-2}2\varphi + \ldots + c_0, \end{aligned} \tag{3.2}$$

причем $c_n \neq 0$.

Эта лемма доказывается индукцией по n. В самом деле, то, что она верна при $n=1$, видно из формул (3.1); пусть теперь она верна для n, то есть $u(p)$ имеет координаты, указанные в (3.2). Посмотрим, какие будут координаты у точки, скажем, $B^2A(p)$ (для $BA(p)$ рассуждение совершенно аналогично). Если обозначить их (x', y', z'), то, согласно формулам (3.1), выполнено равенство

$$x' = \lambda\cos 2\varphi\cdot x - \mu y + \lambda\sin 2\varphi\cdot z. \tag{3.3}$$

Так как по предположению индукции x и y являются многочленами степени $\leqslant n-1$ от $\cos 2\varphi$, умноженными на $\sin 2\varphi$, а z является многочленом степени n от $\cos 2\varphi$, получаем следующее:
— первое слагаемое в правой части (3.3) есть многочлен степени $\leqslant n$ от $\cos 2\varphi$, умноженный на $\sin 2\varphi$;
— второе слагаемое в правой части (3.3) есть многочлен степени $\leqslant n-1$ от $\cos 2\varphi$, умноженный на $\sin 2\varphi$;
— третье слагаемое в правой части (3.3) есть многочлен степени n от $\cos 2\varphi$, умноженный на $\sin 2\varphi$.

Стало быть, и координата x' есть многочлен степени $\leqslant n$ от $\cos 2\varphi$, умноженный на $\sin 2\varphi$, так что для x' индуктивный переход мы провели.

Индуктивный переход для y' проводится совершенно аналогично; проведем теперь индуктивный переход для z'. Ввиду формул (3.1) имеем

$$z' = -\sin 2\varphi\cdot x + \cos 2\varphi\cdot z, \tag{3.4}$$

причем предположение индукции говорит нам, что z — многочлен степени n от $\cos 2\varphi$ с ненулевым старшим коэффициентом, а x равно $\sin 2\varphi$, умноженному на многочлен степени $\leqslant n-1$ от $\cos 2\varphi$. Стало быть, второе слагаемое в правой части (3.4) есть многочлен степени $n+1$ от $\cos 2\varphi$ (с ненулевым первым коэффициентом), а первое слагаемое равно произведению $-\sin^2 2\varphi = \cos^2 2\varphi - 1$ на многочлен степени $\leqslant n-1$ от $\cos 2\varphi$; это произведение есть, очевидно, многочлен степени $\leqslant n$ от $\cos 2\varphi$. Так как сумма многочлена степени $\leqslant n$ и многочлена степени ровно $n+1$ есть многочлен степени ровно $n+1$, индуктивный переход для z' произведен и лемма доказана.

Пусть теперь произведение n преобразований вида BA или B^2A является тождественным; тогда оно, в частности, переводит в себя точку $P=(0;0;1)$, так что z-координата ее образа равна 1. Ввиду леммы это означает, что некоторый многочлен степени n от $\cos 2\varphi$ равен 1, и это тригонометрическое уравнение имеет конечное число решений с точностью до кратных 360°, что и требовалось.

4. Обсуждение

Итак, теорема Банаха—Тарского наконец доказана; в частности, доказано, что равносоставлены любые два многогранника, независимо от их объемов. Давайте попробуем разобраться, почему ее формулировка столь шокирует и в чем конкретно выражается то обстоятельство, что этот результат противоречит интуиции.

Самый первый источник недоумения: «Как же так можно, что ж, выходит, из одной картофелины может сделать две?». На это есть простой ответ: картофелины — не математический объект, и ни одна математическая теорема ничего о картошке (равно как и о других реальных предметах) не говорит. Подмножества в $\mathbb{R}^3$ являются, конечно, «моделями» предметов, встречающихся в реальном мире, но сходство модели с оригиналом простирается не до бесконечности. Для демонстрации этой особенности взаимоотношений математики и реальности вообще незачем привлекать результаты Банаха и Тарского: всякий школьник понимает, что куб с ребром 1 сантиметр можно разрезать на любое количество частей сколь угодно малого размера, в том числе меньших по размерам, чем атомы; физически это ничуть не менее нелепо, чем изготовление двух картофелин из одной. Кроме того, и не вдаваясь в раздумья о том, до какой степени простирается сходство между картофелиной и ограниченным подмножеством пространства, можно отметить, что картошку мы обычно разрезаем на такие куски, где точки одного множества «соединены» друг с другом, а в наших множествах точки «разрозненные».

Давайте теперь попробуем разобраться, чем именно удивительна теорема Банаха—Тарского, не выходя за рамки математики. Видимо, основная причина недоумения следующая. Представляется «естественным», что всякое (по крайней мере ограниченное) подмножество пространства имеет объем. Если облечь это «естественное» представление в аккуратную математическую форму, то получится примерно следующее.

Мечта об объемах. *Всякому ограниченному множеству $X \subset \mathbb{R}^3$ можно сопоставить неотрицательное число $V(X)$, обладающее следующими тремя свойствами*:

1) *если $X \cong Y$, то $V(X) = V(Y)$*;
2) *если $X = Y \sqcup Z$, то $V(X) = V(Y) + V(Z)$*;
3) *если X — куб с ребром 1, то $V(X) = 1$.*

Если эта мечта осуществима, то равносоставленные множества обязаны иметь одинаковый объем: в самом деле, если $X = X_1 \sqcup \ldots \sqcup X_n$, $Y = Y_1 \sqcup \ldots \sqcup Y_n$ и $X_j \cong Y_j$ при $1 \leqslant j \leqslant n$, то

$$V(X) = V(X_1) + \ldots + V(X_n) = V(Y_1) + \ldots + V(Y_n) = V(Y).$$

Это, однако, теореме Банаха—Тарского противоречит: можно показать, что из условий 1—3 вытекают все известные из школьного курса формулы для объемов многогранников — а согласно Банаху и Тарскому любые два многогранника (в том числе и разных объемов) равносоставлены; кроме того, можно и не ссылаться на недоказанные утверждения, а попросту заметить, что из свойств (2) и (3) непосредственно вытекает, что дизъюнктное объединение двух кубов с ребром 1 должно иметь объем 2 — а согласно Банаху и Тарскому это множество равносоставлено с одним кубом с ребром 1!

Единственный логически возможный вывод из этих несообразностей таков.

Наша «мечта об объемах» неосуществима: *попытка приписать объем каждому ограниченному подмножеству пространства, не нарушая условий* 1—3*, ведет к противоречию.*

Таким образом, никакого «потрясения основ» не произошло, и называть теорему Банаха—Тарского парадоксом оснований нет: она всего-навсего свидетельствует о том, что не проходит наивная попытка формализовать понятие объема.

Из того, что наивная попытка формализации не удалась, не следует, что формализовать понятие объема вообще невозможно: это возможно, но действовать приходится менее прямолинейно. Именно, можно добиться выполнения свойств 1—3, если отказаться от идеи приписать объем *каждому* ограниченному множеству: сначала определяется класс «хороших» подмножеств пространства (этот класс очень широк, он включает в себя все подмножества $\mathbb{R}^3$, которые можно получить более или менее «явной» конструкцией — но все-таки не все ограниченные подмножества), а затем определяется объем для этих и только этих подмножеств, всем же прочим («плохим») приписать какой-либо объем мы даже не пытаемся. Мы не будем ничего говорить о том, как именно определяется класс «хороших» множеств и как определяется их объем: вы все это узнаете, когда будете изучать в университете курс «теории меры». Отметим только, что множества, возможность разбиения на которые утверждается теоремой Банаха—Тарского, являются, очевидно, «плохими» — по крайней мере некоторые из них.

Возможно, вам доводилось слышать и высказывания наподобие такого: «Неудивительно, что у Банаха и Тарского получился такой странный и нелепый результат — ведь в его доказательстве используется аксиома выбора». Прокомментируем их.

Аксиома выбора гласит, что для всякого семейства непустых множеств можно выбрать по элементу в каждом из них. Мы заведомо воспользовались ей при доказательстве предложения 15: когда мы разбивали каждую орбиту E_p на множества X_p, Y_p и Z_p, нужно было отождествить эту орбиту с множеством E, а для такого отождествления в ней надо было выбрать точку, «с которой начинать» (точка p на рис. 9).[1] С использованием этой аксиомы проводятся, в частности, очень «неявные» конструкции, которые с непривычки трудно «пощупать» и вообразить — наподобие наших множеств $\mathscr{X}$, $\mathscr{Y}$ и $\mathscr{Z}$.

Сто лет назад, когда теория множеств только создавалась и математики ее всячески осмысляли, аксиома выбора подверглась серьезному критическому анализу, были даже попытки развивать математику без ее участия, но на данный момент «мэйнстримная», если воспользоваться пришедшим в русский язык новым словечком, математика относится к аксиоме выбора с полным равнодушием — так же, как и ко всем прочим аксиомам теории множеств. Отголоски былого настороженного отношения к этой аксиоме можно до сих пор встретить в учебной и особенно популярной литературе.

Теперь поставим следующий вопрос: насколько принципиально то, что мы работали именно с пространственными телами? Может быть, и любые два плоских многоугольника равносоставлены? Существует ли «плоский» аналог теоремы Банаха—Тарского?

Теорема Банаха—Тарского, которую мы доказали, повергает в прах «мечту об объемах». Давайте сформулируем аналогичную мечту о площадях.

Мечта о площадях. *Всякому ограниченному множеству $X \subset \mathbb{R}^2$ можно сопоставить неотрицательное число $S(X)$, обладающее следующими тремя свойствами*:

1) *если $X \cong Y$, то $S(X) = S(Y)$;*

---

[1] Не исключено, что мы пользовались аксиомой выбора и еще где-то: сводить сколько-нибудь серьезные математические рассуждения к ссылкам на аксиомы — занятие не для человека, и точно проследить, какие аксиомы используются, а какие нет, не всегда просто.

2) *если* $X = Y \sqcup Z$, *то* $S(X) = S(Y) + S(Z)$;

3) *если* X — *квадрат со стороной* 1, *то* $S(X) = 1$.

Удивительным образом оказывается, что эта мечта как раз осуществима: есть способ сопоставить каждому ограниченному множеству $X \subset \mathbb{R}^2$ «площадь» $S(X)$, обладающую свойствами 1—3! Правда, этот способ не единственен, и для «плохих» множеств X их «площадь» $S(X)$ однозначно не определена, но для достаточно хороших множеств X (в частности, для всех многоугольников) число $S(X)$ всегда совпадает с площадью в школьном смысле. Из этого, в частности, следует, что два многоугольника разной площади равносоставлены быть не могут, так что никакого аналога теоремы Банаха—Тарского для плоскости нет. Напомним для полноты, что по теореме 8 два многоугольника равной площади, напротив, всегда равносоставлены.

В последнем разделе 5 мы постараемся объяснить, почему теорема Банаха—Тарского имеет место в $\mathbb{R}^3$ (и в пространствах большего числа измерений, коль на то пошло), но не на плоскости. Забегая вперед, скажем, что «виновато» в этом разное строение групп движений плоскости и пространства.

5. В пространстве можно, на плоскости нельзя

В этом заключительном разделе мы объясним, почему на плоскости аналога теоремы Банаха—Тарского нет. Многие доказательства будут только намечены или вовсе опущены, а познаний сверх школьной программы будет требоваться больше, чем в основной части книжки.

Для иллюстрации отсутствия теоремы Банаха—Тарского на плоскости мы покажем, что круг не равносоставлен объединению двух непересекающихся кругов. Именно, верен такой факт.

Предложение 21. *Пусть* D — *круг на плоскости. Тогда невозможно разбить* D *на попарно непересекающиеся подмножества* $X_1, \ldots, X_n$ *таким образом, чтобы множества, конгруэнтные некоторым* k *из них, образовывали разбиение круга* D, *а множества, конгруэнтные оставшимся* $n - k$ *из них, также образовывали разбиение круга* D.

Из этого предложения сразу вытекает, что круг D неравносоставлен дизъюнктному объединению кругов $D_1 \sqcup D_2$. В самом деле, рассуждая от противного, пусть D можно разбить на множества $X_1, \ldots, X_m$, из которых можно сложить $D_1 \sqcup D_2$. Разбив при необходимости какие-то из этих множеств на две части, можно считать, что каждое X_j конгруэнтно либо подмножеству в D_1, либо подмножеству в D_2; теперь ясно, что как из частей, конгруэнтных подмножеству в D_1, так и из частей, конгруэнтных подмножеству в D_2, можно сложить круг, и это противоречит предложению 21.

Прежде чем пытаться доказать предложение 21, разберем аналогичный, но более простой вопрос: невозможность удвоения окружности.

Предложение 22. *Пусть C — окружность с центром O. Тогда невозможно разбить C на попарно непересекающиеся подмножества $X_1, \ldots \ldots, X_n$ таким образом, чтобы некоторые k из них после подходящих поворотов относительно O образовывали разбиение окружности C и оставшиеся $n-k$ из них после подходящих поворотов относительно O также образовывали разбиение C.*

То, что мы разрешаем подмножества окружности только поворачивать, но не отражать, несущественно: если разрешить еще и отражать, удвоения все равно не получится, только доказательство будет чуть сложнее. Нам сейчас важно продемонстрировать общий принцип.

Начнем с совершенно элементарной задачи.

Задача о блохах. На плоскости в каждой точке с целыми координатами сидит по блохе. Все блохи одновременно совершают прыжки, длины которых в совокупности ограничены (скажем, не превосходят 10000), и оказываются в другой точке с целыми координатами (или в той же самой). Могут ли они прыгнуть так, чтобы в результате в каждой точке сидело по две или более блохи?

Ответ на этот вопрос «нет», и решение очень простое. Изначально в квадрате, состоящем из точек с координатами, не превосходящими по модулю натурального числа n, сидели $(2n+1)^2$ блох, а после прыжка там должно оказаться как минимум вдвое больше, то есть по меньшей мере $2(2n+1)^2$. С другой стороны, все эти блохи могли прийти только из квадрата, образованного точками с координатами, не превосходящими (опять по модулю) $n+10000$. Следовательно,

$$2(2n+1)^2 \leqslant (2(n+10000)+1)^2,$$

но при достаточно больших n это неравенство выполняться не может — противоречие.

Поучительно сравнить этот результат со следующей задачей.

Задача 5. В каждой вершине ab-графа (см. рис. 4) сидит блоха. Все блохи одновременно совершают прыжок и оказываются в другой (или той же самой) вершине, причем длина прыжка (т. е. длина пути по ребрам из старой вершины в новую) для каждой блохи не превосходит 10000. Могут ли они прыгнуть так, чтобы в результате в каждой вершине сидело по две или более блохи? (Подсказка: ответ положителен, и ограничение 10000 сильно завышено.)

С другой стороны, результат задачи о блохах обобщается на n-мерные решетки для любого n: точно так же, как и выше, доказывается, что если в каждой точке решетки $\mathbb{Z}^n$ сидит по блохе и все они одновременно прыгают на расстояние $\leqslant K$, где K — некоторая константа, то в результате в каждой точке не может оказаться по две или более блохи.

Теперь применим задачу о блохах к делу. Рассуждая от противного, предположим, что разбиение окружности, о котором идет речь в предложении 22, возможно. Пусть при этом подмножество X_k пришлось поворачивать на угол φ_k; обозначим для краткости поворот на угол φ_k через R_k. Пусть теперь Γ — множество всевозможных поворотов (относительно точки O), являющихся композициями поворотов R_k и обратных к ним (в любом количестве; о любом порядке в данном случае вопрос не стоит, поскольку порядок выполнения поворотов относительно одной и той же точки не важен[1]). Равносильно можно сказать, что Γ — это множество поворотов на всевозможные углы вида $a_1\varphi_1 + \ldots + a_n\varphi_n$, где $a_1, \ldots, a_n$ — целые числа. Ясно, что Γ — группа преобразований окружности C. Зафиксируем какую-нибудь точку $p \in C$ и обозначим через E орбиту точки p относительно группы Γ, то есть множество точек $g(p)$ для всех $g \in \Gamma$. Поскольку поворот, не являющийся тождественным преобразованием, неподвижных точек на C не имеет, соответствие $g \mapsto g(p)$ задает биекцию Γ и E. Посадим теперь на каждую точку орбиты E по блохе и заставим их одновременно прыгнуть следующим образом: если точка $q \in E$ лежит в множестве X_k, то блоха прыгает из q в $R_k(q)$. В результате этих прыжков в каждой точке орбиты E окажется, очевидно, по две блохи. Теперь заметим, что Γ — коммутативная группа, порожденная конечным числом своих элементов. Если нам повезло и Γ изоморфна $\mathbb{Z}^r$ для некоторого r, то мы оказываемся в точности в условиях задачи о блохах: множество E находится во взаимно однозначном соответствии с решеткой $\mathbb{Z}^r$, длины прыжков ограничены в совокупности, так как блохи совершают прыжки лишь на векторы, соответствующие поворотам $R_1, \ldots, R_n$, и задача о блохах говорит, что такая ситуация невозможна. В общем же случае можно применить классическую теорему о коммутативных группах, порожденных конечным числом элементов: они изоморфны или $\mathbb{Z}^r$, или прямой сумме $\mathbb{Z}^r$ и некоторой конечной группы. На нашем «блошином» языке это означает, что блохи сидят не на одном, а на нескольких одинаковых листах r-мерной клетчатой бумаги, а прыжки проходят так: сначала каждая блоха прыгает

[1] Если в некоторой группе композиция элементов не зависит от их порядка, такую группу называют коммутативной.

в пределах своего листа (длины прыжков ограничены), а затем она по желанию может перескочить в точку с теми же координатами на другом листе.

Задача 6. Покажите, что и в этом случае невозможно организовать прыжки так, чтобы на каждом листе в каждой вершине сидело по две или более блохи.

Итак, мы доказали невозможность «удвоения» окружности (предложение 22). Основную роль в этом доказательстве сыграло то, что группа поворотов плоскости коммутативна. Поэтому доказать предложение 21 буквально так же не удастся: ничего не стоит привести примеры движений плоскости, результат композиции которых зависит от порядка ее выполнения. Тем не менее оказывается, что группа движений плоскости хоть и не коммутативна, но в некотором смысле «близка» к коммутативности. Сейчас мы объясним, в каком именно.

Пусть G — группа движений плоскости. Сопоставим каждому движению число $+1$, если оно сохраняет ориентацию, и число -1, если оно ориентацию обращает.[1] Обозначим через H_1 группу, состоящую из чисел $+1$ и -1 (операция — умножение); эта группа, естественно, коммутативна. Отображение, ставящее в соответствие движению (указанным выше образом) число 1 или -1, обозначим $f_1\colon G \to H_1$; это отображение является гомоморфизмом (то есть образ произведения равен произведению образов).

Обозначим через G_1 «ядро» отображения f_1, то есть множество элементов группы G, переходящих при отображении f_1 в единицу. Группа G_1 состоит из движений, сохраняющих ориентацию, то есть из поворотов и параллельных переносов. Если каждому сохраняющему ориентацию движению плоскости сопоставить угол, на который оно поворачивает векторы, то при композиции движений эти углы складываются. Поэтому если обозначить группу углов (она же, если угодно, группа поворотов относительно начала координат) через H_2, то отображение $f_2\colon G_1 \to H_2$, сопоставляющее каждому движению угол, является гомоморфизмом; заметим, что группа H_2 коммутативна. Пусть, наконец, G_2 — ядро гомоморфизма f_2; эта группа состоит из движений, сохраняющих ориентацию и не поворачивающих векторы, то есть из параллельных переносов, и она уже коммутативна. Если группу можно подобным образом «развинтить» на коммутативные, ее называют разрешимой. Вот точное определение.

[1] Одно из возможных формальных определений: движение сохраняет ориентацию, если оно — композиция четного числа симметрий относительно прямых (т. е. поворот или параллельный перенос), и обращает в противном случае (т. е. если движение — симметрия или скользящая симметрия).

Определение 4. Пусть $G = G_0$ — группа. Если существует конечная последовательность коммутативных групп $H_1, \ldots, H_n$ и гомоморфизмов $f_1, \ldots, f_n$, в которой $f_j : G_{j-1} \to H_j$, G_j — ядро гомоморфизма f_j и группа G_n, являющаяся ядром f_n, уже коммутативна, то группу G называют *разрешимой*.

Естественно, согласно этому определению всякая коммутативная группа автоматически разрешима: можно взять $n = 1$, а в качестве гомоморфизма f_1 — гомоморфизм в группу, состоящую только из единичного элемента. Из обсуждения выше явствует, что группа движений плоскости разрешима. Группа же движений пространства разрешимой уже не является. В самом деле, эта группа содержит подгруппу, которую мы обозначили через $\mathrm{SO}(3, \mathbb{R})$ — группу поворотов пространства относительно осей, проходящих через данную точку O.

Задача 7. Докажите, что если $f : \mathrm{SO}(3, \mathbb{R}) \to H$ — гомоморфизм, то либо множество $f(\mathrm{SO}(3, \mathbb{R}))$, либо ядро гомоморфизма f состоит только из единичного элемента.[1]

Указание. Эта довольно трудная задача. Начать решение можно так. Покажите, что если R_1 и R_2 — повороты на один и тот же угол относительно разных осей, то найдется такой поворот S, что $R_2 = S^{-1} R_1 S$ («R_1 и R_2 сопряжены» — см. с. 23). Следовательно, если $f(R_1) = 1$, то и

$$f(R_2) = f(S^{-1} R_1 S) = f(S)^{-1} \circ 1 \circ f(S) = 1.$$

Обратите также внимание, что для получения поворота на любой угол хватает умения получить поворот на «сколь угодно малый» угол.

Задача 8. Выведите из результата предыдущей задачи, что группа движений пространства неразрешима.

Теперь сформулируем следующее свойство разрешимых групп.

Предложение 23. *Если G — разрешимая группа, то каждому подмножеству $X \subset G$ можно сопоставить неотрицательное число $\mu(X)$ таким образом, чтобы выполнялись следующие свойства*:

1) $\mu(X_1 \sqcup X_2) = \mu(X_1) + \mu(X_2)$;

2) *если подмножества $X_1 \subseteq G$ и $X_2 \subseteq G$ конгруэнтны, то* $\mu(X_1) = \mu(X_2)$;

3) $\mu(G) = 1$.

Здесь подмножества X_1 и X_2 называются конгруэнтными, если найдется такой элемент $g \in G$, что $X_2 = \{gh \mid h \in X_1\}$.

Отображение μ, существование которого утверждается в этом предложении, называется «конечно-аддитивной левоинвариантной вероят-

[1] Группы, обладающие таким свойством, называются простыми.

ностной мерой». Группы, на которых существует такая мера, называются «аменабельными».

Если группа G конечна, то легко видеть, что она аменабельна: достаточно в качестве $\mu(X)$ взять отношение количества элементов в множестве X к количеству элементов во всей группе. А вот с бесконечными группами все гораздо сложнее.

Задача 9. Проверьте, что группа G из раздела 2 не является аменабельной.

Задача 10. Из предыдущей задачи и предложения 23 вытекает, что группа G не является разрешимой. Можете ли вы установить это непосредственно?

Имея в своем распоряжении предложение 23, предложение 21 доказать уже нетрудно. Мы будем действовать в основном так же, как при доказательстве невозможности удвоения окружности, только в конце вместо задачи о блохах воспользуемся аменабельностью.

Итак, рассуждая от противного, пусть $D = X_1 \sqcup \ldots \sqcup X_n$ — разбиение диска, невозможность которого утверждается предложением. Пусть при этом множество X_j переводится в новое подмножество D с помощью движения F_j. Пусть Γ — множество, состоящее из тождественного преобразования и всевозможных композиций движений $F_1, \ldots, F_n$ и обратных к ним, в любом количестве и в любом порядке (иными словами, Γ — группа, порожденная движениями $F_1, \ldots, F_n$). Множество неподвижных точек у всякого нетождественного движения плоскости либо пусто, либо состоит из одной точки, либо, наконец, является прямой. Поскольку группа Γ, очевидно, счетна, в круге D есть точка p, не являющаяся неподвижной ни для одного нетождественного преобразования из Γ — это доказывается точно так же, как предложение 17 (годятся оба доказательства). Пусть E — орбита точки p; соответствие $g \mapsto g(p)$ задает биекцию между Γ и E. Если положить $Y_j = E \cap X_j$, то $E = Y_1 \sqcup \ldots \sqcup Y_n$; можно выбрать k из этих множеств так, что множества, конгруэнтные им относительно группы Γ, образуют разбиение E и множества, конгруэнтные оставшимся $n-k$ из них относительно группы Γ, также образуют разбиение E.

Заметим теперь, что так как Γ — подгруппа в разрешимой группе движений плоскости, она также разрешима и тем самым аменабельна. С помощью биекции $g \mapsto g(p)$ перенесем инвариантную меру с Γ на E. Теперь немедленно получаем противоречие: так как $E = Y_1 \sqcup \ldots \sqcup Y_n$, имеем $\mu(Y_1) + \ldots + \mu(Y_n) = \mu(E) = 1$; с другой стороны, если из множеств, конгруэнтных, скажем, $Y_1, \ldots, Y_k$, можно сложить E, то $\mu(Y_1) + \ldots$ $\ldots + \mu(Y_k) = 1$, а поскольку из множеств, конгруэнтных оставшимся

$Y_{k+1}, \ldots, Y_n$, также можно сложить E, то и $\mu(Y_{k+1}) + \ldots + \mu(Y_n) = 1$, откуда $\mu(Y_1) + \ldots + \mu(Y_n) = 2$ — противоречие.

Итак, нам осталось доказать предложение 23. Заметим для начала, что существование (левоинвариантной конечно-аддитивной вероятностной) меры μ равносильно возможности интегрирования для всех ограниченных функций: если указанная мера существует, то для каждой ограниченной функции f на Γ можно определить число $\int_\Gamma f(x)\,dx$ («интеграл»), обладающее следующими свойствами:

1) если f и g — ограниченные функции, а a и b — произвольные числа, то

$$\int_\Gamma (af(x) + bg(x))\,dx = a \cdot \int_\Gamma f(x)\,dx + b \cdot \int_\Gamma f(x)\,dx$$

(«линейность»);

2) если $f(x) \geqslant 0$ для всех x, то $\int_\Gamma f(x)\,dx \geqslant 0$ («положительность»);

3) для всякого $g \in \Gamma$ и всякой ограниченной функции f имеем

$$\int_\Gamma f(gx)\,dx = \int_\Gamma f(x)\,dx$$

(«инвариантность»);

4) $\int_\Gamma dx = 1$ («нормировка»).

Наоборот, если на Γ задан «интеграл» с такими свойствами, то можно определить и инвариантную меру по формуле $\mu(X) = \int_\Gamma \chi_X(x)\,dx$, где χ_X — характеристическая функция множества X, значение которой в точке x равно 1, если $x \in X$, и нулю в противном случае.

Для знатоков скажем, что интеграл по мере определяется с помощью той же конструкции, что интеграл Лебега. То, что мера всего лишь конечно аддитивна, ничему не мешает: счетная аддитивность нужна только для доказательства специфически «лебеговских» теорем о предельном переходе под знаком интеграла, наподобие теоремы Беппо Леви о монотонной сходимости или теоремы Лебега об ограниченной сходимости.

Имея в виду сказанное выше, наметим доказательство предложения 23.

Сначала покажем, как все сводится к случаю коммутативной группы G.

Предложение 24. *Пусть $f\colon G \to H$ — сюръективный гомоморфизм с ядром G_1. Если группы G_1 и H аменабельны, то G также аменабельна.*

В самом деле, пусть $X \subset G$ — произвольное подмножество. Если ν — инвариантная мера на G_1, то определим функцию φ_X на H следующим образом. Для всякого $h \in H$ рассмотрим множество $f^{-1}(h) \cap X$; если это множество пусто, положим $\varphi_X(h)=0$, в противном случае выберем какой-нибудь $g\in f^{-1}(h)$ и положим

$$\varphi_X(h)=\nu(g^{-1}(f^{-1}(h)\cap X));$$

множество в правой части — подмножество в G_1, а от выбора $g\in f^{-1}(h)$ число в правой части не зависит, поскольку мера ν на группе G_1 инвариантна. Если теперь положить

$$\mu(X)=\int_H \varphi_X(h)\,dh,$$

то нетрудно проверить, что так определенная функция μ является инвариантной мерой на G.

Из предложения 24 вытекает, что нам достаточно доказать аменабельность коммутативных групп. Дело это непростое даже для такой бесхитростной коммутативной группы, как группа целых чисел $\mathbb{Z}$. Можно, конечно, попытаться, по аналогии с конечными группами, взять в качестве меры подмножества $\mathbb{Z}$ его «плотность», положив

$$\mu(x)=\lim_{n\to\infty}\frac{\text{число элементов в }\{m: m\in X, |m|\leqslant n\}}{2n+1}. \qquad (*)$$

Так, однако же, не выходит:

Задача 11. Приведите пример подмножества $X\subset\mathbb{Z}$, для которого предел в правой части $(*)$ не существует.

На самом деле строить инвариантные меры на коммутативных группах приходится гораздо менее конструктивным способом. Собственно, мы и не будем строить меру — мы всего лишь докажем ее существование.

Итак, пусть G — коммутативная группа. Мы будем сразу строить интеграл по инвариантной мере. Для этого обозначим через $L(G)$ векторное пространство, состоящее из всех ограниченных функций на G. Для $f_1, f_2\in L(G)$ положим $\rho(f_1,f_2)=\sup_{x\in G}|f_1(x)-f_2(x)|$; очевидно, число $\rho(f_1,f_2)$ обладает всеми свойствами расстояния (это то, что называется «метрикой» на $L(G)$). Пусть

$$V=\{f\in L(G)\mid \inf_{x\in G} f(x)>0\}.$$

Через K обозначим векторное подпространство в $L(G)$, порожденное функциями вида $x\mapsto f(x)-f(gx)$ для всевозможных $f\in L(G)$, $g\in G$. Легко видеть, что пространство K состоит из функций вида

$$x\mapsto(f_1(x)-f_1(g_1x))+(f_2(x)-f_2(g_2x))+\ldots+(f_k(x)-f_k(g_kx)),$$

где $f_1,\ldots,f_k\in L(G)$, $g_1,\ldots,g_k\in G$, k произвольно.

Нетрудно проверить, что интеграл по инвариантной мере на V — то же самое, что линейное отображение $I\colon L(G)\to\mathbb{R}$, обладающее следующими свойствами:

1) $I(\varphi)>0$, если $\varphi\in V$;

2) $I(\varphi)=0$, если $\varphi\in K$;

3) $I(1)=1$, где 1 — функция, тождественно равная единице.

Если выполнены условия (1) и (2), то выполнения условия (3) можно добиться, умножив I на подходящую константу, так что про это условие можно вообще забыть. А теперь докажем такую лемму (именно в ней используется коммутативность группы G!).

Лемма 25. *Пусть G — коммутативная группа, $f_1,\dots,f_k$ — ограниченные функции на G, и пусть $g_1,\dots,g_k\in G$. Тогда*

$$\inf_{x\in G}((f_1(x)-f_1(g_1x))+(f_2(x)-f_2(g_2x))+\ldots+(f_k(x)-f_k(g_kx)))\leqslant 0.$$

(Заметим, что если бы эта лемма была неверна, то интеграла по инвариантной мере заведомо не существовало бы: условия (1) и (2) противоречили бы друг другу.)

Для простоты приведем доказательство леммы для случая $k=2$ (общий случай отличается только более громоздкой записью). Итак, пусть даны ограниченные функции f_1 и f_2 и элементы $g_1,g_2\in G$. Положим

$$F(x)=f_1(x)-f_1(g_1x)+f_2(x)-f_2(g_2x);$$

нам надо доказать, что $\inf F(x)\leqslant 0$. Для этого выберем натуральное число n и просуммируем значения функции F в точках вида $g_1^ig_2^j$, где $0\leqslant i<n$, $0\leqslant j<n$. Обозначим эту сумму через S. Имеем

$$S=\sum_{\substack{0\leqslant i<n\\0\leqslant j<n}}(f_1(g_1^ig_2^j)-f_1(g_1^{i+1}g_2^j)+f_2(g_1^ig_2^j)-f_2(g_1^ig_2^{j+1}));\tag{5.1}$$

после раскрытия скобок в (5.1) сокращаются все слагаемые, в которых $0<i<n-1$ и $0<j<n-1$, так что S равно сумме $8n$ слагаемых вида $f_1(g_1^j)$, $f_2(g_1^j)$, $f_1(g_1^jg_2^n)$, $f_2(g_1^jg_2^n),\dots$, $f_2(g_1^ng_2^j)$ (с какими-то знаками, но это уже неважно). Если функции f и g ограничены по модулю числом M, то отсюда вытекает, что $S\leqslant 8nM$. С другой стороны, $S\geqslant n^2\inf_{x\in G}F(x)$, так как S — сумма значений функции F в n^2 точках. Следовательно,

$$n^2\inf_{x\in G}F(x)\leqslant 8nM;$$

деля на n^2 и устремляя n к бесконечности, получаем, что $\inf F(x)\leqslant 0$, и лемма доказана.

Наша лемма показывает, что $K \cap V = \varnothing$. Поскольку множество V, очевидно, открыто и выпукло, существование линейного отображения I вытекает из следующего классического результата, называемого теоремой Хана—Банаха.

Теорема 26. *Пусть V — открытое выпуклое подмножество в нормированном*[1] *векторном пространстве L, и пусть $K \subset L$ — векторное подпространство, не пересекающееся с V. Тогда существует линейное отображение $I\colon V \to \mathbb{R}$, положительное на V и тождественно равное нулю на K.*

Объясним в заключение, как доказывается эта теорема. Во-первых, все легко сводится (с помощью перехода к факторпространству) к случаю, когда подпространство K состоит только из нуля. Разберем теперь простейший случай: $L = \mathbb{R}^2$, $V \subset L$ — открытое выпуклое подмножество, не содержащее начала координат O. Нам достаточно построить прямую ℓ, проходящую через O и не пересекающуюся с V. Для этого рассмотрим всевозможные лучи, выходящие из O. Если каждый из них имеет непустое пересечение с V, то получаем противоречие: если P и Q — точки из V, лежащие на двух лучах, принадлежащих одной прямой, то точка O лежит на отрезке PQ, и в силу выпуклости V получаем, что O принадлежит V. Значит, найдется луч λ, выходящий из O и не пересекающийся с V. Пусть m — перпендикуляр к λ, проходящий через O; рассмотрим два прямых угла, образованных λ и m. Оба этих угла одновременно иметь непустое пересечение с множеством V не могут: если точки $P, Q \in V$ лежат в этих двух углах, то отрезок PQ пересекается с лучом λ, и точка пересечения обязана лежать в V — противоречие (рис. 10а); стало быть, существует (открытый) прямой угол с вершиной в O, не пересекающийся с V.

Теперь будем «расширять плацдарм». Пусть μ — биссектриса указанного прямого угла и n — перпендикуляр к μ, проходящий через O. Стороны прямого угла и прямая n образуют два угла величиной $\pi/4$; оба этих угла не могут иметь непустое пересечение с V по тем же причинам, что и выше (рис. 10б); присоединив тот из них, что не пересекается с V, к нашему прямому углу, получаем угол величиной $3\pi/4$, не пересекающийся с V; проведя перпендикуляр к его биссектрисе, находим не пересекающийся с V угол величиной $7\pi/8$ и т. д. Объединение всех этих углов — полуплоскость, не пересекающаяся с V.

Переход от этого частного случая к общему осуществляется с помощью стандартной математической техники. Сначала выбирается про-

[1] Нормированное векторное пространство — это пространство, в котором задана метрика со свойствами, аналогичными свойствам нашей метрики ρ на пространстве $L(G)$; подробности вы узнаете в курсе анализа.

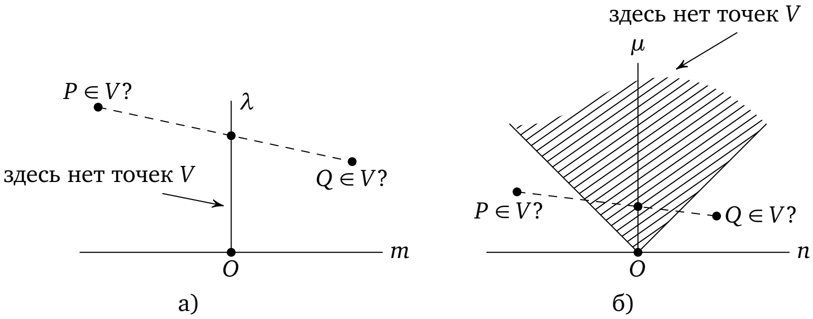


Рис. 10. Теорема Хана—Банаха для плоскости

извольное двумерное подпространство в L и находится в нем прямая, не пересекающаяся с V; затем показывается, что если уже есть подпространство $L' \subset L$ и в нем «гиперплоскость» (т. е. ядро линейного отображения в $\mathbb{R}$) H', не пересекающаяся с V, то можно найти большее подпространство $L'' \supset L$ и в нем гиперплоскость $H'' \supset H'$, не пересекающуюся с V (это делается сведением к рассмотренному нами двумерному случаю), а затем, «переходя к пределу», делают аналогичный вывод и для всего пространства V. Обоснование такого предельного перехода проводится с помощью так называемой леммы Цорна; для любопытствующих отметим, что лемма Цорна также основывается на аксиоме выбора.

Содержание