

МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ

Третья серия

ВЫПУСК 3

МЦНМО

ЧеРо

Москва 1999

Редакционная коллегия

Бугаенко В.О.	Васильев Н.Б.	Винберг Э.Б.
Вялый М.Н.	Глейзер Г.Д.	Гусейн-Заде С.М.
Егоров А.А.	Ильяшенко Ю.С.	Канель-Белов А.Я.
Константинов Н.Н.	Прасолов В.В.	Розов Н.Х.
Савин А.П.	Соловьев Ю.П.	Сосинский А.Б.
Тихомиров В.М.	Шарыгин И.Ф.	Яценко И.В.

ГЛАВНЫЙ РЕДАКТОР: В. М. Тихомиров

ОТВ. СЕКРЕТАРЬ: М. Н. Вялый

АДРЕС РЕДАКЦИИ:

121002, Москва, Б. Власьевский пер., д. 11, к. 211
(с пометкой «Математическое просвещение»)

EMAIL: matpros@mccme.ru

М34 Математическое просвещение. Третья серия, вып. 3. — М.: МЦНМО, «ЧеРо», 1999. — 240 с.

Очередной сборник «Математическое просвещение» содержит материалы, посвященные памяти Н. Б. Васильева и Е. М. Ландиса; рассказ В. И. Арнольда о работах филдсовского лауреата 1998 года М. Л. Концевича; очерк о роли Ф. Клейна и его знаменитой эрлангенской программы в истории математики. В разделе, посвященном проблемам современной математики, помещен цикл статей о теории узлов. В разделе «Студенческие чтения» опубликована лекция А. А. Разборова, в которой популярно излагаются основы теории сложности вычислений. Помимо этого, сборник содержит ряд статей, посвященных ярким и интересным математическим фактам и миниатюрам: доказательству теоремы Жордана, моделям неевклидовых геометрий, характеристизации эллипсоидов, решению проблемы Борсука, связи основной теоремы алгебры с теорией непрерывных групп. В разделе «Конкурсы и олимпиады» обсуждаются наиболее интересные задачи, предлагавшиеся на математических соревнованиях в Москве в 1998 году.

ISBN 5-900916-29-4

©МЦНМО, 1999 г.

Выпуск данного сборника поддержан грантом
Российского Фонда Фундаментальных Исследований
(номер проекта 99-01-14033)

*Посвящается памяти
Николая Борисовича Васильева,
замечательного человека
и поистине выдающегося
деятеля математического просвещения.*

*Кроме всего прочего,
именно благодаря его энтузиазму
и настойчивости было возобновлено
издание серии сборников, один из номеров
которой вы держите сейчас в руках.*

Николай Борисович Васильев

(8.8.1940 – 28.5.1998)

СОДЕРЖАНИЕ

Математический мир

В. И. Арнольд <i>Филдсовская медаль — воспитаннику московской математической школы</i>	7
А. А. Егоров <i>Николай Борисович Васильев</i>	21
А. М. Леонтович <i>Научная деятельность Н. Б. Васильева</i>	29
М. Л. Гервер <i>Памяти Е. М. Ландиса</i>	34
<i>Из выступлений на заседании Московского математического общества, посвященном памяти Е. М. Ландиса</i>	43
Н. Х. Розов <i>Феликс Клейн и его эрлангенская программа</i>	49

Тема номера: теория узлов в конце XX века

С. В. Дужин, С. В. Чмутов <i>Узлы и их инварианты</i>	59
Ю. М. Бурман <i>Длинные кривые, гауссовы диаграммы и инварианты</i>	94
В. В. Прасолов <i>Поверхность Зейферта</i>	116

Студенческие чтения

А. А. Разборов <i>О сложности вычислений</i>	127
---	-----

По-новому о старом: фрагменты классической математики

А. В. Чернавский <i>Теорема Жордана</i>	142
Р. Р. Пименов <i>Образжения сферы и неевклидовы геометрии</i>	158

Наш семинар: математические сюжеты

М. Л. Гервер <i>О разбиении множеств на части меньшего диаметра: теоремы и контрпримеры</i> . . .	168
А. Б. Скопенков <i>n-мерный куб, многогрены и решение проблемы Борсука</i>	184
П. А. Бородин, В. М. Тихомиров <i>Критерии гильбертовости банахова пространства, связанные с теорией приближений</i>	189
Г. Райтер, И. Сонин <i>Наглядная интерпретация некоторых алгоритмов на графах</i>	208

Конкурсы и олимпиады

В. О. Бугаенко <i>Избранные задачи математических соревнований 1998 года в Москве</i>	214
--	-----

Отклики

Б. Р. Френкин <i>Связь основной теоремы алгебры с теорией непрерывных групп</i>	229
--	-----

Задачный раздел	232
----------------------------------	-----

Новые издания	234
--------------------------------	-----

Математический мир

Филдсовская медаль — воспитаннику московской математической школы

В. И. Арнольд

Во время церемонии открытия Международного Математического Конгресса, состоявшегося в Берлине с 18 по 27 августа 1998 года, были оглашены имена 4 лауреатов медали Филдса. Один из них — воспитанник Московского Университета Максим Львович Концевич (сейчас работающий в Институте Высших Научных Исследований в Бюр-сюр-Иветт под Парижем).

Филдсовскими эти медали названы в честь канадского математика Дж. Ч. Филдса (1863 – 1962 гг.), бывшего президентом Математического Конгресса в Торонто в 1924 году. Филдс предложил использовать сэкономленные организаторами Конгресса 2500 канадских долларов для того, чтобы во время каждого Конгресса (которые проводятся раз в четыре года) награждать двух математиков золотой медалью в знак признания их выдающихся заслуг мировым математическим сообществом. При этом заслугой считается как решение давно стоявшей проблемы, так и создание теории, пролагающей новые пути. Формулируя принципы, по которым будут отбираться лауреаты, Филдс подчеркнул, что медаль должна быть не только признанием уже достигнутого (что, конечно, играет основную роль при выборе медалистов), но и залогом будущих достижений (как самого медалиста, так и его последователей).

Хотя подобные слова о будущих успехах были и в завещании Нобеля, их трактовка нобелевским и филдсовским комитетами различна. Филдсовские медали присуждаются только молодым математикам (от старых ожидать новых достижений рискованно). Начиная с московского Конгресса 1966 года, термин «молодой математик» точно определен как «математик, которому в год присуждения исполняется не более сорока лет». М. Концевич (см. фото 1) родился 25 августа 1964 года.

Следует заметить, что используемые в разных обстоятельствах определения молодого математика сильно различаются. Например, Московское математическое общество ежегодно присуждает две чрезвычайно престижных премии молодым математикам, которым исполняется в год присуждения не более тридцати лет. Видимо, такого же принципа придерживается Европейское математическое

*Рис. 1. М. Л. Концевич**Рис. 2. Филдсовская медаль*

общество (премии которого Концевич был удостоен на первом Европейском Математическом Конгрессе в Париже в 1992 году).

Группа французских математиков, объединенных псевдонимом Бурбаки, считает математика молодым до пятидесяти лет (если он не кокотизирован раньше). Кокотизация¹⁾ состоит в том, что математика, молодость которого вызывает сомнения, заставляют выслушать, в присутствии предупрежденных коллег, длиннейшее определение нового математического понятия, составленное так, что ничто, кроме нуля, этому определению не удовлетворяет. Если испытуемый вскричит: «Но ведь это ноль!» — то он спасен, если нет — кокотизирован.

Р. Курант, знаменитый немецкий (впоследствии американский) математик ввел свое определение молодого математика: математик считается молодым, если он в самой неподходящей обстановке говорит о математике. По этому определению Курант оставался молодым до самой смерти (в возрасте восьмидесяти четырех лет).

После смерти Филдса его завещание увеличило предназначенный для награждения фонд, и медаль (которую впервые присудили на Конгрессе в Осло в 1936 году) стала сопровождаться премией. Первоначально размер премии составлял 1 500 канадских долларов. Начиная с московского Конгресса 1966 г., благодаря анонимному вкладу в Филдсовский фонд, число присуждаемых на

¹⁾ Название, видимо, происходит от древнего обычая полинезийских племен испытывать стариков, способны ли они еще приносить пользу племени: старик должен достать кокосовый орех с вершины кокосовой пальмы, которую все племя при этом трясет...

каждом Конгрессе медалей увеличено до четырех и размер премии сейчас составляет уже 15 000 канадских долларов.

Из советских математиков Филдсовской медали были удостоены С. П. Новиков (1970), Г. А. Маргулис (1978), В. Г. Дринфельд (1990), Е. И. Зельманов (1994).

Филдсовский комитет назначается Исполнительным Комитетом Международного Математического Союза, избираемым Генеральной Ассамблеей Союза раз в четыре года. До 1994 года президент Союза был и председателем Филдсовского комитета.

Филдсовская медаль изготавливается из золота (14 карат) по проекту скульптора Р. Т. Маккензи (см. фото 2). На лицевой стороне медали имеется (приписываемая Архимеду) надпись на латыни “*Transire suum pectus mundoque poltri*” («превзойти свою человеческую ограниченность и покорить вселенную») и изображение Архимеда, а на обороте надпись — “*Congregati ex toto orbe mathematici ob scripta insignia tribuere*” («математики, собравшиеся со всего света, чувствуют замечательный вклад в познания»). Имя Филдса нигде не упомянуто (Филдс возражал бы и против названия медали по его имени — оно принято уже после его смерти).

Концевич является ярким представителем московской математической школы. Попытаюсь дать хотя бы некоторое представление о математических достижениях М. Концевича. Большая часть этих достижений относится к молодой области, которую можно назвать *квантовой математикой*.

Квантовая математика возникла под влиянием квантовой физики. В большинстве случаев речь идет здесь не о применениях математических результатов к исследованию физических явлений (хотя это и не исключается), а об обратном влиянии физических идей на фундаментальнейшие области математики.

Обычным признаком квантовости является присутствие в теории одного или нескольких параметров. В большинстве случаев параметр один, он обозначается буквой h (и часто называется постоянной Планка). При стремлении параметра к нулю квантовая математика вырождается в обычную. Удивительным выводом из исследований последних десятилетий является тот факт, что более или менее вся математика допускает согласованную деформацию, после которой все математические понятия приобретают совершенно новый смысл, но соотношения между ними (с соответствующими изменениями) сохраняются и в новой, продеформированной теории.

Начиная с обычных формул и производя последовательное квантование, математики пришли к массе новых формул, догадаться о которых иным путем вряд ли возможно.

Таким образом, большая часть квантовой математики является далеко идущим развитием теории возмущений. Со времен работ Р. Фейнмана стало ясно, что фундаментальным объектом всякой теории возмущений являются графы (в простейшем случае — так называемые диаграммы Фейнмана). На языке графов алгебра становится геометрией, что открывает новые возможности.

Простейшими примерами графов являются многоугольники или квадрат с его двумя диагоналями. Граф (см. рис. 3), по определению, состоит из вершин (точек) и ребер (соединяющих эти точки отрезков); число тех и других обычно конечно.

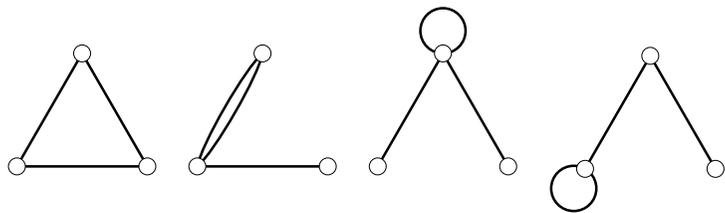


Рис. 3. Связные графы с тремя вершинами и тремя ребрами

Важнейшим достижением М. Концевича является введение им в математику фундаментального нового объекта — комплекса графов. Комплексами в математике называются многомерные обобщения графов. Подобно детскому конструктору, комплекс состоит не только из вершин и ребер, но включает также двумерные, трехмерные и т. д. элементы.

Математики называют *точками* (отрезками и т. д.) объекты любой природы (по словам Гильберта — «например, пивные кружки и лавки»), важны лишь соотношения между ними.

«Точками» комплекса графов Концевича являются сами всевозможные графы (с математической педантичностью можно упомянуть, что отрезки, соединяющие вершину с самой собой, здесь не допускаются, хотя в математической общей теории графов такие петли и рассматриваются).

Определение «отрезков», соединяющих такие «точки» (графы), «граней», соединяющих «отрезки», и т. д. вносит в множество всевозможных графов чрезвычайно важную (и ранее не изученную!) комбинаторную структуру — структуру комплекса Концевича, оказавшуюся очень полезной в большом числе задач, на первый взгляд не имеющих к графам и комбинаторике никакого отношения.

Перечислю лишь некоторые задачи, которые Концевичу удалось решить на этом пути. Первая задача относится к классической теории узлов. Узлы²⁾ (математики обычно завязывают их на *замкнутой* веревке (см. рис. 4), рассматривая всевозможные вложения окружности в трехмерное пространство) начали систематически изучаться в середине 19-го века, но в искусстве встречались задолго до этого (см. фото 5).

Теория узлов возникла в связи с предположением Кельвина, что различия атомов определяются различиями входящих в их состав микроскопических узлов. Это был явный шаг вперед по сравнению с древней теорией Лукреция, который объяснял различия атомов тем, что математики теперь описывают как различия между тремя типами поверхностей — гладких, комбинаторных (имеющих узлы) и топологических (допускающих, по выражению Лукреция, также и заостренные крючки).

Для различения узлов, которые нельзя перевести друг в друга непрерывной деформацией без разрезания веревки, математики изобрели множество разнообразных инвариантов. Инварианты узлов — это *локально постоянные функции*

²⁾ Теория узлов — тема этого номера «Математического просвещения». См. с. 56 – 126. — *Прим. ред.*

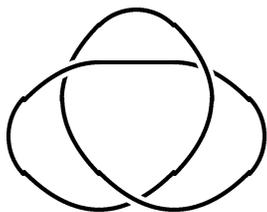


Рис. 4. Простейший узел

Фото 5. Узел на саламандре — тотеме французского короля Франциска I (замок Азай ле Ридо)

на пространстве вложений окружности в трехмерное пространство. Хотя пространство вложений и является бесконечномерным, исследование функций на нем оказывается достаточно интересной и не вполне безнадежной математической задачей.

В 1989 году московский математик В. А. Васильев выделил среди всех инвариантов узлов *инварианты конечного порядка*, занимающие в теории инвариантов узлов такое же фундаментальное место, какое занимают многочлены в обычной теории функций. Теперь эти инварианты называются *инвариантами Васильева*, и им посвящено множество глубоких работ во всем мире. Между прочим, филдсовский комитет Конгресса в Цюрихе (из которого я вышел как бывший научный руководитель своего ученика, являвшегося кандидатом на медаль) почему-то пропустил Васильева при выборе медалистов, а в 1998 году Васильеву уже было больше сорока лет.

Технические достижения Концевича позволили ему дать явное выражение для всех инвариантов Васильева в виде кратных интегралов по соответствующим конфигурационным пространствам. Эти формулы являются грандиозным обобщением электродинамической формулы Гаусса, выражающей коэффициенты зацепления (см. рис. 6) двух замкнутых кривых в трехмерном пространстве в виде двойного интеграла (коэффициент зацепления определяет (алгебраическое) число прохождений одной кривой сквозь другую при их растаскивании в удаленные части пространства).

Между прочим, в этой теории до сих пор остается открытым острейший вопрос: позволяют ли найденные инварианты различить *любые* два узла (недеформируемые друг в друга без самопересечений)?

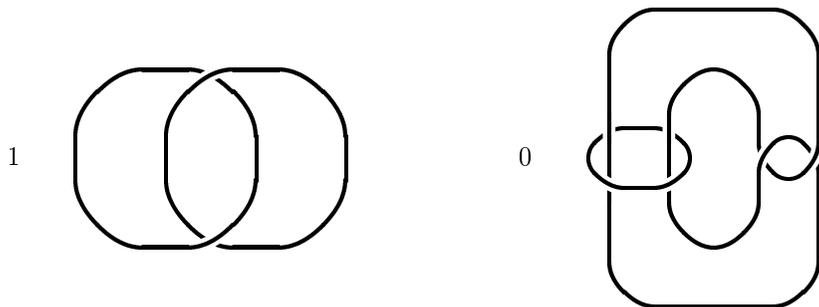


Рис. 6. Зацепления двух кривых с коэффициентами зацепления один и нуль

Официально научным руководителем диссертации Концевича считается работающий в Бонне Дон Загир, но сам Концевич назвал своими учителями (при награждении его премией Европейского математического общества молодым математикам на первом Европейском Математическом Конгрессе 1992 года в Париже) И. М. Гельфанда, Ю. И. Манина, В. И. Арнольда. Позже Концевич подчеркивал, как многому он научился также у более молодых московских математиков: А. Гончарова, А. Бейлинсона, В. А. Гинзбурга, Б. Фейгина, М. Капранова и у многих других.

Однажды в Математическом Институте Макса Планка в Бонне директор Института, профессор Ф. Хирцебрух (иностраный член РАН и один из самых знаменитых математиков нашего времени), спросил меня: почему это Концевич до сих пор не защитил диссертации? Давайте завтра устроим защиту — Загир, я и Вы составим жюри, все мы хорошо знаем его замечательные работы и напишем отзывы. Но через два дня после защиты Максим говорит: знаете, а работа-то об узлах была неверная! Только что из Гарварда пришел контрпример Бар Натана, он прав — моя интегральная формула для инвариантов Васильева не может быть верной! Исправление ошибки заняло еще несколько дней; исправленная формула верна, доказывается почти так же, как и исходная неверная (которая просто решает немного другую задачу). В условиях Бонна все эти волнения не имели никакого практического значения для оценки работ Концевича: всем было ясно, что они замечательны и что было бы нелепо требовать от него тратить время на оформление диссертации и нудные бюрократические процедуры.

Я всегда вспоминаю эту защиту, когда сталкиваюсь с нашими московскими порядками. Теперь, когда все московские зарплаты стали неразличимыми бесконечно малыми, многие талантливейшие молодые математики предпочитают не тратить время на бюрократические процедуры оформления и защиты диссертаций и остаются неостепененными; высвободившееся время они используют для новых открытий. Именно к этой категории молодых московских математиков относится и Максим Концевич.

Поразительно, что идеи одного происхождения позволяют математикам решать совершенно не сходные между собой на вид задачи.

Построенная Концевичем теория позволила ему дать явные интегральные формулы не только для инвариантов Васильева узлов, но и для квантования

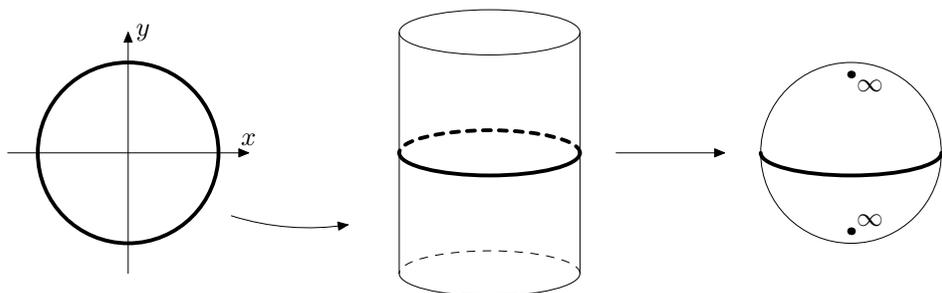


Рис. 7. Вещественная и комплексная окружность

скобок Пуассона. Члены рядов, выражающих такое квантование, нумеруются графами. Коэффициент, с которым в ряд входит данный член, выражается интегралом по конфигурационному пространству, составленному из вложений множества вершин графа в плоскость Лобачевского. Подинтегральное выражение, которое нужно интегрировать, поразительно похоже на подинтегральное выражение, доставляющее инварианты Васильева узлов, и интегрировать в обоих случаях приходится по конфигурационному пространству.

Задача о нахождении квантования скобок Пуассона (т. е. классической механики) — одна из старейших задач квантовой математики, которая возникла — в слегка другой формулировке — вместе с квантовой механикой (т. е. уж никак не позже тридцатых годов). При этом до работы Концевича не только не были известны найденные им явные формулы универсального квантования, но подвергалось сомнению и само его существование.

Быть может, наиболее яркие достижения современной математики связаны с идеями квантовой теории поля, пронизывающими столь разные области науки, как теория чисел, теория четырехмерных многообразий, алгебраическая геометрия и симплектическая топология (т. е. топология фазовых пространств классической механики).

Концевичу принадлежат и здесь фундаментальные результаты, позволившие решить трудные задачи. Уже в своей диссертации он доказал поразительную гипотезу американского физика Э. Виттена, связавшую топологические характеристики пространств модулей алгебраических кривых с отмеченными точками со знаменитым уравнением Кортевега — де Фриза теории мелкой воды.

Алгебраические кривые — это кривые, задаваемые алгебраическими уравнениями, вроде окружности, задаваемой уравнением $x^2 + y^2 = c$ (рис. 7). Комплексные векторы, удовлетворяющие подобному уравнению, образуют в плоскости двух комплексных переменных (x, y) (т. е. в вещественно четырехмерном пространстве) вещественно двумерную *риманову поверхность*. Например, «комплексная окружность», определенная тем же уравнением, что и вещественная, но с комплексными x и y , представляет собой с топологической точки зрения цилиндр. Добавляя две «бесконечно-удаленные» точки, математики получают риманову поверхность окружности, топологически эквивалентную обычной сфере.

Более сложные уравнения приводят к более сложным поверхностям. Например, уравнение третьей степени $y^2 = x^3 + ax + b$ приводит к поверхности тора —

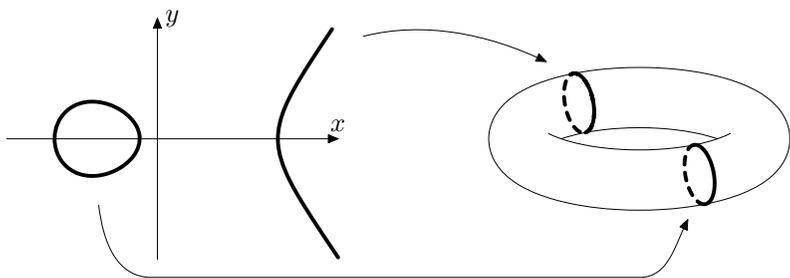


Рис. 8. Вещественная и комплексная эллиптическая кривая

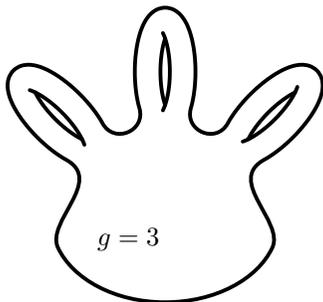


Рис. 9. Поверхность рода три (сфера с тремя ручками)

так называемой *эллиптической кривой*, геометрия которой играет решающую роль в недавнем решении знаменитой проблемы Ферма, данном английским математиком А. Уайлсом (рис. 8).

При изменении коэффициентов уравнения кривая может вырождаться. Например, тор, заданный приведенным выше уравнением третьей степени, вырождается в сферу при $4a^3 - 27b^2 = 0$ (когда многочлен в правой части уравнения имеет кратные корни).

Все комплексные кривые (вещественные поверхности) с топологией сферы одинаковы (переводятся друг в друга взаимнооднозначными комплексно-гладкими отображениями). Но торов (а также и более сложных кривых рода g , получаемых из сферы приклеиванием g ручек; для тора $g = 1$) много (рис. 9). Все кривые рода g топологически одинаковы, но они различаются комплексными структурами: одну можно отобразить гладко на другую в смысле вещественной геометрии, но, вообще говоря, не в смысле комплексной.

Пространство существенно различных кривых рода g имеет вещественную размерность $6g - 6$ при $g > 1$ и 2 при $g = 1$. Это пространство имеет естественную комплексную структуру, и его комплексная размерность равна $3g - 3$ при $g > 1$ и 1 при $g = 1$, 0 при $g = 0$.

Пространство классов эквивалентных кривых называется *пространством модулей* кривых. Это пространство не компактно: при изменении кривой она может вырождаться (в этот момент род g уменьшается, например, тор может выродиться в сферу). Концевичу принадлежит важнейшее понятие стабильного

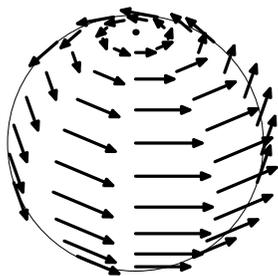


Рис. 10. Векторное поле с двумя особыми точками на сфере

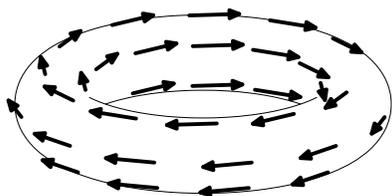


Рис. 11. Векторное поле без особых точек на торе

отображения комплексной кривой с отмеченными точками в комплексное пространство, позволившее правильно компактифицировать пространство отображений (путем добавления «бесконечно-удаленных» вырожденных кривых) и затем применить к полученному компактному (хотя и не совсем гладкому) комплексному многообразию могучие методы современной топологии.

Простейший пример здесь — пространство сфер с четырьмя (упорядоченными) отмеченными точками. Это пространство естественно отождествляется со сферой без трех точек (которыми можно считать $0, 1$ и ∞ на сфере Римана). Действительно, первые три точки можно поместить в $0, 1$ и ∞ соответственно при помощи подходящего дробно-линейного преобразования сферы Римана (переводящего комплексное число z в $(az + b)/(cz + d)$). Четвертая точка определяет тогда класс сферы с отмеченными точками. При компактификации сферы без трех точек получается просто обычная сфера.

В классической топологии гладкие многообразия классифицируются при помощи их инвариантов — характеристических чисел. Простейшим инвариантом является число Эйлера (или эйлерова характеристика), измеряющая препятствия к одновременной нигде не исчезающей деформации многообразия. Деформация задается векторным полем на многообразии (рис. 10), сопоставляющим каждой точке скорость ее движения. При попытке построить ненулевое поле на сфере (или на любой поверхности рода большего единицы) мы обнаруживаем, что поле без особых точек построить не удастся. Более того, алгебраическое (считаемое со знаками) число особых точек оказывается не зависящим от поля, но зависящим лишь от многообразия. Например, на поверхности сферы с g ручками это число равно $2 - 2g$. Для тора ($g = 1$) это число равно нулю. И действительно, тор можно сдвигать по себе без неподвижных точек (рис. 11). На сфере число особых точек любого поля (считаемых со знаками) равно двум. Поле ровно с двумя особыми точками изображено на рис. 10.

Число особых точек поля (считаемых со знаками) называется характеристическим именно потому, что оно не зависит от поля и является характеристикой многообразия. Кроме эйлеровой характеристики, имеются и другие характеристические числа. Они определяются более или менее таким же способом (грубо говоря, полей берется несколько и особые точки определяются несколько иначе).

В описанном выше простейшем случае многообразия модулей сфер с четырьмя отмеченными точками единственным характеристическим числом является эйлерова характеристика (равная двум), единственным полем — касательное сфере поле, характеризующее скорость движения четвертой точки (при всевозможных ее положениях).

Если вместо четырех точек взять больше или вместо сферы взять поверхность большего рода, то возникнет много характеристических чисел, обобщающих эйлерову характеристику сферы. Для описания всех этих чисел Виттен предложил построить по ним функцию, используя их в качестве коэффициентов степенного ряда (после подходящей нормировки). Между частными производными полученной функции было экспериментально обнаружено странное соотношение

$$u_t = 6uu_x + u_{xxx}.$$

Это соотношение было за сто лет до того обнаружено Д. Кортвегом и Г. де Фризом в совершенно другой теории — в теории распространения волн на мелководье. Одно из поразительных открытий математики последних десятилетий (у истоков которого стоят попытки Э. Ферми использовать компьютер для моделирования и проверки гипотез статистической физики) состоит в универсальности этого «вполне интегрируемого» уравнения математической физики. Открытие неожиданных связей между разнородными математическими теориями играет в математике такую же роль, как, например, открытие сходства между восточным берегом Америки и западным берегом Африки (приведшее к современной глобальной тектонике плит³⁾).

В случае связи между характеристическими числами пространств модулей кривых с отмеченными точками и дифференциальным уравнением теории мелководья, роль глобальной тектоники играет квантовая гравитация. Однако физик Виттен (кстати, тоже удостоенный математиками Филдсовской медали) не доказал существование этой связи, а лишь подметил ее (что, конечно, тоже очень важно и в случае геологии соответствует вкладу Вегенера).

Доказательство гипотезы Виттена (потребовавшее построения большой новой математической теории) — заслуга М. Концевича.

Достижения Концевича существеннейшим образом использовались в ряде недавних работ, посвященных решению еще нескольких важнейших задач из разных областей современной математики. Одной из таких задач явилось доказательство гипотезы «зеркальной симметрии», выдвинутой физиками. Эта гипотеза устанавливает удивительную связь двух совершенно различных (и на первый, и на второй взгляд) областей математики. Одна из областей — это симплектическая геометрия, т. е. классическая механика гамильтоновых консервативных систем (без диссипации энергии). Другая область — теория комплексных многообразий (многомерных обобщений римановых поверхностей, о которых шла речь выше).

Явление зеркальной симметрии состоит в том, что симплектическая геометрия одного многообразия оказывается связанной с комплексной геометрией другого. Симплектическая геометрия этого второго многообразия оказывается при этом связанной таким же образом с комплексной геометрией исходного.

³⁾ Истоком которой послужила теория дрейфа материков немецкого геофизика Вегенера (1880 – 1930), создавшего ее в 1912 г.

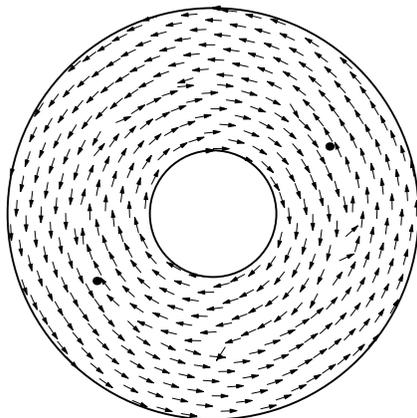


Рис. 12. *Неподвижные точки симплектического отображения кольца в себя*

Математическая теория этого явления создана другим воспитанником московской математической школы, А. Б. Гивенталем (которому я бы тоже присудил Филдсовскую медаль), ныне работающим в Беркли. Он использовал, среди прочего, достижения Концевича.

Другой сенсацией последнего года явилось доказательство (несколькими группами математиков, работающих в разных странах) старых (1965) гипотез о неподвижных точках симплектоморфизмов, т. е. о периодических движениях в соответствующих гамильтоновых системах (например, в небесной механике или в теории ускорителей). Теория стабильных отображений кривых, построенная Концевичем ради доказательства гипотезы Витгена, оказалась полезнейшим техническим средством при исследовании этого вопроса симплектической топологии.

Тот факт, что комплексная геометрия доставляет полезный инструмент исследования вопросов симплектической топологии, был впервые обнаружен ленинградским математиком М. Л. Громовым, ныне работающим во Франции. То, что он — не филдсовский лауреат, всегда меня удивляло. Думаю, что это объясняется причинами не научного характера (сыграло роль, что Россия не поддержала российского математика Громова; филдсовский комитет всегда был очень политизирован, как это и описано в недавно опубликованных мемуарах члена этого комитета, известного антисемита и замечательного математика Л. С. Понтрягина).

Простейший пример теоремы о неподвижных точках в симплектической топологии был указан еще Пуанкаре (в его посмертной «последней геометрической теореме»): если круговое кольцо отображено на себя так, что граничные окружности поворачиваются в разные стороны, а площади областей сохраняются, то такое отображение имеет не менее двух неподвижных точек (рис. 12).

Небольшое размышление показывает, что этот факт вытекает из следующей более общей теоремы: сохраняющее площади отображение тора в себя имеет не менее трех неподвижных точек, если оно не сдвигает с места «центр тяжести» тора. При этом число неподвижных точек даже не меньше четырех, если

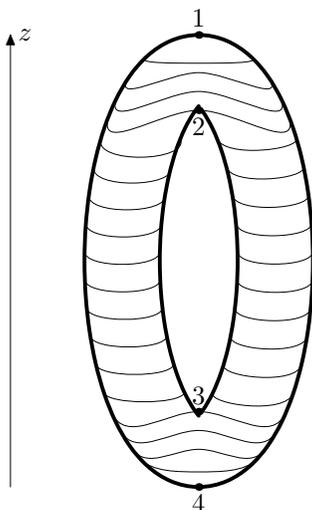


Рис. 13. Функция высоты на торе и ее четыре критические точки

учитывать вырожденные точки (в которых сливается несколько неподвижных точек) с их кратностями. Числа 3 и 4 здесь — это наименьшие числа критических точек (считая и максимумы, и минимумы, и седла и т. п.) гладкой функции на торе (их не меньше четырех, если считать с кратностями, и не менее трех из них геометрически различны).

Наука о критических точках гладких функций называется *теорией Морса – Люстерника – Шнирельмана* по имени американского (М. Морс) и российских (Л. А. Люстерник и Л. Г. Шнирельман) математиков, получивших оценки снизу для чисел критических точек любой функции (на поверхности или на многообразии бóльшего числа измерений) через топологические инварианты — так называемые числа Э. Бетти в случае теории Морса. Например, сумма чисел Бетти тора равна четырем, поэтому функция на торе имеет не менее четырех (учитывая кратности) критических точек (рис. 13).

Гипотезы симплектической топологии, о которых шла речь, утверждают, что точно такие же оценки снизу, как для числа критических точек гладких функций, справедливы и для неподвижных точек симплектоморфизмов фазовых пространств, определяемых системами гамильтоновых дифференциальных уравнений классической механики. Эти неподвижные точки или замкнутые орбиты оказываются критическими точками соответствующей «функции» на *бесконечномерном* многообразии. Стабильные отображения Концевича (вместе с теорией Громова и его последователей) открывают замечательную возможность распространения теории Морса на бесконечномерный случай в ситуации, где (как и в квантовой теории поля, для нужд которой эти стабильные отображения были изобретены) обычная техника вариационного исчисления и функционального анализа не работает.

Еще три недавние работы Концевича также доставляют неожиданные связи между совсем разными областями математики. Перекладывания отрезков —

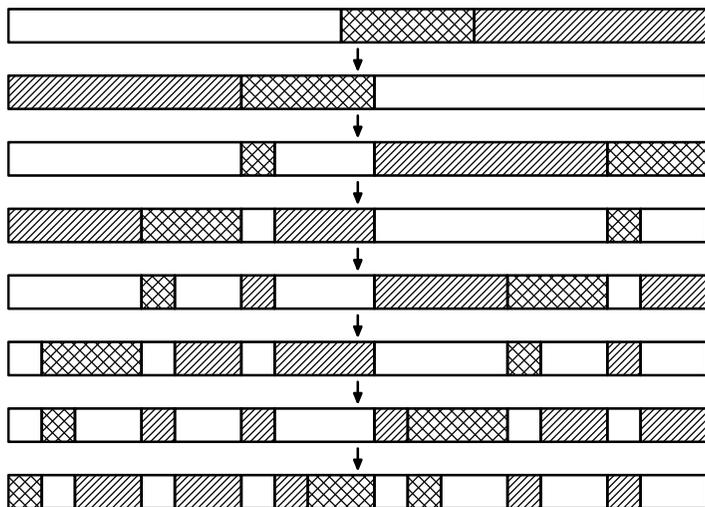


Рис. 14. Перекладывание трех отрезков

классический пример в теории динамических систем, перемешивания и «хаоса». Разобьем отрезок на несколько частей, переставим их в каком-либо порядке и снова склеим в отрезок. Повторяя эту процедуру много раз, мы заставим каждую точку отрезка блуждать по нему более или менее хаотическим образом (рис. 14).

Московский математик А. В. Зорич (сейчас работающий в Ренне, в Бретани) обнаружил в компьютерных экспериментах поразительные асимптотические статистические свойства этого перемешивания. Эти свойства оказались универсальными в том смысле, что асимптотическое поведение отклонений от средних оказалось не зависящим от длин отрезков разбиения (при почти любом их выборе), хотя и сильно зависящим от вида перестановки частей исходного отрезка и от числа разрезов.

Анализируя вместе с Зоричем эти асимптотики, Концевич обнаружил, что они теснейшим образом связаны с геометрией $(6g-6)$ -мерных пространств модулей римановых поверхностей подходящего рода g , о которых говорилось выше. Эта связь совершенно поразительна, даже если учесть, что перекладывания были введены (в 1963 г.) именно в качестве модели для исследования эргодических свойств псевдопериодических динамических систем, заданных гамильтоновыми векторными полями на римановых поверхностях рода g , большего единицы. Это была одна из исходных задач псевдопериодической топологии — молодой области науки, значение которой для теории поверхностей Э. Ферми в физике твердого тела было впоследствии обнаружено в замечательных работах учителя А. В. Зорича, московского математика С. П. Новикова, филдсовского медалиста 1970 г.

Столь же поразительна недавно открытая М. Концевичем (вместе с московским математиком Ю. М. Суховым, ныне работающим в Кембридже) универсальность статистических свойств многомерных цепных дробей.

Цепной дробью называется выражение вида $a + 1/(b + 1/(c + \dots))$. Вопрос о статистических свойствах целых чисел (a, b, c, \dots) , входящих в разложение в цепную дробь случайного числа, был решен К. Ф. Гауссом и Р. О. Кузьминым. До этого вопрос обсуждался шведским астрономом Х. Гильденом и шведскими математиками А. Виманом и Т. Броденом. Важность этого вопроса объясняется тем, что статистика разложений случайных чисел в цепные дроби сказывается в небесной механике на резонансах между движениями планет и астероидов, будучи существенной для установления устойчивости как планетных систем, так и пучков частиц в ускорителях и даже плазмы в токамаках, управляя заодно и потоками космических частиц, ответственных за полярные сияния.

Концевич с Суховым объяснили, как аналогичные теории Гаусса – Кузьмина вопросы об универсальной статистике для многомерных цепных дробей связаны с геометрией пространств модулей решеток в евклидовом пространстве (изучавшейся, между прочим, другим московским лауреатом медали Филдса, Г. А. Маргулисом, ныне работающим в Йельском университете в США). Поразительно, что, в отличие от Гаусса, авторы доказывают *существование* универсальной статистики, вовсе ее не находя. Вопрос о явном вычислении статистики Концевича – Сухова (либо с помощью суммирования соответствующих рядов по полилогарифмам, либо хотя бы прямым компьютерным моделированием соответствующих задач целочисленного линейного программирования, к которому по существу относится теория многомерных цепных дробей) остается пока открытым и ждет энергичных исследователей.

Упомяну в заключение совместную работу Концевича с московским математиком С. А. Баранниковым (ныне работающим во Франции) о геометрии каустик и волновых фронтов задач геометрической оптики, вариационного исчисления и теории оптимального управления. Эта работа объяснила загадочные связи этой области математики с теорией фробениусовых структур (развитой московским математиком Б. А. Дубровиным, ныне работающим в Триесте), выросшей, опять-таки, из квантовой теории гравитации.

Французы обычно сразу же выбирают своих филдсовских лауреатов в Академию наук. Пример М. Концевича показывает, что московская математическая школа продолжает активнейшим образом работать, причем вокруг таких лидеров, как Концевич, собираются, несмотря на географическую разобщенность, большие и сильные коллективы. Жаль только, что полученные ими замечательные достижения (которые я попытался хотя бы отчасти описать выше) у нас мало известны, а Концевич даже не был избран в РАН.

Николай Борисович Васильев

А. А. Егоров

28 мая 1998 года скончался Николай Борисович Васильев, замечательный математик, внесший неопценимый вклад в дело математического просвещения. Это был человек удивительной доброты, обаяния и неугасающей молодости. Все его звали по имени — Коля.

Коля Васильев родился в Москве 8 августа 1940 года в семье с глубокими интеллигентными корнями. Его отец Борис Федорович Васильев был известным инженером-строителем, основателем школы инженеров-строителей промышленных сооружений. Борис Федорович пережил сына, он скончался в возрасте 95 лет через три с небольшим месяца после его смерти.

Мать Коли, Нина Николаевна Лессиг, также была инженером-строителем, кандидатом технических наук.

Н. Б. Васильев родился и прожил всю свою жизнь в одном из красивейших уголков Москвы — на Софийской набережной, прямо напротив Кремля. Его детство прошло среди книг, в кругу широко образованных людей, любящих и глубоко знающих искусство, особенно музыку. В традициях этой семьи сохранялась даже такая диковинка для нашего времени, как домашнее музицирование.

Вся обстановка семьи и семейное окружение привили Коле безупречный вкус и тонкое понимание красоты. Он ценил красоту и в искусстве, и в математике, увлечение которой пришло к нему в старших классах. Коля занимался математикой самостоятельно, не испытывая постороннего влияния, не посещая математические кружки и не участвуя в олимпиадах.

Закончив с серебряной медалью школу в 1957 году и одновременно музыкальное училище при Московской консерватории, он оказался перед выбором — математика или музыка.

Победила математика, и Коля стал студентом первого курса механико-математического факультета МГУ. Все пять лет он учился блестяще.

Окончив университет в 1962 году, он сразу же поступил в аспирантуру мехмата МГУ, а после окончания аспирантуры стал работать в межфакультетской лаборатории математических методов в биологии МГУ, где и трудился до последних дней своей жизни.

Кружки и олимпиады

Начиная с 1 курса, Коля включился в работу Оргкомитета Московской математической олимпиады. И с тех пор все годы (за исключением 1980 – 87 гг.) он был среди организаторов олимпиад. Он придумывал задачи, участвовал в составлении вариантов, проверке работ, обсуждении результатов и присуждении премий. О том, как это происходило, очень хорошо и ярко написано в книгах [7] и [8].

Молодые математики (студенты и аспиранты), увлеченно работавшие со школьниками, всегда отличались большой свободой взглядов, независимостью

суждений и огромной преданностью науке. Свободная атмосфера собраний оргкомитета олимпиады, независимость его от администрации часто вызывали неудовольствие партийного и комсомольского начальства. Оно отказывалось считать олимпиадную деятельность общественной работой. Один из таких начальников прямо заявил: «Какая же это общественная работа, если вы ее делаете ради собственного удовольствия?»

В те годы Московская олимпиада как бы подводила итоги работы школьных математических кружков при МГУ в течение учебного года [7]. Руководили кружками студенты и аспиранты, многие из которых стали ныне известнейшими математиками. Достаточно назвать В. И. Арнольда, А. А. Кириллова, Н. Н. Константинова, А. М. Олевского, Д. Б. Фукса. Почти все участники кружков считали свой долгом явиться на олимпиаду и добиться там успеха. По окончании олимпиады кружковцы подсчитывали количество премий и сравнивали успехи своего кружка с успехами «конкурентов».

Осенью 1958 года Коля начал вместе с автором этой статьи вести одну из секций школьного математического кружка. Кружок оказался сильным. Во всяком случае, он хорошо выступал на олимпиадах, многие из его участников стали весьма успешно работающими математиками. Это С. Гельфанд, Р. Зигангиров, Д. Каждан, А. Каток, И. Д. Новиков и другие.

* * * * * *

Биография Коли Васильева с 1958 до 1979 года неотделима от истории не только Московских, но и Всероссийских, а затем Всесоюзных олимпиад. Подробно с историей олимпиад можно познакомиться по предисловию к книге «Задачи Всесоюзных математических олимпиад» [6]. Работа в Оргкомитете Московской, а затем в жюри и методической комиссии Всероссийской и Всесоюзной олимпиад сблизила Колю Васильева с величайшим математиком современности А. Н. Колмогоровым. Андрей Николаевич долгие годы возглавлял методические комиссии математических олимпиад и неоднократно руководил жюри Олимпиады. Коле повезло не раз быть заместителем Андрея Николаевича.

С середины 60-х годов Центральный оргкомитет Всесоюзной олимпиады рассылал в области и республики брошюры с задачами, рекомендованными для областных и республиканских олимпиад. Многие такие подборки были составлены Н. Б. Васильевым. Это, конечно, не значит, что все задачи были им придуманы. Его авторство — в их составлении. Это был огромный труд. В подборках содержались оригинальные задачи разных уровней — от легких до весьма трудных. Среди задач были задачи по алгебре и геометрии, не забывалась и традиционная «олимпиадная» тематика. Практически во всех областях и республиках СССР на олимпиадах использовались задачи из этих подборок. Мне кажется, что было бы полезно собрать вместе эти маленькие брошюры и издать их отдельной книжкой.

ЖУРНАЛ «КВАНТ»

В 1970 году А. Н. Колмогоров стал первым заместителем главного редактора журнала «Квант», а Коля — членом редколлегии и руководителем раздела «Задачник Кванта» — одного из главных и, пожалуй, самых знаменитых разделов этого журнала.

Надо сказать, что идея создания популярного физико-математического журнала для школьников родилась еще в середине 60-х годов в коллективе математиков и физиков, объединившихся вокруг олимпиады. Коля Васильев был одним из самых настойчивых и упорных сторонников этой идеи, так что создание «Кванта» — это во многом и его заслуга.

Как руководитель «Задачника Кванта», он придумывает задачи, причем его задачи выделяются отточенностью формулировок и решений, глубиной и связями с «большой» математикой. Олимпиады и занятия математикой были тогда делом престижным, и очень многие школьники, студенты, преподаватели присылали в редакцию придуманные ими задачи. Поэтому он имеет дело с огромной почтой. Естественно, большинство из присылаемых задач были, скажем так, не оригинальны и мало интересны. Однако среди них попадались и вполне достойные внимания. Коля как никто мог «обработать» задачу, найти наиболее привлекательную формулировку, обнаружить возможные обобщения и дальнейшее развитие сюжета задачи (очень часто задачи «Задачника» состоят из нескольких пунктов, последовательно усложняющих и развивающих фабулу задачи). Наконец, часто бывало так, что авторские решения вполне хороших задач оказывались бездейственными и, что называется, тупыми — например, лобовой счет в задачах по геометрии, не позволяющий по-настоящему разобраться в существе дела. После Колиной обработки многие задачи становились настоящими сокровищами (разумеется, под формулировками их условий стояла подпись автора, так что участие Коли оставалось «за кадром»). Очень многие написанные им решения сопровождаются комментарием и литературными отсылками для тех, кто захочет глубже разобраться в проблеме.

Колина работа в «Кванте» не ограничивалась «Задачником». Он — автор более тридцати статей, без сомнения, входящих в число лучших публикаций «Кванта». Некоторые из них собраны в вышедшей посмертно книге «Избранные статьи» [10]. Многие статьи написаны им в соавторстве с известными математиками. Часто бывало так, что присланная в редакцию статья была интересна по теме, но по исполнению никак не соответствовала ни уровню «Кванта», ни стилю этого журнала. В таких случаях Николай Борисович принимался за переделку, в результате которой появлялась яркая, интересная статья.

Умение видеть главное и тонкий вкус сочетались в нем с литературными способностями. Его статьи написаны превосходным языком. Коля обладал редкостным даром прозрачно и четко излагать самые сложные математические сюжеты. Зная это его умение, многие математики приносили в редакцию свои статьи и просили Колю помочь довести их «до ума». Он никогда в этом не отказывал, а в некоторых случаях полностью переписывал статью, причем нередко его роль настолько превышала чисто редакторскую, что авторы сами просили его стать их соавтором.

В то же время он умел, не причиняя обиды, но достаточно твердо отклонять статьи в тех случаях, когда они были малоинтересны или непригодны для публикации в журнале для школьников.

С первых дней существования «Кванта» не было ни одного крупного события в жизни журнала, в котором он не принимал бы участия. «Квант» был его вторым домом, а он был его душой.

Николай Борисович считал, что в системе научно-популярных изданий должен существовать журнал, промежуточный по своему уровню между «Квантом» и чисто научными журналами. Именно поэтому он стал одним из главных инициаторов возрождения сборников «Математическое просвещение», был держателем гранта РФФИ, предназначенного для поддержки этого издания, вошел в редколлегию новой серии «Математического просвещения».

ВЗМШ и другие школы

Во второй половине 1963 года началась подготовка к созданию Всесоюзной заочной математической школы. Ее создатель И. М. Гельфанд привлек своего аспиранта Васильева к разработке программ и вступительного задания. С тех пор и до последних дней своей жизни Николай Борисович был без преувеличения ключевой фигурой в ВЗМШ. До последних дней своей жизни он был членом Научного совета и методической комиссии. Им написаны десятки учебных заданий и брошюр, огромное количество статей для учителей и школьников. Ежегодно в течение 35 лет под его руководством разрабатывались вступительные задания в ВЗМШ.

Книги Николая Борисовича, написанные для ВЗМШ, известны не только в России. В частности, книга «Прямые и кривые» [4], написанная им вместе с В. Л. Гутенмахером, в 1980 году была переведена на английский и испанский языки издательством «Мир», а в 1982 году издана в Чехословакии на чешском языке. Книга Н. Б. Васильева, В. Л. Гутенмахера, Ж. М. Раббота и А. Л. Тоома «Заочные математические олимпиады» [1] — превосходный образец популярной литературы для школьников. Книга тех же авторов «Задачи устного экзамена по математике» [9] оказалась очень удачной. В течение нескольких лет она рассылалась ученикам ВЗМШ и сильно помогла им при подготовке к устным вступительным экзаменам в вузы самого высокого уровня. В дальнейшем она послужила образцом для других авторов, писавших на эту тему. К сожалению, эта книга с тех пор ни разу не переиздавалась и сейчас является чрезвычайно библиографической редкостью.

В течение многих лет Николай Борисович принимал участие во вступительных экзаменах в школу-интернат при МГУ (ныне школа им. А. Н. Колмогорова). Он придумывал задачи для экзамена, был одним из лучших доброжелательных экзаменаторов. Характеристики перспектив абитуриентов, даваемые им после экзамена, очень точны и практически всегда подтверждались при дальнейшем их обучении в школе.

В 70-е – 80-е годы Николай Борисович принял участие в работе многих летних школ: под Москвой, в Карелии, в Таджикистане и других. Он великолепно читал лекции и проводил занятия. Каждый раз он придумывал способ изложения, соответствующий уровню аудитории, но не наносящий ущерба глубине и содержанию лекции. Я неоднократно присутствовал на его лекциях и всегда получал истинное удовольствие от мастерства лектора.

Турнир Городов и другие соревнования

В 1979 году чиновники от просвещения разогнали все тогдашнее жюри Всесоюзной олимпиады. И тогда Н. Н. Константинов предложил проводить новую олимпиаду, организация которой была бы полностью в руках математической

общественности и не зависела бы от произвола чиновников. Так родился Турнир Городов — ныне одно из самых уважаемых соревнований в нашей стране и за рубежом. Подробно о нем рассказано в статье Н. Н. Константинова «Турнир городов и математическая олимпиада»¹⁾.

Николай Борисович сразу же подхватил идею и стал одним из ведущих организаторов этого международного турнира. Именно он сыграл основную роль в выработке формы проведения турнира, во многом определил стиль и уровень задач, неоднократно участвовал в работе летних конференций Турнира Городов.

С 1994 года в России проводится Соросовская олимпиада школьников по математике, физике, химии и биологии. Николай Борисович до последних месяцев своей жизни активнейшим образом работал и в этой олимпиаде. Он придумал несколько красивых задач, принимал участие в составлении вариантов и проверке работ.

В последние десятилетия стали весьма популярными научные конференции школьников. Николай Борисович на многих таких конференциях был членом жюри. Его суждение всегда было доброжелательным и четким. Он всегда беседовал с докладчиками, объяснял им возможности развития тем их докладов. Наиболее интересные доклады по его представлению были опубликованы в журнале «Квант».

В этих кратких заметках ни слова не сказано о научной работе Н. Б. Васильева. Ей посвящена статья А. М. Леонтовича в этом номере «Математического просвещения»²⁾.

ИЗБРАННЫЕ ЗАДАЧИ Н. Б. ВАСИЛЬЕВА

Коля придумал огромное количество задач для олимпиад самого разного уровня — от областных и республиканских до Московских, Всесоюзных и Всероссийских. Некоторые из его задач стали олимпиадной классикой, известны почти всем участникам и организаторам олимпиад (правда, далеко не все знают, что их автор — Н. Б. Васильев).

Здесь подобраны сравнительно старые задачи Николая Борисовича. Более новые нетрудно найти, заглянув в «Задачник Кванта». Практически в каждом номере журнала есть либо его задачи, либо написанные им решения, очень часто раскрывающие и углубляющие истинную суть задач, их связи с другими проблемами. Предлагаемые задачи имеют красивые формулировки и очень элегантные остроумные решения.

1. («Автобусная задача».) k человек ехали в автобусе без кондуктора. Известно, что ни у кого из пассажиров не было монет крупнее 20 коп. Известно также, что каждый пассажир уплатил за проезд и получил сдачу. Докажите, что наименьшее число монет, которое могло для этого потребоваться, равно $k + \left[\frac{k+3}{4} \right]$. (В 1961 году в обращении были медные монеты достоинством в 1, 2, 3 и 5 коп., а также «серебряные» — белые монеты в 10, 15 и 20 коп. В автобусе без кондуктора полагалось бросить в кассу 5 коп. в

¹⁾ Математическое просвещение. М.: МЦНМО, сер. 3, вып. 1. 1997. С. 164 – 174.

²⁾ См. с. 29–32.

любом наборе и оторвать билет. Можно было также бросить, например, 20 коп. и оторвать 4 билета.)

(*XXIII Московская математическая олимпиада*)

2. («Коробочка Васильева».) Как надо расположить в пространстве прямоугольный параллелепипед, чтобы площадь его проекции на горизонтальную плоскость была наибольшей.

(*XXV Московская математическая олимпиада*)

3. («Гуляющие джентльмены».) По аллее длиной 100 м идут три джентльмена со скоростями 1, 2 и 3 км/ч соответственно. Дойдя до конца аллеи, каждый из них поворачивает и идет назад с той же скоростью. Докажите, что найдется отрезок времени в 1 мин, когда все трое будут идти в одном направлении.

4. («Квадраты в прямоугольнике».) В прямоугольник со сторонами 20 и 25 бросают 120 квадратов со стороной 1. Докажите, что в прямоугольник можно поместить круг диаметра 1, не пересекающийся ни с одним из квадратов.

(*Первая Всероссийская олимпиада, 1961 год, Москва*)

5. («Жук».) На плоскости нарисована сеть, образованная из правильных шестиугольников со стороной 1. Жук, двигаясь по линиям сети, прополз из

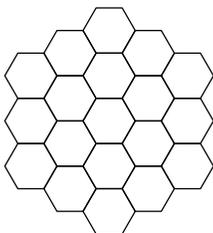


Рис. 1. Сеть из шестиугольников

узла A в узел B по кратчайшему пути, равному 100. Докажите, что половину всего пути он полз в одном направлении.

(*Четвертая Всероссийская олимпиада, 1964 год, Москва*)

6. («Игра Васильева».) Написано 20 чисел: 1, 2, ..., 20. Двое играющих по очереди ставят перед этими числами знаки «+» или «-» (знак можно ставить перед любым свободным числом). Первый стремится к тому, чтобы после расстановки всех 20 знаков сумма всех чисел была как можно меньше по модулю. Какую наибольшую по модулю сумму может обеспечить себе второй игрок?

(*Шестая Всероссийская олимпиада, 1966 год, Воронеж*)

7. («Фигуристы».) После выступлений 20 фигуристов каждый из 9 судей по своему усмотрению распределяет среди них места с 1-го по 20-е. Оказалось, что у каждого фигуриста места, присвоенные ему разными су-

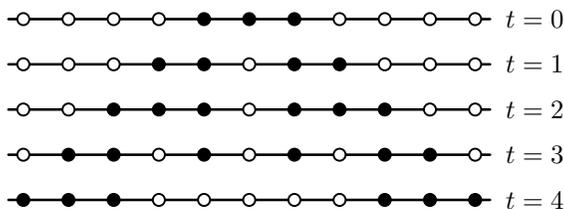


Рис. 2.

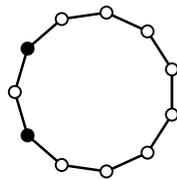


Рис. 3.

дьями, отличаются не более чем на 3. Подсчитаем суммы мест, полученных каждым фигуристом, и расположим эти числа в порядке возрастания: $c_1 \leq c_2 \leq c_3 \leq \dots \leq c_{20}$. Какое наибольшее значение может иметь c_1 ?

(Вторая Всесоюзная олимпиада, 1968 год, Ленинград)

8. («Разноцветный многоугольник».) Вершины правильного n -угольника покрашены несколькими красками (каждая одной краской) так, что точки одного и того же цвета служат вершинами правильного многоугольника. Докажите, что среди этих многоугольников найдутся два равных.

(Четвертая Всесоюзная олимпиада, 1970 год, Симферополь)

9. («Нервные сети».) В бесконечной цепочке нервных клеток каждая клетка может находиться в одном из двух состояний: «покой» и «возбуждение». Если в данный момент клетка возбудилась, то она посылает сигнал, который через единицу времени (скажем, через одну миллисекунду) доходит до обеих соседних с ней клеток. Каждая клетка возбуждается в том и только в том случае, если к ней приходит сигнал от одной из соседних клеток; если сигналы приходят одновременно с двух сторон, то они гашаются, и клетка не возбуждается. Например, если в начальный момент времени $t = 0$ возбудить три соседние клетки, а остальные оставить в покое, то возбуждение будет распространяться, как показано на рис. 2. (Возбужденные клетки — черные.)

Пусть в начальный момент времени $t = 0$ возбуждена только одна клетка. Сколько клеток будет находиться в возбужденном состоянии через 15 мсек? через 65 мсек? через 1000 мсек? через t мсек?

Что будет в том случае, если цепочка не бесконечная, а содержит всего N клеток, соединенных в окружность (рис. 3), — будет ли возбуждение поддерживаться бесконечно долго или затухнет?

(Задачник Кванта, М19)

10. («Параллелепипеды».) В пространстве задали 4 точки, не лежащие в одной плоскости. Сколько существует различных параллелепипедов, для которых эти точки служат вершинами?

(Седьмая Всесоюзная олимпиада, 1973 год, Кишинев)

11. («Векторы».) На плоскости даны векторы a, b, c, d , сумма которых равна 0. Докажите, что

$$|a| + |b| + |c| + |d| \geq |a + d| + |b + d| + |c + d|.$$

(Десятая Всесоюзная олимпиада, 1976 год, Душанбе)

* * *

Всю свою творческую жизнь Николай Борисович работал: писал статьи, книги, решал и придумывал задачи, занимался наукой. Трудно себе представить, что все, им созданное, — дело рук одного человека. Его уход — огромная потеря для всех его друзей, для дела математического просвещения.

СПИСОК ЛИТЕРАТУРЫ

- [1] *Васильев Н. Б., Гутенмахер В. Л., Раббот Ж. М., Тоом А. Л.* Заочные математические олимпиады. М.: Наука, 1981.
- [2] *Васильев Н. Б., Егоров А. А.* Сборник подготовительных задач к Всероссийской олимпиаде юных математиков / под ред. А. Н. Колмогорова. М.: Учпедгиз, 1963.
- [3] *Васильев Н. Б., Савин А. П.* Избранные задачи математических олимпиад. М.: МГУ, 1968.
- [4] *Васильев Н. Б., Гутенмахер В. Л.* Прямые и кривые. М.: Наука, 1978.
- [5] *Васильев Н. Б., Молчанов С. А., Розенталь А. Л., Савин А. П.* Математические соревнования (геометрия). М.: Наука, 1974.
- [6] *Васильев Н. Б., Егоров А. А.* Задачи Всесоюзных математических олимпиад. М.: Наука, 1988.
- [7] Сборник задач московских математических олимпиад / Сост. А. А. Леман. М.: Просвещение, 1965.
- [8] *Гальперин Г. А., Толыго А. К.* Московские математические олимпиады / под ред. А. Н. Колмогорова. М.: Просвещение, 1986.
- [9] *Тоом А., Гутенмахер В., Васильев Н., Раббот Ж.* Задачи устного экзамена по математике. М.: МГУ, 1970.
- [10] *Васильев Н. Б.* Избранные статьи / Приложение к «Кванту», №6. М.: Бюро Квантум, 1998.

Научная деятельность Н. Б. Васильева

А. М. Леонтович

Первые научные результаты были получены Н. Б. Васильевым, когда он еще был студентом. Тогда его интересы в основном относились к функциональному анализу. Уже на первом курсе он стал активным участником семинара А. Н. Колмогорова. Затем его учителями стали Г. Е. Шилов и И. М. Гельфанд. В это время была сделана его первая работа о C^* -алгебрах.

В 1966 году Н. Б. Васильев поступил на работу в межфакультетскую лабораторию математических методов в биологии. Он стал одним из первых ее сотрудников. Лаборатория была образована по инициативе И. М. Гельфанда (который и стал ее заведующим) при активной поддержке ректора университета И. Г. Петровского¹).

Место работы оказалось очень удачным; недаром Н. Б. Васильев проработал в лаборатории (затем институте) до самой своей кончины. И. М. Гельфанд собрал здесь вместе ряд сильных математиков. Можно назвать имена С. Г. Гиндикина, Д. Б. Фукса; позже появились Д. А. Каждан, И. Н. Бернштейн. В лаборатории царил очень дружеская, доброжелательная и в то же время творческая атмосфера.

Н. Б. Васильев стал работать в группе, возглавляемой замечательным математиком И. И. Пятецким-Шапиро, несомненно, оказавшим большое влияние на его творческий почерк. Эта группа занималась теорией автоматов и математическими моделями биологических процессов. Одной из центральных тем для этой группы стало изучение марковских цепей определенного типа — случайных сред с локальными однородными взаимодействиями. Активным участником здесь, наряду с О. Н. Ставской, А. Л. Тоомом, стал Н. Б. Васильев. Он получил ряд ярких результатов, часть из которых составили содержание его кандидатской диссертации, защищенной в 1971 году.

Конструкция случайных сред с локальными однородными взаимодействиями вначале появилась как математическая модель работы нейронных сетей. Но затем эти задачи стали чисто математическими. По своему содержанию они близки к таким приложениям теории марковских цепей как математическое обоснование статистической физики и теория просачивания. Почти одновременно случайными средами с локальными однородными взаимодействиями занималась

¹Одновременно была создана и Межфакультетская лаборатория биоорганической химии и молекулярной биологии под руководством академика А. Н. Белозерского (к ее образованию также активное отношение имели И. Г. Петровский и И. М. Гельфанд). Впоследствии (в 1974 году) эти лаборатории объединились, малочисленная межфакультетская лаборатория стала отделом математических методов в биологии большой Межфакультетской лаборатории (сейчас она называется институтом физико-химической биологии им. А. Н. Белозерского МГУ). Идея создания этих лабораторий оказалась очень плодотворной, и в настоящее время институт им. А. Н. Белозерского имеет, быть может, сильнейший состав биологов в нашей стране.

школа Р. Л. Добрушина, но в ней рассматривались не дискретные сети, а непрерывные марковские поля.

Приведем простейший пример задачи о случайных средах с локальными однородными взаимодействиями — так называемую «задачу о лампочках».

В этой задаче автоматы («лампочки») расположены в целых точках прямой. Время дискретно. Автоматы могут находиться в двух состояниях: 0 и 1 («лампочка не горит» или «лампочка горит»). Если в предыдущий момент данная лампочка и соседняя слева горели, то в текущий момент она обязательно продолжает гореть. Если же в предыдущий момент хотя бы одна из них не горела, то в текущий момент она будет гореть с вероятностью p , $0 < p < 1$ («лампочка спонтанно загорается с вероятностью p »). Предполагается, что состояния всех автоматов в настоящий момент принимаются независимо друг от друга (а только в зависимости от состояний автоматов в предыдущий момент). Такие стохастические правила поведения автоматов определяют марковский оператор P (другими словами, марковскую цепь) с континуальным числом состояний; каждое из таких состояний цепи — это совокупность состояний всех автоматов прямой, т. е. бесконечная последовательность из нулей и единиц. Очевидно, что состояние всей целочисленной прямой «все единицы» поглощающее (переводится марковским оператором P в себя). Этому поглощающему состоянию отвечает дискретная (атомарная) мера, сосредоточенная в этом состоянии «все единицы»; эта мера инвариантна относительно оператора P^* , двойственного P (этот оператор P^* действует на множестве мер, определенных на множестве всех состояний целочисленной прямой).

Спрашивается, есть ли другие инвариантные меры, кроме этой? Оказывается, ответ зависит от значения вероятности p . Именно, для больших p есть только эта инвариантная мера (сосредоточенная в состоянии «все единицы»), и если в нулевой момент состояния всех автоматов равны 0 («все лампочки не горят»), то с течением времени доля единиц будет стремиться к 1. Совсем не такая ситуация при малых p (это было доказано вначале М. Г. Шнирманом, затем А. Л. Томом). В этом случае, если мы начинаем из состояния «все нули», то доля единиц в любой момент времени остается малой. В пределе, когда время стремится к бесконечности, мы выходим на некий стационарный «мигающий» режим работы лампочек. Этому режиму отвечает некоторая инвариантная мера. Эта мера не совпадает с мерой, сосредоточенной в состоянии «все единицы», она будет не дискретна, а непрерывна. Таким образом, имеет место явление типа фазового перехода. Критическое значение p , при котором происходит этот фазовый переход (изменение числа инвариантных мер), приблизительно равно 0,3.

«Задача о лампочках» естественным образом обобщается в разных направлениях: (1) число автоматов, от которых зависит (вероятностным образом) состояние данного автомата i , может быть больше, чем 2 (но всегда конечно — условие локальности), и для каждого набора состояний этих соседних автоматов задается своя вероятность того, что данный автомат окажется в состоянии 1 (причем эти вероятности не зависят от номера i на прямой — условие однородности); (2) вместо целочисленной прямой можно рассматривать многомерные решетки; (3) число состояний автомата может быть больше двух. Для данной случайной среды из таких автоматов (т. е. при фиксированных вероятностях переходов) возникает ряд естественных вопросов: имеет ли место эргодичность

(единственность инвариантной меры); если нет, то каково множество инвариантных мер; характер зависимости инвариантных мер от вероятностей перехода (имеет ли место аналитичность); обладают ли эти инвариантные меры «хорошими» свойствами типа марковости (или близкого к марковости, но более сложно определяемого свойства гиббсовости).

В задачах о случайных средах с локальными однородными взаимодействиями Н. Б. Васильевым были получены результаты в двух направлениях. Во-первых, им было доказано, что при достаточно широких предположениях инвариантная мера аналитически зависит от вероятностей перехода. В частности, это справедливо в случае «задачи о лампочках» — непрерывная инвариантная мера при малых p аналитически зависит от p . Метод, которым это было доказано, можно назвать «геометрическим», он близок к контурным методам в модели Изинга, которые использовали в своих работах Р. Л. Добрушин, Я. Г. Синай и другие (в частности, именно таким образом А. Л. Тоом доказал существование нетривиальной инвариантной меры при малых p в «задаче о лампочках», о чем упоминалось выше). Во-вторых, Н. Б. Васильевым было изучено, при каких вероятностях перехода инвариантная мера будет марковской или гиббсовской (заметим, что для «задачи о лампочках» это не так — непрерывная инвариантная мера в этом случае довольно-таки «плохая» и никаким условиям типа марковости или гиббсовости не удовлетворяет).

У Н. Б. Васильева есть еще ряд работ по марковским цепям, не связанных непосредственно со случайными средами с локальными однородными взаимодействиями, но близких к ним по духу.

Н. Б. Васильев работал в биологическом институте и потому многие его работы инициированы практическими задачами — относящимися к биологии, химии, физиологии. В математическом отношении некоторые из них носят вероятностный характер, другие относятся к динамическим системам. Иногда необходимо было применить численное моделирование.

Изложим вкратце результаты последней работы Н. Б. Васильева (оттиски которой появились за месяц до его смерти, когда он уже лежал в больнице). Ее возникновение связано с практической задачей о клеточных культурах, именно, выяснении того, имеют ли клетки в однослойных клеточных культурах тенденцию к налеганию друг на друга или, наоборот, к отталкиванию. Математическая постановка состоит в следующем. Пусть имеется некоторое число фигур, случайно расположенных в поле зрения (чтобы сделать математическую постановку более точной, предполагается, что поле зрения является тором — квадратом площади 1 с отождествленными противоположными сторонами). Спрашивается, каково распределение суммы площадей попарных пересечений фигур. Легко видеть, что среднее значение площади пересечения двух случайно расположенных фигур равно произведению площадей этих фигур (напомним: предполагается, что площадь поля зрения равна 1). Поэтому среднее значение суммы площадей попарных пересечений n фигур равно сумме произведений площадей этих фигур. Доверительные интервалы для суммы площадей попарных пересечений случайно расположенных фигур находятся через дисперсии площадей пересечений двух фигур. В отличие от средних значений, эти дисперсии зависят не только от площадей фигур, но и от их формы. Встает вопрос об оценках этих дисперсий. Здесь возникает естественная гипотеза о том, что для двух фигур фиксированной

площади дисперсия максимальна в том случае, когда фигуры являются кругами. Это утверждение («изопериметрического типа») было доказано, правда, немного в ослабленном варианте, Н. Б. Васильевым. В частности, из его результата вытекает, что это утверждение справедливо, если одна из фигур выпукла (хотя на самом деле это доказано для более широкого класса фигур). Доказательство теоремы носит в основном геометрический характер, но при этом используются красивые точные формулы, иногда очень неожиданные.

Скажем несколько слов о стиле Н. Б. Васильева как математика.

Его отличала исключительно сильная геометрическая интуиция и в то же время склонность и к комбинаторным соображениям. Очень важна для него была эстетическая сторона. В его доказательствах все время проявлялось стремление к красоте, своего рода музыкальности. (Надо сказать, что его отличала и музыкальность в обычном смысле слова; вообще, музыкальность у него проявлялась во всем — и в литературных вкусах, и в жизни, и в математике. Пожалуй, он был самый музыкальный человек из моих знакомых — хотя формальные музыкальные знания у некоторых из них могли быть и большие.)

Незадолго до смерти, когда Коля лежал в больницу (из которой он так и не вышел), врач при осмотре спросил его, хорошо ли он слышит. Коля ответил, что со слухом у него все в порядке, потом помолчал и сказал: «Слух у меня абсолютный» (и это действительно было так). Так вот, в математике у него был абсолютный и слух, и вкус.

Н. Б. Васильева отличала замечательная способность выделить главное в рассуждениях, увидеть в нем геометрическую основу, даже если формально геометрии в доказательстве и не фигурировало. Поэтому очень часто после обсуждений с ним доказательства сильно упрощались.

Вот типичная картина. Происходит семинар. Докладчик рассказывает плохо и непонятно. Коля вроде бы не слушает, чуть ли не спит. Слушатели не понимают, недовольны этим, требуют, чтобы докладчик говорил понятнее, объяснил суть дела. Но у докладчика это не получается (часто потому, что докладчик и сам не понимает, в чем эта суть — увы, это часто бывает). И вдруг раздается голос Коли, он поясняет сказанное докладчиком, и все всем становится понятным.

По-видимому, такая способность увидеть суть дела приводила к тому, что часто в его работах возникали красивые точные формулы.

А вот громоздкие оценки и рассуждения — этого он не любил (хотя часто они бывают необходимыми).

Н. Б. Васильев был очень контактным человеком, очень доброжелательным к другим, с ним было очень приятно вместе работать. Наверное, поэтому большое количество из его работ совместные. А еще имеется значительно большее число работ, в создании которых он принимал участие при их обсуждениях, но формально автором не числился.

Автор этой статьи хорошо это знает по своему опыту, поскольку, наверно, был самым близким к Коле (в отношении математики) человеком, самым частым его соавтором. Но главное при этом даже не эти совместные работы. Фактически все работы автора — совместные, несовместные — обсуждались с Колей, происходила их своеобразная «обкатка». И это приводило, несомненно, к улучшению качества работ (опять же из-за способности Коли увидеть суть дела, из-за его эстетического чувства).

Евгений Михайлович Ландис

Евгений Михайлович Ландис был одним из тех, благодаря кому шестидесятые называются «золотыми годами московской математики».

Он поступил на мехмат за месяц до своего восемнадцатилетия в 1939 г. и был призван с первого курса в армию. Началась Финская война, и Геня Ландис участвовал в ней с первых дней. Он был обморожен и ранен на Карельском перешейке. Евгений Михайлович прошел всю Великую Отечественную Войну, награжден орденом Красного Знамени и медалями. В мае 1945 года он вошел в Берлин и запомнил его, превращенным в развалины. В том же году он вернулся в Университет.

Обычно перерыв в занятиях математикой даже на два года военной службы пагубно сказывается на творческих способностях студента. Евгений Михайлович, после шести лет войны, вернулся на мехмат, полный сил. Он рассказывал, что чувствовал себя в то время, как джинн из «1001 ночи»: «Что прикажет мой повелитель: построить дворец или разрушить город?»

Он начал работать под руководством А. С. Кронрода, своего друга и сверстника, который был демобилизован из-за ранений и учился на пятом курсе, когда Ландис вернулся на первый. На старших курсах Евгений Михайлович стал учеником и другом Ивана Георгиевича Петровского. Соединявшая их дружба продолжалась до конца дней Ивана Георгиевича (1973 г.). В 1945 году Евгений Михайлович сделал свою первую научную работу; к окончанию факультета в 1950 г. он написал уже пять работ и в 1951 г. получил премию Московского математического общества.

Однако в эти годы, в разгар «борьбы с космополитизмом», путь в аспирантуру Евгению Михайловичу был закрыт. Не удалось ему устроиться и на работу по специальности. Он был наконец принят преподавателем математики в школу рабочей молодежи. Сразу после смерти Сталина, в марте 1953 г., Петровский предложил ему защитить кандидатскую диссертацию, которую Евгений Михайлович и защитил в мае того же года. После этого почти 45 лет Евгений Михайлович работал на мехмате МГУ. Несколько лет он работал параллельно в группе Кронрода в ИТЭФ (Институт Теоретической и Экспериментальной Физики). Это была одна из первых групп, разрабатывавших в России основы науки, которая теперь называется «Computer sciences».

Евгений Михайлович создал большую школу; к числу его учеников принадлежит и автор этих строк. Он учил нас не только математике, но и тому, как быть Человеком в математике. Учил не словами, а своей повседневной жизнью.

В последние годы Евгений Михайлович много болел. Уже на восьмом десятке, с разрушенным здоровьем, он продолжал нести полную профессорскую нагрузку, как нес ее всю жизнь. Евгений Михайлович скончался 12 декабря 1997 г. Представленные далее материалы — скромная дань его памяти.

Памяти Е. М. Ландиса

М. Л. Гервер

Ландис любил Моцарта и Пушкина. И в нем самом была удивительная легкость и простота, напоминавшая моцартовскую музыку или пушкинские стихи.

Эта легкость и простота, мягкость, отзывчивость, редкая доброта и истинная интеллигентность создавали вокруг него особую ауру и привлекали к нему людей.

Мы познакомились в 55-м году, с 56-го года я стал его учеником и с тех самых пор неизменно ощущал необычайно внимательное и благожелательное отношение к себе.

С ним всегда было хорошо. В работе он был требователен, но щедр на похвалы. В беде можно было твердо рассчитывать на его поддержку и участие. Вместе отдыхать (например, в Тракае, где во время долгих лодочных прогулок он неспешно рассказывал о жизни и о войне, о математике и математиках, или беседовал о музыке, о прочитанных книгах, да мало ли о чем еще) было приятно и никогда не скучно.

Иногда замечательные качества Ландиса оборачивалась комической стороной. В 50-е годы с ним совершенно невозможно было ходить по мехмату: перед каждой дверью он останавливался и пропускал тебя вперед. Я выработал специальную стратегию: я вставал почти что по стойке «смирно» в метре от двери, и он выучился входить первым. Но подчас он об этом забывал, а я уже знал, что он умеет, и тогда мы сталкивались в дверях как Манилов и Чичиков.

Про Пушкина и про стихи. Первым учителем Ландиса в МГУ был А. С. Кронрод. Он был выдающимся, талантливейшим математиком, но потом бросил неопубликованными ряд готовых работ и говорил: «пусть сушеные селедки доказывают теоремы, нужно заниматься делом», т. е. заниматься вычислительными машинами. У Кронрода была мечта — написать программу, которая могла бы понимать неадаптированные тексты. Он формулировал задачу так: в машину нужно вложить текст «Онегина», а программа должна быть такой умной, чтобы отвечать на вопросы «по анкете» Онегина. Речь шла о том, чтобы научить машину выделять смысл из текста.

Когда Евгений Михайлович с Борей Кацем (студентом МГУ), с дочерью — Леной Ландис (тогда школьницей, — но уже не такой маленькой, как на фото, а постарше) и с ее одноклассниками писали программу, сочинявшую стихи, до этого они, конечно, не дошли. Они пошли по более простому пути, и в словарь машины заложили много необыденных, чисто поэтических, слов; машина фабриковала довольно приятные вирши, где часто употреблялось слово «Чу!».

Разносторонне способный человек, Ландис обладал незаурядным художественным талантом. В 60-м году он привез с Кавказа картину «Кронрод, Эльбрус и абрикос». И наверное у многих хранятся их портреты и портреты их детей

Е. М. Ландис с дочерью Леной

работы Ландиса. У меня есть нарисованный им в Тракае портрет моей дочки Иры.

Чуть дальше я покажу страничку из черновиков Ландиса, и там будет набросок к портрету скрипачки — знакомой Евгения Михайловича.

Ландис ценил хорошую шутку. Сам любил шутить, причем способом, нетрадиционным для мехмата. На мехмате умеют шутить, но шутки эти носят часто издевательский характер. А у Ландиса были добрые шутки. Вот любимый анекдот Евгения Михайловича «про короля петербургского бильярда»:

— Короля петербургского бильярда интервьюировали по поводу проливов Босфор и Дарданеллы: как незаметно провести по ним русский флот? Король бильярда спрашивает: «Может быть, ночью?» Нет, нельзя потому-то и потому-то. «А если потихоньку?» Нет, невозможно. «Ну, тогда ничем не могу помочь...»

К Ландису постоянно обращались за советами — самыми разными. Он помогал безотказно. Но иногда, — когда вопрос был уж совсем не по адресу, — чувствуя себя бессильным посоветовать что-нибудь путное, рассказывал этот анекдот.

Многие помнят картинку Ландиса «Операторы в свертках», висевшую на кафедре дифференциальных уравнений. Речь шла об операции свертки *, а Ландис нарисовал девочек-операторов за пультом ЭВМ и свертки, перевязанные ленточками. Ученица Ландиса Лена Ситникова посвятила этой картинке стихи, кончавшиеся так:

Машинное время я трачу не зря,
то в поисках, то в находках.
И лишь операторы мучат меня:
они остаются в свертках.

У Евгения Михайловича была привычка — завязывать на носовом платке узелки на память. Нередко в МГУ он вынимал из кармана платок и с недоумением разглядывал узелки: с вечера на завтра было намечено что-то сделать — но что?! 6 октября 1971 года мы подарили Ландису юбилейный альбом к 50-летию, где с молодым азартом откликнулись на эту привычку. На одной из страниц альбома — под заголовком

«Выписка из протокола заседания юбилейной комиссии»

шли пункты 1, 65 и 50 (именно в таком порядке), в которых содержалось следующее:

п. 1 гласил: «Помнить о том, что один человек не может все помнить» (вторая — чуть переиначенная — часть этого афоризма была хорошо известна Ландису еще со студенческих лет от Саши Лунца: дед Лунца, военный врач, сердился на денщика, который один ботинок поваксил черной, а другой — коричневой краской; денщик оборонялся, говоря, что один человек не может все знать);

в п. 65 давалось обещание — к 65-летию Евгения Михайловича сконструировать полуавтомат «Саморазвязывающиеся узелки»;

в п. 50 говорилось: к 50-летию подарить юбиляру 7 узелков на память и действующую модель полуавтомата «Сигнал».

Подарки были преподнесены вместе с альбомом: «Сигнал» — это были наручные часы, которые звонили, когда нужно; 7 узелков действительно были завязав-

ны на большом носовом платке, и в каждом узелке была бумажка с какой-нибудь полезной информацией, иногда совершенно несусветной.

Например, там был рецепт, как выводить с шелка пятна от пива. Абсолютно равнодушный к алкоголю Ландис очень смеялся, развязав этот узелок и прочитав вложенную в него записку. Еще Ландису чрезвычайно понравилась наша рекомендация: «Если вы поймаете слона за заднюю ногу, а он станет брыкаться, не упорствуйте».

На другой странице нашего альбома Евгений Михайлович обнаружил вот что:

... Хорошо в теплых морях. Особенно на корабле. Особенно между Одессой и Севастополем. Когда море спокойно. И когда еще — на всякий случай — работают успокоители качки ...

Это — «История о том, как Ландис в Индийский океан ходил». Она подлинная, т. е. он и впрямь отправился в такую экспедицию на океанском лайнере «Академик Курчатов». Однако неподалеку от Одессы на корабле (в спокойную погоду!) испытали противокачечные устройства. Получилась полная иллюзия океанского шторма, и — вполне насладившись этим искусственным штормом — Ландис дальше не поехал ... Но, между прочим, за время пребывания на «Курчатове» они с Игорем Фараджевым (работавшим вместе с Ландисом в ИТЭФе) написали для судовой ЭВМ навигационную программу «Штурман», которая использовалась и в том плавании, и после.

Плавание, с точки зрения Евгения Михайловича, было подготовлено из рук вон плохо. Он говорил мне, что «страшно разругался» с человеком (не буду называть его фамилию), ответственным за подготовку экспедиции. На самом-то деле, довольно резко прервав телефонный разговор, Ландис вежливо попрощался и повесил трубку.

А вот, кстати, еще одна история, начало которой тоже совершенно не характерно для неизменно деликатного Евгения Михайловича. Эпизод связан с «леммой Гервера – Ландиса», обобщением лагранжевской теоремы о среднем на случай функций нескольких переменных. Я пришел к Ландису сдавать аспирантский отчет про дивизоры и теорему Римана-Роха. Привыкший к тому, что Евгений Михайлович всегда проявлял неподдельный интерес к собеседнику, на этот раз я с удивлением обнаружил — он явно не слушает меня.

Вскоре дивизоры были отставлены в сторону, и Ландис сообщил мне задачу, над которой он размышлял. Задача мне понравилась, и через пару дней я предложил идею решения: нужно провести много тонких трубочек (интеграл по поверхности каждой равен нулю), и каждую трубочку перегородить в своем месте. Ландис ужасно обрадовался этой конструкции. Так возникла наша «макаронная лемма».

В альбоме к 50-летию были не только шутки. Например, там была фотография Булата Окуджавы с гитарой и подпись: БУДЬ ЗДОРОВ, ШКОЛЯР! Для Ландиса (попавшего на финскую войну практически со школьной скамьи, а потом прошедшего и всю отечественную) и военные песни Булата Окуджавы, и его повесть «Будь здоров, школяр» из «Тарусских страниц» были дороги.

Были в альбоме и 23 ПРОБЛЕМЫ (23 — как у Гильберта!), довольно точно отражавшие разнообразие интересов Ландиса:

1. Почему у человека левая рука такая же, как правая?
2. Как лечить рак?
3. Подготовлен ли эволюцией мозг кошки для решения математических задач?
4. Как устроена память?
5. Математические школы и судьбы их воспитанников.
6. Что такое здравый смысл?
7. Анкета по тексту «Евгений Онегин».
8. Программа «Обобщенная Березанская».
9. Программа «Геометрия».
10. Диагностика.
11. «Кошечка».
12. Как записывать доказательство?
13. Обратная задача сейсмологии.
14. Почему бывают землетрясения и как их предсказывать.
15. Механизм движения континентов.
16. Биогеоценоз.
17. Дзета-функция.
18. Четыре краски.
19. Циклы.
20. Число действий.
21. Создание быстродействующих машин.
22. Какую из оставшихся проблем назвать 22-й?
23. Существуют ли в природе датчики случайных чисел?

Примечание к проблемам 8, 9, 11, 19 и 22. Березанская — автор старого школьного задачника по арифметике, проблема 8 — написать программу, которая решала бы арифметические задачи по текстам условий (хотя бы адаптированным). Проблема 9 — написать программу, которая решала бы школьные задачи по геометрии. Проблема 11 — объяснить, как движется (ходит, бегаёт, прыгает) домашняя кошечка (и, в частности, нарисовать на ЭВМ мультфильм, иллюстрирующий это движение). Циклы в проблеме 19 — это предельные циклы систем дифференциальных уравнений с полиномиальными правыми частями. Наконец, проблема 22 напоминает: как ни разнообразны были интересы Ландиса, все-таки в основном в зрелые годы он занимался уравнениями с частными производными — а это направление в приведенном списке не отражено вовсе . . .

При всей шутливости некоторых формулировок, по-моему, наши 23 проблемы неплохо смотрятся и сейчас . . .

... Теперь — еще одна страничка из альбома на 50-летие. Рукой Ландиса написано, кого (каких *школьников* — участников *студенческого* семинара) Ландис просит пропустить на мехмат 13 сентября 1961 года. В списке — по алфавиту: Иосиф Бернштейн, Саша Геронимус, Ленья Лиманов (это Леонид Макара-Лиманов, фамилия сокращена), Гриша Маргулис, Боря Пранов, Волик Фишман. Все они стали математиками, некоторые — всемирно известными; Саша Геронимус сделался впоследствии священнослужителем, Гриша Маргулис стал филдсовским лауреатом, Волик Фишман, автор τ -метода в математической геофизике, к несчастью, умер совсем молодым.

Отмечу, кстати, что все перечисленные школьники были участниками школьного математического кружка «альфа», который вел Ник. Ник. Константинов. В «альфа», а потом в «бета» ходили и другие сильные кружковцы, например, Сеня Вишик.

Обращаю внимание на связь этого сюжета с 5-ым пунктом (проблемой математических школ и их выпускников).

... В альбоме на 50-летие мы слегка легкомысленно дали обещание сконструировать к 65-летию Ландиса полуавтомат «Саморазвязывающиеся узелки». И вот, прошло 15 лет, и надо было обещание как-то выполнять.

Ну что ж, мы придумали, — как. Во-первых, мы обзвонили друзей и знакомых юбиляра и попросили каждого послать Ландису поздравительную открытку — с обязательным условием: там должен был присутствовать «узелок на память». Каких только узлов и узелков не принесли почтальоны Евгению Михайловичу в те дни: здесь были и морские узлы, и узлы — особые точки дифференциальных уравнений... Открытки приходили несколько дней, пачками — это было здорово!

Во-вторых, мы преподнесли Ландису деревянный меч и спицы с пояснительным текстом:

Одни узлы рубить с плеча сподручней с помощью меча,

А для других меч не годится — их развязать помогут спицы.

И, наконец, мы приготовили новый альбом с фотокарточками самого Ландиса, его друзей и учеников. Это был, собственно, не альбом, а тетрадь для рисования. На обложке мы написали:

Тетрадь
ученика 1-го класса Московской математической школы
Е. М. Ландиса.

Ландис оценил эту надпись слету: недаром он был математиком *первого класса*. Здесь уместно вспомнить некоторые его сентенции про работы разных людей. Я помню, как про Адельсона-Вельского он очень серьезно сказал: «Это

математик милостию Божией». А про некоторые работы говорил: «Не диссертация Римана». Или: «Не хуже других»...

Альбом к 65-летию Ландиса открывался большим портретом Евгения Михайловича, и под ним было написано: «Поздравляем!» Потом шли две уникальные фотографии: молодые Кронрод и Ландис (вскоре после возвращения с фронта) и курс Ландиса у старого здания университета (конец 40-х годов).

Еще один портрет Ландиса располагался на странице с заголовком

«Аспиранты, защитившие диссертации.»

Там рукой Ландиса были выписаны 26 фамилий — слева перечислялись 23 кандидата, справа — 3 доктора: Худай-Веренов, Гервер, Новрузов. Впоследствии появился дополнительный список — слева добавилось еще 9 кандидатов, справа — 6 докторов: Ибрагимов, Мамедов, Григорьян, Надирашвили, Черемных, Ильяшенко.

Вот картинка, которую я уже упоминал раньше. Эту «почеркушку» мы включили в альбом так же, как пушкинисты включают в собрания сочинений страницы рукописей с рисунками на полях. Только Ландис рисовал не на полях, а на листке с каким-то случайным печатным текстом — просто потому, что там еще оставалось много свободного места.

Видите, кроме эскиза к портрету знакомой скрипачки, там фамилии: Арнольд, Олейник, Ильин, Вайнберг... А ниже интересная загадка:

$$\begin{array}{ccccccc} & & 5 & 7 & 5 & & \\ 7 & 9 & 5 & 9 & 19 & 9 & 5 & 9 & 7, \end{array}$$

что бы это значило — что это за странные числа? А это — число труб органа в Московской консерватории — столько, сколько их видно из зала.

Среди прочих в альбоме были и мои 2 фотографии: я на Дальнем Востоке, на фоне огромной океанской волны, и я на вертикальной стене (я был тогда тоненький и умел лазать по стенам). И под ними — обращение к Ландису:

Когда захлестывает жизнь меня огромною волною
или отвесною стеною перегораживает путь,
стремлюсь смятенною душою
я к Вам (а не к кому-нибудь).

В каждой шутке есть доля шутки, но теперь — через много лет — я говорю эти слова еще серьезней, чем прежде...

Из выступлений на заседании Московского математического общества, посвященном памяти Е. М. Ландиса

М. И. Вишик: Я был другом Евгения Михайловича Ландиса. Мы встретились с ним в 40-х годах. Он вернулся из армии, в гимнастерке, на старый мехмат. Его математическое дарование сразу было замечено, и я почувствовал необыкновенного человека в нем. С тех пор наша дружба продолжалась. Мы работали сначала в разных институтах, потом на мехмате, на кафедре дифференциальных уравнений, и всегда вместе обсуждали все возникающие вопросы.

Евгений Михайлович, как сегодня было видно из докладов, был выдающимся математиком, очень талантливым. В его работах всегда чувствовалось стремление решать трудные и глубокие задачи. Он не любил простых задач и не стремился к тому, чтобы у него было много работ. Он очень ценил новые направления, новые мысли, что-то новое, с изюминкой, чего раньше не было, т. е. то, что как раз характерно для нашей московской математической школы.

Евгений Михайлович, кроме того, что был выдающимся математиком, был действительно большим человеком. Надо сказать, что мы с ним часто разговаривали на разные математические темы и, конечно, консультировались по другим вопросам, которые у нас возникали.

У нас в Москве была тогда большая, выдающаяся математическая школа. Консультацию по возникающим математическим вопросам можно было получить в личном разговоре, или по телефону, или на мехмате. Я помню, что когда что-то нужно было узнать по дифференциальным уравнениям, мы обращались к Евгению Михайловичу. Он, когда к нему обращались с вопросом, очень это серьезно воспринимал. Помню, я перед ним поставил задачу, связанную с нелинейными уравнениями второго порядка. Он очень обрадовался, когда ее решил. Позвонил, и сам прибежал ко мне домой, и рассказал решение этой задачи. Вот тут виден Евгений Михайлович, как он относился к другим, когда его о чем-то спрашивали, и насколько он готов был всегда помочь, и как-то ответить на вопросы, которые ему задавали.

Мы, когда преподавали на мехмате, на кафедре дифференциальных уравнений, часто совещались по телефону, как читать курс уравнений с частными производными, как читать спецкурсы и т. д. У нас по телефону образовался такой своеобразный семинар. Мы в начале учебного года обычно обсуждали то, как будем читать свои курсы, и затем, когда возникали в течение года другие вопросы, мы тоже их с ним обсуждали. И всегда это было для меня очень полезно.

Мы часто вспоминали слова Ивана Георгиевича Петровского, который говорил нам, как надо читать: «Не читайте много, а читайте внятно». Особенно часто мы их вспоминали, когда начинали спорить, что лучше читать. Мы не укладывались по времени, потому что программа курса дифференциальных уравнений с частными производными была очень обширной. (И действительно, это видно из книги Ивана Георгиевича и других пособий, что трудно было уложиться, основательно излагая весь материал.) Когда мы обсуждали вопрос, что можно опустить, а что нет, Евгений Михайлович, вспоминая эти слова, говорил, что не так важно прочитать подробно, а важно прочитать так, чтобы это дошло, и воспитывало аудиторию должным образом. Вот тут виден Евгений Михайлович, как последователь Ивана Георгиевича, виден его глубокий и вдумчивый подход.

Евгений Михайлович Ландис всегда вел не только основной семинар (впоследствии он вел его с Владимиром Александровичем Кондратьевым), но также все время (и этому он меня научил) вел просеминар для студентов, начиная с 1-го и 2-го курсов. На просеминаре Евгений Михайлович предлагал простые задачи, потом обсуждал решения этих задач, — и очень много сам рассказывал из разных областей математики. Это могли быть не только дифференциальные уравнения. Это была теория функций, о которой сегодня много говорилось, теория функций комплексного переменного, геометрия, дифференциальная геометрия — целые разделы из этих областей математики он излагал. Молодые люди из года в год росли, многие из них поступали к Евгению Михайловичу в аспирантуру, становились учатниками его основного семинара, а он сам начинал все сначала на новом просеминаре для студентов первых курсов. Таким образом создавалась математическая школа Е. М. Ландиса.

Евгений Михайлович был настолько внимателен к молодым людям, что это стало уже известно всей нашей молодежи, и очень много студентов записывалось к нему как к научному руководителю, чтобы учиться у него. Им созданная школа получила мировое признание, а сам Евгений Михайлович, как руководитель этой школы, стал образцом внимательного и чуткого руководителя молодых людей.

Я хотел бы сказать также, что Евгению Михайловичу был присущ художественный талант. Он любил и подмечал все красивое. Известно, что он умел хорошо рисовать и даже писать маслом. Когда нужно было сделать кому-нибудь юбилейное послание, то всегда на кафедре просили, чтобы Евгений Михайлович нарисовал портрет юбиляра, что он и делал с большой охотой.

Я должен сказать, что Евгений Михайлович никогда не делал замечаний каким-то людям. И он навсегда остался в нашей памяти человеком очень добрым. У нас всегда перед глазами его улыбка, а в памяти — его благожелательность. И этот образ Евгения Михайловича является для нас, в каком-то смысле, неповторимым образцом человека, выдающегося математика, очень талантливого, очень глубокого и вместе с тем очень доброго и благожелательного ко всем тем, с кем он имел дело. Все, что он создал в математике, безусловно, останется навсегда в науке, а его образ останется всегда с нами, на всю жизнь.

В. Л. АРЛАЗАРОВ: Это было в 60-е годы. Евгений Михайлович работал в ИТЭФ, в математической лаборатории, которой заведовал А. С. Кронрод. Он был профессором, старшим научным сотрудником. Задачи появлялись у физиков, а математики должны были их как-бы «обматемачивать», и это не сводилось

к тому, чтобы решить правильно поставленную задачу (как раз с этим физики часто сами неплохо справлялись). В сложных случаях хороший результат получался только, если математик и физик ставили задачу совместно. Важнейший тип задач, которые решались в ИТЭФе, это — обсчет траекторий частиц в водородных камерах, выяснение по траекториям, какие реакции происходили при столкновении частиц. Программы для этих задач были довольно сложные. И Евгений Михайлович проявлял при этом ту самую изобретательность, которую мы видели у него в чисто математических работах. Этими программами, которые делались в 63-м, 64-м годах, продолжали пользоваться и в 80-х. Это фантастический срок жизни для программ.

Работа об AVL-дереве до сих пор входит в любой достаточно солидный учебник информатики. Смешно, что и у нас теперь эти буквы пишут по-латински, хотя AVL — это Адельсон-Вельский и Ландис.

Мы тогда занимались программными задачами, в которых надо было запоминать числа и другие объекты, и потом их же и искать. Можно запоминать объекты в порядке поступления, но тогда долго будет их искать. А можно запоминать по-хитрому. Конечно, лучше всего запоминать объекты, которые упорядочены. Тогда поиск занимает ровно логарифм от общего числа объектов. Но тут возникает трудность: искать легко, а добавить новый объект так, чтобы не разрушить порядок, трудно. В те годы появлялось много работ о том, как устроить такую систему, чтобы было и искать быстро, и вставлять быстро — мы называли такие системы «текущими справочными». На семинаре Кронрода, в котором мы все участвовали, кто-то из аспирантов рассказал американскую работу, в которой «справочные» строились в виде дерева. Впрочем, и до этого представлять «справочную» иначе, чем в виде дерева, никому и в голову не приходило. А в таком дереве очень хорошо искать.

На семинаре нам рассказали, что если все исходы равновероятны, то, вообще говоря, гулять по такому дереву, и его строить, и добавлять можно достаточно быстро. Глубина дерева будет, как правило, равняться логарифму от числа вершин, и эта же величина будет равняться количеству операций, уходящих как на поиск, так и на перестройку дерева.

Евгений Михайлович сказал, что если в среднем это так, то должен быть способ построить дерево, чтобы всегда это было так. И он бродил по лестницам, коридорам и переходам нашего здания, думал и время от времени говорил: «О, уже придумал?» За ним ходил Адельсон-Вельский, и каждый раз, когда Ландис что-то придумывал, он догонял его и строил контрпример. И так они ходили две недели друг за другом и придумывали: Ландис — алгоритм, Адельсон — противоречащий пример. Под конец у обоих создалось впечатление, что они дошли до таких алгоритмов, для которых контрпример придумывать все труднее и труднее. И после этого они вдвоем построили правильный алгоритм. Сейчас это можно рассказать за три минуты, так это просто и красиво. Рассказать так, чтобы было видно, что по-другому и быть не может.

У Ландиса была такая присказка, такое любимое выражение, что математика можно узнать по тому, сколько времени он согласен думать над одной задачей. Если он согласен думать достаточно долго, то он либо уже математик, либо потенциальный математик. А если нет, то и говорить не о чем.

Н. Н. Константинов: Я познакомился с Ландисом в 50-м году, на втором курсе, когда занимался в семинаре, который Кронрод с Ландисом вели для второкурсников. Семинар был, в основном, по теории функций действительного переменного, хотя часто выходил за эти рамки. У Евгения Михайловича была тогда любимая тема «Дескриптивная теория множеств и функций». Иногда на семинаре бывали моменты, когда одна задача кончилась, а другая не началась. И тогда Ландис говорил: «Вот я хочу рассказать про дескрипцию».

Позже мне для аспирантского экзамена нужно было изучить статью А. А. Ляпунова про A -множества в Успехах Математических Наук. Она была написана довольно сложно. Ландис мне тогда рассказал удивительно красиво и просто основной стержень этой теории — эквивалентность трех определений A -множества. Мне это очень понравилось. Помню, как Ландис охарактеризовал эту теорию. Она, говорил он, красива и изящна, как музыка Моцарта. Но подумав, сказал: «Нет, пожалуй, музыка Моцарта слишком серьезна для такого сравнения. Скорее — как музыка Россини».

В семидесятые годы группа молодых московских математиков, в основном, преподавателей математических школ, организовала ежегодный летний трудовой математический лагерь в Эстонии. Однажды туда приехал Е. М. Ландис с дочерью Леной. Они жили в палатках, как и все. Жизнь в палатке Евгения Михайловича нисколько не смущала. Было ощущение, что он воспринимает это как нечто естественное. В лагере было довольно много математических занятий, и на одном из них Ландис прочитал для школьников лекцию: «Применение золотого сечения в библиотечном деле». Речь шла о сбалансированном дереве — недавней тогда работе Г. М. Адельсон-Вельского и Е. М. Ландиса. Надо сказать, что эту лекцию не очень-то поняли школьники. Единственный, кто ее с первого раза понял, был Гриша Рыбников, которого, я думаю, многие знают. Говорят, что эту работу можно рассказать за три минуты. Может быть, это и верно, но все-же, видимо, желательно, чтобы у слушателя была к этому времени серьезная подготовка, например, высшее математическое образование.

Отмечу интересную деталь. Колхоз дал нашему лагерю лошадь и телегу, чтобы мы могли ездить за продуктами. И тут выяснилось, совершенно неожиданно, что Ландис большой специалист по лошадям — он разбирался в них не хуже колхозных конюхов. А дело в том, что Ландис во время войны был командиром артиллерийского взвода. Тягловой силой были лошади. И в период некоторого затишья основная забота была о лошадях. Ну, и волей-неволей Ландису пришлось стать большим специалистом по уходу, подковыванию и чистке. Кстати, школьник, которому в нашем лагере было поручено работать с лошадью — кормить, запрягать и ездить за продуктами — имел звание «клячмейстер», и это очень забавляло Ландиса.

Однажды мы с Ландисом переходили улицу. Улица была совсем пустая, но где-то ехала машина. Я говорю: «Пойдем, успеем». «Да, конечно, успеем, но у меня есть привычка пропускать машину. Я не хочу эту привычку разрушать». Такой подход показался мне интересным. И я подумал, что ведь, надо же, человек прошел две войны, и остался целым. Ну, понятно, что здесь, может быть, везения 90%, но еще сколько-то и умения быть осторожным.

Однажды Ландис, будучи студентом, должен был сдать спецкурс Л. А. Люстернику. Они встретились в назначенное время, Люстерник задал вопрос и куда-то вышел. Вскоре он вернулся, но не начал экзамена, а стал разговаривать с Ландисом о разных посторонних вещах. Это длилось около часа. Наконец Люстерник нетерпеливо посмотрел на часы и сказал: «Ну куда же он пропал?». «Кто-то должен подойти?» — спросил Ландис. «Ну конечно, где же тот студент, которого мы с вами должны экзаменовать?». Получился очередной анекдот про Люстерника. А дело в том, что Ландис, во-первых, выглядел намного старше большинства однокурсников, а во-вторых, он постоянно выступал на научных семинарах, и в нем привыкли видеть взрослого математика.

Л. А. Кронрод: Из всех присутствующих здесь, наверное, я больше всех по времени знаю Геню Ландиса. Мы с ним поступили в университет в 1939 году. Но наших мальчишек в 39-м, в сентябре, забрали в армию. И Геня стал учиться в университете уже после войны.

Я была жена (теперь называюсь вдова) Кронрода А. С. Мы с Геней были очень близки, в каком-то смысле даже суеверно близки. День рождения Гени Ландиса 6 октября, А. С. Кронрод умер 6 октября. Были и другие полные совпадения.

Одна совместная работа Кронрода и Ландиса осталась неизвестной, потому что она нигде не публиковалась. Они вдвоем написали очень хороший учебник для школьников и первокурсников по теории функций действительного переменного. Этот учебник хранится у меня дома, и я думаю, что когда-нибудь его опубликую.

Из воспоминаний Евгения Михайловича о войне я могу рассказать один эпизод. Как известно, он был в расчете, где была пушка и лошади. Однажды немцы наступали и уже вошли в одну деревню. А пушка наших стояла на краю деревни. Наши спрашивают у своего начальства: «Что нам делать с пушкой?». И вот, наконец, получают приказ взорвать пушку. Ее взорвали. «А потом, — говорит Геня, — я ничего не помню». Его контузило при этом взрыве. Когда он очнулся, он увидел, что лежит у обочины. Вокруг стонут раненые, а немцы идут по дороге и их пристреливают. Геня притворился мертвым. И если бы не было этой контузии, Геня Ландис был бы сейчас с нами. Его здоровье было очень сильно подорвано этой контузией.

А. Д. Мышкис: Я знаю Евгения Михайловича с 1946 г. Тогда я был ассистентом на кафедре дифференциальных уравнений мехмата МГУ и вел практические занятия, а Евгений Михайлович был студентом второго курса. Я уже слышал от моего друга А. С. Кронрода, чрезвычайно яркого человека и талантливого математика, что у него появился весьма способный ученик, который подключился к его научной работе. Евгений Михайлович был даже старше, чем Александр Семенович (на 16 дней), но из-за войны задержался с учебой и, таким образом, на первом этапе оказался учеником Александра Семеновича.

Авторитет Евгения Михайловича среди студентов был общепризнан. Например, помню такой эпизод. Лектор в курсе дифференциальных уравнений доказывал теорему Пеано по методу Перрона, но как-то очень уж громоздко. Евгений Михайлович придумал упрощение этого доказательства и рассказал об

этом своим однокурсникам. Когда я принимал экзамен, одна студентка, которой попался этот вопрос, спросила: «А можно я буду рассказывать по Ландису?» Получив согласие, она хорошо изложила доказательство.

После публикации работ в области «чистой» ТФДП, Евгений Михайлович довольно быстро перешел к дифференциальным уравнениям — особенно, с частными производными. Это направление и стало в дальнейшем для него преобладающим. Он поступил в аспирантуру к И. Г. Петровскому. Помню, как Иван Георгиевич, понимая формальный характер этой процедуры, поручил О. А. Олейник и мне принять у Евгения Михайловича аспирантский экзамен по теории уравнений с частными производными. А мы к тому времени были уже все трое друзьями и давно перешли на «ты». Мы с Ольгой Арсеньевой задавали вопросы, стараясь «подловить» Евгения Михайловича на чем-нибудь, но долго никак не могли этого сделать. Однако нападающих было двое против одного, и в конце концов мы поймали его на тонком вопросе, связанном с характером изолированных особых точек гармонической функции, чем и были полностью удовлетворены.

Отмечу, что у Ивана Георгиевича учеников было мало, а сумевших проявить себя в математике — и совсем мало. Сам он как-то сказал: «У меня только человек пять вышло в люди.» — и, конечно, Евгений Михайлович был одним из них. Иван Георгиевич не был руководителем в узком смысле, он давал только самые общие установки, предоставляя своим ученикам полную свободу; скорее, он старался передать свои взгляды на различные вопросы математики и жизни. Такой метод хорош только для более сильных аспирантов. Отчасти поэтому, а отчасти и из-за своих выдающихся способностей, Евгений Михайлович довольно быстро стал самостоятельным ученым.

Надо сказать, что Евгений Михайлович Ландис, будучи студентом, довольно быстро перешел к исследовательской деятельности. У него были серьезные работы, совместные с Кронродом. Потом он довольно быстро переключился на уравнения с частными производными и стал учеником И. Г. Петровского.

В то время на мехмате активно работал замечательный семинар по теории уравнений с частными производными, которым руководили С. Л. Соболев, И. Г. Петровский и А. Н. Тихонов (обычно их перечисляли в таком порядке, по росту). Это был центр, вокруг которого группировались исследователи не только из Москвы, но и из других городов. Наряду с корифеями, в нем активно работал дружный коллектив молодых математиков: М. И. Вишик, О. А. Ладженская, О. А. Олейник и другие. Это был романтический период: мы активно работали, немедленно делились друг с другом результатами, обсуждали их, старались помочь друг другу.

Евгений Михайлович быстро вошел в этот коллектив, сохраняя в работе свой характерный стиль, о котором я уже упомянул. Вообще, я думаю, что в мировой математике среди специалистов по теории уравнений с частными производными трудно указать людей, столь совершенно владеющих качественными методами и умеющими их эффективно применять, как Евгений Михайлович.

Евгений Михайлович отличался ровным, доброжелательным характером, он был истинно интеллигентным человеком, я ни разу не слышал, чтобы он повысил голос или кому-нибудь строго выговаривал. Светлый облик Евгения Михайловича всегда в моей памяти.

Феликс Клейн и его эрлангенская программа

(к 150-летию со дня рождения ученого)

Н. Х. Розов

Эрлангенская программа открыла в понятии группы преобразований те узлы, которые скрепляют все разновидности геометрии и вместе с тем определяют отличительные особенности каждой из них; в этой программе вопрос «что такое геометрия?» ставится и получает основополагающий ответ.

Герман Вейль

В истории геометрии не много имен и событий, принципиально изменивших наши представления об этой ветви математики и радикально повлиявших на ее дальнейшее развитие. К числу таковых безо всяких сомнений следует отнести Феликса Клейна и его эрлангенскую программу.

XIX век с полным правом можно назвать веком расцвета геометрии. Именно к этому веку относятся выдающиеся геометрические работы К. Гаусса, Б. Римана, Э. Бельтрами, А. Пуанкаре... Но даже на фоне всех их блистательных достижений два события оказались поистине эпохальными: открытие неевклидовой геометрии Гауссом, Лобачевским, Бойаи и разработка Клейном теоретико-группового подхода к геометрии.

В октябре 1872 года молодой немецкий математик Феликс Клейн (1849 – 1925) вступал в должность профессора на философском факультете Эрлангенского университета. Согласно существовавшей в те времена (и сохраняющейся поныне в ряде университетов Германии) традиции, новый профессор должен был прочитать вступительную лекцию, в которой надлежало изложить программу его научных исследований. Этой лекции Клейна, озаглавленной «Сравнительное обозрение новейших геометрических исследований» (и сразу же выпущенной в виде небольшой брошюры), выпала завидная доля навсегда занять почетное место в истории математической науки. Она известна ныне под именем эрлангенской программы. Известный американский геометр Джулиан Кулидж сказал как-то, что эрлангенская программа Клейна повлияла на геометрическое мышление больше, чем любая другая работа со времен Евклида, за исключением работ Гаусса и Римана.

(Лекция Ф. Клейна на русском языке впервые опубликована в 1895 г. Ее полный текст был помещен также в малодоступной сейчас замечательной книге

из серии «Классики естествознания»¹). С тех пор, насколько нам известно, на русском языке эрлангенская программа больше не переиздавалась.)

До Ф. Клейна разнообразные геометрические факты и теории не объединялись единой концепцией. Сам Ф. Клейн так писал об этом:

«Геометрия, единая по существу своему, при быстром своем развитии в последнее время слишком уж раздробилась на ряд почти раздельных дисциплин, которые продолжают развиваться в значительной степени независимо друг от друга.»

Универсальную концепцию «единства геометрий», общий принцип, охватывающий все ее ветви, впервые увидел и четко сформулировал Феликс Клейн. В основе его концепции лежит идея, связанная с понятием группы преобразований и классификацией групповых свойств. Это фундаментальное открытие представляется тем более замечательным и удивительным, что сама по себе теория групп как отдельный раздел алгебры в то время фактически еще не сложилась.

Вот как формулирует Ф. Клейн в эрлангенской программе центральную проблему геометрии:

«Дано многообразие и в нем группа преобразований; нужно исследовать те свойства образов, принадлежащих многообразию, которые не изменяются от преобразований группы. . . Требуется развить теорию инвариантов этой группы.»

Итак, ветви геометрии различаются инвариантами соответствующей группы преобразований. Г. Вейль писал:

«Эрлангенская программа . . . достигает своего полного проявления именно в теории групп линейных преобразований и их инвариантов.»

При этом имеется своеобразная иерархия инвариантов: в группе движений евклидовой плоскости инвариантом связаны две точки, в группе подобия — три точки в общем положении, в группе аффинных преобразований — три коллинеарные точки, в группе проективных преобразований — четыре коллинеарные точки . . . Таким образом, на основе группового подхода оказываются объединенными такие, казалось бы, различные главы геометрии, как евклидова, аффинная и проективная геометрии, геометрии Лобачевского и Римана и т. д. Мы не будем здесь более подробно углубляться в принципиальные научные детали концепции Ф. Клейна, рекомендуя читателю (как опытному, так и начинающему) познакомиться с ними по недавно появившейся блестящей книге В. В. Прасолова и В. М. Тихомирова²).

Идеи Ф. Клейна были восприняты многими выдающимися математиками. Так, в 1886 г. Софус Ли (1842 – 1899) в своей статье «Замечания на работу Гельмгольца „О фактах, лежащих в основании геометрии“» писал:

«Знаменитая работа Гельмгольца . . . рассматривает задачу, стоящую в теснейшей связи с новой теорией групп преобразований. Побуждаемый Клейном, я поэтому решился применить к этой . . . задаче методы моей теории преобразований.»

¹ Об основаниях геометрии. Сборник классических работ по геометрии Лобачевского и развитию ее идей. – Редакция и вступительная статья А. П. Нордена. – М.: Гостехиздат, 1956. 528 с.

² Прасолов В. В., Тихомиров В. М. Геометрия. М.: МЦНМО, 1997. 352 с.

(Упомянутая работа Германа Гельмгольца (1821 – 1894) была опубликована в 1868 г. и представляла собой одну из первых попыток продвинуться по пути аксиоматического обоснования геометрии.) Через 15 лет после публикации эрлангенской программы мысль о равноправности геометрий Евклида и Лобачевского становится отчетливо ясной. Анри Пуанкаре (1854 – 1912) в своем сочинении «Об основных гипотезах геометрии» (1887) выразил ее «по Клейну» так:

«... Геометрия есть не что иное, как изучение некоторой группы движений, и в этом смысле можно сказать, что справедливость геометрии Евклида нисколько не противоречит справедливости геометрии Лобачевского, так как существование одной группы вполне совместимо с существованием другой.»

Следующий шаг в развитии геометрии сделал Эли Картан (1870 – 1953). В его трактате «Теория групп и геометрия» на основе синтеза идей Клейна и Римана развиваются новые направления в геометрии. Представляется уместным воспроизвести здесь высказывание Э. Картана, определяющее суть открытия Ф. Клейна:

«... Основная идея Ф. Клейна может быть связана с наиболее древними понятиями науки. Элементарная геометрия изучает свойства фигур, не зависящие от их положения в пространстве, ... остающиеся инвариантными относительно некоторой совокупности преобразований, образующей группу движений ... Проективная геометрия, ... с точки зрения Клейна, изучает свойств фигур, инвариантные относительно некоторой совокупности ... проективных преобразований, образующих группу. Вообще всякая группа непрерывных преобразований определяет самостоятельную геометрию. Эта геометрия ... изучает свойства фигур, инвариантные относительно преобразований группы ...»

Особую роль в эволюции геометрии XIX столетия сыграло развитие проективной геометрии. Эрлангенская программа открывается следующей фразой:

«Между приобретениями, сделанными в области геометрии за последние пятьдесят лет, развитие проективной геометрии занимает первое место.»

В ней не следует усматривать недооценку значения открытия геометрии Лобачевского (напомним, что труд «О началах геометрии» Н. И. Лобачевского вышел в свет в 1829 г., за 43 года — менее, чем за полвека — до речи Ф. Клейна). Тем более, что именно Ф. Клейн впервые научно обосновал независимость евклидова пятого постулата. В 1871 г. вышла его статья «О так называемой неевклидовой геометрии», в которой он так описывает историю рождения новой геометрии:

«Гаусс назвал эту геометрию неевклидовой; он много занимался ею, но, к сожалению, ничего не опубликовал о ней, кроме некоторых намеков. К этой же неевклидовой геометрии пришли Лобачевский, профессор математики Казанского университета, и, несколькими годами позже, венгерский математик Я. Больyai.»

(Кстати, уже в этой статье Ф. Клейна можно усмотреть явные попытки связать геометрические свойства со свойствами линейных преобразований, именно их положить в основу геометрических рассуждений.) Как известно, Ф. Клейну

принадлежит проективная интерпретация неевклидовой геометрии (плоскости Лобачевского). В 1889 – 1890 гг. он читал в Геттингенском университете лекции по неевклидовой геометрии; возможно, сегодня не все помнят, что эти лекции издавались и на русском языке³⁾.

Нисколько не умаляя значения поворотного в истории науки открытия неевклидовой геометрии, следует согласиться, что проективная геометрия является одной из наиболее принципиально важных и неожиданно причудливых математических теорий. Есть все основания предполагать, что к своей концепции Ф. Клейн пришел в результате обстоятельных и плодотворных исследований в этой области. Нетрудно заметить, что сама эрлангенская программа в определенном смысле является гимном проективной геометрии, бурно развивавшейся в XIX веке.

Следует констатировать, что в настоящее время проективная геометрия отодвинулась на второй план. И вот результат: на многих математических факультетах университетов систематическое изложение проективной геометрии (как, впрочем, и высшей геометрии) давно уже не дается, а для беглого ознакомления с ее базовыми понятиями выделяются две – три лекции в курсе аналитической геометрии. Это очень жаль, поскольку проективная геометрия и ныне могла бы иметь исключительно важное образовательное значение для математиков (и даже общеобразовательное значение для гуманитариев). Например, в упомянутой выше книге В. В. Прасолова и В. М. Тихомирова рассказывается о том, какую роль сыграла проективная геометрия в научной жизни А. Н. Колмогорова.

Эрлангенская программа представляет собой блестящий образец того, как можно и нужно доступно излагать глубокие мысли. Это произведение, как и многие работы классиков науки второй половины XIX – первой половины XX веков, содержит какой-то непередаваемый аромат личной неторопливой беседы с читателем, красочных ассоциаций и попутных замечаний, неожиданных аналогий и тонких намеков, сравнительного анализа фактов, аргументов, соображений. При изложении органично описывается история научного поиска и специфика математического мышления, детально рассказывается о предшественниках и путях оформления идей, раскрывается процесс математического познания.

Ф. Клейн как бы ведет доверительную спокойную беседу со своими коллегами, сопоставляет результаты, обсуждает точки зрения. С. Ли и Р. Клебш, Л. Гессе и К. Жордан, А. Кэли и Э. Лагерр, М. Шаль и Ю. Плюккер, К. Штаудт и Б. Рейнхардт, Г. Грассман и А. Бриль — все они, сегодня известные каждому или почти забытые, — были современниками или почти современниками, друзьями и заинтересованными понимающими людьми. Вот один только пример — дополнение, которое Ф. Клейн посчитал обязательным сделать уже позже, при переиздании:

«Завершая эту перепечатку эрлангенской программы, я охотно укажу еще на работы Мебиуса (которые я сам осмыслил в их внутренней связи лишь после того, как в 1885 – 1887 годах принял участие в издании собрания его сочинений. . .). Мебиус еще не знал общего определения группы, а также многих геометрических преобразований . . . ; однако он, руководствуясь верным чутьем, расположил свои следую-

³⁾ Клейн Ф. Неевклидова геометрия. М.–Л.: ГТТИ, 1936. 355 с.

щие друг за другом геометрические работы именно так, как это соответствует основным идеям программы.»

В таком же удивительном стиле теплого рассказа о личностях математиков и обстоятельного изложения их идей написана и книга «Лекции о развитии математики в XIX столетии»⁴⁾, которую по праву можно было бы назвать «историей математики с человеческим лицом».

Эрлангенская программа, помимо своего чисто научного значения, оказала заметное влияние на развитие геометрического образования. Выработанная Ф. Клейном концепция геометрической науки во многом определяют построение и содержание программ курсов геометрии в высшей школе.

Ф. Клейн был не только выдающимся ученым, но и талантливым популяризатором математической науки, блестящим педагогом, глубоким и оригинальным исследователем проблем методики преподавания математики в средней школе, последовательным реформатором всего математического образования. Достаточно напомнить, что замечательная книга «Элементарная математика с точки зрения высшей»⁵⁾, представляющая собой курс его лекций для будущих учителей математики, и сегодня остается классическим руководством для каждого, кто пожелает всерьез разобраться в научных основах «школьной математики». Эта книга по своему содержанию, манере подачи материала, стилю изложения и сегодня могла бы оказать неоценимую помощь всем, кто занимается математическим образованием школьников, и очень жаль, что в учебном процессе наших педагогических вузов она не занимает достойного места.

Прекрасной иллюстрацией всемирного признания авторитета Ф. Клейна как педагога и высочайшей оценки его работ, касающихся преподавания математики, является следующий исторический факт. 90 лет тому назад, в апреле 1908 г., на IV Международном конгрессе математиков, проходившем в Риме, была образована Международная комиссия по преподаванию математики. И ее первым Президентом стал именно Феликс Клейн.

Впрочем, в эрлангенской программе педагогическим проблемам автор не уделяет особого внимания. Тем более интересны несколько имеющихся там замечаний, которые в какой-то мере раскрывают точку зрения Ф. Клейна на отдельные вопросы преподавания. Он весьма определенно высказывается о важности при обучении опираться на историю науки, указывая, что следовать тому пути,

«по которому шла наука в своем развитии, . . . при изложении является обычно самым выгодным».

Впоследствии он еще точнее сформулировал эту мысль:

«. . . Преподавание математики, как и вообще всякое преподавание, . . . должно идти по тому же самому пути, по которому все человечество, начиная со своего наивного первобытного состояния, дошло до вершин современного знания! . . . Научно обучать — значит учить человека научно думать, а не оглушать его с самого начала холодной, научно напряженной систематикой. Существенное препятствие к распространению такого естественного и поистине научного метода

⁴⁾ Клейн Ф. Лекции о развитии математики в XIX столетии. М.: Наука, 1989. 456 с.

⁵⁾ Клейн Ф. Элементарная математика с точки зрения высшей. М.: Наука, 1987. Т. I. «Арифметика, алгебра, анализ». 431 с. Т. II. «Геометрия». 416 с.

обучения представляет собой, несомненно, недостаток в знакомстве с историей математики.»

Ф. Клейн совсем бегло затрагивает в эрлангенской программе и один из центральных вопросов методики преподавания математики — вопрос о гармоническом соотношении между наглядным, интуитивным и формальным, аналитическим. Следуя логике профессионального ученого-математика, он совершенно справедливо замечает, что

«математический предмет нельзя еще считать окончательно разработанным, пока он не стал логически очевидным».

И в то же самое время со всей убежденностью опытного преподавателя подчеркивает, что наглядность

«в педагогическом отношении нужно ценить очень высоко. Геометрическая модель, например, с этой точки зрения очень поучительна и интересна».

Позже Ф. Клейн не раз возвращался к этому вопросу, обстоятельно объясняя, что конкретно он понимает под принципом наглядности и как видит его реализацию в процессе преподавания.

В своих работах Ф. Клейн подробно и обстоятельно развивает оригинальные взгляды на математическое образование. В их основе лежат такие фундаментальные положения, как обязательность приспособления к «природным склонностям молодого человека» (выражение Ф. Клейна) и постоянной ориентации на психологические закономерности усвоения математических знаний, необходимость подлинного усвоения идеи функциональной зависимости и глубокого развития функционального мышления, целесообразность реализации фузионизма в преподавании школьной математики, прежде всего — школьного курса геометрии, полезность широкого ознакомления с приложениями математики и т. д. Например, он выступал за свободный выбор учеником гуманитарного или естественно-научного и математического образования и тем самым стоял фактически у истоков идеи гуманизации образования. Многие его основополагающие идеи и концепции перестройки среднего и высшего математического образования хорошо известны, общепризнаны, сохраняют злободневность и в наше время — но все еще очень далеки от воплощения в жизнь.

И в заключение — еще два интереснейших высказывания из эрлангенской программы. Весьма актуально звучит призыв ее автора расширять связи между различными разделами математики и тем самым обогащать возможности творчества и углублять понимание результатов:

«Специалист по математической физике постоянно уклоняется от тех преимуществ, которые может доставить ему во многих случаях сколько-нибудь выработанное проективное воззрение.»

Идея необходимости тесного взаимодействия и взаимопонимания различных ветвей математики, математиков и прикладников всегда была очень близка Ф. Клейну и всячески им пропагандировалась. Очень глубоким представляется и другое положение, касающееся взаимоотношения научного и психологического в геометрическом восприятии:

«Нельзя с математической точки зрения ... помешать кому бы то ни было утверждать, что пространство имеет собственно четыре или

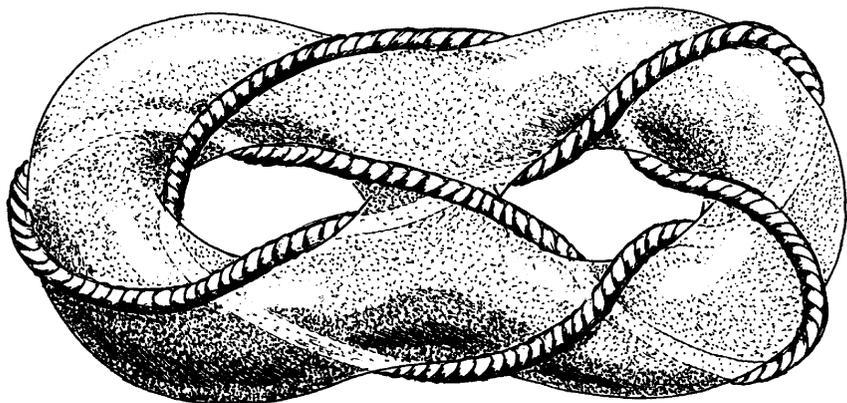
произвольно много измерений, но ... мы в состоянии воспринимать только три.»

23 октября 1897 года российское периодическое издание «Новости и биржевая газета» познакомила своих читателей со следующим сообщением из университетской жизни:

«КАЗАНЬ. В годовщину рождения Лобачевского в актовом зале университета состоялось торжественное заседание физико-математического общества для первого присуждения премии имени Лобачевского профессору Лейпцигского университета Садису, золотой медали — профессору Геттингенского университета Феликсу Клейну. Член общества Рейнгардт предложил войти с ходатайством в Министерство народного просвещения о введении неевклидовой геометрии в курс общего образования.»

Сегодня, сто лет спустя, это сообщение вызывает особый интерес как живое историческое свидетельство примечательной встречи имен Н. Лобачевского и Ф. Клейна, причудливого переплетения судьбы их фундаментальных работ и того специального внимания, которое российская научно-педагогическая общественность всегда специально уделяла геометрическому школьному образованию.

Тема номера: теория узлов в конце XX века



Математическая теория узлов вот уже несколько лет находится в центре внимания не только математиков, но и физиков, биологов, биохимиков. Модное увлечение? Возможно — не без этого, но это увлечение, как в свое время увлечение теорией относительности или квантовой теорией поля, основано на серьезных научных достижениях.

Математиков к теории узлов привлекает скорей всего сам предмет изучения — *узел*, сочетание его наглядной простоты и таинственности. Говоря о таинственности, мы не имеем в виду до сих пор не расшифрованное религиозно-символическое значение изображений узлов и зацеплений узлов, появляющихся в различных древних культурах, ни тем более «магические» превращения узлов в руках эстрадных фокусников. Речь здесь идет о том, что наиболее просто формулируемые задачи, относящиеся к этому простому объекту, до сих пор не имеют удовлетворительного решения.

Узел — действительно очень простой и наглядный объект: его можно себе представить зрительно (или сделать модель) в виде запутанного шнурка, концы которого склеены между собой. Шнурком разрешается манипулировать в пространстве (не разрывая), при этом различные его положения, полученные таким образом, относят к одному и тому же типу узла, и именно *тип узла* нас и интересует. Беда в том, что два положения шнурка, относящиеся к одному и тому же типу узла, могут оказаться совершенно не похожими. И представьте себе —

такая простенькая игрушка вот уже почти полтора века не желает раскрыть свои тайны, несмотря на объединенные усилия многих выдающихся ученых¹⁾.

Так, самая первая и естественная задача теории узлов — найти

▷ *общий способ распутывания узлов*

до сих пор не имеет удовлетворительного решения²⁾. Тем более не имеет удовлетворительного решения

▷ *проблема классификации узлов*

(т. е. составления списка k_1, k_2, k_3, \dots всех узлов без повторений вкупе с алгоритмом, относящим каждому узлу его номер в списке). Не найдена и

▷ *полная система инвариантов узлов*

(т. е. способ сопоставления узлу набора алгебраических характеристик, совпадение которых для двух конкретных узлов k', k'' происходило бы, если и только если k' и k'' относятся к одному и тому же типу узла).

Тем не менее на пути к решению этих задач математики достигли значительных успехов и, что особенно приятно, в большинстве случаев их удается описать совершенно элементарно. Таким образом, здесь математикам есть чем похвастаться, но есть и чем заниматься дальше.

А что же физики? Они первыми начали систематическое изучение узлов около 140 лет тому назад, но больших успехов не достигли, после чего лет на сто потеряли всякий интерес к этой теме. (Возможно, эта потеря интереса связана с тем, что уж очень сильно они здесь оконфузились: исходная идея Томсона (лорда Кельвина) состояла в том, что атомы представляют собой «маленькие завихрения» в «мировом эфире», и эти завихрения (*vortex atoms*) следует моделировать узлами. Таким образом, классификация узлов должна была доставить и классификацию химических элементов. Однако (сама по себе интересная) модель Томсона никакой пользы ни физике, ни химии не принесла, а задача классификации химических элементов была блистательно решена на другом — более арифметическом, чем геометрическом — пути нашим великим соотечественником Менделеевым.)

Но после того, как в середине восьмидесятых годов математики (Воан Джонз и его последователи) изобрели новое семейство инвариантов узлов, физики просто набросились на эту область, и в этот раз достигли серьезных результатов. Парадоксально, но здесь наблюдалось не приложение теории узлов к физике, а скорей наоборот — приложение физических идей к построению этой математической теории. Впрочем, инварианты узлов все же послужили материалом для построения некоторых весьма интересных физических моделей (в работах Э. Виттена, М. Атья, П. Вожеля и др.), но пока вопрос о связи этих моделей с реальной действительностью даже и не ставится. Не успев исчерпать возможности инвариантов типа Джонса, физики не менее рьяно взялись за еще более

¹⁾ На начальных порах среди них были такие авторитеты, как К. Ф. Гаусс, лорд Келвин и А. Пуанкаре, а в самые последние годы — не менее четырех (!) лауреатов премии Филдса (самой престижной премии для математиков): Эдвард Виттен, Воан Джонз, Владимир Дринфельд и Максим Концевич.

²⁾ Формально говоря, алгоритм распутывания узлов существует, но он работает слишком медленно и поэтому его бессмысленно реализовывать на компьютере.

сильные *инварианты конечного порядка*, придуманные московским математиком Виктором Васильевым, но и здесь пока рано судить о естественно-научном значении этих работ.

Узлы стали также обсуждаться в других естественных науках:

▷ в *генетике, молекулярной биологии, и генной инженерии*

в связи с тем, что ДНК и другие длинные молекулы зацепляются и заузливаются, что очень существенно влияет на их поведение;

▷ в *гидродинамике*,

где изучение устойчивых вихрей, образующих узлы, привело А. Моффата к построению замечательного инварианта, называемого *энергией узла*;

▷ в *ферромагнетизме*,

где тоже возникают заузленные потоки, но не жидкости, а магнитных линий.

Однако не нужно наивно думать, что теория узлов напрямую прилагается к практическим вопросам этих наук, доставляет коммерческую прибыль. Нет, пока здесь речь идет (как и в физике) скорей о связях, чем о непосредственных приложениях. Однако эти связи достаточно глубоки, и вызывают интерес не только из-за «моды» на теорию узлов.

Все же мы не сочли возможным более подробно останавливаться на этих вопросах в настоящем выпуске «Математического Просвещения» — все три публикуемые ниже статьи на тему выпуска относятся непосредственно к математической теории узлов.

Основной является статья С. В. Дужина и С. В. Чмутова, в которой дается обзор состояния этой науки на конец XX века, притом обзор, для понимания которого не только не требуется предварительных специальных знаний, но и вообще нужна лишь минимальная общематематическая подготовка. В статье В. Прасолова обсуждается ряд вопросов, связанных с одной из самых красивых конструкций теории узлов — *поверхности Зейферта*, которая позволяет определить классический инвариант узла — его *род* и строить удивительные *разветвленные накрытия* над узлами. Статья же Ю. Бурмана посвящена найденной в самое последнее время конструкции — инвариантам конечного порядка так называемых *длинных кривых* (включая *странность*, придуманную В. И. Арнольдом); отсюда, в частности, получается простой способ для вычисления некоторых инвариантов Васильева.

А. Б. Сосинский

Узлы и их инварианты

С. В. Дужин

С. В. Чмутов

СОДЕРЖАНИЕ

1. Введение	59
2. Понятие узла	60
3. Диаграммы	64
4. Инварианты узлов	70
5. Инварианты Васильева	77
6. Интеграл Концевича	85
7. Знаменитые проблемы теории узлов	89
Приложение. Таблица узлов	91

1. ВВЕДЕНИЕ

Прародителем теории узлов является К. Ф. Гаусс, который вывел замечательную интегральную формулу для числа оборотов одной замкнутой кривой в пространстве вокруг другой, а также оставил в своих записных книжках многочисленные рисунки узлов без комментариев. Систематическое развитие теории узлов началось с 1880-х годов, когда несколько математиков и физиков с большим энтузиазмом занялись составлением таблиц узлов под влиянием идей знаменитого физика У. Томпсона (впоследствии известного как лорд Кельвин) о том, что узлы должны взаимно однозначно соответствовать химическим элементам. Вскоре выяснилось, что эта идея была ошибочной, в химии восторжествовала теория Д. И. Менделеева, и лишь много позже были открыты реальные естественно-научные приложения теории узлов (строение молекул ДНК, гидродинамика, ферромагнетики и т. д.). Впрочем, в нашей статье мы не собираемся затрагивать прикладных тем и будем рассматривать узлы как чисто математический объект.

Теория узлов — одна из наиболее привлекательных областей математики. Для этого есть несколько причин:

- ▷ Теории узлов немногим более 100 лет, что, по-видимому, является возрастом расцвета научной теории.
- ▷ Предметом теории является узел — простой и наглядный объект.
- ▷ Теория узлов тесно связана почти со всеми областями математики, и в ней успешно работают геометрические, алгебраические и аналитические понятия и методы.

▷ В теории узлов доказано много содержательных и трудных теорем, но до сих пор существует много открытых проблем с простой формулировкой, иные из которых стоят уже десятки лет.

Цель этой статьи — рассказать о том, что такое узел и какие есть способы распознать, одинаковы или различны два данных узла. Наряду с понятиями и теоремами классической теории узлов, в статье упоминаются некоторые достижения последних лет (инварианты Васильева, лежандровы узлы). Не претендуя никоим образом на полноту изложения, мы отсылаем заинтересованного читателя к книгам Адамса [1], Прасолова и Сосинского [22], Кроуэлла и Фокса [7].

2. ПОНЯТИЕ УЗЛА

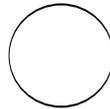
Узел — это гладкая замкнутая несамопересекающаяся кривая в пространстве. Мы будем всегда предполагать эту кривую параметризованной и более точно определение узла формулируем так.

ОПРЕДЕЛЕНИЕ. *Узел* — это гладкое вложение окружности S^1 в трехмерное пространство \mathbb{R}^3 , т. е. бесконечно дифференцируемое отображение $f : S^1 \rightarrow \mathbb{R}^3$, которое переводит разные точки окружности в разные точки пространства, причем вектор скорости $f'(u)$ не обращается в 0 ни при каком $u \in S^1$.

Если выбрать систему координат в пространстве, то узел можно задать набором из трех гладких периодических функций $x(u)$, $y(u)$, $z(u)$.

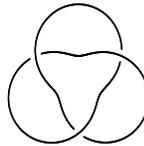
ПРИМЕР 1. *Тривиальный узел*

$$\begin{aligned}x &= \cos u, \\y &= \sin u, \\z &= 0.\end{aligned}$$



ПРИМЕР 2. *Трилистник*

$$\begin{aligned}x &= (2 + \cos 3u) \cos 2u, \\y &= (2 + \cos 3u) \sin 2u, \\z &= \sin 3u.\end{aligned}$$



Мы будем всегда считать узлы *ориентированными*, считая положительным направлением обхода узла то, которое соответствует вращению против часовой стрелки на параметризующей окружности $S^1 = \{(\cos u, \sin u) | u \in \mathbb{R}\} \subset \mathbb{R}^2$.

2.1. Изотопия

Гладкая замкнутая кривая, рассматриваемая сама по себе, может служить предметом дифференциальной геометрии, но не топологии. *Топологическое* понятие узла подразумевает, что узлы рассматриваются с точностью до естественного отношения эквивалентности, моделирующего те преобразования, которые можно физически проделать с куском веревки, концы которой связаны.

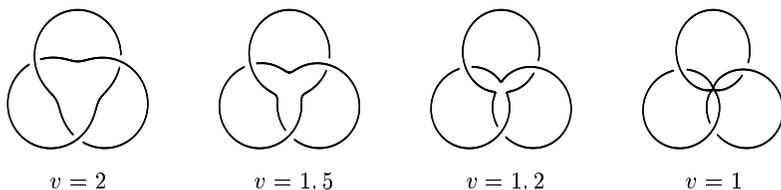
ОПРЕДЕЛЕНИЕ. *Гладкая изотопия узла K , заданного как отображение $f : S^1 \rightarrow \mathbb{R}^3$, — это гладкое семейство узлов f_v такое, что $f_0 = f$.*

Гладкое семейство отображений, или отображение, гладко зависящее от параметра, — это, говоря формально, гладкое отображение $F : S^1 \times I \rightarrow \mathbb{R}^3$, где I — отрезок числовой оси. Если зафиксировать второй аргумент отображения F , придав ему значение v , то мы получим отображение $f_v : S^1 \rightarrow \mathbb{R}^3$.

Например, формулы

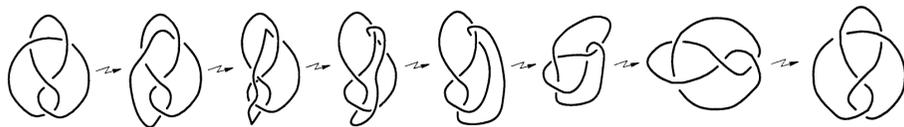
$$\begin{aligned}x &= (v + \cos(3u)) \cos(2u), \\y &= (v + \cos(3u)) \sin(2u), \\z &= \sin(3u)\end{aligned}$$

задают при $v \in (1, +\infty)$ гладкую изотопию рассмотренного выше трилистника, отвечающего значению $v = 2$:



При всех значениях $v > 1$ это будет несамопересекающаяся вложенная кривая, а при $v = 1$ на ней появится тройная точка, а именно, совпадут значения (x, y, z) для $u = \pi/3$, $u = \pi$ и $u = 5\pi/3$.

Два узла называются *эквивалентными*, если от одного к другому можно перейти гладкой изотопией. На рисунке изображен пример гладкой изотопии узла, называемого *восьмеркой*:



Деформация восьмерки в ее зеркальный образ

Тривиальный узел (или *неузел*) — это узел, эквивалентный плоской окружности из примера 1.

Имея перед глазами рисунок тривиального узла, иной раз бывает довольно трудно понять, что он тривиален.



Тривиальные узлы

Алгоритм распознавания тривиального узла впервые разработал В. Хакен [14] в 1961 году. Алгоритм Хакена работает слишком медленно и поэтому никогда не реализовывался на компьютере. Совсем недавно (препринт 1998 года [5])

Дж. Бирман и М. Хирш придумали другой алгоритм, который, судя по всему, уже имеет смысл запрограммировать.

2.2. ОРИЕНТАЦИЯ

Узел — это ориентированная кривая в ориентированном пространстве. Отражение относительно плоскости меняет ориентацию пространства. Изменение направления обхода кривой на противоположное меняет ориентацию кривой.

Естественно выделить классы узлов, которые остаются эквивалентными себе при таких преобразованиях.

ОПРЕДЕЛЕНИЕ. Узел называется *зеркальным*, если он эквивалентен своему зеркальному отражению.

ПРИМЕР 1. Узел «восьмерка» зеркален. Это доказывает изотопия, приведенная на рисунке выше.

ПРИМЕР 2. Трилистник не является зеркальным узлом:



Левый трилистник



Правый трилистник

Этот факт будет доказан в разделе 4.2.

ОПРЕДЕЛЕНИЕ. Узел называется *обратимым*, если он эквивалентен своему обратному, т. е. тому же узлу, проходящему в обратном направлении.

ПРИМЕР. Трилистник обратим, так как направление обхода можно заменить на обратное плавным поворотом на 180° вокруг оси Ox (в приведенном выше тригонометрическом представлении).

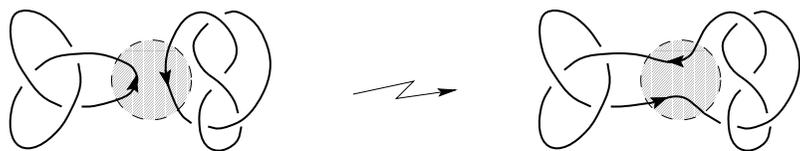
УПРАЖНЕНИЕ. Проверьте, что узел «восьмерка» также обратим.

Среди узлов, имеющих не более 8 пересечений, есть только один необратимый: это узел 8_{17} (см. таблицу узлов в приложении).

Распознать, обратим ли данный узел, непросто. Первое строгое доказательство необратимости было проведено только в 1962 году (пример Троттера [26]). Общего алгоритма для решения этой проблемы не найдено до сих пор.

2.3. КОМПОЗИЦИЯ УЗЛОВ

Существует естественный способ строить более сложные узлы из двух более простых. Для этого каждый из двух узлов нужно разрезать в каком-то месте, а потом приклеить два конца одного узла к двум концам другого. Все это надо делать достаточно осторожно: сначала нужно путем гладкой изотопии продеформировать оба данных узла так, чтобы для какой-то регулярной проекции на плоскость нашелся диск, внутри которого наши узлы имели бы вид, указанный на следующем рисунке слева, а затем изменить их внутри этого диска, как показано справа. Поскольку мы рассматриваем ориентированные узлы, склейку можно произвести только одним способом.



Связная сумма узлов

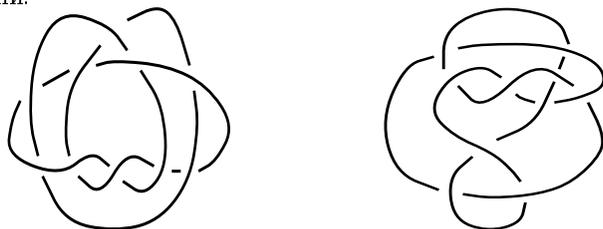
Легко понять, что сумма $K_1 \# K_2$ двух узлов K_1 и K_2 не зависит от точек разреза (вообще, все точки узла равноправны между собой относительно гладких изотопий).

Если представлять себе связную сумму узлов как умножение, то тривиальный узел играет, очевидно, роль единицы: $E \# K = K$ для любого K .

Узел называется *простым*, если его нельзя представить в виде суммы двух нетривиальных узлов. Простые узлы выполняют роль простых чисел в арифметике: оказывается, что любой узел можно представить в виде связной суммы нескольких простых, и притом однозначно.

По этой причине в таблицах узлов, составлением которых математики занимаются с 1880-х годов, приводятся только простые узлы. К настоящему времени составлены полные таблицы всех простых узлов, которые допускают регулярную проекцию с не более чем 16 точками пересечения (Дж. Хост, М. Тислуэйт, Дж. Уикс, 1997). В таблицах узлов приняты обозначения вида n_m , где n — число пересечений, m — номер по порядку среди всех узлов с n пересечениями. Узлы, получающиеся друг от друга заменой ориентации или зеркальным отражением, в таблицах узлов не различаются.

При составлении таблиц узлов люди часто допускали ошибки (и немудрено). Самый знаменитый случай такого рода — это *пара Перко*, два эквивалентных друг другу узла, которые в течение 75 лет, начиная с 1899 года, считались различными. Вот они:



Пара Перко

Обнаружил их эквивалентность юрист из Нью-Йорка К. А. Перко в 1974 году. Сочетание «юрист-профессионал и одновременно математик-любитель» не первый раз встречается в истории математики (вспомните Пьера Ферма!) и всегда приводит к новым открытиям.

В Приложении приведена таблица простых узлов до 8 самопересечений включительно. Для работы с узлами удобно пользоваться компьютерными программами и таблицами, которые имеются на нескольких [www-серверах](#) (см., например, [24]).

3. ДИАГРАММЫ

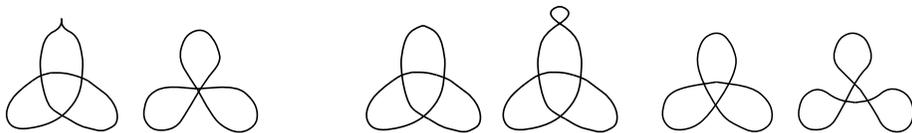
В принципе, любой узел можно задать явными формулами $x(t)$, $y(t)$, $z(t)$ в виде тригонометрических многочленов. Пример представления в таком виде одного узла (трилистника) был приведен выше. Тот факт, что для любого узла найдется эквивалентный ему тригонометрический, вытекает из того, что любые периодические функции $x(t)$, $y(t)$, $z(t)$ можно с любой степенью точности приблизить тригонометрическими многочленами, т. е. комбинациями $\sin kt$ и $\cos kt$.

Надо, однако, заметить, что не существует никакого разумного способа распознать эквивалентность или доказать неэквивалентность узлов, заданных явными формулами. Все способы задания узлов, с которыми можно реально работать, имеют комбинаторный характер. В этом разделе мы рассмотрим два таких способа: (классические) диаграммы узлов и лежандрово представление.

3.1. ДИАГРАММА УЗЛА

Диаграмма узла — это тот естественный способ изображения узла на плоскости, которым мы уже пользовались на предыдущих рисунках. Говоря более формально, диаграмма узла представляет собой гладкую кривую на плоскости, не имеющую других особых точек, кроме конечного числа простых двойных точек (там, где происходит трансверсальное самопересечение), и снабженную в каждой двойной точке одним битом дополнительной информации, указывающим, какая из двух ветвей в этой точке является *проходом* (т. е. лежит снизу), а какая *переходом* (т. е. лежит сверху).

Чтобы нарисовать диаграмму данного узла $f : S^1 \rightarrow \mathbb{R}^3$, нужно его немного пошевелить, т. е. подвергнуть изотопии, близкой к тождественной, с тем, чтобы его проекция $\pi \circ f : S^1 \rightarrow \mathbb{R}^3 \rightarrow \mathbb{R}^2$ на выбранную плоскость стала *регулярной*, т. е. (1) вектор производной $(\pi \circ f)'$ нигде не обращался бы в 0 и (2) никакая точка не имела бы более двух прообразов. Очевидно, что этого всегда можно добиться малой изотопией, не выводящей за пределы класса эквивалентности узла.



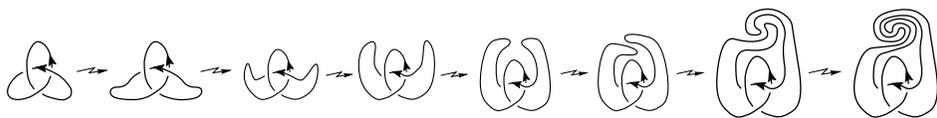
Нерегулярные проекции

Регулярные проекции

При изотопии узла в пространстве его диаграмма может меняться. Эти изменения могут быть двух принципиально различных видов. Во-первых, диаграмма узла, как гладкая замкнутая плоская кривая, может претерпевать *плоскую изотопию*, т. е. подвергаться воздействию семейства взаимно однозначных преобразований плоскости, гладко зависящего от параметра.

Во-вторых, возможны *перестройки диаграммы*, когда она меняет свой тип, например, показанные на следующей странице *перестройки Рейдемейстера*.

К. Рейдемейстер — автор первой книги, посвященной целиком теории узлов (1932). В ней доказана следующая теорема: *две диаграммы задают один и тот*

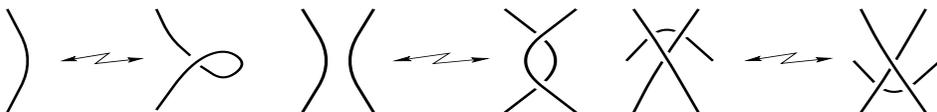


Плоская изотопия диаграммы

I)

II)

III)



Перестройки Рейдемейстера

же узел тогда и только тогда, когда от одной из них можно перейти к другой посредством плоских изотопий диаграммы и перестроек Рейдемейстера.

С точки зрения современной теории особенностей теорема Рейдемейстера довольно очевидна: ясно, что любую гладкую изотопию узла в пространстве можно немного пошевелить так, что она сведется к последовательности элементарных перестроек указанных трех типов. Подробного доказательства этого утверждения мы приводить не будем, отсылая заинтересованного читателя, например, к книге [22].

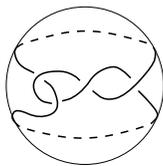
ПРИМЕР. Превращение восьмерки в ее зеркальный образ, показанное на рисунке выше (с. 61), помимо плоских изотопий, содержит следующую последовательность перестроек Рейдемейстера: II, III, III, I, I, III, II.

Объясним, в каком смысле диаграмма может рассматриваться как чисто комбинаторный способ задания класса эквивалентности узла.

Диаграмма узла, рассматриваемая с точностью до плоских изотопий (гладких деформаций в плоскости), содержит следующую информацию: (1) в каком порядке вдоль кривой следуют точки самопересечения, (2) какая ветвь кривой в точке самопересечения является проходом, а какая переходом.

Для кодирования плоских замкнутых кривых удобным инструментом являются *гауссовы диаграммы*. На самом деле, гауссовы диаграммы больше приспособлены для описания замкнутых кривых на сфере, нежели замкнутых кривых на плоскости. Чтобы к ним перейти, мы сделаем следующее простое наблюдение.

Вместо плоских диаграмм узла можно рассматривать сферические, и эти два подхода абсолютно эквивалентны. Сферическая диаграмма получается как проекция узла на сферу $S^2 \subset \mathbb{R}^3$, содержащую наш узел внутри, из центра сферы, если эта проекция регулярна. Прокалывая сферу в какой-то точке, мы получаем из сферической диаграммы плоскую.

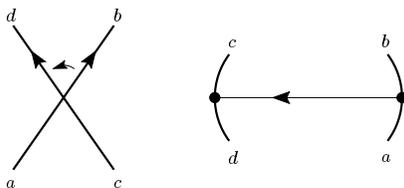


Сферическая диаграмма восьмерки

Как впервые заметил еще К. Ф. Гаусс, а впоследствии (вплоть до конца XX века) заново открывали многие другие математики, замкнутые самопересекающиеся ориентированные сферические кривые с точностью до гладкой изотопии на поверхности сферы однозначно кодируются посредством диаграмм следующего вида.

ОПРЕДЕЛЕНИЕ. Гауссова диаграмма порядка n — это ориентированная окружность, на которой отмечены $2n$ различных точек, разбитых на пары, и на каждой из n хорд, соединяющих парные точки, указано направление.

Опишем построение гауссовой диаграммы кривой. На окружности, которая параметризует данную кривую, нужно отметить прообразы всех двойных точек и соединить каждую пару прообразов одной точки хордой. Направление на хордах выбирается так. На поверхности сферы есть вполне определенное положительное направление вращения (против часовой стрелки, если на сферу смотреть снаружи). Два конца хорды соответствуют двум ветвям кривой в окрестности двойной точки. Положительные направления этих ветвей образуют угол. Так вот, направление на хорде соответствует переходу от одной стороны угла к другой в положительном направлении.



Выбор направления хорды

Информация о проходах и переходах содержится в *оснащенной гауссовой диаграмме*: это значит, что дополнительно на каждой хорде ставится знак «+», если она направлена от перехода к проходу, и знак «-» в обратном случае. Легко видеть, что этот знак совпадает со знаком *локальной закрученности* двойной точки, которая определяется следующим рисунком.



Локальная закрученность

Вот пример диаграммы узла и соответствующей оснащенной гауссовой диаграммы:

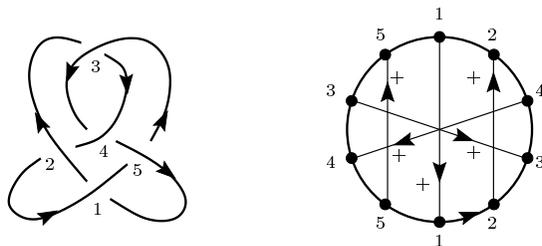


Диаграмма узла и гауссова диаграмма

Некоторые математики, например, М. Поляк и О. Виро, применяют другое правило расстановки стрелок на хордах (см. статью Ю. Бурмана в данном выпуске, с. 94–115). Оба правила равносильны, но наше имеет то преимущество, что, забывая о знаках на хордах, но сохраняя нарисованные на них стрелки, мы получаем классическую гауссову диаграмму, являющуюся полным инвариантом ориентированной сферической кривой.

3.2. ЛЕЖАНДРОВЫ УЗЛЫ

Представление узлов плоскими диаграммами — самый естественный и исторически первый способ наглядного задания узлов. В этом разделе мы приведем другое представление, которое в последние годы все больше привлекает внимание специалистов.

Предположим, что на плоскости выделено некоторое направление, которое мы будем называть вертикальным (например, направление оси ординат в декартовой системе координат).

ОПРЕДЕЛЕНИЕ. *Фронт* называется замкнутая кривая на плоскости без вертикальных касательных, все особенности которой исчерпываются полукубическими точками возврата (*клювами*) и простыми двойными точками.

Полукубическая точка возврата — это точка, вблизи которой наша кривая локально выглядит как полукубическая парабола $y^2 = x^3$ в окрестности начала координат. Это самый общий вид точки на параметризованной кривой $r(t)$, в которой вектор r' обращается в нуль, а векторы r'' и r''' неколлинеарны.

ЗАМЕЧАНИЕ. Название «фронт» используется по той причине, что эквидистанты для кривой общего положения на плоскости (например, эллипса) имеют только особенности указанных типов. А эквидистанта — это как бы фронт световой волны, исходящей из данной кривой.

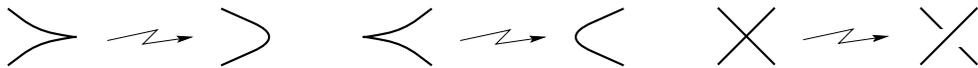
Объясним, в чем состоит связь между фронтами и узлами. Заодно мы расскажем, откуда берется требование отсутствия вертикальных касательных.

Фронт — это кривая на плоскости, скажем, с координатами (x, y) . Добавляя к ним еще третью координату z , равную тангенсу угла наклона кривой в данной точке, мы получим кривую в трехмерном пространстве \mathbb{R}^3 . Если на исходной кривой есть только простые двойные точки и клювы, то полученная кривая в пространстве будет гладким узлом. Тангенс угла наклона определен всюду, поскольку нет вертикальных касательных.

Если отказаться от требования отсутствия вертикальных касательных, то угол наклона кривой будет принимать значения не в точках прямой, а в точках окружности. Таким образом, в этом случае мы получим узел не в \mathbb{R}^3 , а в трехмерном многообразии $\mathbb{R}^2 \times S^1$, представляющем собой полноторие.

Опишем теперь соответствие между фронтами и узлами в терминах диаграмм.

По фронту диаграмма представляемого им узла строится так. Все точки возврата нужно просто сгладить, а проходы и переходы в двойных точках определить так: *ветвь кривой, идущая из левого нижнего угла в правый верхний, всегда является переходом, а другая ветвь — проходом*. Так как фронт не имеет вертикальных касательных, такое разделение ветвей корректно определено.



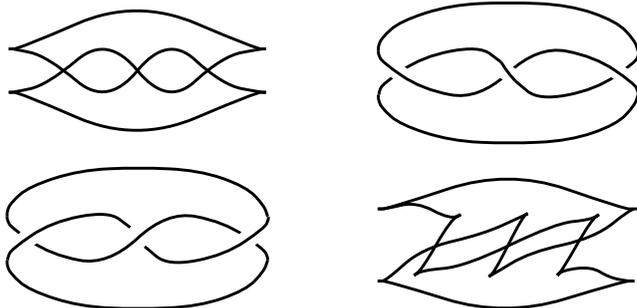
Построение диаграммы по фронту

Обратно, по диаграмме можно построить фронт представляемого ею узла. Приведем диаграмму в общее положение так, что в двойных точках ни одна из касательных не будет вертикальной. После этого заменим все точки с вертикальными касательными на точки возврата, а двойные точки, в которых переход имеет направление «северо-запад–юго-восток», преобразуем так, как показано на рисунке.



Построение фронта по диаграмме

Тем самым мы показали, что всякий узел можно представить фронтом. Например, на следующем рисунке изображены фронты и диаграммы левого и правого трилистника (слева по фронту построена диаграмма, а справа, наоборот, по диаграмме — фронт):

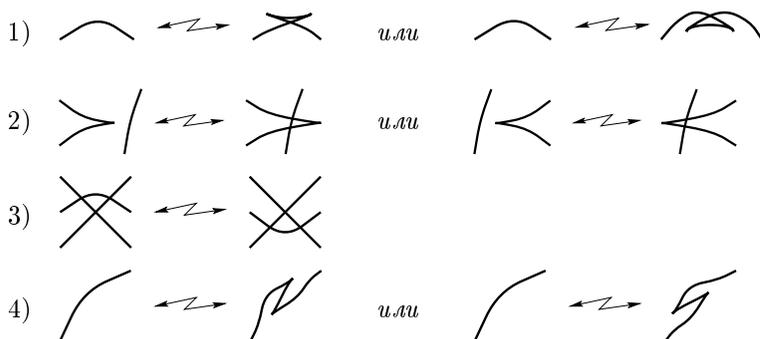


Левый трилистник

Правый трилистник

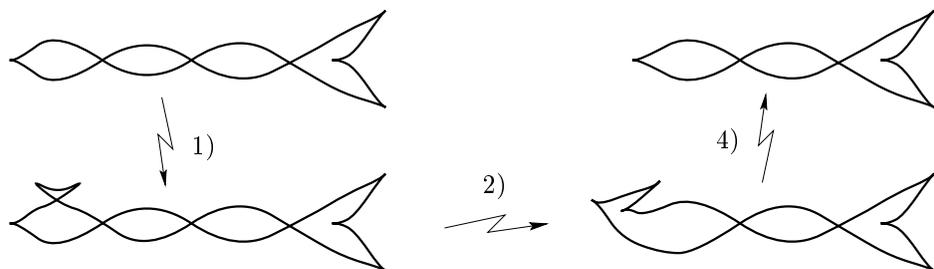
Узел, представленный фронтом, называется *лежандровым* узлом. Аналог перестроек Рейдемейстера для лежандровых узлов, описывающих топологическую эквивалентность двух узлов, в явном виде был выписан совсем недавно Д. Б. Фуком и С. Л. Табачниковым. Поэтому перечисленные в следующей теореме перестройки мы будем называть *перестройками Фукса–Табачникова*.

ТЕОРЕМА (Д. Б. ФУКС, С. Л. ТАБАЧНИКОВ [13]). *Два фронта представляют эквивалентные узлы тогда и только тогда, когда от одного фронта к другому можно перейти конечной последовательностью следующих перестроек:*



ОПРЕДЕЛЕНИЕ. В отличие от приведенной выше эквивалентности (топологической) скажем, что два лежандровых узла *лежандрово эквивалентны*, если от фронта одного из них можно перейти к фронту другого при помощи только первых трех перестроек Фукса–Табачникова.

В каждом классе топологически эквивалентных узлов существует бесконечно много лежандрово не эквивалентных узлов. Например, следующие два фронта представляют два неузла (тривиальных узла) и потому топологически они эквивалентны. На рисунке показана также последовательность перестроек Фукса–Табачникова, переводящая один фронт в другой.



Топологическая эквивалентность двух неузлов

В разделе 4.4 мы докажем, что эти два неузла лежандрово не эквивалентны. Это означает, что любая последовательность перестроек Фукса – Табачникова, переводящая один фронт в другой, обязательно содержит хотя бы одну

перестройку типа 4). Мы приведем также теорему Я. Элиашберга, описывающую лежандрову эквивалентность среди топологически тривиальных узлов.

4. ИНВАРИАНТЫ

Для того, чтобы доказать, что два узла эквивалентны, достаточно предъявить изотопию между ними, заданную, например, последовательностью перестроек Рейдемейстера для соответствующих диаграмм.

Стандартным приемом доказательства неэквивалентности, как в теории узлов, так и вообще в математике, является использование инвариантов.

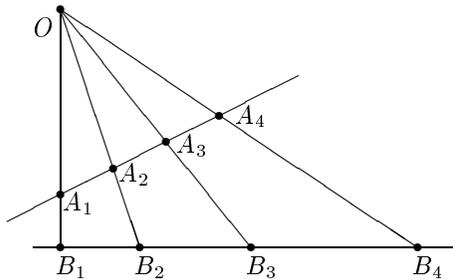
4.1. ОБЩЕЕ ПОНЯТИЕ ИНВАРИАНТА

Допустим, в каком-то множестве X задано отношение эквивалентности. Отображение $X \rightarrow A$ называется *инвариантом*, если оно принимает равные значения на элементах X , эквивалентных между собой. В качестве A можно использовать множество произвольной природы. Если A состоит из чисел, говорят о числовых инвариантах; если из многочленов — о полиномиальных инвариантах, и т. д.

Понятие инварианта реально работает в тех случаях, когда в множестве X отношение эквивалентности устроено достаточно сложно, а в множестве A проблема равенства решается достаточно просто.

ПРИМЕР 1. Двойное отношение как инвариант четверки точек на прямой относительно отображений центрального проектирования.

Рассмотрим множество X , элементами которого являются упорядоченные четверки точек на прямой. Две четверки, A_1, A_2, A_3, A_4 и B_1, B_2, B_3, B_4 , будем считать эквивалентными, если прямые, их содержащие, можно расположить на плоскости таким образом, чтобы центральная проекция из некоторой точки плоскости переводила A_i в B_i для всех $i = 1, \dots, 4$.



Эквивалентность двух четверок относительно центральной проекции

ЗАДАЧА. Докажите, что одну четверку можно перевести в другую тогда и только тогда, когда выполняется пропорция

$$\frac{A_1 - A_3}{A_2 - A_3} : \frac{A_1 - A_4}{A_2 - A_4} = \frac{B_1 - B_3}{B_2 - B_3} : \frac{B_1 - B_4}{B_2 - B_4}.$$

Мы видим, что величина $\frac{A_1 - A_3}{A_2 - A_3} : \frac{A_1 - A_4}{A_2 - A_4}$ (*двойное отношение четырех точек*) является инвариантом упорядоченной четверки точек на прямой. Результат

предыдущей задачи показывает, что этот инвариант очень сильный: если его значения на двух элементах совпадают, то сами элементы эквивалентны между собой. Инвариант, обладающий таким свойством, называется *полным*.

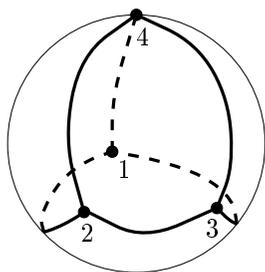
Таким образом, двойное отношение четырех точек является полным инвариантом отношения проективной эквивалентности.

ПРИМЕР 2. Эйлерова характеристика двумерной поверхности, представленной в виде триангуляции.

Всякую компактную двумерную поверхность (сферу, тор, диск, ленту Мёбиуса и т. п.) можно разрезать на треугольники так, что любые два треугольника либо вообще не пересекаются, либо имеют одну общую вершину, либо одну общую сторону. Такое разрезание называется *триангуляцией* поверхности.

Комбинаторное задание триангуляции состоит в том, что все вершины нумеруют, а затем указывают, какие тройки номеров отвечают граням триангуляции.

Пример. Триангуляция $(1, 2, 3)$, $(1, 3, 4)$, $(1, 4, 2)$, $(2, 4, 3)$ задает сферу:



Триангуляция сферы

Эйлеровой характеристикой триангуляции называется число $f - e + v$, где f — число граней (треугольников), e — число ребер (сторон треугольников), v — число вершин.

Фундаментальная теорема комбинаторной топологии гласит, что эйлерова характеристика не зависит от выбора конкретной триангуляции, а определяется топологическим типом самой поверхности. Следовательно, эйлерова характеристика является инвариантом поверхности, рассматриваемой как топологическое пространство.

Известно, что любая компактная ориентируемая поверхность без края представляет собой сферу с ручками, причем если g — это число ручек, то эйлерова характеристика равна $2 - 2g$. Таким образом, эйлерова характеристика является полным топологическим инвариантом компактной ориентируемой поверхности без края. Для произвольных компактных двумерных поверхностей мы получим полный инвариант, если к эйлеровой характеристике добавим еще две величины: *число компонент края* (принимает натуральные значения) и *ориентируемость* (принимает значения в множестве {да, нет}).

ЗАДАЧА. Проверьте, что следующие две триангуляции задают поверхности с одинаковой эйлеровой характеристикой, но топологически неэквивалентные:

- 1) (1, 2, 6), (2, 3, 5), (2, 5, 6), (1, 4, 6), (3, 5, 7), (4, 6, 7), (4, 6, 7), (1, 3, 7), (2, 4, 7), (1, 2, 7);
 2) (1, 2, 3), (1, 3, 4), (1, 4, 5), (1, 5, 6), (1, 6, 7), (1, 7, 8), (1, 8, 9), (2, 3, 6), (2, 6, 7), (4, 5, 8), (4, 8, 9).

4.2. ИНВАРИАНТЫ УЗЛОВ

Инварианты, придуманные людьми для различения узлов, поистине неисчислимы.

Начнем с краткого описания нескольких замечательных инвариантов, играющих по разным причинам важную роль в теории узлов.

- ▷ *Число пересечений*, т. е. минимальное число двойных точек на диаграмме узла.

В каком-то смысле число пересечений можно считать мерой сложности узла. Узлы с числом пересечений 1 или 2 все тривиальны. Простейший нетривиальный узел — трилистник — имеет число пересечений, равное 3, и с точностью до зеркального отражения это единственный узел с тремя пересечениями.

В таблицах узлы традиционно упорядочиваются по числу пересечений.

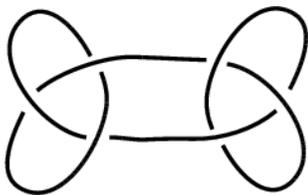
- ▷ *Топологический тип дополнения* — полный инвариант неориентированного узла.

Это теорема Гордана-Люке, доказанная только в 1989 году, хотя проблема была поставлена еще в начале XX века. Теорема звучит так: *если дополнения $\mathbb{R}^3 \setminus K_1$ и $\mathbb{R}^3 \setminus K_2$ двух узлов гомеоморфны, то сами узлы эквивалентны с точностью до ориентации узла и ориентации пространства.*

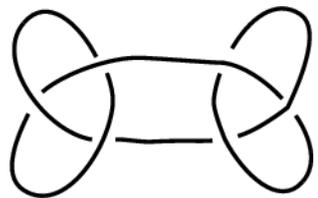
- ▷ *Род узла* — наименьший род его поверхности Зейферта (см. статью В. Прасолова в настоящем сборнике, с. 116–126).

- ▷ *Группа узла* — фундаментальная группа топологического пространства $\mathbb{R}^3 \setminus K$ (дополнения к узлу в пространстве). Эта группа состоит из классов эквивалентности замкнутых путей, начинающихся и заканчивающихся в одной и той же точке (групповая операция — последовательное прохождение двух путей). Существует простой алгоритм задания группы узла при помощи образующих и определяющих соотношений (см., например, [7]).

Неэквивалентные узлы могут иметь изоморфные группы. Например, группы бабушкина узла и сквер-узла изоморфны:



Бабушкин узел



Сквер-узел

Оба эти узла составные: они получаются композицией двух трилистников (двух левых или левого и правого). Интересно отметить, что для простых узлов подобный пример невозможен. Это вытекает из теоремы Гордана–Люке, сформулированной выше, и следующей теоремы [29]: *Если группы двух простых узлов изоморфны, то эти узлы имеют гомеоморфные дополнения.*

4.3. Полином Джонса

Особое значение в теории узлов играют инвариантные полиномы, т. е. инварианты узлов, принимающие значения в кольце многочленов от одного или нескольких переменных. Первый из полиномиальных инвариантов был введен Дж. Александером в 1920-х годах. Конструкция Александера довольно сложна. В последние годы появилось множество различных инвариантных полиномов, для которых существуют простые алгоритмы вычисления их значений по диаграмме узла. Из них мы приведем определение только одного: полинома Джонса. (Определение и основные свойства других полиномов можно найти, например, в книге [22]).

Полином Джонса определяется по диаграмме узла посредством нескольких простых правил. Заметим, что полином Джонса необходимо определять сразу для узлов и *зацеплений*, поскольку при перестройках диаграмм, которые фигурируют в приведенном ниже рекуррентном соотношении (1), диаграмма узла легко может превратиться в диаграмму зацепления. Зацепление — это многокомпонентный узел, т. е. набор нескольких замкнутых гладких кривых в пространстве, рассматриваемых с точностью до гладкой изотопии. Каждая компонента зацепления может быть заузлена как сама с собой, так и с другими компонентами. Зацепления изображаются на плоскости при помощи таких же точно диаграмм, что и узлы, с единственной разницей, что плоская кривая, составляющая диаграмму, может состоять из нескольких компонент связности.

ОПРЕДЕЛЕНИЕ. *Полином Джонса* V — это лорановский полином от переменной \sqrt{t} , определяемый соотношениями:

$$(1) \quad t^{-1}V \left(\begin{array}{c} \nearrow \\ \searrow \end{array} \right) - tV \left(\begin{array}{c} \searrow \\ \nearrow \end{array} \right) = (t^{1/2} - t^{-1/2})V \left(\begin{array}{c} \curvearrowright \\ \curvearrowleft \end{array} \right)$$

$$(2) \quad V \left(\bigcirc \right) = 1.$$

Лорановским полиномом от переменной, скажем, x , называется полином от переменных x и x^{-1} . Таким образом, полином Джонса представляет собой некоторую линейную комбинацию положительных и отрицательных полуцелых степеней буквы t .

Соотношение (1) связывает между собой полиномы Джонса трех зацеплений, которые внутри небольшого шара выглядят, как нарисовано на картинках, и совершенно одинаковы вне этого шара. Соотношение (2) гласит, что полином Джонса тривиального узла равен 1. Используя соотношение (1) рекуррентно, можно вычислить полином Джонса любого узла или зацепления путем сведения его к начальному условию (2). Разумеется, надо еще доказать, что результат не зависит от последовательности действий, а также, что полином Джонса определяется изотопическим классом узла, т. е. не меняется при перестройках Рейдемейстера. Это действительно так, но доказательства мы приводить не будем.

Полином Джонса обладает многими замечательными свойствами. Упомянем следующие.

- ▷ Для зацеплений с нечетным числом компонент (в частности, для узлов) все степени переменной t в полиноме Джонса целые, а для зацеплений с четным числом компонент — полуцелые.
- ▷ Полином Джонса связной суммы узлов равен произведению полиномов Джонса слагаемых.
- ▷ Полином Джонса не меняется при обращении узла, т. е. при замене направления обхода на противоположное.

Несмотря на то, что полином Джонса «не чувствует» ориентации самого узла, он чувствует ориентацию пространства. Например, он позволяет отличить левый трилистник от правого, чего нельзя сделать при помощи полинома Александера. Сейчас мы проделаем необходимые для этого выкладки, начиная с вычисления полиномов Джонса простейших зацеплений.

$$(i) \ V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right) = -t^{1/2} - t^{-1/2}.$$

Действительно, согласно (2) $V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right) = V \left(\begin{array}{c} \circlearrowright \\ \circlearrowleft \end{array} \right) = 1$. Тогда (1) дает $t^{-1} - t = (t^{1/2} - t^{-1/2})V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right)$, откуда и следует ответ.

$$(ii) \ V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowleft \end{array} \right) = -t^{1/2} - t^{5/2}.$$

Применим (1) к исходному зацеплению:

$$t^{-1}V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowleft \end{array} \right) - tV \left(\begin{array}{c} \circlearrowright \\ \circlearrowright \end{array} \right) = (t^{1/2} - t^{-1/2})V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right).$$

Но $V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowleft \end{array} \right) = -(t^{1/2} + t^{-1/2})$ по пункту (i), а $V \left(\begin{array}{c} \circlearrowright \\ \circlearrowright \end{array} \right) = 1$ согласно (2). Следовательно,

$$V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowleft \end{array} \right) = t(t^{1/2} - t^{-1/2}) - t^2(t^{1/2} + t^{-1/2}) = -t^{1/2} - t^{5/2}.$$

$$(iii) \ V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right) = -t^{-5/2} - t^{-1/2}.$$

Аналогично предыдущему:

$$t^{-1}V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right) - tV \left(\begin{array}{c} \circlearrowright \\ \circlearrowleft \end{array} \right) = (t^{1/2} - t^{-1/2})V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right).$$

Следовательно,

$$V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right) = -t^{-1}(t^{1/2} - t^{-1/2}) - t^{-2}(t^{1/2} + t^{-1/2}) = -t^{-5/2} - t^{-1/2}.$$

$$(iv) \ V \left(\begin{array}{c} \circlearrowleft \\ \circlearrowleft \\ \circlearrowright \end{array} \right) = -t^{-4} + t^{-3} + t^{-1}.$$

Используя (1), получаем:

$$t^{-1}V \left(\text{left trefoil} \right) - tV \left(\text{right trefoil} \right) = (t^{1/2} - t^{-1/2})V \left(\text{left trefoil} \right).$$

Далее, $V \left(\text{left trefoil} \right) = 1$ по (2), как полином Джонса тривиального узла, а

$$V \left(\text{right trefoil} \right) = V \left(\text{link of two circles} \right) = -t^{-5/2} - t^{-1/2} \text{ по (iii)}. \text{ Следовательно,}$$

$$V \left(\text{left trefoil} \right) = t^{-2} - t^{-1}(t^{1/2} - t^{-1/2})(-t^{-5/2} - t^{-1/2}) = -t^{-4} + t^{-3} + t^{-1}.$$

$$(v) \quad V \left(\text{left trefoil} \right) = t + t^3 - t^4.$$

$$\text{Аналогично, } t^{-1}V \left(\text{left trefoil} \right) - tV \left(\text{right trefoil} \right) = (t^{1/2} - t^{-1/2})V \left(\text{left trefoil} \right).$$

$$V \left(\text{right trefoil} \right) = 1 \text{ по (2), } V \left(\text{link of two circles} \right) = -t^{1/2} - t^{5/2} \text{ по (ii)}. \text{ Поэтому}$$

$$V \left(\text{left trefoil} \right) = t^2 + t(t^{1/2} - t^{-1/2})(-t^{1/2} - t^{5/2}) = t + t^3 - t^4.$$

Из различия полиномов Джонса левого и правого трилистников вытекает их неэквивалентность.

4.4. ЛЕЖАНДРОВЫ ИНВАРИАНТЫ

Приведем определение двух простейших инвариантов лежандровой эквивалентности фронтов (т. е. лежандровых узлов, см. раздел 3.2): число Маслова μ и число Беннекена β , которые не являются топологическими инвариантами в том смысле, что их значения на топологически эквивалентных узлах могут различаться. При смене ориентации фронта число Маслова меняет знак, а число Беннекена не изменяется.

ОПРЕДЕЛЕНИЕ. Точку возврата назовем *положительной*, если в ее окрестности ориентация идет с нижней ветви на верхнюю, и *отрицательной* в противном случае.



Положительные точки возврата

Отрицательные точки возврата

Числом Маслова $\mu(L)$ лежандрова узла L называется половина разности между числом положительных и числом отрицательных точек возврата фронта L .

$$\mu(L) = \frac{1}{2} \left(\# \left(\begin{array}{c} \nearrow \\ \nwarrow \\ \swarrow \end{array} \right) + \# \left(\begin{array}{c} \nwarrow \\ \nearrow \\ \swarrow \end{array} \right) - \# \left(\begin{array}{c} \nwarrow \\ \swarrow \\ \nearrow \end{array} \right) - \# \left(\begin{array}{c} \swarrow \\ \nwarrow \\ \nearrow \end{array} \right) \right).$$

ОПРЕДЕЛЕНИЕ. Двойную точку фронта назовем *положительной*, если ориентации двух ветвей фронта в ней направлены в разные полуплоскости, и *отрицательной* в противном случае. Поскольку фронт не имеет вертикальных касательных, такое определение знака двойной точки корректно.



Положительные двойные точки

Отрицательные двойные точки

Числом Беннекена $\beta(L)$ лежандрова узла L называется разность между числом положительных и числом отрицательных двойных точек, увеличенная на половину общего числа точек возврата фронта L :

$$\beta(L) = \# \left(\begin{array}{c} \times \\ \oplus \end{array} \right) - \# \left(\begin{array}{c} \times \\ \ominus \end{array} \right) + \frac{1}{2} \# \left(\begin{array}{c} \langle \end{array} \right) + \frac{1}{2} \# \left(\begin{array}{c} \rangle \end{array} \right).$$

УПРАЖНЕНИЕ. Докажите, что числа $\mu(L)$ и $\beta(L)$ — лежандровы инварианты.

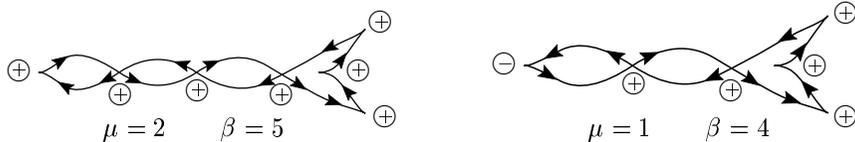
(Указание. Проверьте, что они не меняются при перестройках Фукса–Табачникова.)

Д. Беннекен [6] доказал, что число $\beta(L) + |\mu(L)|$ не меньше эйлеровой характеристики произвольной поверхности Зейферта узла L (см. статью В. Прасолова в данном выпуске, с. 116–126). Усиление этого неравенства см. в статьях [12, 23, 9]. (Предостережение. В статьях [6, 12, 23] используется противоположное нашему соглашению об ориентации объемлющего пространства, поэтому соответствующие неравенства в них имеют противоположный знак. Мы используем здесь соглашение, общепринятое в теории узлов.)

Для лежандровых узлов можно построить теорию инвариантов конечного порядка, аналогичную теории инвариантов Васильева из следующего раздела (см. [12]). При этом μ имеет порядок 0, а β — порядок 1. И, что самое удивительное, все остальные лежандровы инварианты конечного порядка, отличные от μ и β , являются инвариантами топологической эквивалентности, т. е. инварианты также и относительно четвертой перестройки Фукса–Табачникова.

Числа Маслова и Беннекена двух неузлов (топологически тривиальных узлов), приведенных в конце раздела 3.2, вычисляются на рисунке на следующей странице.

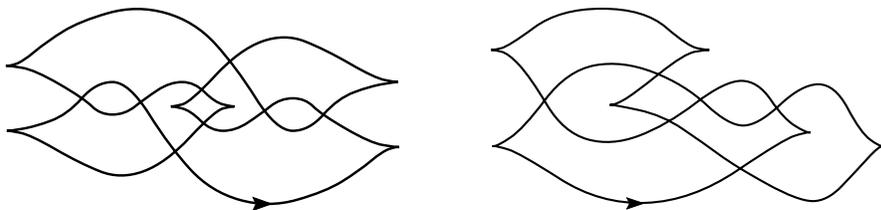
Различие μ и β для этих неузлов доказывает их лежандрову неэквивалентность.



Числа Маслова и Беннекена двух узлов

ТЕОРЕМА (ЭЛИАШБЕРГ[11]). *Два лежандровых узла лежандрово эквивалентны тогда и только тогда, когда они имеют одинаковые числа Маслова и Беннекена.*

Довольно долго был открыт вопрос об аналоге теоремы Элиашберга для произвольных узлов. Недавно Ю. Чеканов [8] предложил новый мощный инвариант лежандровой эквивалентности и доказал, что следующие два узла лежандрово неэквивалентны, хотя топологически они эквивалентны и имеют одинаковые числа Маслова и Беннекена.



Пара Чеканова

Мы оставляем читателю в качестве упражнения проверку того, что эти два узла топологически эквивалентны и что их числа Маслова равны 0, а числа Беннекена равны -1 .

5. ИНВАРИАНТЫ ВАСИЛЬЕВА

Инварианты Васильева представляют собой самый мощный и, в известном смысле, самый простой класс численных инвариантов узлов.

5.1. ИНВАРИАНТЫ ВАСИЛЬЕВА

Инварианты Васильева (инварианты конечного типа) появились около 1990 года в работах В. Васильева (см., например, [27]) по топологии дискриминантов в бесконечномерных пространствах отображений. Это не один какой-то инвариант, как полином Джонса, а целый класс инвариантов, выделяемый определенным свойством.

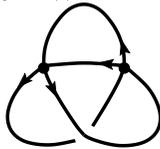
Исходная идея заключается в том, что нужно ввести в рассмотрение, помимо обычных узлов, еще так называемые *особые узлы*, и определить продолжение инвариантов, определенных первоначально для обычных узлов, на множество всех особых узлов.

ОПРЕДЕЛЕНИЕ. *Особым узлом* называется гладкое отображение $K : S^1 \rightarrow \mathbb{R}^3$, являющееся вложением всюду, кроме конечного числа простых двойных точек (т. е. точек трансверсального самопересечения).

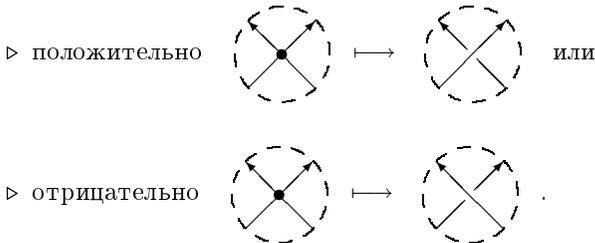
Множество всех особых узлов с n двойными точками, рассматриваемых с точностью до изотопии, мы обозначим через \mathcal{X}_n . В частности, \mathcal{X}_0 — это множество (классов эквивалентности) обычных узлов. Буквой \mathcal{X} без индекса мы будем обозначать объединение всех \mathcal{X}_n .

Диаграмма особого узла отличается от диаграммы обычного узла тем, что на ней, кроме точек прохода и перехода, есть еще точки самопересечения, которые на рисунках мы будем изображать жирными точками. Поскольку касательные к двум ветвям кривой в двойной точке узла различны, их можно считать различными и на диаграмме.

Вот пример особого узла с двумя двойными точками:



Каждую двойную точку особого узла можно разрешить двояким образом:



УПРАЖНЕНИЕ. Путем вращения головы и движения листа бумаги в пространстве убедитесь в том, что обе операции корректно определены на изотопических классах узлов и при этом не зависят от выбранной на узле ориентации.

ОПРЕДЕЛЕНИЕ. Пусть $f : \mathcal{X}_0 \rightarrow A$ — некоторый инвариант узлов. *Продолжение инварианта f на особые узлы* — это функция $\hat{f} : \mathcal{X} \rightarrow A$, совпадающая на \mathcal{X}_0 с f и удовлетворяющая соотношению Васильева

$$\hat{f}(\text{left diagram}) = \hat{f}(\text{right diagram}) - \hat{f}(\text{middle diagram}).$$

В этом соотношении фигурируют три особых узла, диаграммы которых совпадают между собой всюду, кроме указанного фрагмента. Оба узла, стоящие в правой части, имеют на одну двойную точку меньше, чем узел, стоящий в левой части.

Пользуясь этим соотношением рекуррентно, всякий инвариант, заданный первоначально на обычных узлах, можно продолжить на множество всех особых узлов.

ПРИМЕР. Сосчитаем значение продолженного по Васильеву полинома Джонса на особом узле, приведенном выше в качестве примера.

$$\begin{aligned}
\hat{v} \left(\text{Diagram 1} \right) &= v \left(\text{Diagram 2} \right) - v \left(\text{Diagram 3} \right) - v \left(\text{Diagram 4} \right) + \\
&+ v \left(\text{Diagram 5} \right) = v \left(\text{Diagram 6} \right) - v \left(\text{Diagram 7} \right) \\
&= t + t^3 - t^4 - 1.
\end{aligned}$$

ОПРЕДЕЛЕНИЕ. Инвариант $f : \mathcal{X}_0 \rightarrow A$ называется *инвариантом Васильева* порядка $\leq n$, если его продолжение на множество особых узлов обращается в нуль на всех узлах, имеющих более чем n точек самопересечения.

Множество \mathcal{V}_n всех инвариантов Васильева порядка $\leq n$ образует *линейное пространство*, так как линейная комбинация нескольких таких инвариантов всегда принадлежит \mathcal{V}_n . Нам в первую очередь будет интересовать вопрос, чему равна размерность этого пространства $d_n = \dim \mathcal{V}_n$, т. е. сколько существует линейно независимых (существенно различных) инвариантов Васильева данного порядка.

ПРИМЕР. Пространство \mathcal{V}_0 одномерно и состоит только из констант.

В самом деле, если $f \in \mathcal{V}_0$, то f обращается в нуль на любом особом узле, имеющем хотя бы одну двойную точку. В силу определения, это значит, что значение f на обычном узле не меняется при замене любого прохода на переход. Но такими действиями любой узел можно распутать, т. е. свести к тривиальному узлу. Значит, значение нашего инварианта f на любом узле равно его значению на тривиальном узле и, таким образом, f есть константа.

Аналогичное утверждение имеет место и для инвариантов Васильева порядка 1. Его доказательство ненамного сложнее предыдущего, и мы оставляем его читателю в качестве упражнения.

ЗАДАЧА. Докажите, что пространство \mathcal{V}_1 совпадает с \mathcal{V}_0 .

После этих двух примеров следующий факт кажется неожиданным, не так ли?

ПРИМЕР. Пространство \mathcal{V}_2 двумерно. Кроме констант, оно содержит еще один нетривиальный базисный элемент.

Явное описание нетривиального инварианта Васильева порядка 2 можно найти в статье Ю. Бурмана в этом выпуске, с. 94–115.

Причину, по которой нетривиальные инварианты Васильева появляются только в порядке 2, можно объяснить следующим образом. Пусть v — инвариант Васильева порядка n . Рассмотрим его значения на особых узлах, имеющих ровно n двойных точек. В силу соотношения Васильева и ввиду того, что v обращается в нуль на любом узле, у которого больше, чем n двойных точек, значение $v(K)$ не изменится, если узел K подвергнуть произвольной деформации (включая замены проходов на переходы и обратно), при которой двойные точки остаются на месте. Следовательно, значение $v(K)$ зависит лишь от порядка, в котором при обходе узла на нем встречаются двойные точки.

Если двойная точка одна (a), то она может встретиться только так: aa . Если же двойных точек две (a и b), то есть две существенно различных возможности, а именно, $aabb$ и $abab$. Различные варианты чередования двойных точек при обходе узла удобно изображать графически посредством хордовых диаграмм.

ОПРЕДЕЛЕНИЕ. *Хордовая диаграмма* степени n — это ориентированная окружность, в которой проведены n хорд, все концы которых различны.

Иными словами, хордовая диаграмма — это неориентированная гауссова диаграмма. Хордовые диаграммы рассматриваются как чисто комбинаторный объект: расстояние между концами хорд и форма хорд не имеют никакого значения, важен лишь порядок, в котором пары точек, соединенных хордами, следуют по кругу. Хордовая диаграмма — это то же самое, что слово в алфавите из n букв a_1, \dots, a_n , в котором каждая буква встречается ровно два раза. Такие слова рассматриваются с точностью до циклических перестановок входящих в них букв и произвольной перенумерации переменных a_1, \dots, a_n .

ПРИМЕРЫ. Существует

1. одна хордовая диаграмма степени 1: 

2. две хордовые диаграммы степени 2: 

3. пять хордовых диаграмм степени 3: 

(Подразумевается, что на всех этих рисунках окружность ориентирована против часовой стрелки.)

Каждому особому узлу K , имеющему n двойных точек, отвечает определенная хордовая диаграмма $\chi(K)$ степени n .

ПРИМЕР. Для особого узла, приведенного выше в качестве примера, имеем:

$$\chi\left(\text{[Diagram of a knot with 3 crossings]}\right) = \text{[Diagram of a circle with 3 chords]}.$$

Пусть CD_n — множество всех хордовых диаграмм степени n . Их число $\#(\text{CD}_n)$ дает оценку сверху на зазор между размерностью пространства \mathcal{V}_n и размерностью пространства $\mathcal{V}_{n-1} \subset \mathcal{V}_n$. В самом деле, мы только что объяснили, как по инварианту Васильева порядка n построить функцию на множестве хордовых диаграмм порядка n . Если обозначить пространство всех функций на множестве хордовых диаграмм порядка n через \mathcal{F}_n , то мы получаем линейное отображение

$$\sigma : \mathcal{V}_n \rightarrow \mathcal{F}_n.$$

Чему равно ядро этого линейного отображения, т. е. какие инварианты Васильева принимают значение 0 на всех узлах, имеющих ровно n двойных точек? По определению, это инварианты Васильева порядка $\leq n-1$. Итак, ядро нашего отображения равно в точности \mathcal{V}_{n-1} , и мы имеем линейное вложение факторпространства $\mathcal{V}_n/\mathcal{V}_{n-1}$ (инвариантов Васильева порядка $\leq n$, рассматриваемых с точностью до прибавления инвариантов порядка $\leq n-1$):

$$\bar{\sigma} : \mathcal{V}_n/\mathcal{V}_{n-1} \rightarrow \mathcal{F}_n.$$

Отсюда следует, что размерности всех пространств \mathcal{V}_n конечны, причем

$$\dim \mathcal{V}_n / \mathcal{V}_{n-1} = \dim \mathcal{V}_n - \dim \mathcal{V}_{n-1} \leq \dim \mathcal{F}_n = \#(\text{CD}_n).$$

А чему равен образ отображения $\bar{\sigma}$? Как можно охарактеризовать функции на множестве хордовых диаграмм, принадлежащие образу этого отображения, т. е. происходящие из инвариантов Васильева? Такие функции называются *весовыми системами*, а ответ на заданный вопрос дает следующая теорема.

ТЕОРЕМА ВАСИЛЬЕВА–КОНЦЕВИЧА. (1) (*В. Васильев*) Всякая весовая система удовлетворяет

(а) одночленным соотношениям

$$f \left(\text{Diagram with one chord} \right) = 0,$$

(б) четырехчленным соотношениям

$$f \left(\text{Diagram 1} \right) - f \left(\text{Diagram 2} \right) + f \left(\text{Diagram 3} \right) - f \left(\text{Diagram 4} \right) = 0.$$

(2) (*М. Концевич*) Любая функция на множестве CD_n , удовлетворяющая а) и б), происходит из некоторого инварианта Васильева порядка $\leq n$.

На рисунках хордовых диаграмм пунктирные линии означают, что в этих местах могут встречаться еще какие-то (ненарисованные) хорды, одни и те же для всех диаграмм, входящих в данную формулу. На участках окружности, изображенных сплошной линией, все имеющиеся хорды честно указаны.

Итак, чтобы определить число $d_n = \dim \mathcal{V}_n - \dim \mathcal{V}_{n-1}$, нужно составить и решить систему линейных уравнений, в которой неизвестные — это значения весовой системы на хордовых диаграммах степени n , а уравнения получаются из всевозможных 1- и 4-членных соотношений. Вместо того, чтобы говорить о значениях функции на множестве диаграмм, можно в качестве неизвестных взять сами эти диаграммы, рассматривая их как буквы некоторого нового алфавита \bigcirc , $\bigcirc \bigcirc$, $\bigcirc \times$, ... (в самом деле, чем эти буквы хуже латинских или греческих?)

ОПРЕДЕЛЕНИЕ. *Пространство хордовых диаграмм* \mathcal{A}_n — это линейное пространство, порожденное всеми хордовыми диаграммами степени n по модулю одночленных и четырехчленных соотношений

$$\begin{aligned} \text{Diagram with one chord} &= 0, \\ \text{Diagram 1} - \text{Diagram 2} + \text{Diagram 3} - \text{Diagram 4} &= 0. \end{aligned}$$

Иначе говоря, \mathcal{A}_n — это пространство решений однородной системы линейных уравнений, состоящей из всех указанных соотношений, относительно набора неизвестных, соответствующих хордовым диаграммам степени n .

ПРИМЕРЫ. 1. Пространство \mathcal{A}_1 нульмерно, ибо оно порождено одним элементом , который равен нулю в силу 1-членного соотношения.

2. Пространство \mathcal{A}_2 имеет размерность 1, поскольку  = 0, а нетривиальных 4-членных соотношений в этом случае нет. В качестве базисного элемента можно взять .

3. Размерность пространства \mathcal{A}_3 равна числу свободных переменных в системе линейных уравнений

$$\begin{aligned} \text{Diagram 1} &= 0, \\ \text{Diagram 2} &= 0, \\ \text{Diagram 3} &= 0, \\ \text{Diagram 4} - \text{Diagram 5} &= 0, \\ \text{Diagram 6} - 2 \text{Diagram 7} + \text{Diagram 8} &= 0, \end{aligned}$$

то есть 1. Первые три уравнения выписанной системы суть одночленные соотношения, последние два — четырехчленные. Базисным элементом может служить диаграмма  или .

4. Для пространства \mathcal{A}_4 мы приведем лишь ответ, не выписывая систему уравнений: $\dim \mathcal{A}_4 = 3$, а в качестве базиса можно взять диаграммы ,  и .

С ростом n число переменных и уравнений в системе для определения размерности пространства \mathcal{A}_n очень быстро растет. Используя описанный прямой подход, на компьютере удалось сосчитать размерность и базис \mathcal{A}_n только для $n \leq 9$ (чтобы найти последнее значение $\dim \mathcal{A}_9$, Д. Бар-Натану пришлось в январе 1993 года одолжить для своей рабочей станции на несколько дней дополнительно 32 Мб оперативной памяти). К настоящему времени точные значения размерности пространств \mathcal{A}_n и, следовательно, \mathcal{V}_n , известны до $n = 12$:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
$\dim \mathcal{P}_n$	0	0	1	1	2	3	5	8	12	18	27	39	55
$\dim \mathcal{A}_n$	1	0	1	1	3	4	9	14	27	44	80	132	232
$\dim \mathcal{V}_n$	1	1	2	3	6	10	19	33	60	104	184	316	548

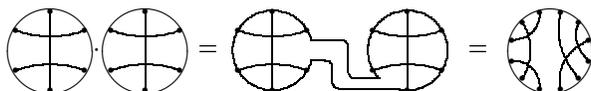
Ответ для $n = 10, 11$ и 12 получен Яном Кнайссером [16] в 1997 году с использованием более тонкой техники, о которой мы расскажем в следующем разделе. Там же будет объяснено, что такое \mathcal{P}_n и как по числам $\dim \mathcal{P}_n$ можно найти числа $\dim \mathcal{A}_n$.

5.2. АЛГЕБРА ХОРДОВЫХ ДИАГРАММ

Оказывается, что лучше понять устройство линейных пространств \mathcal{A}_n и дальше продвинуться в конкретных вычислениях помогает мультипликативная структура, которая имеется в прямой сумме всех пространств хордовых диаграмм

$$\mathcal{A} = \bigoplus_{n=0}^{\infty} \mathcal{A}_n = \mathcal{A}_0 \oplus \mathcal{A}_1 \oplus \mathcal{A}_2 \oplus \dots$$

Произведение двух хордовых диаграмм определяется так: две окружности разрываются и склеиваются в одну:



Хордовая диаграмма, которая получается в правой части этого соотношения, зависит, вообще говоря, от того места, где разрываются окружности. Замечательно, что с учетом 4-членных соотношений (в факторпространстве \mathcal{A}) умножение определено корректно, т. е. не зависит от произвола в определении.

ЗАДАЧА. Докажите этот факт.

Таким образом, бесконечномерное пространство \mathcal{A} является алгеброй (скажем, над полем рациональных чисел).

Оказывается, что в этой алгебре можно выделить счетное семейство элементов, такое что всякий элемент алгебры однозначно записывается в виде многочлена от этих элементов. Значит, \mathcal{A} фактически представляет собой алгебру многочленов от бесконечного числа переменных.

Выбор такого семейства образующих основан на следующих соображениях. Для любого n в пространстве \mathcal{A}_n есть замечательное линейное подпространство, называемое *пространством примитивных элементов* и обозначаемое \mathcal{P}_n . Если в каждом пространстве \mathcal{P}_n , $n = 1, 2, \dots$ выбрать любой базис, то их объединение может служить системой образующих полиномиальной алгебры \mathcal{A} . Опишем один из способов построения примитивного пространства \mathcal{P}_n .

ОПРЕДЕЛЕНИЕ. *Диаграмма Фейнмана* степени n — это регулярный трехвалентный граф с $2n$ вершинами, в котором выделен некоторый ориентированный цикл, называемый *петлей Уилсона*, и в каждой вершине, не лежащей на петле Уилсона, задан циклический порядок выходящих из нее ребер.

Хордовые диаграммы являются частным случаем диаграмм Фейнмана (у них все трехвалентные вершины лежат на петле Уилсона). Каждую диаграмму Фейнмана f можно превратить в линейную комбинацию хордовых диаграмм $\rho(f)$, многократно применяя следующее соотношение для разрешения тройных точек, смежных с петлей Уилсона:



ПРИМЕР.

$$\begin{aligned}
 & \text{Diagram with a loop} \mapsto \text{Diagram with two lines meeting at a vertex} - \text{Diagram with two lines meeting at a vertex and a loop} \\
 & \mapsto (\text{Diagram with two lines meeting at a vertex} - \text{Diagram with two lines meeting at a vertex and a loop}) - (\text{Diagram with two lines meeting at a vertex} - \text{Diagram with two lines meeting at a vertex and a loop}) \\
 & = -2 (\text{Diagram with two parallel lines} - \text{Diagram with two lines crossing}) = -2 \text{Diagram with two lines crossing}
 \end{aligned}$$

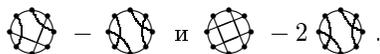
Степень связной диаграммы Фейнмана — это половина общего числа вершин графа. Если Φ_n — множество всех диаграмм Фейнмана степени n и $f \in \Phi_n$, то $\rho(f)$ можно рассматривать как элемент пространства \mathcal{A}_n . Диаграмма Фейнмана называется *связной*, если соответствующий граф остается связным после отбрасывания петли Уилсона.

ОПРЕДЕЛЕНИЕ. Пространство примитивных элементов \mathcal{P}_n ($n \geq 1$) — это линейная оболочка элементов $\rho(f)$ для всех связных диаграмм $f \in \Phi_n$.

ПРИМЕРЫ. Пространство $\mathcal{P}_1 = \mathcal{A}_1$ нулевое.

Пространства \mathcal{P}_2 и \mathcal{P}_3 одномерны и совпадают, соответственно, с пространствами \mathcal{A}_2 и \mathcal{A}_3 .

В трехмерном пространстве \mathcal{A}_4 примитивное подпространство \mathcal{P}_4 двумерно

и порождено элементами  $- 2$ .

Работать с примитивными подпространствами \mathcal{P}_n проще, потому что они намного меньше, чем \mathcal{A}_n . Предположим, что мы нашли размерности всех \mathcal{P}_n до какого-то места. Спрашивается, как тогда найти начальный участок последовательности $\dim \mathcal{A}_n$? Поясним это на примере.

Допустим, мы знаем, что значения $\dim \mathcal{P}_n$ для $n = 1, 2, 3, 4, 5, 6, 7$ равны $0, 1, 1, 2, 3, 5, 8$. Пусть p_2 — базисный элемент \mathcal{P}_2 , p_3 — базисный элемент \mathcal{P}_3 , а базисы пространств \mathcal{P}_4 и \mathcal{P}_5 суть соответственно p_{41}, p_{42} и p_{51}, p_{52}, p_{53} . Степенью элемента $p \in \mathcal{P}_n$ называется число n , а степень произведения равна сумме степеней сомножителей. Посмотрим, какие можно составить мономы из данных базисных элементов до степени 7 включительно.

Мономы степени 4: p_2^2 (всего 1).

Мономы степени 5: $p_2 p_3$ (всего 1).

Мономы степени 6: $p_3^2, p_2^2 p_2, p_2 p_{41}, p_2 p_{42}$ (всего 4).

Мономы степени 7: $p_2^2 p_3, p_2 p_{51}, p_2 p_{52}, p_2 p_{53}, p_3 p_{41}, p_3 p_{42}$ (всего 6).

Базис пространства \mathcal{A}_n составляют сами примитивные элементы и их произведения соответствующей степени. Для $n = 4, 5, 6, 7$ получаем: $\dim \mathcal{A}_4 = 2 + 1 = 3$, $\dim \mathcal{A}_5 = 3 + 1 = 4$, $\dim \mathcal{A}_6 = 5 + 4 = 9$, $\dim \mathcal{A}_7 = 8 + 6 = 14$.

Оказывается (это открыл П. Вожел [28] в 1995 году), что в пространстве всех примитивных элементов $\mathcal{P} = \bigoplus_{n=1}^{\infty} \mathcal{P}_n$ есть еще внутренняя мультипликативная структура. Именно она и позволила Я. Кнайслеру в 1997 году разработать новый, более эффективный алгоритм нахождения базиса в пространстве инвариантов Васильева и превзойти рекорд Д. Бар-Натана 1993 года.

6. ИНТЕГРАЛ КОНЦЕВИЧА

Интеграл Концевича изобретен в 1992 году [18] как средство доказательства сформулированной выше (с. 81) теоремы Васильева–Концевича. Он является далеко идущим обобщением интегральной формулы Гаусса, выражающей коэффициент зацепления двух замкнутых кривых в пространстве.

Представим трехмерное пространство \mathbb{R}^3 как прямое произведение комплексной прямой \mathbb{C} с координатой z и вещественной прямой \mathbb{R} с координатой t . Вложим узел K в пространство $\mathbb{R}^3 = \mathbb{C}_z \times \mathbb{R}_t$ так, чтобы координата t была функцией Морса на K . Это значит, что во всех точках, где t как функция параметра на кривой имеет нулевую производную, ее вторая производная не должна обращаться в нуль, а значения t во всех таких точках (*критические значения*) должны быть различны между собой.

ОПРЕДЕЛЕНИЕ. *Интеграл Концевича* узла K — это следующий элемент пополнения $\overline{\mathcal{A}}$ алгебры хордовых диаграмм:

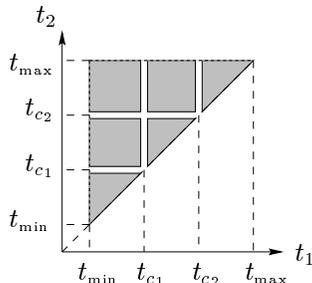
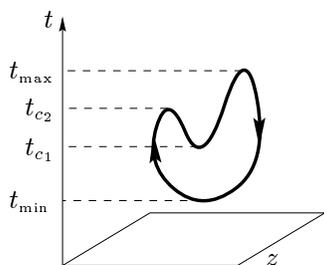
$$Z(K) = \sum_{m=0}^{\infty} \frac{1}{(2\pi i)^m} \int_{\substack{t_{\min} < t_1 < \dots < t_m < t_{\max} \\ t_j \text{ не критические}}} \sum_{P=\{(z_j, z'_j)\}} (-1)^{\downarrow} D_P \prod_{j=1}^m \frac{dz_j - dz'_j}{z_j - z'_j}.$$

Пополнение $\overline{\mathcal{A}}$ алгебры \mathcal{A} — это, по определению, множество всех бесконечных формальных рядов вида $\sum_m c_m a_m$, где $c_m \in \mathbb{C}$ — численные коэффициенты, а $a_m \in \mathcal{A}_m$ — элементы пространства хордовых диаграмм степени m .

6.1. WHO IS WHO В ИНТЕГРАЛЕ КОНЦЕВИЧА

Действительные числа t_{\min} и t_{\max} суть минимальное и максимальное значения функции t на K .

Область интегрирования представляет собой m -мерный симплекс $t_{\min} < t_1 < \dots < t_m < t_{\max}$, разделенный критическими значениями на некоторое число *связных компонент*. Например, для следующего вложения неузла и $m = 2$ область интегрирования состоит из шести компонент связности и выглядит так:



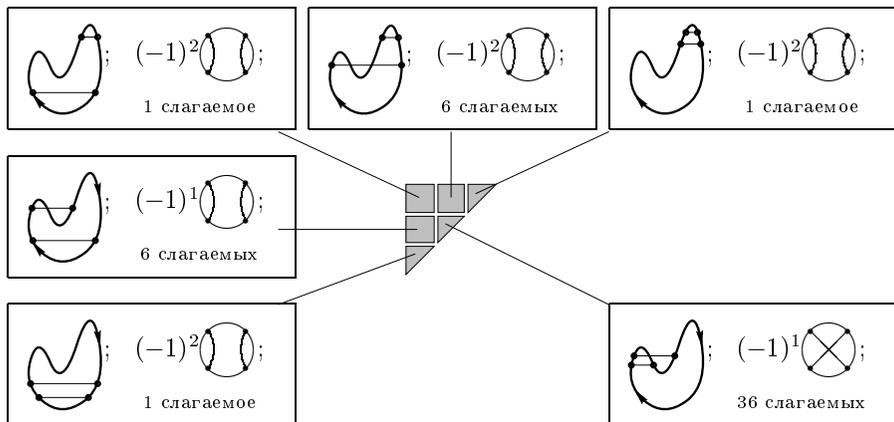
Число слагаемых в подынтегральном выражении постоянно для каждой компоненты связности, но может быть различным для разных компонент. Часть узла в полосе, заключенной между двумя критическими уровнями функции t , представляет собой набор кривых, каждая из которых однозначно параметризована переменной t . В каждой плоскости $\{t = t_j\} \subset \mathbb{R}^3$ выберем неупорядоченную пару различных точек (z_j, t_j) и (z'_j, t_j) , лежащих на узле K . Обозначим

через $P = \{(z_j, z'_j)\}$ систему из m таких пар для $j = 1, \dots, m$. Под интегралом стоит сумма по всем способам выбора системы P . В произвольной связной компоненте области интегрирования фиксация системы P означает, что для каждой переменной t_j выбрана некоторая пара ветвей узла в соответствующей полосе пространства.

В приведенном выше примере для связной компоненты $\{t_{\min} < t_1 < t_{c_1}, t_{c_2} < t_2 < t_{\max}\}$ есть лишь одна пара точек на уровнях $\{t = t_1\}$ и $\{t = t_2\}$. Следовательно, для этой компоненты сумма под интегралом состоит только из одного слагаемого. Напротив, для компоненты $\{t_{\min} < t_1 < t_{c_1}, t_{c_1} < t_2 < t_{c_2}\}$ у нас по-прежнему есть лишь один выбор на уровне $\{t = t_1\}$, но плоскость $\{t = t_2\}$ пересекает узел K уже по четырем точкам, поэтому мы имеем $\binom{4}{2} = 6$ возможных пар (z_2, z'_2) и общее число слагаемых равно шести.

Для данного выбора системы P символ ' \downarrow ' обозначает число точек (z_j, t_j) или (z'_j, t_j) из P , в которых координата t убывает при движении по узлу в положительном направлении.

Зафиксируем систему пар P . Рассмотрим узел K как ориентированную окружность и соединим на ней точки (z_j, t_j) и (z'_j, t_j) хордами. Получится хордовая диаграмма порядка m . Эта хордовая диаграмма является элементом алгебры \mathcal{A} , который мы обозначаем через D_P . На следующем рисунке для каждой связной компоненты показано одно из возможных разбиений на пары, соответствующая хордовая диаграмма со знаком $(-1)^{\downarrow}$ и указано число слагаемых подынтегрального выражения (некоторые из которых равны нулю в силу одночленного соотношения).



На каждой компоненте связности переменные z_j и z'_j являются гладкими функциями переменной t_j . Подставляя эти функции в выражение $\prod_{j=1}^m \frac{dz_j - dz'_j}{z_j - z'_j}$, мы получаем дифференциальную форму от переменных t_1, \dots, t_m , определенную на симплексе в m -мерном пространстве этих переменных. Ее и следует проинтегрировать. Это даст коэффициент при соответствующей хордовой диаграмме.

Описанная конструкция работает при всех значениях $m > 0$. Для $m = 0$ по определению полагают, что получится хордовая диаграмма порядка 0 (окружность без хорд, представляющая единичный элемент алгебры \mathcal{A}) с коэффициентом 1.

6.2. УНИВЕРСАЛЬНЫЙ ИНВАРИАНТ ВАСИЛЬЕВА

Можно доказать, что интеграл $Z(K)$, во-первых, сходится для любого узла, расположенного в пространстве указанным выше способом, а во-вторых, не меняется при гладких изотопиях узла, при которых сохраняется число критических точек функции t . Ввиду того, что узел — замкнутая кривая, появляться и исчезать критические точки могут только парами. Оказывается, что при этом выражение $Z(K)$ меняется вполне определенным образом, а именно, если узел K_1 получается из узла K добавлением пары соседних критических точек, то имеет место равенство

$$Z(K_1) = Z(K) \cdot Z\left(\text{шляпка}\right).$$

Обозначим через H входящий в эту формулу тривиальный узел, вложение которого в пространство имеет два максимума и два минимума, и рассмотрим выражение

$$I(K) = \frac{Z(K)}{Z(H)^{c/2}},$$

где c обозначает число критических точек функции t на K , а частное — это деление в алгебре \mathcal{A} , осуществляемое на основе правила $(1+a)^{-1} = 1 - a + a^2 - a^3 + \dots$.

Из сказанного выше ясно, что $I(K)$ не меняется также и при добавлении (и удалении) пар критических точек. Следовательно, $I(K)$ представляет собой настоящий инвариант узла. Он называется *универсальным инвариантом Васильева*, потому что (в известном смысле) все инварианты Васильева входят в него как коэффициенты.

Исходный интеграл $Z(K)$ называют еще *предварительным интегралом Концевича*, а подправленное выражение $I(K)$ — *окончательным интегралом Концевича*.

ПРИМЕРЫ. Пусть O — неузел, вложенный в пространство как окружность, стоящая вертикально. Тогда:

$$Z(O) = 1,$$

$$I(O) = 1/Z(H),$$

$$I(H) = Z(H)/(Z(H)^2) = 1/Z(H).$$

Последние два ответа совпадают. Это неудивительно, ведь I — инвариант узла, а узлы O и H эквивалентны. В следующем пункте речь пойдет именно об этом выражении.

6.3. ИНТЕГРАЛ КОНЦЕВИЧА ДЛЯ НЕУЗЛА

Интеграл Концевича — довольно-таки сложный объект. В течение нескольких лет никто не умел вычислять его даже для тривиального узла: были известны

лишь коэффициенты при некоторых хордовых диаграммах в бесконечной сумме. (Конечно, здесь имеется ввиду окончательный интеграл Концевича $I(O)$, ибо предварительный $Z(O)$ равен просто 1, как мы уже отмечали.)

В 1997 году появилась гипотеза Д. Бар-Натана с соавторами [3] (доказанная в 1998 году в работе [4]) о том, чему равна вся сумма:

$$I(O) = \exp \sum_{n=0}^{\infty} b_{2n} w_{2n} = 1 + \left(\sum_{n=0}^{\infty} b_{2n} w_{2n} \right) + \frac{1}{2} \left(\sum_{n=0}^{\infty} b_{2n} w_{2n} \right)^2 + \dots$$

Здесь b_{2n} — *модифицированные числа Бернулли*, т. е. коэффициенты следующего ряда Тейлора:

$$\sum_{n=0}^{\infty} b_{2n} x^{2n} = \frac{1}{2} \ln \frac{e^{x/2} - e^{-x/2}}{x/2},$$

($b_2 = 1/48$, $b_4 = -1/5760$, ...), а w_{2n} — *колеса*, т. е. диаграммы вида

$$w_2 = \text{circle with 2 dots}, \quad w_4 = \text{circle with 4 dots}, \quad w_6 = \text{circle with 6 dots}, \quad \dots$$

Произведения колес, возникающие при вычислении экспоненты через ряд Тейлора, понимаются как несвязное объединение диаграмм, а сами колеса (и их объединения) интерпретируются как линейные комбинации диаграмм Фейнмана, равные среднему арифметическому по всем $(2n)!$ способам приклеить петлю Уилсона к одновалентным вершинам колеса.

Например,

$$\begin{aligned} \text{circle with 4 dots} &= \frac{1}{24} \left(\begin{array}{c} \text{circle with 4 dots} \\ \text{rectangle with 4 dots} \end{array} \right) \\ &+ \begin{array}{c} \text{circle with 4 dots} \\ \text{rectangle with 4 dots} \end{array} \\ &+ \begin{array}{c} \text{circle with 4 dots} \\ \text{rectangle with 4 dots} \end{array} \\ &+ \begin{array}{c} \text{circle with 4 dots} \\ \text{rectangle with 4 dots} \end{array} \end{aligned}$$

Если пристально посмотреть на эти рисунки, можно убедиться в том, что 16 слагаемых выписанной суммы эквивалентны первой диаграмме, а еще 8 — второй. Приводя подобные члены, получаем:

$$\text{circle with 4 dots} = \frac{1}{3} \begin{array}{c} \text{circle with 4 dots} \\ \text{rectangle with 4 dots} \end{array} + \frac{2}{3} \begin{array}{c} \text{circle with 4 dots} \\ \text{rectangle with 4 dots} \end{array}.$$

Разрешение внутренних тройных точек диаграмм Фейнмана по правилу, описанному на с. 83, приводит к такому ответу:

$$\text{Diagram} = \text{Diagram}_1 - \frac{10}{3} \text{Diagram}_2 + \frac{4}{3} \text{Diagram}_3.$$

Подобным образом (только проще) можно найти выражение через хордовые диаграммы также для мономов w_2 и w_2^2 . Подставляя найденные выражения в ряд

$$I(O) = 1 + \frac{1}{48} w_2 + \frac{1}{4608} w_2^2 - \frac{1}{5760} w_4 + \dots,$$

мы получим следующий результат.

Бесконечный ряд, выражающий интеграл Концевича от неузла, начинается так:

$$I(O) = 1 - \frac{1}{24} \text{Diagram}_1 - \frac{1}{5760} \text{Diagram}_2 + \frac{1}{1152} \text{Diagram}_3 + \frac{1}{2880} \text{Diagram}_4 + \dots$$

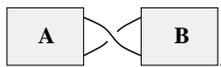
Все коэффициенты этого ряда рациональны. Этот факт верен для любого узла. Он показывает удивительную связь между теорией узлов и арифметикой.

7. ЗНАМЕНИТЫЕ ПРОБЛЕМЫ ТЕОРИИ УЗЛОВ

В этом разделе мы приводим формулировки нескольких знаменитых проблем теории узлов. Некоторые из них не поддавались решению в течение десятков лет и были решены сравнительно недавно, а некоторые открыты до сих пор.

7.1. РЕШЕННЫЕ ПРОБЛЕМЫ

1. *Существуют ли необратимые узлы?* (Решена положительно в 1962 году [26]).
2. *Определяется ли класс эквивалентности неориентированного узла K топологическим типом его дополнения $\mathbb{R}^3 \setminus K$?* (Решена положительно в 1989 году [13]).
3. В конце прошлого века П. Тейт сформулировал несколько знаменитых гипотез, две из которых совсем недавно были доказаны, а одна опровергнута. В первых двух гипотезах речь идет об альтернированных диаграммах (таких, у которых при движении по узлу проходы и переходы чередуются).
Гипотеза 1. *Редуцированная, т. е. не имеющая вида*



альтернированная диаграмма имеет наименьшее возможное число точек самопересечения среди всех диаграмм данного узла. Эта гипотеза доказана независимо в трех работах [17], [25] и [21].

Гипотеза 2. (Tait flipping conjecture). *От одной альтернированной диаграммы узла к другой можно перейти последовательностью переворачиваний следующего вида:*



Эта гипотеза доказана в 1993 году [20].

Гипотеза 3. Узел, для которого минимальное число пересечений нечетно, всегда зеркален, т. е. эквивалентен своему зеркальному отражению. Эта гипотеза просуществовала 104 года и была опровергнута только в 1994 году М. Тислтуэйтом (опубликовано в [15]).

4. Существует ли алгоритм распознавания, эквивалентны ли два данных узла, по их диаграммам? Проблема решена положительно. В последние годы изобретено несколько таких алгоритмов (С. В. Матвеев, Дж. Бирман, Е. Фоминых, И. Дынников). Надо только отметить, что полное доказательство правильности алгоритма проведено лишь для одного из них, самого трудоемкого и имеющего чисто теоретическое значение. Подробности см. в статье [19].

7.2. НЕРЕШЕННЫЕ ПРОБЛЕМЫ

1. Как распознать обратимость узла?
2. Пусть $C(n)$ — число различных простых неориентированных узлов с минимальным числом пересечений n . Как ведет себя функция $C(n)$ с ростом n ? Про нее очень мало известно. Например, до сих пор нет строгого доказательства, что $C(n)$ монотонно возрастает с ростом n , хотя этот факт кажется очевидным. На компьютере найдены значения $C(n)$ для $n \leq 16$ (последние три значения нашли в 1997 году Хост, Тислтуэйт и Уикс [15]):

3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	2	3	7	21	49	165	552	2176	9988	46972	253293	1388705

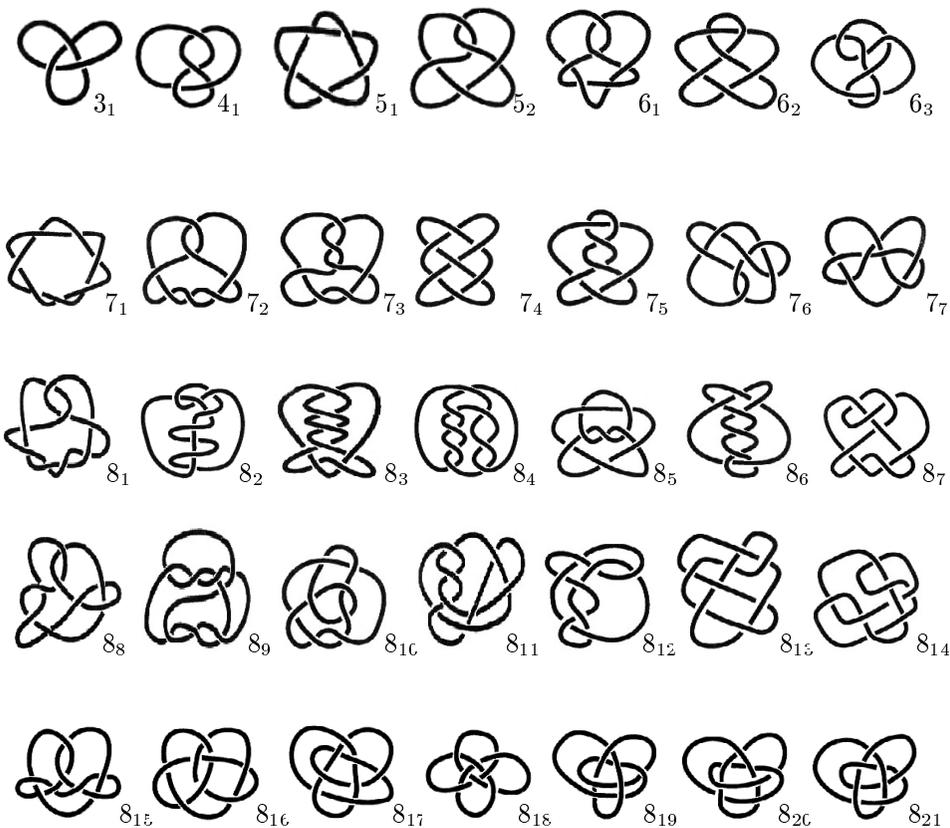
3. Если некоторый узел имеет две диаграммы K_+ и K_- , которые получаются друг из друга заменой одного прохода на переход, то верно ли, что удаление этой двойной точки приводит к распадению узла на две компоненты связности?
4. Правда ли, что число пересечений аддитивно относительно связной суммы узлов: $c(K_1 \# K_2) = c(K_1) + c(K_2)$?
5. Существует ли нетривиальный узел, полином Джонса которого равен 1, т. е. полиному Джонса тривиального узла? Доказано, что пример, если он существует, должен быть достаточно сложным: число пересечений такого узла не меньше 18 (см. [10]).
6. Найти ограничение сверху на число перестроек Рейдемейстера, переводящих одну из двух данных диаграмм узла в другую.
7. Найти критерий простоты узла в терминах его диаграммы.
8. Является ли универсальный инвариант Васильева полным инвариантом (ориентированного) узла? То есть, верно ли, что если два узла K_1 и K_2 различны, то найдется инвариант конечного порядка v такой, что $v(K_1) \neq v(K_2)$?

9. Верно ли, что инварианты Васильева различают обращение узла? Иными словами, существует ли пара взаимно обратных узлов K, K' , неэквивалентных друг другу, и такой инвариант Васильева v , что $v(K) \neq v(K')$?

Этот вопрос допускает следующую комбинаторную формулировку: *существует ли хордовая диаграмма, которая неэквивалентна своему зеркальному отражению по модулю 4-членных соотношений?* Известно, что такая диаграмма, если она существует, должна иметь степень не меньше 13.

ПРИЛОЖЕНИЕ. ТАБЛИЦА УЗЛОВ

В таблице приведены все простые неориентированные узлы с числом пересечений, не превосходящим 8.



СПИСОК ЛИТЕРАТУРЫ

- [1] *Colin C. Adams*. The knot book. NY: W.H.Freeman and Company, 1994.
- [2] *D. Bar-Natan*. On the Vassiliev knot invariants // *Topology*, V. 34, 1995. P. 423–472.
- [3] *D. Bar-Natan, S. Garoufalidis, L. Rozansky, D. Thurston*. Wheels, wheeling, and the Kontsevich integral of the unknot. Preprint, 1997.
- [4] *D. Bar-Natan, Thang Le, Dylan Thurston*. The Kontsevich integral of the unknot. In preparation.
- [5] *Joan S. Birman, Michael D. Hirsch*. A new algorithm for recognizing the unknot. <http://xxx.lanl.gov/abs/math.GT/9801126>.
- [6] *D. Bennequin*. Entrelacements et équations de Pfaff. *Astérisque*, V. 107–108, 1983. P. 87–161. Русский перевод: УМН, т. 44, вып. 3, 1989. С. 3–53.
- [7] *Р. Кроуэлла, Р. Фокс*. Теория узлов. М.: Мир, 1967.
- [8] *Yu. Chekanov*. Differential algebras of Legendrian links. Preprint, 1997.
- [9] *S. Chmutov, V. Goryunov*. Polynomial invariants of Legendrian links and plane fronts // In *Topics in Singularity Theory. V.I. Arnold's 60-th Anniversary Collection* (A. Khovanskii, A. Varchenko, V. Vassiliev, eds.). Amer. Math. Soc. Transl. Ser. 2, V. 180, 1997. P. 25–44.
- [10] *O. T. Dasbach, S. Hougardy*. Does the Jones Polynomial Detect Unknottedness? *J. of Experimental Mathematics*, Vol. 6, No. 1, 1997. P. 51–56.
- [11] *Ya. Eliashberg*. Legendrian and transversal knots in tight contact 3-manifolds // In *Topological methods in modern mathematics*, Publish or Perish, Houston, 1993. P. 171–193.
- [12] *D. Fuchs, S. Tabachnikov*. Invariants of Legendrian and transverse knots in the standard contact space. *Topology*, V. 36, 1997. P. 1025–1053.
- [13] *C. Gordan, J. Luecke*. Knots are determined by their complements. *J. Amer. Math. Soc.* Vol. 2, 1989. P. 371–415.
- [14] *W. Haken*. Theorie der Normalflächen. *Acta Math.*, Vol. 105, 1961. P. 245–375.
- [15] *J. Hoste, M. Thistlethwaite and J. Weeks*. The first 1,701,936 knots. *Mathematical Intelligencer*. Vol. 20, 1998. P. 33–48.
- [16] *Jan Kneissler*. The number of primitive Vassiliev invariants up to degree twelve. Preprint (June 1997).
- [17] *L. Kauffman*. State models and the Jones polynomial. *Topology*. Vol. 26, 1987. P. 395–407.
- [18] *M. Kontsevich*. Vassiliev's knot invariants. *Adv. in Soviet Math.* Vol. 16, Part 2, 1993. P. 137–150.
- [19] *С. В. Матвеев*. Классификация достаточно больших трехмерных многообразий. УМН. Т. 52, вып. 5, 1997. С. 147–174.

- [20] *W. Menasco, M. Thistlethwaite*. A classification of alternating links. *Annals of Mathematics*. Vol. 138, 1993. P. 113–171.
- [21] *K. Murasugi*. Jones polynomial and classical conjectures in knot theory. *Topology*. Vol. 26, 1987. P. 184–194.
- [22] *В. В. Прасолов, А. Б. Сосинский*. Узлы, зацепления, косы и трехмерные многообразия. М: МЦНМО, 1997.
- [23] *S. Tabachnikov*. Estimates for the Bennequin number of legendrian links from state models for knot polynomials. *Mathematical Research Letters*. Vol. 4, 1997. P. 143–156.
- [24] *J. Hoste, M. Thistlethwaite, J. Weeks*. Таблица узлов и программа KnotScape. <http://www.math.utk.edu/~morwen> (домашняя страница М.Тислтуэйта).
- [25] *M. Thistlethwaite*. A spanning tree expansion for the Jones polynomial. *Topology*. Vol. 26, 1987. P. 297–309.
- [26] *H. F. Trotter*. Non-invertible knots exist. *Topology*. Vol. 2, 1964. P. 275–280.
- [27] *V. A. Vassiliev*. Cohomology of knot spaces // *Theory of Singularities and Its Applications* (ed. V. I. Arnold), *Advances in Soviet Math*. Vol. 1, 1990. P. 23–69.
- [28] *P. Vogel*. Algebraic structures on modules of diagrams. Institut de Mathématiques de Jussieu, Prépublication 32, August 1995.
- [29] *W. Whitten*. Knot complements and groups. *Topology*. Vol. 26, 1987. P. 41–44.

Длинные кривые, гауссовы диаграммы и инварианты

Ю. М. Бурман

Кривая есть след движущейся точки. Это утверждение кажется нам сегодня простым и естественным, но для его формулировки математикам потребовалось не одно тысячелетие. Попытка объяснить, что такое кривая, имеется в «Началах» Евклида, но только в конце XIX века французский математик Камилл Жордан сформулировал приведенное выше определение — первое, удовлетворяющее современным стандартам математической строгости (правда, следует заметить, что сами эти стандарты родились примерно в то же время). То же самое происходит и с изучением свойств кривых — о них немало было известно уже в древности, и вместе с тем большинство результатов, излагаемых в этой статье, было получено три–пять лет назад.

Наша статья посвящена, в основном, теории деформаций плоских кривых. Мы изучаем вопрос, когда две кривые можно непрерывно продеформировать друг в друга, и какие события — например, тройные самопересечения — могут или обязаны произойти при такой деформации. Эта задача оказывается весьма интересной и непростой, по сложности не уступающей теории узлов (об узлах см. статью С. В. Дужина и С. В. Чмутова в настоящем сборнике, с. 59–93). Узлы и плоские кривые вообще связаны между собой множеством разнообразных связей — на некоторых из них мы коротко остановимся в конце статьи.

Основным «героем» нашей статьи являются так называемые *длинные кривые*. Их определение приводится в начале раздела 1; в этом же разделе вводится описание кривых с помощью специальных схем, называемых *гауссовыми диаграммами*. В разделах 2 и 3 мы определяем с помощью гауссовых диаграмм разнообразные инварианты кривых; в большей общности полученные инварианты исследуются в разделе 4. Заключительный раздел 5 посвящен обобщениям введенных инвариантов на другие объекты, тесно связанные с длинными кривыми — замкнутые кривые, плоские лежандровы фронты и узлы.

Мы предполагаем, что читатель данной статьи знаком с основами дифференциального исчисления и линейной алгебры. Текст, набранный мелким шрифтом, — в основном это доказательства технических утверждений — порой требует знакомства с более сложным материалом. Некоторые простые, но скучные доказательства мы опускаем вовсе, отсылая читателя за подробностями к специальной литературе.

1. ГАУССОВЫ ДИАГРАММЫ ДЛИННЫХ КРИВЫХ

Начнем с определений. Назовем *длинной кривой* гладкое отображение прямой в плоскость $\gamma(t) = (x(t), y(t))$, $t \in \mathbb{R}$, обладающее следующими свойствами:

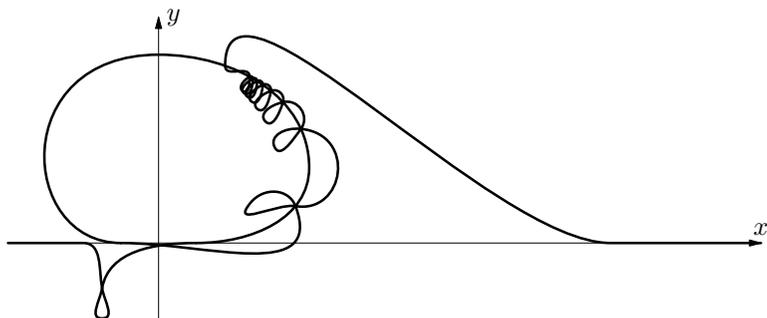


Рис. 1. Пример длинной кривой

1. Производные функций $x(t)$ и $y(t)$ ни при каком t не обращаются одновременно в нуль: $\forall t (x'(t))^2 + (y'(t))^2 > 0$.
2. При $t \leq -1$ и $t \geq 1$ выполнены равенства $x(t) = t$ и $y(t) = 0$.
3. Если $t_1 \in \mathbb{R}$ и $t_2 \in \mathbb{R}$ обладают тем свойством, что $x(t_1) = x(t_2)$ и $y(t_1) = y(t_2)$, то $-1 \leq t_1, t_2 \leq 1$.

Свойство 1 означает, что образ длинной кривой (множество точек $\{(x(t), y(t)) \mid t \in \mathbb{R}\}$, т. е. то, что мы обычно привыкли называть «кривой») является гладким. Действительно, в каждой его точке $(x(t), y(t))$ определен ненулевой касательный вектор $(x'(t), y'(t))$, определяющий однозначно направление касательной прямой. Свойство 2 означает, что вне отрезка $t \in [-1, 1]$ длинная кривая ведет себя стандартным образом — движется вдоль оси абсцисс с единичной скоростью (поэтому кривую и называют длинной).

Пары t_1, t_2 , о которых идет речь в свойстве 3, называются, по очевидным причинам, *точками самопересечения* длинной кривой γ . Свойство 3 означает, что все точки самопересечения лежат на отрезке $[-1, 1]$, — длинная кривая не пересекает себя на том участке, где она идет вдоль оси абсцисс. Разумеется, ничто не запрещает длинной кривой пересекаться с осью абсцисс и при $t \in [-1, 1]$, и вообще, — при $t \in [-1, 1]$ длинная кривая может вести себя весьма непростым образом, см., например, рис. 1.

Читатель, ознакомившийся со статьей С. В. Дужина и С. В. ЧмUTOва, знает, какую важную роль играют в геометрии соображения «общего положения», показывающие, что интересующий нас объект (узел, зацепление, лежандров фронт, длинную кривую и т. п.) можно сколь угодно малым шевелением привести к виду, не содержащему сложных особенностей. Для длинных кривых необходимое утверждение выглядит следующим образом:

ПРЕДЛОЖЕНИЕ 1. *Всякую длинную кривую γ можно сколь угодно мало пошевелить так, чтобы получить длинную кривую γ_0 , обладающую следующими свойствами:*

1. Кривая γ_0 имеет конечное (возможно, равное нулю) число точек самопересечения.

2. Все точки самопересечения кривой γ_0 простые — равенства

$$x(t_1) = x(t_2) = x(t_3), y(t_1) = y(t_2) = y(t_3)$$

не выполнены ни для какой тройки различных чисел $t_1, t_2, t_3 \in \mathbb{R}$.

3. Если t_1, t_2 — точка самопересечения кривой γ , то векторы $(x'(t_1), y'(t_1))$ и $(x'(t_2), y'(t_2))$ не коллинеарны.

Иными словами, в каждой точке самопересечения пересекаются лишь две ветви кривой и касательные к этим ветвям не совпадают (говорят, что самопересечение *трансверсально*). Кривые, обладающие свойствами кривой γ_0 из предложения 1, будем называть *кривыми общего положения*.

Доказывается предложение 1 с помощью одного чрезвычайно важного утверждения — теоремы Тома о трансверсальности. Читатель может познакомиться с теоремой Тома по книге [1].

Пусть теперь γ — кривая общего положения. Изобразим вещественную ось \mathbb{R} в виде прямой и отметим на этой прямой все точки t_1, t_2 самопересечения кривой γ — их будет, по свойству 1 кривых общего положения, конечное число пар. Точки, принадлежащие одной и той же паре, соединим дугой. По свойству 2 разные дуги не имеют общих концов.

Свойство 3 позволяет приписать каждой дуге число ± 1 . А именно, рассмотрим числа t_1, t_2 , соединенные дугой — т. е. относящиеся к одной точке самопересечения. Упорядочим их по возрастанию — пусть, например, $t_1 < t_2$. Векторы $(x'(t_1), y'(t_1))$ и $(x'(t_2), y'(t_2))$ не коллинеарны и поэтому образуют на плоскости базис. Припишем дуге число $+1$, если этот базис «правый», т. е. вектор $(x'(t_2), y'(t_2))$ получается из вектора $(x'(t_1), y'(t_1))$ поворотом (на угол, меньший развернутого) *против часовой стрелки* (и соответствующим растяжением, конечно). В противном случае — если базис «левый» — припишем дуге число -1 .

Полученная конструкция — прямая с отмеченными точками, которые попарно соединены дугами с приписанными им числами ± 1 — называется *гауссовой диаграммой*¹⁾ длинной кривой γ .

Пусть Γ_1 и Γ_2 — две гауссовы диаграммы и $a_1 < a_2 < \dots < a_N, b_1 < b_2 < \dots < b_N$ — соответствующие наборы концов дуг. Мы будем говорить, что Γ_1 и Γ_2 *эквивалентны*, если точки a_i и a_j соединены дугой тогда и только тогда (при тех же i и j), когда дугой соединены точки b_i и b_j , причем знаки обеих дуг совпадают. Иными словами, эквивалентные гауссовы диаграммы не отличаются ни знаками, ни взаимным расположением дуг (хотя и могут отличаться конкретными значениями параметра t , при которых происходят самопересечения).

Оказывается, гауссова диаграмма, взятая с точностью до эквивалентности, аккумулирует в себе практически всю существенную информацию о длинной кривой:

ПРЕДЛОЖЕНИЕ 2. Пусть γ_1 и γ_2 — две длинные кривые общего положения, гауссовы диаграммы которых эквивалентны. Тогда существуют гладкое взаимно однозначное преобразование плоскости Φ и гладкое взаимно однозначное

¹⁾ По имени К.-Ф. Гаусса — знаменитого математика XVIII – начала XIX в., заслужившего при жизни наименование «короля математиков». Насколько известно автору этой статьи, К.-Ф. Гаусс исследованием длинных кривых не занимался.

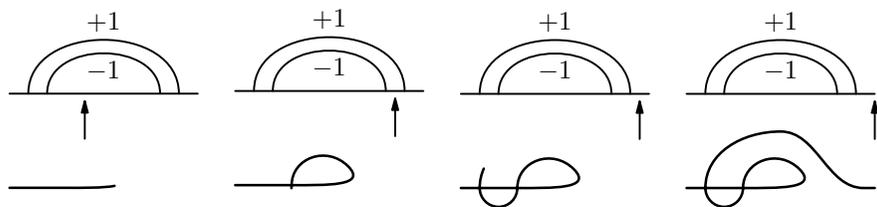


Рис. 2. Построение кривой по ее гауссовой диаграмме

преобразование прямой φ , для которых выполнено равенство

$$\forall t \in \mathbb{R} \quad \gamma_2(t) = \Phi(\gamma_1(\varphi(t))). \quad (1)$$

По-другому равенство (1) можно выразить так: на плоскости и на прямой существуют, помимо стандартных, другие системы координат — на плоскости в общем случае система координат криволинейная — такие что кривая γ_2 записывается в этих системах координат теми же выражениями $(x(t), y(t))$, что кривая γ_1 — в стандартной. Иными словами, предложение 2 утверждает, что гауссова диаграмма содержит всю информацию о длинной кривой на плоскости, не зависящую от выбора систем координат.

Аккуратное доказательство предложения 2 довольно трудоемко. Идею его, однако, понять несложно. Действительно, пусть нам задана гауссова диаграмма кривой γ . Будем строить кривую γ следующим образом. При $t \leq -1$ кривая проводится однозначно — движется вдоль оси абсцисс с единичной скоростью. Продолжим движение до тех пор, пока на гауссовой диаграмме не встретится конец A какой-нибудь дуги. Если конец A — левый, то сделаем на кривой отметку и продолжим движение, не пересекая уже проведенной части кривой. Если конец A — правый, то найдем отметку, соответствующую левому концу B той же самой дуги. За небольшое время до точки A загнем кривую так, чтобы приблизиться к этой отметке, по-прежнему не пересекая уже проведенной части кривой; при этом знак дуги AB показывает, с какой стороны нужно приближаться — справа или слева. В момент времени A пересечем кривую в точке, отмеченной B , и продолжим движение, следуя тем же правилам.

Конечно, в процессе такого построения у нас остается значительная свобода действий — «приблизиться к отметке» и «продолжать движение» можно по-разному. Утверждение предложения 2 заключается в том, что эту свободу действий можно скомпенсировать выбором подходящего преобразования плоскости Φ . Преобразование прямой φ отвечает за то, что для эквивалентных гауссовых диаграмм пересечение может происходить в разные моменты времени. Пример построения кривой по (очень простой) гауссовой диаграмме приведен на рис. 2.

Предложение 2, однако, не утверждает, что по *всякой* гауссовой диаграмме (т. е. прямой с точками, дугами и знаками) можно построить кривую. Причины этого более или менее ясны из приведенной идеи доказательства: может оказаться невозможным достичь отметки B , не пересекая при этом кривую. На рис. 3 приведен пример гауссовой диаграммы, не соответствующей никакой кривой.



Рис. 3. Нереализуемая гауссова диаграмма

Гауссовы диаграммы, получающиеся как гауссовы диаграммы кривых, называются *реализуемыми*.

Представляет интерес вопрос о том, как по гауссовой диаграмме определить, реализуема она или нет. Ответ дается следующей теоремой:

ТЕОРЕМА 1. *Гауссова диаграмма Γ реализуема тогда и только тогда, когда для каждой ее дуги AB выполнено равенство*

$$\sum_{\lambda} \varepsilon(\lambda) = 0$$


Рисунок под знаком суммирования означает, что сумма берется по всем дугам λ , правый конец которых находится между точками A и B (а левый — где угодно); $\varepsilon(\lambda)$ — знак дуги λ .

Идея доказательства теоремы 1 такова. Каждой точке пересечения отрезка кривой γ между точкой B и точкой, расположенной непосредственно перед A , соответствует правый конец дуги, лежащий между B и A . Сумма знаков таких дуг равна индексу пересечения отрезка кривой и уже проведенного участка кривой. Очевидно, что для того, чтобы точки A и B можно было соединить, не пересекая уже проведенную кривую дополнительно, необходимо и достаточно, чтобы этот индекс равнялся нулю.

Таким образом, далеко не все гауссовы диаграммы реализуемы. Общее число классов эквивалентности гауссовых диаграмм с n дугами равно $(2n)!/n!$; вычисление общего числа классов эквивалентности реализуемых гауссовых диаграмм с n дугами (т. е. числа типов длинных кривых с n самопересечениями) — нерешенная задача. Приведем для сравнения результаты компьютерных вычислений, принадлежащие С. М. Гусейн-Заде и Ф. С. Дужину:

Число дуг	1	2	3	4	5	6	7
Все диаграммы	2	12	120	1680	30240	665280	17297280
Реализуемые диаграммы	2	8	42	260	1796	13396	105706

Число дуг	8	9	10
Все диаграммы	518918400	17643225600	670442572800
Реализуемые диаграммы	870772	7420836	65004584

2. Гомотопии и индекс Уитни

Гомотопией длинных кривых называется семейство длинных кривых γ_τ , непрерывно зависящих от параметра $\tau \in [0, 1]$. Иными словами, речь идет о непрерывной деформации кривой γ_0 в кривую γ_1 . Подчеркнем, что гомотопия рисунка 4 «незаконна», поскольку в процессе ее (на средней кривой) возникает острое — точка, в которой $x'(t) = y'(t) = 0$.



Рис. 4. «Незаконная» гомотопия

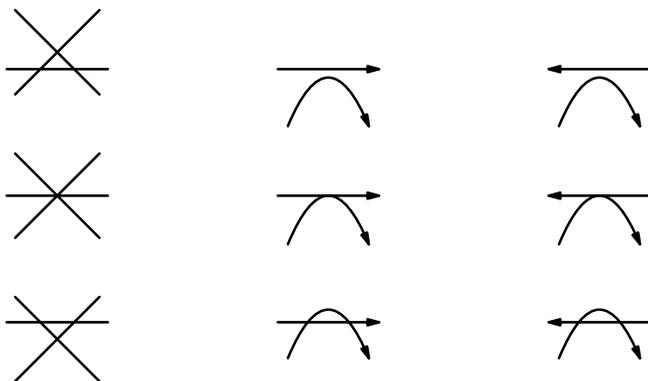


Рис. 5. Локальные катастрофы при гомотопиях

Кривые γ_τ в общем случае не предполагаются кривыми общего положения. Более того, как мы вскоре увидим, даже если γ_0 и γ_1 — кривые общего положения, то, как правило, невозможно «пошевелить» гомотопию γ_τ так, чтобы все кривые стали кривыми общего положения: нарушение общности положения, которое, согласно предложению 1, можно шевелением удалить для конкретной кривой γ_{τ_0} , неизбежно при этом появится у кривой γ_τ с другим, близким к τ_0 , значением параметра τ . Тем не менее, «теорема об общем положении», аналогичная предложению 1, существует и для гомотопий:

ПРЕДЛОЖЕНИЕ 3. *Всякую гомотопию γ_τ длинных кривых можно сколь угодно мало пошевелить так, чтобы выполнялись следующие условия:*

1. *Для всех значений параметра τ , кроме конечного набора τ_1, \dots, τ_n «катастрофических значений», кривые γ_τ — общего положения.*
2. *При проходе каждого из катастрофических значений нарушение общности положения возникает только в одной точке кривой и принадлежит к одному из трех типов, изображенных на рис. 5.*

Катастрофы, изображенные в столбцах на рис. 5, называются соответственно прохождением тройной точки, попутным касанием и противокасанием. Имеется в виду, что эти катастрофы могут происходить в любом из двух возможных направлений — так, при касании число точек самопересечения может как увеличиваться на 2, так и уменьшаться на 2.

Предложение 3, как и предложение 1, является следствием теоремы Тома о трансверсальности.

Рассмотрим теперь влияние гомотопий на гауссовы диаграммы кривых. Очевидно, что на участках параметра $\tau \in (\tau_k, \tau_{k+1})$, где кривая γ_τ остается кривой

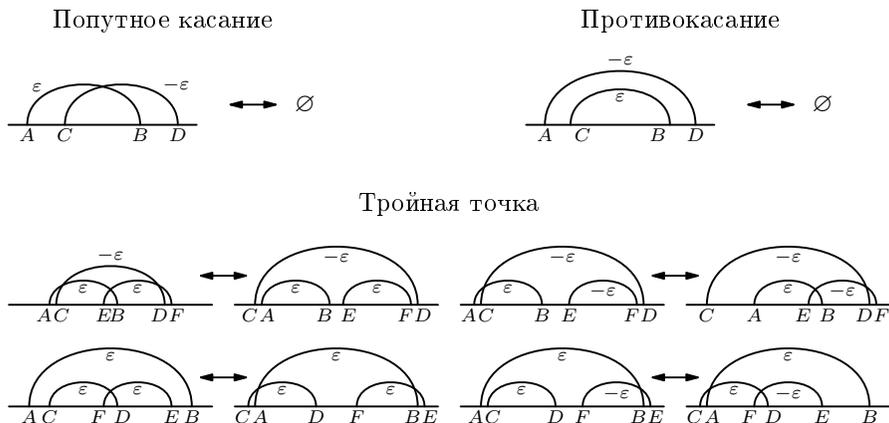


Рис. 6. Изменение гауссовых диаграмм при катастрофах

общего положения, ее гауссова диаграмма заменяется на эквивалентную. Мы не будем в дальнейшем различать эквивалентные гауссовы диаграммы; так что будем считать, что на интервалах между катастрофами гауссова диаграмма кривой не меняется.

Изменение гауссовой диаграммы при катастрофе зависит от типа катастрофы. Начнем со случая касания — слияния и исчезновения двух точек самопересечения. Две близкие точки самопересечения изображаются на диаграмме двумя близкими дугами, AB и CD , причем между A и C (левыми концами), а также между B и D других концов дуг нет. После катастрофы обе дуги исчезают. Нетрудно видеть — это обстоятельство будет неоднократно использовано в дальнейшем — что знаки дуг AB и CD обязательно противоположны. Существует два варианта расположения концов сливающихся дуг на прямой (рис. 6). При одном из них дуги пересекаются — это соответствует случаю попутного касания. При другом одна из дуг «объемлет» другую — это случай противокасания. Разумеется, процесс может происходить и в обратном направлении — вместо исчезновения получится рождение двух близких дуг.

В катастрофе прохождения тройной точки участвуют три точки самопересечения, и, как следствие, три дуги гауссовой диаграммы, AB , CD и EF . Эти дуги близки в том смысле, что их концы разбиваются на «близкие» пары (на рис. 6 — (A, C) , (B, E) и (D, F)), между которыми нет концов «посторонних» дуг. При прохождении тройной точки концы дуг в каждой «близкой» паре меняются местами; знаки дуг сохраняются. Конкретное расположение и знаки дуг при прохождении тройной точки могут быть различными — они зависят от ориентации сторон «исчезающего треугольника» на рис. 5 и порядка, в котором кривая проходит эти стороны. Перебор возможных случаев показывает, что всего существует 8 вариантов — они показаны на рис. 6; здесь везде $\varepsilon = \pm 1$.

Примером применения предложения 3 (и образцом того, как вообще применяются подобные «теоремы об общем положении») является следующее важное утверждение.

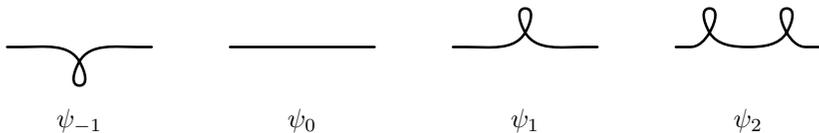


Рис. 7. Примеры кривых с различным индексом Уитни

ТЕОРЕМА УИТНИ²⁾. Пусть γ_0 и γ_1 — две гомотопные длинные кривые, а Γ_0 и Γ_1 — соответствующие гауссовы диаграммы. Тогда сумма знаков всех дуг на диаграммах Γ_0 и Γ_1 совпадает.

ДОКАЗАТЕЛЬСТВО. Пусть γ_τ — гомотопия, соединяющая кривые γ_0 и γ_1 . Можно считать, что эта гомотопия удовлетворяет условиям предложения 3. Тем самым в процессе гомотопии гауссова диаграмма кривой претерпевает конечное количество перестроек, типы которых описаны выше. При перестройке, соответствующей прохождению тройной точки, знаки дуг на гауссовой диаграмме не меняются вовсе — меняется только их взаимное расположение. При перестройках, соответствующих самокасанию, появляются или пропадают две дуги разных знаков — следовательно, сумма знаков все равно остается прежней. \square

Построенный инвариант — сумма знаков дуг гауссовой диаграммы — называется *индексом Уитни* кривой. Мы показали, что он одинаков у гомотопных кривых; отсюда немедленно следует, что кривые, изображенные на рисунке 7, попарно негомотопны.

Оказывается, индекс Уитни можно вычислить и не обращаясь к гауссовым диаграммам (в частности, из этого следует, что он определен не только для кривых общего положения). А именно, рассмотрим касательный вектор $v(t) = (x'(t), y'(t))$ к кривой γ . Согласно свойству 1 этот вектор никогда не обращается в нуль. По свойству 2 длинной кривой вектор $v(t)$ при $t = 1$ и $t = -1$ одинаков: $v(-1) = v(1) = (1, 0)$. Тем самым, пока t пробегает отрезок $[-1, 1]$, вектор $v(t)$ описывает на плоскости некоторую замкнутую кривую, не проходящую через начало координат. Полное число оборотов, которое делает эта кривая вокруг начало координат (обороты против часовой стрелки считаются с плюсом, против — с минусом), называется *числом вращения* кривой γ .

Иными словами, число вращения можно определить как элемент фундаментальной группы $\pi_1(\mathbb{R}^2 \setminus \{0\}) = \mathbb{Z}$, соответствующий кривой $v(t)$. Читатель, незнакомый с понятием фундаментальной группы, может познакомиться с ним по любому учебнику топологии (например, по книге [3] или по замечательной книге [4], изданной в серии «Библиотечка Квант»).

ТЕОРЕМА УИТНИ (ПРОДОЛЖЕНИЕ). Число вращения кривой равно ее индексу Уитни.

Мы докажем эту часть теоремы несколько позднее.

Важность индекса Уитни (числа вращения) как инварианта длинной кривой подчеркивается тем обстоятельством, что он является, по сути, единственной

²⁾ На самом деле Уитни в своей знаменитой работе [2] рассматривал не длинные кривые, а замкнутые, однако его теорема, вместе с доказательством, остается верной и для нашего случая. О замкнутых кривых мы немного скажем в разделе 5.1.

характеристикой длинной кривой, сохраняющейся при гомотопиях. А именно, имеет место утверждение, обратное теореме Уитни:

ТЕОРЕМА УИТНИ (ОКОНЧАНИЕ). Пусть длинные кривые γ_0 и γ_1 таковы, что их числа вращения совпадают. Тогда кривые γ_0 и γ_1 гомотопны.

ДОКАЗАТЕЛЬСТВО. Нетрудно видеть, что вектор-функция $v(t) = (a(t), b(t))$ тогда и только тогда может быть представлена в виде $(x'(t), y'(t))$ для некоторой длинной кривой $\gamma(t) = (x(t), y(t))$, когда выполнены следующие условия:

$$\forall t \in [-1, 1] \quad v(t) \neq (0, 0), \quad (2)$$

$$v(-1) = v(1) = (1, 0), \quad (3)$$

$$\int_{-1}^1 a(t) dt = 2, \quad \int_{-1}^1 b(t) dt = 0. \quad (4)$$

Пусть теперь $v_0(t)$ и $v_1(t)$ — вектор-функции, соответствующие длинным кривым γ_0 и γ_1 и имеющие одинаковое число вращения. Требуется доказать, что можно построить гомотопию $v_\tau(t)$, $\tau \in [0, 1]$, в которой каждая кривая $v_\tau(t)$ удовлетворяет условиям (2) и (4).

Рассмотрим для каждого вектора (p, q) вектор-функцию $V_{p,q}(t)$, определенную на отрезке $[1 - |p| - |q|, 1]$ и обладающую следующими свойствами:

$$\forall t \quad V_{p,q}(t) \neq (0, 0),$$

$$V_{p,q}(1 - |p| - |q|) = V_{p,q}(1) = (1, 0),$$

$$\int_{1 - |p| - |q|}^1 V_{p,q}(t) dt = (p, q),$$

$$\text{число вращения } V_{p,q} = 0.$$

Поскольку числа вращения вектор-функций $v_0(t)$ и $v_1(t)$ совпадают, можно построить гомотопию $w_\tau(t)$ между ними, обладающую свойствами (2) и (3), но не обладающую, вообще говоря, свойством (4). Пусть $\int_{-1}^1 w_\tau(t) dt = (p_\tau, q_\tau)$. Без ограничения общности можно считать, что $w_\tau(t) \equiv (1, 0)$ при $t \in [1 - |p_\tau| - |q_\tau|, 1]$. Тогда определим новую гомотопию $v_\tau(t)$ формулой

$$v_\tau(t) = \begin{cases} w_\tau(t), & t \in [-1, 1 - |p_\tau| - |q_\tau|], \\ V_{p_\tau, q_\tau}(t), & t \in [1 - |p_\tau| - |q_\tau|, 1]. \end{cases}$$

Очевидно, построенная гомотопия удовлетворяет всем условиям (2)–(4). \square

Теперь мы можем доказать вторую часть теоремы Уитни — она следует из третьей. Действительно, рассмотрим набор кривых $\dots, \psi_{-1}, \psi_0, \psi_1, \dots$, изображенных на рис. 7. Индекс Уитни кривой ψ_k равен k ; нетрудно видеть, что таково же и ее число вращения. Пусть теперь k — число вращения какой-либо длинной кривой γ . Из третьей части теоремы Уитни следует, что γ гомотопна кривой ψ_k . Но тогда из первой части теоремы следует, что индекс Уитни кривой γ такой же, как и у ψ_k , т.е. тоже равен k . Теорема Уитни доказана. \square

3. СТРАННОСТЬ, J^+ И J^-

Таким образом, инварианты длинных кривых с точностью до гомотопии исчерпываются индексом Уитни. Новые инварианты появляются, если мы рассмо-

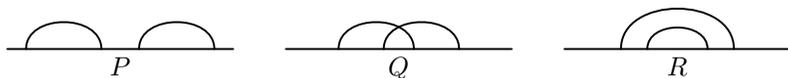


Рис. 8. Взаимное расположение пары дуг

трим более узкие классы кривых и, соответственно, уменьшим число разрешенных гомотопий. В этом случае кривые, которые ранее считались гомотопными, могут стать негомотопными, и теория станет более содержательной.

Чаще всего рассматривают три варианта таких теорий:

1. Теорию кривых без тройных точек.
2. Теорию кривых без попутных касаний.
3. Теорию кривых без противокасаний.

Покажем, что все эти три теории не сводятся к теории раздела 2, иными словами, что в них имеются инварианты, отличные от индекса Уитни.

Все эти инварианты будут выражаться через гауссовы диаграммы формулами, похожими на формулу для индекса Уитни, но только суммирование будет происходить по всем *парам* дуг на диаграмме. Существуют три способа взаимного расположения пары дуг — они указаны на рис. 8.

Пусть Φ — один из типов расположения пары дуг, а γ — кривая общего положения. Рассмотрим величину

$$\Phi(\gamma) = \sum_{(\lambda_1, \lambda_2) = \Phi} \varepsilon(\lambda_1)\varepsilon(\lambda_2). \quad (5)$$

Здесь сумма берется по всем парам дуг (λ_1, λ_2) на гауссовой диаграмме кривой γ , относящимся к типу Φ ; $\varepsilon(\lambda)$ означает знак дуги λ . Таким образом, мы имеем для каждой кривой γ величины $P(\gamma)$, $Q(\gamma)$ и $R(\gamma)$. Введем еще величину $S(\gamma)$, равную количеству дуг на диаграмме, независимо от их знаков. Величину $S(\gamma)$ можно рассматривать как частный случай величин (5) — в этом случае нужно считать, что сумма в (5) берется по всем парам *совпадающих* дуг λ_1 и λ_2 .

Величины $\Phi(\gamma)$ представляют собой функции на множестве классов эквивалентности кривых общего положения. Все функции на этом множестве, очевидно, образуют линейное пространство. Рассмотрим в нем подпространство (которое по причинам, излагаемым позже, назовем \mathcal{G}_2), натянутое на функции P , Q , R , S ; произвольная функция из этого подпространства имеет вид $f = aP + bQ + cR + dS$. Рассмотрим, как меняется число $f(\gamma)$, когда кривая γ претерпевает одну из катастроф, описанных в разделе 2.

При катастрофе попутного касания (рождение двух близких точек самопересечения) появляются две новые дуги, λ_1 и λ_2 , расположенных по типу Q . Разобьем все пары дуг на диаграмме кривой γ (после катастрофы) на 4 типа:

1. Пары, в которые не входит ни λ_1 , ни λ_2 .
2. Пары, в которые входит λ_1 и еще одна дуга, отличная от λ_1 и λ_2 .
3. Пары, в которые входит λ_2 и еще одна дуга, отличная от λ_2 и λ_1 .
4. Пары (λ_1, λ_2) , (λ_1, λ_1) и (λ_2, λ_2) .

Пары типа 1 входят в сумму, определяющую значение функции f , как до, так и после катастрофы. Поэтому при вычислении разности $f(\gamma_{\text{после}}) - f(\gamma_{\text{до}})$ такие пары можно не учитывать.

Между парами типа 2 и парами типа 3 можно установить взаимно однозначное соответствие, заменяя в каждой паре типа 2 дугу λ_1 на дугу λ_2 . При этом соответствии тип пары не меняется, поскольку дуги λ_1 и λ_2 — «близкие», между их соответствующими концами нет концов других дуг. Произведение же знаков меняется на противоположное, поскольку знаки дуг λ_1 и λ_2 различны. Поэтому вклады, вносимые парами типов 2 и 3 в разность $f(\gamma_{\text{после}}) - f(\gamma_{\text{до}})$, сокращаются.

Тем самым единственный вклад в разность вносят пары типа 4, не имеющие соответствий в диаграмме кривой до катастрофы. Поскольку $\varepsilon(\lambda_1)\varepsilon(\lambda_2) = -1$, вклад равен $2d - b$.

Для катастрофы противокасания рассуждения совершенно аналогичны, за исключением того, что дуги λ_1 и λ_2 расположены по типу R . Изменение значения функции f при такой катастрофе равно, тем самым, $2d - c$.

При катастрофе прохождения тройной точки существует несколько вариантов расположения и знаков дуг, см. рис. 6. Так же, как и для катастроф самокасания, устанавливается, что вклад в разность вносится только парами дуг, участвующих в катастрофе. Нетрудно проверить, что если катастрофа происходит в направлении «слева направо» на рис. 6 (мы будем говорить, что происходит положительное прохождение тройной точки), то изменение значения функции равно $a + b - 2c$, независимо от конкретного расположения дуг.

Приравнивая нулю соответствующие разности, мы можем отобразить в пространстве \mathcal{G}_2 функции, остающиеся неизменными при определенных типах катастроф, и, тем самым, построить примеры инвариантов в трех теориях, указанных выше.

Результаты этих вычислений мы сведем в теорему:

ТЕОРЕМА 2. 1. *Пространство функций $f \in \mathcal{G}_2$, остающихся инвариантными при перестройке самокасания, двумерно и порождено функциями $W^2 = P + Q + R + S$ и $St = Q + R + S/2$. Значение первой из этих функций не изменяется также и при прохождении тройной точки, а значение второй увеличивается на единицу при положительном прохождении.*

2. *Пространство функций $f \in \mathcal{G}_2$, остающихся инвариантными при перестройке прохождения тройной точки и при перестройке попутного касания, двумерно и порождено функциями W^2 и $J^- = 4Q + 2R + 2S$. Функция J^- увеличивается на 2 при противокасании с появлением двух новых точек самопересечения.*

3. *Пространство функций $f \in \mathcal{G}_2$, остающихся инвариантными при перестройке прохождения тройной точки и при перестройке противокасания, двумерно и порождено функциями W^2 и $J^+ = -4Q - 2R - S$. Функция J^+ увеличивается на 2 при попутном касании с появлением двух новых точек самопересечения.*

Теорема 2 показывает, что теории кривых без тройных точек, без попутных касаний и без противокасаний богаче, чем просто теория длинных кривых —

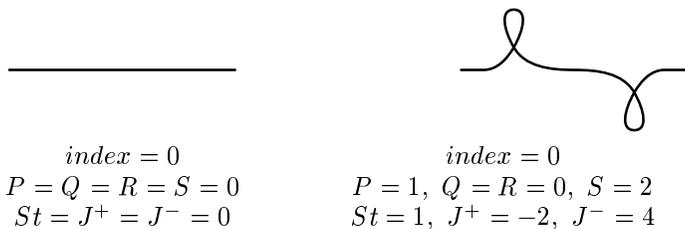


Рис. 9. Инварианты St , J^+ и J^-

здесь есть новые инварианты. На рис. 9 изображены две кривые, имеющие одинаковый индекс Уитни, но разные инварианты St , J^+ и J^- . Отсюда следует, что они негомотопны ни в одной из трех теорий. Иными словами, их можно соединить гомотопией, но в процессе такой гомотопии обязательно встретятся тройные точки, попутные касания и противокасания.

Функция W^2 представляет собой, как нетрудно видеть, квадрат индекса Уитни. Это объясняет ее инвариантность при всех катастрофах, а также то, что она является единственной, с точностью до пропорциональности, функцией из \mathcal{G}_2 , обладающей таким свойством. Что касается выбора второй базисной функции, St , J^+ и J^- , в соответствующих подпространствах, то он обусловлен следующим обстоятельством.

На множестве длинных кривых можно определить операцию *прямой суммы*. Прямая сумма кривых $\gamma_1(t) = (a_1(t), b_1(t))$ и $\gamma_2(t) = (a_2(t), b_2(t))$ это кривая $\gamma(t)$, задаваемая равенством

$$\gamma(t) = \begin{cases} (t, 0), & t \leq -1 \text{ и } t \geq 1, \\ \left(\frac{1}{2}a_1(2t+1) - \frac{1}{2}, b_1(2t+1) \right), & -1 \leq t \leq 0, \\ \left(\frac{1}{2}a_2(2t-1) + \frac{1}{2}, b_2(2t-1) \right), & 0 \leq t \leq 1. \end{cases}$$

Иными словами, кривая $\gamma \stackrel{\text{def}}{=} \gamma_1 + \gamma_2$ проходит сначала γ_1 , а затем γ_2 . Коэффициенты в формуле вызваны необходимостью соблюдать условия нормировки. Гауссова диаграмма прямой суммы получается, с точностью до эквивалентности, «приставлением» гауссовой диаграммы кривой γ_2 справа от гауссовой диаграммы кривой γ_1 .

Вычислим теперь значение функции $f \in \mathcal{G}_2$ на кривой $\gamma_1 + \gamma_2$. Пары, обе дуги которых лежат в гауссовой диаграмме кривой γ_1 , дадут в сумме $f(\gamma_1)$; пары, обе дуги которых лежат в гауссовой диаграмме кривой γ_2 , дадут $f(\gamma_2)$. Если одна дуга пары лежит в диаграмме кривой γ_1 , а другая — в диаграмме кривой γ_2 , то они образуют пару типа P . Поэтому получаем

$$f(\gamma_1 + \gamma_2) = f(\gamma_1) + f(\gamma_2) + aW(\gamma_1)W(\gamma_2),$$

где W — индекс Уитни. Функции $f \in \mathcal{G}_2$, для которых $a = 0$ (и только они), обладают свойством аддитивности: $f(\gamma_1 + \gamma_2) = f(\gamma_1) + f(\gamma_2)$. Функции St , J^+ и J^- аддитивны. Более того, свойство аддитивности вместе со свойством

инвариантности при катастрофах соответствующих типов определяет функции $St, J^+, J^- \in \mathcal{G}_2$ однозначно с точностью до пропорциональности.

Первоначально функции St (называемая *странностью* кривой) и J^\pm были определены В. И. Арнольдом в статье [5] аксиоматически — как функции, обладающие свойствами, указанными в теореме 2, и свойством аддитивности. Гауссовы диаграммы в определении не участвовали; проблема при этом состояла, во-первых, в том, чтобы описать без их участия, что такое «положительное» и «отрицательное» прохождение тройной точки, и, во-вторых, в доказательстве существования и единственности таких функций. Первоначальное доказательство существования не давало явной формулы, позволявшей вычислить инварианты St и J^\pm по виду конкретной кривой. Первые явные формулы были получены в работах О. Виро [6] и А. Шумаковича [7] (в частном случае также Ф. Аикарди [8]). Подход, излагаемый в данной статье, принадлежит М. Поляку; значительная часть содержания статьи заимствована из его работы [9].

4. ГАУССОВЫ ИНВАРИАНТЫ ВЫСШИХ ПОРЯДКОВ

По аналогии с инвариантами P, Q, R, S предыдущего раздела мы можем определить инварианты высших порядков. Для этого следует рассмотреть вместо пар дуг наборы из n дуг. Эти n дуг могут быть расположены на прямой различными способами; в число способов расположения мы будем, как и раньше, включать и такие способы, где некоторые дуги совпадают. Эти способы естественным образом изображаются в виде гауссовых диаграмм, где, в отличие от гауссовых диаграмм кривых, допускаются кратные дуги. Обозначим множество гауссовых диаграмм с дугами суммарной кратности n символом \mathcal{A}_n . Каждой диаграмме $\Phi \in \mathcal{A}_n$ соответствует функция

$$\Phi(\gamma) = \sum_{(\lambda_1, \dots, \lambda_n) = \Phi} \varepsilon(\lambda_1) \dots \varepsilon(\lambda_n). \quad (6)$$

Здесь γ — длинная кривая общего положения; суммирование производится по всем наборам из n дуг, расположение которых соответствует диаграмме Φ ; $\varepsilon(\lambda)$ — знак дуги λ . Как и в предыдущем разделе, мы рассмотрим линейное пространство \mathcal{G}_n , порожденное всеми функциями типа (6) для $\Phi \in \mathcal{A}_n$; элементы пространства \mathcal{G}_n будем называть *гауссовыми функциями* порядка n . Можно также рассмотреть линейное пространство $G_n = \bigoplus_{k=0}^n \mathcal{G}_k$ гауссовых функций порядка не выше n . Если гауссова функция является инвариантом одной из теорий, описанных в разделе 3, то мы будем называть ее гауссовым инвариантом.

Теория гауссовых инвариантов n -ого порядка весьма напоминает теорию 2-го порядка, рассмотренную в разделе 3. Теорию удобно строить последовательно, «индукцией по n », ввиду следующего факта:

ПРЕДЛОЖЕНИЕ 4. Пусть γ_0 и γ_1 — две гомотопные кривые общего положения, такие что при гомотопии γ_τ между ними происходит ровно одна катастрофа типов, описанных в предложении 3. Пусть $\Phi \in \mathcal{G}_n$. Тогда существует гауссова функция $\Psi \in G_n$ такая, что $\Phi(\gamma_1) - \Phi(\gamma_0) = \Psi(\gamma_0)$. Функция Ψ не зависит от кривых γ_0 и γ_1 , а только от функции Φ и типа катастрофы. Для катастроф самокасания можно выбрать $\Psi \in \mathcal{G}_{n-2}$.

Доказательство предложения состоит в явном построении функции Ψ . Очевидно, что достаточно разобрать случай, когда Φ — функция типа (6).

ДОКАЗАТЕЛЬСТВО. Предположим сначала, что происходит катастрофа путного самокасания с возникновением пары точек самопересечения, которым соответствуют дуги λ_1 и λ_2 . Как и в разделе 3 заметим, что вклад в разность $\Phi(\gamma_1) - \Phi(\gamma_0)$ вносят только наборы из n дуг, содержащие как λ_1 , так и λ_2 . Поскольку знаки дуг λ_1 и λ_2 противоположны, вклад от каждого такого набора равен взятому со знаком минус произведению знаков всех дуг набора, кроме λ_1 и λ_2 . Поэтому в качестве функции Ψ в этом случае можно взять сумму (со знаком минус) всех поддиаграмм диаграммы Φ , получающихся из нее выкидыванием пары дуг, расположенных по типу Q и таких, что между их соответствующими концами нет концов других дуг (ибо только такие пары дуг могут соответствовать λ_1 и λ_2). В случае катастрофы противокасания ситуация аналогична, только выкидывать нужно пары, расположенные по типу R .

Для катастрофы прохождения тройной точки рассуждения аналогичны. Трудность здесь заключается в том, что перестройки гауссовой диаграммы при такой катастрофе могут быть различными — см. рис. 6. Небольшой перебор показывает, тем не менее, что функция Ψ во всех случаях получается одна и та же. \square

Предложение 4 позволяет в принципе перечислить гауссовы инварианты порядка n в каждой из трех теорий, описанных в начале раздела 3. Действительно, для каждой гауссовой диаграммы $\Phi \in \mathcal{A}_n$ можно вычислить функцию Ψ для каждого типа катастрофы. Тогда условие того, что функция $f = \sum_{\Phi \in \mathcal{A}_n} \alpha_\Phi \Phi$ инвариантна при катастрофах данного типа (т.е. является инвариантом соответствующей теории) превращается в систему линейных уравнений на коэффициенты α_Φ . Условие аддитивности, описанное в разделе 3, приводит к дополнительным уравнениям.

Следует, однако, иметь ввиду некоторые трудности, которые «не видны» для функций второго порядка, но появляются, когда порядок становится большим.

Во-первых, две функции различных порядков могут совпадать. Это связано с тем, что знак дуги $\varepsilon(\lambda) = \pm 1 = \varepsilon^3(\lambda)$, и поэтому тройная дуга эквивалентна одинарной. Например, функция из \mathcal{G}_3 , соответствующая диаграмме с одной тройной дугой, является инвариантом во всех трех теориях — это просто индекс Уитни. Чтобы избежать этого, мы будем рассматривать впредь гауссовы диаграммы с кратностью дуг не более 2.

Во-вторых, на пространстве $\bigoplus_{n=0}^{\infty} \mathcal{G}_n$ всех гауссовых функций имеется естественное умножение. Нетрудно видеть, что функция $\Phi\Psi$, где $\Phi \in \mathcal{G}_k$, а $\Psi \in \mathcal{G}_l$, принадлежит \mathcal{G}_{k+l} . Точнее, $\Phi\Psi$ есть сумма всех диаграмм из $k+l$ дуг, получающихся объединением диаграмм Φ и Ψ ; при этом концы дуг, принадлежащих Φ и Ψ , чередуются между собой всеми возможными способами. Таким образом, в $\bigoplus_{n=0}^{\infty} \mathcal{G}_n$ имеется структура градуированного кольца. Если функции Φ и Ψ являются инвариантами какой-нибудь теории, то $\Phi\Psi$ тоже, разумеется, инвариант. Этот инвариант, однако, не является в разумном смысле «новым».

Третья, и наиболее существенная, трудность заключается в том, что сумма $f = \sum_{\Phi \in \mathcal{A}_n} \alpha_\Phi \Phi$ с ненулевыми коэффициентами может порождать функцию на кривых, тождественно равную нулю. Например, равны нулю функции Z_+ и Z_- третьего порядка, соответствующие диаграммам на рис. 10 (символ 2 означает

дугу кратности 2), — это следует из условия реализуемости гауссовой диаграммы (теорема 1). Соответственно равны нулю и все функции вида $Z_{\pm}\Phi$, для всех гауссовых функций Φ .

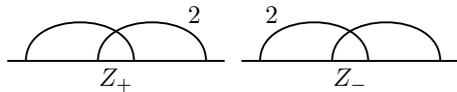


Рис. 10. Функции, равные нулю

размерности 7. В нем содержится одномерное подпространство, порожденное кубом индекса Уитни; дополнительное шестимерное пространство образуют функции, тождественно равные нулю.

Инварианты третьего порядка типа странности (не изменяющиеся при самокасаниях, но изменяющиеся при прохождении тройных точек) образуют пятимерное пространство. Двумерное подпространство в нем образуют аддитивные инварианты. Двумерное подпространство порождено «старыми» инвариантами — кубом индекса Уитни и произведением индекса Уитни на странность. Остальные 4 инварианта, включая оба аддитивных, — новые.

Инварианты третьего порядка типа J^+ (выдерживающие все перестройки, кроме попутных касаний) образуют трехмерное пространство. Базис в нем составляют куб индекса Уитни, произведение индекса Уитни и инварианта J^+ (в смысле раздела 3), и новый аддитивный инвариант.

Инварианты третьего порядка типа J^- (изменяются только при противокасаниях) образуют двумерное пространство. Базис в нем составляют куб индекса Уитни и произведение индекса Уитни и J^- . Таким образом, новых инвариантов типа J^- в третьем порядке не появляется; также нет и аддитивных инвариантов.

Полученный результат показывает, что различие между катастрофами попутного касания и противокасания введено не даром. На самом деле теории кривых без попутных касаний и без противокасаний различаются коренным образом. Теория кривых без попутных касаний тесно связана с другим замечательным разделом геометрии малых размерностей — теорией лежандровых узлов. Аналогичные интерпретации J^- -теории неизвестны.

С ростом n размерности пространств и размеры систем уравнений для описания гауссовых инвариантов быстро растут; точный ответ неизвестен (по крайней мере, автору данной статьи) даже при $n = 4$. Тем более замечателен результат, полученный в 1997 г. А. Мерковым (см. [12]) и показывающий, что в теории кривых без тройных точек гауссовых инвариантов достаточно для полной классификации:

ТЕОРЕМА 3. Пусть γ_0 и γ_1 — две длинные кривые общего положения. Гомотопия γ_T без тройных точек, соединяющая γ_0 и γ_1 , существует тогда и только тогда, когда для каждого гауссова инварианта X типа странности выполнено равенство $X(\gamma_0) = X(\gamma_1)$.

Неизвестно, верен ли аналогичный результат для теории кривых без самокасаний.

5. ОБОБЩЕНИЯ ГАУССОВЫХ ИНВАРИАНТОВ

Схема определения инвариантов кривых с помощью гауссовых диаграмм оказывается чрезвычайно плодотворной. Подобным образом можно определить инварианты не только для длинных кривых, но и для других объектов, так или иначе связанных с кривыми.

Чтобы определить гауссовы функции на каком-либо множестве объектов, нужно проделать следующие действия:

1. Доказать для этих объектов аналог предложения 1, т.е. описать совокупность объектов «общего положения».
2. Для объектов общего положения построить комбинаторное описание, аналогичное гауссовым диаграммам. Желательно при этом доказать аналог предложения 2 — это даст надежду, что построенный класс гауссовых функций будет достаточно богатым.
3. Описать гомотопии общего положения (т.е. доказать аналог предложения 3) и изучить, как при таких гомотопиях меняются гауссовы диаграммы.

После этого вычисление гауссовых инвариантов, т.е. отбор из числа гауссовых функций тех, что не изменяются при гомотопиях, сводится, как и для длинных кривых, к решению системы линейных уравнений на коэффициенты.

Приведем примеры объектов, для которых эту программу удастся в той или иной мере осуществить.

5.1. ЗАМКНУТЫЕ КРИВЫЕ

Замкнутой кривой на плоскости называется гладкое отображение окружности в плоскость $\gamma(t) = (x(t), y(t)), t \in S^1$, обладающее тем свойством, что $\forall t \quad (x'(t))^2 + (y'(t))^2 > 0$. Последнее условие, как и для длинных кривых, обеспечивает возможность однозначного проведения касательной в любой точке образа кривой, т.е. гладкость кривой в обычном смысле слова.

Теория замкнутых кривых весьма близка к теории длинных кривых. В частности, для замкнутых кривых остаются верными, вместе с доказательствами, предложения 1 и 3. Для замкнутых кривых можно также определить число вращения способом, описанным на странице 101 (без использования гауссовых диаграмм), и вторая часть теоремы Уитни для них также верна — число вращения является единственной характеристикой замкнутой кривой, сохраняющейся при всех гомотопиях.

Для построения гауссовой диаграммы замкнутой кривой общего положения нарисуем окружность (которую пробегает параметр t в отображении $\gamma(t)$) и отметим на ней пары значений (t_1, t_2) , для которых $\gamma(t_1) = \gamma(t_2)$. Отмеченные пары значений соединим хордами. Для полной аналогии с длинными кривыми нужно еще приписать хордам знаки. Тут, однако, мы сталкиваемся с трудностью: для определения знака точки самопересечения использовался линейный порядок на прямой, иными словами, учитывалось, какое из значений t_1, t_2 больше, а какое меньше. На окружности естественного линейного порядка точек не существует, а взамен имеется циклический порядок: для *трех* различных точек

A, B, C всегда можно сказать, расположены они по часовой стрелке или против. Это обстоятельство подсказывает следующий выход. Отметим на окружности специальную точку O , и определим знак хорды (t_1, t_2) способом, описанным на странице 96, где вместо условия $t_1 < t_2$ потребуем, чтобы точки O, t_1, t_2 были расположены против часовой стрелки. Этим завершается построение гауссовых диаграмм для замкнутых кривых. Для корректности следует, конечно, добавить в условие общности положения кривой требование, чтобы отмеченная точка O не была точкой самопересечения — аналог предложения 1 при этом остается верным. К списку «типичных катастроф» гомотопии общего положения (предложение 3) в этом случае добавится еще одна — прохождение отмеченной точки O через точку самопересечения.

Приведенная конструкция гауссовых диаграмм позволяет определить гауссовы функции для замкнутых кривых. Однако если мы попытаемся выразить через гауссову диаграмму кривой хотя бы ее индекс Уитни (число вращения), то мы натолкнемся на неожиданную трудность: предложение 2 для замкнутых кривых неверно! Действительно, рассмотрим на плоскости две кривых. Первая кривая задается формулой $\gamma_1(t) = (\cos 2\pi t, \sin 2\pi t)$ и представляет собой единичную окружность, пробегаемую точкой против часовой стрелки за единичное время. Вторая кривая является образом первой при осевой симметрии, $\gamma_2(t) = (\cos 2\pi t, -\sin 2\pi t)$, и представляет собой ту же окружность, но пробегаемую по часовой стрелке. Число вращения у кривой γ_1 равно 1, а у кривой γ_2 оно равно -1 и, следовательно, эти кривые не гомотопны. Обе кривые, однако, не имеют точек самопересечения, и поэтому их гауссовы диаграммы выглядят одинаково — не содержат ни одной хорды.

Чтобы преодолеть эту трудность, введем понятие индекса отмеченной точки. Для этого рассмотрим вектор $r(t)$, соединяющий отмеченную точку O с точкой $\gamma(t)$, где t пробегает всю окружность, начиная с точки O и заканчивая в ней же. Поскольку O не является точкой самопересечения, $r(t)$ обращается в нуль только в начале и в конце пути. Поэтому можно определить количество оборотов вектора $r(t)$ вокруг начала координат, которое в данном случае будет *полуцелым* числом. Это число и называется индексом отмеченной точки O . При гомотопиях кривой, не задевающих отмеченную точку, ее индекс не меняется, а при прохождении отмеченной точки через точку самопересечения индекс увеличивается на ± 1 . Правило выбора знака в зависимости от знака точки самопересечения и направления прохождения читатель без труда установит самостоятельно.

ТЕОРЕМА УИТНИ ДЛЯ ЗАМКНУТОЙ КРИВОЙ. *Число вращения замкнутой кривой общего положения равно сумме знаков всех хорд на ее гауссовой диаграмме плюс удвоенный индекс отмеченной точки.*

Доказательство теоремы производится по схеме доказательства теоремы Уитни в разделе 2. Сперва следует проверить, что правая часть равенства в теореме Уитни не изменяется при типичных гомотопиях кривой — для этого нужно разобраться со знаками при прохождении отмеченной точки через точку самопересечения, как указано выше. После этого необходимо выбрать для каждого $k \in \mathbb{Z}$ стандартную кривую индекса k и проверить для нее утверждение теоремы.

Для замкнутых кривых существуют аналоги понятий странности и инвариантов J^\pm (фактически, первоначально эти инварианты были определены имен-

но для замкнутых кривых). Явные формулы для этих инвариантов включают, помимо суммирования по определенным парам хорд на гауссовой диаграмме, также и «негауссовы» члены, зависящие от индекса отмеченной точки. Явные формулы, а также обсуждение теории инвариантов высших порядков можно найти в статье [9]. Также существует развитая теория гауссовых инвариантов многокомпонентных кривых, см. статьи [9] и [11].

5.2. ЛЕЖАНДРОВЫ ФРОНТЫ

Понятие лежандрова фронта является одним из ключевых понятий контактной геометрии. Читатель, не знакомый с этим замечательным разделом современной математики, может пропустить первую фразу следующего абзаца, где дается определение, и считать определением то, что написано в дальнейшем. Лучшим выходом, однако, будет ознакомиться с основами симплектической и контактной теории — например, по книге [13]; см. также статью С. В. Дужина и С. В. Чмутава.

Лежандровы фронты на плоскости это, по определению, проекции лежандровых кривых в \mathbb{R}^3 со стандартной контактной структурой $dz - pdq$ на плоскость (q, z) . Эти проекции в общем положении представляют собой замкнутые плоские кривые с конечным числом точек негладкости (т.е. таких точек t , что $x'(t) = y'(t) = 0$). Все точки негладкости представляют собой простые точки возврата, устроенные, как точка $(0, 0)$ на кривой $x^2 = y^3$ (в подходящей системе координат, разумеется). Кроме того, фронт имеет в общем положении конечное число точек трансверсального самопересечения. Типичные гомотопии лежандровых фронтов описываются аналогично предложению 3. Помимо тройных точек и самокасаний, у фронтов бывают катастрофы еще двух типов — самопересечение в точке возврата и рождение (или, напротив, уничтожение) двух точек возврата (рис. 11).

Гауссовы диаграммы лежандровых фронтов определяются аналогично диаграммам плоских кривых и представляют собой окружности, на которые нанесены данные двух типов — во-первых, обозначены хордами точки самопересечения, и во-вторых, отмечены точки возврата. Знак хорды определяется так же, как для замкнутых кривых — с помощью вспомогательной отмеченной точки O .

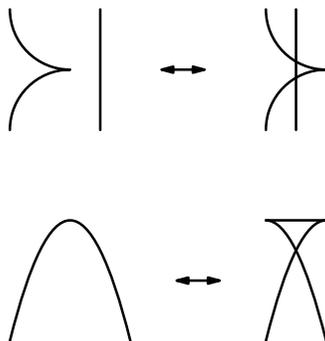


Рис. 11. Дополнительные катастрофы для лежандровых фронтов

Точкам возврата также можно приписать знаки. В точке возврата у фронта однозначно определена касательная прямая. Если рассмотреть касательный вектор к лежандровой кривой в \mathbb{R}^3 , и спроектировать его на плоскость, то проекция в точке возврата будет менять направление на противоположное. Рассмотрим теперь две точки t_1, t_2 , лежащие вблизи точки возврата t_* . Упорядочим номера точек так, чтобы t_1, t_*, t_2 были расположены против часовой стрелки. В общем положении касательные векторы к фронту $\gamma'(t_1)$ и $\gamma'(t_2)$ неколлинеарны, и, следовательно, образуют на плоскости базис. Припишем точке возврата t_* знак $+1$ или -1 в зависимости от того, является ли этот базис «правым» или «левым». Очевидно, от выбора точек t_1, t_2 (достаточно близких к t_* !) этот знак не зависит. Заметим также, что знак точки возврата не зависит и от выбора отмеченной точки O .

Построенные гауссовы диаграммы позволяют определять гауссовы функции по схеме, изложенной в предыдущих разделах. При этом перемножаются знаки как хорд, так и точек возврата, и учитывается взаимное расположение тех и других. Поэтому кольцо гауссовых функций в данном случае обладает биградуировкой — по числу хорд и по числу точек. В частности, существует единственная, с точностью до пропорциональности, гауссова функция бистепени $(0, 1)$ — сумма знаков всех точек возврата.

ТЕОРЕМА 4. *Сумма знаков всех точек возврата лежандрова фронта не изменяется при гомотопиях.*

ДОКАЗАТЕЛЬСТВО. При катастрофе рождения (уничтожения) двух точек возврата (см. рис. 11) две вновь появляющиеся (уничтожаемые) точки обязательно имеют разные знаки. При катастрофах прочих типов сохранение суммы знаков очевидно. \square

Построенный инвариант называется *индексом Маслова* лежандрова фронта. Можно доказать, что индекс Маслова и индекс Уитни составляют полную систему инвариантов гомотопий лежандровых фронтов на плоскости — см., например, книгу [13].

Дополнительные сведения о гауссовых инвариантах фронтов читатель может почерпнуть из статьи [9]. В частности, там доказано, что инвариант Беннекена (см. о нем в статье С. В. Дужина и С. В. Чмутова) также является гауссовым.

5.3. Узлы и зацепления

С основами теории узлов читатель может ознакомиться по статье С. В. Дужина и С. В. Чмутова в настоящем сборнике; мы будем пользоваться обозначениями этой статьи без подробных разъяснений.

Идея определения гауссовых инвариантов узлов состоит в том, чтобы рассмотреть гауссовы инварианты их диаграмм (проекций общего положения на плоскость), не изменяющиеся при преобразованиях Рейдемейстера. При этом необходимо учитывать, какая из ветвей кривой в точке самопересечения диаграммы проходит сверху. Чтобы отразить на гауссовой диаграмме соответствующую информацию, ориентируем каждую дугу — нарисуем на ней стрелку, направленную от верхней ветви к нижней.

Диаграмма узла представляет собою замкнутую кривую. Мы, однако, используем различие между «проходами» и «переходами» для определения знака

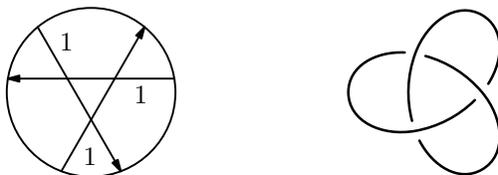


Рис. 12. Трилистник и его гауссова диаграмма

точки самопересечения (хорды на гауссовой диаграмме) без использования отмеченной точки. А именно, пусть (t_1, t_2) — точка самопересечения диаграммы: $x(t_1) = x(t_2), y(t_1) = y(t_2)$, причем $z(t_1) > z(t_2)$ (т. е. ветвь t_1 идет сверху). Тогда определим знак пересечения так же, как в разделе 1, считая точку t_1 «первой», а точку t_2 «второй». Такой знак часто называется английским словом “writhe” (to writhe — букв. «корчиться» (от боли); имеется в виду, что в районе точки самопересечения узел «извивается» в том или ином направлении). Таким образом, мы получаем гауссову диаграмму диаграммы узла — набор ориентированных дуг, снабженных знаками. На рис. 12 изображен узел «трилистник» и гауссова диаграмма его стандартной проекции.

Рассмотрим влияние преобразований Рейдемейстера на гауссовы диаграммы узлов. Первое преобразование Рейдемейстера сводится, очевидно, к тому, что на диаграмме появляется (или, наоборот, исчезает) дуга, между концами которой других дуг нет. Для длинных кривых такого преобразования не было, и быть не могло, поскольку сумма знаков всех дуг — индекс Уитни — сохраняется при любой гомотопии. Тем самым для узлов не существует аналога индекса Уитни — пространство гауссовых инвариантов порядка 1 пусто.

Два других преобразования Рейдемейстера представляют собой уже известные катастрофы кривых. Единственное отличие состоит в том, что они происходят только при определенных сочетаниях направлений и знаков участвующих в преобразовании дуг. Так, при самокасании (второе преобразование Рейдемейстера) возникающие (или исчезающие) дуги имеют одно и то же направление и разные знаки. Список возможных комбинаций знаков и направлений дуг, участвующих в третьем преобразовании Рейдемейстера (прохождение тройной точки), более длинен; читатель, желающий разобраться в доказательствах приводимых ниже результатов, может составить его самостоятельно.

Как и в случае длинных кривых, не всякая диаграмма со знаками и стрелками реализуется как гауссова диаграмма узла. А именно, имеет место такое утверждение.

ТЕОРЕМА 5. Пусть Γ — гауссова диаграмма некоторого узла и χ — хорда на ней. Тогда имеет место равенство:

$$\sum \varepsilon = \sum \varepsilon$$

Здесь суммирование ведется по всем хордам, расположенным так, как показано на рисунках под знаками суммирования; $\varepsilon = \pm 1$ означает знак хорды.



Рис. 13. Инвариант Васильева порядка 2

Теорема 5 является прямым аналогом (фактически следствием) теоремы 1 и доказывается аналогично.

Все известные на настоящий момент инварианты узлов являются инвариантами конечного типа (инвариантами Васильева) или выражаются через таковые — см. об этом в статье С. В. Дужина и С. В. Чмутова в настоящем сборнике. Не составляют исключения и гауссовы инварианты, ввиду следующей теоремы:

ТЕОРЕМА 6. *Всякий гауссов инвариант узлов порядка n является инвариантом конечного типа порядка не выше n .*

Доказательство этой теоремы несложно, но довольно длинно из-за необходимости перебора большого количества случаев. Мы не будем останавливаться на нем подробно.

Недавно (в 1998 г.) М. Гусаров доказал чрезвычайно трудную обратную теорему:

ТЕОРЕМА 7. *Всякий инвариант узлов конечного типа является гауссовым.*

Явное описание кольца гауссовых инвариантов узлов представляет собой задачу, аналогичную соответствующей задаче для кривых. Размерности пространств (и, соответственно, сложность вычислений) в случае узлов выше, так как дуги гауссовых диаграмм имеют направление, и тем самым число диаграмм с заданным количеством дуг больше. В настоящее время известны явные формулы для инвариантов Васильева порядков 2 и 3 (см. [10]); мы приведем здесь результат для инварианта порядка 2. Любопытно, что формулу можно существенно упростить, если ввести в гауссову диаграмму узла дополнительный элемент — отмеченную точку. Выбор этой точки не влияет на знаки хорд (в отличие от отмеченной точки на гауссовых диаграммах замкнутых кривых), но при вычислении гауссовой функции появляется возможность учитывать расположение этой точки относительно хорд. Разумеется, гауссовы инварианты узлов не должны, в частности, зависеть от выбора отмеченной точки.

ТЕОРЕМА 8. *Гауссова функция Φ , соответствующая диаграмме на рис. 13, является инвариантом порядка 2, принимающим значение 0 для тривиального узла и значение 1 для трилистника.*

Отметим, что существует ровно один инвариант порядка 2 с такими свойствами (доказательство см. в статье С. В. Дужина и С. В. Чмутова).

ДОКАЗАТЕЛЬСТВО. Докажем сначала, что гауссова функция Φ действительно является инвариантом. Независимость Φ от выбора отмеченной точки непосредственно следует из теоремы 5. При первом преобразовании Рейдемейстера появляется или исчезает хорда, не пересекающаяся с прочими хордами на диаграмме, — ясно, что такая операция не влияет на значение функции Φ .

Рассмотрим теперь второе преобразование Рейдемейстера с возникновением пары новых хорд. Возникающие хорды обязательно сонаправлены и имеют разные знаки; поэтому вклад в величину $\Phi(\gamma_{\text{после}}) - \Phi(\gamma_{\text{до}})$ может внести только слагаемое, в котором обе хорды — вновь возникшие. Однако такого слагаемого быть не может, так как при гомотопии общего положения самокасание не происходит в отмеченной точке.

Инвариантность Φ при третьем преобразовании Рейдемейстера проверяется аналогично (хотя и требует разбора большего количества случаев).

Таким образом, гауссова функция Φ — инвариант узла. Утверждение про порядок следует из теоремы 6. Очевидно, $\Phi(\text{тривиальный узел}) = 0$; равенство $\Phi(\text{трилистник}) = 1$ проверяется непосредственным вычислением. \square

СПИСОК ЛИТЕРАТУРЫ

- [1] В. И. Арнольд, А. Н. Варченко, С. М. Гусейн-Заде. Особенности диффеоморфизмов отображений. В 2-х томах. М.: Наука, 1984.
- [2] H. Whitney. *On regular closed curves in the plane*, *Comp. Math.* Vol. 4, 1937. P. 276–284.
- [3] А. Т. Фоменко, Д. Б. Фукс. Курс гомотопической топологии. М.: Наука, 1989.
- [4] В. Г. Болтянский, В. А. Ефремович. Наглядная топология. М.: Наука, 1983.
- [5] V. I. Arnold. Plane curves, their invariants, perestroikas and classification / In *Singularities and Bifurcations. Ser. Advances in Soviet Mathematics*. Vol 21, 1994. Amer. Math. Soc., Providence, RI. P. 33–91.
- [6] O. Viro. First degree invariants of generic curves on surfaces. Preprint Uppsala University, 1994.
- [7] A. Shumakovich. Formulas for strangeness of plane curves. Preprint Uppsala University, 1994.
- [8] F. Aicardi. Classification and invariants of tree-like curves / In *Singularities and Bifurcations. Ser. Advances in Soviet Mathematics*. Vol. 21, 1994. Amer. Math. Soc., Providence, RI. P. 1–31.
- [9] M. Polyak. Invariants of curves and fronts via Gauss diagrams // *Topology*. Vol. 37, No 5, 1998. P. 989–1009.
- [10] M. Polyak, O. Viro. Gauss diagram formulas for Vassiliev invariants.
- [11] V. A. Vassiliev. Invariants of ornaments / In *Singularities and Bifurcations. Ser. Advances in Soviet Mathematics*. Vol. 21, 1994. Amer. Math. Soc., Providence, RI. P. 225–262.
- [12] A. B. Merkov. Vassiliev invariants classify plane curves and doodles. Preprint 1996–98; submitted to *J. of Knot Theory and Ramifications*; <http://www.botik.ru/~duzhin/as-papers/ficd-dvi.zip>.
- [13] В. И. Арнольд. Особенности каустик и волновых фронтов. М.: Фазис, 1996.

Поверхность Зейферта

В. В. Прасолов

При изучении свойств узлов и зацеплений часто оказывается полезной вспомогательная ориентированная поверхность, натянутая на узел (зацепление). Таких поверхностей много, но обычно результат конструкции, основанной на одной из этих поверхностей, не зависит от выбора поверхности. Точное определение таково. Пусть L — ориентированное зацепление в $S^3 = \mathbb{R}^3 \cup \{\infty\}$ (ориентированность означает, что на каждой компоненте зацепления задана ориентация). *Поверхностью Зейферта* зацепления L называют связную компактную двумерную поверхность в S^3 , краем которой служит L , причем ориентация L индуцирована ориентацией поверхности.

ТЕОРЕМА 1. *У любого зацепления в S^3 есть поверхность Зейферта.*

ДОКАЗАТЕЛЬСТВО. Рассмотрим диаграмму D ориентированного зацепления L и уничтожим перекрестки диаграммы, как это показано на рис. 1. А именно, мы заменяем сплошные линии на пунктирные, которые проводятся единствен-

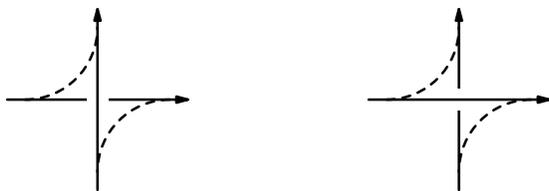


Рис. 1. Уничтожение перекрестков диаграммы

ным способом согласованно с ориентацией зацепления L . После уничтожения перекрестков получается набор замкнутых несамопересекающихся и попарно не пересекающихся кривых на плоскости. Эти кривые называют *окружностями Зейферта*.

Плоские диски, натянутые на окружности Зейферта, могут пересекаться (одна окружность Зейферта может лежать внутри другой). Но внутренние точки этих дисков можно сдвинуть в направлении, перпендикулярном плоскости диаграммы, так, чтобы диски стали попарно не пересекающимися (окружности Зейферта остаются при этом на месте). На каждом перекрестке соединим диски перекрученной полоской, краями которой служат ветви зацепления над этим перекрестком (на рис. 2 такая полоска заштрихована).

Естественные ориентации дисков согласованы с естественными ориентациями полосок, поэтому в результате мы получаем ориентированную поверхность, краем которой служит ориентированное зацепление L . Но эта поверхность не обязательно связна. Чтобы получить связную поверхность, возьмем две связные

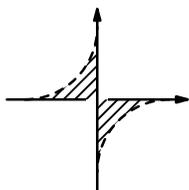


Рис. 2. Приклеивание полоски к дискам

компоненты, вырежем в каждой из них по маленькому диску и соединим возникшие края тонкой трубкой. В результате число компонент связности уменьшится, поэтому после нескольких таких операций получим связную поверхность. \square

Более подробное обсуждение этого доказательства и других взаимосвязей между узлами и двумерными поверхностями приведено в [1].

Род узла и простые узлы

Поверхность Зейферта узла K представляет собой компактную двумерную ориентированную поверхность, краем которой служит одна окружность. Приклеив к краю этой поверхности диск, получим двумерную поверхность рода g (сферу с g ручками). Родом узла K называют минимальный род поверхности, полученной таким образом (минимум здесь берется по всем поверхностям Зейферта узла K). Род узла K будем обозначать $g(K)$.

УПРАЖНЕНИЕ 1. Доказать, что узел тривиален тогда и только тогда, когда его род равен 0.

УПРАЖНЕНИЕ 2. Доказать, что род трилистника равен 1.

УПРАЖНЕНИЕ 3. Доказать, что род торического узла типа (p, q) не превосходит $(p-1)(q-1)/2$. (Торический узел типа (p, q) , где p и q — взаимно простые числа, задается в $S^3 \subset \mathbb{C}^2$ системой уравнений $|z| = |w|$, $z^p = w^q$. В действительности род торического узла типа (p, q) в точности равен $(p-1)(q-1)/2$.)

Для двух ориентированных узлов K_1 и K_2 можно определить их композицию, или связную сумму, $K_1 \# K_2$ посредством операции, изображенной на рис. 3.

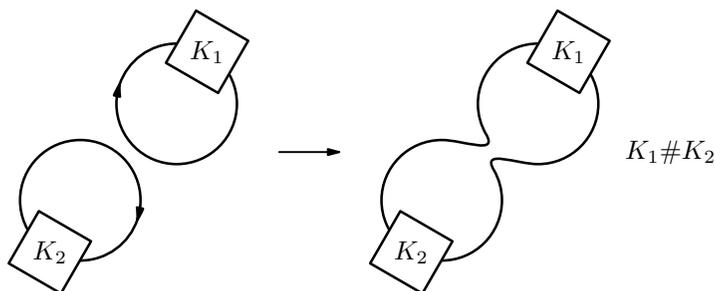


Рис. 3. Связная сумма узлов

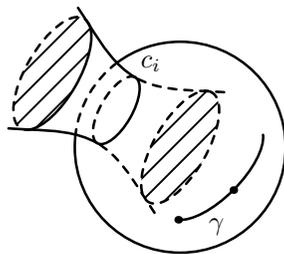


Рис. 4. Два параллельных диска

Неформально $K_1 \# K_2$ можно представлять так: мы берем веревку и завязываем на одном конце узел K_1 , а на другом конце узел K_2 ; затем концы веревки соединяем.

УПРАЖНЕНИЕ 4. Доказать, что узлы $K_1 \# K_2$ и $K_2 \# K_1$ эквивалентны.

ТЕОРЕМА 2. $g(K_1 \# K_2) = g(K_1) + g(K_2)$.

ДОКАЗАТЕЛЬСТВО. Пусть F_1 и F_2 — поверхности Зейферта минимального рода для узлов K_1 и K_2 . Разведем узлы K_1 и K_2 в разные стороны, так чтобы поверхности F_1 и F_2 не пересекались. Поверхности F_1 и F_2 не разбивают S^3 , поэтому их края K_1 и K_2 можно соединить длинной узкой полоской; чтобы ориентации поверхностей оказались согласованными, полоску можно перекрутить, в результате получим поверхность Зейферта рода $g(K_1) + g(K_2)$ для узла $K_1 \# K_2$. Это означает, что $g(K_1 \# K_2) \leq g(K_1) + g(K_2)$.

Пусть теперь F — поверхность Зейферта минимального рода для узла $K_1 \# K_2$. Из определения узла $K_1 \# K_2$ следует, что существует сфера $S^2 \subset S^3$, которая пересекает $K_1 \# K_2$ ровно в двух точках и разделяет узлы K_1 и K_2 . После малого шевеления поверхностей F и S^2 можно считать, что их пересечение $F \cap S^2$ состоит из нескольких гладких замкнутых кривых c_1, \dots, c_n и дуги γ , соединяющей точки пересечения S^2 и $K_1 \# K_2$. Каждая окружность c_i разбивает S^2 на два диска, один из которых не содержит γ . Этот диск заменяем на два параллельных ему диска, расположенных по разные стороны от S^2 ; на рис. 4 эти диски заштрихованы. Цилиндр, заключенный между новыми дисками, уничтожаем.

Сделаем такие перестройки поверхности F для всех окружностей c_1, \dots, c_n . В результате получим поверхность, одна из связных компонент которой является поверхностью Зейферта узла $K_1 \# K_2$. Общее число ручек всех связных компонент полученной поверхности не превосходит $g(F)$. Действительно, если перестройка по диску не нарушает связности поверхности, то число ручек уменьшается на 1, а если перестройка по диску разбивает поверхность на две связные компоненты, то общее число ручек не изменяется. Итак, мы построили для $K_1 \# K_2$ поверхность Зейферта F' , которая пересекает S^2 только по дуге γ и для которой выполняется неравенство $g(F') \leq g(F)$. Из минимальности рода поверхности F следует, что $g(F') = g(F) = g(K_1 \# K_2)$. Сфера S^2 разбивает F' на поверхности Зейферта для узлов K_1 и K_2 . Поэтому $g(K_1) + g(K_2) \leq g(F') = g(K_1 \# K_2)$. \square

СЛЕДСТВИЕ 1. Если узлы K_1 и K_2 нетривиальны, то узел $K_1 \# K_2$ нетривиален.

Прежде чем сформулировать еще два следствия теоремы 2, введем понятие простого узла. Нетривиальный узел P называют *простым*, если его нельзя представить в виде $P = K_1 \# K_2$, где K_1 и K_2 — нетривиальные узлы.

СЛЕДСТВИЕ 2. Если $g(K) = 1$, то K — простой узел.

СЛЕДСТВИЕ 3. Любой узел K можно представить в виде $K = P_1 \# \dots \# P_n$, где P_1, \dots, P_n — простые узлы.

ДОКАЗАТЕЛЬСТВО. Если узел K не простой, то его можно представить в виде $K = K_1 \# K_2$, где $1 \leq g(K_i) < g(K)$. Теперь можно воспользоваться индукцией по роду узла, основываясь на том, что если род узла равен 1, то узел простой. \square

Теоремы о существовании и единственности разложения на простые множители возникают в разных ситуациях. Наиболее известный пример теоремы такого рода — основная теорема арифметики. Как правило, существование разложения доказывается проще, чем единственность. Например, в основной теореме арифметики существование разложения очевидно (если $n = ab$ и $a > 1$, $b > 1$, то $a < n$ и $b < n$), а единственность требует нетривиального доказательства (наиболее распространенное доказательство основано на алгоритме Евклида). Более того, бывают ситуации, когда разложение на простые множители существует, но не единственно. Например, любая замкнутая двумерная поверхность представляется в виде связанной суммы некоторого количества торов T^2 и проективных плоскостей P^2 (которые уже не разлагаются в связанные суммы нетривиальным образом), но при этом $T^2 \# P^2 = P^2 \# P^2 \# P^2$.

Единственность разложения на простые множители нарушается также в кольцах целых алгебраических чисел. Первым это отметил Куммер, но его статья некоторое время не привлекала внимания. Вскоре после ее публикации французский математик Ламе предложил «доказательство» Последней теоремы Ферма, по поводу которого Лиувиль возразил, что оно опирается на единственность разложения на простые множители, которая для комплексных чисел не очевидна. Но Коши поддержал Ламе и некоторое время они пытались доказать единственность разложения. Это продолжалось до тех пор, пока Куммер не прислал Лиувиллю свою статью, в которой было показано, что разложение не единственно.

Вернемся, однако, к узлам. Представление узла K в виде $K = P_1 \# \dots \# P_n$, где P_1, \dots, P_n — простые узлы, единственно в том смысле, что если $P_1 \# \dots \# P_n = Q_1 \# \dots \# Q_m$ и все узлы P_i, Q_j простые, то $m = n$ и $P_i = Q_{\sigma(i)}$, где $\sigma \in S_n$ — некоторая подстановка. Единственность разложения узла K доказывается сложнее, чем существование. Для доказательства единственности нам потребуется следующее утверждение.

ТЕОРЕМА 3. Пусть $K = P \# Q = K_1 \# K_2$, причем узел P простой. Тогда либо $K_1 = P \# K'_1$ и $Q = K'_1 \# K_2$, либо $K_2 = P \# K'_2$ и $Q = K_1 \# K'_2$.

ДОКАЗАТЕЛЬСТВО. Пусть S^2_{PQ} и S^2_{12} — сферы в S^3 , которые пересекают узел K ровно в двух точках и разделяют узлы P и Q , K_1 и K_2 . После малого шевеления этих сфер можно считать, что множество $S^2_{PQ} \cap S^2_{12}$ состоит из конечного

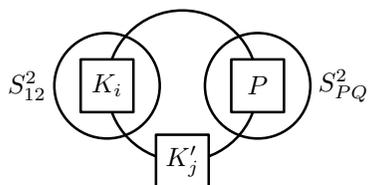


Рис. 5. Сферы S_{PQ}^2 и S_{12}^2

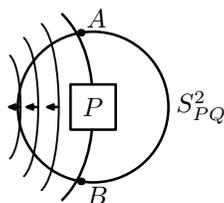


Рис. 6. Уничтожение кривых, не разделяющих точки A и B

числа гладких замкнутых несамопересекающихся (и попарно не пересекающихся) кривых c_1, \dots, c_n . Наша задача заключается в том, чтобы уничтожить все кривые c_1, \dots, c_n и сделать сферы S_{PQ}^2 и S_{12}^2 непересекающимися (рис. 5).

Пусть A и B — точки пересечения сферы S_{PQ}^2 и узла K . Каждая кривая c_i лежит на S_{PQ}^2 , поэтому она либо разделяет, либо не разделяет точки A и B . Кривые, которые не разделяют точки A и B , легко уничтожаются; начинать нужно с тех кривых, которые не содержат внутри себя других кривых (рис. 6).

При уничтожении кривых c_i , разделяющих точки A и B , нужно воспользоваться простотой узла P . Сфера S_{PQ}^2 разбивает S^3 на два трехмерных диска. Пусть D_P^3 — тот из этих дисков, который содержит узел P . Рассмотрим компоненту связности множества $D_P^3 \cap S_{12}^2$, содержащую кривую c_i . Эта компонента связности может быть либо двумерным диском (рис. 7 (а)), либо цилиндром (рис. 7 (б)). Во втором случае кривая c_i уничтожается вместе с еще одной кривой (вторым краем цилиндра). В первом случае двумерный диск D^2 пересекает узел K ровно в одной точке (мы предполагаем, что оба узла K_1 и K_2 нетривиальны;

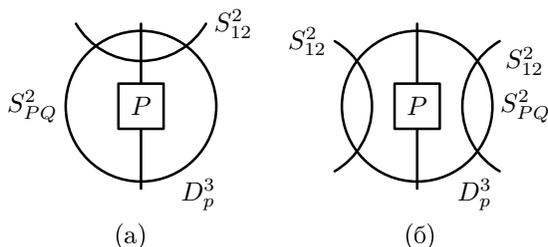


Рис. 7. Поверхность, натянутая на кривую c_i

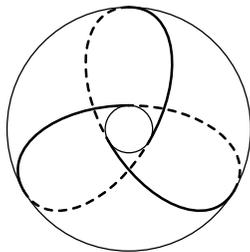


Рис. 8. Трилистник на торе

если один из этих узлов тривиален, то утверждение теоремы очевидно). Это следует из того, что сфера S_{12}^2 пересекает узел K ровно в двух точках, а если бы диск D^2 пересекал узел K в двух точках, то один из узлов K_1 и K_2 был бы тривиальным, а другой совпадал бы с P . Из простоты узла P следует, что в D_P^3 он целиком расположен по одну сторону от диска D^2 , поэтому диск D^2 можно уничтожить, уничтожив тем самым кривую c_i . \square

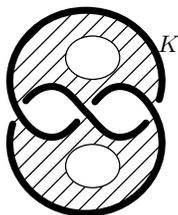
РАССЛОЕННЫЕ УЗЛЫ

Узел $K \subset S^3$ называют *расслоенным*, если существует семейство поверхностей Зейферта F_φ , $\varphi \in [0, 2\pi)$, которые целиком заполняют сферу S^3 , причем две разные поверхности пересекаются только по K .

ТЕОРЕМА 4. *Трилистник — расслоенный узел.*

ДОКАЗАТЕЛЬСТВО. Сферу $S^3 = \{(z, w) \in \mathbb{C}^2 \mid |z|^2 + |w|^2 = 1\}$ можно представить в виде объединения двух множеств, задаваемых неравенствами $|z| \leq |w|$ и $|z| \geq |w|$. Каждое из этих множеств является полноторием; их общая граница — тор. На этом торе расположен трилистник K , заданный уравнением $z^2 = w^3$ (рис. 8). Для одного из полноторий мы построим семейство заполняющих его поверхностей G_φ , каждая из которых гомеоморфна поверхности, изображенной на рис. 9. После этого требуемое семейство поверхностей F_φ легко будет построить, заклеив две круглые дырки в поверхности G_φ дисками.

Одна из поверхностей G_φ изображена на рис. 10 (а); ее край состоит из узла K и двух окружностей C_1 и C_2 . Чтобы более детально описать поверхность G_φ ,

Рис. 9. Часть поверхности F_φ

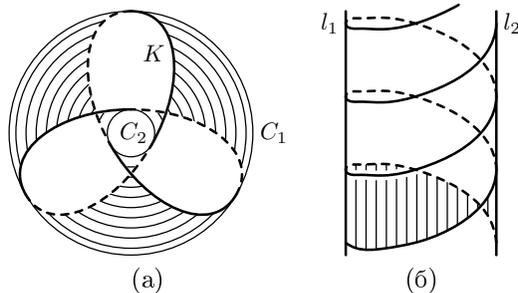


Рис. 10. Строение поверхности G_φ

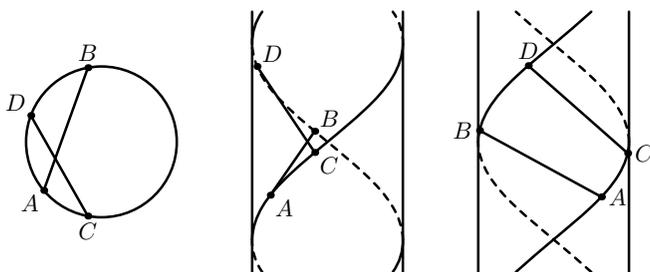


Рис. 11. Три проекции на координатные плоскости

мы воспользуемся тем, что полноторие можно получить из бесконечного прямого кругового цилиндра очевидной факторизацией. На поверхности этого цилиндра трилистнику K соответствуют две винтовые линии (рис. 10 (б)). Фиксируем для цилиндра две параллельные опорные плоскости Π_1 и Π_2 и рассмотрим поверхность G' , состоящую из отрезков, параллельных опорным плоскостям; концы отрезков лежат при этом на соседних витках винтовых линий. Часть поверхности G' изображена на рис. 10 (б). Граница поверхности G' состоит из двух винтовых линий и прямых l_1 и l_2 , по которым опорные плоскости касаются поверхности цилиндра.

Поворачивая опорные плоскости, мы получим семейство поверхностей G'_φ (угол φ соответствует удвоенному углу поворота опорных плоскостей). Две винтовые линии служат общим краем всех этих поверхностей. Рассмотрев проекции на три координатные плоскости, можно показать, что две разные поверхности G'_φ не имеют общих точек помимо их общего края. А именно, если проекции отрезков AB и CD , из которых состоят поверхности G'_φ , на две координатные плоскости пересекаются, то их проекции на третью координатную плоскость не пересекаются (рис. 11).

Очевидно, что каждая точка поверхности цилиндра принадлежит некоторой поверхности G'_φ . Покажем, что и каждая внутренняя точка x цилиндра тоже принадлежит некоторой поверхности G'_φ . Рассмотрим для этого часть поверхности G' , изображенную на рис. 10 (б). Прямая l_x , проведенная через точку x параллельно оси цилиндра, пересекает эту часть поверхности G' ровно в одной

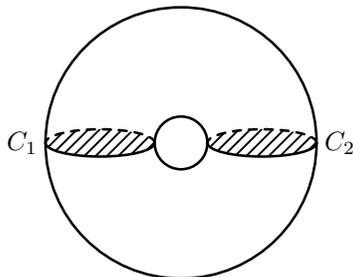


Рис. 12. Семейство поверхностей в другом полнотории

точке $x(0)$. Поворачивая поверхность G' так, чтобы ее края двигались по винтовым линиям, мы будем получать точки $x(\varphi) \in l_x \cap G'_{\varphi \pmod{2\pi}}$, $\varphi \in \mathbb{R}$. Точка $x(\varphi)$ непрерывно зависит от φ и любая точка прямой l_x расположена между точками $x(2\pi n)$ и $x(2\pi(n+1))$ для некоторого $n \in \mathbb{Z}$. Поэтому $x = x(\varphi)$ для некоторого $\varphi \in \mathbb{R}$.

Мы разобрались с одним полноторием. На другом полнотории окружности C_1 и C_2 выглядят так, как показано на рис. 12. Ясно, что диски, заклеивающие семейство таких окружностей, заполняют полноторие и попарно не пересекаются. \square

БЕСКОНЕЧНОЕ ЦИКЛИЧЕСКОЕ НАКРЫТИЕ

Оставшаяся часть статьи предназначена для тех, кто знаком с элементами теории накрытий.

Если X — достаточно хорошее линейно связное топологическое пространство, то для любой подгруппы $\pi \subset \pi_1(X, x_0)$ существует такое накрытие $p: \tilde{X} \rightarrow X$, что $p_*\pi_1(\tilde{X}, \tilde{x}_0) = \pi$, $p(\tilde{x}_0) = x_0$. Достаточно хорошим является, например, любое локально линейно связное локально односвязное пространство. Нас будет в основном интересовать случай $X = S^3 \setminus K$, где K — узел, но циклические накрытия удобно обсудить сразу в общей ситуации.

Накрытие $p: \tilde{X} \rightarrow X$ называют *циклическим*, если подгруппа $p_*\pi_1(\tilde{X}, \tilde{x}_0) \subset \pi_1(X, x_0)$ нормальная и факторгруппа $G = \pi_1(X, x_0)/p_*\pi_1(\tilde{X}, \tilde{x}_0)$ циклическая. Если $G = \mathbb{Z}_n$, то накрытие p называют *циклическим порядка n* , а если $G = \mathbb{Z}$, то накрытие p называют *бесконечным циклическим*. Из стандартных теорем теории накрытий следует, что накрытие $p: \tilde{X} \rightarrow X$ циклическое тогда и только тогда, когда группа автоморфизмов этого накрытия циклическая, причем группа автоморфизмов действует на накрытии транзитивно, т. е. для любых точек $\tilde{x}_1, \tilde{x}_2 \in p^{-1}(x_0)$ существует автоморфизм накрытия, переводящий \tilde{x}_1 в \tilde{x}_2 .

Для циклических накрытий группа $\pi_1(X, x_0)/p_*\pi_1(\tilde{X}, \tilde{x}_0)$ абелева, поэтому группа $p_*\pi_1(\tilde{X}, \tilde{x}_0)$ содержит коммутант π' группы $\pi_1(X, x_0)$. Это означает, что если $\pi_1(X, x_0)/\pi' \cong \mathbb{Z}$, то бесконечное циклическое накрытие $p: \tilde{X} \rightarrow X$ существует и единственно (с точностью до эквивалентности накрытий). Действительно, если $\pi_1(\tilde{X}, \tilde{x}_0) \supset \pi'$ и $\pi_1(X, x_0)/\pi' \cong \pi_1(X, x_0)/p_*\pi_1(\tilde{X}, \tilde{x}_0)$, то $p_*\pi_1(\tilde{X}, \tilde{x}_0) = \pi'$, а подгруппа $p_*\pi_1(\tilde{X}, \tilde{x}_0) \supset \pi_1(X, x_0)$ однозначно задает накрытие.

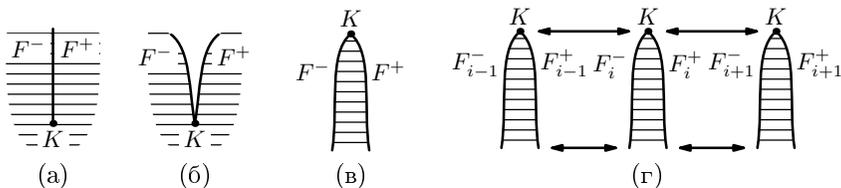


Рис. 13. Построение бесконечного циклического накрытия

Циклическое накрытие порядка n в таком случае тоже существует и единственно, потому что в \mathbb{Z} есть ровно одна группа H , для которой $\mathbb{Z}/H \cong \mathbb{Z}_n$, а именно, $H = n\mathbb{Z}$.

Если K — узел в S^3 , то фундаментальная группа пространства $X = S^3 \setminus K$ как раз и обладает тем свойством, что факторгруппа ее по коммутанту изоморфна \mathbb{Z} . Поэтому для $X = S^3 \setminus K$ однозначно определено бесконечное циклическое накрытие $p: \tilde{X} \rightarrow X$. Пространство \tilde{X} удобно строить с помощью поверхности Зейферта узла K . Независимость этой конструкции от выбора поверхности Зейферта следует из того, что пространство \tilde{X} единственно (с точностью до гомеоморфизма).

Пусть F — поверхность Зейферта узла K . Разрежем сферу S^3 по F . В результате получим 3-мерное многообразие M^3 с краем $F^+ \cup F^-$; каждой точке $a \in F \setminus K$ соответствуют две точки $a^+ \in F^+$ и $a^- \in F^-$ (возможность выбрать знаки $+$ и $-$ для всех точек согласованным образом следует из ориентируемости поверхности F). Возьмем счетное множество экземпляров M_i^3 , $i \in \mathbb{Z}$, таких многообразий и отождествим F_{i-1}^+ с F_i^- (отождествляются точки a_{i-1}^+ и a_i^- , соответствующие одной и той же точке). Схематично эта конструкция изображена на рис. 13. Полученное пространство \tilde{X} покрывает $X = S^3 \setminus K$. Отображения $M_i^3 \rightarrow M_{i+1}^3$, $i \in \mathbb{Z}$, индуцируют автоморфизм σ этого накрытия, причем группа автоморфизмов накрытия порождается элементом σ и изоморфна \mathbb{Z} . Кроме того, группа автоморфизмов действует на накрытии транзитивно. Поэтому $p: \tilde{X} \rightarrow X$ — бесконечное циклическое накрытие.

Бесконечное циклическое накрытие пространства $S^3 \setminus K$ используется для построения важного инварианта узла K — полинома Александра. Кроме того, одно из многих возможных определений коэффициента зацепления тоже использует бесконечное циклическое накрытие.

КОЭФФИЦИЕНТ ЗАЦЕПЛЕНИЯ

Пусть J и K — ориентированные замкнутые кривые в S^3 . Предполагается, что эти кривые не пересекаются и не имеют самопересечений. Рассмотрим диаграмму ориентированного зацепления, составленного из кривых J и K . Будем обращать внимание лишь на те перекрестки, где кривая K проходит над J . Такие перекрестки бывают двух типов (см. рис. 14).

Для каждого рассматриваемого перекрестка возьмем соответствующее значение $\varepsilon_i = \pm 1$ и сложим все числа ε_i . Полученное в результате число называют коэффициентом зацепления замкнутых ориентированных кривых J и K в S^3 и

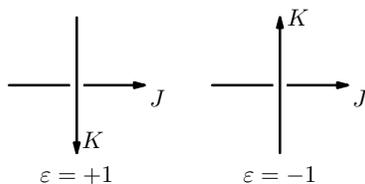


Рис. 14. Выбор знака

обозначают $lk(J, K)$. В [2] показано, что коэффициент зацепления не зависит от выбора диаграммы зацепления $\{J, K\}$; кроме того, $lk(J, K) = lk(K, J)$.

Известно много других (эквивалентных) определений коэффициента зацепления. Мы приведем два определения коэффициента зацепления, первое из которых использует поверхность Зейферта явно, а второе — неявно, через посредство бесконечного циклического накрытия.

ОПРЕДЕЛЕНИЕ 1. Пусть F — поверхность Зейферта ориентированного узла J . Будем считать, что ориентированный узел K трансверсально пересекает поверхность F в конечном числе точек. Каждой точке пересечения сопоставим число $\varepsilon_i = \pm 1$ по следующему правилу. Зададим на поверхности Зейферта ориентацию, считая положительно ориентированным репер $\{e_1, e_2\}$, где e_1 — вектор скорости кривой J , e_2 — вектор внутренней нормали (он перпендикулярен e_1 , касается поверхности F и направлен внутрь F). Пусть e_3 — вектор скорости кривой K в точке $a \in K \cap F$, $\{e'_1, e'_2\}$ — положительно ориентированный репер на поверхности F в точке a . Тогда $\varepsilon_i = +1$, если ориентация репера $\{e'_1, e'_2, e_3\}$ совпадает с ориентацией сферы S^3 ; в противном случае $\varepsilon_i = -1$. Сумму всех чисел ε_i называют коэффициентом зацепления кривых J и K .

ОПРЕДЕЛЕНИЕ 2. Пусть $p : \tilde{X} \rightarrow X$ — бесконечное циклическое накрытие пространства $X = S^3 \setminus J$, σ — образующая группы автоморфизмов этого накрытия (таких образующих две; какую именно нужно выбрать, мы обсудим чуть позже). Кривую K можно рассматривать как петлю в X с началом и концом в точке $x_0 \in K$. Пусть $\tilde{x}_1 \in p^{-1}(x_0)$. Рассмотрим поднятие петли K с началом в точке \tilde{x}_1 . Пусть \tilde{x}_2 — конец этого поднятия. Тогда $\tilde{x}_2 = \sigma^m(\tilde{x}_1)$. Число m называют коэффициентом зацепления кривых J и K .

Покажем, что определения 1 и 2 эквивалентны исходному определению коэффициента зацепления. Отметим, что в определении 1 есть зависимость от выбора поверхности Зейферта F , а в определении 2 такой зависимости нет. Поэтому достаточно доказать следующие утверждения:

а) для некоторой поверхности Зейферта F определение 1 совпадает с исходным определением коэффициента зацепления;

б) для любой поверхности Зейферта F определение 1 совпадает с определением 2.

Начнем с доказательства утверждения а). Можно считать, что поверхность Зейферта F узла J расположена над плоскостью его диаграммы (здесь мы не обращаем внимания на перекрестки диаграммы узла J , потому что проекция кривой K не пересекает их). Можно также считать, что кривая K расположена в плоскости диаграммы и лишь на перекрестках проходит под или над

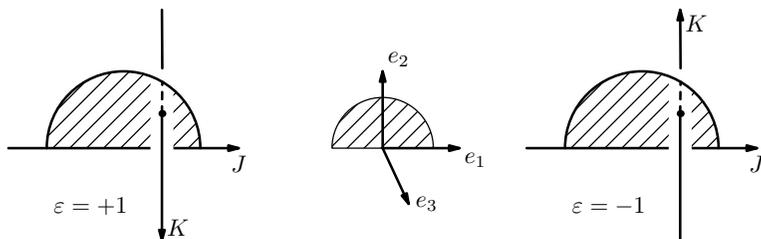


Рис. 15. Пересечение узла и поверхности Зейферта

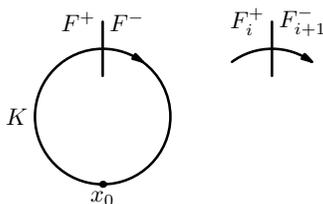


Рис. 16. Выбор знака

плоскостью. В такой ситуации кривая K пересекает поверхность F в точности на тех перекрестках, на которых кривая K проходит над J . При этом знаки \pm , соответствующие точкам пересечения K и F , совпадают со знаками, соответствующими перекресткам диаграммы зацепления $\{J, K\}$ (рис. 15).

Займемся теперь доказательством утверждения б). По ходу доказательства мы выясним, какую именно из двух образующих циклической группы автоморфизмов накрытия нужно выбирать. Построим пространство \tilde{X} с помощью произвольной поверхности Зейферта F узла J . Пусть сначала кривая K пересекает поверхность F ровно в одной точке, причем пересекает в направлении от F^+ к F^- (рис. 16). Тогда если \tilde{x}_1 и \tilde{x}_2 — начало и конец поднятия петли K , то $\tilde{x}_2 = \sigma(\tilde{x}_1)$, где автоморфизм σ индуцирован отображениями $M_i^3 \rightarrow M_{i+1}^3$. Поэтому знаки поверхностей F^+ и F^- нужно выбрать так, чтобы положительно ориентированный репер на поверхности F вместе с вектором, направленным от F^+ к F^- , образовывал положительно ориентированный репер в S^3 . (Выбор знаков поверхностей F^+ и F^- эквивалентен выбору образующей σ .) При таком выборе знаков в общей ситуации количество точек пересечения кривой K с поверхностью F (с учетом знаков) равно как индексу зацепления, вычисленному по определению 1, так и индексу зацепления, вычисленному по определению 2.

СПИСОК ЛИТЕРАТУРЫ

- [1] Прасолов В. В. Наглядная топология. М.: МЦНМО, 1995.
- [2] Прасолов В. В., Сосинский А. Б. Узлы, зацепления, косы и трехмерные многообразия. М.: МЦНМО, 1997.

Студенческие чтения

О сложности вычислений

А. А. Разборов

Лекция, прочитанная студентам Математического Колледжа НМУ
23 апреля 1998 г.

Эта лекция рассчитана на тех, кто не знаком с теорией сложности вычислений. Поэтому здесь будет рассказано только об основах этой теории и самых первых ее результатах. При этом мы постараемся передать основные идеи, которыми руководствуются исследователи в этой области науки.

Сюжетной канвой дальнейшего рассказа послужат истории, происходящие с некоторым персонажем — назовем его M (от слова «математик»). И начнем рассказ со следующей истории.

1. ПРЕДЫСТОРИЯ

Однажды M сидел дома и пытался доказать некоторую (возможно важную, возможно нет) теорему T . Он пытался доказать эту теорему в течение недели, двух, месяца, ..., но ничего у него не получалось. В конце концов он не выдержал и задал вполне естественный вопрос:

а можно ли в принципе доказать эту теорему?

Вопрос был обращен неизвестно к кому, и, скорее всего, улетел бы в пространство, если бы мимо не проходил другой наш персонаж, которого мы обозначим буквой L , от слова «логик».

L услышал вопрос, зашел в комнату и объяснил, что такими вопросами математики начали интересоваться примерно с начала 20-го века. В общем виде этот вопрос входит в знаменитую «программу Гильберта», посвященную понятию математического доказательства.

Эта программа, в частности, включала три следующих пункта.

1) ФОРМАЛИЗАЦИЯ ПОНЯТИЯ ДОКАЗАТЕЛЬСТВА.

Перед тем как задавать вопросы о том, можно ли доказать то или иное утверждение, необходимо дать строгое математическое определение доказуемости.

2) Полнота.

После формализации понятия доказательства нужно установить *полноту* построенной *формальной теории*. Это означает, что любое истинное утверждение T должно быть доказуемым в нашей формализации. В частности (в силу некоторых исторических причин), самого Гильберта преимущественно интересовал вопрос о доказуемости утверждения о непротиворечивости математики, формализованного подходящим образом.

3) РАЗРЕШИМОСТЬ.

Следующей целью программы было построение такого вычислительного устройства, которое по внешнему виду теоремы T , записанной в некотором формальном языке, определяло бы, является ли эта теорема доказуемой (предполагалось, что в силу пункта 2 программы это будет эквивалентно истинности).

Хорошо известно, что первый пункт программы был успешно выполнен. В настоящее время под словом «теорема» большинство математиков (хотя не все это признают) понимает то, что можно доказать в теории множеств Цермело – Френкеля.

Хуже дело обстояло с последующими двумя пунктами. Первым ударом, потрясшим в свое время все математическое сообщество, был полученный в 30-е годы результат Курта Гёделя, утверждающий, что никакая достаточно сильная теория с задаваемым явным списком множеств аксиом не может быть полной. Более точно, если наша теория непротиворечива и в ней можно формализовать рассуждения о натуральных числах, то существует утверждение, которое нельзя ни доказать, ни опровергнуть, более того, в качестве примера можно взять интересовавшее Гильберта утверждение о непротиворечивости рассматриваемой теории.

Мы сегодня занимаемся вычислениями, поэтому для нас более интересна теорема Чёрча (1936). Он доказал, что не существует никакого алгоритма, который по утверждению автоматически проверял бы, является это утверждение доказуемым или нет.

Итак, наш математик M узнал от логика L , что в полной общности ответить на вопрос о доказуемости математического утверждения невозможно.

Впрочем, к этому моменту у него уже начали появляться сомнения в том, что теорема T истинна, и он решил заняться поиском контрпримера. M пошел в компьютерную комнату и написал программу P , проверяющую последовательно все слова до тех пор, пока она не найдет данные, на которых T не верна. Он запустил свою программу и стал ждать окончания ее работы.

Программа работала час, два, день, неделю, . . . И снова у математика возник вопрос, на этот раз:

остановится ли P когда-нибудь?

На этот раз математик уже знал, к кому обратиться с таким вопросом. Он нашел логика и спросил существует ли способ узнать, прекратит ли данная программа свое выполнение или же будет работать до бесконечности. И снова L дал вполне квалифицированный ответ. По теореме, доказанной Тьюрингом в 30-ые годы, не существует алгоритма, определяющего, остановится ли когда-нибудь данная программа или нет, т. е. проблема остановки неразрешима.

2. ОТПРАВНАЯ ТОЧКА

После этого наш математик отправился домой. Как это иногда бывает, его ребенок, изучающий геометрию в седьмом или восьмом классе, попросил папу помочь сделать домашнее задание. В народе бытует мнение, что люди, занимающиеся математикой, умеют хорошо считать, решать квадратные уравнения и разные задачки из элементарной геометрии, что не всегда соответствует действительности.

Итак, наш математик начал решать задачу, предложенную своим чадом, и понял, что напрочь забыл геометрию, которую учил в школе. Он однако помнил, что любую задачу геометрии можно записать в координатах на языке действительных чисел.

И ему в голову пришел вопрос, а существует ли универсальный алгоритм, решающий школьные задачи по геометрии. После разговора с логиком, он знал, что если в теории выразимы натуральные числа, то она неразрешима. Интуиция подсказывала, что теория, работающая с действительными числами, которых гораздо больше, должна быть уж тем более неразрешима. Все же, для очистки совести, он решил позвонить логике и удостовериться в этом. Как ни странно, выяснилось противоположное. Классический результат Тарского, доказанный в 1948 году, утверждал существование алгоритма, проверяющего доказуемость утверждений элементарной геометрии¹⁾.

Математик обрадовался. Ему уже начинало казаться, что от логики нет никакой пользы, а тут оказалось, что логика имеет весьма практичное применение в реальной жизни.

Итак, М попросил ребенка подождать, а сам отправился к торговцам программного обеспечения. Там он обнаружил два CD-диска, посвященных решению задач по геометрии, которые назывались соответственно

`Tarsky for Windows 95` и `Collins for Windows 95`.

Первый диск стоил 30 у.е., второй — 300 у.е. Естественно, М попытался понять, чем же вызвана такая разница в цене. Никаких объяснений он не получил, поэтому он купил более дешевый диск `Tarsky for Windows 95`.

М пришел домой, вставил CD в компьютер и решил проверить программу на задачке из школьного учебника. Однако стала повторяться все та же история. Программа работала час, два, . . . , но не проявляла никакого желания выдать решение. Математик прервал выполнение программы и попросил ее доказать какую-нибудь элементарную теорему, например о том, что сумма углов в треугольнике равна 180 градусам. Результат не сильно изменился, и программа продолжала думать час, второй, сутки, другие, . . . Тогда М позвонил логике и несколько повышенным тоном спросил, что же происходит. Л ответил, что это не его проблема. Тарский доказал теорему о существовании алгоритма, он совершенно уверен, что авторы этого диска запрограммировали алгоритм правильно, а что происходит дальше — не имеет никакого отношения ни к математике, ни к логике.

¹⁾ И вообще любых утверждений о вещественных числах, использующих арифметические операции, элементарные логические связи (отрицание, конъюнкция, дизъюнкция, импликация) и кванторы по множеству всех вещественных чисел: «для всех (вещественных чисел)», «существует (вещественное число), такое что. . .»; совокупность таких утверждений называется алгеброй Тарского.

И это как раз то место, где начинается теория сложности вычислений. Нас интересует не просто существование алгоритмов для нашей задачи, а то, насколько они эффективны.

Конечно, рассказанная история несколько стилизована, но, в общем-то, так и происходило развитие исследований, которое привело к современному состоянию. В частности, тот этап, когда стало выкристаллизовываться понимание того, что одни алгоритмы могут быть лучше других и это существенно, пришелся на 60-ые годы. Тогда стали появляться настоящие компьютеры (они назывались таким страшным словосочетанием: «электронные вычислительные машины»), например, БЭСМ (многие в аудитории и не знают, наверное, что это такое). Стало понятно, что необходимо построение некоторой математической теории.

3. НАЧАЛА ТЕОРИИ: ОСНОВНЫЕ ПОНЯТИЯ

Давайте пока оставим в стороне истории из жизни математика М, потом мы к нему еще вернемся не раз.

Попробуем дать несколько определений.

Первое наблюдение, которое сделал М, когда пытался разобраться в этой теории, состоит в том, что подавляющее большинство алгоритмических задач можно представить в подходящей кодировке как задачи вычисления некоторого отображения

$$f : \{0, 1\}^* \rightarrow \{0, 1\}^*$$

множества конечных двоичных слов в себя. Именно такими отображениями мы и будем заниматься.

Второй вопрос: на каких устройствах мы собираемся вычислять наше отображение? Есть огромное разнообразие и алгоритмических языков, и компьютерных архитектур, но насколько нам это важно? Именно ответ на этот вопрос отличает теорию сложности вычислений от различных смежных областей. Поэтому рассмотрим этот вопрос более подробно.

В качестве примера возьмем программу, с которой чаще всего и сталкивается большинство математиков, которые не программируют. Это TeX процессор. Выбор TeX здесь никакого особого значения не имеет, то же самое можно сказать и о других программах.

Итак, имеется некоторое отображение f , которое в данном случае представляет собой отображение

$$\boxed{\text{paper.tex}} \rightarrow \boxed{\text{paper.dvi}},$$

преобразовывающее файл `paper.tex` в файл `paper.dvi` (оба файла мы рассматриваем как два длинных двоичных слова). Это преобразование осуществляется некоторым алгоритмом `texdvi.exe`, закодированным для выполнения на 286-м процессоре (те, кто пытался использовать TeX на 286-м процессоре, помнят, как это было).

Нас сейчас интересует время (быстродействие) работы алгоритма. Именно такую, как говорят в этой науке, *меру сложности* мы будем рассматривать сегодня чаще всего. Бывают и другие характеристики (память — вторая по зна-

чению мера сложности), но за недостатком времени мы их не сможем подробно обсудить.

Развитие техники, относящееся к рассматриваемой нами задаче, шло по двум направлениям. Во-первых, и это хорошо все знают, улучшались процессоры:

Intel 286, Intel 386, Intel 486, . . . ,

появлялись один за другим все более мощные модели, возникали разнообразные конструкторские ухищрения, ускоряющие их работу, и т. д. Во-вторых, улучшался сам алгоритм, появлялись его версии

texdvi1.exe, . . . , texdvi10.exe,

Справедлива, хотя и с некоторыми оговорками, такая формула

⟨общее время работы⟩ = ⟨число операций алгоритма⟩ × ⟨время на одну операцию⟩.

Улучшение обоих членов в этой простой формуле — вещи более-менее независимые. Грубо говоря, программы (software) отвечают за первые сомножитель, а процессоры (hardware) — за второй.

Что здесь происходит с математической точки зрения? Если у нас имеется какой-то алгоритм, который использует t операций, то от улучшения быстродействия нашего процессора время решения нашей задачи умножается на некоторую константу.

Поэтому, и это очень важно, в теории сложности вычислений сложилась традиция измерять время работы алгоритма с точностью до мультипликативной константы («с точностью до $O(\cdot)$ », как говорят математики). Такой подход позволяет нам абстрагироваться от выбора конкретной модели вычислительного устройства, от того, сколько времени занимает выполнение одной операции, какую систему команд имеет используемая нами машина и т. п. Именно этот подход и позволяет нам построить довольно красивую математическую теорию.

Теория сложности вычислений несколько отличается от всеобъемлющей (по определению) теории марксизма-ленинизма (не все в аудитории видимо знают, что это такое, и это хорошо): если в формуле есть два члена, и за второй наша теория никакой ответственности не несет, то об этом заявляется явно. Улучшением процессоров занимаются другие люди, мы занимаемся улучшением алгоритмов с точностью до мультипликативной константы.

Продолжим построение строгой теории. Одним из непосредственных преимуществ введенного выше соглашения является то, что нас не очень заботит выбор точной модели. Как правило, от смены модели (сейчас я имею в виду уже абстрактные математические модели) улучшение или ухудшение происходят с точностью до мультипликативной константы, а мы договорились этого не замечать.

Стандартный способ определения вычислительной модели состоит в том, что для исполнения алгоритма у нас имеется некоторая машина с адресуемой памятью, ячейки которой индексируются натуральными числами; в каждой ячейке можно хранить натуральные числа; можно выполнять арифметические действия и т. п. Детали здесь мало интересны, потому что наше $O(\cdot)$ -соглашение позволяет о них заботиться не слишком. Если же заменить «с точностью до $O(\cdot)$ » на слегка более грубое «с точностью до полинома», то вообще все реалистические вычислительные устройства оказываются эквивалентными.

Итак, фиксируем некоторую вычислительную модель. Если есть машина M , которая вычисляет функцию f , и определены входные данные x — некоторое

двоичное слово, тогда можно определить нашу основную функцию $T(M, x)$ — количество операций (тактов), которые требуются машине M для работы на входном слове x . Функция $T(M, x)$ называется *сигнализирующей функцией по времени*. выполнения любых элементарных операций

4. ТЕОРЕМА ОБ УСКОРЕНИИ

Давайте теперь изучать эту функцию $T(M, x)$. Первое, что приходит в голову: у нас имеется алгоритмическая задача f ; давайте выберем самый хороший алгоритм для решения этой задачи и назовем сложностью задачи сложность этого самого хорошего алгоритма.

Оказывается, что такой интуитивно очевидный способ действий, к сожалению, неосуществим и на этот счет имеется *теорема Блюма об ускорении*. В вольной формулировке эта теорема утверждает, что как бы мы ни пытались определять понятие «наилучшая для данной задачи машина», у нас ничего не получится (по крайней мере для некоторых задач).

Теорема Блюма открывает серию теорем, лежащих в основе современной теории сложности вычислений. Все эти теоремы были доказаны в 70-е годы, например, теорема Блюма об ускорении — это 1971 год. Примерно в то же время появилась концепция NP-полноты, изложением которой и завершится наш вводный рассказ о теории сложности вычислений.

Для дальнейшего нам потребуется еще одно важное определение. Функция $T(M, x)$ устроена очень нерегулярно. Рассмотрим пример TeX процессора. Для подавляющего большинства файлов произойдет остановка в самом начале работы из-за несоответствия входному формату TeX, а на каких-то файлах это время возможно окажется бесконечным из-за заикливания. В общем случае эту функцию изучать никакой возможности не представляется, она слишком рыхлая. Мы хотим извлечь из нее функцию натурального аргумента, т. е. получить функцию из натуральных чисел в натуральные числа, которая отражала бы поведение $T(M, x)$. Есть несколько способов подойти к этой задаче. Мы здесь ограничимся самым распространенным, который называется *сложность в наихудшем случае*. Она определяется следующей формулой

$$t_M(n) = \max_{|x| \leq n} T(M, x).$$

Из всех слов битовой длины, не превышающей n , выбираем то, на котором наша машина работает хуже всего (т. е. дольше всего). Время работы на таком слове и называем сложностью $t_M(n)$. За время $t_M(n)$ машина гарантированно закончит работу на любом входном слове длины, не превышающей n . Может, конечно, случиться, что на каких-то словах работа завершится раньше.

Мы приводим упрощенный вариант теоремы Блюма; на самом деле, вместо $\log t$ в ней можно взять любую «разумную» стремящуюся к бесконечности функцию.

ТЕОРЕМА (БЛУМ, 1971). *Существует такая вычислимая²⁾ функция f , что любую машину M , вычисляющую f , можно ускорить следующим образом: существует другая машина M' , также вычисляющая f и такая, что*

²⁾ Вычислимая функция — это такая функция, которую можно вычислить хотя бы одним алгоритмом.

$$t_{M'}(n) \leq \log t_M(n)$$

для почти всех n .

Функция из теоремы Блюма — это некоторая экзотика (как следует из формулировки теоремы, время ее вычисления в наихудшем случае растет очень быстро для любой машины). Одна из «идеологических» задач в теории сложности вычислений — это строить теорию так, чтобы такие патологические явления в ней по возможности не возникали. Теорема Блюма доказывается построением с помощью техники диагонализации, и получающаяся функция не имеет никакого отношения к реальной практике вычислений или к остальной математике. Но тем не менее, поскольку мы строим математическую теорию, ничего не поделаешь — мы вынуждены признать, что выбранный подход не годится для ее построения. Нужно двигаться дальше.

5. СЛОЖНОСТНЫЕ КЛАССЫ

Итак, мы не можем надеяться на построение для каждой функции самой лучшей машины, вычисляющей эту функцию. Альтернативой является понятие *сложностного класса* — одно из центральных понятий для теории сложности вычислений.

Несколько вольная аналогия может быть здесь проведена с определениями интеграла по Риману и по Лебегу. Если мы не можем проинтегрировать по Риману, то мы меняем ось и начинаем суммировать по другой оси. В нашей ситуации, если мы не можем для функции сказать, что такое самая хорошая машина, то давайте также сменим ось. Рассмотрим множество всех машин, которые нас устраивают, и назовем класс всех функций, вычисляемых такими машинами, *классом сложности*. Пожалуй, проще всего дать сразу конкретный пример, чем долго рассуждать на эту тему:

$$\text{DTIME}(t(n)) \stackrel{\text{def}}{=} \{f \mid \exists M : (M \text{ вычисляет } f) \& (t_M(n) = O(t(n)))\}.$$

Это одно из центральных определений в теории сложности. Буква D означает детерминированные алгоритмы (бывают и другие), TIME означает ровно то, о чем вы подумали. Если у нас есть произвольная функция $t(n)$ от натурального аргумента, то мы образуем класс сложности, состоящий из таких функций, для которых существует вычисляющая f машина M , такая что сигнализирующая функция по времени ограничена исходной функцией $t(n)$ с точностью до мультипликативной константы. Приведенная выше теорема Блюма справедлива только для некоторых специальных функций. Но если мы хотим достичь ускорения, скажем, в 10 раз, то это можно сделать с любой функцией — например, путем увеличения количества машинных команд. Именно поэтому в правой части определения стоит O -большое.

Теперь определим один из самых важных сложностных классов

$$P = \bigcup_{k \geq 0} \text{DTIME}(n^k).$$

Класс P — это те функции, которые можно вычислить на наших машинах, и время их вычисления растет полиномиально с ростом длины слова. Он очень удобен и с практической, и с теоретической точки зрения. С практической точки зрения

это достаточно хорошая аппроксимация (бывают исключения, о которых будет сказано ниже) класса тех функций, которые поддаются вычислению за реальное время на реальных компьютерах. С математической точки зрения этот класс бесконечно удобен тем, что он замкнут относительно суперпозиции. Как мы увидим дальше, именно это обстоятельство позволяет построить теорию вычислимости для этого класса.

Бывают аналогичные классы языков, которые можно распознать за экспоненциальное время

$$\text{EXPTIME} = \bigcup_{k \geq 0} \text{DTIME}(2^{n^k}),$$

также можно определить двойное экспоненциальное время

$$\text{DOUBLEEXPTIME} = \bigcup_{k \geq 0} \text{DTIME}(2^{2^{n^k}})$$

и так далее.

6. ТЕОРЕМА ОБ ИЕРАРХИИ И СЛОЖНОСТЬ ЭЛЕМЕНТАРНОЙ ГЕОМЕТРИИ

Теперь давайте вернемся к нашему математику М. Оказывается, время работы алгоритма Тарского, с помощью которого М пытался решать задачи элементарной геометрии, не лежит ни в одном из приведенных выше классов. Для него есть такая нижняя оценка

$$t_{\text{Tarski}}(n) \geq 2^{2^{\dots^2}} \} \varepsilon n.$$

Как вы помните, для разрешимости алгебры Тарского продавался еще и диск *Collins for Windows 95*.

ТЕОРЕМА (КОЛЛИНЗ). *Алгебра Тарского принадлежит DOUBLEEXPTIME.*

Теперь М смог понять разницу между этими CD-ROM: время работы алгоритма Коллинза неизмеримо меньше времени работы алгоритма Тарского (хотя оно все равно может быть очень велико — оценка времени работы алгоритма двойной экспонентой не гарантирует нам, что даже теорема о сумме углов треугольника будет доказана за время существования Вселенной).

Возникает вопрос: а нельзя ли еще улучшить алгоритм разрешимости для алгебры Тарского? И более общий вопрос: а нельзя ли то же самое сделать вообще с любым алгоритмом? Вдруг, например, всякая вычислимая функция лежит в классе P. Или, по крайней мере, всякая функция из DOUBLEEXPTIME лежит в P.

Другими словами, вопрос в том, есть ли у нас предмет исследования, или, возможно, теорема Блюма об ускорении применима вообще ко *всем* функциям.

Второй краеугольный камень теории сложности вычислений — это *теорема об иерархии*.

Как и в случае с теоремой об ускорении, мы приводим ее далеко не в самой общей форме.

ТЕОРЕМА (ХАРТМАНИС (1965)). $P \neq \text{EXPTIME}$.

Таким образом, не все сложностные классы совпадают, так что предмет исследования есть.

Я не могу отказать себе в удовольствии привести почти полное доказательство этой теоремы. Если вы еще помните нашего математика М, второй момент его злключения состоял в том, что он спрашивал, остановится ли его программа *когда-нибудь*. Сейчас, наученный горьким опытом, он задал такой вопрос:

остановится ли эта программа до Нового Года?

Выясняется, что если до Нового Года осталось в точности экспоненциальное время, то эта задача и отделяет EXPTIME от P. Потому что есть очень простой алгоритм, позволяющий проверить, остановится ли программа до Нового Года, а именно, нужно просто подождать до него, и все само собой решится. Теорема об иерархии утверждает, что эту задачу нельзя решить существенно быстрее, чем описанным выше способом.

Я, конечно, слегка утрирую, есть там и кое-какие технические детали, но смысл состоит именно в этом.

И теперь наш математик оказался вполне подготовлен к восприятию следующей теоремы.

ТЕОРЕМА (ФИШЕР – РАБИН (1974)). *Алгебра Тарского не принадлежит классу P. Время работы любого разрешающего алгоритма для алгебры Тарского не меньше, чем $2^{\epsilon n}$, где ϵ — некоторая константа.*

Такая высокая *нижняя оценка* объясняет причины неудачи нашего математика с практическим решением задач в алгебре Тарского.

Наиболее сложная и, по-видимому, наиболее важная область теории вычислений, связана как раз с доказательством нижних оценок. В англоязычной литературе тот раздел теории сложности, который занимается конструированием алгоритмов, так и называется — “Theory of Algorithms”, а собственно под “Complexity Theory” подразумевается как раз построение нижних оценок. Таким образом, теория сложности пытается доказать, что не существует эффективных алгоритмов.

Уже на этом примере видны две трудности, стоящие перед людьми, которые пытаются доказывать нижние оценки. Смотрите, была теорема Тарского, ее пытались улучшить, но ничего не получалось. Естественно было предположить, что эта оценка оптимальна. После этого придумали алгоритм, основанный на совсем других идеях. Алгоритмов бывает много, алгоритмы бывают сложные, алгоритмы бывают разные. А ведь мы пытаемся доказать, что более эффективный алгоритм не появится никогда, и при этом пытаемся анализировать свойства алгоритмов из достаточно широкого класса.

7. СВОДИМОСТЬ И ПОЛНОТА

Давайте продвинемся дальше в построении нашей теории. Понятие сложностного класса было введено вовсе не для того, чтобы экономить на формулировках теорем (в конце концов, теорему Фишера – Рабина можно сформулировать и без упоминания каких бы то ни было сложностных классов – для этого достаточно оставить в ее формулировке лишь вторую фразу). Понятие сложностного класса становится важным в тот момент, когда у нас появляется понятие *сводимости*,

и это второе центральное понятие современной теории сложности вычислений. Есть несколько вариантов определения сводимости, я расскажу лишь о наиболее важном из них.

Сводимость по Карпу. Эта сводимость устроена очень просто. Для начала заметим, что мы занимаемся вычислением функций, отображающих конечные слова в конечные слова. Но во многих случаях оказывается гораздо удобнее, и это, как правило, не приводит к потере общности, рассматривать так называемые *языки*. Язык L можно рассматривать как множество слов $L \subseteq \{0, 1\}^*$ или интерпретировать его как отображение вида $\varphi : \{0, 1\}^* \rightarrow \{0, 1\}$, тогда $L = \varphi^{-1}(1)$. Переход к языкам не слишком нас ограничивает: если есть произвольная функция из слов в слова, то с ней можно связать такой язык — множество пар (x, i) таких, что i -й бит $f(x)$ равен 1.

ОПРЕДЕЛЕНИЕ. Язык L_1 сводится к языку L_2 по Карпу (обозначается это таким образом $L_1 \leq_p L_2$), если существует такая функция f из P , что

$$\forall x : (x \in L_1 \equiv f(x) \in L_2).$$

Сводимость можно рассматривать как использование некоторой подпрограммы, перерабатывающей исходное слово x в $f(x)$, к которому затем нужно применить алгоритм распознавания принадлежности языку L_2 . В этом варианте понятия сводимости обращение к программе распознавания L_2 осуществляется ровно один раз, если же разрешено использовать в качестве подпрограммы распознавание языка L_2 сколько угодно раз, то получается другая сводимость (по Тьюрингу). Сейчас это различие нам неважно.

Отношение сводимости — это отношение предпорядка. Оно рефлексивно и транзитивно. Самое главное свойство состоит в том, что если $L_2 \in P$, то и $L_1 \in P$. Здесь важно, что класс полиномов замкнут относительно операции суперпозиции.

Например, EXP-сводимость не получится. Не будет замкнутости относительно такой сводимости. Замкнется все лишь на конечных башнях экспонент — классе *элементарных рекурсивных функций*.

Если говорить только о естественных классах, в которых реально есть хотя бы одна разумная функция и которые замкнуты относительно суперпозиции, то есть еще класс квазиполиномов $2^{(\log n)^{O(1)}}$ и класс квазилинейных функций $n \log^{O(1)} n$. Квазилинейными функциями сейчас начинают интенсивно интересоваться, потому что считается, что по-настоящему хороший алгоритм — это алгоритм, который работает как раз квазилинейное время.

В любом случае, класс P является центральным в теории, и мы сейчас будем рассматривать именно полиномиальную сводимость.

Легко видеть, что класс EXPTIME замкнут относительно этой сводимости. Поэтому естественно поставить вопрос, а существуют ли в этом классе самые сложные языки, т. е. такие, что любой другой язык из этого класса к ним сводится. В этом случае можно решить любую задачу из EXPTIME, используя произвольный алгоритм распознавания для такого самого сложного языка и полиномиальную сводимость. Языки из некоторого сложностного класса, к которым сводится любой язык из этого класса, называются *полными* (относительно данного класса и данного типа сводимости). А если мы опустим требование принадлежности нашему классу самого языка, получим определение *трудного* языка.

Теорема Фишера – Рабина доказывается именно так. Вместо того, чтобы доказывать напрямую, что не существует полиномиального алгоритма для алгебры Тарского, доказывается, что алгебра Тарского трудна для класса EXPTIME. Поэтому полиномиальный алгоритм для алгебры Тарского давал бы полиномиальный алгоритм и для всех остальных задач из этого класса. А мы знаем, что в EXPTIME есть задачи, которые нельзя решить за полиномиальное время («остановка программы до Нового Года»).

Такой способ рассуждений типичен для теории сложности. Мы не пытаемся действовать «в лоб», а сводим одну задачу к другой. И, естественно, ситуация наиболее хороша, когда к данной задаче сводится много других задач из данного класса.

Успех этой науки определяется тем практическим обстоятельством, что полных задач оказывается много в самых разных ситуациях и они более естественны, чем проблема остановки.

8. ВСЯКИЙ ЛИ ПОЛИНОМИАЛЬНЫЙ АЛГОРИТМ ХОРОШ?

Теперь самое время вернуться к нашему бедному математику М и поговорить об исключениях из правила «класс P = класс эффективно вычислимых функций». Как-то М понадобилось для каких-то своих целей решать системы линейных неравенств:

$$\sum a_{ij}x_j \leq b_i. \quad (i = 1 \dots m, j = 1 \dots n)$$

Другими словами, ему потребовался пакет программ для задач линейного программирования. Специалисты по теории сложности вычислений перед покупкой программного обеспечения изучают литературу, и М, наученный предыдущим горьким опытом, стал следовать этому правилу. Он, конечно же, слышал про симплекс-метод, которым пользуются для решения задачи линейного программирования почти всюду, включая военную кафедру Московского университета. Но М обнаружил статью, в которой доказывается, что симплекс-метод не полиномиален. А после этого наш М обнаружил и статью Хачияна (1979), в которой был построен полиномиальный алгоритм для решения задачи линейного программирования. Поэтому, приехав на Митинский радиорынок, он попытался найти что-то вроде CD Khachiyana for Windows 95. К его удивлению, ничего похожего не было. Всё, что ему предлагали, было основано на симплекс-методе и его вариациях. Оказывается, что хотя теоретически алгоритм решения задачи линейного программирования симплекс-методом экспоненциальный, а алгоритм Хачияна полиномиальный, на практике первый работает быстрее второго. Чтобы построить пример задачи линейного программирования, на которой симплекс-метод будет работать долго, нужно очень и очень постараться, а полиномиальный алгоритм Хачияна работает примерно одинаково на любом входе (показатель экспоненты 6, что весьма плохо).

Это самое известное исключение из того правила, что полиномиальные алгоритмы хороши, а экспоненциальные — плохи. Но это исключение, на самом деле, подтверждает правило, потому что, хотя алгоритмом Хачияна никто и не пользуется для решения задач линейного программирования, выяснилось, что с

помощью этого алгоритма можно решать такие задачи, к которым с симплекс-методом подступиться в принципе невозможно. Например, пусть у нас есть произвольное выпуклое тело K и задано некоторое направление. Нужно максимизировать значение соответствующей линейной формы на K . Про выпуклое тело вы ничего не знаете: оно задано с помощью черного ящика (или, как иногда говорят, *оракула*). То есть, если вы указываете точку p , то с некоторой погрешностью ε можно сказать, лежит ли точка p в теле K , а если не лежит — получить некоторую отделяющую p от K гиперплоскость. Ответ нужно получить с той же точностью. Никакой симплекс-метод здесь, понятное дело, не работает — вообще нет никаких вершин, перебором которых занимается симплекс-метод. А алгоритм Хачияна и та наука, которая из него выросла, замечательно (т. е. полиномиально) с такими задачами справляется. Роль параметра n , описывающего размер входа, здесь играет $d \cdot (\log \varepsilon^{-1})$, где d — размерность, а ε — точность вычисления.

Таких исключений известно очень мало. Второй известный пример — проверка простоты числа. Как правило, подтверждается тезис, что если у вас имеется алгоритм, который теоретически работает хорошо, и это — алгоритм для нормальной задачи, которая и в самом деле откуда-то возникла, а не сконструирована с хулиганскими намерениями, то этот алгоритм будет хорошо работать и на практике. В частности, показатель k в оценке скорости работы алгоритма n^k обычно оказывается небольшим, а когда он поначалу велик, его можно уменьшить различными ухищрениями. Для подавляющего большинства естественных задач показатель не превосходит 3.

9. НЕДЕТЕРМИНИРОВАННЫЕ ВЫЧИСЛЕНИЯ

По-видимому, те, кто пришел на эту лекцию, хотели узнать что-нибудь про самую известную открытую проблему в этой области $P \stackrel{?}{\neq} NP$. Что такое P , я уже немножко объяснил. Теперь давайте займемся NP .

Для этого опять вернемся к M . Пока он разбирался в теории сложности, его сын поступил в университет на первый курс и стал изучать математическую логику. Изучение математической логики, как известно, начинается с такой знаменитой науки, как исчисление высказываний. У вас имеется некоторая пропозициональная формула $\Phi(p_1, \dots, p_n)$, она называется тавтологией, если она истинна независимо от того, что мы в нее подставим. И одно из неприятных упражнений при занятиях этой наукой состоит в том, что нужно выяснить по выписанной пропозициональной формуле, является она тавтологией или нет. Именно с таким вопросом и обратился сын нашего математика к своему папе. M , естественно, уже никуда не поехал, а стал, как и поступают всегда специалисты по теории сложности, пытаться поместить задачу в один из уже известных классов сложности. Чтобы использовать стандартные обозначения, будем говорить о двойственной ей задаче SAT — выполнимости³⁾: есть ли хотя бы один набор переменных, при котором формула истинна.

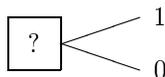
Накладывая эту задачу на нашу картину классов сложности, M увидел первым делом, что $SAT \in EXPTIME$. Алгоритм решения задачи SAT за экспоненциальное время очевиден — всего возможных наборов значений переменных

³⁾ SAT — от слова satisfiability — выполнимость.

2^n , а время вычисления значения формулы при заданных значениях переменных полиномиально. Следующий шаг — классифицировать эту задачу: лежит ли она в P , или полна в $EXPTIME$? Этот вопрос, возникший в начале 70-х годов, до сих пор открыт. Почему у специалистов, потративших на решение этой задачи почти 30 лет, не получается построение полиномиального алгоритма, объяснить сложно (я, во всяком случае, никак это объяснять не берусь). Но тот факт, что не получается доказательство полноты этой задачи в $EXPTIME$, некоторому объяснению поддается. Если мы посмотрим на настоящие экспоненциальные алгоритмы, например, на алгоритм для аналога алгебры Тарского над комплексными числами, и сравним с ними то детское рассуждение, которое мы привели выше, то сразу невооруженным глазом видна разница — приведенный алгоритм для *SAT* использует экспоненциальное время очень слабо — только для перебора экспоненциального числа возможностей, с каждой из которых он справляется за полиномиальное время.

Взрыв в теории сложности вычислений начался с того, что было сформулировано определение класса NP , как класса языков, которые распознаются переборными алгоритмами. Более научно, переборные алгоритмы называют недетерминированными алгоритмами. NP — это класс языков, которые можно распознать за недетерминированное полиномиальное время.

Теперь дадим определение этого класса. В этом определении, как и в определении класса P , будет фигурировать слово «машина», сокращенно мы будем обозначать ее НМТ («недетерминированная машина Тьюринга»); впрочем, использование слова «машина» в этом контексте может вызывать у людей, имеющих отношение к реальным машинам, некоторое чувство неудовольствия. Под недетерминированной машиной мы понимаем такую машину, которая работает как обычно, но в какой-то момент она может нарисовать в какой-то ячейке знак вопроса и после этого ее работа разделяется на две ветви 0 или 1:



(в клетке окажется записанным либо 0, либо 1). После этого машина продолжает работу. В какой-то момент она может раздвоиться еще раз. Появляется дерево вычислений. Вдоль каждой ветви дерева вычислений НМТ работает как обычное вычислительное устройство, но результат работы НМТ зависит от результатов работы вдоль всех ветвей. Определим результат работы НМТ в случае проверки принадлежности слова x к языку L . На каждой из ветвей вычисления получается один из двух возможных ответов: «да» или «нет». НМТ распознаёт язык L , если всякое слово x принадлежит L тогда и только тогда, когда *хотя бы на одной* ветке вычислений получен ответ «да».

Проиллюстрируем возможности НМТ на примере с упражнениями на проверку выполнимости пропозициональных формул. Частенько, когда студент (полиномиальная детерминированная машина Тьюринга) проваливается и не может решить задачу, преподаватель выступает в роли недетерминированной машины — он показывает (загодя приготовленный) ответ, который легко может быть проверен. Другими словами можно сказать так: НМТ «стремится» доказать утверждение $x \in L$, а в момент раздвоения она обладает неограниченными

интеллектуальными возможностями и выбирает наилучший вариант. Если существует ветка вычислений, при которых ответ «да», то $x \in L$. В противном случае никакое ветвление не приведет к положительному результату (слово не принадлежит языку лишь тогда, когда у НМТ нет никакой возможности доказать обратное).

Не спрашивайте меня о физике процесса: как происходит такое раздвоение, где находится такая машина, можно ли ее посмотреть... Таких машин в реальности нет. Одним из самых значительных достижений в теории сложности за последние годы стало то, что была сформулирована квантовая модель вычислений. Квантовые компьютеры являются одним из кандидатов на роль такой машины в реальном мире. По крайней мере, существование квантовых компьютеров не вызывает принципиальных возражений у физиков, а надежда на имитацию недетерминированных машин квантовыми компьютерами пока не вызывает категорических возражений со стороны специалистов в теории сложности. Более того, абстрактные квантовые компьютеры уже умеют решать некоторые важнейшие переборные задачи, такие, как факторизация чисел⁴⁾.

Подчеркнем еще раз, что недетерминированная машина является чисто теоретическим понятием. Удобства от введения такого понятия получаются совершенно фантастические.

Это самая первая вымышленная модель, появившаяся в теории сложности вычислений. В настоящий момент есть масса таких моделей, гораздо более сложных: интерактивные доказательства и т. п. Эти модели появляются не сами по себе, а с целью определения сложных классов и для классификации естественных задач.

Интуитивно ясно, что понятие НМТ прекрасно приспособлено к моделированию переборных алгоритмов. Собственно, уже показано, как это делать. Если у нас есть алгоритм, который пытается перебрать некоторое количество возможностей, то наша машина может угадать, какая из возможностей является хорошей, а потом имитировать вторую часть (полиномиальную проверку конкретного варианта). В обратную сторону это несколько сложнее, но все равно достаточно просто. Если у нас есть НМТ, то она порождает дерево вычисления, и перебор нужно производить по всем возможным ветвям.

Современная теория сложности вычислений началась с результатов Кука, Карпа, Левина (который получил их независимо) в начале 70-х годов (1970–1972). Эта серия теорем состоит вот в чем.

1. Задача *выполнимость* (*SAT*) полна для класса NP (теорема Кука). Таким образом, задача нашего математика M (можно ли придумать полиномиальный алгоритм для *SAT*) эквивалентна вопросу: совпадают ли классы P и NP ($P \stackrel{?}{\neq} NP$). Если полиномиального алгоритма нет, то эти классы не совпадают — их отделяет задача *SAT*, если же он есть, то можно эффективно решить любую задачу из класса NP. Задача выполнимости отвечает за класс NP. Грубо говоря, NP — это не что иное, как задачи, которые могут быть сведены к выполнимости пропозициональных формул.

2. Естественно, что такой результат может вызвать некоторую реакцию отторжения — не настолько уж важна задача выполнимости, чтобы строить целую

⁴⁾ На последнем Международном Конгрессе (Берлин, август 1998г.) американский математик П. Шор был удостоен за эти исследования премии им. Неванлинны.

теорию ее решения. Но следующим шагом была статья Карпа (1971), в которой уже была указана 21 полная задача для класса NP. Все они между собой эквивалентны.

В настоящее время список NP-полных задач, которые возникают буквально во всех областях математики, содержит тысячи задач. Везде, где возникают алгоритмы, возникают и переборные задачи. Это неудивительно, потому что большая часть программистской работы и состоит в выборе варианта получше. Удивительно то, что часто, а практически — всегда (с некоторыми исключениями) — если у вас есть конкретная переборная задача, то ее можно сравнительно легко классифицировать — либо для нее есть полиномиальный алгоритм, либо она полна.

Таким образом строится теория переборных задач. Из-за того, что переборные алгоритмы возникают почти всюду, она приобрела большое значение. По мнению американского тополога Смейла, вопрос $P \stackrel{?}{\neq} NP$ будет одним из самых важных вопросов математической науки следующего столетия.

Вот и подошел к концу рассказ о том, что является, так сказать, ядром теории сложности вычислений. Не стоит понимать этот рассказ так, что вся теория сложности занимается только лишь соотношением $P \stackrel{?}{\neq} NP$. Изложенная схема исследований — которая объединяет задачи в сложностные классы и потом исследует эти задачи с помощью их сводимости друг к другу — оказалась удивительно эффективна плодотворна в самых разных ситуациях.

Перечислим коротко несколько самых важных возможностей (помимо уже упомянутых выше квантовых вычислений).

Бывает сложность алгебраическая. Если нас интересует вычисление некоторого полинома, и мы абстрагируемся не только от деталей битового вычисления, но и от того, как производятся арифметические операции, то все зависит лишь от того, с какими числами производятся арифметические операции. Тогда длина вычисления определяется количеством сложений и умножений.

Бывает сложность геометрическая — диаграммы Вороного и близкие к ним вещи, о которых я говорить не буду.

Бывает сложность булева — она отличается от того, чем мы занимались сегодня, тем, что нас интересуют функции $f : \{0, 1\}^n \rightarrow \{0, 1\}$, определенные на словах фиксированной длины.

В любом из этих разделов возникает масса постановок интереснейших задач именно в рамках этой общей идеологии.

БЛАГОДАРНОСТИ

Я выражаю глубокую благодарность М. Н. Вялому и М. В. Алехновичу за квалифицированную помощь в приведении данной лекции к пригодному для чтения виду и проф. Ј. Крајісек за ряд ценных замечаний.

По-новому о старом: фрагменты классической математики

Теорема Жордана

А. В. Чернавский

I. СВОЙСТВО «РАЗБИВАТЬ» И ДРУГИЕ ТОПОЛОГИЧЕСКИЕ СВОЙСТВА

1. ЦЕЛОЕ НЕ ЕСТЬ СУММА СВОИХ ЧАСТЕЙ. Это набившее оскомину высказывание известный кибернетик М. Минский освежил такой иллюстрацией: никакая часть ограды не есть ограда.

На языке, которым мы будем пользоваться, это значит: топологический образ окружности разбивает плоскость, но никакая его часть плоскости не разбивает.

2. ЧТО ЗНАЧИТ «РАЗБИВАТЬ»? Термин «разбивать» связан с одним из двух основных топологических свойств пространств. Мы не будем прибегать к общему понятию топологического пространства, чтобы не отклоняться далеко от нашей темы, и ограничимся подмножествами евклидовых пространств \mathbb{R}^n . Собственно, нам нужна плоскость, но мы по мере возможности будем касаться и общего случая топологического образа сферы S^{n-1} в \mathbb{R}^n . Плоскость будем обозначать Π .

Подмножество *разбивает* связное множество, если дополнение к нему не связно.

3. СВЯЗНОСТЬ. Напомним, что связность $X \subset \mathbb{R}^n$, по определению означает, что X нельзя представить объединением двух непустых и не пересекающихся подмножеств, которые *открыты в X* . Подмножество $A \subset X$ открыто в X , если оно есть пересечение X с открытым подмножеством \mathbb{R}^n , а открытость в \mathbb{R}^n означает, что каждая точка подмножества \mathbb{R}^n лежит в этом подмножестве вместе со своей окрестностью, например, с шаром с центром в этой точке.

Открытое подмножество \mathbb{R}^n связно тогда и только тогда, когда оно *линейно связно*, т. е. любые две его точки соединимы лежащей в нем *простой дугой*. В качестве такой дуги можно взять ломаную.

Свойство связности и понятие «разбивать» являются топологическими и потому позволяют доказывать в простых случаях топологическую неэквивалентность множеств. С помощью аксиомы Дедекинда легко проверяется, что связными подмножествами числовой оси \mathbb{R} являются в точности интервалы (включая точки и неограниченные интервалы — полупрямые и саму \mathbb{R}).

Открытые интервалы топологически отличаются от замкнутых и полуоткрытых тем, что каждая их точка разбивает любую их связную окрестность. Окружность отличается от любого интервала тем, что ни одна ее точка ее не разбивает. Две точки окружность разбивают. Это можно назвать одномерным случаем теоремы Жордана. Прямая не гомеоморфна плоскости, так как каждая ее точка разбивает каждую свою связную окрестность, а точки плоскости, очевидно, не разбивают кругов, центрами которых они служат.

Напомним, что компонентой связности (или просто компонентой) множества называется любое его связное подмножество, не лежащее в большем связном подмножестве. Всякое множество распадается в сумму попарно непересекающихся своих компонент связности.

4. КОМПАКТНОСТЬ. Второе важное топологическое свойство — компактность, оно определяется с помощью открытых покрытий. Покрытие множества — это система его подмножеств, дающая в объединении все это множество. Открытое покрытие множества X состоит из открытых в X подмножеств. Множество компактно, если из каждого его открытого покрытия можно выделить конечную подсистему, являющуюся покрытием. Из интервалов компактны только замкнутые интервалы (и точки). Это — известная лемма Гейне–Бореля–Лебега. Но на прямой имеется много других компактных подмножеств, например, знаменитое канторово множество.

Вообще, компактные подмножества \mathbb{R}^n это в точности подмножества, являющиеся одновременно замкнутыми и ограниченными. (Замкнутые подмножества — это дополнения к открытым. Они характеризуются тем, что содержат все свои предельные точки, т. е. такие, в любой окрестности которых имеются точки данного множества. Ограниченное множество — множество, содержащееся в шаре какого-нибудь радиуса.)

5. КОНТИНУУМЫ. Оба свойства — связность и компактность — выражают непрерывность множества (в отличие от непрерывности отображения). Они являются топологическими, т. е. сохраняются при гомеоморфизмах (взаимно однозначных и в обе стороны непрерывных отображениях). На самом деле они сохраняются при любом непрерывном отображении. Обратное, вообще говоря, неверно, в чем читатель может убедиться на простых примерах: разрывное отображение может сохранять связность и компактность.

Множество и связное, и компактное одновременно называется *континуумом*. Непрерывный образ континуума есть снова континуум. (Например, если функция определена на замкнутом интервале, то и область ее значений есть замкнутый интервал. Из этого вытекают теоремы Больцано и Вейерштрасса о непрерывных функциях.)

В частности, топологический образ C окружности S^1 на плоскости является континуумом. Мы будем называть такой образ простой замкнутой кривой. То же верно и для топологического образа C^{n-1} сферы S^{n-1} в \mathbb{R}^n (такой образ будем называть топологической сферой).

6. ТЕОРЕМЫ ЖОРДАНА, ЖОРДАНА-БРАУЭРА И ШЕНФЛИСА. В своей основной части теорема Жордана утверждает, что простая замкнутая кривая C на плоскости \mathbb{P} разбивает \mathbb{P} .

Более того, $\mathbb{P} \setminus C$ состоит из двух компонент и C является общей границей этих компонент. Последнее означает, что каждая точка C служит предельной точкой для каждой из этих компонент.

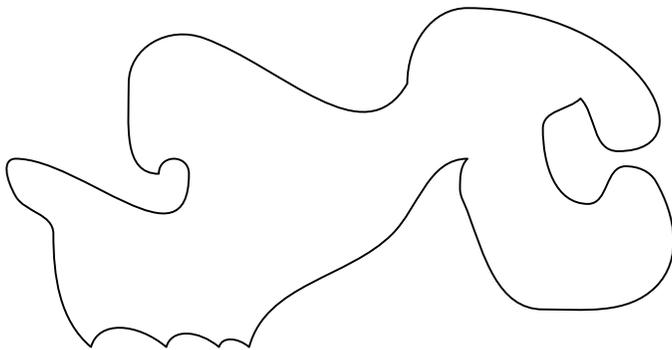


Рис. 1.

Это же утверждение справедливо и для обобщения теоремы Жордана, доказанного Брауэром в 1912 году: $\mathbb{R}^n \setminus C^{n-1}$ состоит из двух компонент и C^{n-1} является их общей границей.

Но следующее уточнение теоремы Жордана, данное А. Шенфлисом в 1906 г., не справедливо в общем случае при $n \geq 3$:

Замыкания компонент $\mathbb{R}^2 \setminus C$ гомеоморфны замыканиям компонент дополнения к S^1 . (Это эквивалентно утверждению, что гомеоморфизм S^1 на C может быть продолжен до гомеоморфизма плоскости на себя.)

Мы не будем доказывать здесь это уточнение и отошлем читателя к книге Л. В. Келдыш «Топологические вложения в евклидово пространство». Но в конце мы приведем несложный пример (Е. Артина и Р. Фокса, подробнее описанный в этой книге), из которого будет видно, почему этот результат не переносится на более высокие размерности.

7. РЕТРАКЦИИ. Мы будем опираться на понятие ретракции. Оно имеет общематематическое употребление (теперь это значит «общекатегорное»). В нашем случае подмножество A множества X называется *ретрактом* X , если имеется непрерывное отображение $r : X \rightarrow A$, при котором все точки A неподвижны.

(В более общей форме: пусть даны отображения $f : B \rightarrow X$ и $g : X \rightarrow B$ с условием $gf = 1|_B$ — тождественное отображение B . Тогда отображение $r = fg : X \rightarrow X$ обладает свойством $r^2 = rr = fgfg = fg = r$. Более того, если $x = r(y)$, т. е. x лежит в образе A отображения r , то $fg(x) = r(x) = r(r(y)) = r(y) = x$. Значит, r есть ретракция X на A и также f есть отображение, обратное к g , ограниченному на A . Иными словами, f есть гомеоморфизм B на A .)

Например, ретракцией является проекция плоскости на прямую. Нам понадобится ретракция следующего типа. В шаре B^n выкинем произвольную точку O

(не обязательно центр). Каждый полуинтервал $(O, x]$, где $x \in S^{n-1}$, отображим на свой конец x . Мы, очевидно, получим ретракцию $B \setminus O$ на границу шара S^{n-1} .

8. **ОКРУЖНОСТЬ — АБСОЛЮТНЫЙ ОКРЕСТНОСТНЫЙ РЕТРАКТ.** Для нас существенным будет свойство окружности, состоящее в ее принадлежности к классу *абсолютных окрестностных ретрактов*. Не давая общего определения, мы отметим только, что согласно этому свойству, любой гомеоморфный образ C в любом пространстве \mathbb{R}^n является ретрактом некоторой своей окрестности. Это же верно для всех сфер. Мы покажем это для гомеоморфных образов S^1 в \mathbb{R}^2 .

9. **ОТРЕЗОК — АБСОЛЮТНЫЙ РЕТРАКТ.** Замкнутый интервал является *абсолютным ретрактом*, что, в частности, означает, что любая простая дуга (гомеоморфный образ отрезка) в плоскости является ее ретрактом и, значит, любого подмножества плоскости, содержащего эту дугу.

Абсолютные ретракты обладают также двойственным свойством: непрерывное отображение подмножества множества X в такое пространство может быть распространено на все X . Для отрезка (и, значит, для любой простой дуги) это есть утверждение леммы Урысона, с которой начинается гомотопическая топология. Мы докажем ее только в очень частном, нужном нам дальше, случае.

ЛЕММА. *Непрерывное отображение f границы квадрата Q в отрезок I может быть продолжено до непрерывного отображения $F : Q \rightarrow I$.*

ДОКАЗАТЕЛЬСТВО. Возьмем внутри Q точку O и отображим ее в любую точку A отрезка. Затем каждый отрезок $[O, x]$, где x лежит на границе Q , линейно отображим на отрезок $[A, f(x)]$. То, что получится непрерывное отображение, продолжающее f , проверяется непосредственно. \square

10. **ТЕОРЕМА О НЕРЕТРАГИРУЕМОСТИ КРУГА НА ГРАНИЦУ И ТЕОРЕМА О НЕПОДВИЖНОЙ ТОЧКЕ.** Наше доказательство опирается на теорему о невозможности ретракции круга на его граничную окружность. Эта теорема справедлива для шара любой размерности. Мы приведем ниже простое доказательство для круга. Оно может быть «продолжено» по индукции на высшие размерности, но при этом обрастает техническими деталями и оказывается ненамного проще обычных доказательств с помощью симплициальной или гладкой аппроксимации.

Как известно, эта теорема эквивалентна теореме о существовании неподвижной точки у любого непрерывного отображения шара в себя. Нам этот факт не потребуется и мы оставим доказательство этой эквивалентности читателю в качестве упражнения.

11. **СХЕМА ДОКАЗАТЕЛЬСТВА ТЕОРЕМЫ ЖОРДАНА И СОПУТСТВУЮЩИХ УТВЕРЖДЕНИЙ.** Если $\mathbb{P} \setminus C$ связно, то можно построить ретракцию плоскости на C , значит, и содержащего C круга, а тогда и ретракцию этого круга на его границу. Мы полностью проводим доказательство только в двумерном случае, но будет вполне очевидно, что оно прямо переносится и на многомерный случай.

Неретрагируемость круга на границу влечет за собой более общее утверждение: замыкание ограниченной компоненты дополнения к компактному множеству на плоскости не ретрагируется на свою границу. Отсюда, в частности, вытекает, что простая дуга не разбивает плоскости, а также, что гомеоморфный образ C окружности является общей границей всех компонент своего дополнения.

Наконец, мы приведем доказательство утверждения, что компонент дополнения к S ровно две. Оно оказывается более сложным и в отличие от предыдущих существенно использует свойства Π , хотя, утверждение, конечно, верно и в n -мерной ситуации.

Для полноты мы выведем из теоремы Жордана-Брауера другие классические результаты Брауера: теорему об инвариантности области и размерности \mathbb{R}^n — и сделаем краткий исторический экскурс.

II. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ЖОРДАНА

Начнем с доказательства теоремы о неретрагируемости круга B на его границу S .

1. ТЕОРЕМА. *Не существует ретракции B на S .*

ДОКАЗАТЕЛЬСТВО. Пусть $r : B \rightarrow S$ — ретракция. Возьмем точку $x_0 \in S$. Пусть K полный прообраз точки x_0 . Это компактное множество, причем $K \cap S = x_0$. Для каждой точки $x \in K$ возьмем малый треугольник Δ_x , содержащий x внутри себя. Треугольник Δ_{x_0} для точки x_0 возьмем так, чтобы его одна сторона лежала на S и содержала x_0 внутри себя. Треугольники для остальных точек не должны пересекать S . Выберем конечное число треугольников, объединение внутренностей которых содержит K . Немного пошевелив вершины,

можно добиться, чтобы вершины каждого треугольника не лежали на прямых, содержащих стороны других треугольников. Тогда граница объединения оставленных треугольников состоит из конечного числа ломаных, из которых только одна, назовем ее l , может быть незамкнутой — она должна соединять два конца стороны треугольника Δ_{x_0} , лежащей на S . В самом деле, конец такой ломаной не может лежать внутри B . (Отрезок не разбивает связной окрестности своего конца.)

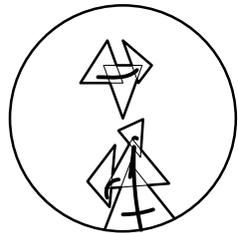


Рис. 2.

Если треугольники брались достаточно малыми, то расстояние точек l от K столь мало, что расстояние ее образа $r(l)$ от x_0 меньше заданного ε . Но этот образ есть связное множество, которое содержит оба конца указанной стороны Δ_{x_0} . Так как x_0 разбивает свою окрестность, этот образ должен содержать x_0 , т. е. l пересекает K , что невозможно по построению. \square

Это доказательство довольно непосредственно продолжается по индукции на высшие размерности. Мы укажем только схему этого рассуждения, так как подробное его проведение, как сказано, имеет примерно тот же порядок сложности, что и обычные доказательства.

Рассуждая от обратного, допустим, что имеется ретракция $r_n : Q^n \rightarrow S^{n-1}$ n -мерного куба на его край. Возьмем внутри $(n-1)$ -мерной грани $Q^{n-1} \subset S^{n-1}$ этого куба точку x_n . Пусть K_n — полный прообраз этой точки. Он пересекает S^{n-1} только по точке x_n . Покроем K_n конечным числом кубов, из которых только один пересекает S^{n-1} своей гранью, содержащей x_n внутри себя. Мы не будем доказывать элементарный факт комбинаторной топологии: некоторая симплициальная окрестность этого объединения кубов имеет границы

$(n - 1)$ -мерное многообразие. Такая окрестность называется регулярной, в следующем пункте мы рассмотрим элементарный случай регулярной окрестности дерева. Обозначим границу регулярной окрестности объединения построенных кубов через M^{n-1} . Ее край совпадает с краем S^{n-2} $(n - 1)$ -мерного куба, содержащего \mathbf{x}_n . Образ $r_n(M^{n-1})$ лежит в малой окрестности этого куба и не содержит \mathbf{x}_n . В силу этого имеется ретракция r'_n этого образа на S^{n-2} . Композиция $r'_n r_n$ определяет ретракцию $r_{n-1} : M^{n-1} \rightarrow S^{n-2}$.

Возьмем теперь точку \mathbf{x}_{n-1} в какой-либо грани S^{n-2} и обозначим ее полный прообраз через $K_{n-1} \subset M^{n-1}$. Далее повторяем проведенное только что построение с тем только отличием, что вместо кубов берем объединения симплексов, содержащих точки K_{n-1} . Мы можем построить регулярную окрестность этого прообраза так, что ее граница M^{n-2} (взятая относительно M^{n-1}) пересекает S^{n-2} по границе S^{n-3} куба, содержащего \mathbf{x}_{n-1} , и при этом ее образ не содержит этой точки. Поэтому, как и выше, строится ретракция $r_{n-2} : M^{n-2} \rightarrow S^{n-3}$ и рассуждение может быть продолжено дальше.

Так как размерность многообразий M^i уменьшается на единицу с каждым шагом, через n шагов мы придем к ретракции ломаной без пересечений на ее два конца, что невозможно в силу ее связности.

Докажем теперь два предварительных утверждения.

2. РЕГУЛЯРНАЯ ОКРЕСТНОСТЬ ДЕРЕВА. *Деревом* называется связный граф (объединение отрезков, пересекающихся только в концах), у которого нет циклов, т. е. в нем нет замкнутых цепочек ребер, в которых каждое ребро встречается не более одного раза. Эквивалентным свойством является невозможность удаления одного (открытого) ребра так, чтобы граф остался связным. Еще иначе можно сказать, что дерево — это граф, каждая точка которого или его разбивает или является конечной только для одного ребра.

Дерево, очевидно, имеет ребро, один из концов которого принадлежит только этому ребру, причем после отбрасывания этого ребра снова останется дерево. Иначе, смещаясь по ребрам, мы обязательно получим цикл.

Последовательно отбрасывая ребра, мы в конце концов оставим только одну вершину. Возьмем многоугольник вокруг этой точки и, присоединяя ребра в обратном порядке, будем каждый раз добавлять содержащий очередное ребро прямоугольник, граничащий с построенной на предыдущем шаге фигурой по отрезку, вместе с многоугольником вокруг другой вершины этого ребра (см. рис. 3).

Достаточно очевидно, что каждый раз мы будем получать фигуру, гомеоморфную кругу. Таким образом, дерево на плоскости имеет сколь угодно тесную окрестность, гомеоморфную кругу.

Построенная окрестность называется регулярной. Индукцией по числу ребер легко строится ретракция такой окрестности на дерево, причем ясно, что если окрестность достаточно тесная, то и ретракцию можно сделать сдвигающей точки достаточно мало.

3. ОКРУЖНОСТЬ — ОКРЕСТНОСТНЫЙ РЕТРАКТ ПЛОСКОСТИ.

ПРЕДЛОЖЕНИЕ. Пусть дан гомеоморфизм $h : S \rightarrow C$ окружности S на подмножество C плоскости Π . Тогда для некоторой окрестности $U(C)$ в плоскости имеется ретракция $U(C)$ на C .

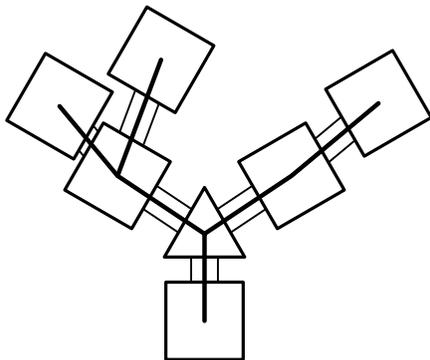


Рис. 3.

ДОКАЗАТЕЛЬСТВО. Рассмотрим последовательность Z_i разбиений плоскости на квадраты (как в тетрадке «в клетку») со стороной 2^{-i} . Они строятся с помощью пар семейств прямых: прямые в одном семействе отстоят друг от друга на расстояния кратные 2^{-i} , а прямые разных семейств ортогональны. Кроме того, разбиение Z_{i+1} получается добавлением новых прямых к прямым разбиения Z_i .

Каждой вершине v каждого квадрата каждого разбиения поставим в соответствие одну из ближайших точек C , которую обозначим x_v . Заметим, что если мы выберем последовательность v_i этих вершин, сходящуюся к некоторой точке x_0 на C , то и последовательность точек x_{v_i} будет сходиться к этой же точке.

Рассмотрим теперь ребра квадратов, не пересекающих C . Сначала каждое такое ребро $[v', v'']$ подразделения Z_1 гомеоморфно отобразим на ту из двух дуг, на которые C разбито точками $x_{v'}$ и $x_{v''}$, которая имеет меньший диаметр.

Затем для подразделения Z_2 возьмем те квадраты, которые не пересекают C и не лежат в отобранных на первом шаге квадратах и поступим с их ребрами (не лежащими на уже отображенных) точно так же.

Продолжим это построение на все Z_i , выбирая на очередном шаге те квадраты, которые не пересекают C и не лежат в выбранных ранее. Заметим снова, что если мы возьмем последовательность точек w_i на ребрах квадратов этих подразделений, которая сходится к точке на C , то и последовательность их образов на C будет сходиться к той же точке.

Мы фактически построили непрерывную ретракцию на C множества, полученного присоединением к C бесконечного графа, состоящего из ребер квадратов этих подразделений, номер которых растет по мере приближения к C . Каждая точка плоскости вне C принадлежит квадрату одного из подразделений Z_i , который не пересекает C и граница которого уже отображена в C .

Теперь мы постараемся распространить построенное отображение на такие квадраты. Вообще говоря, это невозможно. (Граница квадрата могла быть отображена в C с целым ненулевым числом обходов, но мы не хотим останавливаться на строгом определении этого понятия.)

Однако, если квадрат достаточно близок к C , продолжение может быть построено!

Подробнее. Для всякого $\varepsilon > 0$ найдется номер i такой, что границы квадратов подразделений Z_j для $j > i$, отобранных в нашем построении, отображены в подмножества C , диаметры которых меньше ε . Образы границ таких квадратов содержатся в C в гомеоморфных образах отрезка и, согласно п. 9 предыдущего раздела, мы можем продолжить отображение на эти квадраты.

Ясно также, что при этом диаметры образов квадратов стремятся к нулю при возрастании i .

Рассмотрим теперь все квадраты, для которых нами было определено отображение в C . Во-первых, ясно, что отображение их объединения непрерывно и непрерывно продолжается тождественным отображением C на себя, т. е. снова мы имеем ретракцию. Отображение определено, возможно, не для всех квадратов. Но объединение тех, для которых оно определено, содержит некоторую окрестность C в плоскости и поэтому мы в действительности построили ретракцию окрестности C на C , что и требовалось. \square

ЗАМЕЧАНИЕ. Мы построили ретракцию на C множества, полученного из C добавлением ее окрестности и границ некоторых квадратов подразделений Z_i . Мы используем ее дальше. Обозначим ее \bar{r} .

4. ОТРЕЗОК — АБСОЛЮТНЫЙ РЕТРАКТ ПЛОСКОСТИ. Для отрезка I справедливо более сильное утверждение, чем мы воспользуемся при доказательстве дополнений к теореме Жордана.

ЛЕММА. Если $q : I \rightarrow J \subset \Pi$ гомеоморфизм, то J — ретракт Π .

ДОКАЗАТЕЛЬСТВО ЛЕММЫ. Ретракция строится так же, как для C , но здесь не возникает того *препятствия*, которое возникало для гомеоморфного образа окружности в виде ненулевых степеней отображений границ квадратов на C . Иными словами, мы для каждого квадрата можем пользоваться леммой п. 9 предыдущего раздела. \square

5. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ЖОРДАНА. Мы построили выше ретракцию \bar{r} на C множества, полученного присоединением к C некоторой ее окрестности U и границ квадратов подразделений Z_i , причем каждая точка вне C лежит в одном из отобранных квадратов.

Возьмем большой квадрат Q , содержащий окрестность U целиком внутри себя. Можно считать, что это один из квадратов первого подразделения Z_1 . Квадраты подразделений Z_i либо лежат в нем целиком, либо не пересекают его внутренность.

Возьмем точку v_0 на границе Q . Занумеруем отобранные выше квадраты подразделений Z_i , обозначив их q_k , и возьмем в каждом из них по точке v_k . Если предположить, что C не разбивает Π , т. е. что дополнение к C имеет одну компоненту связности, то можно для каждой точки v_k построить ломаную, соединяющую с ней точку v_0 , которая проходит в Q и не пересекает C . Объединение этих ломаных образует граф и можно допустить, что этот граф — дерево (иначе выкинем из него несколько отрезков).

Возьмем регулярную окрестность R этого дерева, считая, что она не пересекает C , лежит в Q , притом пересекается с границей Q только по отрезку d , содержащему v_0 . R гомеоморфна кругу. Ее граница состоит из d и дополнительной дуги d' . Легко построить ретракцию R на d' . Продолжим эту ретракцию

тождественно на остальную часть квадрата Q и получим ретракцию r_1 квадрата на дополнение к R .

Это дополнение не содержит точек v_k . Для каждого из отобранных квадратов q_k построим ретракцию дополнения к точке v_k на его границу. Получим ретракцию объединения квадратов $q_k \setminus v_k$ на сумму их ребер. Дополним эту ретракцию тождеством на U и обозначим получившуюся ретракцию r_2 . Наконец, мы имеем ретракцию $r_3 = \bar{r}$, построенную ранее.

Композиция трех ретракций $r_3 r_2 r_1$ дает ретракцию $\hat{r} : Q \rightarrow C$. Возьмем теперь гомеоморфизм $q : S^1 \approx C$ и продолжим его до непрерывного отображения $\bar{q} : B \rightarrow Q$ круга, ограниченного S^1 , в квадрат Q (радиус $[Ox]$ линейно отображается на отрезок $[A, q(x)]$, где A — какая-нибудь точка в Q). Композиция $q^{-1} \hat{r} \bar{q}$ дает ретракцию круга B на его границу, что невозможно. Теорема доказана. \square

III. ДОКАЗАТЕЛЬСТВО СОПУТСТВУЮЩИХ ПРЕДЛОЖЕНИЙ

Мы докажем теперь три утверждения, дополняющих теорему Жордана: мы покажем, что простая замкнутая кривая на плоскости разбивает плоскость в точности на две компоненты и является общей границей обеих этих компонент своего дополнения, в то время как простая дуга плоскости не разбивает. Этого недостаточно, чтобы утверждать, что гомеоморфные образы окружности и отрезка на плоскости расположены с точки зрения топологии так же, как и стандартные окружность и отрезок. Эти утверждения справедливы, но требуют для доказательства дополнительной специальной техники, и мы отсылаем интересующегося читателя к книге Л. В. Келдыш.

1. ТЕОРЕМА 1. *Гомеоморфный образ J отрезка $I = [0, 1]$ в плоскости Π (т. е. простая дуга) не разбивает Π .*

Мы должны доказать связность дополнения к $J = q(I)$ в плоскости, где q — гомеоморфизм отрезка $I = [0, 1]$ с подмножеством плоскости J . Мы докажем сначала лемму, которая касается вообще компактных подмножеств плоскости.

2. ЛЕММА. *Если компактное множество $K \subset \Pi$ разбивает Π , то оно не является ретрактом замыкания ни одной из ограниченных компонент своего дополнения в Π .*

ДОКАЗАТЕЛЬСТВО ЛЕММЫ. Имеется только одна неограниченная компонента дополнения, содержащая дополнение к достаточно большому кругу, который содержит наше множество K (дополнение к кругу связно, а компоненты, лежащие в круге, ограничены). Возьмем большой квадрат Q , содержащий внутри себя K . Если U — одна из ограниченных компонент, то она также лежит внутри Q . Пусть r — ретракция замыкания U на свою границу (лежащую в K). Дополним ее тождеством вне U . Мы получим отображение Q в себя, тождественное на границе и такое, что точки U не лежат в образе. Существует ретракция Q с одной выкинутой точкой (любой точкой из U) на границу Q . Композиция этой ретракции и r дает ретракцию Q на границу, что невозможно. \square

3. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1. В п. 5 раздела I показано, что имеется ретракция Π на любое подмножество, гомеоморфное отрезку, значит, и ретракция на него любого большего подмножества Π . Теорема вытекает из предыдущей леммы. \square

Итак, часть окружности утрачивает важное свойство, которое имеет целая окружность, — разбивать плоскость при любом ее (гомеоморфном) расположении на плоскости. Однако, чтобы целое имело новое свойство, которое не имеют ее части, все же нужно, чтобы эти части имели какие-то свои свойства, которые обеспечат новое свойство целого. Таким свойством отрезка является его способность *локально разбивать* плоскость вблизи его внутренних точек. Мы предлагаем читателю самому определить это свойство.

4. Границы компонент дополнения к C .

ТЕОРЕМА 2. *Граница каждой компоненты дополнения к гомеоморфному образу C окружности совпадает с C .*

ДОКАЗАТЕЛЬСТВО. Границы компонент дополнения во всяком случае лежат в C . Пусть Q — компонента дополнения, граница которой не совпадает с C . Тогда эта граница лежит в простой дуге, содержащейся в C , и Π , а тогда и Q , ретрагируется на эту дугу. Значит, $Q \cup C$ ретрагируется на C . По лемме Q может быть только неограниченной компонентой. В результате C является общей границей всех ограниченных компонент своего дополнения. (На самом деле такая компонента одна, но мы этого не доказали.)

Неограниченная компонента Q_∞ превращается в ограниченную с помощью следующего приема. Пополним Π одной точкой в бесконечности. Получится сфера (топологически). Выкинем одну точку O из другой компоненты — получится плоскость. При этих операциях граница Q_∞ не изменится, а компонента станет ограниченной. \square

5. ЛЕММА. *Пусть, как и раньше, C — гомеоморфный образ окружности. Пусть U — произвольная компонента дополнения к C , b — точка в U и a — точка на C . Для любой окрестности O точки a в плоскости существует простая дуга l , соединяющая b с точкой a' , лежащей в $O \cap C$, причем $l \setminus a'$ лежит в U . Дугу L можно взять конечнозвенной ломаной.*

ДОКАЗАТЕЛЬСТВО. Как показано в предыдущем пункте, каждая точка C является предельной для каждой компоненты дополнения. Поэтому существует ломаная \tilde{l} , соединяющая в U точку b с точкой c , лежащей в O . Считая, что O выпукла (например, круг), соединим a и c отрезком. Пусть a' — первая точка из C на этом отрезке, считая от c . (Такая точка существует по принципу дедекиндова сечения.) Тогда объединение \tilde{l} с отрезком $[c, a']$ дает требуемую ломаную. \square

6. ТЕОРЕМА 3. *Гомеоморфный образ окружности разбивает плоскость ровно на два связанных открытых множества.*

ДОКАЗАТЕЛЬСТВО. Пусть снова дан гомеоморфизм $h : S \rightarrow C$ окружности на компактное подмножество плоскости C , которое, как мы уже знаем, разбивает плоскость Π . Допустим, что компонент дополнения к C больше двух.

Только одна из этих компонент неограниченная и тогда, согласно условию, имеются по крайней мере две ограниченные компоненты (их замыкания компактны). Обозначим две такие компоненты U_1 и U_2 . В каждой из областей U_1 и U_2 возьмем по точке b_1 и b_2 соответственно. Кроме того, возьмем две различные точки a_1 и a_2 на C .

Используя лемму, соединим точку b_i с двумя различными точками a_{ij} , лежащими в малых окрестностях на C соответственно точек a_j , ломаными l_{ij} , причем так, что за исключением конца a_{ij} ломаная l_{ij} лежит в области U_i . Кроме того, можно считать, что l_{i1} и l_{i2} пересекаются только в b_i .

Обозначим малую дугу, содержащуюся в C с концами a_{1j} и a_{2j} через l_j . Замыкание дополнения к сумме двух дуг l_j на C также распадается на две дуги l^1 и l^2 .

Четыре ломаных l_{ij} вместе с дугами l_j на C образуют простую замкнутую кривую Z . Пересечение Z с C состоит из дуг l_j . Эта кривая разбивает плоскость ровно на две компоненты.

В самом деле, каждая точка простой замкнутой кривой на плоскости, как мы показали, является предельной для каждой компоненты дополнения. Поэтому, если такая кривая имеет в качестве своей части хотя бы один отрезок, ее дополнение состоит из двух компонент (так как вблизи точек отрезка точки дополнения распадаются на два подмножества так, что точки каждого из подмножеств можно соединить отрезком, не пересекающим этой кривой).

Из двух компонент, на которые Z разбивает плоскость, одна неограниченная, обозначим ее W , другая ограниченная, обозначим ее V .

Мы утверждаем, что две дуги l^1 и l^2 лежат в разных компонентах дополнения.

В самом деле, если они лежат в одной компоненте, скажем, в W , то V не пересекается с C . Возьмем в V точку v и, пользуясь той же леммой, проведем две дуги, соединяющие ее в V с точками в кругах с центрами b_1 и b_2 , не пересекающимися с C . Затем добавим два отрезка, соединяющие эти точки с b_1 и b_2 . Мы получим дугу, которая лежит в V , за исключением отрезков около своих концов, и соединяет точки b_1 и b_2 , минуя C , что невозможно, так как эти точки выбраны в разных компонентах дополнения к C . Значит, одна из дуг l^i , скажем, l^1 , лежит в ограниченной компоненте V .

Возьмем на l^1 точку a и малый круг o с центром в a , целиком лежащий в V . Как угодно близко от a , в частности, в o , можно найти точку u из неограниченной компоненты дополнения к C .

Эту точку можно связать с как угодно далекой точкой дугой L , не пересекающейся с C . Эта дуга должна пересекать границу области V и, значит, одну из четырех дуг l_{ij} .

Но эти дуги лежат (за исключением своих концов на C) в ограниченных компонентах дополнения к C . Значит, существует дуга, которая, минуя C , соединяет точку неограниченной компоненты дополнения к C с точкой ограниченной компоненты. Это невозможно и предположение, что имеется две ограниченные компоненты, приведено к противоречию. \square

7. ПРИМЕРЫ КНАСТЕРА И ФРАКТАЛЫ. Возможно, читателю трудно допустить, что более двух областей на плоскости могут иметь совпадающую границу. Такие континуумы, которые разбивают плоскость на заданное число областей, границей каждой из которых они являются, были построены впервые польским топологом Кнастером в 1922 году. Их построение см. в книгах [8, 9]

Интересно, что такие примеры возникают в вычислительной математике. Достаточно применить алгоритм метода Ньютона для решения уравнения $z^n = 1$

на комплексной плоскости: $z_{i+1} = z_i - \frac{z_i^n - 1}{nz_i^{n-1}}$. Плоскость разобьется на n областей с общими границами, причем точки из одной области будут порождать последовательность, сходящуюся к одному и тому же корню из 1, а из разных областей к разным корням. Это один из простейших примеров возникновения так называемых *фракталов*.

8. ПРИМЕРЫ АРТИНА—ФОКСА. Приведем простейший пример гомеоморфного образа двумерной сферы в трехмерном пространстве, для которой не существует гомеоморфного преобразования пространства, переводящего этот образ в «круглую» сферу, т. е. для которого не выполнена теорема Шенфлиса. Впервые такой (более сложный) пример был построен в 20-х годах американским топологом Дж. Александером и называется «рогатой сферой Александера». Его построение приведено в книге Л. В. Келдыш. Большая серия разнообразных примеров была построена в 1949 году А. Артином и Р. Фоксом. Мы приведем самый простой из них. Другие примеры и подробности можно найти также в книге Л. В. Келдыш.

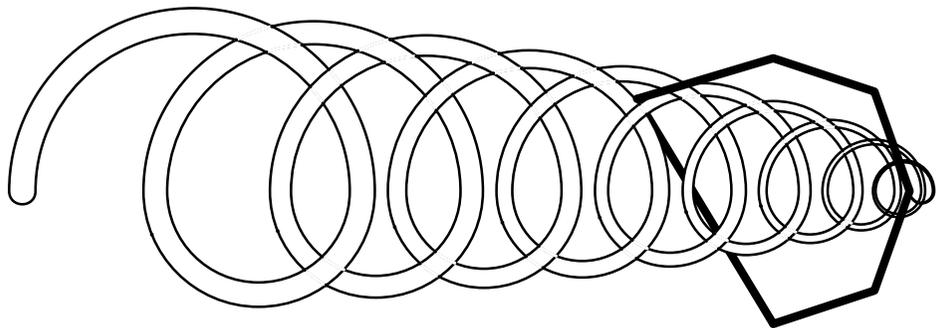


Рис. 4.

Построение примера, надеемся, ясно из рис. 4. Строится дуга, «заузленная в одном конце». Строго это означает отрицание свойства, которое называется локальной односвязностью и которое для стандартного отрезка в трехмерном пространстве удовлетворено для его конечных точек: для «заузленного конца» неверно, что для всякого положительного ε найдется такое положительное δ , что любое отображение окружности в δ -окрестность конца дуги, при котором образ не пересекает дугу, может быть продеформировано в точку в дополнении к дуге в ε -окрестности этого ее конца.

Внимательно рассмотрев рисунок, читатель сообразит, что это так и есть в данном случае. Маленькую петлю, «надетую» на дугу вблизи заузленного конца нельзя снять с дуги, не протаскив ее к другому концу. Строгое доказательство требует знакомства с техникой фундаментальной группы, изложенной, в применении к такого рода примерам, в книге Л. В. Келдыш.

9. ЛОКАЛЬНАЯ ПЛОСКОСТНОСТЬ. Правильная формулировка обобщения теоремы Шенфлиса, как видно из приведенного примера, требует определенных локальных ограничений. Можно доказать, например, что если данный гомеоморфизм сферы на подмножество K пространства можно продолжить до

гомеоморфизма окрестности каждой точки, то он может быть продолжен и до гомеоморфизма пространства на себя. Это условие называется локальной плоскостностью образа сферы.

Достаточно, как показал американский тополог Р. Бинг в 50-х годах, потребовать, чтобы ни в одной точке образа не был возможен феномен примера Артина и Фокса.

IV. ПРИМЕНЕНИЯ И ИСТОРИЯ

1. РЕТРАКЦИИ НА C . В нашем доказательстве теоремы Жордана мы фактически построили ретракцию квадрата Q , содержащего C , на замыкание ограниченной компоненты $\Pi \setminus C$. Ее композиция с ретракцией Π на Q , дает ретракцию всей плоскости на это замыкание. Выкидывая из Π точки ограниченной компоненты дополнения к C , мы получаем ретракцию замыкания неограниченной компоненты на C . Добавляя точку в бесконечности и выкидывая точку из ограниченной компоненты, мы превратим ее в неограниченную компоненту. В результате получаем следующее утверждение, которое мы используем в доказательстве теоремы об инвариантности области:

ЛЕММА. Пусть C — гомеоморфный образ окружности S^1 на плоскости Π , U — ограниченная и U' — неограниченная компоненты дополнения к C . Тогда замыкание $U' \cup C$ неограниченной компоненты ретрагируется на C и также ретрагируется на C множество $U \cup C \setminus x_0$, где x_0 — произвольная точка в U .

В n -мерном случае мы не доказали, что топологический образ $(n-1)$ -мерной сферы разбивает \mathbb{R}^n ровно на две компоненты. Однако, проведенное при доказательстве теоремы Жордана рассуждение (после соответствующей модификации для n -мерного случая) прямо доказывает утверждение этой леммы. Суть в том, что для остовов до размерности $(n-1)$ соответствующего разбиения на кубы не возникает препятствия при построении ретракции, а для n -мерных построения не требуется после проведения операции выметания, как в этом доказательстве.

2. ТЕОРЕМЫ ОБ ИНВАРИАНТНОСТИ ОБЛАСТИ И РАЗМЕРНОСТИ \mathbb{R}^n . Докажем теперь, опираясь на теорему Жордана-Брауэра и теорему о неретрагируемости шара на границу, еще одну фундаментальную теорему Брауэра — *теорему об инвариантности области*. Теорема о негомеоморфности пространств \mathbb{R}^n и \mathbb{R}^k при $k \neq n$ является ее прямым следствием.

ТЕОРЕМА. Если $h: U \rightarrow V$ — гомеоморфное отображение открытого подмножества $U \subset \mathbb{R}^n$ на подмножество $V \subset \mathbb{R}^n$, то V также является открытым подмножеством \mathbb{R}^n .

ДОКАЗАТЕЛЬСТВО. Компоненты открытых подмножеств \mathbb{R}^n открыты. Поэтому можно считать, что U связно. Тогда связно и V .

Пусть $x_0 \in U$ и $y_0 = h(x_0) \in V$. Возьмем малый шар $B \subset U$ с центром в x_0 и границей S . Рассмотрим $C = h(S)$ и покажем, что y_0 лежит в ограниченной компоненте дополнения к C , которая целиком лежит в V . Отсюда будет вытекать, что y_0 — внутренняя точка V и, значит, V — открытое подмножество \mathbb{R}^n .

Во-первых, образ $\text{Int } B$ целиком лежит в одной компоненте, скажем W , дополнения к C , в силу связности.

Если W не лежит в V , то имеется точка $z \in W$, которая не принадлежит V и, значит, образу B . Поэтому, если мы построим ретракцию $(W \setminus z) \cup C$ на C , то с помощью h она перейдет в ретракцию B на S , которой не существует.

Теорема доказана. \square

3. ТЕОРЕМА ЖОРДАНА И ТОПОЛОГИЧЕСКАЯ ХАРАКТЕРИЗАЦИЯ ДВУМЕРНЫХ ПОВЕРХНОСТЕЙ.

ТЕОРЕМА УАЙЛДЕРА. *Локально связный континуум, содержащий простую замкнутую кривую, разбивающийся каждой своей простой замкнутой кривой и не разбивающийся никакой своей парой точек, гомеоморфен сфере.* (См. книгу Уайлдера «Топология многообразий» [11]).

Континуум *локально связан*, если каждая его точка имеет в каждой своей окрестности связную окрестность. В классе таких континуумов двумерные многообразия выделяются тем, что они не разбиваются парами точек и теорема Жордана справедлива для них локально, т. е. каждая *достаточно малая* простая замкнутая кривая их разбивает.

Между собой компактные двумерные многообразия различаются, грубо говоря, тем, сколько можно провести непересекающихся замкнутых кривых, объединение которых не разбивает многообразия (род многообразия). Мы не приводим деталей определений, так как этот материал хорошо доступен в различной учебной литературе. (См., например, книгу Дж. Масси, Дж. Столлинга «Алгебраическая топология» [17].)

Одномерные многообразия (т. е. интервалы, окружность и их дизъюнктные объединения) характеризуются также с помощью понятия «разбивать». Например, окружность характеризуется как локально связный континуум, который не разбивается одной, но разбивается двумя своими точками.

4. Приведем теперь краткий обзор различных доказательств теоремы Жордана и сопутствующих утверждений за ее столетнюю историю.

Первое упоминание о свойстве замкнутой кривой без самопересечений можно отметить в работе Неймана 1865 года. Первое доказательство было дано Жорданом в 1887 году. Оно основывалось на аппроксимации кривой с двух сторон полигональными кривыми, для которых утверждение Жордан считал очевидным и не доказывал. Однако, его дальнейшие рассуждения также были не полны. Фактически он показал, что полигоны можно построить так, что каждая точка данной кривой будет ε -близка к одному из них, а требовалось доказать, что близость будет к обоим полигонам. Тем не менее его доказательство может быть доведено до конца. По существу это было сделано Шенфлисом в 1906 году на основе анализа понятия достижимой точки (т. е. точки кривой, для которой имеется полигональная кривая, пересекающаяся с данной кривой только в этой точке). При этом Шенфлис дал существенное усиление результата Жордана, о котором сказано выше.

Первое строгое доказательство было дано годом раньше Вебленом, который также анализировал понятие достижимости.

Доказательство, данное Брауэром в 1910 году, привело Дж. Александера к идее доказательства, основанного на гомологических свойствах множеств. Суть состоит в том, что если два множества связны и пересекаются по двум связным множествам, то их объединение разбивает плоскость на две компоненты.

В дальнейшем теорема Жордана доказывалась много раз. Обзор имеется в статье Достала и Гиндела [7]. Мы даем неполный список, стараясь привести более или менее оригинальные работы. В основном доказательства следовали либо пути Жордана – Шенфлиса, т. е. аппроксимации полигонами, либо пути Александера, т. е. гомологическому. Автору представляется, что доказательство, данное здесь, является новым.

Оригинальный метод доказательства теоремы Жордана был найден Э. Мойсом [15]. В [16] он был усовершенствован японским математиком Р. Маехара. Метод опирается на следующую лемму, которую Маехара сводит к теореме о неподвижной точке, но которую нетрудно доказать непосредственно: если в прямоугольнике две дуги соединяют пары точек на противоположных сторонах, то они пересекаются.

Простой замкнутый контур C может быть представлен в виде суммы двух дуг u и d (верхняя и нижняя), соединяющих две точки в прямоугольнике на паре противоположных сторон. На отрезке I , соединяющем другую пару сторон, возьмем отрезок с одним концом на u так, что ниже этого конца точек u нет. Тогда точки этого отрезка лежат в ограниченной компоненте дополнения к C , так как иначе легко строится дуга, соединяющая концы I и не пересекающая одну из дуг u или d в противоречие с леммой.

Это изящное рассуждение, однако, не переносится на многомерный случай.

По-видимому, доказательство, данное Александером, привело его к идее обобщения теоремы Жордана – Брауэра на подмножества сферы произвольной размерности — к теореме *двойственности Александера*. Теперь этот результат является элементарным следствием точной последовательности пары и двойственности Пуанкаре. В 1927 году Понтрягин сделал существенное добавление к этому результату, показав, что двойственность Александера реализуется как двойственность зацепления. Грубо говоря, каждому k -мерному циклу в подмножестве сферы отвечает в дополнении к этому подмножеству (однозначно с точностью до гомологии) цикл дополнительной размерности $n - k - 1$, зацепленный с данным. В случае теоремы Жордана - Брауэра основному циклу $(n - 1)$ -мерной сферы отвечает в дополнении цикл из двух точек (по одной в каждой компоненте).

СПИСОК ЛИТЕРАТУРЫ

- [1] *Camille Jordan*. Cours d'Analyse, 1887, 1893.
- [2] *Schoenflies A.* Beiträge zur Theorie der Punktenmenge // Math. Ann. Bd. 62, 1906. P. 86-328.
- [3] *Veblen O.* Theory of plane curves in nonlinear analysis situs // Trans. Amer. Math. Soc. Vol. 6, 1905. P. 83-98.
- [4] *Alexander J. W.* A proof and extension of the Jordan–Brouwer separation theorem // Trans. Amer. Math. Soc. Vol. 23, 1922. P. 333-349.
- [5] *Brouwer L. E. J.* Beweis des Jordanschen Kurvensatzes // Proc. d. Akad. Wet. Amsterdam. Bd. 12, 1910.
- [6] *Brouwer L. E. J.* Über den natürlichen Dimensionsbegriff // Journ. f. Math. Bd. 142, 1913, 146-152.

- [7] *Dostal M., Tindell R.* The Jordan Curve Theorem Revisited // Jahresber. Deutsche Math. Ver. Bd. 80, 1978, 111-128.
- [8] *Пархоменко А. С.* Что такое линия? М., 1954. С. 140.
- [9] *Александров П. С.* Введение в теорию множеств и общую топологию. М., 1977. С. 127.
- [10] *Келдыш Л. В.* Топологические вложения в евклидово пространство // Труды МИАН, №81, 1966.
- [11] *Wilder R. L.* Topology of manifolds. 1949.
- [12] *Вольперт А. И.* Элементарное доказательство теоремы Жордана // УМН. Т. 5, вып. 5, 1950. С. 168-172.
- [13] *Филиппов А. Ф.* Элементарное доказательство теоремы Жордана // УМН. Т. 5, вып. 5, 1950. С. 173-176.
- [14] *van Kampen E. R.* On some characterizations of 2-manifolds // Duke Math. J. Vol 1, 1935. P. 74-93.
- [15] *E. Moise.* Geometric Topology in Dimensions 2 and 3. 1977.
- [16] *R. Maehara.* The Jordan curve theorem via the Brouwer fixed point theorem // Amer. Monthly. Vol 91, 1984. P. 641-643.
- [17] *Масси У., Столлингс Дж.* Алгебраическая топология. Введение. М.: Мир, 1977.

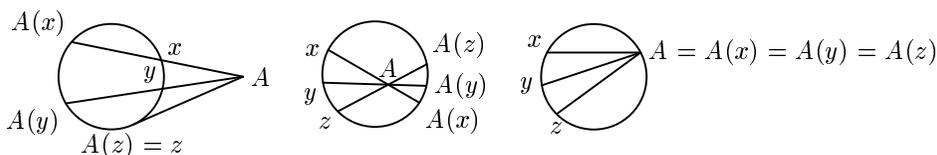
Отображения сферы и неевклидовы геометрии

Р. Р. Пименов

Хотелось бы надеяться на то, что эта статья будет интересна любителям геометрии. Во-первых, в ней единообразно строятся модели всех трех геометрий — Евклида, Римана и Лобачевского. Сама конструкция модели, конечно, связана с идеями Клейна и Пуанкаре, но, в отличие от них, геометрии строятся единообразно и несколько наглядней. Причем все приемы построения были известны не только Паскалю, но, вероятно, и Аполлонию Пергскому. Во-вторых, основную роль в этих моделях играют осевые и точечные симметрии, что продолжает линию Бахмана в основаниях геометрии, иллюстрируя материал его книги «Построение геометрии на основании понятия симметрии». В-третьих, по ходу дела нам удастся проиллюстрировать некоторые важные понятия теории групп и строение бесконечно удаленных точек.

Модели трех геометрий базируются на одном специальном отображении сферы в себя. Естественно начать с соответствующего отображения окружности.

Пусть S^1 — окружность на плоскости и A — некоторая точка. Прямая, проходящая через A и точку $x \in S^1$ и не являющаяся касательной, пересекает S^1 в другой точке, которую мы обозначим $A(x)$. Если A лежит на S^1 , то можно положить $A(x)$ тождественно равной A для всех x . Если же прямая — касательная и x — точка касания, то положим $A(x) = x$. Мы построили отображение $A: S^1 \rightarrow S^1$, которое назовем A -отображением.

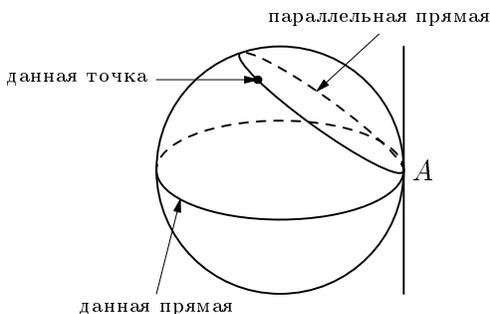


Теперь мы уже готовы к построению моделей разных геометрий. Это значит, что мы должны объяснить, что понимается под словами «точка», «прямая», «угол», «расстояние» и т. п.

Рассмотрим сферу S^2 в трехмерном пространстве и фиксированную точку A . Точка A аналогично плоскому случаю определяет A -отображение этой сферы в себя.

Назовем «точками» моделируемой геометрии пары $(x, A(x))$ где $x \neq A(x)$, т. е. пересечения сферы S^2 прямыми, проходящими через A , а «прямыми» моделируемой геометрии — окружности сферы, лежащие в одной плоскости с A , т. е. пересечения сферы S^2 плоскостями, проходящими через A . Если точка A лежит на S^2 , получается модель евклидовой геометрии, если она лежит внутри S^2 , то — римановой, а если вне — то геометрии Лобачевского.

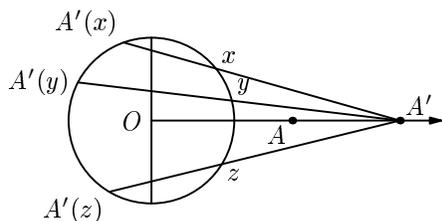
То, что с аксиомой о параллельности в наших моделях все в порядке, видно сразу: если A лежит на S^2 , то через точку вне прямой можно провести единственную параллельную (соответствующую плоскость через данную точку надо провести через прямую, касательную к окружности, являющуюся «данной прямой»).



Если же A лежит внутри сферы, то, как и должно быть, нет параллельных прямых, а если A — вне сферы, то их много. Теперь осталось определить углы и расстояния. Мы сделаем это, как и Бахман, определяя геометрию на основании симметрии. Описанные ранее A -отображения будут являться симметриями геометрий. Свойства же A -отображений тесно связаны с полярным соответствием и другими идеями, восходящими к Паскалю, например, с его знаменитой теоремой о вписанном шестиугольнике. Для осуществления этой программы нам снова придется вернуться к окружности.

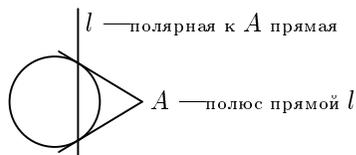
1.

Отметим несколько простых свойств A -отображений. Из самого построения очевидно, что $A(A(x)) = x$. Такие отображения, которые дважды примененные оставляют объекты неподвижными, называют инволютивными. Это очень широкий класс отображений, например, к ним относятся точечные и осевые симметрии. Они удобны, например, тем, что обратное к композиции нескольких инволютивных отображений записывается просто. Например, если P, F, R инволютивны, то обратное к $PPFPR$ запишется просто справа налево — $RPFP$. Изучим теперь некоторые свойства A -отображений.



Пусть точка A' удаляется по лучу $[OA)$, где O — центр окружности S^1 . A' -отображения (т. е. отображения, построенные не из точки A , а из точки A'), будут тогда все меньше отличаться от осевой симметрии относительно l , где l — диаметр S^1 , перпендикулярный к (OA) , поскольку пучок прямых из A' по мере удаления A' все меньше будет отличаться от пучка параллельных, каждая из которых перпендикулярна l . Заметим, что если точка A'' удаляется по лучу $[AO)$ (в другую сторону), то A'' -отображения будут все меньше отличаться от A' -отображений и в пределе тоже совпадут с симметрией относительно l . Мы можем считать, что

то же самое происходит с сферой. Если точка A' удаляется по лучу $[OA)$, то A' -отображения будут все меньше отличаться от осевой симметрии относительно l , где l — диаметр сферы, перпендикулярный к (OA) .



Прежде чем ответить, когда

$$A(B(x)) = B(A(x)),$$

дадим определение. Если точка A лежит вне окружности, то *полярной* к A прямой (или *полярной* к A) называется прямая l , проходящая через две точки окружности, лежащие на касательных, проведенных из A к этой окружности. При этом точку A называют полюсом прямой l . Прежде всего отметим несколько простых свойств полярных прямых: если A лежит на поляре к B , то B лежит на поляре к A ; если полярные трем точкам прямые пересекаются в одной точке, то эти три точки лежат на одной прямой. Хотя эти свойства хорошо известны, мы приводим в приложении их доказательства, поскольку на их примере можно еще раз увидеть роль инволютивных преобразований.

Две точки называются *взаимно полярными*, если одна из них лежит на поляре к другой. В дальнейшем нам понадобится определить в пространстве полярную к точке плоскость относительно сферы S^2 . Это делается совершенно аналогично плоскому случаю: полярной к точке A называют плоскость, проходящую через основание касательного к S^2 конуса с вершиной в A . Полярные плоскости обладают такими же свойствами, что и полярные прямые.

Сформулируем теперь геометрический смысл тождества

$$A(B(x)) = B(A(x)).$$

Оно означает в точности то, что для любой точки x на S точки $A(x)$ и $A(B(x))$ лежат на одной прямой с B . В проективной геометрии доказывается (можно доказать это и используя декартовы координаты, а вот найти красивое наглядное доказательство мне не удалось), что точки пересечения продолжения сторон четырехугольника, вписанного в круг, взаимно полярны. Отсюда следует, что B необходимо лежит на поляре к A . С другой стороны, раз прямая $(x, B(x))$ пересекается с полярной к A в одной точке, то отсюда следует, что в этой же точке пересекается с полярной к A и прямая $(A(x), B(x))$. Отсюда следует достаточность принадлежности точки B поляре к A для выполнения изучаемого тождества. Итак, A -отображение коммутирует с B -отображением тогда и только тогда, когда B лежит на поляре к A . Иначе говоря, когда A и B взаимно полярны.

В заключение наших планиметрических рассуждений заметим, что A -отображения можно свести к инверсии. Именно, инверсией с центром в O называют отображение F , которое отображает произвольную точку плоскости x в точку $F(x)$ так, что:

- ▷ Центр инверсии, точка и ее образ лежат на одной прямой.
- ▷ Произведение расстояний от центра инверсии до точки x и до ее образа не зависит от x . Корень квадратный из этого произведения называют радиусом инверсии.
- ▷ Центр инверсии не лежит между точкой и ее образом.

Инверсии обладают рядом интересных свойств. Они отображают окружности в окружности (или прямые). Каждая инверсия оставляет неподвижной все

точки некоторой окружности S . Если мы рассмотрим все инверсии, отображающие данную окружность S в себя, то они как раз совпадут с A -отображениями (если A вне S , то центр инверсии совпадет с A , поскольку $|A(x), A| \cdot |x, A|$ постоянно; случай, когда A внутри S можно считать инверсией, сопряженной с точечной симметрией). Наконец, инверсии оставляют неизменными углы между пересекающимися кривыми (если углом между кривыми считать угол между их касательными в точке пересечения).

2.

Теперь мы готовы к построению трехмерных моделей различных геометрий. Рассмотрим, как было сказано вначале, A -отображения сферы S^2 в трехмерном евклидовом пространстве в себя.

Отметим, прежде всего, что образ окружности при A -отображении есть снова окружность. Чтобы не заблудиться в стереометрии, сошлемся на известные факты. Во-первых, хорды (или секущие) делятся точкой пересечения на отрезки (другие концы которых — точки пересечения с окружностью), произведение длин которых постоянно и равно квадрату длины касательной из A к S^2 . Во-вторых, это позволяет нам продолжить данное A -отображение сферы до отображения в себя всего пространства, при котором сфера S^2 , разумеется, переходит сама в себя. Это отображение, как и в плоском случае, называется инверсией, его свойства известны и среди них то, что инверсия переводит окружности в окружности. Заметим, что как и в плоском случае, при удалении точки Q , вдоль произвольной прямой l пучок прямых, проходящих через Q , все меньше отличается вблизи сферы S^2 от параллельных к прямой l и Q -отображение все меньше отличается от симметрии относительно плоскости, проходящей через центр S^2 и перпендикулярной l . В пределе его можно считать совпадающим с этой симметрией.

Множество всех точек P , таких что для всех x на сфере $A(P(x)) = P(A(x))$ (т. е. A и P определяют коммутирующие отображения), есть полярная к A плоскость. Это тривиально следует из плоского случая. В самом деле, проведем через P , A и B плоскость γ . γ пересекается с S^2 по некоей окружности T , а с полярной к A плоскостью (обозначим ее π) по некоей прямой l . Поскольку касательные в плоскости γ из A к T есть одновременно и касательные к сфере S^2 , то точки касания лежат в π и, следовательно, полярная прямая к A в плоскости γ совпадает с l . Поскольку в плоскости γ коммутативность отображений, задаваемых A и P , выполняется (P лежит на поляре к A), и все построения с точкой x не выходят за пределы плоскости γ , то требуемое тождество установлено.

Как уже было сказано, «точки» моделируемой геометрии это пары $(x, A(x))$ где $x \neq A(x)$, т. е. пересечения сферы S^2 прямыми, проходящими через A , а «прямые» моделируемой геометрии — окружности сферы, лежащие в одной плоскости с A , т. е. пересечения сферы S^2 плоскостями, проходящими через A . Выясним, какое отображение «точек» и «прямых» моделируемой геометрии задает точка P , лежащая на полярной к A плоскости, но не принадлежащая S^2 .

P -отображение переводит пару $(x, A(x))$ в пару $(P(x), P(A(x))) = (P(x), A(P(x)))$, т. е. «точку» модели в «точку». Поскольку P -отображение переводит окружности на сфере S^2 в окружности, а пару точек, лежащих на одной прямой с A , в пару точек, снова лежащую на одной прямой с A , то «прямые»

модели перейдут в «прямые». Если P лежит вне сферы S^2 , то P -отображение оставляет неподвижным основание касательного конуса к S^2 . Эта окружность лежит на полярной к P плоскости, проходящей через A , и следовательно, является «прямой» модели. Обратно, всякая «прямая» l модели задает некую полярную к ней точку P , лежащую вне S^2 на полярной к A плоскости и такую, что P -отображение оставляет неподвижной l . Пусть теперь P лежит внутри S^2 . Проведем через P и A прямую. Эта прямая пересекает сферу S^2 в двух точках x и $A(x)$, определяющих «точку» модели. Очевидно, эта «точка» перейдет в себя при P -отображении, так как $P, x, A(x)$ лежат в данном случае на одной прямой.

Итак, каждая точка P на полярной к A плоскости π определяет инволютивное P -отображение модели геометрии в себя, причем, если P лежит вне сферы S^2 , то имеется неподвижная «прямая», а если внутри S^2 , то имеется неподвижная «точка». Напрашивается мысль, что мы имеем дело с осевыми и точечными симметриями модели. Прежде чем выяснить, так ли это, надо проверить, как обстоит дело в этих моделях со свойствами взаимного расположения точек и прямых (пересечение и параллельность прямых). Для этого удобнее мыслить «точки» и «прямые» геометрии не как пары точек и окружности на сфере, а как прямые и плоскости, проходящие через эти точки и окружности (и, разумеется, проходящие через A).

Про то, что в указанных моделях все в порядке со свойствами параллельности, было сказано в самом начале. Приступим к изучению метрических свойств модели, т. е. к тому, как можно определить расстояния и углы. При этом желательно, чтобы P -отображения задавали осевые и точечные симметрии. Заметим, что стоящая перед нами задача противоположна возникающей в школьном курсе или в книге Гильберта «Основания геометрии». Обычно предполагается, что мы умеем измерять длину отрезков и величину углов. Исходя из этого мы выясняем, какие отображения плоскости в себя сохраняют расстояния и углы, в частности, убеждаемся, что расстояние между парой точек не меняется при осевой симметрии. Сейчас же мы предполагаем, что некоторые отображения «точек» и «прямых» модели являются симметриями, и хотим узнать, как можно определить длины и углы. Понятно, что определяемое расстояние между «точками» и величина угла между «прямыми» не должны меняться при P -отображении.

Можно поступать различными способами. Можно определить, что отрезок $[AB]$ конгруэнтен отрезку $[A'B']$ тогда и только тогда, когда существует последовательность P -отображений, при которой A перейдет в A' и B в B' , и затем проверить выполнение аксиом конгруэнтности, выделенных Гильбертом в особую группу. Можно, и это проще и красивей, следуя Бахману определять геометрию на основе симметрии и проверить, что свойства P -отображений в точности соответствуют аксиомам Бахмана (при этом, например, свойства P -отображений будут красиво соответствовать теоремам о вписанном в круг шестиугольнике). Но проще всего воспользоваться (и в этом есть методологическое преимущество единообразия, при котором мы не рассматриваем отдельно каждую геометрию, а по возможности оперируем общими свойствами сечений сферы) известными свойствами инверсий.

Именно, оказывается, что проще, чем расстояние между «точками» модели, определить величину угла между «прямыми» модели. Прежде всего, определим угол между двумя пересекающимися окружностями на сфере, как угол между

касательными к ним в точке пересечения (точек пересечения две, но указанные углы одинаковы, из-за симметрии). P -отображения, как было сказано, — частный случай инверсии в трехмерном пространстве. Инверсии хорошо изучены и известно, что они сохраняют определенные таким образом углы между окружностями (и вообще всякими кривыми). Поэтому P -отображение оставляет неизменными углы между окружностями сферы S^2 , в частности, между окружностями, лежащими на одной плоскости с A , т. е. с «прямыми». Мы можем тогда назвать углом между «прямыми» модели угол между задающими их окружностями, поскольку он не меняется при P -отображении. Нужно показать лишь, что через данную «точку» L данной «прямой» l можно отложить под заданным углом φ (и в заданном направлении) одну и только одну «прямую» l' .

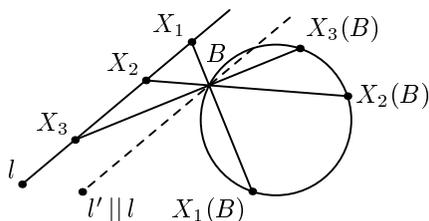
Одну такую «прямую» построить просто. Касательная к окружности, определяющей l (для краткости будем тоже называть ее l , как и далее будем пользоваться одним символом для обозначения окружности и определенной ею «прямой» модели), в точке L лежит на касательной к S^2 в точке L плоскости, там же лежит и касательная к l' . Поскольку угол между этими касательными может быть любым, то проведя через эти касательные и A плоскости, получим «прямые» модели, пересекающиеся под данным углом. Докажем теперь, что второй такой «прямой» нет. Пусть есть l'' такая, что $\angle l, l' = \angle l, l''$. Тогда $\angle l', l'' = 0$ (следует например, из того, что все касательные к трем рассматриваемым «прямым» лежат в одной плоскости и проходят через точку общую касания). Угол между двумя окружностями равен нулю лишь если они касаются друг друга, т. е. имеют всего одну общую точку. Такой случай невозможен для «точки» модели из-за условия $x \neq A(x)$. Поэтому между пересекающимися «прямыми» модели угол не может быть нулевым. Следовательно, $l'' = l'$, что и требовалось. Кстати, заметим, что если окружности l'' и l' пересекаются в полярной к A плоскости, то определяемые ими «прямые» приближаются друг к другу и пересекались бы, если неподвижные точки входили бы в модель.

Итак, откладывание углов заданной величины возможно и единственно. Теперь мы можем сказать, что P -отображение сохраняет углы между прямыми модели. Если геометрия неевклидова (т. е. A не лежит на поверхности сферы), то в ней не существует подобных и не конгруэнтных треугольников, каждый треугольник (и длины его сторон) вполне определяется величинами своих углов. Поэтому, раз P -отображение сохраняет углы в этих геометриях, то оно сохраняет и длины. Что и требовалось. В евклидовом случае возможно подобие треугольников, поэтому сохранение углов не влечет автоматически сохранение длин. Но P -отображение не только сохраняет углы, но и оставляет неподвижным «прямую», а это возможно лишь при осевой симметрии.

Посмотрим теперь, как выглядят окружности в этих моделях. Построение окружностей производится единообразно для всех геометрий и не зависит от расположения S^2 и A . На плоскости окружность можно определить так: будем проводить всевозможные прямые через данную точку O и будем смотреть куда отображается при осевых симметриях относительно этих прямых фиксированная точка A . Множество всех этих образов и составит окружность с центром в O и радиусом $|OA|$. Аналогично поступим для всех трех моделей.

P -отображения оставят неподвижными «точку» O и некоторую «прямую» модели, проходящую через нее, только если P лежит на касательной к O плоскости.

(Строго говоря, «точка» модели есть пара точек, но в следующих построениях можно взять какую-нибудь одну точку из этой пары). Эта плоскость пересекается с полярной к A плоскости по некоторой прямой l . Пусть B лежит на S^2 и вкупе с B' определяет «точку» модели (как уже упоминалось, наличие точки B' не существенно для этих построений), найдем все точки, куда может перейти B при P -отображениях, оставляющих неподвижными O . Пусть точка $x \in l$, тогда $x(B)$ лежит на плоскости β , определенной l и B , причем, если x пробегает всю l , то прямые, проходящие через x и B , заполняют всю β , кроме одной единственной прямой, проходящей через B и параллельной l . Отсюда следует, что множество $x(B)$ заполняет окружность, по которой β пересекается с S^2 , кроме одной точки, лежащей на прямой, проходящей через B и параллельной l . В плоскости β это выглядит, как показано на рисунке.



Но, как мы уже говорили, добавляя бесконечно-удаленную точку (или рассмотрев симметрию относительно перпендикулярной к l плоскости, проходящей через центр S^2) мы получим недостающую точку окружности. Теперь можно рассмотреть $A(O)$ и $A(B)$ и выполнить точно такие же построения. Их итог совпадет с образом уже построенной окружности при P -отображении.

Итак, окружности в исследуемых моделях изображаются парой окружностей на сфере. Можно в нашей модели рассматривать не все точки сферы, а лишь сегмент сферы (как, рассматривая сферическую геометрию, мы ограничиваемся полусферой). Тогда каждая точка этого сегмента будет однозначно соответствовать «точке» модели и обратно, окружность же модели будет изображаться на этом сегменте одной окружностью, а не парой. Как легко видеть, центры этих окружностей модели не совпадают с их обычными центрами на сфере, что неудивительно, поскольку расстояния на сфере и в этих моделях не совпадают (в отличие от углов).

Свяжем теперь эти модели с уже хорошо известными.

Пусть A в центре сферы. Тогда получается общеизвестная модель сферической геометрии. Заметим, что в роли осевых симметрий этой геометрии выступают симметрии относительно плоскостей, проходящих через центр сферы. Эти симметрии совпадают с P -отображениями, где все P — бесконечно удалены. Если A внутри сферы, но отлично от ее центра, то можно перейти через инверсию к случаю, когда A совпадает с центром.

Пусть A на поверхности сферы. Рассмотрим стереографическую проекцию, т. е. проекцию сферы S^2 на плоскость, касающуюся сферы в точке, диаметрально противоположной A . «Прямые» и «точки» определенной нами модели перейдут в прямые и точки плоскости с сохранением всех метрических свойств.

Наиболее интересен случай, когда A — вне сферы. Рассмотрим сечение S^2 полярной к A плоскостью. Обозначим окружность сечения T . Можно сопоста-

вить каждой точке круга, ограниченного T , «точку» геометрии, проведя прямую через нее и A , а каждой хорде в круге — «прямую», проведя плоскость через хорду и A . Тем самым получим модель Клейна для геометрии Лобачевского. Как известно, в этой модели две хорды представляют перпендикулярные прямые, если прямые (на коих лежат хорды) взаимно полярны, что согласуется с проведенными выше построениями.

Указанная связь с уже существующими моделями делает предыдущие рассуждения о том, определяют ли P -отображения конгруэнтные длины и углы, избыточными. Тем не менее, мне хотелось показать, как рассмотрение пучков прямых и плоскостей, пересеченных со сферой, позволяет определить свойства неевклидовых геометрий, одновременно обрабатывая разные геометрические модели (ведь в большей части рассуждений взаиморасположение A и S^2 было несущественно).

ПРИЛОЖЕНИЕ

Докажем приведенные в основном тексте свойства полярных прямых. Доказательство интересно тем, что использует инволютивное преобразование в пространстве.

Пусть радиус исходной окружности R . Пусть тогда точка W расположена над центром окружности O на расстоянии R . Рассмотрим пучок прямых и плоскостей, проходящих через W (для краткости — W -пучок). Каждой прямой этого пучка можно однозначно сопоставить плоскость пучка, перпендикулярную ей. Обратно, каждой плоскости пучка можно однозначно сопоставить перпендикулярную ей прямую пучка. Если прямая W -пучка l лежит на некоей плоскости δ , то перпендикулярная к l плоскость включает в себя перпендикулярную к δ прямую. Если три прямых W -пучка лежат на одной плоскости, то перпендикулярные им плоскости пересекаются по одной прямой.

Теперь выясним, как пересекаются перпендикулярные прямые и плоскости пучка с исходной плоскостью, где лежит окружность. Пусть прямая пересекается в точке A , а плоскость по прямой m . Опустим перпендикуляр из A на m , пусть основание этого перпендикуляра A' . Треугольник $A'WA$ — прямоугольный, $|WO|$ — его высота, отсюда $|WO| \cdot |WO| = |OA'| \cdot |OA|$. Поэтому, если мы отобразим симметрично относительно O прямую m , то получим как раз полярную к точке A . Тем самым точке и полярной к ней прямой можно сопоставить прямую и перпендикулярную ей плоскость W -пучка. Отсюда сразу следуют указанные свойства поляр. Мы рассматривали лишь случай, когда A лежит вне окружности. Если же A лежит внутри окружности, то удобно назвать ее полярной такую прямую, что для всякой точки B на этой прямой, полярная к B проходит через A (из сказанного выше следует, что множество таких точек есть прямая).

Доказательство свойств полярных плоскостей можно получить, рассматривая нужные сечения плоскостями сферы S^2 , или, рассматривая четырехмерное пространство, пучок прямых и трехмерных пространств, проходящих через данную точку, и определяя соответствие прямой и перпендикулярного ей трехмерного пространства.

Наш семинар: математические сюжеты

Вниманию читателя предлагаются две статьи — М. Л. Гервера и А. Б. Скопенкова, посвященные знаменитой проблеме К. Борсука: *верно ли, что любое ограниченное множество в d -мерном пространстве можно разбить на $d + 1$ частей меньшего диаметра?*

Видный польский математик Карол Борсук сформулировал этот вопрос в 1933 году и сам получил положительный ответ на него при $d = 2$ (для плоских множеств). Проблемой Борсука занимались многие выдающиеся математики нашего времени. В частности, П. Эрдеш — человек, решивший за свою жизнь великое множество проблем олимпиадного типа.

Большинство исследователей полагало, что проблема Борсука имеет положительное решение в любой размерности. Но Эрдеш высказывал сомнения. И он оказался прав. Спустя 60 лет после того, как Борсук выдвинул свою гипотезу — в 1993 году — Д. Кан и Г. Калаи построили контрпример к ней: *конечное множество в 1325-мерном пространстве, которое нельзя разбить на 1326 частей меньшего диаметра!* Вскоре в работе А. Нилли был указан более простой путь к построению контрпримеров, и самая маленькая размерность, при которой гипотеза Борсука оказалась опровергнутой, стала равна 946.

В статье Скопенкова приводится упрощенное построение примера А. Нилли.

Существенный шаг был сделан в 1997 году студентом механико-математического факультета А. Райгородским. Он понизил размерность с 946 до 561.

В статье М. Л. Гервера приводятся модификации примеров А. Нилли и А. Райгородского. М. Л. Гервер составил по мотивам проблемы Борсука цикл задач для X летней конференции Турнира Городов, состоявшейся в 1998 году в Гамбурге. Цикл Гервера вызвал большой интерес у школьников, участвовавших в этой конференции. Ими были получены новые содержательные результаты по проблеме Борсука. В частности, было доказано, что проблема Борсука имеет отрицательное решение для любого $d \geq 561$. Этот результат (который вскоре после конференции получили независимо друг от друга школьники Саша Гайфуллин из России и Дима Гуревич из Израиля) публикуется ниже в статье Гервера¹⁾. Независимо от школьников тот же результат был получен Райгородским²⁾.

Вопрос о том, как разрешается проблема Борсука в пространствах размерности $4 \leq d \leq 560$ остается открытым. Может быть читатели нашего журнала смогут сдвинуть его с мертвой точки?

В. М. Тихомиров

¹⁾ В этой статье содержатся и другие интересные факты, установленные школьниками, в частности, новое положительное решение проблемы Борсука в размерности 3, придуманное на конференции Максимом Давыдовым и Русланом Батршиным из Украины.

²⁾ Он публикуется в журнале «Успехи математических наук»; помимо отмеченного выше факта, в этой статье Райгородского имеются и другие важные результаты.

О разбиении множеств на части меньшего диаметра: теоремы и контрпримеры

М. Л. Гервер

Диаметром ограниченного множества M называется *наименьшее* число D , обладающее тем свойством, что *для любых двух точек из M расстояние между ними не превосходит D* . В частности, для множеств M , состоящих из конечного числа точек, диаметр D равен *наибольшему из попарных расстояний между точками M* .

Нетрудно проверить, что любую фигуру (любое ограниченное множество) на плоскости *можно разбить на три части меньшего диаметра*^{<1>*} и что некоторые плоские фигуры *нельзя разбить на две части меньшего диаметра*.

Простейшая из таких фигур состоит из трех точек — трех вершин правильного треугольника. Четыре вершины правильного тетраэдра дают аналогичный пример множества в трехмерном пространстве (его нельзя разбить на три части меньшего диаметра), а $d + 1$ вершин правильного d -мерного симплекса — пример множества, которое нельзя разбить на d частей меньшего диаметра.

ГИПОТЕЗА БОРСУКА

В 1933 г. польский математик Карол Борсук высказал гипотезу:

Любое ограниченное множество в трехмерном пространстве можно разбить на четыре части меньшего диаметра. И вообще: любое ограниченное множество в d -мерном пространстве можно разбить на $d + 1$ частей меньшего диаметра.

В 1955 г. Х. Эгглстон и в 1957 г. Б. Грюнбаум и А. Хепшеш доказали гипотезу Борсука при $d = 3$. А еще раньше она нашла подтверждение для всех центрально-симметричных множеств и всех *гладких выпуклых тел*^{<2>} (Г. Хадвигер, 1946 г.) [1].

Так как выпуклая оболочка любого ограниченного множества M имеет тот же диаметр, что и само множество M , то предположение о выпуклости *не является ограничением*: гипотезу Борсука достаточно проверить для выпуклых множеств. Казалось бы, оставалось сделать немного — избавиться от допущения о *гладкости* ...

КОНТРПРИМЕРЫ К ГИПОТЕЗЕ БОРСУКА

Однако в 1993 г. случилось неожиданное: Д. Кан и Г. Калаи построили контрпример к гипотезе Борсука при $d = 1\ 325$ и для всех $d > 2\ 014$ [2].

*Здесь и далее ссылки типа <1> отсылают к Комментариям — заключительному разделу статьи, расположенному перед Приложениями 1-3.

В [3] и [4] были построены новые контрпримеры — при $d = 946$ и $d = 561$ ^{<3>}.

Ниже предлагается модификация этих контрпримеров. Будет показано, что гипотеза Борсука неверна при всех $d \geq 561$. Кроме того, будет приведено новое — более простое, чем прежние, — доказательство гипотезы Борсука при $d = 3$. Вопрос «Что верно — теорема или контрпример — при $4 \leq d \leq 560$?» остается открытым.

ИДЕЯ КОНТРПРИМЕРА

Положим $d = 561$ и построим конечное множество $Y \subset \mathbb{R}^d$ диаметра D , которое нельзя разбить на $d + 1$ частей диаметра меньше D .

Множество Y будет состоять из $|Y| = H$ точек. Превратим Y в граф Γ , соединяя вершины $y, y' \in Y$ ребром в том и только в том случае, когда $|yy'| < D$, и исследуем клики, или полные подграфы графа Γ , т. е. такие подграфы, в которых каждые две вершины соединены ребром. Нам удастся оценить число q вершин максимальной (содержащей наибольшее возможное число вершин) клики графа Γ , указав такое Q , что

$$q \leq Q, \quad H/Q > d + 1. \quad (1)$$

Из (1), очевидно, следует ^{<4>}: если множество Y разбито на части, диаметр каждой из которых меньше D , то число таких частей обязательно больше $d + 1$, так что Y — контрпример к гипотезе Борсука.

ПЛАН ПОСТРОЕНИЯ КОНТРПРИМЕРА

Возьмем $n = 33$, так что $d = 561 = C_{n+1}^2$. Наряду с множеством $Y \subset \mathbb{R}^d$ будет построено вспомогательное множество $X \subset \mathbb{R}^{n+1}$.

Введем в \mathbb{R}^{n+1} прямоугольную систему координат x_0, x_1, \dots, x_n , рассмотрим куб с вершинами $\{x_j\}_{j=0}^n$, $x_j = \pm 1$, и определим X как следующее подмножество вершин n -мерной грани $x_0 = 1$ этого $(n + 1)$ -мерного куба $[-1, 1]^{n+1}$:

Вершина $\{x_j\}_{j=0}^n$, $x_j = \pm 1$, принадлежит к X , если $x_0 = 1$ и число минус единиц среди $\{x_j\}_{j=1}^n$ четно: $\prod_{j=1}^n x_j = 1$. Тем самым X состоит из $2^{n-1} = 2^{32}$ точек ^{<5>}.

Столько же точек будет и в Y : $|X| = |Y| = H = 2^{n-1} = 2^{32}$. При этом Y будет подмножеством вершин d -мерного куба с вершинами $\{y_k\}_{k=1}^d$, $y_k = \pm 1$: $Y \subset [-1, 1]^d$.

Между X и Y будет установлено такое взаимно однозначное соответствие

$$x = (x_0, x_1, \dots, x_n) \longleftrightarrow y, \quad x' = (x'_0, x'_1, \dots, x'_n) \longleftrightarrow y', \quad (2)$$

что максимально удаленным друг от друга — находящимся на расстоянии D — точкам $y, y' \in Y$ будут соответствовать $x, x' \in X$ со скалярным произведением ± 2 :

$$(x, x') = x_0 x'_0 + x_1 x'_1 + \dots + x_n x'_n = \pm 2 \longleftrightarrow |yy'| = D. \quad (3)$$

Именно соответствие (2) со свойством (3) позволит проверить неравенства (1).

С этой проверки мы и начнем, а уже потом построим множество Y и установим между X и Y соответствие (2), удовлетворяющее условию (3).

МАКСИМАЛЬНАЯ КЛИКА ГРАФА Γ

Ввиду (2), (3), граф Γ можно получить по-новому (не из Y , а из X), соединив ребрами все пары $x, x' \in X$, для которых скалярное произведение $(x, x') \neq \pm 2$.

ЗАМЕЧАНИЕ. Все прочие значения (x, x') при $x \neq x'$ — это ^{<6>} числа $\pm 6, \pm 10, \pm 14, \pm 18, \pm 22, \pm 26$ и ± 30 , а скалярный квадрат (x, x) равен 34. Итак, в графе Γ две различные вершины x и x' соединяются ребром тогда и только тогда, когда

$$(x, x') = \pm 6, \pm 10, \pm 14, \pm 18, \pm 22, \pm 26, \pm 30. \quad (4)$$

Положим $Q = \sum_{k=0}^7 C_{33}^k$ (вскоре станет ясно, откуда берется это число) и докажем, что число вершин q максимальной клики $\{a_1, \dots, a_q\}$ графа Γ не превосходит Q .

МНОГОЧЛЕНЫ g_1, \dots, g_q

Каждой вершине $a_j, 1 \leq j \leq q$, сопоставим многочлен $g_j(x)$ по следующему правилу. Фиксируем a_j . При $x = (x_0 = 1, x_1, \dots, x_n) \in X$ обозначим через $s = s_j(x)$ скалярное произведение (a_j, x) и положим

$$F(s) = (s + 6)(s + 10)(s + 14)(s + 18)(s + 22)(s + 26)(s + 30), \\ G_j(x) = F(s_j(x)), \quad 1 \leq j \leq q. \quad (5)$$

Записав $G_j(x)$ в виде линейной комбинации одночленов и последовательно применяя соотношение $x_i^2 = 1, 1 \leq i \leq n$, ^{<7>} получаем новый, равный $G_j(x)$ на X , многочлен $g_j(x)$, являющийся целочисленной линейной комбинацией функций

$$x_{j_1}^{h_1} x_{j_2}^{h_2} \dots x_{j_7}^{h_7}, \quad 1 \leq j_1 < j_2 < \dots < j_7 \leq n = 33, \quad h_m = 0 \text{ или } 1, \quad 1 \leq m \leq 7. \quad (6)$$

Всего в (6), очевидно, имеется Q функций, образующих базис, $Q = \sum_{k=0}^7 C_{33}^k$, и неравенство $q \leq Q$ будет доказано, если установить, что многочлены g_1, \dots, g_q линейно независимы над кольцом целых чисел, т.е. если проверить, что тождество

$$c_1 g_1(x) + \dots + c_q g_q(x) \equiv 0, \quad \text{где } c_1, \dots, c_q \text{ — целые числа}, \quad (7)$$

возможно лишь тогда, когда все коэффициенты c_1, \dots, c_q в (7) равны нулю ^{<8>}.

ЛИНЕЙНАЯ НЕЗАВИСИМОСТЬ g_1, \dots, g_q

Многочлен $F(s)$ в (5) выбран так, что среди чисел $s = (x, x')$ из (4) все $s < 0$ являются корнями $F(s)$, а все $s > 0$ таковы, что $F(s)$ делится на 27: $F(s) \equiv 0 \pmod{27}$ при $s = 6, 10, 14, 18, 22, 26, 30$; последнее свойство легко проверяется с помощью таблицы 1, содержащей все кратные числу 3 сомножители $F(s)$ для указанных значений s . Из той же таблицы видно, что $F(34) \not\equiv 0 \pmod{27}$. В терминах $g_j(x)$ для x из клики $\{a_1, \dots, a_q\}$ перечисленные свойства $F(s)$ означают:

$$\text{Если } j \neq k, \text{ то } g_j(a_k) \equiv 0 \pmod{27}; \quad g_j(a_j) \not\equiv 0 \pmod{27}. \quad (8)$$

Линейная независимость $g_1(x), \dots, g_q(x)$ легко следует отсюда: допустим, что в (7) не все c_j равны нулю, и докажем, что это допущение противоречит (8).

s	$s + 6$	$s + 10$	$s + 14$	$s + 18$	$s + 22$	$s + 26$	$s + 30$
6	12			24			36
10			24			36	
14		24			36		
18	24			36			48
22			36			48	
26		36			48		
30	36			48			60
34			48			60	

Таблица 1.

Дополнительно предположим, что не все c_j кратны числу 3 (если это не так, можно разделить все c_j на общий множитель). Пусть, для определенности, c_1 не делится на 3. Подставляя в (7) $x = a_1$, получаем: $c_1g_1(a_1) + \dots + c_qg_q(a_1) \equiv 0$. Так как, ввиду (8),

$$c_2g_2(a_1) + \dots + c_qg_q(a_1) \equiv 0 \pmod{27}, \text{ то, тем самым, и } c_1g_1(a_1) \equiv 0 \pmod{27},$$

а значит, снова ввиду (8), $c_1 \equiv 0 \pmod{3}$, — противоречие.

Итак, неравенство $q \leq Q$ (первое из неравенств (1)) доказано. Второе из неравенств (1) проверяется непосредственным подсчетом:

$$H = 2^{32} = 4\,294\,967\,296, \quad Q = \sum_{k=0}^7 C_{33}^k = 5\,663\,890, \quad H/Q > 758. \quad (9)$$

Таким образом, для обоснования контрпримера остается построить множество Y и установить между X и Y соответствие (2), удовлетворяющее условию (3).

СООТВЕТСТВИЕ МЕЖДУ МНОЖЕСТВАМИ X И Y

Пусть, как и прежде, $n = 33$. Построим квадратную таблицу $(n+1) \times (n+1)$. Ее строки и столбцы занумеруем числами $i, j = 0, 1, \dots, n$. Рассмотрим клетки (i, j) в верхней половине таблицы (над диагональю (i, i) , $0 \leq i \leq n$, идущей из левого верхнего в правый нижний угол). Их число равно $d = (34 \cdot 34 - 34)/2 = C_{34}^2 = 561$.

Произвольно фиксируем $x = \{x_j\}_{j=0}^n \in X$ и в клетку на пересечении строки с номером $i - 1$ и столбца с номером j поместим произведение $x_{i-1}x_j$, $1 \leq i \leq j \leq n$. Эти произведения объявим координатами y_k , $1 \leq k = k(i, j) \leq d$, точки $y = y(x) \in Y$.

Поочередно перебрав так все $x \in X$, получаем некоторое множество Y , являющееся подмножеством вершин d -мерного куба $[-1, 1]^d$. Так как $x_0 = 1$ для любого $x \in X$, то $y_{k(1,j)} = x_j$ при всех $j = 1, \dots, n$, т.е. одновременно установлено взаимно однозначное соответствие (2) между X и Y . Докажем, что оно удовлетворяет условию (3).

Квадрат расстояния между $y = \{y_k\}$ и $y' = \{y'_k\}$ равен сумме квадратов разностей $y_k - y'_k$, $1 \leq k \leq d$. Каждое слагаемое в этой сумме равно либо 0 (если $y_k = y'_k$), либо 4 (если $y_k \neq y'_k$), т.е. расстояние между y и y' тем больше,

чем больше несовпадающих y_k и y'_k . Без ограничения общности можно считать, что первые s координат x_j и x'_j совпадают (включая $x_0 = x'_0 = 1$), а остальные $n + 1 - s$ координат различны.

Эквивалентное допущение (сравним с ^{<6>}):

а) все x_j (а значит, и все y_k) равны 1,

б) первые s координат x'_j равны 1, а остальные $n + 1 - s$ координат x'_j равны -1 .

Тогда расстояние между y и y' тем больше, чем больше координат y'_k , равных -1 .

При условиях а) и б) координаты $y'_k = -1$ заполняют прямоугольник в верхней половине нашей таблицы. «Площадь» этого прямоугольника (число *минус единиц* в нем) тем больше, чем больше он похож на квадрат. В точности квадрат получиться не может ^{<9>}, а прямоугольник, наиболее похожий на квадрат, получается, если $s = 18$, $n + 1 - s = 16$, либо $s = 16$, $n + 1 - s = 18$ (т. е. если $(x, x') = \pm 2$).

Итак, расстояние $|yy'|$ между вершинами $y = y(x)$, $y' = y(x') \in Y$ тогда и только тогда равно диаметру D множества Y , когда скалярное произведение $(x, x') = \pm 2$.

Построение контрпримера Y закончено. Трактую Y как подмножество d -мерного пространства при $561 \leq d \leq 757$, вследствие (9) получаем, что *гипотеза Борсука неверна при всех d , $561 \leq d \leq 757$.*

ЗАМЕЧАНИЕ. Вскоре нам понадобится точное значение диаметра D множества Y :

$$D^2 = 4 \cdot 16 \cdot 18, \quad D = \text{diam } Y = 24\sqrt{2}. \quad (10)$$

КЛИКА, СОДЕРЖАЩАЯ Q ВЕРШИН

Неравенство $q \leq Q$ (связывающее число многочленов g_1, \dots, g_q с числом Q многочленов в базисе (6)), служит, оказывается, точной (*неулучшаемой*) оценкой числа вершин максимальной клики графа Γ : опираясь на (4), предьявим клику Π графа Γ , содержащую в точности Q вершин.

Для каждой точки $x = \{x_j\}_{j=0}^n \in X$ среди x_j имеется p координат, равных $+1$, и m координат, равных -1 , где $p > 0$ и $m \geq 0$ — четные числа, $p + m = n + 1$, так что

$$m = 0, 2, \dots, n - 1 = 32. \quad (11)$$

В соответствии с (11) рассмотрим подмножества X_m множества X , $m = 0, 2, \dots, \dots, 32$. К X_0 относится одна точка: все $x_j = 1$; к X_2 относятся C_{33}^2 точек: все x_j , кроме двух, равны 1; к X_4 относятся C_{33}^4 точек: все x_j , кроме четырех, равны 1 и т. д.

Нетрудно проверить, что подграф Π , содержащий все вершины

$$x \in X_0 \cup X_2 \cup X_4 \cup X_6 \cup X_{26} \cup X_{28} \cup X_{30} \cup X_{32}, \quad (12)$$

является кликой: $(x, x') \neq \pm 2$ для любых $x, x' \in \Pi$. При этом число вершин (12) равно $Q = \sum_{k=0}^7 C_{33}^k$ (поскольку $C_{33}^k = C_{33}^{33-k}$, $k = 32, 30, 28, 26$).

ЗАМЕЧАНИЕ. Хотя сама оценка $q \leq Q$ неулучшаема, можно попытаться усовершенствовать способ ее применения: вряд ли при разбиении графа Γ на клики *все они* могут оказаться максимальными, содержащими — *каждая* — по

Q вершин, так что здесь, видимо, есть потенциальная возможность построить контрпример в \mathbb{R}^d для размерностей $d < 561$ ^{<10>}.

РАЗМЕРНОСТИ $d \geq 860$

Вместо пары $(n = 33, d = 561)$ рассмотрим пару $(n = 41, d = 861)$. Так же, как прежде ^{<11>}, построим множества $X \subset [-1, 1]^{41}$, $Y \subset [-1, 1]^{861}$, $|Y| = |X| = H = 2^{40}$, при этом — вместо (9) — получатся следующие соотношения:

$$Q = \sum_{k=0}^9 C_{41}^k = 473\,732\,328, \quad H = 2^{40} = 1\,099\,511\,627\,776, \quad H/Q > 2\,320. \quad (13)$$

Поскольку Y можно трактовать как подмножество \mathbb{R}^d при $861 \leq d \leq 2\,319$, то из неравенства (13) следует, что гипотеза Борсука неверна при всех d , $861 \leq d \leq 2\,319$, а также при $d = 860$ ^{<12>}. Значит, с учетом [2], она неверна при всех $d \geq 860$.

Итак, контрпримеры к гипотезе Борсука построены при всех d , $561 \leq d \leq 757$ и $d \geq 860$. До августа 1998 г. вопрос «Что верно — теорема или контрпример — при $757 < d < 860$?» оставался открытым.

РЕЗУЛЬТАТЫ ШКОЛЬНИКОВ — ПОБЕДИТЕЛЕЙ XIX ТУРНИРА ГОРОДОВ

В августе 1998 г. в Гамбурге состоялась X летняя конференция Турнира Городов ^{*}). На конференции я предложил ее участникам, победителям XIX Турнира Городов, ряд задач о разбиении множеств на части меньшего диаметра, включая некоторые нерешенные вопросы (в том числе, о размерностях d , $757 < d < 860$).

В Гамбурге украинские школьники Руслан Батршин и Максим Давыдов (г. Львов) нашли новое — более простое, чем прежние, — доказательство гипотезы Борсука при $d = 3$ (доказательство приведено в конце этой статьи, см. Приложение 3).

А вскоре после конференции школьники Дима Гуревич (г. Тель-Авив, Израиль) и Саша Гайфуллин (г. Жуковск, Россия) независимо друг от друга получили ответ на вопрос о размерностях d , $757 < d < 860$: для всех этих d они построили контрпримеры к гипотезе Борсука. Их построение базируется на следующей, совсем простой, лемме:

ЛЕММА. Пусть в прямоугольном треугольнике ABC гипотенуза AB ($|AB| = D$) и катеты AC и BC ($|AC| = r$, $|BC| = R$) связаны соотношением $r \leq R$, или, что то же самое, $2r^2 \leq D^2$. Пусть точка T катета BC такова, что $|TA| = |TB| = \rho$. Тогда $2\rho^2 \leq D^2$.

Лемма сразу следует из подобия треугольников ABC и TBS , где S — середина AB (рис. 1), а применение ее основано на том, что Y (наш 561-мерный контрпример к гипотезе Борсука) лежит на 560-мерной сфере радиуса r ($r^2 = 561$) и, с учетом (10), r и $\text{diam } Y = D$ связаны соотношением $2r^2 \leq D^2$ ($r^2 = 561 = 17 \cdot 33 < 18 \cdot 32 = D^2/2$).

^{*}Подробный рассказ о Турнире Городов (международном математическом соревновании школьников старших классов, которое проводится, начиная с 1980 г.) можно найти в [5].

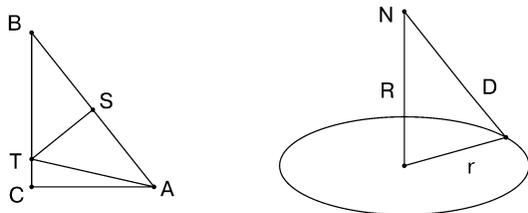


Рис. 1.

Точная формулировка утверждения Д. Гуревича–А. Гайфуллина такова:

УТВЕРЖДЕНИЕ. Пусть множество M диаметра D лежит в k -мерном пространстве на $(k-1)$ -мерной сфере с радиуса r , причем

$$r^2 \leq D^2/2, \quad (14)$$

и пусть M является контрпримером к гипотезе Борсука: M нельзя разбить на $k+1$ частей, диаметр каждой из которых меньше D . Тогда контрпример к гипотезе Борсука в \mathbb{R}^d можно построить при любом $d > k$.

ДОКАЗАТЕЛЬСТВО УТВЕРЖДЕНИЯ. На рис. 1 $(k-1)$ -мерная сфера условно изображена в виде окружности радиуса r . Множество $M = M_1$ диаметра D (контрпример к гипотезе Борсука в \mathbb{R}^k) лежит на s , причем $2r^2 \leq D^2$. Добавляя одну размерность и добавляя к M_1 одну точку N , расположенную на расстоянии D от всех точек s , получаем множество $M_2 \in \mathbb{R}^{k+1}$, которое по-прежнему имеет диаметр D и которое, очевидно^{<13>}, нельзя разбить на $k+2$ части диаметра меньше D .

Тем самым, M_2 является контрпримером к гипотезе Борсука в \mathbb{R}^{k+1} .

Так как (по лемме) M_2 лежит на k -мерной сфере, для радиуса которой выполняется (14), то добавление — по одной — новых точек (с увеличением размерности на единицу) можно продолжить по индукции.

А. Гайфуллин доказал также^{<14>}, что доказанное утверждение остается верным, если заменить в нем (14) неравенством

$$r^2 \leq 3D^2/4. \quad (15)$$

НЕКОТОРЫЕ НЕРЕШЕННЫЕ ВОПРОСЫ

Несмотря на успехи последних лет, остается ряд нерешенных вопросов, связанных с задачей Борсука. Вот — основные два из них:

1. Что верно — теорема или контрпример — при $4 \leq d \leq 560$?

2. Обозначим через $f(d)$ минимальное из таких чисел m , что любое ограниченное множество в d -мерном пространстве можно разбить на m частей меньшего диаметра. В [2] получена оценка $f(d) \geq (1,2)^{\sqrt{d}}$, начиная с некоторого d . Недавно она уточнена в [7]: существует такая функция $r(d)$, $1 \leq r(d) = o(e^{\sqrt{d}})$, что

$$f(d) \geq r(d) \cdot (2/\sqrt{3})^{\sqrt{2d}} \geq (1,2255)^{\sqrt{d}}.$$

Как на самом деле растет $f(d)$ с ростом d ?

КОММЕНТАРИИ

1. Пусть M — произвольное множество диаметра D на плоскости. Проведем горизонтальную прямую L_1 , содержащую хотя одну граничную точку M и такую, что *над* L_1 нет точек M . *Под* L_1 параллельно L_1 проведем прямую L'_1 на расстоянии D от L_1 . Под L'_1 , очевидно, нет точек M , т.е. мы заключили M в полосу P_1 ширины D .

Построив так же еще 2 полосы P_2 и P_3 ширины D , расположенные под углом 60° к P_1 , заключим M в 6-угольник S (являющийся пересечением трех полос ширины D).

У такого 6-угольника S противоположные стороны попарно параллельны, а длины сторон a и b ($a \leq b$) чередуются: $a - b - a - b - a - b$. Опуская из центра S перпендикуляры на стороны длины b , мы разбиваем S (а значит, и M) на три части диаметра меньше D , рис. 2.

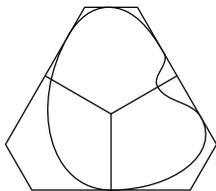


Рис. 2.

ЗАМЕЧАНИЕ. Стандартное решение этой задачи содержит дополнительный шаг: вращая 6-угольник S , добиваются, чтобы он стал правильным. Для разбиения S и M на 3 части диаметра меньше D этот шаг не нужен. Другие решения см. в Приложении 1.

2. Через каждую граничную точку выпуклого множества V в трехмерном пространстве можно провести *хотя бы одну опорную плоскость* (т.е. такую плоскость, что V лежит по одну сторону от нее). Через вершину выпуклого многогранника (например, через вершину куба) проходит бесконечно много опорных плоскостей. Выпуклое множество (или выпуклое тело) V называется *гладким*, если через каждую граничную точку V проходит *единственная* опорная плоскость.

В Приложении 2 доказано, что любое *гладкое* выпуклое тело в трехмерном пространстве можно разбить на 4 части меньшего диаметра.

Теорема Хадвигера в d -мерном случае доказывается аналогично.

3. В [3] контрпримеры построены для всех размерностей $d = C_{4p}^2$, где p — простое число, большее или равное 11; в частности, при $p = 11$ получается $d = = 946 = C_{44}^2$.

4. Как бы ни разбить множество Y , состоящее из N точек, на $d + 1$ частей Y_1, \dots, Y_{d+1} , хотя бы в одной из них (скажем, в Y_j) окажется, ввиду (1), больше, чем q точек; значит, Y_j не является кликой графа Γ , и — по определению графа Γ — в Y_j найдутся точки y, y' , отстоящие друг от друга на расстояние $D = \text{diam } Y$. Тем самым диаметр Y_j равен диаметру Y , т.е. (1) дает контрпример к гипотезе Борсука.

5. Знаки $\{x_j\}_{j=1}^{32}$ выбираются произвольно, а знак x_{33} определяется однозначно — так, чтобы $\prod_{j=1}^{33} x_j = 1$. Тем самым X состоит из 2^{32} точек.

6. Для любых $x = \{x_j\}$ и $x' = \{x'_j\}$ из X скалярное произведение

$$(x, x') = x_0 x'_0 + x_1 x'_1 + \dots + x_n x'_n,$$

очевидно, совпадает со скалярным произведением (e, e') , где $e = \{e_j\}$ и $e' = \{e'_j\}$ — точки в \mathbb{R}^{n+1} с координатами e_j , тождественно равными 1, и координатами e'_j , равными $x_j x'_j$ при всех j . Обе точки — и e , и e' — принадлежат X , поскольку

$$e_0 = e'_0 = 1 \quad \text{и} \quad \prod_{j=1}^n e'_j = \prod_{j=1}^n x_j \cdot \prod_{j=1}^n x'_j = 1.$$

Если $e'_j = 1$ при всех j , то $(e, e') = 34$; если $e'_j = 1$ при всех j , кроме двух, то $(e, e') = 30$; если $e'_j = 1$ при всех j , кроме четырех, то $(e, e') = 26$ и т. д.

7. Напомню: все $x_i = \pm 1$, поэтому $x_i^2 = 1$, $1 \leq i \leq n$.

8. Каждый из многочленов $g_j(x)$ является линейной комбинацией с целыми коэффициентами одночленов (6), т.е. одночленов

$$1, x_1, x_2, \dots, x_n, x_r x_s \quad (1 \leq r < s \leq n), x_r x_s x_t \quad (1 \leq r < s < t \leq n), \dots$$

Обозначим их e_1, e_2, \dots, e_Q , где $Q = \sum_{k=0}^7 C_n^k$, $n = 33$.

Смысл каждого слагаемого в этой сумме ясен: $C_{33}^0 = 1$ — имеется 1 одночлен, тождественно равный 1; $C_{33}^1 = 33$ одночлена являются линейными функциями; C_{33}^2 одночленов — попарные произведения $x_r x_s$, $1 \leq r < s \leq n$; C_{33}^3 одночленов — произведения $x_r x_s x_t$, $1 \leq r < s < t \leq n$; и т. д.

В этих обозначениях

$$g_j(x) = K_{1,j} e_1 + \dots + K_{Q,j} e_Q, \quad K_{1,j}, \dots, K_{Q,j} \text{ — целые числа, } 1 \leq j \leq q.$$

Если показать, что *тождество (7) возможно только при том условии, что все коэффициенты c_1, \dots, c_q равны нулю*, то для тех, кто знает линейную алгебру, неравенство $q \leq Q$ становится очевидным:

e_1, \dots, e_Q образуют базис из Q элементов, и число элементов q в системе линейно независимых g_1, \dots, g_q не может быть больше Q .

Докажем неравенство $q \leq Q$, не опираясь на этот факт. Переписывая (7) в виде

$$c_1(K_{1,1} e_1 + \dots + K_{Q,1} e_Q) + \dots + c_q(K_{1,q} e_1 + \dots + K_{Q,q} e_Q) = \\ (c_1 K_{1,1} + \dots + c_q K_{1,q}) e_1 + \dots + (c_1 K_{Q,1} + \dots + c_q K_{Q,q}) e_Q \equiv 0,$$

видим, что (7) выполняется тогда и только тогда, когда

$$c_1 K_{i,1} + \dots + c_q K_{i,q} = 0 \quad \text{при всех } i, 1 \leq i \leq Q. \quad (16)$$

Будем трактовать (16) как систему Q линейных уравнений с целыми коэффициентами $K_{i,j}$ относительно q неизвестных c_1, \dots, c_q . Задача будет решена, если показать, что при условии $q > Q$ (т.е. при условии, что неизвестных больше, чем уравнений) система (16) имеет ненулевое целочисленное решение c_1, \dots, c_q .

А это, действительно, так, поскольку для любой совместной (имеющей хоть одно решение) системы уравнений верно следующее более общее

s	$s + 6$	$s + 10$	$s + 14$	$s + 18$	$s + 22$	$s + 26$
6	$3 \cdot 2^2$	2^4	$5 \cdot 2^2$	$3 \cdot 2^3$	$7 \cdot 2^2$	2^5
10	2^4	$5 \cdot 2^2$	$3 \cdot 2^3$	$7 \cdot 2^2$	2^5	$9 \cdot 2^2$
14	$5 \cdot 2^2$	$3 \cdot 2^3$	$7 \cdot 2^2$	2^5	$9 \cdot 2^2$	$5 \cdot 2^3$
18	$3 \cdot 2^3$	$7 \cdot 2^2$	2^5	$9 \cdot 2^2$	$5 \cdot 2^3$	$11 \cdot 2^2$
22	$7 \cdot 2^2$	2^5	$9 \cdot 2^2$	$5 \cdot 2^3$	$11 \cdot 2^2$	$3 \cdot 2^4$
26	2^5	$9 \cdot 2^2$	$5 \cdot 2^3$	$11 \cdot 2^2$	$3 \cdot 2^4$	$13 \cdot 2^2$
30	$9 \cdot 2^2$	$5 \cdot 2^3$	$11 \cdot 2^2$	$3 \cdot 2^4$	$13 \cdot 2^2$	$7 \cdot 2^3$

Таблица 3.

ца 3, позволяющая утверждать, что значения (нового — 6-й степени) многочлена

$$F(s) = (s + 6)(s + 10)(s + 14)(s + 18)(s + 22)(s + 26)$$

при $s = 6, 10, 14, 18, 22, 26$ делятся на 2^{17} , а $F(30)$ делится лишь на 2^{16} и не делится на 2^{17} ; однако — вместо (9) — можно получить лишь неравенства

$$q \leq Q = \sum_{k=0}^6 C_{29}^k = 621\,616, \quad H = 2^{28} = 268\,435\,456, \quad H/Q > 431.$$

Оценка $H/Q > 431$ чуть хуже, чем требуется для обоснования контрпримера в \mathbb{R}^d при $d = 435$. Тем не менее, пара ($n = 29, d = 435$) представляется «перспективной»: едва ли при разбиении соответствующего графа на клики *все они* могут оказаться максимальными, и если бы удалось сосчитать, сколько может быть непересекающихся клик, содержащих — *каждая* — по Q вершин, то число 431 в последней оценке, возможно, удалось бы заменить числом 436.

11. Подробнее эти слова означают следующее.

а) При $n = 41$ вложим n -мерный куб $[-1, 1]^n$ в $(n + 1)$ -мерное пространство с координатами $\{x_j\}_{j=0}^n$, и, трактуя $[-1, 1]^{41}$ как 41-мерную грань $x_0 = 1$ куба $[-1, 1]^{42}$, определим X подобно тому, как это было сделано прежде: вершина $\{x_j\}_{j=0}^n, x_j = \pm 1$, принадлежит к X , если $x_0 = 1$ и число минус единиц среди $\{x_j\}_{j=1}^n$ четно.

Ясно, что X содержит $H = 2^{n-1}$ точек: $|X| = H = 2^{40}$.

б) Так же, как прежде, отобразим X на подмножество Y вершин d -мерного куба $[-1, 1]^d, d = C_{42}^2 = 861$.

Рассмотрим множество P всех пар $(i, j), 1 \leq i \leq j \leq n = 41$ (число таких пар равно d , так что их можно перенумеровать числами $k = k(i, j), 1 \leq k \leq d$) и вершине $x = \{x_j\}_{j=0}^n \in X$ сопоставим точку $y = y(x) \in Y$ с координатами

$$y_k = y_{k(i, j)} = x_{i-1} x_j, \quad (i, j) \in P, \quad 1 \leq k = k(i, j) \leq d.$$

Множество Y , как и X , содержит H точек: $|Y| = |X| = H = 2^{40}$.

с) Для $x, x' \in X$ скалярное произведение $(x, x') = \sum_{j=0}^n x_j x'_j$ принимает значения

$$42, \pm 38, \pm 34, \pm 30, \pm 26, \pm 22, \pm 18, \pm 14, \pm 10, \pm 6, \pm 2. \quad (17)$$

Связанный с ними многочлен 9-й степени

$$F(s) = (s + 6)(s + 10)(s + 14)(s + 18)(s + 22)(s + 26)(s + 30)(s + 34)(s + 38) \quad (18)$$

обладает следующими свойствами:

$F(s) \equiv 0 \pmod{11}$ при всех $s > 0$ из (17), кроме $s = 42$, и $F(42) \not\equiv 0 \pmod{11}$.

d) Расстояние $|yy'|$ между вершинами $y = y(x)$, $y' = y(x') \in Y$ тогда и только тогда равно диаметру D множества Y , когда скалярное произведение $(x, x') = \pm 2$.

e) Превратим X в граф Γ , соединив ребрами все пары вершин $x, x' \in X$, для которых $(x, x') \neq \pm 2$. Число вершин максимальной клики графа Γ равно $Q = \sum_{k=0}^9 C_{41}^k$.

То, что их не больше Q , проверяется с использованием свойств многочлена $F(s)$ из (18); то, что их не меньше Q , показывает (сравним с (12)) пример клики

$$X_0 \cup X_2 \cup X_4 \cup X_6 \cup X_8 \cup X_{32} \cup X_{34} \cup X_{36} \cup X_{38} \cup X_{40}.$$

12. Фиксируя любую из координат y_k , получаем разбиение Y на два подмножества Y^+ и Y^- размерности 860 (на которых $y_k = +1$ и $y_k = -1$). Хоть одно из них нельзя разбить на 861 часть диаметра меньше D — иначе (вопреки (13)) Y разбилась бы на 1722 таких части.

Это дает контрпример (Y^+ или Y^-) к гипотезе Борсука для $d = 860$.

13. При разбиении M_2 на части диаметра меньше D точка N не может попасть в одну часть ни с одной из точек M_1 , так как удалена от каждой из них на расстояние D .

14. В d -мерном пространстве введем декартовы координаты $\{u_j\}_{j=1}^d$. Сферу s в k -мерном подпространстве $u_j = 0$, $k + 1 \leq j \leq d$, зададим уравнением

$$\sum_{j=1}^k u_j^2 = r^2.$$

Пусть $d > k + 1$ (случай $d = k + 1$ будет рассмотрен отдельно). Положим $n = d - k - 1$ (так что $n \geq 1$ и $k + n = d - 1$), определим $h \geq 0$ из условия

$$r^2 + D^2/4 + h^2 = D^2 \quad (19)$$

(ввиду (15) такое h существует) и в $(n + 1)$ -мерном подпространстве $u_1 = \dots = u_k = 0$ возьмем $(n - 1)$ -мерную сферу s' диаметра D

$$\sum_{j=k+1}^{k+n} u_j^2 = D^2/4, \quad u_d = h.$$

Условие (19) означает, что расстояние от любой точки $P \in s$,

$$P = (u_1, \dots, u_k, 0, \dots, 0, 0), \quad \sum_{j=1}^k u_j^2 = r^2,$$

до любой точки $Q \in s'$,

$$Q = (0, \dots, 0, u_{k+1}, \dots, u_{k+n}, u_d = h), \quad \sum_{j=k+1}^{k+n} u_j^2 = D^2/4,$$

равно D :

$$|PQ|^2 = \left(\sum_{j=1}^k u_j^2 \right) + \left(\sum_{j=k+1}^{k+n} u_j^2 \right) + h^2 = r^2 + D^2/4 + h^2 = D^2. \quad (20)$$

Объединение множества $M \subset s$ со сферой s' обозначим через M' ; по построению диаметр M' равен D . По условию M нельзя разбить на $k+1$ частей диаметра меньше D . Сферу s' нельзя разбить на n частей диаметра меньше D . Иными словами, при разбиении M и s' на части меньшего диаметра минимальное число частей равно соответственно $p(M) \geq k+2$ и $n+1$. Поэтому (вследствие (20)) для M' минимальное число частей меньшего диаметра равно $p(M) + n + 1 \geq k + n + 3 = d + 2$, то есть M' — контрпример к гипотезе Борсука в d -мерном пространстве:

$$M' \text{ нельзя разбить на } d+1 \text{ частей диаметра меньше } D, \quad (21)$$

и это доказано — при условии (15) — для любого $d > k + 1$.

Чтобы получить (21) для $d = k + 1$, достаточно, чтобы выполнялось условие $r \leq D$. В этом случае M' получается из M добавлением одной точки

$$Q = \left(0, \dots, 0, u_d = u_{k+1} = \sqrt{D^2 - r^2} \right),$$

лежащей на расстоянии D от всех точек сферы s .

ПРИЛОЖЕНИЕ 1. ТЕОРЕМА БОРСУКА В \mathbb{R}^2 И В \mathbb{R}^3 .

ДОКАЗАТЕЛЬСТВА ТЕОРЕМЫ ДЛЯ КОНЕЧНЫХ МНОЖЕСТВ

В Приложении 3 приведено новое доказательство гипотезы Борсука в \mathbb{R}^3 .

Чтобы разобраться в нем, полезно самостоятельно перенести (переформулировать и доказать) леммы из Приложения 3 на случай \mathbb{R}^2 .

Поскольку *контрпримеры* к гипотезе Борсука были построены для *конечных* множеств, особый интерес вызывают и *доказательства* для таких множеств.

На конференции в Гамбурге уже упоминавшийся в этой статье Дима Гуревич предложил следующий план решения задачи Борсука для конечных множеств.

Пусть M — конечное множество в d -мерном пространстве, $\text{diam } M = D$. Превратим M в граф $G = G(M)$, соединяя точки P, Q из M тогда и только тогда, когда $|PQ| = D$ (иначе говоря, проведем все «отрезки — диаметры» в M). Назовем ситуацию *особой*, если степень каждой вершины графа G не меньше $d + 1$.

ЛЕММА. Если в \mathbb{R}^d особая ситуация невозможна, то гипотеза Борсука верна для любого конечного множества $M \subset \mathbb{R}^d$.

ДОКАЗАТЕЛЬСТВО. Индукция по числу точек M .

ТЕОРЕМА. Особая ситуация невозможна в \mathbb{R}^2 и в \mathbb{R}^3 .

Сформулированная теорема

а) верна и без труда проверяется в \mathbb{R}^2 ;

б) верна и довольно сложно доказывается в \mathbb{R}^3 (на конференции в Гамбурге доказать ее в \mathbb{R}^3 никому не удалось);

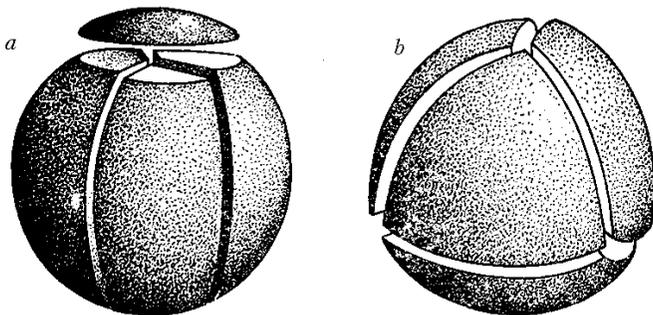


Рис. 3.

с) заведомо не обобщается на \mathbb{R}^d при $d \geq 4$; для этих d легко построить контр-примеры к теореме, но они не дают никакого продвижения в задаче Борсука.

Из теоремы и из леммы следует решение задачи Борсука в \mathbb{R}^2 и в \mathbb{R}^3 для конечных множеств.

Сама теорема является следствием утверждений А и В:

А. Пусть $M \subset \mathbb{R}^2$ состоит из n точек; тогда в $G(M)$ максимум n ребер.

В. Пусть $M \subset \mathbb{R}^3$ состоит из n точек; тогда в $G(M)$ максимум $2n - 2$ ребра.*)

ПРИЛОЖЕНИЕ 2. ТЕОРЕМА ХАДВИГЕРА

Сначала разобьем трехмерный шар на 4 части меньшего диаметра.

Одно нужное разбиение шара показано на рис. 3а. Другое (более симметричное) можно получить так: впишем в шар с центром O правильный тетраэдр $ABCD$; четыре трехгранных угла $OBCD$, $OACD$, $OABD$ и $OABC$, под которыми видны из центра грани тетраэдра, рассекают шар на 4 части меньшего диаметра (рис. 3б).

Теперь, следуя [1], докажем, что верна

ТЕОРЕМА ХАДВИГЕРА в \mathbb{R}^3 . Любое гладкое выпуклое тело V диаметра D в трехмерном пространстве можно разбить на 4 части меньшего диаметра.

ДОКАЗАТЕЛЬСТВО. Разобьем шар U диаметра D на четыре части U_1, U_2, U_3, U_4 меньшего диаметра. Каждой граничной точке $v \in V$ сопоставим такую граничную точку $u = u(v) \in U$, что опорная плоскость к V в точке v параллельна касательной плоскости к U в точке u (при этом u выбирается так, чтобы U и V лежали по одну сторону от этих плоскостей). Точку v на границе V отнесем к V_j , если $u(v) \in U_j$, $1 \leq j \leq 4$. Докажем, что

$$\text{diam } V_j < D, \quad 1 \leq j \leq 4. \quad (22)$$

Допустим, что при каком-нибудь j это не так и $\text{diam } V_j = D$. Пусть A и B — две граничные точки V_j , для которых $|AB| = D$. Проведем через A и B две плоскости, перпендикулярные отрезку AB . Ясно, что V лежит в полосе между ними (иначе $\text{diam } V > D$). Поэтому проведенные плоскости являются *опорными*

*) Утверждения А и В доказаны в [6, задачи 87 а, б]; см. также [6, задачи 101, 102].

к V . Значит, касательные плоскости к шару в точках $u(A)$ и $u(B)$ (параллельные проведенным опорным плоскостям) *параллельны* друг другу, т.е. $u(A)$ и $u(B)$ — диаметрально противоположные точки шара U . Поэтому расстояние между ними равно D . С другой стороны, по построению $u(A)$ и $u(B)$ принадлежат U_j , так что расстояние между ними меньше D . Полученное противоречие доказывает (22).

Пусть теперь O — любая точка внутри V . Соединим O отрезками со всеми точками $v \in V_j$ и объединение этих отрезков обозначим через W_j , $1 \leq j \leq 4$. Ясно, что, ввиду (22), $\text{diam } W_j < D$. Построенные «конусы» W_j заполняют все выпуклое тело V , т.е. образуют разбиение V на 4 части диаметра меньше D .

ПРИЛОЖЕНИЕ 3. ТЕОРЕМА БОРСУКА В \mathbb{R}^3

В августе 1998 года на X летней конференции Турнира Городов в Гамбурге львовские школьники Руслан Батршин и Максим Давыдов получили

новое доказательство гипотезы Борсука в \mathbb{R}^3 .

ТЕОРЕМА. Пусть M — замкнутое выпуклое множество в \mathbb{R}^3 , D — диаметр M . Тогда M можно разбить на 4 части — каждая диаметра меньше D .

ДОКАЗАТЕЛЬСТВО. Введем следующие обозначения:

B_0 — наименьший шар, содержащий M ,

S_0 — его сфера,

R_0 — его радиус,

V — замкнутая выпуклая оболочка $M \cap S_0$ (т.е. минимальное замкнутое выпуклое множество, содержащее все точки множества M , попавшие на сферу S_0).

ЛЕММА 1. V содержит центр O шара B_0 .

ЛЕММА 2. Возможны три случая:

(а) V — отрезок длины $2R_0$,

(б) V — треугольник в экваториальной плоскости B_0 ,

(с) V содержит вписанный в шар B_0 тетраэдр $T \ni O$.

ЛЕММА 3 (О МИНИМАКСЕ). Пусть T_0 — правильный тетраэдр, вписанный в шар B_0 , ρ_0 — длина его ребра; T — неправильный тетраэдр, вписанный в шар B_0 и содержащий центр O шара B_0 , ρ — длина наибольшего ребра T . Тогда $\rho > \rho_0$.

Используя конструкцию, изображенную на рис. 3б (см. Приложение 2), разделим шар B_0 на четыре части C_1, C_2, C_3, C_4 .

ЛЕММА 4. $\text{diam } C_j = \rho_0$, $1 \leq j \leq 4$.

В случаях (а) и (б) (см. лемму 2) теорема Борсука легко следует из леммы 4.

В случае (с), если тетраэдр T в лемме 2 — неправильный, теорема Борсука следует из лемм 3 и 4.

Наконец, пусть T в случае (с) — правильный тетраэдр $ABCD$ (с ребром ρ_0), так что $\text{diam } M \geq \rho_0$. Тогда построим тетраэдр $T' = A'B'C'D'$, симметричный T относительно центра O шара B_0 . Исходя из T' , проведем такое построение, как на рис. 3б, и применим лемму 4.

Если $\text{diam}(M \cap C_j) < \rho_0$, $1 \leq j \leq 4$, то теорема доказана.

Если $\text{diam}(M \cap C_j) = \rho_0$ при некотором j , то M содержит ребро T' . Пусть для определенности это — ребро, выходящее из вершины A' . Тогда $M \supset AA'$, так что $\text{diam} M = 2R_0 > \rho_0$, и теорема в этом случае тоже доказана.

ВАЖНОЕ ЗАМЕЧАНИЕ. Леммы 1–3 обобщаются на все размерности $d > 3$, и лишь при обобщении леммы 4 доказательство не проходит уже для $d = 4$.

СПИСОК ЛИТЕРАТУРЫ

- [1] *Болтянский В. Г., Гохберг И. Ц.* Теоремы и задачи комбинаторной геометрии. М.: Наука, 1965.
- [2] *Kahn J., Kalai G.* // Bull. AMS (N. S.) Vol. 29. No 1. 1993. P. 60-62.
- [3] *Nilli A.* // Contemp. Math. Vol. 178. 1994. P. 209-210.
- [4] *Райгородский А.* // УМН. Т. 52. Вып. 6. 1997. С. 181-182.
- [5] *Константинов Н. Н.* // Математическое просвещение. Сер 3, вып. 1, 1997. С. 164-174.
- [6] *Шклярский Д. О., Ченцов Н. Н., Яглом И. М.* Геометрические оценки и задачи из комбинаторной геометрии. М.: Наука, 1974.
- [7] *Райгородский А.* // УМН, 1999. В печати.

n -мерный куб, многочлены и решение проблемы Борсука

А. Б. Скопенков

1. ВВЕДЕНИЕ

В 1993 г. Д. Кан и Г. Калаи [5], следуя идеям Болтянского, Эрдеша и Лармана, построили контрпример к гипотезе Борсука. Они показали, что для некоторых d некоторое подмножество вершин d -мерного куба может быть разбито на части меньшего диаметра, только если количество частей растет вместе с d примерно как $1,2\sqrt{d}$. Это, конечно, больше $d + 1$ для больших d . Конкретно, для $d = 1325$ гипотеза Борсука неверна. Построение Кана и Калаи (см. [5, 8]) было основано на комбинаторной теореме П. Фрэнкла и Р. Вильсона [3]. Израильский математик Алон Нилли разобрался в доказательстве теоремы Фрэнкла – Вильсона. Это позволило ему упростить построение Кана и Калаи, а также снизить наименьшую известную размерность, в которой гипотеза Борсука заведомо неверна, с 1325 до 946 [6]. Дальнейшие результаты в этом направлении были получены А. Гайфуллиным, Д. Гуревичем и А. Райгородским [4, 7].

Цель настоящей статьи — воспроизвести (упрощенное) доказательство Нилли¹⁾. Это доказательство — самое простое из известных сегодня, хотя другие доказательства дают более сильные результаты [4, 7]. Наиболее элементарные шаги сформулированы в виде задач. Звездочкой отмечены задачи, не используемые в дальнейшем.

2. ПОСТРОЕНИЕ КОНТРПРИМЕРА

Начнем с необходимых определений. *Точкой* (или, что то же самое, *вектором*) $x = (x_1, \dots, x_n)$ n -мерного пространства называется упорядоченный набор n чисел. *Расстояние* между точками $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$ определяется формулой $|x, y| = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$. Нам понадобится также *скалярное произведение* векторов $x = (x_1, \dots, x_n)$ и $y = (y_1, \dots, y_n)$, определяемое формулой $(x, y) = x_1 y_1 + \dots + x_n y_n$. Полезность скалярного произведения для изучения расстояний (а, следовательно, и диаметров) иллюстрируется формулой

$$(*) \quad |x, y|^2 = x_1^2 + \dots + x_n^2 + y_1^2 + \dots + y_n^2 - 2(x, y).$$

Пусть $E_2^n = \{(x_1, \dots, x_n) \mid x_i = \pm 1\}$ — множество вершин n -мерного куба. Начиная с этого момента, через x и y будем обозначать вершины n -мерного куба (а

¹⁾ Построение контрпримера к гипотезе Борсука — удивительный пример важного результата в современной математике, не требующего для полного понимания полугодового специального университетского курса (после двухгодичного обязательного курса).

не произвольные точки *n*-мерного пространства). Для вершин *n*-мерного куба формула (*) выглядит особенно просто:

$$(**) \quad |x, y|^2 = 2n - 2(x, y).$$

Контрпример к гипотезе Борсука будет строиться не в E_2^n , а в $E_2^{n^2}$. Вершины n^2 -мерного куба удобно задавать наборами (z_{ij}) , в которых индексы i, j пробегают независимо числа от 1 до n (вместо наборов (z_i) , в которых индекс i пробегает числа от 1 до n^2). Поставим в соответствие каждой вершине $x = (x_1, \dots, x_n)$ из *n*-мерного куба вершину $fx = (x_i x_j)$ из n^2 -мерного куба.

ЗАДАЧА 2.1.а) Найдите $f(1, -1, -1)$ и $f(-1, 1, 1)$.

б) $fx = f(-x)$, где $-(x_1, \dots, x_n) = (-x_1, \dots, -x_n)$.

с) Если $fx = fy$, то $y = \pm x$.

д) Пусть $M' = \{(x_1, \dots, x_n) \in E_2^n \mid x_1 = 1\}$ — $(n - 1)$ -мерная грань *n*-мерного куба. Тогда $fx \neq fy$ для любых двух различных точек $x, y \in M'$.

е) Если $(z_{ij}) = fx$, то $z_{ii} = 1$ и $z_{ij} = z_{ji}$.

Определение отображения f мотивируется следующим его красивым свойством.

ЗАДАЧА 2.2. $(fx, fy) = (x, y)^2$.

Контрпримером к гипотезе Борсука является f -образ множества

$$M = \{x \in E_2^n \mid x_1 = 1 \text{ и среди чисел } x_1, \dots, x_n \text{ число минус единиц четно}\}$$

(для некоторых n , делящихся на 4). Задача 2.1.d иллюстрирует, почему в качестве элементов множества M мы взяли не все вершины *n*-мерного куба, а только вершины x с $x_1 = 1$. Следующая задача поясняет, почему мы взяли n кратным 4, и взяли только вершины $x = (x_1, \dots, x_n)$ с четным числом минус единиц среди x_1, \dots, x_n .

ЗАДАЧА 2.3. Если n делится на 4 и в каждом из наборов $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in E_2^n$ четное число минус единиц, то (x, y) делится на 4.

ЗАДАЧА 2.4. $|M| = 2^{n-2}$, где через $|X|$ обозначается число элементов множества X .

ТЕОРЕМА. Для достаточно большого простого числа p и $n = 4p$ множество fM в n^2 -мерном пространстве нельзя разбить на $n^2 + 1$ часть меньшего диаметра.

3. ИДЕЯ ДОКАЗАТЕЛЬСТВА

Теорема вытекает из Наблюдения, Основной Леммы и Задачи 3.2.

НАБЛЮДЕНИЕ. Если множество fM можно разбить на k частей меньшего диаметра, то M можно разбить на k частей, каждая (одна) из которых не содержит пары ортогональных векторов.

Чтобы сделать Наблюдение, начнем с изучения расстояний между точками множества fM и изучения его диаметра. Из (***) и задачи 2.2 вытекает, что $|fx, fy|^2 = 2n^2 - 2(x, y)^2$. Поэтому $|fx, fy| \leq n\sqrt{2}$ для любых $x, y \in E_2^n$. Более того,

$$(***) \quad |fx, fy| = n\sqrt{2} \text{ тогда и только тогда, когда } (x, y) = 0$$

(такие векторы x, y называются *ортгоналными*). Начиная с этого момента, пусть n чётно. Тогда в M найдётся пара ортгоналных векторов (заметим, хотя это и не используется в дальнейшем, что при нечётном n в M не найдётся пары ортгоналных векторов). Из этого и (***) вытекает, что $\text{diam } fM = n\sqrt{2}$. Более того, расстояние между точками $fx, fy \in fM$ равно $\text{diam } fM$ тогда и только тогда, когда $(x, y) = 0$.

Если множество fM разбито на k частей A_1, \dots, A_k , то их прообразы $f^{-1}A_1, f^{-1}A_2, \dots, f^{-1}A_k$ образуют разбиение множества M на k частей. Из задачи 2.1.d следует, что $|A_i| = |f^{-1}A_i|$. Из (***) и равенства $\text{diam } fM = n\sqrt{2}$ следует, что каждая из частей A_i имеет диаметр, меньший $\text{diam } M$, тогда и только тогда, когда в каждой (одной) из частей $f^{-1}A_i$ никакие два вектора не ортгоналны. Итак, Наблюдение сделано.

Мы хотим доказать, что k велико (точнее, $k > n^2 + 1$). Для этого достаточно доказать, что $|f^{-1}A_i| = |A_i|$ мало (точнее, $|A_i| < \frac{|M|}{n^2 + 1} = \frac{2^{n-2}}{n^2 + 1}$) для каждого $i = 1, \dots, k$. Оказывается, что замеченного свойства множества $f^{-1}A_i$ (никакие два его вектора не ортгоналны) достаточно для нужной верхней оценки на $|f^{-1}A_i|$. Приведем простой пример. Пусть $A \subset M$ и никакие два вектора из A не ортгоналны. Тогда из задач 2.4 и 3.1 следует, что $|A| \leq 2^{n-2} - 1$ (позтому $k \geq 2$). Следующая лемма усиливает эту тривиальную оценку.

ОСНОВНАЯ ЛЕММА. Пусть p — простое (не обязательно большое!), $n = 4p$, $A \subset M$ и никакие два вектора из A не ортгоналны. Тогда

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

ЗАДАЧА 3.2. (Оценка) $\alpha(n) < \frac{n}{4} C_{n-1}^{\frac{n}{4}-1} < \frac{2^{n-2}}{n^2 + 1}$ для достаточно больших n (указание: используйте формулу Стирлинга $\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} n^n e^{-n}} = 1$).

4. МНОГОЧЛЕНЫ И ДОКАЗАТЕЛЬСТВО ОСНОВНОЙ ЛЕММЫ

При доказательстве Основной Леммы можно забыть про конструкцию n^2 -мерного куба и отображения f , зато нужно будет проделать новую конструкцию. Чтобы сформулировать данное свойство множества A (никакие два вектора из A не ортгоналны) на удобном для доказательства языке многочленов, введем определения. Начиная с этого момента, пусть p — простое (не обязательно большое). Положим $G(t) = (t-1)(t-2)\dots(t-p+1)$.

ЗАДАЧА 4.1. Для целых t , $G(t)$ делится на p тогда и только тогда, когда t не делится на p .

Для каждого вектора $a \in A$ определим многочлен от $n-1$ переменной x_2, \dots, x_n формулой $F_a(x_2, \dots, x_n) = G((a, x))$, где $x = (1, x_2, \dots, x_n)$. Раскроем скобки в произведении $G((a, x))$ и в каждом из полученных одночленов будем заменять x_i^2 на 1, пока это возможно. Полученный многочлен обозначим $\tilde{F}_a(x_2, \dots, x_n)$. Он будет *свободен от квадратов*, т.е. будет суммой одночленов $x_{i_1} \dots x_{i_s}$, где i_1, \dots, i_s — различные числа от 2 до n . Его степень не будет превосходить $p-1$ (степенью одночлена $x_1^{b_1} \dots x_s^{b_s}$ называется число $b_1 + \dots + b_s$; степенью многочлена, являющегося непустой суммой различных одночленов с

числовыми коэффициентами, называется максимум степеней одночленов из этой суммы). Пусть A' — множество свободных от квадратов многочленов степени не более $p - 1$. Основная Лемма следует из нижеследующих Леммы 1 и задач 4.2 и 4.5.

ЛЕММА 1. Пусть p — простое, $n = 4p$, $A \subset M$ и никакие два вектора из A не ортогональны. Тогда семейство многочленов $\{\tilde{F}_a(x_2, \dots, x_n)\}_{a \in A}$ линейно независимо.

Многочлен \tilde{F} называется *линейно выражающимся* через многочлены $\tilde{F}_1, \dots, \tilde{F}_s$, если существуют рациональные числа $\lambda_1, \dots, \lambda_s$, такие что $\tilde{F} = \lambda_1 \tilde{F}_1 + \dots + \lambda_s \tilde{F}_s$. Например, многочлен x_2 линейно выражается через многочлены $2x_1$, 1 и $x_1 + x_2$. Семейство многочленов называется *линейно независимым*, если ни один из них не выражается линейно через остальные. Например, семейство из n многочленов $1, x_2, x_3, \dots, x_n$ является линейно независимым.

ЗАДАЧА 4.2. а) Семейство многочленов $x_{i_1} \dots x_{i_s}$, где $s = 0, \dots, p - 1$ и i_1, \dots, i_s — различные числа от 2 до n , является линейно независимым.

б) Любой многочлен из A' линейно выражается через многочлены системы, указанной в а) (такие линейно независимые системы называются *базисами* множества A').

с) В системе из а) ровно $\alpha(n)$ многочленов.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ 1. Пусть, напротив, существуют рациональные числа $\lambda_1, \dots, \lambda_s$, такие что

$$(***) \quad \tilde{F}_a = \lambda_1 \tilde{F}_{a_1} + \dots + \lambda_s \tilde{F}_{a_s}$$

для некоторых $a, a_1, \dots, a_s \in A$. Разберем сначала случай целых $\lambda_1, \dots, \lambda_s$. Подставим в равенство (***) значения $x_2 = a_2, \dots, x_n = a_n$. Из $(a, a) = n = 4p$ и задачи 4.1 вытекает, что левая часть равенства (***) не делится на p . Из задачи 4.1 и нижеследующей задачи 4.3 вытекает, что правая часть равенства (***) делится на p . Полученное противоречие доказывает лемму.

ЗАДАЧА 4.3. Если $a, b \in M$ различны и не ортогональны, то (a, b) не делится на p . (Указание: в противном случае $(a, b) \in \{\pm p, \pm 2p, \pm 3p\}$, что невозможно по задаче 2.3.)

ЗАДАЧА 4.4. Приведите аналогично к противоречию общий случай, когда $\lambda_1, \dots, \lambda_s$ рациональны. (Указание: домножьте равенство (***) на общий знаменатель и используйте метод спуска.)

ЗАДАЧА 4.5. Если $\tilde{F}_1, \dots, \tilde{F}_k$ — линейно независимая система в A' , а Q_1, \dots, Q_s — базис в A' , то $k \leq s$. (Указание: сначала докажите, что систему $\tilde{F}_1, \dots, \tilde{F}_k$ можно дополнить до базиса. Поэтому можно с самого начала считать, что $\tilde{F}_1, \dots, \tilde{F}_k$ — базис. Поскольку \tilde{F}_1 линейно выражается через Q_1, \dots, Q_s , так что не все коэффициенты нулевые, то в системе $\tilde{F}_1, \dots, \tilde{F}_k$ можно заменить \tilde{F}_1 на один из Q_i , так что полученная система тоже будет базисом. Повторяя такие замены несколько раз, мы получим базис Q_{i_1}, \dots, Q_{i_k} . Значит, $k \leq s$.)

БЛАГОДАРНОСТИ

Я хочу поблагодарить Н. П. Долбилина и А. М. Райгородского, от которых я узнал контрпримеры к гипотезе Борсука, учеников физ.-мат. школы им.

А. Н. Колмогорова и школы №57 г. Москвы, которые узнали эти контрпримеры от меня, а также В. Н. Дубровского за полезные обсуждения настоящей заметки.

СПИСОК ЛИТЕРАТУРЫ

- [1] *Болтянский В. Г., Гохберг И. Ц.* Теоремы и задачи комбинаторной геометрии. М.: Наука, 1965.
- [2] *K. Borsuk.* Fund. Math. Vol. 20, 1933. P. 177–190.
- [3] *P. Frankl, R. Wilson.* Combinatorica. Vol. 1, 1981. P. 259–286.
- [4] *М. Л. Гервер.* Математическое Просвещение, сер. 3, вып. 3. М.: МЦНМО, 1999. С. 168 – 183.
- [5] *Kahn J., Kalai G.* // Bull. AMS (N. S.) Vol. 29. No 1. 1993. P. 60–62.
- [6] *Nilli A.* // Contemp. Math. Vol. 178. 1994. P. 209-210.
- [7] *Райгородский А.* // УМН. Т. 52. Вып. 6. 1997. С. 181-182.
- [8] *A. Skopenkov.* Quantum Vol 7, No 1, 1996. P. 16–21, 63.

Критерии гильбертовости банахова пространства, связанные с теорией приближений*

П. А. Бородин

В. М. Тихомиров

0. ВВЕДЕНИЕ

Эллипсоидом называют образ сферы при линейном преобразовании объемлющего пространства. В трехмерном пространстве в подходящей декартовой системе координат эллипсоид имеет уравнение

$$\left(\frac{x_1}{a_1}\right)^2 + \left(\frac{x_2}{a_2}\right)^2 + \left(\frac{x_3}{a_3}\right)^2 = 1.$$

Эллипсоиды выделяются из всех выпуклых центрально-симметричных поверхностей многими уникальными свойствами. Приведем некоторые из них.

1. Если все плоские сечения выпуклой центрально-симметричной поверхности суть эллипсы, то поверхность — эллипсоид.
2. Если при освещении выпуклой центрально-симметричной поверхности произвольным параллельным пучком света она отбрасывает на плоскость тень в виде внутренности эллипса, то поверхность — эллипсоид.
3. Если при освещении выпуклой центрально-симметричной поверхности произвольным параллельным пучком света *граница тени* (точки поверхности, в которых лучи света ее касаются) плоская, то поверхность — эллипсоид.
4. Если для всякого плоского сечения выпуклой центрально-симметричной поверхности найдется направление параллельного освещения, при котором граница тени будет совпадать с этим плоским сечением, то поверхность — эллипсоид.

Ниже мы сформулируем эти утверждения более тщательно, а некоторые из них докажем.

Знаменитый немецкий математик Герман Минковский по каждому выпуклому ограниченному центрально-симметричному множеству в конечномерном пространстве построил метрику, с помощью которой можно мерить расстояние между точками, а американец Винер (изобретатель кибернетики) и поляк Банах распространили определение Минковского на бесконечномерные пространства. При этом изначальное множество является единичным шаром, расстояния от точек которого до начала координат не превосходят единицы. Эти пространства с введенной в них метрикой (и еще одним дополнительным свойством полноты)

*Работа обоих авторов выполнена при поддержке РФФИ (первого — проект № 96-01-01366, второго — проект № 96-01-00325 и проект Программы поддержки вед. научных школ № 96-15-96072).

называются банаховыми пространствами. Среди банаховых пространств особую, уникальную роль играют пространства, у которых в конечномерном случае границами единичных шаров являются эллипсоиды. Они называются гильбертовыми пространствами.

Наличие метрики позволяет ставить задачи о приближении элементов множествами. Приближения в гильбертовом пространстве устроены особенно совершенно. В данной работе показывается, что эти совершенные свойства приближений характеризуют гильбертовы пространства.

Эту статью мы постарались сделать «замкнутой в себе». Вначале делается экскурс в теорию банаховых и гильбертовых пространств, а затем доказываются критерии гильбертовости банахова пространства (в простейшем случае — критерии эллипсоидности единичной сферы этого пространства), связанные с теорией приближений. Излагаются все необходимые понятия и факты теории нормированных пространств, теории приближений и геометрии, так что статья доступна студентам-математикам 1 курса и подготовленным школьникам старших классов.

1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

В линейном пространстве X (пространстве, элементы которого можно складывать и умножать на вещественные числа) часто бывает необходимо ввести метрику, то есть расстояние $d(x, y)$ между элементами $x, y \in X$. При этом естественно потребовать, чтобы это расстояние не менялось при сдвиге элементов на один и тот же вектор: $d(x + z, y + z) = d(x, y)$. Поэтому достаточно определить расстояние от каждого элемента $x \in X$ до нулевого элемента, или *норму* элемента x : $\|x\| = d(x, 0)$. Эта норма, то есть функция $\|\cdot\| : X \rightarrow \mathbf{R}$, должна обладать следующими свойствами:

- а) $\forall x \in X \quad \|x\| \geq 0; \|x\| = 0 \iff x = 0$ (свойство неотрицательности);
- б) $\forall x \in X \forall \lambda \in \mathbf{R} \quad \|\lambda x\| = |\lambda| \cdot \|x\|$ (свойство однородности);
- в) $\forall x \in X \forall y \in X \quad \|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника).

Если теперь положить $d(x, y) := \|x - y\|$, то метрика d будет обладать всеми требуемыми свойствами: неотрицательностью ($d(x, y) \geq 0$ и $d(x, y) = 0 \iff x = y$), симметричностью ($d(x, y) = d(y, x)$) и неравенством треугольника ($d(x, y) \leq d(x, z) + d(y, z)$), так что (X, d) , по определению, — *метрическое* пространство. Кроме того, для введенной таким образом метрики справедливы равенства $d(x + z, y + z) = d(x, y)$ и $d(\lambda x, \lambda y) = |\lambda|d(x, y)$, естественные в линейном пространстве. Пара $(X, \|\cdot\|)$ называется *нормированным* пространством.

В одном и том же линейном пространстве X норму можно ввести многими разными способами. Приведем несколько примеров норм в обычном n -мерном вещественном пространстве \mathbf{R}^n , состоящем из векторов $x = (x_1, \dots, x_n)$ с координатами $x_k \in \mathbf{R}$:

- I. $\|x\| = (\sum_{k=1}^n x_k^2)^{1/2}$;
- II. $\|x\| = \sum_{k=1}^n |x_k|$;
- III. $\|x\| = \max_{1 \leq k \leq n} |x_k|$.

Можно рассматривать и бесконечномерные линейные пространства. Так, например, в бесконечномерном линейном пространстве $C[0, 1]$ вещественных функций f , непрерывных на отрезке $[0, 1]$, норму также можно ввести разными способами:

$$\text{IV. } \|f\| = \left(\int_0^1 f^2(x) dx\right)^{1/2};$$

$$\text{V. } \|f\| = \int_0^1 |f(x)| dx;$$

$$\text{VI. } \|f\| = \max_{x \in [0, 1]} |f(x)| \text{ (равномерная, или чебышевская норма).}$$

Проверка свойств а), б), в) нормы в каждом из примеров I – VI предоставляется читателю.

Единичный шар $B = \{x \in X : \|x\| \leq 1\}$ линейного нормированного пространства X всегда является *выпуклым центрально-симметричным* (относительно нуля) *поглощающим* ($\forall x \in X \exists \lambda > 0 : \lambda x \in B$) *ограниченным* ($\forall x \in X \exists \lambda > 0 : \lambda x \notin B$) телом в X . (Выпуклость следует из того, что если $x, y \in B$, то для $0 \leq \alpha \leq 1$ имеем

$$\|\alpha x + (1 - \alpha)y\| \leq^a \|\alpha x\| + \|(1 - \alpha)y\| =^b \alpha \|x\| + (1 - \alpha)\|y\| \leq \alpha + (1 - \alpha) = 1,$$

т. е. весь отрезок с концами x, y целиком лежит в B ; остальные из перечисленных свойств достаточно очевидны.) Например, в случае нормы из примера I и прямоугольной системы координат в $X = \mathbf{R}^n$ единичный шар совпадает с обычным шаром с центром в 0 и радиусом 1 (если система координат произвольна, то B является эллипсоидом). Единичные шары для норм II и III — многогранники; в частном случае $n = 3$ и прямоугольной системы координат в $X = \mathbf{R}^3$ норма II дает правильный октаэдр с вершинами на осях координат, а норма III — куб с ребрами длины 2, параллельными координатным осям.

Верно и обратное (об этом говорилось во введении): *каждое выпуклое центрально-симметричное поглощающее ограниченное множество B в линейном пространстве X задает в X норму $\|\cdot\|_B$, в которой оно является «почти» единичным шаром: $\{x : \|x\|_B < 1\} \subseteq B \subseteq \{x : \|x\|_B \leq 1\}$. Именно, надо положить $\|x\|_B = \inf\{\lambda > 0 : \frac{x}{\lambda} \in B\}$. Проверка свойств нормы а), б), в) и приведенного двойного включения предоставляется читателю.*

Линейное нормированное пространство $(X, \|\cdot\|)$ называется *полным*, если оно содержит все свои предельные точки (т. е. если для всякой фундаментальной последовательности $\{x_n\} \in X$, т. е. такой последовательности, что $\|x_n - x_m\| \rightarrow 0$ при $n, m \rightarrow \infty$, существует элемент $x \in X$, к которому эта последовательность сходится: $\|x_n - x\| \rightarrow 0$). Конечномерное нормированное пространство всегда полно, а бесконечномерное — необязательно. Например, пространство $C[0, 1]$ полно относительно нормы VI, но не полно относительно норм IV или V (проверьте!). Любое неполное пространство X можно «пополнить», добавив предельные точки, или, другими словами, вложить в полное пространство \bar{X} с сохранением нормы, так что X плотно в \bar{X} (см., напр., [1, гл. II, § 3]).

Полные нормированные пространства называются *банаховыми*, в честь выдающегося польского математика С. Банаха. В дальнейшем мы будем рассматривать только банаховы пространства.

Норма в линейном пространстве X обладает особенно хорошими свойствами, если она порождена скалярным произведением. *Скалярное произведение* — это функция $(\cdot, \cdot) : X \times X \rightarrow \mathbf{R}$, удовлетворяющая следующим требованиям:

- i) $\forall x \in X \quad (x, x) \geq 0$ и $(x, x) = 0 \iff x = 0$ (свойство неотрицательности);
- ii) $\forall x, y \in X \quad (x, y) = (y, x)$ (свойство симметричности);
- iii) $\forall x, y, z \in X \quad \forall \alpha, \beta \in \mathbf{R} \quad (\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ (свойство билинейности).

Скалярное произведение порождает норму $\|x\| = \sqrt{(x, x)}$ (проверка свойств а), б), в) нормы предоставляются читателю). Так, норма из примера I порождена стандартным скалярным произведением в \mathbf{R}^n : $(x, y) = \sum_{k=1}^n x_k y_k$, а норма из примера IV — скалярным произведением функций: $(f, g) = \int_0^1 f(x)g(x)dx$. Любое скалярное произведение на плоскости \mathbf{R}^2 задается так: если $x = (x_1, x_2)$, $y = (y_1, y_2)$, то $(x, y) = a_{11}x_1y_1 + a_{12}x_1y_2 + a_{21}x_2y_1 + a_{22}x_2y_2$, где $a_{12} = a_{21}$ и (необходимое и достаточное условие неотрицательности) $a_{11} > 0$, $a_{11}a_{22} - a_{12}^2 > 0$. При этом единичный шар, т. е. совокупность тех $x \in \mathbf{R}^2$, для которых $(x, x) \leq 1$, — внутренность эллипса. Вообще, любое скалярное произведение в \mathbf{R}^n задается положительно определенной симметрической билинейной формой $A(x, y) = \sum_{i,j=1}^n a_{ij}x_iy_j$ ($a_{ij} = a_{ji} \in \mathbf{R}$, $\forall x \neq 0 \quad A(x, x) > 0$). Имеются критерии положительной определенности формы A — например, в терминах главных миноров матрицы a_{ij} (критерий Сильвестра — см., напр., [2]). Единичный шар в соответствующей норме — $\{x \in \mathbf{R}^n : \sum a_{ij}x_ix_j \leq 1\}$ — эллипсоид с центром в 0, вместе со своей внутренностью. Обратно, любой эллипсоид с центром в нуле задает норму, порожденную некоторым скалярным произведением.

Банахово пространство, в котором норма задана скалярным произведением, называется *гильбертовым пространством*, в честь великого немецкого математика Д. Гильберта. Норма в гильбертовом пространстве обладает многими замечательными свойствами, которые зачастую являются характеристическими для гильбертовости. Приведем классический пример такого свойства.

ЛЕММА А (П. ЙОРДАН, ДЖ. ФОН НЕЙМАНН, [3]). *Банахово пространство X является гильбертовым тогда и только тогда, когда для любых двух векторов $x, y \in X$ выполнено равенство параллелограмма:*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

(сумма квадратов диагоналей параллелограмма равна сумме квадратов его сторон).

ДОКАЗАТЕЛЬСТВО. Выполнение равенства параллелограмма в гильбертовом пространстве проверяется непосредственно.

Обратно, пусть в банаховом пространстве выполнено равенство параллелограмма. Положим

$$(x, y) := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

и покажем, что эта функция двух переменных обладает всеми свойствами скалярного произведения. Свойства а) и б) очевидны. Для доказательства с) установим вначале аддитивность:

$$(x + y, z) = (x, z) + (y, z),$$

которая в нашем случае равносильна равенству

$$\|x + y + z\|^2 - \|x + y - z\|^2 = \|x + z\|^2 - \|x - z\|^2 + \|y + z\|^2 - \|y - z\|^2. \quad (1)$$

Запишем равенства параллелограмма для параллелограммов, натянутых на пары векторов $x + z$ и y , $y + z$ и x , $x - z$ и y , $y - z$ и x :

$$\begin{aligned}\|x + z + y\|^2 + \|x + z - y\|^2 &= 2(\|x + z\|^2 + \|y\|^2), \\ \|y + z + x\|^2 + \|y + z - x\|^2 &= 2(\|y + z\|^2 + \|x\|^2), \\ \|x - z + y\|^2 + \|x - z - y\|^2 &= 2(\|x - z\|^2 + \|y\|^2), \\ \|y - z + x\|^2 + \|y - z - x\|^2 &= 2(\|y - z\|^2 + \|x\|^2).\end{aligned}$$

Вычитая из суммы первых двух равенств сумму последних двух равенств, получаем (1).

Оставшееся свойство однородности $(\lambda x, y) = \lambda(x, y)$ получается так: для целых λ оно следует из аддитивности; от целых λ легко перейти к рациональным λ , а от рациональных, по непрерывности, ко всем $\lambda \in \mathbf{R}$.

Теперь нетрудно проверить, что введенное таким образом скалярное произведение (x, y) порождает исходную норму. Лемма А доказана.

СЛЕДСТВИЕ. Если каждое двумерное (или каждое n -мерное, $n \geq 2$) подпространство банахова пространства X гильбертово, то и само X гильбертово.

Отсюда вытекает одна из упомянутых во введении характеристик эллипсоида: если любое плоское сечение выпуклой центрально-симметричной поверхности есть эллипс, то сама поверхность — эллипсоид.

С помощью леммы А легко показать, что пространства из примеров II, III, V, VI гильбертовыми не являются.

Гильбертовы пространства изучаются и применяются в самых разных областях математики, и очень часто возникают утверждения, подобные лемме А, — когда какое-либо свойство гильбертова пространства оказывается характеристическим и кладется в основу критерия гильбертовости банахова пространства. Доказательству некоторых из этих критериев — именно, критериев, возникающих из теории приближений, — и посвящена настоящая работа.

Несколько слов об истории. Скалярное произведение (под названием inner product — «внутреннее произведение», это название сохранилось в англоязычной литературе до сих пор) ввел Гамильтон в 1853 г. Гильбертово пространство было определено в начале XX века в работах Гильберта и Шмидта. Последний в 1908 г. ввел обозначение $\|\cdot\|$ для нормы. Нормированные пространства были определены Банахом и Винером в 1922 г. Гильбертовы пространства выделяются среди банаховых пространств не только богатством свойств, но и «одинаковостью»: любые два гильбертовых пространства одинаковой размерности изометрически изоморфны друг другу (т. е. между ними существует линейное взаимно-однозначное соответствие, сохраняющее норму). Так, любое n -мерное вещественное пространство со скалярным произведением $\sum a_{ij}x_iy_j$ изометрически изоморфно стандартному (евклидову) пространству \mathbf{R}^n со скалярным произведением $\sum x_k y_k$ (дело в том, что любую положительно определенную симметрическую билинейную форму можно привести к диагональному виду линейной заменой переменных — см., напр., [2]). Любое бесконечномерное сепарабельное (т. е. обладающее счетным всюду плотным подмножеством) гильбертово пространство — в частности, пополнение пространства $C[0, 1]$ по норме IV — изометрически изоморфно пространству l_2 , состоящему из векторов (x_1, x_2, \dots) со

счетным числом координат, удовлетворяющих неравенству $\sum_{k=1}^{\infty} x_k^2 < \infty$ (см., напр., [1, гл. III, § 4]). Скалярное произведение в этом пространстве задается так: $(x, y) = \sum_{k=1}^{\infty} x_k y_k$. Такая уникальность гильбертова пространства дает еще один повод к отысканию критериев гильбертовости.

Нельзя не упомянуть еще об одном банаховом пространстве, которое по своей значимости, быть может, сравнимо с гильбертовыми, — именно, о пространстве $C[0, 1]$ непрерывных функций с равномерной нормой (пример VI). Дело в том, что *любое* сепарабельное банахово пространство изометрически изоморфно некоторому (своему) подпространству пространства $C[0, 1]$ (теорема Банаха – Мазура). Такая универсальность пространства $C[0, 1]$ делает его противоположным гильбертову пространству по многим свойствам (см. замечание в конце следующего пункта).

В заключение этой вводной части приведем точную формулировку двух геометрических характеристик эллипсоидов в трехмерном пространстве, о которых говорилось во введении. На них мы будем опираться далее при доказательстве критериев гильбертовости.

ТЕОРЕМА А. Пусть в \mathbf{R}^3 задана двумерная выпуклая замкнутая поверхность S , которая касается каждого описанного вокруг нее цилиндра по множеству, лежащему в двумерной плоскости. Тогда S — эллипсоид (т. е. задается в некоторой системе координат как множество уровня положительно определенной квадратичной формы).

Эта теорема обычно называется теоремой Бляшке об эллипсоиде. Сам В. Бляшке [4] доказал ее для так называемого овалоида (регулярной аналитической замкнутой выпуклой поверхности с везде отличной от нуля кривизной). В общем случае теорема А доказана А. Д. Александровым [5]. Если интерпретировать множество, по которому поверхность S пересекается с описанным вокруг нее цилиндром, как границу «тени» S при освещении параллельным пучком света, направленным по образующей этого цилиндра, то теорема А будет звучать так: *всякая поверхность с плоскими границами теней — эллипсоид* (этой аналогией и объясняется название статьи [5]). Верно и обратное: любой эллипсоид имеет плоские границы теней (докажите!).

ТЕОРЕМА В. Пусть в \mathbf{R}^3 задана двумерная выпуклая замкнутая поверхность S , симметричная относительно начала координат O . Если для каждой двумерной плоскости Γ , проходящей через O , существует направление $\tau(\Gamma)$, опорное к S во всех точках пересечения $S \cap \Gamma$, то S — эллипсоид.

Эту теорему обычно называют теоремой Бляшке – Какутани. Сформулировавший ее С. Какутани [6] не дал полного доказательства, лишь указав на возможность сведения этой теоремы к теореме А. Доказательство теоремы В содержится в [7, гл. 5]. Вывод теоремы В из теоремы А осуществлен в [8].

2. МЕТРИЧЕСКАЯ ПРОЕКЦИЯ И ГИЛЬБЕРТОВОСТЬ

Пусть $(X, \|\cdot\|)$ — банахово пространство, Y — некоторое подмножество X , $\rho(x, Y) = \inf_{y \in Y} \|x - y\|$ — расстояние от элемента $x \in X$ до множества Y , $P_Y(x) = \{y \in Y : \|x - y\| = \rho(x, Y)\}$ — метрическая проекция x на Y , т. е. множество ближайших к x элементов в Y (элементов наилучшего приближения). Для

данного x метрическая проекция $P_Y(x)$ может не содержать ни одного элемента или содержать более одного элемента, а отображение $P_Y : x \rightarrow P_Y(x)$, называемое *оператором метрического проектирования*, вообще говоря, определено не на всем X , многозначно и даже разрывно.

Если же X — гильбертово пространство, а Y — его замкнутое (т. е. полное относительно нормы) линейное подпространство (вообще, везде ниже слово подпространство будет обозначать замкнутое линейное подпространство), то свойства P_Y предельно совершенны. В этом случае, как мы сейчас покажем, оператор P_Y *определен на всем X , однозначен, линеен и имеет норму 1*. Действительно, пусть $x \in X$, $\rho = \rho(x, Y)$ и последовательность $\{y_n\}$ такова, что $\|x - y_n\| \rightarrow \rho$. По равенству параллелограмма,

$$\begin{aligned} \|y_n - y_m\|^2 &= 2(\|x - y_n\|^2 + \|x - y_m\|^2) - \|2x - y_n - y_m\|^2 = \\ &= 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4\|x - \frac{1}{2}(y_n + y_m)\|^2 \leq \\ &\leq 2(\|x - y_n\|^2 + \|x - y_m\|^2) - 4\rho^2 \rightarrow 0 \quad \text{при } n, m \rightarrow \infty, \end{aligned}$$

так что последовательность y_n сходится по норме к некоторому элементу $y \in Y$ (подпространство Y замкнуто). Очевидно, $\|x - y\| = \rho$, т. е. $y \in P_Y(x)$. Кроме того, для любого $y' \in Y$ и любого числа λ имеем

$$\|x - y + \lambda y'\|^2 = \|x - y\|^2 + |\lambda|^2 \|y'\|^2 + 2\lambda(x - y, y') \geq \rho^2 = \|x - y\|^2.$$

При малых λ это неравенство может выполняться лишь в том случае, если $(x - y, y') = 0$. Отсюда $\|x - y + \lambda y'\| > \|x - y\| = \rho$ при всех $y' \neq 0$ и $\lambda \neq 0$. Итак, метрическая проекция $P_Y(x)$ состоит ровно из одного элемента y , причем $\forall y' \in Y \quad (x - y, y') = 0$. Положим $Y^\perp = \{z \in X : \forall y' \in Y \quad (z, y') = 0\}$ — ортогональное дополнение к подпространству Y , являющееся, очевидно, подпространством. Как мы только что показали, $X = Y \oplus Y^\perp$, т. е. любой элемент $x \in X$ однозначно разлагается в сумму элемента $y = P_Y(x) \in Y$ и элемента $x - y \in Y^\perp$. Линейность отображения $x \mapsto y = P_Y(x)$ следует из линейности Y^\perp . Кроме того, оператор P_Y имеет норму 1: $\|P_Y\| := \sup_{\|x\|=1} \|P_Y(x)\| = 1$, поскольку $1 = \|x\|^2 = \|x - y + y\|^2 = \|x - y\|^2 + \|y\|^2 + 2(x - y, y) = \|x - y\|^2 + \|y\|^2 \geq \|y\|^2$, а с другой стороны, если $x \in Y$, то $P_Y(x) = x$ и $\|P_Y(x)\| = \|x\|$.

Итак, в данном случае оператор метрического проектирования $P_Y : X \rightarrow Y$ совпадает с оператором ортогонального проектирования на подпространство Y , однозначен, линеен и имеет норму 1. (Конечно, в нашем трехмерном пространстве это полностью соответствует нашим представлениям: ортогональная проекция точки x на плоскость Y является единственной ближайшей к x точкой в Y .) Возникает естественный вопрос: *не является ли какое-нибудь из этих свойств — скажем, линейность оператора P_Y или существование линейного проектора $X \rightarrow Y$ нормы 1 — характеристическим для гильбертова пространства? Или вдруг существуют другие банаховы пространства с такими замечательными свойствами?*

Рассмотрим трехмерное гильбертово пространство $X = \mathbf{R}^3$, норма в котором задается единичным шаром — некоторым эллипсоидом с центром в начале координат. Для любого одномерного подпространства $Y \subset X$, порождаемого некоторым вектором v , оператор метрического проектирования P_Y линеен. Это означает, что множество $\text{Ker}(P_Y) = \{q \in X : P_Y(q) = 0\} = Y^\perp$ — двумерное

(линейное) подпространство. В то же время, $\text{Ker}(P_Y)$ — это в точности те элементы $q \in X$, для которых $\|q + \lambda v\| \geq \|q\|$ при любом числе λ , т. е. те точки q , в которых параллельные Y прямые являются опорными к сфере $\{x \in X : \|x\| = \|q\|\}$. Вот мы и доказали, что любой цилиндр, описанный вокруг эллипсоида, касается его по плоской кривой. Но по теореме А из п. 1 это есть характеристическое свойство эллипсоида! Таким образом, если в трехмерном банаховом пространстве оператор метрического проектирования на каждое одномерное подпространство линеен, то единичная сфера в этом пространстве удовлетворяет условиям теоремы А, а значит, это пространство гильбертово.

Ниже мы покажем, что существование линейного проектора нормы 1 на любое двумерное подпространство в трехмерном банаховом пространстве X равносильно тому, что единичная сфера $S(X)$ удовлетворяет условиям теоремы В — т. е. опять же, равносильно гильбертовости X .

Какие же аналогичные условия будут критериями гильбертовости в банаховом пространстве произвольной размерности?

Один из авторов этой статьи (В. М. Тихомиров) в начале 60-х годов столкнулся с такой проблемой, поставленной С. М. Никольским на конгрессе в Амстердаме ([17, стр. 263]). Пусть X — банахово пространство, Y — подпространство в X , C — подмножество в X . Определим *уклонение* C от Y в X : $d(C, Y, X) = \sup_{x \in C} \rho(x, Y)$. Если X гильбертово, то $\rho(x, Y) = \|x - P_Y(x)\|$, и, следовательно, уклонение $d(C, Y, X)$ равно верхней грани $\sup_{x \in C} \|x - P_Y(x)\|$, где P_Y — линейный оператор метрического проектирования на Y . В других пространствах, не гильбертовых, дело обстоит сложнее, так как ближайший к x элемент пространства Y находится при помощи, вообще говоря, нелинейной операции. И здесь возникает проблема (пишет С. М. Никольский): нельзя ли все же построить такой линейный оператор Λ , который отображал бы множество C в подпространство Y так, чтобы $d(C, Y, X) = \sup_{x \in C} \|x - \Lambda(x)\|$? Конечно, для отдельной пары C, Y такой оператор может найтись и в негильбертовом пространстве X — например, в случае, когда оператор P_Y допускает *линейную выборку* (т. е. существует такой линейный оператор $P : X \rightarrow Y$, что $P(x) \in P_Y(x)$ для любого $x \in X$). Скажем, в пространстве $C[-1, 1]$ с равномерной нормой (оно, как мы знаем, не гильбертово) существует линейная выборка из оператора метрического проектирования на подпространство всех четных функций из $C[-1, 1]$: $f(x) \mapsto \frac{f(x) + f(-x)}{2}$ (проверьте!). Поэтому вопрос С. М. Никольского надо уточнить: *как велика в негильбертовом пространстве X может быть совокупность пар $\{C, Y\}$, для каждой из которых найдется (свой) оператор Λ с указанными свойствами?*

На эти и некоторые другие вопросы частично отвечает

ТЕОРЕМА 1. Пусть X — вещественное банахово пространство, $\dim X \geq 3$. Следующие условия эквивалентны:

0) X — гильбертово;

1) существует такое число $n \in \mathbb{N}$, $n \leq \dim X - 2$, что для любого n -мерного подпространства $Y_n \subset X$ и любого подмножества $C \subset X$ найдется линейный оператор $\Lambda : X \rightarrow Y_n$, для которого $\sup_{x \in C} \|x - \Lambda(x)\| = d(C, Y_n, X)$;

- 2) существует такое число $n \in \mathbf{N}$, $n \leq \dim X - 2$, что для любого n -мерного подпространства $Y_n \subset X$ оператор метрического проектирования P_{Y_n} допускает линейную выборку;
- 3) существует такое число $k \in \mathbf{N}$, $2 \leq k \leq \dim X - 1$, что для любого k -мерного подпространства $Y_k \subset X$ найдется линейный проектор $\pi_{Y_k} : X \rightarrow Y_k$ нормы 1.

В этой теореме утверждение о равносильности 0) и 1) частично отвечает на вопрос С. М. Никольского и вытекает, как мы увидим, из эквивалентности 0) \iff 2), доказанной в [8] (соответствующая теорема 3 в [8] формулируется в других терминах). До этого Рудин и Смит [9] нашли необходимое и достаточное условие гильбертовости более сильное, чем 2): они требовали, чтобы сам оператор P_{Y_n} был однозначным и линейным для каждого n -мерного подпространства $Y_n \subset X$. Утверждение о равносильности 0) и 3) доказано в [8] и обобщает теорему Какутани [6], в которой требовалось существование проектора π_{Y_n} нормы 1 на любое подпространство $Y \subset X$. Для $k = 2$ эквивалентность 0) \iff 3) была доказана Филлипсом [10]. В работе [8] содержатся также обобщения этих критериев и близкие им утверждения.

ДОКАЗАТЕЛЬСТВО. Выполнение каждого из условий 1), 2) и 3) в гильбертовом пространстве уже отмечалось выше. Докажем обратные утверждения по следующей схеме: 1) \implies 2) \implies 0) и 3) \implies 0).

1) \implies 2). Возьмем произвольное n -мерное подпространство $Y_n \subset X$ и положим $C = Y_n + B$, где $B = \{x \in X : \|x\| \leq 1\}$ — единичный шар пространства X . Согласно 1), $1 = d(C, Y_n, X) = \sup_{x \in C} \|x - \Lambda(x)\|$, где $\Lambda : X \rightarrow Y_n$ — некоторый ограниченный линейный оператор. Докажем, что Λ является выборкой из оператора P_{Y_n} метрического проектирования на подпространство Y_n . Возьмем любой элемент $x \in X$, какой-нибудь его элемент наилучшего приближения $y \in P_{Y_n}(x)$ и положим $q = x - y$. Элемент $x' = x/\|q\| = q/\|q\| + y/\|q\| = q' + y'$ принадлежит множеству C , а значит, $1 \geq \|x' - \Lambda(x')\| = \|q' + (y' - \Lambda(x'))\|$. Нулевой элемент 0 является в Y_n элементом наилучшего приближения для q , а значит, и для q' , поэтому $\|q' + (y' - \Lambda(x'))\| \geq \|q' - 0\| = \|q'\| = 1$. Вместе с предыдущим неравенством это дает $\|x' - \Lambda(x')\| = 1$, откуда $\|x - \Lambda(x)\| = \|q\| \|x' - \Lambda(x')\| = \|q\| = \rho(x, Y_n)$, т. е. $\Lambda(x) \in P_{Y_n}(x)$.

Итак, для любого n -мерного подпространства $Y_n \subset X$ мы нашли линейную выборку из оператора P_{Y_n} , так что X удовлетворяет условию 2) при данном n .

2) \implies 0). Можно считать, что $\dim X = n + 2$. Действительно, в противном случае возьмем произвольное подпространство $X_{n+2} \subset X$ размерности $n+2$. Оно также будет удовлетворять условию 2). Доказав гильбертовость любого такого X_{n+2} , мы в силу следствия из леммы А докажем и гильбертовость исходного пространства X .

Итак, считаем $\dim X = n + 2$. Возьмем произвольное подпространство $Z \subset X$ размерности $n - 1$. В факторпространстве X/Z , состоящем из классов смежности $[x] = \{x + z : z \in Z\}$, как обычно (см., напр., [1], гл. III, § 3), введем норму $\|[x]\| = \inf_{z \in Z} \|x - z\|$ (доказательство того, что это действительно норма, предоставляется читателю). Любое одномерное подпространство в X/Z есть Y/Z для некоторого подпространства $Y \subset X$, $\dim Y = n$. По условию, существует линейный проектор $P : X \rightarrow Y$, сопоставляющий каждому $x \in X$ какой-то

элемент наилучшего приближения $P(x) \in P_Y(x)$. Покажем, что возникающий линейный проектор $\tilde{P} : X/Z \rightarrow Y/Z$, $\tilde{P}([x]) = [P(x)]$, обладает тем же свойством в факторпространстве X/Z . Действительно, для любого $[y] \in Y/Z$ имеем

$$\begin{aligned} \|[x] - \tilde{P}([x]) - [y]\|_{X/Z} &= \inf_{z \in Z} \|x - P(x) - y - z\| \geq \inf_{z \in Z} \|x - P(x)\| = \|x - P(x)\| = \\ &= \inf_{z \in Z} \|x - P(x) - z\| = \|[x] - [P(x)]\|_{X/Z} = \|[x] - \tilde{P}([x])\|_{X/Z}. \end{aligned}$$

Таким образом, для любого $(n-1)$ -мерного подпространства $Z \subset X$ трехмерное пространство $\tilde{X} = X/Z$ удовлетворяет 2) с $n=1$. Если мы докажем, что каждое такое \tilde{X} гильбертово, то гильбертовость X будет вытекать из следующей леммы.

ЛЕММА В ([8]). Пусть X — конечномерное банахово пространство, и существует такое натуральное число $d \leq \dim X - 2$, что для любого d -мерного подпространства $Z \subset X$ факторпространство X/Z гильбертово. Тогда гильбертовым является и само X .

ДОКАЗАТЕЛЬСТВО ЛЕММЫ В. Достаточно показать, что для любых двух элементов $f, g \in X$ выполняется равенство параллелограмма. Возьмем в X гиперплоскость F , опорную к шару $\{x \in X : \|x\| \leq \|f\|\}$ в точке f , и аналогичную гиперплоскость G для элемента g . В пересечении $F \cap G$ выберем d -мерную плоскость, и пусть Z — d -мерное подпространство, параллельное этой плоскости. Плоскость $Z + f$ содержится в плоскости F , т. е. является опорной к шару $\{x : \|x\| \leq \|f\|\}$ в точке f . Это означает, что $\|f+z\| \geq \|f\|$ для любого $z \in Z$. Аналогично, $\|g+z\| \geq \|g\|$ для любого $z \in Z$. Поэтому в факторпространстве X/Z нормы классов смежности $[f]$ и $[g]$ совпадают с нормами исходных элементов f и g : $\|[f]\|_{X/Z} = \|f\|$, $\|[g]\|_{X/Z} = \|g\|$.

Пространство X/Z — гильбертово, поэтому

$$\|[f] + [g]\|^2 + \|[f] - [g]\|^2 = 2(\|[f]\|^2 + \|[g]\|^2)$$

или

$$\inf_{z \in Z} \{\|f + g + z\|^2\} + \inf_{z \in Z} \{\|f - g + z\|^2\} = 2(\|f\|^2 + \|g\|^2).$$

Отсюда

$$\|f + g\|^2 + \|f - g\|^2 \geq 2(\|f\|^2 + \|g\|^2) \quad \forall f, g \in X.$$

Запишем последнее неравенство с заменой f и g , соответственно на $(f+g)/2$ и $(f-g)/2$:

$$\|f\|^2 + \|g\|^2 \geq \frac{\|f+g\|^2 + \|f-g\|^2}{2}.$$

Последние два неравенства вместе дают искомое равенство параллелограмма. Лемма В доказана.

Геометрически лемму В можно интерпретировать так: если выпуклая центрально-симметричная поверхность $S \subset \mathbf{R}^n$ при проектировании вдоль любого d -мерного подпространства L ($d \leq n-2$ — заданное натуральное число) на некоторую $(n-d)$ -мерную плоскость $\Pi = \Pi(L)$ дает эллипсоид в Π , то она сама является эллипсоидом (ср. с одной из приведенных во введении характеристик эллипсоидов в \mathbf{R}^3).

Вернемся к доказательству теоремы. Осталось доказать гильбертовость трехмерного пространства \tilde{X} , удовлетворяющего условию 2) с $n = 1$. Докажем, что единичная сфера S этого пространства — эллипсоид. Возьмем произвольное одномерное подпространство $Y \subset \tilde{X}$ и соответствующий проектор $P : \tilde{X} \rightarrow Y$, $\forall x \in \tilde{X} \quad P(x) \in P_Y(x)$. Ядро этого проектора — некоторое двумерное подпространство $U = U(Y)$. Для каждого элемента $u \in U$ одним из ближайших элементов в Y является 0, т. е. $\|u + y\| \geq \|u + 0\| = \|u\|$ при любом $y \in Y$. Это означает, что прямая, параллельная Y и проходящая через u , является опорной к сфере $\{x : \|x\| = \|u\|\}$ в точке u . В частности, пересечение единичной сферы S и описанного вокруг нее цилиндра $C(Y)$ с образующей Y целиком содержит кривую $S \cap U$, т. е. $(S \cap C(Y)) \supseteq (S \cap U)$. Для окончательного приведения к условиям теоремы А (см. п. 1) надо доказать, что $(S \cap C(Y)) = (S \cap U)$.

Пусть, вопреки этому, прямая l параллельна подпространству Y и является опорной к сфере S в точке $p \notin U$. Эта прямая пересекается с U в некоторой точке q . Как мы знаем, l является опорной к сфере $\{x : \|x\| = \|q\|\}$ в точке q , и в то же время, по предположению, она является опорной к $S = \{x : \|x\| = 1\}$ в точке p . Отсюда $\|q\| = 1$, отрезок $pq \subset S$ (единичный шар — выпуклое множество), а прямая l является опорной к S во всех точках этого отрезка.

Пусть Π — двумерная плоскость, опорная к S во всех точках отрезка pq , а Γ — двумерное подпространство, порожденное векторами p и q . Существует одномерное подпространство $Y' \neq Y$, параллельное Π и такое, что соответствующее ему подпространство U' (такое, что $(S \cap C(Y')) \supseteq (S \cap U')$) не совпадает с Γ . Действительно, если $(S \cap \Gamma) \subseteq (S \cap C(Y'))$ для любого $Y' \parallel \Pi$, $Y' \neq Y$, то по непрерывности $(S \cap \Gamma) \subseteq (S \cap C(Y))$, что невозможно, так как $Y \subset \Gamma$ (отрезок $pq \parallel Y$). Таким образом, прямая $\Pi \cap U'$ и отрезок pq , лежащие в одной плоскости Π , не параллельны. Как и выше, доказывается, что всякий отрезок, параллельный подпространству Y' и имеющий один конец на отрезке pq и другой на прямой $\Pi \cap U'$, целиком лежит на сфере S . Следовательно, пересечение $\Pi \cap S$ имеет внутренние точки относительно плоскости Π , т. е. представляет собой невырожденную (выпуклую замкнутую) плоскую область D .

Возьмем вектор v , не параллельный плоскости Π , порожденное им одномерное подпространство $Y(v)$ и соответствующее двумерное подпространство $U(v)$ (для которого $(S \cap C(Y(v))) \supseteq (S \cap U(v))$). Это подпространство $U(v)$ не может пересекать область D по внутренности, иначе часть D (а значит, и сферы S) лежит вне цилиндра $C(Y(v))$, описанного по кривой $S \cap U(v)$. Итак, для любого вектора v , не параллельного Π , область D лежит в Π по одну сторону от прямой $\Pi \cap U(v)$ (при этом, если вектор v , будучи приставлен к какой-нибудь точке из $\Pi \cap U(v)$, «смотрит» в то из двух подпространств, где лежит S , то его проекция на плоскость Π и область D лежат в Π по разные стороны от прямой $\Pi \cap U(v)$). По непрерывности мы можем утверждать это и для любого вектора v , параллельного Π . Но тогда, как уже было дважды показано выше в подобных случаях, область D вместе с каждой своей точкой p содержит отрезок pq , параллельный v и соединяющий p с точкой $q \in \Pi \cap U(v)$ (при этом вектор \overline{pq} сонаправлен с v).

Итак, мы получили невырожденную выпуклую замкнутую область $D \subset \Pi$, которая по любому направлению v , параллельному плоскости Π , сама по себе проектируется на некоторую прямую $l(v) = \Pi \cap U(v)$, не пересекающую внутренность D . Такой области не существует (докажите!).

Итак, наше предположение неверно, $(S \cap C(Y)) = (S \cap U(Y))$ для любого одномерного подпространства $Y \subset \tilde{X}$, и по теореме А из п. 1 S — эллипсоид, а \tilde{X} — гильбертово пространство.

Импликация 2) \implies 0) полностью доказана.

3) \implies 0). Предположим вначале, что пространство X удовлетворяет условию 3) с $k = 2$ (случай Филлипса [10]). Тогда, очевидно, любое трехмерное подпространство $X_3 \subset X$ также удовлетворяет этому условию. Доказав гильбертовость любого такого X_3 , мы в силу следствия из леммы А докажем и гильбертовость X .

Итак, надо доказать гильбертовость трехмерного банахова пространства, в котором на каждое двумерное подпространство Y есть проектор π_Y нормы 1. Существование π_Y равносильно существованию такого вектора $\tau(Y)$, определяющего направление проектирования, что $\|y\| = \|\pi_Y(y + \lambda\tau(Y))\| \leq \|y + \lambda\tau(Y)\|$ для любого $y \in Y$ и любого числа λ . Это означает, что направление $\tau(Y)$ является опорным к единичной сфере $S = S(X_3)$ во всех точках пересечения $S \cap Y$. По теореме В из п. 1, S — эллипсоид, а значит, X_3 — гильбертово пространство.

Пусть теперь исходное пространство X удовлетворяет условию 3) с $k \geq 3$. Возьмем в X произвольное подпространство Y размерности k , а в Y — произвольное $(k-1)$ -мерное подпространство Y_1 . Возьмем также какой-нибудь элемент $c \in \text{Ker}(\pi_Y)$, $c \neq 0$, и порожденное им подпространство $\langle c \rangle$. Подпространство $Z := \langle c \rangle \oplus Y_1$ имеет в X размерность k — ту же, что и Y . По условию существует проектор $\pi_Z : X \rightarrow Z$ нормы 1. Рассмотрим действие композиции $\pi_Y \circ \pi_Z$ на подпространстве Y . Для всех $y \in Y$ имеем $\pi_Z(y) = y_1 + \lambda c$, где $y_1 \in Y_1$, а λ — число. Отсюда $\pi_Y \circ \pi_Z(y) = \pi_Y(y_1) + \lambda\pi_Y(c) = y_1 + \lambda \cdot 0 = y_1$, т. е. оператор $\pi_Y \circ \pi_Z$ переводит Y в его собственное подпространство Y_1 . Поскольку π_Y и π_Z не увеличивают норму, то их композиция также обладает этим свойством. При этом для любого $y_1 \in Y_1$ имеем $\pi_Y \circ \pi_Z(y_1) = y_1$. Таким образом, $\pi_Y \circ \pi_Z = \pi_{Y_1} : Y \rightarrow Y_1$ — проектор нормы 1, т. е. Y как банахово пространство удовлетворяет условию 3) с $k = 1$. При этом в рассматриваемом случае Y не менее чем трехмерно.

Точно так же доказывается, что в любом $(k-1)$ -мерном подпространстве существует проектор нормы 1 на любое его $(k-2)$ -мерное подпространство. Продолжая так далее, за конечное число шагов мы покажем, что в любом трехмерном пространстве $X_3 \subset X$ существует проектор нормы 1 на любое двумерное подпространство, т. е. приходим к уже разобранным случаю.

Можно обойтись и без индукции, показав, что для любого $(k-3)$ -мерного подпространства U в k -мерном Y в факторпространстве $Y_3 = Y/U$ существует проектор $\pi_V : Y_3 \rightarrow V$ нормы 1 на любое двумерное подпространство $V \subset Y_3$. После этого остается только применить лемму В.

Теорема 1 полностью доказана.

Заметим, что условия на n и k в теореме 1 существенны. Если $\dim X \neq \infty$ и $n = \dim X - 1$, то для *любого* n -мерного подпространства $Y \subset X$ оператор метрического проектирования P_Y допускает линейную выборку (именно, проектирование вдоль любого направления $v \neq 0$, для которого $P_Y(v) \ni 0$). В условии 3) равенство $k = 1$ также не годится: в любом банаховом пространстве X есть проектор π_Y нормы 1 на любое одномерное подпространство Y (именно, проектирование идет вдоль гиперплоскости, опорной к единичной сфере S в точках $S \cap Y$).

Существуют «антиподы» гильбертовых пространств — банаховы пространства, в которых нет собственных подпространств Y коразмерности > 1 с линейным оператором метрического проектирования $P_Y : X \rightarrow Y$. Таким является, например, пространство $C[0, 1]$ с равномерной нормой [11]. Можно построить и (даже трехмерное) пространство X , в котором нет собственных подпространств Y коразмерности > 1 с линейной выборкой из оператора P_Y . В справедливости этого последнего факта мы предлагаем читателю убедиться самостоятельно.

3. ПОПЕРЕЧНИКИ И ГИЛЬБЕРТОВОСТЬ

Пусть $(X, \|\cdot\|)$ — банахово пространство, C — подмножество X , $\Phi = \{F\}$ — некоторое семейство непрерывных отображений из C в X . Рассмотрим величину

$$p_\Phi(C, X) = \inf_{F \in \Phi} \sup_{x \in C} \|x - F(x)\|. \quad (2)$$

Если Φ — совокупность линейных (на линейной оболочке C) отображений, образ которых n -мерен, величина (2) называется *линейным поперечником* $\lambda_n(C, X)$; при том же Φ величина $d^n(C, X) = \frac{1}{2} \inf_{F \in \Phi} \sup_{x \in C} \text{diam}[F^{-1}(F(x))]$ называется *поперечником по Гельфанду* множества C . Если Φ — совокупность всех непрерывных отображений F , образ каждого из которых принадлежит некоторой n -мерной плоскости $\Pi = \Pi_n(F)$, величину (2) называют *поперечником по Колмогорову* $d_n(C, X)$. Если же Φ — совокупность всех непрерывных отображений, образ которых — n -мерный комплекс¹⁾, расположенный в X , то (2) называется *александровским поперечником* (в честь П. С. Александрова) и обозначается $a_n(C, X)$.

О свойствах поперечников и их роли в теории приближений можно подробно узнать в гл. 3 обзора [12]. Здесь же мы займемся поперечниками подмножеств гильбертова пространства и теми их свойствами, которые являются характеристическими для гильбертовости.

Для любого отображения F множества C в n -мерную плоскость L величина $\sup_{x \in C} \|x - F(x)\|$ не меньше уклонения C от L . Поэтому $\lambda_n(C, X) \geq d_n(C, X) \geq \inf_{\dim L=n} d(C, L, X)$. Но, как мы знаем, в гильбертовом пространстве $d(C, L, X) = \sup_{x \in C} \|x - P_L(x)\|$, где P_L — линейный оператор ортогонального проектирования на L , откуда $\lambda_n(C, X) \leq \inf_{\dim L=n} d(C, L, X)$. Следовательно, $\lambda_n(C, X) = d_n(C, X)$ для любого множества C в гильбертовом пространстве X .

Далее, для линейного P_L имеем

$$\frac{1}{2} \sup_{x \in C} \text{diam}[P_L^{-1}(P_L(x))] \leq \frac{1}{2} \sup_{x \in C} 2 \sup_{z \in C: P_L(z)=P_L(x)} \|z - P_L(x)\| = \sup_{x \in C} \|x - P_L(x)\|,$$

откуда $d^n(C, X) \leq d_n(C, X)$ (в случае произвольного банахова пространства X поперечники d^n и d_n , вообще говоря, никак не связаны).

Если Y — некоторое подпространство в гильбертовом пространстве X , то оператор P_Y имеет норму 1, откуда $\sup_{x \in C} \|x - F(x)\| \geq \sup_{x \in C} \|P_Y(x) - P_Y \circ F(x)\|$. Отсюда для любого поперечника $p_n = \lambda_n, d_n, d^n, a_n$ следует, во-первых,

¹⁾ Объединение в некотором числе не более чем n -мерных симплексов: точек (нульмерные симплексы), отрезков (одномерные симплексы), треугольников (двумерные симплексы), тетраэдров (трехмерные симплексы) и т. д.

что $p_n(P_Y(C), X) \leq p_n(C, X)$, а во-вторых, что если $C \subset Y$, то $p_n(C, Y) = p_n(C, X)$.

При этом для множеств $C \not\subset Y$ поперечник не обязан строго уменьшаться после ортогонального проектирования на подпространство. Докажем, например, что для единичного шара B и любого $(n+1)$ -мерного подпространства Y выполняется равенство $a_n(B \cap Y, X) = a_n(B, X) = 1$. Прежде всего, $a_n(B \cap Y, Y) = a_n(B \cap Y, X) \leq a_n(B, X) \leq 1$ (последнее неравенство следует из того, что для отображения $F : B \rightarrow \{0\}$ шара B в нульмерный комплекс $\{0\}$ имеем $\sup_{x \in B} \|x - F(x)\| = 1$). Осталось доказать равенство $a_n(B \cap Y, Y) = 1$. Следуя М. И. Стесину ([13, теорема 1.1.7]), предположим противное: существует такое непрерывное отображение F шара $B_Y = B \cap Y$ в n -мерный комплекс $K_n \subset Y$, что $\|x - F(x)\| \leq 1 - \delta$, $\delta > 0$, для любого $x \in B_Y$. Возьмем произвольный элемент $y \in Y$, $\|y\| \leq \delta/2$, и покажем, что $y \in F(B_Y)$. Если $y \notin F(B_Y)$, то $G(x) := \frac{y - F(x)}{\|y - F(x)\|}$ — непрерывное отображение шара B_Y в себя, точнее, в сферу $S_Y = \{y : \|y\| = 1\}$. По теореме Брауэра о неподвижной точке у отображения G существует неподвижная точка x_0 : $G(x_0) = x_0$, $\|x_0\| = 1$. Отсюда $F(x_0) = y - \|y - F(x_0)\|x_0$, так что

$$\begin{aligned} \|x_0 - F(x_0)\| &= \|x_0(1 + \|y - F(x_0)\|) - y\| \geq \|x_0\|(1 + \|y - F(x_0)\|) - \|y\| \geq \\ &\geq \|x_0\| - \delta/2 > 1 - \delta, \end{aligned}$$

что противоречит предположению. Поэтому образ $F(B_Y)$ содержит $\left(\frac{\delta}{2}\right)$ -окрестность нуля и не может быть n -мерным комплексом (напомним, наше пространство Y имеет размерность $n+1$), так что действительно $a_n(B_Y, Y) = 1$. Заметим, что это рассуждение Стесина годится для любого $(n+1)$ -мерного банахова пространства.

Если X — произвольное банахово пространство с $\dim X \geq n+1$, то, вообще говоря, $a_n(B \cap Y, X) \leq 1$. В то же время для колмогоровских поперечников всегда $d_n(B \cap Y, X) = 1$.

Как показывает следующая теорема, выполнение любого из выявленных нами (не)равенств при каком-то фиксированном n является характеристическим свойством гильбертовых пространств.

ТЕОРЕМА 2. Пусть X — вещественное банахово пространство, $\dim X \geq 3$. Следующие условия эквивалентны:

0) X — гильбертово;

1) существует такое число $n \in \mathbb{N}$, $n \leq \dim X - 2$, что для любого множества $C \subset X$ справедливо равенство $d_n(C, X) = \lambda_n(C, X)$;

2) существует такое число $n \in \mathbb{N}$, $n \leq \dim X - 2$, что для любого множества $C \subset X$ справедливо неравенство $d^n(C, X) \leq d_n(C, X)$;

3) существует такое число $n \in \mathbb{Z}_+$, $n \leq \dim X - 3$, что для любого $(n+2)$ -мерного подпространства $Y \subset X$ и любого множества $C \subset Y$ справедливо равенство $d_n(C, Y) = d_n(C, X)$;

4) существует такое число $n \in \mathbb{N}$, $n \leq \dim X - 2$, что для любого $(n+1)$ -мерного подпространства $Y \subset X$ справедливо равенство $a_n(B \cap Y, X) = 1$.

В этой теореме утверждение о равносильности 0) и 3) доказано в [14]. Статья одного из авторов (В. М. Тихомиров), в которой были сформулированы эквивалентности 0) \iff 1) \iff 2), осталась неопубликованной. Равносильность условий 0) и 4) доказана М. И. Стесиним [13].

Доказательство. Выполнение каждого из условий 1), 2), 3) и 4) в гильбертовом пространстве уже отмечалось выше. Докажем обратные утверждения.

1) \implies 0). Пусть Y_n — произвольное n -мерное подпространство в X . Положим $C = Y_n + B = \{x \in X : \rho(x, Y_n) \leq 1\}$. Очевидно, $d_n(C, X) = d(C, Y_n, X) = 1$. Отсюда $\lambda_n(C, X) = 1$, и подпространство Y_n является экстремальным также и для этого поперечника (если π_k — линейные проекторы множества C на n -мерные плоскости Y_{n_k} и $\sup_{x \in C} \|x - \pi_k(x)\| \rightarrow 1$ при $k \rightarrow \infty$, то все $Y_{n_k} \parallel Y_n$, $d(Y_{n_k}, Y_n, X) \rightarrow 1$, а $\pi_k(x) \rightarrow P(x)$, где $P : C \rightarrow Y_n$ — линейный проектор, $\sup_{x \in C} \|x - P(x)\| = 1$). Итак, существует линейный проектор $P : X \rightarrow Y_n$ с условием $\forall x \in C \quad \|x - P(x)\| \leq 1$. Рассуждая точно так же, как при доказательстве импликации 1) \implies 2) в теореме 1, доказываем, что P является линейной выборкой из оператора метрического проектирования P_{Y_n} . По теореме 1 X — гильбертово.

2) \implies 0). Возьмем произвольное $(n+2)$ -мерное подпространство $X_{n+2} \subset X$, а в нем — произвольное n -мерное подпространство Y_n . Положим опять $C = Y_n + B$. Имеем $1 = d^n(B, X) \leq d^n(C, X) \leq d_n(C, X) = 1$, так что $d^n(C, X) = 1$. Пусть $\rho_k = \frac{1}{2} \sup_{x \in C} \text{diam}[F_k^{-1}(F_k(x))] \rightarrow 1$, где $F_k : X \rightarrow Y_n^{(k)}$ — линейные проекторы на некоторые n -мерные подпространства $Y_n^{(k)}$ (при определении поперечника по Гельфанду можно ограничиться отображениями только на n -мерные подпространства). Для любого элемента $q \in \text{Ker}(F_k) \cap C$ весь отрезок $[-q, q] \subset (\text{Ker}(F_k) \cap C)$, поэтому $\frac{1}{2} \text{diam}[-q, q] \leq \rho_k$, т. е. $\|q\| \leq \rho_k$. Ядра $\text{Ker}(F_k)$ трансверсальны Y_n , поэтому в X_{n+2} можно рассмотреть линейные проекторы $\pi_k : X_{n+2} \rightarrow Y_n$ вдоль двумерных подпространств $X_{n+2} \cap \text{Ker}(F_k)$. Для любого элемента $x \in X_{n+2}$, $x = y + q$, $y \in Y_n$, $q \in \text{Ker}(F_k)$ имеем

$$\|\pi_k(x) - x\| = \|q\| \leq \rho(q, Y_n) \left\| \frac{q}{\rho(q, Y_n)} \right\| \leq \rho(q, Y_n) \rho_k = \rho(x, Y_n) \rho_k \rightarrow \rho(x, Y_n).$$

Далее, в X_{n+2} из последовательности π_k можно выделить сходящуюся подпоследовательность, и для предельного проектора $\pi : X_{n+2} \rightarrow Y_n$ будет $\|\pi(x) - x\| \leq \rho(x, Y_n)$, так что π — линейная выборка из оператора метрического проектирования $P_{Y_n} : X_{n+2} \rightarrow Y_n$. По теореме 1 X_{n+2} — гильбертово, а значит, X тоже гильбертово.

В доказательстве оставшихся утверждений теоремы 2 существенную роль будет играть один результат А. Л. Гаркави. Прежде чем формулировать его, напомним, что нулевой колмогоровский поперечник $d_0(C, X)$ называется также *чебышевским радиусом* множества C (infimum радиусов шаров, содержащих C), а точка $t = t(C)$, для которой $d_0(C, X) = \sup_{x \in C} \|x - t\|$, (если она существует) — *чебышевским центром* множества C .

ЛЕММА С (А. Л. ГАРКАВИ, [15]). *Вещественное банахово пространство X , $\dim X \geq 3$, является гильбертовым тогда и только тогда, когда для*

любых трех точек некоторый их чебышевский центр принадлежит их линейной оболочке.

Как видим, здесь критерием гильбертовости выступает ослабленное условие 3) при $n = 0$ из теоремы 2.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ С. Необходимость уже отмечалась. Докажем достаточность.

Возьмем произвольное трехмерное подпространство $X_3 \subset X$, двумерное подпространство $Y_2 \subset X_3$ и точку $x \in X_3 \setminus Y_2$. Обозначим $S_r = \{y \in Y_2 : \|x - y\| = r\}$. Если $y_1, y_2, y_3 \in S_r$, то шары $B(y_j, r) = \{z \in X_3 : \|z - y_j\| \leq r\}$, $j = 1, 2, 3$, имеют общую точку в плоскости Y_2 (иначе никакой чебышевский центр трех точек $\{y_j\}_{j=1}^3$ не лежит в Y_2). Отсюда по теореме Хелли ([16, гл. I, § 7]) существует общая точка $z_r \in Y_2$, принадлежащая всем шарам $B(y, r)$ при $y \in S_r$.

Докажем теперь, что все шары вида $B(y, \|x - y\|)$, $y \in Y_2$, имеют общую точку. Достаточно доказать это для любых трех шаров $B(y_j, \|x - y_j\|)$, $j = 1, 2, 3$ (а потом опять применить теорему Хелли). Положим $a = \max\{\|x - y_j\| : j = 1, 2, 3\}$, и пусть y'_j — такая точка из S_a , что отрезок $[z_a, y'_j]$ содержит y_j , $j = 1, 2, 3$ (при этом, конечно, $y'_j = y_j$ по крайней мере для одного j). Имеем $\|z_a - y_j\| + \|y_j - y'_j\| \leq a = \|x - y'_j\| \leq \|x - y_j\| + \|y'_j - y_j\|$, т. е. $\|z_a - y_j\| \leq \|x - y_j\|$, а значит, $z_a \in B(y_j, \|x - y_j\|)$, $j = 1, 2, 3$, что и требовалось.

Если точка $z \in B(y, \|x - y\|) \cap Y_2$ для всех $y \in Y_2$ (существование такой точки только что доказано), то $\|y - z\| \leq \|y - x\|$ для всех $y \in Y_2$. А это означает, что существует проектор $\pi : X_3 \rightarrow Y_2$ нормы 1. Действительно, каждый элемент $v \in X_3$ можно однозначно представить в виде $v = \alpha(x - z) + y$, где $y \in Y_2$, α — некоторое число, и можно определить отображение $\pi : v \rightarrow y$. Это линейный проектор, и его норма равна 1:

$$\|v\| = \|\alpha(x - z) + y\| = |\alpha| \|x - (z - \frac{y}{\alpha})\| \geq |\alpha| \|z - (z - \frac{y}{\alpha})\| = \|y\| = \|\pi(v)\|.$$

Итак, в пространстве X_3 существует проектор нормы 1 на каждое двумерное подпространство Y_2 . По теореме 1 X_3 — гильбертово. Следовательно (по следствию из леммы А), гильбертовым является и исходное пространство X .

Лемма С доказана.

Продолжим доказательство теоремы 2.

3) \implies 0). Пусть X не гильбертово. По лемме С существуют три точки $x_1, x_2, x_3 \in X$, никакой чебышевский центр которых не лежит в их линейной оболочке. Можно считать, что эта линейная оболочка — двумерное подпространство Y_2 , 0 — чебышевский центр множества $C_0 = \{x_1, x_2, x_3\}$ относительно Y_2 и $d_0(C_0, Y_2) = 1 \geq d_0(C_0, X)$.

Покажем, что $\|x_j\| = 1$, $j = 1, 2, 3$. Менее чем для двух точек это равенство выполняться не может, иначе 0 не является чебышевским центром C_0 в Y_2 . Предположим, что $\|x_1\| = \|x_2\| = 1$, а $\|x_3\| < 1$. Если $\|x_1 - x_2\| < 2$, то для $z(\alpha) = \alpha \frac{x_1 + x_2}{2}$, $\alpha \in (0, 1)$, при $j = 1, 2$ имеем

$$\|x_j - z(\alpha)\| \leq \|\alpha(x_j - \frac{x_1 + x_2}{2})\| + \|(1 - \alpha)x_j\| = \alpha \|\frac{x_1 - x_2}{2}\| + 1 - \alpha < 1,$$

в то время как $\|x_3 - z(\alpha)\| < 1$ при всех достаточно малых α , так что опять получаем, что 0 — не чебышевский центр. Если же $\|x_1 - x_2\| = 2$, то $d_0(C_0, X) \geq$

$\geq d_0(\{x_1, x_2\}, X) = 1$, что опять противоречит условиям на точки x_j . Поэтому наше предположение неверно, и все три точки лежат на единичной сфере $S(Y_2)$ подпространства Y_2 .

Если в Y_2 провести прямые l_j , опорные к $S(Y_2)$ в точках x_j , $j = 1, 2, 3$ (в одной точке может быть много опорных прямых — выберем какую-нибудь), то эти прямые будут образовывать треугольник T , содержащий внутри себя сферу $S(Y_2)$ (докажите!).

Возьмем в X какое-нибудь $(n + 3)$ -мерное подпространство X' (напомним, $\dim X \geq n + 3$), содержащее Y_2 , а в нем — гиперплоскости Γ_j , опорные к единичной сфере $S(X')$ в точках x_j и содержащие в себе прямые l_j , $j = 1, 2, 3$. Пересечение трех подпространств $\Gamma_j \subset X'$, параллельных Γ_j , по крайней мере n -мерно. Возьмем в этом пересечении n -мерное подпространство Y_n и положим $C = C_0 + Y_n$, $Y = Y_2 + Y_n$. Размерность Y равна $n + 2$, поскольку Y_n трансверсально Y_2 .

Очевидно, $d_n(C, X) \leq d_0(C, X) < 1$. Для получения противоречия с условием 3) осталось доказать, что $d_n(C, Y) = d_0(C_0, Y_2) = 1$.

Если $d_n(C, Y) < 1$, то существует такая n -мерная плоскость $\Pi \subset Y$, параллельная Y_n , что $\forall x \in C \quad \rho(x, \Pi) \leq 1 - \delta$, где $\delta > 0$ (плоскость Π не может быть не параллельна Y_n , иначе $d(C, \Pi, Y) = \infty$). Это равносильно тому, что расстояние от точки $p = \Pi \cap Y_2$ до множества C не превосходит $1 - \delta$, откуда $\rho(p, \Gamma_j) \leq \rho(p, x_j + Y_n) \leq 1 - \delta$, $j = 1, 2, 3$. Следовательно, точка p при каждом j лежит в той из двух полуплоскостей в $Y_2 \setminus (Y_2 \cap \Gamma_j^0)$, где лежит x_j . Но эти три полуплоскости не пересекаются именно в силу того, что треугольник T , образованный прямыми l_j , содержит внутри себя сферу $S(Y_2)$ (проверьте!).

4) \implies 0). Будем опять доказывать от противного: предположим, что X не гильбертово, возьмем те же точки $x_1, x_2, x_3 \in Y_2$ и треугольник T , что и в предыдущем пункте.

Как и в предыдущем пункте, возьмем в X какое-нибудь $(n + 2)$ -мерное подпространство X' (напомним, $\dim X \geq n + 2$), содержащее Y_2 , а в нем — гиперплоскости Γ_j , опорные к единичной сфере $S(X')$ в точках x_j и содержащие в себе прямые l_j , $j = 1, 2, 3$. Пересечение трех подпространств $\Gamma_j^0 \subset X'$, параллельных Γ_j , по крайней мере $(n - 1)$ -мерно. Возьмем в этом пересечении $(n - 1)$ -мерное подпространство Y_{n-1} и положим $Y = Y_2 + Y_{n-1}$. Размерность Y равна $n + 1$, поскольку Y_{n-1} трансверсально Y_2 .

Из включения $B \cap Y \subset T + Y_{n-1}$ вытекает неравенство $a_n(B \cap Y, X) \leq a_1(T, X)$ (действительно, если для какого-то непрерывного отображения F треугольника T в одномерный комплекс K_1 имеем $s = \sup_{t \in T} \|t - F(t)\|$, то для непрерывного отображения $\tilde{F} : x = t + z \mapsto F(t) + z$, $t \in T, z \in Y_{n-1}$ пересечения $B \cap Y$ в некоторый n -мерный комплекс из суммы $K_1 + Y_{n-1}$ будет справедливо неравенство $\sup_{t \in B \cap Y} \|t - \tilde{F}(t)\| \leq s$).

Для получения противоречия с условием 4) осталось доказать неравенство $a_1(T, X) < 1$. Возьмем какую-нибудь точку $w \in X$ с условием $\|w - x_j\| = \rho < 1$, $j = 1, 2, 3$. Такая точка существует, поскольку $d_0(C_0, X) < 1$, и заведомо не лежит в Y_2 . Обозначим через y_k вершины треугольника T — так, чтобы стороны l_i и l_j пересекались в точке y_k с $k \neq i, j$. Построим теперь отображение F треугольника T в комплекс $[w, y_1] \cup [w, y_2] \cup [w, y_3]$ следующим образом: в маленьком треугольнике $x_1 x_2 x_3 \subset T$ положим $F(x) = w$, а в каждом из трех оставшихся

маленьких треугольников $x_i x_j x_k \subset T$, где i, j, k — отличные друг от друга числа из набора $\{1, 2, 3\}$, определим $F(x)$ как такую точку на отрезке $[w, y_k]$, для которой отрезок $[F(x), x]$ параллелен плоскости треугольника $w x_i x_j$. Нетрудно видеть, что отображение F непрерывно и $\|x - F(x)\| \leq \rho$ для любой точки $x \in T$. Отсюда следует требуемое неравенство $a_1(T, X) < 1$.

Теорема 2 полностью доказана.

Известно множество других критериев гильбертовости банахова пространства; в настоящее время их число достигает нескольких сотен и постоянно увеличивается. В связи с этим ценность каждого нового критерия невелика, да и формулировки их, как правило, весьма естественны. Одним из немногих аргументов в пользу результатов этого типа может служить, пожалуй, возможность получения из них геометрических характеристик эллипсоидов в классе всех выпуклых замкнутых центрально-симметричных поверхностей в многомерных пространствах. Так, теоремы 1 и 2 дают многомерные обобщения теорем Бляшке-Александрова и Бляшке-Какутани (см. [9], [8], [15]).

Мы, впрочем, и не стремились побудить читателя заняться изобретением новых критериев гильбертовости, а лишь хотели познакомить его с несколькими изящными фактами из геометрии и теории приближений, возникающими при доказательстве таких критериев.

СПИСОК ЛИТЕРАТУРЫ

- [1] *Колмогоров А. Н., Фомин С. В.* Элементы теории функций и функционального анализа. М.: Наука, 1989.
- [2] *Гельфанд И. М.* Лекции по линейной алгебре. М., 1971.
- [3] *Jordan P., Neumann J. von.* On inner products in linear metric spaces // *Ann. Math. (2)*. 1935. V. 36. P. 719-723.
- [4] *Бляшке В.* Круг и шар. М., 1967.
- [5] *Александров А. Д.* О выпуклых поверхностях с плоскими границами теней // *Матем. сборник*. 1939. Т. 5, № 2. С. 309-316.
- [6] *Kakutani S.* Some characterizations of euclidian spaces // *Japan. J. Math.* 1939. V. 16, No 2. P. 93-97.
- [7] *Бурбаки Н.* Топологические векторные пространства. М., 1959.
- [8] *Бородин П. А.* Квазиортогональные множества и условия гильбертовости банахова пространства // *Матем. сборник*. 1997. Т. 188, № 8. С. 63-74.
- [9] *Rudin W., Smith K. T.* Linearity of best approximation: a characterization of ellipsoids // *Indag. Math.* 1961. V. 23, No 1. P. 97-103.
- [10] *Phillips R. S.* A characterization of euclidian spaces // *Bull. Amer. Math. Soc.* 1940. V. 46, No 12. P. 930-933.
- [11] *Бородин П. А.* О линейности оператора метрического проектирования на чебышевские подпространства в пространствах L_1 и C // *Матем. заметки*. 1998. Т. 63, вып. 6. С. 812-820.

- [12] *Тихомиров В. М.* Теория приближений. Итоги науки и техники. Сер. Современные проблемы математики. Т. 14. ВИНТИ, 1987. С. 103-260.
- [13] *Стесин М. И.* Александровские поперечники множеств в банаховых пространствах. Канд. дисс. М., 1975.
- [14] *Тихомиров В. М., Исмагилов Р. С., Бабаджанов С. Б.* Геометрия банахова пространства и поперечники множеств // Изв. АН Узбекской ССР. Сер. физ.-мат. 1979. № 4. С. 25-32.
- [15] *Гаркави А. Л.* О чебышевском центре и выпуклой оболочке множества // УМН, 1964. Т. 19, № 6. С. 139-145.
- [16] *Лейтвейс К.* Выпуклые множества. М., 1985.
- [17] *Никольский С. М.* Некоторые вопросы приближений функций полиномами // Международный Математический Конгресс в Амстердаме. М.: Физматгиз, 1961.

Наглядная интерпретация некоторых алгоритмов на графах

Г. Райтер И. Сонин

(Университет Северной Каролины, г. Шарлот)

Большое место в теории графов занимают алгоритмы. Нахождение кратчайших путей, построение остовных деревьев с нужными свойствами, построение паросочетаний, раскраска вершин — вот далеко не полный перечень задач, решение которых состоит в описании алгоритма.

Некоторые студенты сталкиваются с трудностями при понимании алгоритмов. Как правило, алгоритмы на графах основаны на довольно простых идеях. Но их формальное изложение бывает длинным, а понимание такого изложения может требовать определенной математической подготовки. Еще одна трудность состоит в том, что алгоритмы зачастую записываются в псевдокоде — очень удобно для программирования компьютера, но не для восприятия человеком. Поэтому нужны какие-то мнемонические приемы для облегчения запоминания и понимания таких алгоритмов. Наш подход состоит в том, чтобы интерпретировать некоторые алгоритмы из теории графов в знакомом большинству студентов контексте: вступление в клуб.

Пусть $G(V, E)$ — граф с множеством вершин V и множеством ребер E . Всюду далее считаем этот граф связным (в противном случае можно просто рассматривать все его связные компоненты по отдельности). Напомним, что граф H называется *подграфом* G , если он получается выбором некоторых вершин G и части соединяющих эти вершины ребер G . *Дерево* — это связный граф без циклов. *Остовное дерево* графа G — это подграф G , содержащий все вершины G и являющийся деревом. Мы будем рассматривать алгоритмы построения остовного дерева T графа G . Построение остовного дерева начинается с выделенной вершины, которую будем называть *корнем* дерева.

Идея состоит в том, чтобы интерпретировать V как множество людей. Ребро соединяет двух людей, если они знакомы. Есть клуб, в который хотят вступить все эти люди. Постепенно, один за другим, они туда все и вступят. Чтобы человек u мог вступить в клуб, у него должен быть *поручитель*: член клуба v , знакомый с u . Член клуба может быть поручителем, если он знаком хотя бы с кем-то не из клуба, такого члена клуба будем называть *открытым*. Конечно, по мере роста клуба, открытый член клуба может перестать быть таковым, но не наоборот. Приведенные ниже правила определяют для каждого момента, кто вступает в клуб следующим и кто его поручитель.

Наше основное правило порождает широкий класс корневых остовных деревьев. Дополнительные к основному правилу условия позволяют построить такие

остовные деревья как дерево поиска в ширину (Breadth First Search, BFS), дерево поиска в глубину (Depth First Search, DFS), кратчайшая связывающая сеть (Minimal Total, MT), дерево кратчайших путей (Shortest Path, SP). Эти деревья встречаются при решении многих задач. Например, дерево поиска в ширину можно использовать для определения длины кратчайшего (содержащего наименьшее число ребер) пути из корня в любую другую вершину. Дерево поиска в глубину позволяет находить мосты в графе, а в случае их отсутствия — построить сильно связную ориентацию графа. (*Мост* — ребро, удаление которого делает граф несвязным. *Ориентация* — приписывание ребрам направления, ориентированный граф *сильно связан*, если из любой вершины можно пройти в любую по ребрам вдоль указанных на них направлений, пример см. на рис. 3.)

ОСНОВНОЕ ПРАВИЛО. *Есть ровно один основатель клуба. Все остальные вступают в клуб по одному, пока еще есть кому вступать. Каждый вступающий в клуб имеет поручителя.*

Основное правило порождает последовательность подграфов. Начинаем с выделенного корня (основатель u_0), на каждом шаге в подграф добавляется одна вершина (новый член) и одно ребро (от поручителя к новому члену), рис. 1. Алгоритмы, которые мы обсудим, отличаются предписаниями по выбору нового члена и его поручителя. Если две или более пар «новый член – поручитель» имеют равный приоритет вступления в клуб, выбираем одну из них произвольно. Это соглашение применимо ко всем обсуждаемым далее алгоритмам.

Давайте убедимся, что любой подграф, получаемый в соответствии с Основным правилом и любым его уточнением, — остовное дерево. Заметим, что в любой момент множество членов клуба и ребер, соединяющих членов с их поручителями, является связным подграфом графа G . Предположим, что после вступления некоторого члена u образовался цикл. Ясно, что u не является ни чьим поручителем — после него в клуб никто не вступал. Значит, оба члена клуба, связанные с u ребрами цикла, — его поручители, что противоречит Основному правилу. Поэтому в любой момент мы имеем связный ациклический граф. Наконец, если в некоторый момент еще есть нечлены клуба, то — ведь граф G связный — кто-то из членов клуба должен быть открытым. Так что процесс продолжается до тех пор, пока все нечлены вступят в клуб. И в этот момент клуб задает дерево, содержащее все вершины G , — остовное дерево G .

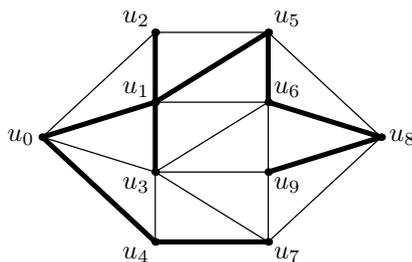


Рис. 1. Применение основного правила приводит к остовному дереву (выделено жирным). Индексы указывают на порядок вступления в клуб

Порядок вступления в клуб задает помеченное дерево: каждой вершине приписано число от 0 до $n - 1$ так, что на пути от основателя до любого члена пометки идут в возрастающем порядке.

Следующие два правила основаны на понятии *старшинства*. Если u вступил в клуб раньше, чем v , то u *старше* v (а v младше u).

ПРАВИЛО СТАРШИНСТВА. *Поручителем при вступлении в клуб выбирают старейшего на данный момент открытого члена.*

Получаемое остовное дерево называется деревом поиска в ширину, рис. 2.

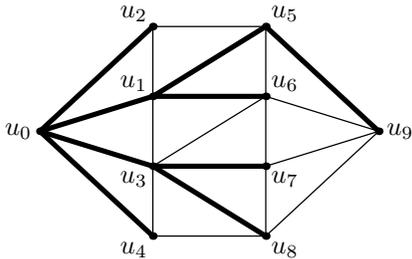


Рис. 2. Применение правила старшинства приводит к дереву поиска в ширину (выделено жирным)

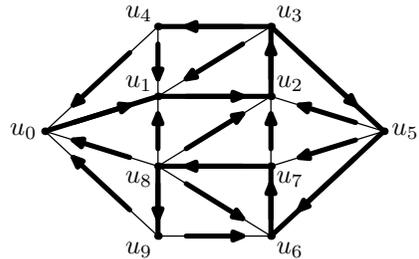


Рис. 3. Применение правила младшинства приводит к дереву поиска в глубину (выделено жирным)

ПРАВИЛО МЛАДШИНСТВА. *Поручителем при вступлении в клуб выбирают самого младшего на данный момент открытого члена.*

Получаемое остовное дерево называется деревом поиска в глубину (DFS). На рис. 3 изображено DFS с дополнительной структурой сильно связанной ориентации. Ребра этого дерева ориентированы от старшего члена к младшему. Остальные ребра ориентированы от младшего члена к старшему. Любой граф без мостов можно сильно связно ориентировать, построив дерево поиска в глубину и задав ориентацию как сказано выше.

Некоторые алгоритмы применяются к *сетям*, т. е. таким графам, каждому ребру e которых приписан неотрицательный вес $f(e)$ (длина ребра). Кратчайший путь из вершины u в вершину v — это путь с наименьшей суммой весов входящих в него ребер (по сравнению с остальными путями). Чтобы сформулировать следующее правило, будем считать, что каждый член платит вступительный взнос, равный весу ребра от этого члена до его поручителя. Обратите внимание, что величина вступительного взноса может зависеть от выбора поручителя.

ПЕРВОЕ ЖАДНОЕ ПРАВИЛО. *В клуб вступает тот, у кого величина вступительного взноса на данный момент является наименьшей. (Поручителем становится, естественно, тот, кто эту величину обеспечивает.)*

Получаемое таким образом дерево T называется кратчайшей связывающей сетью (МТ), см. рис. 4. Его вес — это $\sum_{e \in T} f(e)$. Жадное правило — простейший

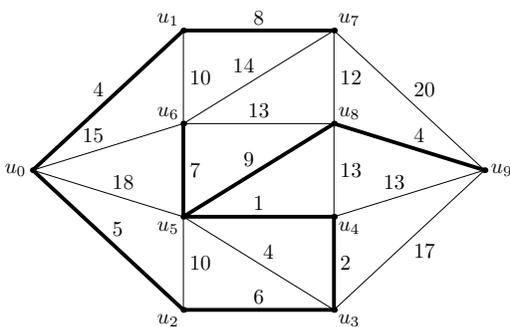


Рис. 4. Применение первого жадного правила приводит к остовному дереву наименьшего суммарного веса

способ описать алгоритм Прима для построения остовного дерева наименьшего веса. Доказательство того, что T в самом деле является МТ, есть в [3] или [5].

Наше заключительное правило описывает построение остовного дерева, задающего кратчайшие пути из корня в любую другую вершину. Чтобы его сформулировать, предположим, что каждый новый член не только должен заплатить вступительный взнос, но и вернуть своему поручителю всю сумму, которую тот заплатил при вступлении в клуб. Поэтому общая сумма, выплачиваемая новым членом, равна сумме весов ребер на пути (единственном) от основателя до этого нового члена. Как и раньше, величина выплаты зависит от выбора поручителя.

ВТОРОЕ ЖАДНОЕ ПРАВИЛО. В клуб вступает тот, кому приходится платить меньше всего на данный момент. (Поручителем выбирается, естественно, тот, кто такой платеж обеспечивает.)

Второе жадное правило — это наш способ сформулировать алгоритм Дейкстры построения остовного дерева кратчайших путей, см. рис. 5.

Все эти алгоритмы можно применять и к ориентированным графам, считая

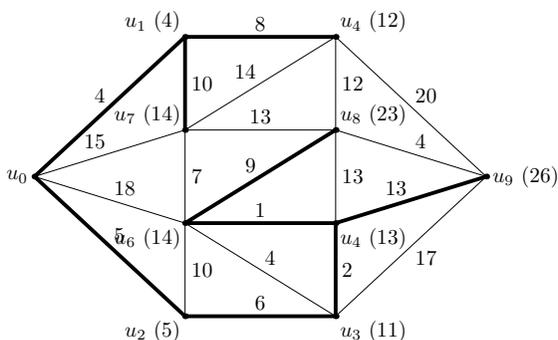


Рис. 5. Применение второго жадного правила приводит к остовному дереву с кратчайшими путями от корня до любой вершины. (В скобках указаны платежи при вступлении.)

что ориентированное ребро $\langle u, v \rangle$ позволяет u быть поручителем v , но не наоборот. Другие алгоритмы на графах также можно объяснять на языке клубов и членов. В частности, алгоритм Форда–Фалкерсона построения максимального потока (или минимального разреза), равно как и алгоритм построения максимального взвешенного паросочетания, использует построение остовных деревьев как часть общей процедуры, см. [3, 5]. Простой способ описания алгоритмов построения остовных деревьев с нужными свойствами облегчает понимание этих, более сложных, алгоритмов. Мы надеемся, что наши читатели с радостью вступят в клуб тех, кто знает и умеет использовать алгоритмы теории графов.

СПИСОК ЛИТЕРАТУРЫ

- [1] *K. P. Bogart*. Introductory Combinatorics. 2nd ed. Harcourt Brace Jovanovich, New York, 1990.
 - [2] *G. Chartand, L. Lesniak*. Graphs and Digraphs. 2nd ed. Wadsworth & Brooks/Cole, Monterey, CA, 1986.
 - [3] *A. Gibbons*. Algorithmic graph theory. Cambridge University Press, Cambridge, 1985.
 - [4] *F. Harary*. Graph theory. Addison–Wesley, Reading, MA, 1969.
 - [5] *J. A. McHugh* Algorithmic graph theory. Prentice Hall, Englewood Cliffs, NJ, 1990.
 - [6] *F. S. Roberts* Applied Combinatorics. Prentice Hall, Englewood Cliffs, NJ, 1984.
- ДОБАВЛЕНО ПРИ ПЕРЕВОДЕ:
- [7] *А. А. Зыков*. Основы теории графов. М.: Наука, 1987.
 - [8] *Ф. Харари*. Теория графов. М.: Мир, 1973.
 - [9] *Н. Кристофидес*. Теория графов. Алгоритмический подход. М.: Мир, 1978.
 - [10] *А. Ахо, Дж. Хопкрофт, Дж. Ульман*. Построение и анализ вычислительных алгоритмов. М.: Мир, 1979.

Конкурсы и олимпиады

Вниманию читателя предлагаются задачи нескольких олимпиад прошедшего 1998 года. Выбраны олимпиады, к проведению которых прямое отношение имеет «московская» группа математиков. Слово «московская» применимо к ней лишь условно: одни разъехались, а другие никогда и не были москвичами. Однако эта группа может называться московской, потому что она следует традициям московских математических олимпиад.

Это сказано не для того, чтобы противопоставить московские олимпиады каким-либо другим. Скажем, между московскими и питерскими олимпиадами на общемировом фоне гораздо больше сходства, чем различий. Главное в московских традициях, на мой взгляд, это стремление, чтобы задачи были интересные и содержательные. Олимпиада должна быть праздником для школьников и всех, кто проводит олимпиаду, — студентов, аспирантов, преподавателей. Это не проверка усвоения, не отбор на олимпиаду следующего уровня, не выделение чемпионов, хотя такие побочные результаты тоже достигаются. Главное — это радость общения с математикой, с интересными задачами, которые хочется решать, рассказать их друзьям.

Конечно, идеал редко достигается, так как задачи должны быть к тому же не слишком трудные, достаточно разнообразные и должны быть к сроку подготовлены. Все же постоянная ориентация московской группы на качество задач дает свои результаты. Эти задачи популярны в России и за рубежом.

Школьникам я советую пробовать решать эти задачи, не заглядывая в решения. В них стоит смотреть, только если уже решил задачу или если после долгих попыток уже потерял всякую надежду. А что значит «долгих» — зависит от характера человека и интереса к решению задач. Если задача интересна, то и времени на нее не жалко.

*Н. Н. Константинов, президент Международного
математического Турнира Городов*

Избранные задачи математических соревнований 1998 года в Москве

В. О. Бугаенко

Мы выбрали задачи, которые, на наш взгляд, являются наиболее сложными или интересными. К ним приведены полные решения, а к некоторым — несколько различных решений. В последнем случае мы старались, чтобы первое решение было более школьным — таким, какое, скорее всего, придумает участник олимпиады, а другие решения показывали задачу с более общей точки зрения, иногда даже выходя за рамки школьной программы.

При подготовке этой статьи многие люди помогали советами и предоставленными материалами. Это в первую очередь: С. С. Анисов, М. Н. Вялый, А. Я. Канель-Белов, Л. Э. Медников, М. Ю. Панов, А. В. Шаповалов, Г. Р. Челноков.

УСЛОВИЯ ЗАДАЧ

1. (LXI Московская математическая олимпиада, 9 кл.)

На отрезке $[0, 1]$ отмечено несколько различных точек. При этом каждая отмеченная точка расположена либо ровно посередине между двумя другими отмеченными точками (не обязательно соседними с ней), либо ровно посередине между отмеченной точкой и концом отрезка. Докажите, что все отмеченные точки рациональны.

В. Произволов

2. (LXI Московская математическая олимпиада, 10 кл.; Десятнадцатый Турнир Городов, весенний тур, 10–11 кл.)

Квадрат со стороной 1 разрезали на прямоугольники, у каждого из которых отметили одну сторону. Докажите, что сумма длин всех отмеченных сторон не может быть меньше 1.

В. Произволов

3. (LXI Московская математическая олимпиада, 10 кл.)

Дорога протяженностью 1 км полностью освещена фонарями, причем каждый фонарь освещает отрезок дороги длиной 1 м. Какое наибольшее количество фонарей может быть на дороге, если известно, что после выключения любого фонаря дорога будет освещена уже не полностью?

С. Агеев

4. (LXI Московская математическая олимпиада, 10 кл.)

На пол положили правильный треугольник ABC , выпиленный из фанеры. В пол вбили три гвоздя (по одному вплотную к каждой стороне треугольника) так, что треугольник невозможно повернуть, не отрывая от пола. Первый гвоздь делит сторону AB в отношении $1 : 3$, считая от вершины A , второй делит сторону BC в отношении $2 : 1$, считая от вершины B . В каком отношении делит сторону AC третий гвоздь?

А. Шень

5. (LXI Московская математическая олимпиада, 10–11 кл.)

Натуральные числа от 1 до n расставляются в ряд в произвольном порядке (число n фиксировано). Расстановка называется плохой, если в ней можно отметить 10

чисел (не обязательно стоящих подряд), идущих в порядке убывания. Остальные расстановки называются хорошими. Докажите, что количество хороших расстановок не превосходит 81^n .

А. Канель-Белов

6. (Десятнадцатый Турнир Городов, весенний тур, 8–9 кл.)

За круглым столом сидят десять человек, перед каждым — несколько орехов. Всего орехов — сто. По общему сигналу каждый передает часть своих орехов соседу справа: половину — если у него было четное число или один орех плюс половину остатка — если нечетное число. Такая операция прodelывается второй раз, затем третий и так далее, до бесконечности. Докажите, что через некоторое время у всех станет по десять орехов.

А. Шаповалов

7. (Десятнадцатый Турнир Городов, весенний тур, 10–11 кл.)

Внутренняя точка M выпуклого четырехугольника $ABCD$ такова, что треугольники AMB и CMD — равнобедренные ($AM = MB$, $CM = MD$) и у каждого угол при вершине M равен 120° . Докажите, что найдется точка N такая, что треугольники BNC и DNA — правильные.

И. Шарыгин

8. (Десятнадцатый Турнир Городов, весенний тур, 10–11 кл.)

а) Двое показывают карточный фокус. Первый снимает пять карт из колоды, содержащей 52 карты (предварительно перетасованной кем-то из зрителей), смотрит в них и после этого выкладывает их в ряд слева направо, причем одну из карт кладет рубашкой вверх, а остальные — картинкой вверх. Второй участник фокуса отгадывает закрытую карту. Докажите, что они могут так договориться, что второй всегда будет угадывать карту.

б) Второй фокус отличается от первого тем, что первый участник выкладывает слева направо четыре карты картинкой вверх, а одну не выкладывает. Могут ли в этом случае участники фокуса так договориться, чтобы второй всегда угадывал невыложенную карту?

Г. Гальперин

9. (Турнир им. Ломоносова, 8–11 кл.)

$\{a_1, a_2, \dots, a_n\}$ — набор целых положительных чисел. Строим новый набор чисел $\{b_0, b_1, b_2, \dots\}$ по следующему правилу: b_0 — количество чисел исходного набора, которые больше 0, b_1 — количество чисел исходного набора, которые больше 1, b_2 — количество чисел исходного набора, которые больше 2, и т. д., пока не пойдут нули. Докажите, что сумма всех чисел исходного набора равна сумме всех чисел нового набора.

фольклор

10. (Двадцатый Турнир Городов, осенний тур, 8–9 кл.)

Шайка разбойников отобрала у купца мешок монет. Каждая монета стоит целое число грошей. Оказалось, что какую бы монету ни отложить, оставшиеся монеты можно разделить между разбойниками так, чтобы каждый получил одинаковую сумму в грошах. Докажите, что если отложить одну монету, то число монет разделится на число разбойников.

А. Шаповалов

11. (Двадцатый Турнир Городов, осенний тур, 10–11 кл.)

Будем называть «размером» прямоугольного параллелепипеда сумму трех его измерений — длины, ширины и высоты. Может ли случиться, что в некотором прямоугольном параллелепипеде поместится больший по размеру прямоугольный параллелепипед?

А. Шень

12. (Двадцатый Турнир Городов, осенний тур, 10–11 кл.)

Дана функция $f(x) = (x^2 + ax + b)/(x^2 + cx + d)$, где трехчлены $x^2 + ax + b$ и $x^2 + cx + d$ не имеют общих корней. Докажите, что следующие два утверждения равносильны:

1) найдется числовой интервал, свободный от значений функции;

2) $f(x)$ представима в виде: $f(x) = f_1(f_2(\dots f_{n-1}(f_n(x))\dots))$, где каждая из функций $f_i(x)$ есть функция одного из видов: $k_i x + b_i, x^{-1}, x^2$. А. Канель-Белов

РЕШЕНИЯ ЗАДАЧ

1. Мы приведем три решения этой задачи. Первое будет чисто «школьным» и не потребует никаких знаний, выходящих за рамки школьной программы. Для понимания второго решения необходимо знание теории систем линейных уравнений, а третье требует знания основ линейной алгебры.

ПЕРВОЕ РЕШЕНИЕ. Обозначим координаты концов отрезка и отмеченных точек через x_0, x_1, \dots, x_{n+1} ($0 = x_0 < x_1 < \dots < x_{n+1} = 1$). Условие задачи означает выполнение n равенств вида

$$x_i = \frac{1}{2}(a_i + b_i) \quad (i = 1, 2, \dots, n),$$

где каждый из символов a_i и b_i означает какое-то из чисел x_j ($j = 0, 1, \dots, n+1$), при этом $a_i < x_i < b_i$.

Во все правые части этих равенств, в которых присутствует x_1 , подставим его значение из первого равенства. Получим новый набор равенств (с теми же левыми частями, что и в старом), правые части которых уже не содержат x_1 . Если при этом в правой части второго равенства появится член вида αx_2 , то перенесем его в левую часть и разделим обе части на $1 - \alpha$ (ниже мы докажем, что $\alpha \neq 1$). Затем во все правые части, содержащие x_2 , подставим его значение из второго равенства (быть может, уже измененного после первой подстановки). Далее то же сделаем для x_3 и т. д.

В любой момент каждое равенство будет иметь вид

$$x_i = \alpha_0 x_0 + \alpha_1 x_1 + \dots + \alpha_n x_n + \alpha_{n+1} x_{n+1}, \quad (*)$$

где все коэффициенты α_j ($j = 0, 1, \dots, n+1$) рациональны, неотрицательны и не равны нулю одновременно. Действительно, для исходного набора это верно. Делая очередную подстановку из j -го уравнения, мы заменяем коэффициент α_j на 0, а любой другой коэффициент α_k на $\alpha_k + \lambda \alpha_j$, где λ — коэффициент при x_k в j -м уравнении. Рациональность и неотрицательность при этом сохраняются, а наибольший номер ненулевого коэффициента не уменьшается, следовательно, всегда больше i . При переносе в левую часть члена $\alpha_i x_i$ получаем в правой части положительное число. Действительно, все x_j положительны, а все коэффициенты α_j неотрицательны, причем по крайней мере один из них строго положителен. Значит, левая часть тоже положительна, поэтому $1 - \alpha_j > 0$. При делении обеих частей уравнения на положительное рациональное число $1 - \alpha_j$ все перечисленные свойства также сохраняются.

Проделав указанные подстановки n раз, мы добьемся, чтобы коэффициенты α_j ($j = 1, \dots, n$) в правых частях равенств (*) равнялись нулю, тогда эти равенства будут иметь вид $x_i = \alpha_0 x_0 + \alpha_{n+1} x_{n+1}$. Значит, x_i — рациональные числа.

ВТОРОЕ РЕШЕНИЕ. Рассмотрим систему линейных уравнений, связывающих координаты отмеченных точек. Докажем, что она не может иметь более одного решения. Пусть (x'_1, \dots, x'_n) и (x''_1, \dots, x''_n) — какие-то два различных ее решения.

Определим смещение i -й точки формулой $d_i = |x'_i - x''_i|$ ($i = 1, 2, \dots, n$). Смещение концов отрезка считаем равным нулю.

Пусть $d > 0$ — максимальное из смещений. Среди всех точек со смещением d выберем самую правую. Она лежит посередине между какими-то двумя другими отмеченными точками, смещение одной из которых не превосходит по абсолютной величине d , а другой — строго меньше d . С другой стороны, ее смещение должно равняться среднему арифметическому смещений этих двух точек. Противоречие.

Если же система линейных уравнений с рациональными коэффициентами имеет единственное решение, то оно выражается по формулам Крамера через ее коэффициенты (возможно, предварительно нужно отбросить часть уравнений), поэтому является рациональным.

ТРЕТЬЕ РЕШЕНИЕ. Предположим, что существует набор точек, удовлетворяющий условию задачи, содержащий иррациональное число α .

Рассмотрим линейное (над множеством рациональных чисел) отображение множества действительных чисел в себя, оставляющее на месте все рациональные числа и переводящее α в число, большее 1 или меньшее 0. Применяя это отображение к множеству отмеченных точек, получим новый набор, удовлетворяющий условиям задачи и содержащий точки вне отрезка $[0, 1]$. Однако, это невозможно, так как крайняя из этих точек (наиболее удаленная от концов отрезка) не может лежать посередине между двумя другими.

2. ПЕРВОЕ РЕШЕНИЕ. Так как длины сторон получившихся прямоугольников не превосходят 1, то площадь каждого из них не превосходит длины меньшей стороны. Значит, сумма длин отмеченных сторон прямоугольников не меньше суммы их площадей, т. е. 1.

ВТОРОЕ РЕШЕНИЕ. Проведем две прямые, содержащие смежные стороны квадрата. Спроецируем каждый прямоугольник на ту из проведенных прямых, которой параллельна его отмеченная сторона. Если проекции полностью покрыли одну из сторон квадрата, то уже сумма параллельных ей отмеченных сторон больше 1. Если же на двух сторонах квадрата есть по крайней мере по точке, не покрытой проекциями прямоугольников, то существует точка внутри квадрата, проекции которой на стороны не покрыты. Тем самым, она не содержится ни в одном прямоугольнике. Это противоречит условию, значит этот случай невозможен.

3. ОТВЕТ: 1998 фонарей.

Будем считать, что фонари стоят в серединах освещаемых ими отрезков. Пронумеруем фонари в порядке их следования вдоль дороги. Тогда отрезки, освещаемые k -м и $(k+2)$ -м фонарями не пересекаются (так как в противном случае при выключении $(k+1)$ -го фонаря освещенная область не изменилась бы). Значит, сумма длин отрезков, освещенных фонарями с нечетными номерами, меньше длины дороги, следовательно, их количество не превосходит 999. Аналогично, фонарей с четными номерами также не больше 999. Поэтому всего фонарей не больше 1998.

С другой стороны, 1998 фонарей быть может. Например, пусть первый и последний фонари стоят на расстоянии полметра от концов дороги, а остальные расположены равномерно вдоль дороги. При этом расстояние между соседними фонарями равно $\frac{999}{1997}$ м, что больше, чем полметра. Значит, области,

освещаемые i -м и $(i+2)$ -м фонарями, не пересекаются. Зазор между этими областями освещается только $(i+1)$ -м фонарем. Поэтому при выключении любого фонаря появятся неосвещенные участки дороги.

4. Найдем прежде всего условие того, что вбитый в точке M на стороне треугольника гвоздь «мешает» повернуть этот треугольник вокруг некоторой внутренней точки O . Проведем окружность с центром в O , проходящую через точку M . Пусть эта окружность пересекает сторону треугольника так, что двигаясь по окружности по часовой стрелке от точки M , мы попадаем внутрь треугольника (рис. 1 а). Тогда при повороте на маленький угол против часовой стрелки,

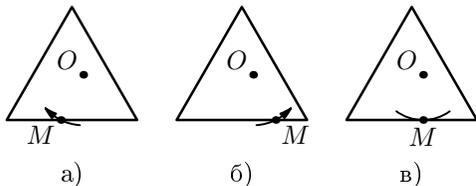


Рис. 1.

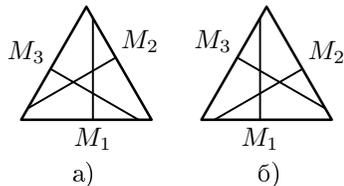


Рис. 2.

в точку, пробитую гвоздем, попадает некоторая внутренняя точка треугольника. Это означает, что гвоздь препятствует такому повороту треугольника. Повернуть же его на маленький угол по часовой стрелке этот гвоздь не мешает. Если же двигаясь по окружности по часовой стрелке, мы попадаем наружу треугольника (рис. 1 б), то треугольник можно повернуть вокруг точки O против часовой стрелки, но нельзя в обратную сторону. Наконец, если проведенная окружность касается стороны в точке M (рис. 1 в), то повернуть треугольник вокруг точки O нельзя ни в какую сторону.

Восставим перпендикуляр к стороне треугольника в точке M . Тогда для любой точки O , лежащей на этом перпендикуляре, поворот треугольника относительно нее невозможен. Для остальных точек возможен поворот в одну или в другую сторону в зависимости от того, по какую сторону от восстановленного перпендикуляра лежит точка. А именно, если точка O лежит по правую сторону от перпендикуляра (если смотреть внутрь треугольника от его стороны), то возможен поворот по часовой стрелке, а если по левую, то против.

Пусть теперь три гвоздя вбиты в точках M_1 , M_2 и M_3 на сторонах AB , BC и CA соответственно. Если три перпендикуляра, восстановленные в этих точках к сторонам треугольника, не пересекаются в одной точке, то среди частей, на которые они разбивают пространство, имеется треугольник, лежащий строго внутри данного треугольника. Каждая точка внутри этого треугольника лежит по одну сторону от всех трех перпендикуляров: либо по левую (рис. 2 а), либо по правую (рис. 2 б). Значит, в этом случае возможен поворот треугольника, что противоречит условию задачи. Следовательно, три восстановленных перпендикуляра пересекаются в одной точке. В этом случае для точки, взятой в любой из шести областей, на которые перпендикуляры делят треугольник, повороту вокруг нее в каждом направлении препятствует какой-то гвоздь.

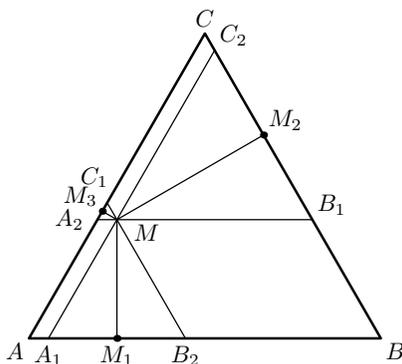


Рис. 3.

Докажем, что если рассматриваемые перпендикуляры пересекаются в одной точке, то

$$AM_1 + BM_2 + CM_3 = M_1B + M_2C + M_3A.$$

Проведем для этого прямые, параллельные сторонам треугольника, через точку M пересечения перпендикуляров. Точки пересечения этих прямых со сторонами треугольника обозначим $A_1, A_2, B_1, B_2, C_1, C_2$, как на рис. 3. Тогда $A_1M_1 = M_1B_2$, $B_1M_2 = M_2C_2$, $C_1M_3 = M_3A_2$, поскольку высоты равнобедренных треугольников A_1MB_2 , B_1MC_2 и C_1MA_2 являются также медианами. Кроме того, $AA_1 = C_2C$, $BB_1 = A_2A$, $CC_1 = B_2B$, как боковые стороны равнобедренных трапеций. Складывая шесть равенств, получаем требуемое.

Отсюда получаем $\frac{M_3A}{M_3C} = \frac{5}{7}$.

5. С каждой расстановкой чисел от 1 до n свяжем две конечные последовательности $\{a_i\}$ и $\{b_j\}$ ($1 \leq i, j \leq n$). Определим a_i (глубину числа i) как наибольшее количество чисел в расстановке, идущих в порядке убывания, последнее из которых равно i , а b_j (глубину места j) как наибольшее количество чисел в расстановке, идущих в порядке убывания, последнее из которых стоит на j -м месте.

Зная последовательности $\{a_i\}$ и $\{b_j\}$, мы можем однозначно восстановить исходную расстановку. Действительно, на каждое место можно поставить только число, имеющее ту же глубину, что и это место. Расставить же числа данной глубины по местам этой же глубины можно единственным способом — в порядке возрастания, иначе мы придем к противоречию с определением глубины.

Для хорошей расстановки все элементы последовательностей $\{a_i\}$ и $\{b_j\}$ могут принимать значения от 1 до 9. Таких пар последовательностей всего имеется $(9^n)^2 = 81^n$. При этом каждая из них может соответствовать не более, чем одной расстановке. Поэтому хороших расстановок не больше.

6. ПЕРВОЕ РЕШЕНИЕ. Обозначим через N максимальное количество орехов, имеющих в данный момент у одного человека. Перекладывая орехов приводят к изменению N . Проследим за этими изменениями. Докажем сначала, что N не может возрасти после перекалывания орехов. Действительно, если N четно, то каждый человек оставляет себе не более $N/2$ орехов, а получает также не

более $N/2$ орехов, а если N нечетно, то каждый человек оставляет себе не более $(N-1)/2$ орехов, а получает не более $(N+1)/2$ орехов.

Предположим, что $N > 10$. Тогда существует человек, имеющий меньше 10 орехов (в противном случае общее количество орехов было бы больше 100).

Докажем, что после какого-то количества переключиваний N уменьшится. Из этого сразу следует утверждение задачи, поскольку в любом случае $N \geq 10$. Рассмотрим серию переключиваний, при которой N не меняется. Чтобы сделать анализ наглядным, предположим, что каждый человек держит флажок: синий, если у него N орехов, белый — если $(N-1)$ орех, красный — в остальных случаях. Человек, имеющий не более 9 орехов, может держать только красный флажок. Поэтому хотя бы один красный флажок обязательно присутствует. Будем доказывать, что количество синих флажков уменьшается, так что серия рано или поздно завершится. Рассмотрим отдельно два случая.

Первый случай: N четно и $N \geq 12$. Синие флажки не добавляются (держатый не синий флажок оставляет у себя не более $(N/2-1)$, а получает не более $N/2$ орехов) и не двигаются, т. е. после очередного переключивания синие флажки будут только у тех, кто их держал до переключивания. Красные флажки «двигаются» (тот, кто сидит после человека с красным флажком, получает не более $(N/2-2)$, а оставляет не более $N/2$ орехов, так что после этой операции он возьмет в руки красный флажок). Так как хотя бы один красный флажок существует, то количество держащих синие флажки обязательно уменьшится после некоторого количества шагов (синий и красный флажок неминуемо «столкнутся»).

Второй случай: N нечетно. В отличие от первого случая, новые синие флажки могут появляться. Изменим несколько правила. Предположим, что перед тем, как осуществить передачу орехов, каждый передал всю имеющуюся у него кучу своему соседу слева. Ясно, что такая операция не влияет на распределение орехов между людьми. При этом рассматриваемый случай становится аналогичным уже разобранному.

ВТОРОЕ РЕШЕНИЕ. Обозначим через $f_+(n)$ количество орехов, которое должен отдать своему правому соседу, человек, имеющий n орехов, а через $f_-(n)$ количество орехов, которое при этом он оставляет себе. Очевидно, что

$$f_+(n) = \begin{cases} \frac{n}{2}, & \text{если } n \text{ четно,} \\ \frac{n+1}{2}, & \text{если } n \text{ нечетно;} \end{cases} \quad f_-(n) = \begin{cases} \frac{n}{2}, & \text{если } n \text{ четно,} \\ \frac{n-1}{2}, & \text{если } n \text{ нечетно.} \end{cases}$$

Пронумеруем людей за столом числами от 1 до 10 против часовой стрелки и обозначим через $a_i = 10 + b_i$ количество орехов у i -го человека. Заметим, что $\sum_{i=1}^{10} b_i = 0$. Пусть после одной операции количество орехов у i -го человека стало $a'_i = 10 + b'_i$. Тогда легко подсчитать, что $a'_i = f_-(a_i) + f_+(a_{i-1})$, $b'_i = f_-(b_i) + f_+(b_{i-1})$ (мы предполагаем, что нумерация людей проведена «по модулю 10»; в частности, нулевой человек суть десятый).

Рассмотрим сумму $S = \sum_{i=1}^{10} |b_i| = \sum_{i=1}^{10} |a_i - 10|$ и проследим, как она изменяется при каждой операции. Воспользуемся при этом неравенством о модуле суммы и

тем, что при любом целом n числа $f_-(n)$ и $f_+(n)$ не могут быть разных знаков, откуда $n = |f_+(n)| + |f_-(n)|$. Докажем, что величина S не возрастает:

$$\begin{aligned} \sum_{i=1}^{10} |b'_i| &= \sum_{i=1}^{10} |f_-(b_i) + f_+(b_{i-1})| \leq \\ &\leq \sum_{i=1}^{10} (|f_-(b_i)| + |f_+(b_{i-1})|) = \sum_{i=1}^{10} |f_-(b_i)| + \sum_{i=1}^{10} |f_+(b_{i-1})| = \\ &= \sum_{i=1}^{10} |f_-(b_i)| + \sum_{i=1}^{10} |f_+(b_i)| = \sum_{i=1}^{10} (|f_-(b_i)| + |f_+(b_i)|) = \sum_{i=1}^{10} |b_i|. \end{aligned}$$

Докажем теперь, что если не все b_i равны нулю, то S уменьшится (быть может, после нескольких операций). Прежде всего заметим, что среди b_i должны быть числа как положительные, так и отрицательные (в противном случае их сумма не могла бы равняться нулю), иными словами, есть люди, имеющие больше 10 орехов, и люди, имеющие меньше 10 орехов. Тогда обязательно найдутся такие j и i , что $a_j > 10$, $a_{j+1} = a_{j+2} = \dots = a_{i-1} = 10$ и $a_i < 10$ (действительно, временно отсадив всех, имеющих по 10 орехов, и двигаясь вдоль стола против часовой стрелки, мы обязательно в какой-то момент перейдем от человека, имеющего больше 10 орехов, к человеку, имеющему меньше 10 орехов). На следующем шагу мы будем иметь $a_{j+1} > 10$ (так как $(j+1)$ -й человек отдал 5 орехов, а получил больше), $a_{j+2} = \dots = a_{i-1} = 10$ (поскольку все эти люди отдали и получили по 5 орехов) и, по-прежнему, $a_i < 10$ (так как i -й человек, отдавая орехи, оставил себе не больше четырех, а получил 5). Значит, через несколько шагов окажется $a_{i-1} > 10$, $a_i < 10$, т. е. $b_{i-1} > 0$, $b_i < 0$. В этом случае $f_+(b_{i-1}) > 0$ и $f_-(b_i) < 0$, поэтому неравенство о модуле суммы, которым мы воспользовались, превращается в строгое. Значит, величина S на следующем шагу уменьшится. Поскольку величина S не может уменьшаться бесконечно долго, в некоторый момент все люди за столом будут иметь по 10 орехов.

7. ПЕРВОЕ РЕШЕНИЕ. По условию $\angle AMB = \angle CMD$, прибавив к обеим частям этого равенства $\angle BMC$, получаем $\angle AMC = \angle BMD$. Кроме того, $AM = BM$ и $CM = DM$, поэтому треугольники AMC и BMD равны по двум сторонам и углу между ними. Значит, $AC = BD$ и $\angle MAC = \angle MBD$. Пусть O — точка пересечения диагоналей четырехугольника $ABCD$ (рис. 4 а). Из равенства $\angle MAO = \angle MBO$ и обратной теоремы о вписанном угле следует, что точки A , B , O и M лежат на одной окружности, и следовательно, $\angle AOB = \angle AMB = 120^\circ$.

Рассмотрим такую точку N (лежащую по ту же сторону от прямой AD , что и четырехугольник $ABCD$), что треугольник ADN равносторонний (рис. 4 б). Соединим ее со всеми вершинами четырехугольника $ABCD$. Поскольку $\angle AND = 60^\circ = 180^\circ - \angle AOB = \angle AOD$, из обратной теоремы о вписанном угле следует, что точки A , D , N и O лежат на одной окружности, значит $\angle OAN = \angle ODN$. Следовательно, треугольники ANC и DNB равны по двум сторонам и углу между ними. Поэтому $\angle DNB = \angle ANC$, вычитая из обеих частей этого равенства $\angle DNC$, получаем $\angle CNB = \angle AND = 60^\circ$. Кроме того, $CN = BN$. Это означает, что треугольник CNB правильный. Точка N является искомой.

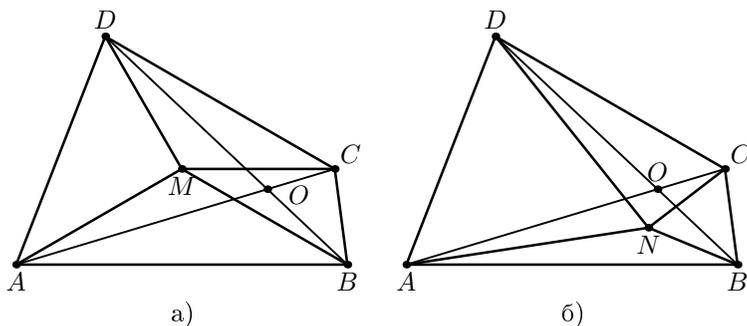


Рис. 4.

ВТОРОЕ РЕШЕНИЕ. Докажем сначала следующее утверждение.

ЛЕММА. Пусть направленные отрезки $\overrightarrow{A_1B_1}$ и $\overrightarrow{A_2B_2}$ имеют равные длины, и угол между ними равен α . Тогда их можно перевести друг в друга поворотом на угол α вокруг некоторой точки.

Переведем направленный отрезок $\overrightarrow{A_1B_1}$ в $\overrightarrow{A_2B_2}$ с помощью композиции двух осевых симметрий. Ось первой симметрии возьмем серединный перпендикуляр к отрезку A_1A_2 , при этом точка A_1 перейдет в A_2 , а B_1 — в некоторую точку B' . Ось второй симметрии — биссектриса угла $B'A_2B_2$. Композиция двух осевых симметрий — это либо параллельный перенос (если оси параллельны), либо поворот на угол α (если оси пересекаются под углом $\alpha/2$). В рассматриваемом случае при применении композиции двух осевых симметрий направление вектора $\overrightarrow{A_1A_2}$ изменилось на угол α , поэтому это мог быть только поворот на угол α . Лемма доказана.

Фактически мы доказали, что любое сохраняющее ориентацию движение плоскости является либо параллельным переносом, либо поворотом. Это утверждение называется теоремой Шаля.

Перейдем теперь к доказательству утверждения задачи. Из условия следует, что при повороте на угол 120° относительно точки M вершина A переходит в B , а C в D . Значит, при этом повороте направленный отрезок \overrightarrow{AC} переходит в направленный отрезок \overrightarrow{BD} , поэтому их длины равны, а угол между ними равен 120° . Значит, угол между направленными отрезками \overrightarrow{AC} и \overrightarrow{DB} равен 60° . Согласно лемме, существует точка N , при повороте вокруг которой на 60° точка A перейдет в D , а C в B . Она и является искомой.

8. а) Пять карт можно выложить в различном порядке 120 способами, а количество карт, из которого нужно отгадать закрытую, равно 48. Поэтому для решения задачи достаточно найти способ закодировать закрытую карту с помощью порядка выкладывания пяти карт.

Прежде всего, фокусники договариваются об упорядочении по какому-либо правилу всех карт колоды, при этом каждая карта получит номер от 1 до 52. Кроме того, они нумеруют все перестановки из пяти элементов числами от 1 до 120.

Пусть первый фокусник произвольным образом выбирает одну из пяти полученных карт (которую он оставит закрытой) и находит ее номер N . Сопоставив оставшимся четырем картам цифры 1, 2, 3 и 4 по старшинству, а закрытой карте — цифру 5, он выкладывает их в порядке, соответствующем перестановке с номером N . Второй видит порядок выкладывания карт и поэтому может отгадать закрытую карту.

б) Если действовать так же, как и в пункте а) — пытаться закодировать номер закрытой карты с помощью порядка выкладывания четырех карт, то этого не получится: закрытой может оказаться одна из 48 карт, а возможных способов последовательного выкладывания четырех карт всего 24. Тем не менее, ответ и в этом случае положительный. Для этого фокусникам нужно использовать то, что выбор одной из пяти карт, которую нужно будет отгадывать, тоже находится в их руках.

Опять считаем, что карты в колоде и перестановки из четырех элементов пронумерованы. Пусть первый фокусник получает пять карт, номера которых равны a_1, a_2, a_3, a_4 и a_5 (в порядке возрастания). Если $a_5 - a_4 \leq 24$, то он откладывает старшую карту (с номером a_5) и выкладывает оставшиеся четыре в порядке, соответствующем числу $a_5 - a_4$. Если же $a_5 - a_4 \geq 25$, то тогда $a_5 - a_1 \geq 28$, поэтому $52 + a_1 - a_5 \leq 24$. В этом случае первый фокусник откладывает младшую карту (с номером a_1) и выкладывает четверку карт в порядке, соответствующем числу $52 + a_1 - a_5$.

Теперь второму фокуснику, чтобы узнать номер закрытой карты, достаточно к номеру карты, старшей в выложенной четверке, прибавить номер выложенной последовательности и, если полученное число больше 52, вычесть 52.

9. Рассмотрим следующую конструкцию. Возьмем набор кубиков и будем строить из них «башенки» — столбцы высотой a_1, a_2, \dots, a_n (рис. 5). Посчитаем двумя способами, сколько кубиков нам для этого понадобится. Считая по столбцам, получаем, что количество кубиков равно сумме чисел первого набора. Другой способ подсчета — по слоям. Количество кубиков в первом слое (стоящих на полу) равно $b_0 = n$, в следующем слое — b_1 , и т. д., количество кубиков в i -м слое равно b_i . Поэтому общее число кубиков равно сумме чисел второго набора. Значит, суммы чисел обоих наборов совпадают.

ЗАМЕЧАНИЕ. Похожая конструкция используется при введении понятия интеграла. Чтобы подсчитать площадь криволинейной фигуры под графиком функции (интеграл), ее приближают ступенчатыми функциями, графики которых состоят из горизонтальных отрезков. Площадь фигуры под ступенчатой функцией (интегральную сумму) можно подсчитать двумя способами, аналогичными приведенным в решении способам подсчета числа кубиков. Получаемые пределы интегральных сумм в этих двух случаях называются *интегралом Римана* и *интегралом Лебега* соответственно.

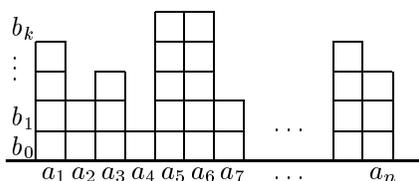


Рис. 5.

10. Можно считать, что стоимости монет в грошах взаимно просты. Действительно, если они имеют наибольший общий делитель $d > 1$, то деноминируем

грош, приравняв один новый к d старым. Тогда все условия задачи по-прежнему выполнены, но новые стоимости монет будут взаимно простыми.

Пусть n разбойников отняли m монет на общую сумму в g грошей. Так как при вычитании из g стоимости любой монеты получим число, кратное n , то стоимости всех монет дают при делении на n один и тот же остаток r . Так как стоимость любого набора из $m - 1$ монеты делится на n , то $(m - 1)r$ делится на n . Кроме того, r и n взаимно просты, поскольку любой их общий делитель является общим делителем стоимостей всех монет. Отсюда следует, что $m - 1$ делится на n .

11. ОТВЕТ: не может.

ПЕРВОЕ РЕШЕНИЕ. Пусть мы имеем прямоугольный параллелепипед P' , лежащий внутри прямоугольного параллелепипеда P . Оценим двумя способами сумму l длин проекций трех взаимно перпендикулярных ребер параллелепипеда P' на три прямых, параллельных ребрам параллелепипеда P .

С одной стороны, сумма проекций трех взаимно перпендикулярных ребер параллелепипеда P' на любое ребро параллелепипеда P не превосходит длины этого ребра. Действительно, проекция внутреннего параллелепипеда на любое ребро внешнего есть отрезок, концы которого — это проекции двух противоположных вершин. От одной из них до другой можно пройти по трем взаимно перпендикулярным ребрам. Поэтому рассматриваемая проекция равна сумме проекций этих трех ребер, она, очевидно, не больше, чем длина ребра, на которое мы проецируем. Значит, l не превосходит размера параллелепипеда P .

С другой стороны, длина любого отрезка не превосходит суммы его проекций на три взаимно перпендикулярных направления. Поэтому размер параллелепипеда P' не больше l .

Итак, размер внутреннего параллелепипеда не больше размера внешнего.

ВТОРОЕ РЕШЕНИЕ. Сначала докажем, что площадь поверхности внутреннего параллелепипеда не больше площади поверхности внешнего. Приведенное ниже доказательство годится для любого выпуклого многогранника, содержащегося внутри другого многогранника.

Наденем на каждую грань внутреннего многогранника «шапку». *Шапкой* мы назовем объединение всех лучей с вершинами на грани, перпендикулярных ей и направленных наружу многогранника. Другими словами, шапка — это множество точек, проекции которых на плоскость грани лежат внутри грани и расположены по другую сторону от этой плоскости, чем многогранник. Из выпуклости многогранника следует, что любая точка пространства лежит не более, чем в одной шапке.

Куски поверхности внешнего многогранника, попавшие внутрь каждой шапки, спроецируем на соответствующую грань. Их площадь не меньше площади этой грани, так как проекции покроют всю грань, а при проецировании площадь не увеличивается. Просуммировав эти неравенства для всех граней внутреннего многогранника, получаем, что площадь его поверхности не превосходит площади части поверхности внешнего многогранника, попавшей внутрь шапок, которая, в свою очередь, не превосходит площади всей поверхности внешнего многогранника.

Пусть прямоугольный параллелепипед с ребрами a' , b' и c' содержится внутри

прямоугольного параллелепипеда с ребрами a , b и c . Тогда согласно доказанному

$$2(a'b' + a'c' + b'c') \leq 2(ab + ac + bc).$$

Кроме того, диаметр (т. е. расстояние между двумя наиболее удаленными точками) внутреннего параллелепипеда не превосходит диаметра внешнего. Диаметром прямоугольного параллелепипеда является длина его диагонали, поэтому, сравнивая квадраты диагоналей, имеем

$$a'^2 + b'^2 + c'^2 \leq a^2 + b^2 + c^2.$$

Сложив эти два неравенства, получаем

$$(a' + b' + c')^2 \leq (a + b + c)^2.$$

Значит, $a' + b' + c' \leq a + b + c$.

ТРЕТЬЕ РЕШЕНИЕ. Назовем ε -окрестностью фигуры F фигуру F_ε , состоящую из точек, отстоящих на расстояние не больше ε от хотя бы одной точки фигуры F (в частности, $F_0 = F$). Любая точка M из ε -окрестности P_ε многогранника P имеет один из следующих четырех типов в зависимости от того, какое из следующих условий выполняется:

- ▷ M лежит внутри P ;
- ▷ ближайшая к M точка многогранника P принадлежит его грани;
- ▷ ближайшая к M точка многогранника P принадлежит его ребру;
- ▷ ближайшая к M точка многогранника P является его вершиной.

Найдем объем $\text{vol } P_\varepsilon$ ε -окрестности P_ε прямоугольного параллелепипеда P с ребрами a , b и c . Для этого разрежем его на четыре части P^0 , P^1 , P^2 и P^3 , являющихся множествами точек указанных выше типов соответственно.

Объем P^0 равен abc . P^1 представляет собой объединение шести прямоугольных параллелепипедов, основания которых являются гранями параллелепипеда P , а высоты равны ε , поэтому $\text{vol } P^1 = 2(ab + ac + bc)\varepsilon$. P^2 — это двенадцать «четвертушек» цилиндров радиуса ε , высоты которых суть ребра параллелепипеда. Значит, $\text{vol } P^2 = (a + b + c)\pi\varepsilon^2$. P^3 — это восемь «осьмушек» шара радиуса ε (каждая «осьмушка» представляет собой пересечение октанта с вершиной в вершине параллелепипеда и шара радиуса ε с центром в ней же). Поэтому $\text{vol } P^3 = \frac{4}{3}\pi\varepsilon^3$ — это объем шара радиуса ε .

Окончательно имеем

$$\text{vol } P_\varepsilon = \frac{4}{3}\pi\varepsilon^3 + (a + b + c)\pi\varepsilon^2 + 2(ab + ac + bc)\varepsilon + abc.$$

Если $F' \subset F$, то $F'_\varepsilon \subset F_\varepsilon$ для любого ε . Пусть P — прямоугольный параллелепипед с ребрами a , b и c , а P' — содержащийся в нем прямоугольный параллелепипед с ребрами a' , b' и c' . Тогда для любого ε фигура P'_ε содержится внутри P_ε , и поэтому $\text{vol } P'_\varepsilon \leq \text{vol } P_\varepsilon$. Обе части этого неравенства — многочлены третьей степени от ε с равными коэффициентами при старшем члене. Значит, коэффициент при ε^2 в левой части не больше, чем коэффициент при ε^2 в правой части. Это означает, что $a' + b' + c' \leq a + b + c$, т. е. размер параллелепипеда P' не превосходит размера параллелепипеда P .

12. ПЕРВОЕ РЕШЕНИЕ. Прежде всего докажем, что любая дробно-линейная (т. е. имеющая вид $f(x) = \frac{ax+b}{cx+d}$) функция, не являющаяся константой, представляется в виде композиции функций двух первых видов (линейных и взятия обратной величины).

Если $c = 0$, то утверждение очевидно (функция является линейной). Если $ad = bc$, то функция является константой. В остальных случаях

$$\frac{ax+b}{cx+d} = f_1(f_2(f_3(x))),$$

где $f_1(t) = t + \frac{a}{c}$, $f_2(t) = t^{-1}$, $f_3(t) = \frac{c^2}{bc-ad}t + \frac{cd}{bc-ad}$.

Обратно, композиция любого количества линейных функций и взятия обратной величины является дробно-линейной функцией. Действительно, сами по себе эти функции являются дробно-линейными, и их композиция с любой дробно-линейной функцией также является дробно-линейной функцией.

Докажем, что 2) \Rightarrow 1). Пусть данная функция $f(x)$ представима в виде композиции линейных функций, взятия обратной величины и возведения в квадрат. Так как $f(x)$ не является дробно-линейной, то возведение в квадрат обязательно присутствует. Поскольку квадрат функции принимает только неотрицательные значения, появляется интервал, свободный от значений функции. С другой стороны, присутствие интервала, свободного от значений функции, сохраняется, если мы применяем к ней линейную функцию или взятие обратной величины.

Теперь докажем, что 1) \Rightarrow 2). Пусть имеется функция $f(x)$, для которой найдется интервал, не содержащий ее значений. Опять считаем, что $f(x) \neq \text{const}$.

Функцией, обратной к $\varphi(t) = t^{-1}$, является она сама, а обратная функция к линейной (но не являющейся константой) — тоже линейная. Поэтому достаточно доказать, что квадрат некоторой дробно-линейной функции представляется как композиция нескольких линейных функций и взятий обратной величины с данной функцией $f(x)$.

Прежде всего вычитаем из $f(x)$ единицу и обращаем дробь, получаем функцию вида $\frac{x^2 + cx + d}{px + q}$. Далее, если $p = 0$, то применяем линейную функцию $qt + \left(\frac{c^2}{4} - d\right)$ и получаем полный квадрат $\left(x + \frac{c}{2}\right)^2$. Если же $p \neq 0$, то делаем линейную замену переменных $x' = x + \frac{q}{p}$ и получаем $\frac{x'^2 + rx' + s}{px'}$. При этом $s \neq 0$, так как по условию числитель и знаменатель не имеют общих корней. Рассмотрим два случая.

Первый случай: $s > 0$, можно считать, что $s = u^2$. Применим функцию $\frac{p}{u}t - \frac{r}{u}$ и сделаем новую замену $x'' = \frac{x'}{u}$. Получим

$$\frac{x''^2 + 1}{x''} = \frac{4}{\left(\frac{2}{x'' - 1} + 1\right)^2 - 1} + 2.$$

Второй случай: $s < 0$. Аналогично, получаем функцию $\frac{x''^2 - 1}{x''}$. Непосред-

ственно проверяем, что при любом λ уравнение $\frac{x''^2 - 1}{x''} = \lambda$ имеет решение. Значит, не существует интервала, свободного от значений функции.

ВТОРОЕ РЕШЕНИЕ. Докажем только импликацию $1) \Rightarrow 2)$.

Для начала сведем задачу к случаю, когда множество значений функции $f(x)$ ограничено. В самом деле, пусть $f(x)$ не принимает значений из интервала (x_1, x_2) , тогда функция $g(x) = \left(f(x) - \left(\frac{x_1 + x_2}{2}\right)\right)^{-1}$ ограничена сверху числом $\frac{2}{x_2 - x_1}$.

Докажем, что если четыре точки с координатами (a, b) , (c, d) , (a', b') и (c', d') лежат на одной прямой, то $\frac{x^2 + ax + b}{x^2 + cx + d} = \varphi \left(\frac{x^2 + a'x + b'}{x^2 + c'x + d'}\right)$, где φ — дробно-линейная функция.

Действительно,

$$\begin{aligned} a &= \mu a' + (1 - \mu)c', & b &= \mu b' + (1 - \mu)d', \\ c &= \nu a' + (1 - \nu)c', & d &= \nu b' + (1 - \nu)d', \end{aligned}$$

откуда

$$\frac{x^2 + ax + b}{x^2 + cx + d} = \frac{\mu \frac{x^2 + a'x + b'}{x^2 + c'x + d'} + (1 - \mu)}{\nu \frac{x^2 + a'x + b'}{x^2 + c'x + d'} + (1 - \nu)}.$$

Пусть $f(x) = \frac{x^2 + ax + b}{x^2 + cx + d}$ — ограниченная функция, удовлетворяющая условию 1). Из ограниченности следует, что знаменатель не обращается в нуль, следовательно, квадратный трехчлен имеет отрицательный дискриминант ($c^2 < 4d$). Это значит, что точка (c, d) лежит «внутри» параболы $y = x^2/4$. Проведем на координатной плоскости прямую, проходящую через точки (a, b) и (c, d) . Она должна пересечь параболу; точки пересечения обозначим (a', b') и (c', d') . Тогда по доказанному утверждению $f(x)$ выражается в виде дробно-линейной функции от $\frac{x^2 + a'x + b'}{x^2 + c'x + d'}$, а последняя дробь является квадратом дробно-линейной функции, так как и в числителе, и в знаменателе содержит квадратный трехчлен с нулевым дискриминантом.

Отдельно следует рассмотреть случай, когда проведенная прямая пересекает параболу лишь в одной точке, т. е. является вертикальной. Тогда $a = c$, и этот случай легко разбирается.

ТРЕТЬЕ РЕШЕНИЕ. Областью значений $E(f)$ функции $f(x)$ может быть либо вся действительная прямая, либо отрезок, либо луч, либо объединение двух лучей. Действительно, $E(f)$ — это множество таких значений y , что уравнение $x^2 + ax + b = y(x^2 + cx + d)$ имеет действительные корни. Это условие означает, что дискриминант получающегося квадратного уравнения относительно x неотрицателен, и может быть записано в виде квадратичного неравенства относительно y . Множество решений квадратичного неравенства может иметь только один из указанных видов. В любом из этих случаев, применив дробно-линейную функцию, мы можем перевести область определения в положительную полуось.

Докажем, что если $E(f) = [0, +\infty)$, то $f(x)$ представляется в виде квадрата дробно-линейной функции. Предположим, что числитель дроби $f(x) = \frac{P(x)}{Q(x)}$ принимает значения разных знаков. Значит, его график пересекает ось абсцисс в некоторой точке x_0 , являющейся его корнем. Поскольку x_0 не является корнем знаменателя, знаменатель $Q(x)$, а значит, и вся дробь, принимает в окрестности точки x_0 значения разных знаков, что противоречит принятому предположению об области определения функции. Аналогично, знаменатель не может принимать значений разных знаков.

С другой стороны, числитель должен принимать нулевое значение, так как нуль входит в область значений дроби. Знаменатель также должен принимать нулевое значение, так как в противном случае значения дроби были бы ограничены. Поскольку числитель и знаменатель — квадратные трехчлены, их области значений суть множества неотрицательных чисел. Это значит, что они оба являются квадратами линейных функций.

ЗАМЕЧАНИЕ. Строго говоря, утверждение задачи не вполне корректно и требует некоторого уточнения. Мы предполагаем, что в правой и левой частях равенства

$$\left(\frac{P(x)}{Q(x)}\right)^{-1} = \frac{Q(x)}{P(x)}, \quad (*)$$

стоят равные функции, хотя на самом деле их области определения не всегда совпадают. Чтобы избежать этой проблемы, можно принять одно из следующих предположений:

1) считать две функции равными, если их значения совпадают на пересечении их областей определения; при этом допустить, что функция может быть не определена для конечного множества значений аргумента;

2) дополнить множество действительных чисел еще одним «числом» ∞ , удовлетворяющим естественным арифметическим свойствам $\frac{a}{\infty} = 0$ и $\frac{a}{0} = \infty$ (при $a \neq 0$ или ∞), и считать встречающиеся в задаче функции определенными и принимающими значения в таком расширении множества действительных чисел;

3) считать встречающиеся в задаче дроби $\frac{P(x)}{Q(x)}$ не функциями, а формальными записями, над которыми можно производить арифметические операции по естественным правилам, в частности, взятие обратной величины определяется формулой (*); равенство функций на их общей области определения влечет в этом случае равенство формальных записей с точностью до умножения числителя и знаменателя на один и тот же множитель.

Связь основной теоремы алгебры с теорией непрерывных групп

Б. Р. Френкин

Темой первого выпуска настоящей серии стала теорема об алгебраической замкнутости поля комплексных чисел [1, с. 30 – 95]. Как отмечено в одной из статей указанного выпуска, «каждый раздел математики имеет собственное доказательство основной теоремы алгебры» [1, с. 71, Е. А. Горин «От спектрального радиуса к основной теореме алгебры»]. Большинство этих доказательств не предназначено, так сказать, для первого знакомства с теоремой, поскольку сами они опираются на более сложные результаты. Их смысл в другом: выявить связь между существенными математическими фактами, на первый взгляд далекими друг от друга. По меньшей мере два таких доказательства, естественных по своей логике, можно провести и в терминах теории непрерывных групп.

Теорема о существовании комплексного корня у любого многочлена ненулевой степени с комплексными коэффициентами равносильна следующему утверждению: поле вещественных чисел не имеет конечномерных расширений размерности выше двух. В такой форме мы и будем ее доказывать.

Пусть S — конечномерное собственное расширение поля вещественных чисел. Иными словами, S - вещественное линейное пространство конечной размерности $N \geq 2$, на котором дополнительно введена операция умножения со следующими свойствами: 1) выполнены аксиомы поля; 2) S содержит подполе R , изоморфное полю вещественных чисел; 3) умножение на элементы R совпадает с умножением на вещественные числа в смысле вещественного линейного пространства (и потому R является одномерным подпространством). Покажем, что операцию с такими свойствами можно ввести лишь при $N = 2$.

Операции умножения и взятия обратного непрерывны в естественной топологии пространства S . Читатель может убедиться в этом, используя непрерывность операций сложения векторов и умножения на вещественные коэффициенты (см. также Дополнение в конце заметки). Таким образом, ненулевые элементы S образуют топологическую группу по умножению G . Она коммутативна, поскольку S является полем. Ввиду свойства 2) в G содержится подгруппа H , изоморфная группе положительных вещественных чисел по умножению. Из свойства 3) следует, что геометрически H является открытым лучом, исходящим из нуля. Поэтому подгруппа H замкнута в G и факторгруппа $F = G/H$ является непрерывной группой. Она коммутативна и при этом

гомеоморфна сфере размерности $N - 1$. Но, как известно, группа с такими свойствами существует лишь при $N = 2$.

Действительно, сфера конечной размерности компактна, связна и локально евклидова (т. е. локально гомеоморфна конечномерному вещественному линейному пространству). Непрерывная коммутативная группа с такими свойствами является прямой суммой конечного числа групп окружности. Этот факт вытекает из следствия теоремы 36 в [2, с. 90]; см. также [3, с. 270, теорема 49]. Его можно рассматривать как топологический аналог теоремы о разложении конечной абелевой группы в прямую сумму конечных циклических групп.

Таким образом, если на сфере задана коммутативная топологическая группа, то сфера одновременно является тором. Но тогда это окружность (на сфере большей размерности любая окружность стягивается в точку, а для тора это неверно). Как следствие, рассматриваемая выше группа F одномерна. Поэтому размерность пространства S равна двум, что и требовалось.

Отметим, что при построении теории непрерывных групп не используется понятие комплексного числа, так что логический круг в изложенном рассуждении невозможен. Вполне естественно с точки зрения геометрической интуиции, что ограничение на возможные расширения поля вещественных чисел возникает из ограничения на топологию их мультипликативных групп. Но последний топологический шаг (сравнение сферы с тором) можно заменить алгебраическим рассуждением, которое выявляет еще одну интересную аналогию.

Предположим, что F является прямой суммой более чем одной группы окружности. Тогда каждая из них содержит два квадратных корня из единицы. Но пересечение этих групп состоит только из единицы, поэтому всего в F таких корней более двух. Пусть r - любой из них, s - его прообраз в G . Тогда s^2 является элементом подгруппы H , и его можно отождествить с положительным вещественным числом. Разделим s на положительный квадратный корень из того же числа. Мы получим квадратный корень из единицы в G , который также является прообразом r . Как следствие, в G более двух квадратных корней из единицы. Но в поле это невозможно.

Приведенное рассуждение (в котором можно было бы использовать корни любой большей степени) напоминает доказательство того факта, что мультипликативная группа конечного поля циклическа [4, с. 144 - 145]. В данном случае место циклическости занимает одномерность, а место конечности поля - его конечномерность над полем вещественных чисел и происходящая отсюда компактность группы F . Таким образом, основная теорема алгебры связана - через посредство теории непрерывных групп - со столь различными вопросами, как топология многообразий и строение конечных полей.

ДОПОЛНЕНИЕ. Непрерывность умножения и взятия обратного элемента в топологии пространства S можно доказать так. Фиксируем в S базис e_1, \dots, e_N . Пусть $a, b \in S$, $x \rightarrow a$. Тогда можно представить $x - a$ в виде $\sum_{j=1}^N \xi_j e_j$, где $\xi_j \rightarrow 0$ ($j = 1, \dots, N$). С учетом свойств 1), 3) и непрерывности операций векторного пространства

$$(x - a)b = \sum_{j=1}^N (\xi_j e_j)b = \sum_{j=1}^N \xi_j (e_j b) \rightarrow 0.$$

Отсюда следует непрерывность умножения.

Пусть теперь $x_i \rightarrow a \neq 0$ при $i \rightarrow \infty$, но x_i^{-1} не стремятся к a^{-1} . Максимум модуля координат x_i^{-1} в фиксированном базисе обозначим M_i . Если M_i ограничены в совокупности, то можно выбрать подпоследовательность, для которой x_i^{-1} стремятся к некоторому пределу $b \neq a^{-1}$. Для нее, с учетом доказанной непрерывности умножения,

$1 = x_i x_i^{-1} \rightarrow ab \neq 1$, что невозможно. Поэтому существует подпоследовательность, для которой $M_i \rightarrow \infty$. У соответствующих элементов $y_i = x_i^{-1} M_i^{-1}$ максимум модуля координат равен 1, и из них можно выбрать подпоследовательность, сходящуюся к некоторому пределу $c \neq 0$. Для нее $0 = \lim_{i \rightarrow \infty} M_i^{-1} = \lim_{i \rightarrow \infty} x_i y_i = ac \neq 0$. Полученное противоречие означает, что в действительности взятие обратного элемента — непрерывная операция.

Автор признателен М. Н. Вялomu за полезное обсуждение.

СПИСОК ЛИТЕРАТУРЫ

- [1] Математическое просвещение. Третья серия, выпуск I. // М.: МЦНМО, 1997.
- [2] Моррис С. Двойственность Понтрягина и строение локально компактных абелевых групп. М.: Мир, 1980.
- [3] Понтрягин Л.С. Непрерывные группы. М.: Наука, 1973.
- [4] Ленг С. Алгебра. М.: Мир, 1968.

Задачный раздел

В этом разделе вниманию читателей предлагается подборка задач разной степени сложности, в основном трудных. Некоторые из этих задач (не обязательно самые сложные!) требуют знания «неэлементарной» математики — анализа, линейной алгебры и т. п.

Составителям этой подборки кажется, что предлагаемые ниже задачи окажутся интересными как для сильных школьников, интересующихся математикой, так и для студентов-математиков.

Помимо *решения* задач, в высшей степени полезно упражняться в *изложении решений*. Мы советуем всем, решившим какую-либо из задач, постараться записать её решение в максимально простом и понятном виде и прислать в редакцию. В последующих номерах мы опубликуем самые изящные из полученных решений.

К сожалению, нам известны авторы далеко не всех предлагаемых ниже задач. Многие из них известны десятилетиями и стали частью «математического фольклора». Одна из целей, преследуемых составителями данного раздела, — записать этот «фольклор», многие части которого стремительно исчезают в наше время.

Мы обращаемся с просьбой ко всем читателям, имеющим свои собственные подборки таких задач, присылать их в редакцию. И, разумеется, мы с удовольствием будем публиковать свежие авторские задачи. Ждем ваших писем.

В скобках после условия задачи приводится фамилия автора (уточнения со стороны читателей приветствуются). Если автор задачи неизвестен, мы указываем того, кто предложил эту задачу.

1. A, B, C — произвольные матрицы размера 2×2 . Докажите тождество Холла: $[[A, B]^2, C] = 0$. (Через $[A, B] \stackrel{\text{def}}{=} AB - BA$ обозначается *коммутатор*). (фольклор)
2. Пусть $|\varepsilon_i| < 1$ и произведение $\prod(1 - \varepsilon_i)$ сходится. Верно ли, что сходится ряд $\sum \varepsilon_i$? (А. Белов)
3. а) Существует ли непрерывная функция $f(x)$, удовлетворяющая функциональному уравнению: $f(f(f(x))) = e^{-x}$.
б) Тот же вопрос для разрывной функции.
в) И для функции, имеющей конечное число точек разрыва. (К. Мальков)
4. а) Пусть $p > 3$ — простое число. Докажите, что на готической шахматной доске размера $p \times p$ можно расставить p ферзей так, чтобы они не били друг друга.
б) Назовем *магараджей* фигуру, которая из клетки $(0, 0)$ за один ход может попасть в клетки $(0, \pm k)$, $(\pm k, 0)$, $(\pm k, \pm k)$, $(\pm k, \pm 2k)$, $(\pm 2k, \pm k)$ (k — целое положительное число). Ответьте на вопрос пункта а) для магарадж и при $p > 7$. (А. Белов)

5. Дан произвольный многочлен с комплексными коэффициентами. Докажите, что корни его производной лежат внутри выпуклой оболочки корней самого многочлена. (фольклор)
6. (Задача на исследование). При каких n и k существует замкнутая n -звенная ломаная, пересекающая каждое свое звено ровно k раз (такую ломаную будем называть ломаной типа (n, k)).
- Если n и k оба нечетны, то это невозможно.
 - Если nk четно, и $n > 3k$, то это возможно.
 - Постройте ломаную типа $(8, 2)$.
 - Существует ли ломаная типа $(6, 2)$.
- Приведите разные другие примеры. (А. К. Ковальджи)
7. Две кривые второго порядка проходят через точки A, B, C, D . Через точку O пересечения прямых AC и BD проведены хорды KM, LN одной кривой и $K'M', L'N'$ другой. Прямые KL и $K'L'$ пересекаются в точке P , MN и $M'N'$ — в точке Q . Доказать, что точки P, Q, O лежат на одной прямой. (А. Заславский)
8. Внутри выпуклого пятиугольника проведены диагонали. Докажите, что они образуют пятиугольник, проективно эквивалентный исходному. (Д. Любшин)
9. Докажите, что на поверхности выпуклого многогранника всегда есть 5 точек, являющихся вершинами правильного пятиугольника. (Из переписки С. Маркелова с исследователем Антарктиды)
10. $a_0 = 1, a_{n+1} = 9^{a_n}, x = \sum \frac{1}{a_i}$. Докажите, что в десятичном разложении числа x встречается любая комбинация цифр. (А. Ерошин и А. Белов)
11. ** а) Плоскость раскрашена в 3 цвета. Докажите, что можно указать такой цвет, что для любого неотрицательного r найдутся две точки плоскости на расстоянии ровно r , окрашенные в этот цвет.
б) Приведите пример раскраски в 6 цветов, когда это не так. (фольклор)

Новые издания

Издания МЦНМО в 1998 году

Московский Центр непрерывного математического образования и Высший Колледж Математики Независимого Московского университета активно издают учебную и научную литературу по математике. Здесь приводится список изданий МЦНМО и ВКМ НМУ, вышедших во второй половине 1998 года и начале 1999 года, а также информация об изданиях, которые должны появиться в ближайшее время.

Совместно с НМУ

- ▷ Математический колледж НМУ — информация для поступающих.
- ▷ *Серия «Студенческие чтения»*
 - Ю. И. Маннин. Рациональные кривые, эллиптические кривые и уравнение Пенлеве. 2-й вып.
 - А. А. Кириллов. Метод орбит и конечные группы. 3-й вып.
- ▷ *Лекционные курсы НМУ*
 - Ю. М. Бурман, Б. Л. Фейгин. Бесконечномерные алгебры Ли.
 - И. М. Парамонова, О. К. Шейнман. Задачи семинара «Алгебры Ли и их приложения».
 - С. М. Натанзон. Геометрия двумерных топологических теорий поля.

ПЕРЕИЗДАНИЯ

- ▷ Э. Б. Винберг. Начала алгебры. (Совместно с УРСС.)
- ▷ Задачи по алгебре / Под ред. Э. Б. Винберга. (Совместно с УРСС.)
- ▷ И. М. Гельфанд. Лекции по линейной алгебре. (Совместно с «Добросвет».)
- ▷ В. А. Зорич. Математический анализ. II часть.

НОВЫЕ ИЗДАНИЯ

- ▷ Э. Картан. Избранные труды.
- ▷ Введение в криптографию / Под ред. В. В. Яценко.
- ▷ Математическое просвещение. Сер. 3, вып. 2.
- ▷ Математический анализ в 57 школе / Под ред. Б. М. Давидовича.

ОЛИМПИАДЫ

- ▷ 4-я Соросовская олимпиада (1997–1998).
- ▷ *В. О. Бугаенко*. Турниры им. Ломоносова. (Переизд., доп.)
- ▷ *И. В. Яценко*. Приглашение на математический праздник.

ДЛЯ АБИТУРИЕНТОВ

В. В. Ткачук. Математика — абитуриенту. (Переизд., доп.)

ДРУГИЕ ИЗДАНИЯ

Летние школьные практики по ботанике. Метод. пособие. —
Под. ред. П. Ю. Жмылева, составитель А. Б. Шипунов.

ГОТОВЯТСЯ К ИЗДАНИЮ

- ▷ *С. П. Новиков*. Уравнение Шредингера и симплектическая геометрия. (Серия «Студенческие чтения».)
- ▷ *А. А. Белавин, А. Г. Кулаков*. Лекции по теоретической физике.
- ▷ *В. В. Шмидт*. Введение в теорию сверхпроводников.
- ▷ Введение в алгоритмы. Перевод под. ред. А. Шеня.
- ▷ *В. А. Васильев*. Лагранжевы и лежандровы характеристические классы.
- ▷ *M. Deza*. Geometry of cuts and metrics. (Перевод.)
- ▷ *Н. Верещагин, А. Шень*. Логика.
- ▷ *С. М. Львовский*. Введение в когомологии пучков.
- ▷ *Д. В. Аносов*. Линейная алгебра (конспект лекций).

ОПЕЧАТКИ, ЗАМЕЧЕННЫЕ В №2

СТРАНИЦА,	СТРОКА	НАПЕЧАТАНО	СЛЕДУЕТ ЧИТАТЬ
137,	3 сверху	$\int_{-1}^1 P_k^d(t) P_l^d(t) (1-t^2)^{\frac{d-3}{2}}$	$\int_{-1}^1 P_k^d(t) P_l^d(t) (1-t^2)^{\frac{d-3}{2}} dt$
181,	2 снизу	спени	степени

Издательство Московского Центра
непрерывного математического образования

Издательство «ЧеРо»

Директор издательского проекта И. В. Ященко
Технический редактор М. Н. Вялый

Лицензия ЛР №071150 от 11.04.95 г.

Подписано в печать .02.99 г. Формат 70 × 100/16

Бумага офсетная №1. Печать офсетная. Печ. л. 15,0

Тираж 1000. Заказ №

МЦНМО

121002, Москва, Большой Власьевский пер., д. 11

ТОО «ЧеРо»

Редакционно-издательский отдел

121002, Москва, Б. Власьевский пер., д. 11, комн. 208

Тел. 241 33 90, 241 18 69

Отдел реализации

118899, Москва, ул. Акад. Хохлова, д. 11

Тел. 939 47 09, 939 34 93