МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ

Третья серия

выпуск 15

Москва
Издательство МЦНМО
2011

УДК 51.009
ББК 22.1
М34

Адрес редакции:
119002, Москва, Б. Власьевский пер., д. 11, к. 301
(с пометкой «Математическое просвещение»)
Email: matpros@mccme.ru   Web-page: www.mccme.ru/free-books

М34 **Математическое просвещение.** Третья серия, вып. 15. —
М.: МЦНМО, 2011. — 248 с.
ISBN 978-5-94057-741-6

В сборниках серии «Математическое просвещение» публикуются материалы о проблемах современной математики, изложенные на доступном для широкой аудитории уровне, заметки по истории математики, обсуждаются проблемы математического образования.

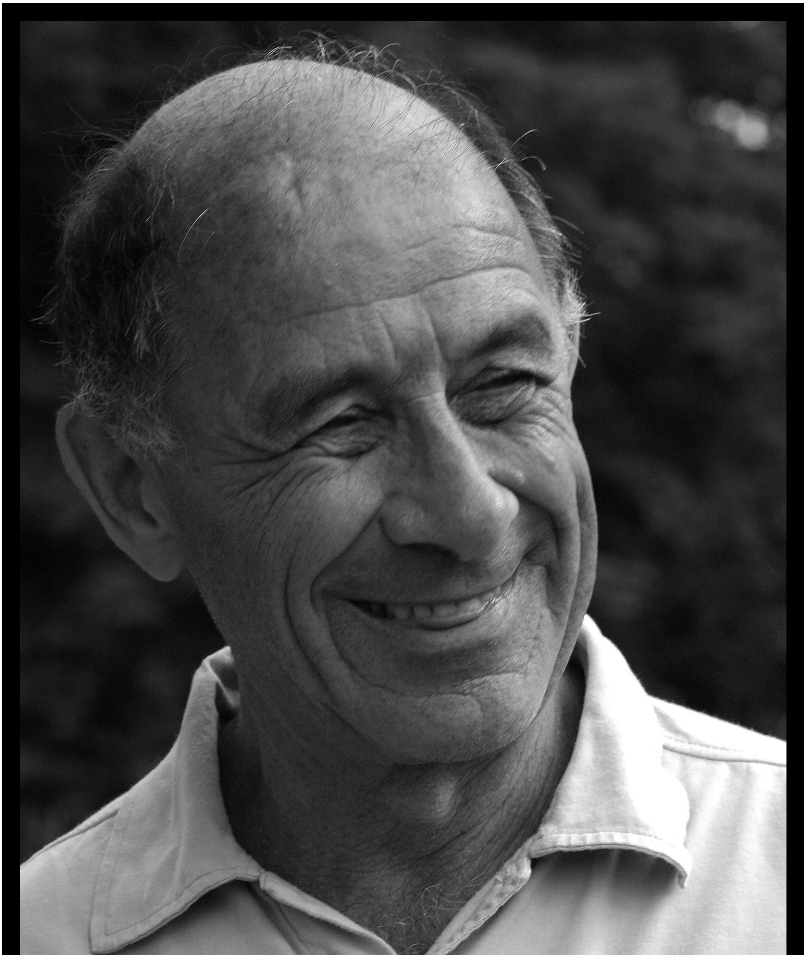
УДК 51.009
ББК 22.1

Фотография на с. 5 сделана С. Третьяковой

ISBN 978-5-94057-741-6

Содержание

Математический мир



В. И. Арнольд (12.06.1937 – 03.06.2010)

Владимир Игоревич Арнольд и математическое просвещение

С. К. Ландо

Иной и учится, но неусердно
и потому живет долго
из письма Геннадия,
архиепископа Новгородского (ок. 1500)

Тонкий яд математического образования ...
Приписывается В. И. Арнольдом Феликсу Клейну

Педагогическое наследие В. И. Арнольда велико и многогранно. Сейчас, спустя лишь несколько месяцев после его ухода, можно позволить себе только самые предварительные заметки — наверняка в его бумагах обнаружатся документы, о которых мы не имеем представления. В любом случае, однако, это наследие не сводится к документам. Впечатление, которое он производил на тех, кто его слушал, магия его личности, влияли на процесс обучения даже сильнее, чем его великие книги.

Владимир Игоревич много думал, говорил и писал о самых разных стадиях обучения математике — от начальной школы до периода вступления в самостоятельные исследования. Сам он продолжал учиться на протяжении всей жизни, учиться страстно. В эпиграф к этой статье вынесены любимые им слова Новгородского архиепископа Геннадия. Усердное же ученичество самого В. И. не позволило ему прожить по-настоящему долгую жизнь.

1. В. И. Арнольд — лектор

При появлении на математическом небосклоне новой звезды математики сразу пытаются выяснить, чьим учеником является этот — чаще всего молодой — человек. Не только область, в которой работает новоиспеченный член сообщества, но и сам его «математический почерк» подсказывают, где искать истоки выдающихся результатов. Владимир Игоревич Арнольд был учеником великого математика Андрея Николаевича

Математическое просвещение, сер. 3, вып. 15, 2011(6–19)

Колмогорова. Многочисленные устные рассказы, записанные воспоминания Арнольда демонстрируют глубину влияния, оказанного Колмогоровым на его жизненный и исследовательский путь, равно как и стремление оторваться от своего учителя, выйти на свою собственную орбиту.

> «*Когда я начинал читать лекции на мехмате, Колмогоров сказал мне*: „*Ни одно слово лекции никакого значения для слушателей не имеет — они всё равно ничего не поймут. Нужно только, чтобы они поняли из курса лекций, какие вопросы будут им заданы на экзамене и как на эти вопросы надо отвечать*“. *Меня поразило здесь то, насколько точно Колмогоров понимал реакцию студентов на его курсы*: *его действительно никто не понимал* (*да и невозможно это было, так как ни одна фраза не была грамматически правильной — то ни одного подлежащего, то сразу три сказуемых, с неразборчивым мычанием вместо дополнения*).» [1)]

Для себя В. И. считал такой подход абсолютно неприемлемым. Каждая его лекция была образцом педагогического мастерства, тщательно продуманной попыткой донести до как можно большего числа слушателей суть излагаемого предмета. Это не означает, что его лекции было легко слушать — он лишь убирал барьеры между студентом и сопротивляющимся познанию материалом, не подменяя этот материал различного рода эрзацами. В результате процент тех, кто стремился понять происходящее на лекциях, был необычайно велик.

Чтобы поддерживать внимание аудитории и контролировать ее способность следить за изложением, В. И. любил вставлять в ключевые места лекции заранее заготовленные ошибки. Обычно за формулировкой неверного утверждения следовала пауза, в течение которой В. И. ожидал, заметит ли кто-нибудь ошибку, и радовался, когда сразу же раздавались вопросы. Если же этого не происходило, то он сам обращался к аудитории с предложением отыскать неточность в формулировке или пробел в доказательстве — тут пробуждались даже самые инертные. В то же время проскакивание ошибки непреднамеренной — неизбежный атрибут любого сколь-нибудь длинного курса — он считал для себя недопустимым и страшно не любил признаваться в чем-либо подобном, предпочитая выдавать такие ошибки за заранее подготовленный трюк. Впрочем, подобные случаи были на удивление редки.

Арнольд также считал себя обязанным не только знать об излагаемом им предмете больше, чем любой из потенциальных слушателей, но и свободнее ориентироваться в нем, находя ответы на возникающие вопросы

[1)] Интервью, взятое В. Б. Демидовичем у В. И. Арнольда в ноябре 2008 г., в сборнике *Мехматяне вспоминают*: 2, М.: Мехмат МГУ, 2009. С. 40.

и решая задачи быстрее, чем многолюдная аудитория. Как правило, ему это удавалось, несмотря на задорную наглость мехматской молодежи, всегда пребывавшей в состоянии готовности подколоть маститого лектора и поймать его на каком-либо промахе.

2. В. И. Арнольд — автор учебников

Любовь к чтению лекций и тщательная подготовка к ним легли в основу умения Арнольда писать учебники. Он объяснял, что нет ничего проще — при подготовке к лекции вы записываете ее будущее содержание, а после лекции — сразу же, пока впечатления не остыли, вносите исправления, дополняете список задач. В результате по окончании чтения курса в вашем распоряжении оказывается готовая книга. Первые варианты книг представляли собой ротапринтированные брошюры, которые студенты получали в библиотеке. Учебник складывался после второго-третьего осуществления лекционного курса, и потом дорабатывался для последующих изданий. Так были созданы знаменитые «Лекции по обыкновенным дифференциальным уравнениям», «Дополнительные главы теории дифференциальных уравнений», «Математические методы классической механики». Иногда подготовленный материал лекционных курсов дополнялся записками лекций, сделанными слушателями (так появился на свет двухтомник «Особенности гладких отображений» — библия теории особенностей, написанный В. И. в сотрудничестве с его учениками А. Н. Варченко и С. М. Гусейн-заде), или эти записки ложились в основу самостоятельных книг (отправной точкой для книги В. Б. Алексеева «Теорема Абеля в примерах и задачах» послужили лекции В. И. Арнольда для школьников колмогоровского интерната).

Нет смысла пытаться определить, какое направление деятельности В. И. оказало большее влияние на развитие мировой математики — его исследовательская деятельность, воздействие его личности или книги. Практически сразу после выхода учебники Арнольда переводились на иностранные языки издательством «Мир» — тогда это было еще большей редкостью, чем немедленное издание книги. Несомненно, именно благодаря книгам, а не статьям, о В. И. слышали в самых отдаленных уголках земли, там, где есть хоть какие-то высшие учебные заведения, преподающие математику.

Не все замыслы Владимиру Игоревичу удавалось реализовать. Несмотря на свои настойчивые просьбы, он не получил возможности прочитать на мехмате курс лекций по уравнениям в частных производных. Впоследствии он прочитал этот курс студентам Независимого Московского университета. Записки лекций были изданы, однако в полноценную книгу они, к сожалению, не превратились, и оригинальный подход Арнольда к теории уравнений в частных производных не получил адекватного изложения. Работать над книгами В. И. продолжал до последних дней — его «Геометрия», в полной мере отразившая его взгляд на математику в целом, еще не сдана в издательство.

При написании книг Арнольд считал необходимым приводить мотивировки даваемых определений. В его книгах введение всякого нового понятия предваряется примерами, в которых это понятие естественно возникает, после чего новые примеры дополняют базу для работы с ним. Ставшая классической «кошка Арнольда» исходно была предназначена лишь для иллюстрации понятия гиперболического преобразования тора.

Когда он вводил в своих учебниках понятие группы, то это всегда была группа преобразований — т. е. набор взаимно однозначных отображений множества в себя, содержащий тождественное преобразование, композицию любых двух своих элементов и обратное преобразование для любого своего элемента. При таком подходе из абстрактного определения группы выпадает требование ассоциативности, которое для человека, впервые знакомящегося с этим понятием, кажется труднее всего поддающимся проверке. Быть может, еще более важно то, что всякая группа оказывается подмножеством вполне понятного объекта — множества всех взаимно однозначных отображений некоторого множества на себя, т. е. множества перестановок.

Многообразие для него было классом диффеоморфности подмногообразий в евклидовых пространствах — т. е., подмножеств, допускающих локальное задание набором независимых гладких уравнений. Теорема Уитни о вложении гарантирует, что при таком подходе мы ничего не упустим. На первый взгляд, это определение вызывает трудности. Во-первых, многообразия можно задавать и по-другому, и далеко не всегда очевидно, каким образом многообразие можно вложить в евклидово пространство. Во-вторых, проверка диффеоморфности двух подмногообразий может представлять значительные сложности. Эти сложности, однако, в значительной степени компенсируются наглядностью определения — при слове «многообразие» возникает зрительный образ поверхности в пространстве. К тому же другие определения — вроде хаусдорфова топологического пространства со счетной базой, снабженного атласом с гладкими функциями перехода, — выглядят гораздо хуже. Они пригодны лишь для доказательства того, что какой-то объект является многообразием и ничуть не способствует пониманию.

И в случае группы, и в случае многообразия любимые им определения были ровно теми, которые давались самими изобретателями понятий. Владимир Игоревич считал, что последующее уточнение и формализация исходных подходов практически всегда приводят к их ухудшению, к затемнению сути. Многие часы провел он за изучением работ Ньютона, Пуанкаре, Якоби, восстанавливая забытые открытия, и побуждал к тому же своих учеников. Распространенные сейчас попытки упрекнуть классиков в нестрогости встречали его яростный отпор.

3. Реформа математического образования

В начале 60-х годов прошлого века Андрей Николаевич Колмогоров предпринял значительные усилия для реформирования школьного математического образования. Его попытки втянуть в это предприятие своего лучшего и самого близкого ученика натолкнулись на упорное сопротивление последнего. И тогда, и позднее, Арнольду нравились использовавшиеся в советской школе стандартные учебники математики — в первую очередь, «Геометрия» Киселёва. Впоследствии он не раз предлагал к ним вернуться. Ниже речь идет только об обучении математике, хотя В. И. высказывался и по гораздо более широкому кругу проблем российской школы.

Направления предлагаемых Колмогоровым изменений также не вызывали у В. И. энтузиазма. Стремление осовременить школьный курс, сводившееся, во многом, к выстраиванию самосогласованной терминологии и подчеркиванию важности логически строгих выводов, не находило у него понимания. Арнольд любил говорить, что у аксиоматического метода много преимуществ по сравнению с традиционным подходом — подобных преимуществам воровства перед честным трудом.

И учитель, и ученик очень болезненно переживали такое взаимонепонимание.

Напротив, деятельность А. Н. по организации в Москве физико-математического интерната для нестоличных школьников (ныне — СУНЦ им. А. Н. Колмогорова, и москвичей тоже принимают туда учиться) Арнольд всячески поддерживал. Эта поддержка не ограничивалась случайными встречами со школьниками — как уже упоминалось выше, В.И. подготовил и прочитал в интернате курс теории Галуа в задачах.

Недавно Арнольд писал по этому поводу[2)]:

> «*Во всяком случае, даже если школьникам в интернате и бывало порой трудно, польза от интерната была и остаётся огромной, неизмеримо, на мой взгляд, большей, чем от попыток Колмогорова модернизировать курсы математических наук с заменой классических учебников А. Киселёва новыми учебниками бурбакистского толка (с их современной терминологией, заменившей классические евклидовы „признаки равенства*

[2)] В. И. Арнольд. *Новый обскурантизм и российское просвещение*. М.: Фазис, 2003.

> *треугольников" малопонятными, хотя и логически предпочтительными, „признаками конгруэнтности").*
> *Это реформирование подорвало авторитет и школы, и учителей, и учебников, создав наукообразную иллюзию псевдознания, прикрывающую полное непонимание простейших фактов, вроде того, что* $5 + 8 = 13$. *В проекте новой реформы заметны такие же тенденции одурачивания школьников, которым предлагается непонятная „геометрия Лобачевского" взамен исключаемых из обучения записи простых дробей десятичными и „текстовых арифметических задач" об экипажах, следующих из пункта A в пункт B, или о купцах, продающих сукно за топоры, или о землекопах и трубах, наполняющих водоёмы, — задач, на которых выучились думать предыдущие поколения.»*

Упомянутая в процитированном тексте «новая реформа» началась около 15 лет назад и продолжается по сей день. Среди вызвавших ее причин — видимое невооруженным глазом несоответствие целей, стоящих перед школьным математическим образованием, и результатами этого образования. За время реформы ее направление и цели неоднократно пересматривались, но глобальная тенденция, пожалуй, не изменилась. Состоит она в том, чтобы подогнать процесс обучения под западный образец. Здесь стоит заметить, что единого такого образца просто не существует (подходы к обучению во Франции, Финляндии и Южной Корее мало похожи друг на друга и все они не имеют ничего общего с обучением в США, которое, видимо, и выступает основной целевой моделью). Арнольд — и среди ведущих математиков он далеко не одинок — с его многообразным опытом преподавания в лучших западных университетах, пришел к недвусмысленному выводу о преимуществах российской системы школьного математического образования перед всеми западными подходами. Востребованность плодов этой системы на западе после 1990 г. — наилучшее тому подтверждение.

Реформы необходимы, альтернатива им — застой. Но заниматься реформированием должны высокопрофессиональные и талантливые люди, хорошо знакомые с предметом обучения, с историей развития системы, с текущим положением дел, с существующими проблемами и представляющие себе весь спектр возможных решений. Слепое следование нелучшим образцам может привести лишь к созданию их ухудшенных копий. Арнольд не чувствовал себя специалистом в организации школьного обучения и не брал на себя смелость выстраивать школьные курсы. Он полагал, что квалифицированные учителя лучше справятся с этой задачей. Однако он считал себя вправе — и у него для этого были все основания — оценивать предлагаемые изменения и избираемые пути, он справедливо полагал себя экспертом высочайшей квалификации. А голоса экспертов экстра-класса не должны тонуть в мельтешении дилетантов.

4. Семинар Арнольда и его школа

Наибольшее влияние Владимир Игоревич Арнольд оказал, несомненно, на постоянных участников своего семинара. Этому семинару около 50 лет и именно его организация стала высшим педагогическим достижением Арнольда. Из семинара вышли и стали лидерами в своих областях такие замечательные и такие разные математики как Александр Варченко, Виктор Васильев, Сабир Гусейн-заде, Александр Гивенталь, Максим Казарян, Аскольд Хованский, и многие другие.

Помимо нескольких десятков постоянных участников, появлявшихся на семинаре еженедельно в течение десятилетий, через него прошли несколько сотен других людей — студенты в поисках научного руководителя, математики, зашедшие послушать интересный доклад, докладчики, приглашенные самим Арнольдом (среди докладчиков нередки были физики). Почувствовав искренний интерес к своим работам, некоторые случайные докладчики становились затем постоянными участниками семинара.

Впервые появившись на семинаре в 1979 году, я обнаружил, что не понимаю в докладах ничего. Несмотря на декларируемое им требование доступности докладов, добиться выполнения этого требования было для В. И. не так просто. Причина этого понятна — в разгар разработки Арнольдом и его школой теории особенностей большинство докладов было посвящено именно ей, а постоянные участники семинара были погружены в нее гораздо глубже, чем зеленые новички. Необходимость же движения вперед требовала и отнесенной вперед отправной точки каждого доклада, иначе одна и та же порция сведений повторялась бы каждый раз. Некоторые проблески понимания возникали у меня, когда за объяснение брался сам В. И. Лишь несколько лет спустя, когда в моем сознании начала складываться общая картина происходящего, я пришел к выводу, что основной причиной моего непонимания было банальное незнание вещей, которым следовало бы обучать каждого студента-математика, — но либо нас им не обучали, либо я их своевременно не усвоил.

Для многих участников семинар начинался значительно раньше официального звонка. Обычным местом предварительного сбора была курилка — раньше на мехматской лестнице дозволялось курить — и приходили туда в том числе и некурящие. Лишь единицы семинаристов работали или учились на мехмате, а для всех остальных это была редкая возможность встретиться и обсудить математические вопросы. После же семинара — если в этот вторник не было заседания Московского математического общества, сначала вице-президентом, а потом президентом которого был Арнольд, — к Владимиру Игоревичу выстраивалась очередь желающих что-то спросить, показать свой или чужой результат, обсудить планы. Эти разговоры могли продолжаться до позднего вечера, а процесс стояния в очереди сам по себе превращался в математическую дискуссию.

Благодаря международной известности В. И., многие исследователи со всего мира почитали за честь послать ему свою статью. Проходные работы, естественно, не посылались — только те, которые представляли ценность с точки зрения самих авторов. Признанные мэтры присылали достойные внимания работы своих учеников. Кроме того Арнольд получал и свежие выпуски журналов, в редколлегии которых он входил. В условиях крайне ограниченного доступа к свежей зарубежной периодике (справедливости ради надо сказать, что по обеспеченности журналами библиотека мехмата в 1960–80е годы была одной из лучших в СССР) участники семинара тем не менее могли знакомиться с только что вышедшими работами. В.И. лично просматривал все источники, отбирая из них то, что, по его мнению, могло представлять ценность, и распределял их между участниками семинара, предлагая разобрать статью и подготовить ее изложение. Случаи, когда, разобрав текст, потенциальный докладчик приходил к выводу о нецелесообразности доклада, были чрезвычайно редки — Арнольд редко ошибался в оценке значимости текста.

Нельзя сказать, чтобы Арнольд специально чему-то учил участников семинара. Скорее показывал, как он думает, предоставляя возможность учиться на собственном примере. И эта повторявшаяся раз за разом в течение многих лет демонстрация и служила стратегическим методическим приемом, позволяя слушателям кропотливо, по крупицам, осваивать — не знания, а подходы к их получению, обретать умение задавать природе правильные вопросы.

5. Математический тривиум и Задачи Арнольда

В отличие от большинства научных сообществ математикам присуща высокая степень согласия относительно того, что означает понимание. Математики едины в том, что понимание предмета требует, не в последнюю очередь, умения решать задачи, его касающиеся. Верхнюю ступень в иерархии задач занимают тем или иным образом оформленные «великие проблемы» — вроде доказанной недавно последней теоремы Ферма, решенной проблемы четырех красок, почти исчерпанного списка проблем Гильберта или наследующего ему спустя век едва початого списка задач тысячелетия. Несмотря на то, что наука прокладывает свои пути, почти не замечая этих заброшенных в будущее маяков, они не перестают привлекать к себе внимание и новообращенных юнцов, и опытных мастеров, и, в особенности, досужих внешних наблюдателей.

Однако, если польза попыток определить с помощью задач развитие науки на годы вперед и оказывается зачастую сомнительной, то на ближайшую перспективу и, в первую очередь, на процесс обучения, правильно подобранные задачи способны повлиять весьма существенно. Арнольд начал вести свой семинар будучи еще аспирантом, в самом конце 50-х годов, и с первого года работы семинара формулировал для него нерешенные задачи, которые представлялись ему важными. Проявив редкостную (и присущую ему в полной мере) организованность, он сохранил списки всех этих задач вплоть до 2010 года — последнего года своей жизни. В книге «Задачи Арнольда» (Москва, Фазис, 1997) и ее расширенном переводе «Arnold's problems» (Springer, 2003) опубликованы все доступные на момент публикации задачи с комментариями. Комментарии написаны самим В. И. и его учениками и отражают современное состояние дел в области, к которой задача относится, ссылки на посвященные ей работы. Многие из задач Арнольда не решены и по сей день — чем не вызов для молодого математика.

На основании многолетних наблюдений Арнольд определил период полураспада своих задач в 7 лет — в среднем именно за этот срок половина поставленных задач оказывалась решенной. Впрочем, он любил говорить, что ученики никогда не решают поставленную задачу — они всегда заменяют ее той, которую могут решить. Решения некоторых из этих задач стали отправными точками новых теорий. Достаточно упомянуть симплектическую топологию, выросшую из вопроса Арнольда про число неподвижных точек симплектоморфизма. Эти теории он вероятно провидел. Однако совершенно сознательно не считал нужным формулировать гипотетические теоремы будущей теории, предпочитая ограничиваться первым содержательным вопросом, ответ на который неизвестен. Он полагал, что после получения ответа на такой вопрос теория произрастет сама без дополнительных усилий — как на самом деле и происходило. Свой подход он противопоставлял стремлению сформулировать задачу в максимально общем виде, присущему, по его мнению, французским математикам. И здесь его точка зрения близка к взглядам И. М. Гельфанда, полагавшего, что теории приходят и уходят, а примеры остаются.

Мне представляется, что одно из главных достоинств книги «Задачи Арнольда» — возможность проследить за тем, как менялись задачи. Нередко на протяжении многих лет В. И. возвращался к одной и той же проблеме, изменяя ее формулировку в соответствии с эволюцией своих взглядов. Эта драгоценная возможность предметно — и на очень сжатом материале — увидеть прорастание смыслов, пожалуй, не имеет аналогов в математической литературе. Как правило, при чтении исследовательских статей нам приходится иметь дело с задачами в окончательной формулировке, а весь процесс развития остается за кадром.

Задачами он был пропитан. Когда спустя полтора года после окончания мехмата я, набравшись храбрости, впервые подошел к нему и попросил задачу, он, отослав меня было к списку задач для семинара (о существовании которого я не имел ни малейшего представления), тут же написал мне 4–5 штук. Через несколько месяцев я принес ему решение одной из них, что и определило его согласие взять меня в ученики.

Сформулированные В.И. задачи очищены от шелухи. За редким исключением понимание их формулировок не требует большого объема предварительных сведений, и попытать свои силы могли и младшекурсники. Важные задачи с такими формулировками редки и ценны, В.И. же в период своего расцвета порождал их дюжинами, чему нельзя не восхищаться. При этом он не считал для себя возможным навязывать ученику — даже своему официальному аспиранту — ту или иную задачу. По его мнению, это было все равно, что навязывать невесту, и при выборе для себя задачи мы пользовались большой свободой.

Среди сформулированных им задач практически не встречаются такие, на которые возможен ответ «да» или «нет». Он считал, что задача с бинарным ответом не может быть интересной. Интересными ему представлялись задачи, открывающие новую область исследований, устанавливающие связи между разными областями, задачи подсчета конкретных чисел геометрического происхождения, вычисления, имеющие физический смысл. Закрытие направлений его не волновало. Он публично высказывал свое неприятие и таких общепризнанно великих проблем, как упомянутая выше последняя теорема Ферма[3)], что нередко вызывало раздражение многих, в том числе ведущих ученых. «*Как он может такое говорить*» — узнав, что я ученик Арнольда, выражал мне свое возмущение году в 96м Серж Ленг, — «*его же СЛУШАЮТ*». Что-что, а свое мнение В.И. действительно охотно высказывал публично и недвусмысленно, и его действительно слушали. Слышали — далеко не всегда.

Задачи служили для него и естественным критерием качества собственно математического обучения.

«*Чем определяется уровень подготовки математика? Ни перечень курсов, ни их программы уровень не определяют. Единственный способ зафиксировать, чему мы действительно научили своих студентов — это*

[3)] «*Ради рекламы современные математики иногда выдают подобные спортивные достижения за последнее слово своей науки. Понятно, что это не только не способствует высокой оценке математики обществом, а, напротив, вызывает здоровое недоверие к необходимости затраты усилий на занятия (типа скалолазания) этими экзотическими и неизвестно зачем и кому нужными вопросами*» В. И. Арнольд. О преподавании математики // УМН, т. 53, вып. 1, 1998. С. 229–234.

перечислить задачи, которые они должны уметь решать в результате обучения. Речь идет здесь не о каких-либо трудных задачах, а о тех простых вопросах, которые составляют строго необходимый минимум. Этих задач не обязательно должно быть много, но уметь решать их нужно требовать жестко.»

Эталонный список из ста задач, которые должен уметь решать выпускник физического факультета, — «Математический тривиум» — он разработал и опубликовал в 1991 г.[4]. Отсылка к физикам была, подозреваю, способом защиты от неизбежных последующих нападок — круг профессоров-математиков, не только владеющих терминологией, необходимой для понимания условий всех задач, но и способных решить, не прикладывая к этому сверхусилий, хотя бы половину из них, крайне ограничен. И это при том, что за пределы университетской программы ни одна из задач не выходит. Знакомство с этим списком, видимо, самый эффективный способ составить представление о взглядах В. И. Арнольда на университетское математическое образование.

6. Независимый Московский университет и летняя школа «Современная математика»

К концу 80-х годов в Москве оформился круг математиков, по большей части молодых, не имевших возможности преподавания сильным студентам, но стремившихся такую возможность получить. За плечами многих из них был опыт преподавания в специализированных школах, входивших в «систему Константинова», включавшую несколько московских школ, объединенных общим подходом к обучению старшеклассников, выработанным в начале 60-х годов Николаем Николаевичем Константиновым, одним из первых кружковцев которого был и Дима Арнольд. Эта система сформировалась благодаря неуемной энергии Константинова и способности посвящать все свое время общению со школьниками, их обучению, организации олимпиад и конкурсов. Набранный опыт и выработанные при общении со своими учителями представления о том, как должно быть организовано университетское образование в области естественных наук и математики, требовали выхода. Неудивительно поэтому, что при первой же возможности, открывшейся при смене общественного строя в России, Николай Николаевич инициировал создание нового университета, ориентированного на подготовку исследователей, и эта его инициатива получила энергичную поддержку.

[4] В. И. Арнольд *Математический тривиум* // УМН, т. 46, вып. 1, 1991. С. 225–232, *Математический тривиум–II* // УМН, т. 48, вып. 1, 1993. С. 211–222.

Поддержка исходила с двух сторон. Молодые математики — среди которых были и исследователи с уже заработанным именем — с энтузиазмом взялись за преподавание выпускникам специализированных классов, поставившим себе цель подготовиться к исследовательской деятельности. Ведущие ученые старшего поколения выступили с формальной инициативой о создании университета, и принимали активное участие в его работе, особенно в годы его становления. Массовые отъезды ученых за границу ослабили образовательную базу НМУ, но не разрушили ее полностью.

Владимир Игоревич Арнольд был первым из тех, к кому Н. Н. Константинов пришел со своей инициативой. Он согласился возглавить Научный Совет Математического Колледжа, поставив условием, что именно Константинов будет его главным организатором и именно он будет нести ответственность за реализацию выработанных подходов к обучению. О поддержке Константинова и Арнольда заявили академики (тогда еще АН СССР) С. П. Новиков, Я. Г. Синай, Л. Д. Фаддеев. Тесно связанные с российской математикой Роберт Макферсон и Пьер Делинь также сыграли ключевую роль в создании НМУ. А. А. Кириллов, В. М. Тихомиров были первыми лекторами, а А. Н. Рудаков стал еще и первым деканом Математического колледжа.

Следует упомянуть, что В. И. испытывал инстинктивное отвращение к занятию руководящих постов. Он тщательно оберегал свою свободу и никогда не был даже заведующим лабораторией. Лишь изредка он брался возглавлять общественные дела, важность которых представлялась ему неоспоримой. Среди таких дел — президентство в Московском математическом обществе, руководство редколлегией журнала «Функциональный анализ и его приложения», руководство Международным математическим союзом (он был вице-президентом), руководство Попечительским Советом Московского центра непрерывного математического образования. Научное руководство Математическим колледжем НМУ относится к их числу.

О создании Независимого Московского университета было объявлено в конце 1990 года, и в сентябре 1991 года начались занятия в двух колледжах — Математическом и Математической физики. Химический колледж существовал несколько в стороне и вскоре его связи с Независимым университетом прервались.

История Независимого Московского университета — отдельная тема, подробное обсуждение которой увело бы нас слишком далеко от предмета разговора. Скажу лишь, что в начале 90-х годов новые университеты рождались как грибы (видимо, подобная идея носилась в воздухе), но до сегодняшнего дня из них дожили считанные единицы. Крепкая человеческая база, на которой создавался НМУ даже при отсутствии какой бы то ни было материальной, придала ему достаточно жизнестойкости, и он продолжает служить целям, для достижения которых был создан, год от года привлекая в свои ряды способную молодежь.

Для Арнольда научное руководство Математическим колледжем НМУ было возможностью реализации выстраданных им взглядов на построение математического образования. Именно он (при активном участии А. Н. Рудакова, Ю. С. Ильяшенко и других) составил базовый план обучения, и именно в соответствии с этим базовым планом строится преподавание в НМУ и сейчас. Принципы построения этого базового плана сохраняются и на факультете математики Высшей школы экономики, созданном ею недавно в сотрудничестве с Независимым университетом. Так получилось — и в этом нет ничего удивительного, — что в первоначальном составе преподавателей Независимого университета лучше всего были представлены школы Арнольда, Новикова, Манина, Гельфанда. В. И. вообще иногда говорил, что эти школы не существуют как отдельные единицы, а есть единая мощная Московская математическая школа. Возможно, именно это единство послужило залогом того, что образовательные планы Арнольда получили адекватное воплощение.

С приглашением В. И. на постоянную позицию в парижском университете Дофин (это произошло в 1991 году) его московский семинар потерял былой блеск. Многие из его ранее активных участников разбрелись по свету в поисках средств, необходимых для поддержания жизни своей семьи. Оговорюсь, что в абсолютном большинстве случаев эти поиски были весьма успешными, и ученики Арнольда сейчас занимают достойные позиции в ведущих университетах и исследовательских центрах многих стран мира. Владимир Игоревич регулярно проводил в Москве весенний семестр, а осенью руководил работой семинара издалека — парижское ответвление семинара было ориентировано на его вновь набранных французских учеников. Преподавание французским студентам явно не удовлетворяло потребность В.И. в общении с молодежью, а его длительные отлучки из Москвы препятствовали появлению у него новых учеников в России.

Отдушиной для него стала летняя школа «Современная математика» в санатории «Ратмино» под Дубной. С 2000 года она проводится ежегодно усилиями Московского центра непрерывного математического образования (директор и инициатор Школы — Иван Валерьевич Ященко, в последние годы основную тяжесть организации несет на себе заместитель директора МЦНМО, племянник В. И. Виталий Дмитриевич Арнольд). Для участия в Школе отбираются со всей страны около 120 старшеклассников и студентов младших курсов, склонных к занятиям математикой.

В рамках Школы Арнольд и молодежь нашли друг друга, и я не берусь определить, для кого из них возможность общения оказалась важнее.

Владимир Игоревич не читал в Школе длинных лекционных курсов — обычно это были 2-3 лекции, в которых он излагал актуальное состояние собственных исследований. Полезно бросить взгляд на темы его выступлений перед школьниками:

- ▷ 2001 г. Астроидальная геометрия и топология
- ▷ 2002 г. Арифметика совершенных квадратичных форм, симметрия их непрерывных дробей и геометрия их мира де Ситтера
- ▷ 2003 г. Топология алгебры и гидродинамика арифметики
- ▷ 2005 г. Статистика топологии и алгебры
- ▷ 2006 г. Тригонометрические многочлены Морса и шестнадцатая проблема Гильберта
- ▷ 2007 г. На сколько частей n прямых делят плоскость?
- ▷ 2007 г. Квадратичные иррациональные числа, их цепные дроби и их палиндромы
- ▷ 2008 г. Цепные дроби квадратных корней из целых чисел
- ▷ 2009 г. Измерение объективной степени случайности конечного набора точек
- ▷ 2009 г. Об истории обобщенных функций

В те же самые годы он писал и публиковал научные статьи, посвященные этим темам, и одной из целей выступлений перед школьниками было совершенствование этих статей.

Неизбежные вопросы после лекций нередко перерастали в длительные индивидуальные беседы, и многие навсегда запомнили Владимира Игоревича, уходящего по тропинке в сторону Волги, объясняя идущему рядом школьнику устройство раскинувшейся над рекой радуги.

С. К. Ландо, Независимый Московский университет; Высшая школа экономики, факультет математики

Математическая жизнь в Казани в годы войны

М. М. Арсланов

В канун Великой Отечественной Войны Казань являлась одним из крупных математических центров России. Будучи одним из старейших университетских центров, где жил и трудился Н. И. Лобачевский, Казань всегда играла значительную роль в математической жизни России. Предвоенный период математической жизни Казани связан с именами А. В. Васильева и Н. Н. Парфентьева, воспитавших в стенах Казанского университета ряд первоклассных математиков. Период наибольшего расцвета математических исследований в Казани приходится на 30–40-е годы, что связано с переездом в Казань Н. Г. Чеботарева.

Николай Григорьевич Чеботарев основал в Казанском университете алгебраическую школу, получившую мировую известность. Прежде всего крупные результаты из самых различных разделов классической математики были получены самим Чеботаревым. Здесь нет возможности их перечислить, отметим лишь, что ему принадлежит наиболее глубокое обобщение теоремы Дирихле о простых числах в арифметической прогрессии. Исследуя проблему преобразования алгебраического уравнения к уравнению с наименьшим числом независимых параметров, известную как «проблема резольвент», Николай Григорьевич получил основополагающие результаты, за которые ему посмертно была присуждена Сталинская премия 1-й степени (1948).

Ученик Н. Г. Чеботарева И. Д. Адо в своей кандидатской диссертации решил трудную проблему точного конечномерного представления конечномерных алгебр Ли над полем характеристики нуль, за этот результат ему была присуждена сразу степень доктора наук.

Работы В. В. Морозова по теории групп Ли вместе с результатами московского математика Е. Б. Дынкина позволили получить полное решение проблемы классификации всех примитивных представлений групп Ли, поставленной выдающимся норвежским математиком Софусом Ли еще в девятнадцатом веке. Знаменитая теорема регулярности Морозова, доказанная в процессе работы над этой проблемой и составившая основу его метода исследования, долгие годы оставалась одним из наиболее значительных достижений теории групп Ли.

В области теории функций и математического анализа были весьма заметны работы профессора Б. М. Гагаева и его ученика Ф. Д. Гахова.

Под влиянием работ академика Н. Н. Лузина Б. М. Гагаев занимался вопросами теории функций Бэра и ортогональными функциями. Им были найдены необходимые и достаточные условия принадлежности предела функций из конкретного класса Бэра к тому же классу. К наиболее значительным работам Б. М. относятся также его работы по проблеме, оставленной открытой Н. Н. Лузиным в его докторской диссертации «Интеграл и тригонометрический ряд»: существуют ли, кроме системы $\{1, \sin x, \cos x, \sin 2x, \cos 2x, \ldots\}$, другие ортогональные системы функций, инвариантные (с точностью до числовых множителей) относительно дифференцирования? Борису Михайловичу удалось получить полное решение этой проблемы, установив, что кроме вышеприведенной системы тригонометрических функций этим свойством могут обладать только системы из конечного числа функций.

Работы в области неевклидовой геометрии и тензорному анализу основателя казанской геометрической школы П. А. Широкова являются одними из самых ярких достижений казанских математиков предвоенных лет.

В области теоретической механики значительные результаты были получены профессором Н. Г. Четаевым, впоследствии членом-корреспондентом АН СССР и директором Института механики АН СССР. Им была создана научная школа, получившая широкую известность как «казанская школа теории устойчивости».

Великая Отечественная война 1941–1945 годов стала для нашей страны тяжелейшим испытанием, наложившим отпечаток на все сферы ее жизнедеятельности. Развитие научных исследований в Казани в годы войны тесно связано с деятельностью эвакуированных в республику научно-исследовательских учреждений АН СССР, в том числе института математики имени В. А. Стеклова («Стекловка», как любовно называют его математики) и института механики. В Казань прибыли 33 из 85 научных учреждений АН СССР, в том числе 15 московских и ленинградских НИИ. В Казань были эвакуированы также многие проектно-конструкторские организации (в том числе знаменитая КБ тюремного типа, так называемая «шарашка», где работал С. П. Королев), ряд университетов и технических вузов, большое число иных научно-исследовательских организаций и предприятий страны. Переехал также президиум АН СССР во главе с вице-президентами О. Ю. Шмидтом, Л. А. Орбели и Е. А. Чудаковым.[1)]

[1)] Президиум Академии наук СССР находился в Казани до 1942 года, затем он был переведен в Свердловск. Следует отметить, что Казань — не единственный город в Татарстане, приютивший в годы Великой Отечественной Войны крупные научные силы страны. Коллективы Ленинградского и Воронежского университетов были эвакуированы в небольшой город Елабуга, расположенный в 200 километрах от Казани, и размещены в здании местного учительского института. Среди большого числа

В Казань были эвакуированы семьи 1884 научных сотрудников (около 5 тысяч человек), в том числе семьи 39 академиков и 44 членов-корреспондентов Академии наук СССР (есть источники, где приводятся другие сведения — всего 93 члена и члена-корреспондента АН СССР). Всего в республике было размещено 226 тысяч эвакуированных, люди разных возрастов и профессий. Среди эвакуированных в Казань ученых были вице-президенты АН СССР О. Ю. Шмидт, Л. А. Орбели, Е. А. Чудаков, академики С. И. Вавилов, И. М. Виноградов, А. Ф. Иоффе, П. Л. Капица, А. Н. Колмогоров, Г. М. Кржижановский, А. Н. Крылов, С. Л. Соболев, члены-корреспонденты А. Д. Александров, А. П. Александров, П. С. Александров, И. М. Гельфанд, А. О. Гельфонд, Б. Н. Делоне, Л. Д. Ландау, Л. С. Понтрягин, И. Е. Тамм, а также Л. В. Келдыш, Ю. В. Линник, Е. М. Лифшиц, Л. А. Люстерник, А. А. Ляпунов, А. И. Мальцев, С. М. Никольский, П. С. Новиков, Д. К. Фаддеев и многие другие. В 1943 году из Ашхабада, куда был сначала эвакуирован МГУ, в Казань переехал И. Р. Шафаревич, тогда молодой кандидат наук, и стал докторантом Стекловки. (Как вспоминает Игорь Ростиславович [4. с. 10], в Казань его телеграммой вызвал Б. Н. Делоне: «Присылайте Шафаревича, я готов взять его членом семьи».) Докторантами Стекловки были также А. И. Мальцев и С. М. Никольский.

Естественно, что вызванные этой эвакуацией перемены коснулись всех сторон жизни университета и сказались на организации учебного процесса, быта и досуга, частной жизни студентов и преподавателей.

Прежде всего перед коллективом университета была поставлена задача уплотниться и разместить на своих площадях несколько ведущих институтов, сохранив при этом относительно нормальные условия для учебы студентов. Размещением эвакуированных руководил вице-президент АН СССР О. Ю. Шмидт. Он скупил все имевшиеся в магазинах и на складах Казани подушки, кровати и матрасы. Их расставили в актовом, предактовом, спортивном и читальном залах университета, а также в помещении бывшей университетской церкви. В считанные дни университет превратился в огромную коммуналку. Студенты того времени Г. Вульфсон и Н. Муньков вспоминают: «Из студенческого общежития сюда перевезли кровати и тумбочки, остальное доделали сами ученые. Непонятно, как скреплялись одеяла, простыни, как появились импровизированные комнаты-шалаши, у которых был единый пол из старинного паркета и единый лепной потолок, сооруженный по проекту М. Пятницкого. В этих шалашах-комнатах потекла необычная жизнь многих ученых, сотрудников Академии и их семей. На одной тумбочке стояла настольная лампа

ученых, приехавших тогда в Елабугу, были академики Владимир Иванович Смирнов (математик), Виктор Амазаспович Амбарцумян (астроном), Владимир Александрович Фок (физик) и другие выдающиеся ученые.

с импровизированной чернильницей, на другой — таз с замоченным бельем, на третьей шумел и пыхтел примус. Когда вдруг гас свет, появлялись свечи или керосиновые лампы, и в ульях актового зала продолжалась жизнь» ([3, с. 50]). Дневниковые записи сотрудника историко-филологического факультета В. А. Климентовского дополняют общую картину тех дней: «За несколько дней университет резко изменил свой всегдашний, строго официальный внешний вид. По коридорам снует много людей: кто с чайником за водой, кто с полотенцем для умывания и т. д. Многочисленная детвора, по-видимому, даже довольна таким оживлением и таким обширным помещением. Ребятишки бегают вдоль длинных коридоров, и один даже выехал на своем трехколесном автомобиле» ([4, с. 368]).

Позднее многих из них расселили (часто насильственно) по квартирам жителей Казани. Больше всего повезло «соединенному семейству» А. Н. Колмогорова и П. С. Александрова (в Казань выехали мать и сестра Павла Сергеевича и тетя Андрея Николаевича Вера Яковлевна, а позднее и другая его тетушка, Варвара Яковлевна). Они нашли приют в роскошной квартире известного казанского аптекаря А. А. Вильде, расположенной на улице Академическая (ныне улица Вишневского) недалеко от Арского поля, где им выделили две большие комнаты. Позднее (летом 1942 года) Колмогоров и Александров предложили А. И. Мальцеву перейти к ним, и Мальцев до конца пребывания в Казани жил у них, «занимая вполне пристойный угол за шкафом» (воспоминания С. М. Никольского [7, с. 227]). До этого А. И. Мальцев и С. М. Никольский сначала занимали маленькую чердачную комнату главного здания университета, потом их переселили в спортивный зал. Оба они приехали в Казань без семей.

Колмогоров вскоре возвращается в Москву к своим обязанностям академика-секретаря Физико-математического отделения Академии и для выполнения работ оборонного характера. В Казань выбирается только временами.

П. А. Широков приютил у себя семью Б. Н. Делоне. Позднее, после переезда из Ашхабада в Казань, у них поселился также И. Р. Шафаревич. Л. С. Понтрягин с матерью и женой поселились в квартире В. В. Морозова, позднее к ним присоединилась семья А. И. Плеснера. Л. С. Понтрягин в своих воспоминаниях о казанском периоде жизни по этому поводу пишет: «Ко мне подошел казанский математик В. В. Морозов с предложением поселиться в квартире, где жил на правах жильца.[2] Хозяева его тоже сочли нас наиболее подходящими. Мать и жена поселились в очень маленькой комнатке, а я поселился в большой комнате с Морозовым. Он как сосед вел себя очень деликатно, совершенно не мешал мне спать, кроме

---

[2] В. В. Морозов жил в квартире своей племянницы Т. В. Семенихиной. — Прим. автора.

одного-единственного способа. Проснувшись утром рано, он тихонечко закуривал, не производя никакого шума, но дым папиросы сразу же будил меня. Морозов, конечно, не мог этого думать, а я стеснялся ему сказать» ([8]).

В квартире Н. Г. Чеботарева поселилась семья академика Г. С. Ландсберга, а также родственники Н. Г., также эвакуированные из Москвы.

Л. А. Орбели занял квартиру, которую в свое время занимал Н. И. Лобачевский. Академики Е. А. Чудаков, А. Н. Крылов, О. Ю. Шмидт, А. Ф. Иоффе, П. Л. Капица, член.-корр. И. Е. Тамм поселились во дворе университета, в комнатах исторического геометрического кабинета («геометрички», как его любовно называют казанцы) и «механички».

В «геометричке» расположился также Стекловский институт. Л. С. Понтрягин в своих воспоминаниях пишет: «Математический институт помещался в здании Казанского университета и занимал небольшое помещение. Рядом с этим помещением была отделена квартира академику Чудакову. А уборная от этой квартиры выходила в Математический институт, но запиралась ключом, который находился у Чудакова. Это дало Люстернику повод сострить. Он дал новое название нашему институту: математический институт имени Стеклова при уборной академика Чудакова» ([8]).

Многих эвакуированных поместили в общежитиях университета, других вузов Казани (проживавшие в общежитиях студенты были переведены в частный сектор), а также в различных помещениях города, в которых были необходимые условия для проживания. Например, большая семья С. Л. Соболева (жена, четверо детей и две бабушки) поселилась в комнатах «Дворца труда». Там же поселился Д. К. Фаддеев с женой и сыном (будущим академиком Л. Д. Фаддеевым). В этом здании проводились заседания некоторых научно-исследовательских семинаров, а также читались лекции для студентов университета.

Положение с продовольствием в городе стало тяжелым. Каждый день сокращалась подача топлива, воды, электричества. Не ходил трамвай, «пешеходный» способ передвижения стал практически единственным. Собственный автомобиль имел лишь О. Ю. Шмидт. Ректор университета К. П. Ситников передвигался по городу в пролетке, запряженной лошадью.

Осенью 1941 г. была введена продуктовая карточная система. Служащие университета получали в день по 400 г хлеба, в месяц по 300 г сахара или кондитерских изделий, 1 200 г мяса и рыбы, 300 г жиров, 800 г круп и макарон. Не работающим членам их семей полагалось в день по 400 г хлеба, в месяц по 200 г сахара, 500 г мяса и рыбы, 200 г жиров, 600 г круп. Столько же получали студенты. Детям преподавателей в возрасте до 12 лет выдавали в день по 400 г хлеба и (в месяц) по 300 г сахара, 400 г мяса и рыбы, 300 г жиров, 800 г круп и макарон. Люди ночами стояли в очереди, чтобы получить свою пайку.

Это было скудное существование, но оно распространялось на всех, а потому не рождало протеста. Тем не менее, существовала и определенная иерархия при распределении продуктов и товаров. Служащие эвакуированных в Казань академических институтов снабжались несколько лучше. Например, математики, занимающиеся прикладными проблемами, одно время получали по 800 г хлеба, а остальные — по 600 граммов. Правда, вскоре распоряжением О. Ю. Шмидта эта дифференциация между «прикладниками» и «теоретиками» была отменена (П. С. Александров [1, с. 260]).

Л. С. Понтрягин вспоминает:

«В эвакуации научные работники были обеспечены пищей наравне с рабочими, несущими тяжелую физическую работу. Именно: я получал 800 граммов хлеба в день, а мои члены семьи по 400. Кроме хлеба мы получали еще какое-то количество водки, а что касается масла, сахара и мяса, то их было ничтожно мало. Иногда Академия наук производила своим членам разовые выдачи пищи. Однажды мы получили восемь килограммов какао-бобов. Мы пропускали их через мясорубку и получалась маслянистая масса. Прибавляя к ней некоторое количество сахара, можно было получить нечто вроде шоколада. Кроме того, что мы получали по карточкам, а это был в основном хлеб, мы могли пользоваться еще столовой, организованной для нас Академией наук. Столовая эта организовывалась несколько раз заново. Менялось помещение, и питание устанавливалось новое. Но каждый раз питание быстро ухудшалось и всегда было почти совершенно отвратительным. Самое лучшее, что я помню в этих столовых, — это так называемый бигорох, т. е. обед, состоящий из горохового супа и гороховой каши. В питании строго соблюдалась табель о рангах. Академики, члены-корреспонденты, доктора, кандидаты — все снабжались по-разному.

В столовой, где нас кормили бигорохом, была отдельная комнатка для академиков, куда не пускали уже и членкоров. Но это только теоретически. Я часто туда проникал. Удобство заключалось в том, что пальто можно было повесить на стену, а не держать на том стуле, где сидишь. Но иногда эту комнатку контролировали академические дамы и выводили членкоров» ([8]).

Весной 1942 г. семьи ученых и сотрудников получили участки земли под огороды в пригородной деревне Лесные Моркваши, на берегу Казанки у Кремля, а также в районе нынешней улицы Гвардейской. Выращивали в основном картошку, помидоры, огурцы, тыкву, кабачки, морковь, капусту, турнепс. Многие приезжие математики, включая академиков, усердно трудились на своих огородах и получали неплохой урожай, обеспечивший их собственными овощами почти на всю зиму. Л. С. Понтрягин вспоминает, с каким усердием и удовольствием они с академиком И. М. Виноградовым ухаживали за своими огородами и, как правило, получали хороший урожай. Однажды осенью им разрешили выкапывать и брать с собой на каком-то поле сколько угодно моркови, так как копать ее было некому. Они копали вчетвером — Л. С. Понтрягин с женой, А. Н. Колмогоров и П. С. Александров.

Иногда находились и «чисто академические» решения продовольственной проблемы. По инициативе академика Л. А. Орбели на Волге был организован сбор двустворчатых моллюсков, восполнявших отсутствие в рационе питания белковой пищи. В университете из них готовили супы, делали фарш. Моллюсков использовали и как лечебное питание для раненых в госпиталях. Член-корреспондент Л. А. Галин написал в честь этого шутливую оду:

Вслед Господу начнем мы песню
о скользких моллюсках,
тех, что питанием служили мужам
благодарной науки...
Скажем лишь только, что из моллюсков
съедобных котлеты
ели и ими остались довольны, и всякого
есть призываем.
В реке сарматской Казанке мы много
моллюсков ловили.
очень больших и превкусных.

Ученых всех рангов постоянно привлекали на многочисленные субботники и воскресники: грузили уголь, разгружали вагоны и баржи, расчищали от снега посадочную полосу аэропорта. С. М. Никольский в своих воспоминаниях пишет:

«Однажды осенью 1942 г. Математический институт в полном составе принял участие в разгрузке дров с баржи на Волге. Явились все, начиная с лаборантов и кончая академиками.

Энтузиазм был невероятный. Иван Матвеевич[3)] выбирал себе самые крупные бревна, и сам клал их себе на плечи. Большие бревна брали С. Л. Соболев, А. Д. Александров, Б. Н. Делоне. Вообще все старались в меру своих сил, объединяясь в случае необходимости. Работа продолжалась целый день, воистине это был субботник ленинского типа» ([7, с. 227]).

Ученые активно участвовали и в сборе средств в Фонд обороны. Газета «Красная Татария» 16 января 1943 года писала: «Члены Академии наук принимают активное участие в строительстве боевого вооружения

[3)] И. М. Виноградов. — Прим. автора.

для Красной Армии. Академик Виноградов внес 85000 рублей своих сбережений на постройку эскадрильи самолетов. По 25000 рублей внесли академики Орбели и Тарле. От академика Кржижановского поступило 22000 рублей, от академика Абросимова — 21000 рублей.» «Красная Татария» в номере от 6 июня 1943 года также отметила, что академик Виноградов подписался на новый заем государству в сумме 7,5 тысяч рублей.

Несмотря на сложные условия, связанные с эвакуацией, и трудности материально-бытового характера, научно-исследовательская работа в Казани в годы войны была организована замечательно, и она ознаменовалась новыми достижениями как эвакуированных, так и казанских ученых.

Перед эвакуацией в Казань, в октябре 1941 года, состоялось собрание Московского математического общества, на котором было принято решение о временном разделении Общества на два отделения: Казанское и Ташкентское (куда был эвакуирован МГУ и некоторые учреждения АН СССР), руководителем Казанского отделения стал президент Общества П. С. Александров, секретарем — А. И. Мальцев. В Казани заседания Казанского отделения Московского математического общества проходили еженедельно по вторникам совместно с Казанским физико-математическим обществом (президентом которого был Н. Н. Парфентьев). Первое совместное заседание состоялось 4 августа 1941 года под председательством Н. Н. Парфентьева. На нем были рассмотрены организационные вопросы и вопросы, связанные с обеспечением жизнедеятельности эвакуированных в Казань математиков и членов их семей. На последующих заседаниях по очереди председательствовали Н. Н. Парфентьев и П. С. Александров. 1 октября 1943 г. Павел Сергеевич вернулся в Москву и объединенные заседания продолжались под председательством Н. Н. Парфентьева.

Всего за период с августа 1941 года до возвращения Стекловки в Москву состоялись 32 совместных заседания, на которых делали доклады П. С. Александров, Л. С. Понтрягин, Н. Г. Чеботарев, А. Я. Хинчин, А. О. Гельфонд, С. М. Никольский, П. С. Новиков, Л. В. Келдыш, Д. К. Фаддеев, А. Н. Колмогоров, И. М. Гельфанд, А. М. Обухов, Ю. В. Линник, Б. Н. Делоне, А. И. Мальцев, Л. А. Люстерник, А. Н. Крылов, А. Д. Александров, А. А. Марков, И. Р. Шафаревич, А. А. Ляпунов и другие. Из казанских математиков на заседаниях Общества делали доклады, кроме Н. Г. Чеботарева, Б. М. Гагаев и В. В. Морозов. Последний на двух заседаниях изложил основные результаты своей докторской диссертации, защищенной весной 1943 года (его оппонентами на этой защите были А. И. Мальцев и Л. С. Понтрягин).

В декабре 1942 г., состоялось (под председательством П. С. Александрова) торжественное заседание Московского общества, посвященное его 75-летию (1867–1942). Как пишет П. С. Александров [1], «на этом заседании с очень интересными докладами общего (историко-математического) характера выступили Алексей Николаевич Крылов и Александр Яковлевич Хинчин».

Перед научными силами Республики и сотрудниками эвакуированных в Казань институтов была поставлена задача консолидации усилий для поиска эффективных решений технологических, производственно-технических и научных проблем. С этой целью в сентябре 1941 года Президиум АН СССР образовал во главе с О. Ю. Шмидтом комиссию для организации и руководства научными работами в республике, получившей название «оборонной». Комиссия по согласованию с республиканскими планирующими органами наметила основные направления в деятельности научных коллективов: организация научно-технической и консультативной помощи промышленности, выявление и мобилизация местных ресурсов, а также выполнение теоретических и прикладных исследований.

Были организованы несколько научно-исследовательских семинаров с участием казанских и эвакуировавшихся математиков. На семинаре по баллистике в аэродинамике, организованном Н. Г. Чеботаревым, активно участвовали, кроме Чеботарева и его аспирантов, Л. С. Понтрягин, Н. Н. Мейман, Д. К. Фаддеев и ряд других математиков и механиков. На этом семинаре предметом особого внимания были вопросы по теории миномета, лобового сопротивления самолетов и вопросы турбулентности, в частности разрабатывалась тема «Вопросы теории подъемной силы и сопротивления самолета». Н. Г. Чеботарев со своими учениками исследовал проблему вибрации стволов морских орудий при выстреле и тесно связанную с ней проблему Гурвица для трансцендентных уравнений. Полное решение проблемы для наиболее важных для технических приложений случаев были получены Н. Г. Чеботаревым и Л. С. Понтрягиным.

В Казани и А. Н. Колмогоров занимается теорией стрельбы. Как пишет А. Н. Ширяев [9, с. 81–82], эти его занятия были ответом на запрос «дать свое заключение по поводу разногласий имеющихся приемов оценки меры точности по опытным данным». К казанскому периоду относится его (как пишет П. С. Александров [1. с. 260]) «знаменитые заметки по теории турбулентности», опубликованные в 1941 году в Докладах Академии наук.

В годы войны Андрей Николаевич со своими сотрудниками по Математическому институту, механико-математическому факультету университета и непосредственными практиками из Артиллерийского научно-исследовательского морского института разворачивает большую теоретическую и расчетную работу по эффективности систем стрельбы. Завершается она появлением отдельного выпуска «Трудов МИАН» (Андрей Николаевич называл его «Стрельбным сборником»).

П. С. Александров в своей знаменитой «казанской» работе, написанной им в 1941–1942 годах, исследовал гомологическую последовательность пары компактных топологических пространств. В ней впервые выписаны все элементы точной последовательности, активно используемого инструмента в тех разделах математики, которые используют алгебраические методы. П. С. Александров об этой своей работе пишет: «Я одновременно писал эту работу по-английски и по-русски, так что один текст представлял собою точный перевод другого. Русский текст был напечатан в „Известиях Академии наук“ (1943 г.) и, естественно, считался основным. Английский перевод был по предложению Лефшеца напечатан в “Transactions of the American Mathematical Society”, — знак большого внимания к этой работе, так как в “Transactions” переводных статей не печатают» ([1, с. 260]). Эта работа Александрова тогда же была удостоена Сталинской премии первой степени.

Работы казанского периода Ю. В. Линника относятся к аналитической теории чисел. Здесь он начал разрабатывать свой знаменитый плотностный метод в теории L-рядов Дирихле, позволивший ему получить решения целого ряда классических проблем теории чисел, в частности решение проблемы о наименьшем простом числе в арифметической прогрессии. Ю. В. Линник в Казань приехал в составе Ленинградского отделения Стекловского института в июне 1942 года, можно сказать, прямо с полей сражений: он летом 1941 года уходит добровольцем во фронт и до демобилизации летом 1942 года участвует в обороне Ленинграда в районе Пулковских высот, командуя артиллерийским взводом. Ю. В. в 1942 году было всего 27 лет, но он успел уже принять участие также и в Финской войне и защитить докторскую диссертацию (в возрасте 24 лет).

Д. К. Фаддеев в Казани занимался задачей погружения полей, являющейся естественным обобщением обратной задачи теории Галуа. В процессе этой работы им была открыта теория когомологий групп. По воспоминаниям Л. Д. Фаддеева (см. [2]), «в один из „казанских вечеров“ в 1943 году отец ходил по комнате весь возбужденный, и восклицавший, что он открыл нечто замечательное (как оказалось позже — это были коциклы)». Одновременно теорию когомологий групп открыли С. Эйленберг и С. Маклейн, которые пришли к ней, исходя из совсем другого вопроса. Создание теории когомологий групп считается одним из самых значительных математических событий середины двадцатого столетия.

Сотрудники кафедры механики проводили исследования по специальной теме Главного артиллерийского управления. С участием А. Н. Колмогорова и С. Л. Соболева решались задачи по расчету артиллерийской стрельбы, бомбометания. В докладе о работе Академии наук СССР в годы войны академик П. Л. Капица отмечал: «Основанное только на теоретических предпосылках улучшение формы снаряда без дополнительной затраты пороха и увеличения прочности ствола орудия позволило увеличить дальность стрельбы примерно на 10 процентов».

В. В. Морозов в своих воспоминаниях об этом периоде жизни казанских математиков отмечает: «У математиков с первых же дней создалось хорошее содружество, и если атмосфера геометрического кабинета была прохладной, если не сказать большего, то это не мешало теплым отношениям и плодотворной работе. Здесь И. М. Гельфанд разрабатывал свою теорию унитарных представлений групп, А. И. Мальцев написал работу о подалгебрах полупростых алгебр Ли, А. Д. Александров в доме у Фуксовского сада писал свою книгу о выпуклых поверхностях, П. С. Александров опубликовал работу, которую он сам цитирует как „казанскую“ и т. д.» ([6, с. 44]).

Коллектив Института физических проблем под руководством П. Л. Капицы работал по созданию новых методов достижения низких температур и получения жидкого кислорода. Прибыв в июле 1941 года в Казань, Институт физических проблем сразу же приступил к монтажу оборудования. И скоро кислород стал поступать в казанские госпитали. В Казани П. Л. Капица создал самую мощную в мире турбинную установку для получения его в больших количествах, необходимых в военной промышленности.

Будущий «отец» атомного оружия в СССР И. В. Курчатов прибыл в Казань в январе 1942 года. Он сразу же заболел сыпным тифом, а затем — воспалением легких. Казанский период жизни Курчатова отмечен тем, что здесь ученый стал отращивать свою знаменитую бороду, и коллеги дали ему прозвище «бородач». В октябре Курчатов выехал в Москву, где получил правительственное задание по проведению ядерных исследований и созданию урановой бомбы. Вернулся в Казань в декабре, а 9 января 1943 г. окончательно выехал в Москву для работы над созданием атомной бомбы. Семья Курчатова временно оставалась в Казани, где его брат завершал учебу на химическом факультете Казанского университета.

В своих воспоминаниях о казанском периоде жизни П. С. Александров пишет: «Заметными событиями в математической казанской жизни той зимы были защиты двух докторских диссертаций (будущими академиками) — А. И. Мальцевым и С. М. Никольским и двух кандидатских (обе по общей топологии) — С. В. Фоминым и Н. А. Шаниным[4)]. Диссертации Мальцева и Никольского завершались скромными по условиям военного времени „банкетами“ в маленькой чердачной комнатушке (под крышей главного университетского здания), в которой проживали оба диссертанта. Гостями на этих „банкетах“ были А. Н. Колмогоров и я. Их атмосфера запомнилась мне своей какой-то особой уютностью и сердечностью, и сами банкеты были, как впрочем и вообще наши встречи в ту зиму в Казани,

4) Н. А. Шанин был эвакуирован в Йошкар-Олу и в Казани бывал наездами. — Прим. автора.

как бы продолжением наших незабвенных лодочных путешествий, только уж в другой, более суровой обстановке» [1, с. 262].

Как уже говорилось, докторскую диссертацию по алгебре в 1943 году защитил и казанский математик — В. В. Морозов.

Большой успех и резонанс имели организованные Академией наук СССР юбилейные мероприятия, посвященные выдающимся деятелям науки: 24–25 ноября 1943 года в стенах Казанского университета прошла научная сессия, посвященная 300-летию со дня рождения Исаака Ньютона. Тогда же, в 1943 году, отмечалось 400-летие со дня смерти Коперника и 300-летие со дня смерти Галилея, а 25–28 ноября 1943 г. в актовом зале университете состоялось празднование 150-летия со дня рождения Н. И. Лобачевского, организованное АН СССР и физико-математическим обществом. Для участия на этих торжествах приехали воспитанники Казанского университета, известные геометры А. П. Котельников и Д. М. Синцов. По итогам этих мероприятий были опубликованы сборники «Галилео Галилей», «Исаак Ньютон», «Николай Коперник», «Николай Иванович Лобачевский». Кроме того, было принято решение об издании полного собрания сочинений Н. И. Лобачевского.

В те годы деканом физико-математического факультета был П. А. Широков. Пользуясь возможностью привлечь оказавшихся в Казани крупнейших математиков и физиков страны к чтению лекций для студентов, аспирантов и научных работников, он сумел организовать для них целый ряд интересных курсов. Почти все из приезжих ученых охотно на это шли. П. С. Александров читал курс по теоретико-множественной топологии, Л. С. Понтрягин — по комбинаторной топологии, А. И. Мальцев по теории непрерывных групп. И. Р. Шафаревич прочитал специальный курс по теории алгебраических полей. С. Л. Соболев читал специальные курсы по математической физике, академик Н. Е. Кочин по аэродинамике, Б. Н. Делоне читал лекции по аналитической геометрии и, по просьбе П. А. Широкова, который глубоко интересовался проблемами кристаллографии, математическим основам кристаллографии. А. Я. Хинчин прочитал несколько спецкурсов по вариационному исчислению и теории интегрирования, Н. Г. Четаев для студентов-механиков и аспирантов читал лекции по теоретической механике.

Начавшиеся осенью 1941 г. занятия вскоре были прерваны, поскольку коллектив университета во главе с ректором К. П. Ситниковым, профессорами М. В. Марковым, Б. А. Арбузовым и А. Н. Вознесенским (всего 750 студентов, преподавателей и сотрудников) 26 октября отправился на строительство защитных рубежей за Волгу, в Кайбицкий район.[5] Профессора

[5] По решению Государственного Комитета Обороны на областную партийную организацию и городской комитет обороны была возложена задача организовать и

провели «на окопах» две недели, а остальные работали там до 8 февраля 1942 г. Студентка Казанского университета тех лет Р. Иванова вспоминает: «Во время метели, когда не было видно впереди идущего человека, ректор брал в руки веревку, и остальные шли за ним, держась за нее, след в след. Зима 1941/42 года выдалась холодная (до минус 50 градусов). У многих не было теплой одежды. Для „утепления“ им привязывали сверху одеяла и повязывали поверх шапок платки. Вскоре из университета прислали воз лаптей с онучами, изготовленными из красных штор, снятых с аудиторных окон».[6)] Известно, что именно в Казани С. Л. Соболев научился вязать и связал себе свитер, затем научил этому ремеслу детей.

К лету 1943 года положение на фронте уже достаточно стабилизировалось, и, как пишет в своих воспоминаниях В. В. Морозов, «если 40-й том ДАН издавался в Казани, то 41-й — уже в Москве». Началась реэвакуация находящихся в Казани научных учреждений.

Научная жизнь в Казани так же интенсивно продолжалась и после отъезда институтов Академии наук из Казани, хотя математическая общественность Казани вскоре понесла тяжелые потери: не выдержало трудностей военного лихолетья больное сердце П. А. Широкова, 26 февраля 1944 года он скончался. Еще раньше, 22 января 1943 года, скончался Н. Н. Парфентьев. Казанское физико-математической общество после его смерти возглавил Н. Г. Чеботарев. Одним из наиболее крупных достижений казанских ученых в последние годы войны является открытие в 1944 г. на физическом факультете Казанского университета профессором Е. К. Завойским, впоследствии академиком АН СССР, явления парамагнитного резонанса, которое послужило началом развития электронной техники. Школа Е. К. Завойского вырастила целую плеяду ученых и приобрела известность в мире. Е. К. Завойский стал Героем Социалистического Труда, лауреатом Ленинской премии. В годы войны и в первые годы после ее окончания в Казани была создана исследовательская база для оборонной и аэрокосмической индустрии. Эвакуация в Казань Президиума АН СССР и ее научных учреждений, успешная совместная работа казанских ученых с их коллегами из этих институтов, разумеется, имели положительное значение для укрепления творческого содружества московских, ленинградских и казанских ученых. Они способствовали тому, что 13 апреля 1945 года было принято постановление правительства СССР об открытии

начать возведение оборонительных сооружений на участке от административной границы ТАССР с Ульяновской областью до районного центра Чувашской АССР Цивильска. В этих работах приняли участие 32,5 тысяч казанцев, среди которых была и колонна университета.

[6)] Иванова Р. Г. «Студенческая жизнь в годы Великой Отечественной войны». Отдел рукописей и редких книг Республиканской научной библиотеки, ед. хр. 10097/1, л. 9.

Казанского филиала АН СССР. В нем были созданы физико-технический, химический, геологический, биологический институты, а также отдел водохозяйственных проблем и энергетики. В состав филиала вошел также созданный в 1939 году институт языка, литературы и истории. Первым председателем Президиума Казанского филиала стал академик А. Е. Арбузов, директором физико-технического института стал Н. Г. Чеботарев. Он же стал и заведующим организованной в институте секции математики.

В Казанском университете бережно хранят память об этих событиях военных лет. В Музее истории КГУ значительное место занимает выставка, названная «Линия научной обороны» и посвященная казанскому периоду жизни ученых московских и ленинградских институтов Академии наук СССР. Выставка содержит богатейшую коллекцию, насчитывающую свыше пятисот единиц. В ней документы, фотографии, книги и рукописи, письма и воспоминания, личные вещи академиков А. Ф. Иоффе, С. И. Вавилова, Л. Д. Ландау, А. Н. Фрумкина и других.

Список литературы

[1] Александров П. С. *Страницы автобиографии*. Часть 2 // Успехи мат. наук. Т. 35, вып. 3(213), 1980. С. 241–278.

[2] Востоков С. В., Шафаревич И. Р. *Гармония в алгебре* (*к столетию со дня рождения члена-корреспондента АН СССР Дмитрия Константиновича Фаддеева*) // Владикавказский мат. журнал. Т. 10, вып. 1, 2008. С. 3–9.

[3] Вульфсон Г. Н., Муньков Н. П. *Страницы памяти* // Во имя отчизны. Казанский университет в годы Великой Отечественной Войны. Казань: КГУ, 1975. С. 42–63.

[4] *История Казанского университета*, 1804–2004. Казань: КГУ, 2004. — 368 с.

[5] *Мехматяне вспоминают*. М.: Мехмат МГУ, 2009. С. 10.

[6] Морозов В. В. *Взгляд назад* (*несколько страниц воспоминаний*) // Избранные вопросы алгебры и логики. Новосибирск: Наука, Сибирское отделение, 1973. С. 314–321.

[7] Никольский С. М. *Отрывки из воспоминаний о А. И. Мальцеве* // Успехи мат. наук. Т. 27, вып. 4(166), 1972. С. 223–230.

[8] Понтрягин Л. С. *Жизнеописание Льва Семеновича Понтрягина, составленное им самим. Рождения* 1908 *г. Москва*. М.: ИЧП «Прима В», 1998. — 304 с.

[9] Ширяев А. Н. *Жизнь и творчество. Биографический очерк* // Колмогоров. Юбилейное издание в трех книгах. Книга первая «Истина — благо». М.: Физматлит, 2003. С. 17–210.

М. М. Арсланов, Казанский государственный университет, механико-математический факультет; НИИ математики и механики им. Н. Г. Чеботарёва

Знаменитые теоремы

Теорема Гёделя о неполноте и четыре дороги, ведущие к ней

(Лекции, прочитанные в июле 2007 г. и в июле 2009 г. на летней школе «Современная математика» в Дубне)

В. А. Успенский

Часть первая: вступление

§1. Что такое доказательство?

Теорема Гёделя о неполноте, пожалуй, — одна из двух (наряду с Теоремой Ферма) самых знаменитых теорем математики. Не самая главная (таковой — по моему мнению, с которым многие не согласятся, — является теорема о существовании фактор-множества для всякого отношения эквивалентности), не самая известная (таковой является Теорема Пифагора), но — самая знаменитая. И это вполне заслуженно: ведь её можно считать теоремой теории познания. Да и автор её — человек достаточно знаменитый. Когда известный американский журнал «Тайм» составлял список ста самых выдающихся деятелей уходящего XX века, на всю науку он отвёл десять мест, из них на математику — одно. Это место и занял Курт Гёдель (Kurt Friedrich Goedel, 28.04.1906–14.01.1978).

Хотя Теорема Ферма, по внематематическим (и отчасти скандальным) причинам, возможно, и знаменитее Теоремы Гёделя, думается, что какое-то представление о Теореме Гёделя имеет более широкий круг людей. В самом первом и самом грубом приближении Теорема Гёделя утверждает, что невозможно доказать все истины, иными словами, что *существуют недоказуемые истины*.

Такая формулировка, однако, содержит в себе очевидное противоречие. Вот оно. Раз Гёдель доказал существование утверждения, которое, хотя и недоказуемо, является истинным, то, значит, он доказал истинность этого утверждения. Но ведь доказать истинность некоторого утверждения — это и значит доказать это утверждение. Получается, что Гёдель доказал обсуждаемое утверждение, — но тогда оно не может быть недоказуемым. Как же быть?

Всё дело в том, что существуют два различных понятия доказательства и, следовательно, два различных понятия доказуемости. Первое из них хорошо известно всем из школы, да и из повседневной жизни тоже; мы будем называть его *содержательным*, или *психологическим*, доказательством. Второе принадлежит математической логике; мы будем называть его *формальным* доказательством.

Понятие психологического доказательства потому названо *психологическим*, что оно принадлежит не математике, а психологии и, отчасти, лингвистике. Психологическое доказательство — это *содержательное рассуждение*, убеждающее нас в истинности какого-либо утверждения настолько, что мы готовы убеждать других в истинности этого утверждения при помощи **этого же рассуждения**. Именно такие доказательства и фигурируют как в школьных, так и в вузовских курсах математики. Да и в университетских курсах используются именно они.

Психологическое доказательство имеет место и тогда, когда что-то доказывают, исходя из каких-то аксиом, — скажем, аксиом геометрии или аксиом кольца. Ведь использование аксиом при доказывании теорем не меняет сути дела: доказательство и в этом случае состоит в рассуждении, имеющим целью убедить, что если предположить аксиомы истинными, то окажется истинным и доказываемое утверждение. Приведём простой пример. Рассмотрим множество, в котором выделен элемент 0 и на котором задана одноместная операция $'$. Известные аксиомы Пеано содержат в своём составе две такие: $(x' = y') \Rightarrow (x = y)$ и $\neg\exists x(x' = 0)$. Исходя из этих аксиом, докажем неравенство $0'''' \neq 0''$. Доказываем от противного. Предположим, что $0'''' = 0''$. В силу первой аксиомы тогда $0''' = 0'$. В силу той же первой аксиомы тогда $0'' = 0$. Но это противоречит второй аксиоме, потому что получается, что 0 есть результат применения операции $'$ к элементу $0'$. Точно таким же р а с с у ж д е н и е м доказывается различие любых двух элементов вида $0''^{\cdots}{}'$, имеющих в своей записи различное количество штрихов. Поэтому выражения

$$0,\ 0',\ 0'',\ \ldots$$

можно использовать в качестве обозначений, или имён, натуральных чисел.

Психологическое доказательство есть продукт социальной истории, и требования к убедительности меняются со временем. Не сомневаюсь, что в Древнем Египте написанный на папирусе исходящий из храма текст, содержащий рецепты, скажем, сложения дробей или вычисления площадей, считался не подлежащим оспариванию доказательством (независимо от того, верными или неверными оказывались указанные рецепты). В средневековой Индии доказательством служил чертёж, под которым было подписано «Смотри». Именно так, например, доказывалась формула $S = lr/2$, связывающая площадь S произвольного круга радиуса r с длиной l его окружности. Сейчас нас вряд ли устроит подобное доказательство (хотя, надо признать, оно довольно убедительно — и уж во всяком случае очень наглядно). Но, вполне возможно, те доказательства, которые мы сейчас считаем строгими, не устроят наших потомков — такую возможность допускал великий математик Пуанкаре.

Психологические, содержательные доказательства часто называют также *неформальными*. Мы видим, что тó, что именуется просто *доказательством* в обычной математической (но не в математикологической!) практике, и есть в точности тó, чтó мы только что назвали психологическим (неформальным, содержательным) доказательством.

В отличие от психологических доказательств, формальные доказательства являются математическими объектами, подобными, скажем, треугольникам или шарам. В реальном мире нет ни треугольников, ни шаров в точном смысле этих слов, но есть нечто, идеальными образами, математическими формализациями чего служат треугольники и шары. Аналогичным образом, понятие формального доказательства служит формализацией понятия психологического доказательства. Важно подчеркнуть, что понятие психологического доказательства допускает различные формализации, подчиняющееся неким общим для всех таких формализаций требованиям. Более точно понятие формального доказательства будет описано в §2. Пока же мы ограничимся заявлением, что формальное доказательство есть конечная последовательность знаков, определённым образом организованная.

Мы в состоянии теперь сформулировать Теорему Гёделя чуть-чуть более аккуратно: *какую бы формализацию понятия доказательства ни предъявить, всегда найдётся истинное утверждение, которое не имеет формального доказательства в рамках этой формализации.* Противоречие, о котором шла речь выше, устранено: истинность утверждения доказывается при помощи психологического доказательства, а его недоказуемость означает отсутствие для него формального доказательства.

Ясно, однако, что столь общая теорема не может быть верной без каких-то ограничений. Во-первых, если запас утверждений, истинность и доказуемость которых исследуется, очень беден, то можно предложить такую формализацию понятия доказательства, при котором все истинные утверждения из этого запаса окажутся доказуемыми. За примером ходить недалеко — именно так будет, скажем, в случае, когда все утверждения суть утверждения о равенстве или неравенстве арифметических выражений, изучаемых в начальной школе, то есть выражений, составленных из натуральных чисел и операций сложения и умножения. Чтобы язык подпадал под действие Теоремы Гёделя, он должен быть богатым, то есть обладать такими выразительными средствами, при помощи которых можно выразить действительно содержательные утверждения. Поэтому формулировку Теоремы Гёделя следует начать с ограничительной клаузулы «если язык достаточно богат». Точное определение термина *достаточно богатый язык* будет дано в конце §4. Во-вторых, нетрудно ввести такое понятие формального доказательства, при котором решительно все утверждения языка, как истинные, так и ложные, будут обладать формальным доказательством. Для этого достаточно, например, объявить, что запись любого утверждения является его формальным доказательством. В советское время именно так обстояло дело со всеми утверждениями, содержащимися в трудах Ленина и Сталина: формальным доказательством истинности такого утверждения служило его предъявление в составе официального издания.

Ясно, что если не позаботится об исключении подобных ситуаций, Теорема Гёделя окажется неверна. Исключение естественнее всего достигается требованием семантической непротиворечивости. Рассматриваемое понятие формального доказательства называется *семантически непротиворечивым*, коль скоро выполнено следующее требование: *никакое ложное утверждение не может обладать формальным доказательством.* Другая ограничительная клаузула, следовательно, должна иметь вид «если понятие формального доказательства семантически непротиворечиво».

Заметим ещё, что тот вариант Теоремы Гёделя, о котором говорилось до сих пор, есть так называемый *семантический* её вариант: он использует в своей формулировке представление об истинности подлежащих доказыванию утверждений. Известен и так называемый *синтаксический* вариант, не использующий указанного представления. Именно в синтаксическом варианте Гёдель и доказал в 1930 г. свою знаменитую теорему. В самом первом и самом грубом приближении Гёдель доказал, что существуют утверждения, которые нельзя ни доказать, ни опровергнуть. В более аккуратной формулировке, синтаксический вариант Теоремы Гёделя гласит, что какую бы формализацию понятия доказательства ни предъявить, всегда найдётся такое утверждение — причём даже в арифметике, то есть среди утверждений о натуральных числах, — что ни его само, ни его отрицание невозможно доказать в рамках предъявленной формализации. Синтаксический вариант требует своих ограничений. Требование семантической непротиворечивости заменяется требованием *синтаксической непротиворечивости*; последняя состоит в том, что отрицание утверждения, обладающего формальным доказательством, само не должно обладать таковым. Требование достаточного богатства языка заменяется предъявлением списка конкретных утверждений, относительно которых выдвигается требование, чтобы они обладали формальными доказательствами. Синтаксический вариант Теоремы Гёделя обсуждается в Дополнении к этому очерку.

Но вернёмся к семантическому варианту. Сформулируем его чуть более точно: какое бы строго определённое понятие формального доказательства ни предъявить, всегда найдётся истинное утверждение (причём даже в арифметике), которое невозможно доказать в рамках этого понятия.

Знаменитый многотомный трактат Николя Бурбаки «Начала математики» начинается со слов: «Со времён греков говорить „математика“ — значит говорить „доказательство“». Древние греки тут не случайны. Именно они создали современную европейскую математику. Именно у них возникло и само понятие доказательства. Дело в том, что первые доказательства, то есть убедительные рассуждения появились у греков в их народных собраниях и в судах — в прениях сторон. С этой точки зрения, математика — младшая сестра юриспруденции.

Математика, как известно, не отменяет, но уточняет содержательные, интуитивные представления. Так, математическое понятие действительного числа уточняет интуитивное, физическое представление о длине отрезка. Да и математическое понятие отрезка уточняет соответствующее физическое понятие. Ведь математических отрезков в реальном мире не бывает. Математическое уточнение того или иного неформального представления состоит в том, что математика предлагает некоторое строгое, точно очерченное понятие, не совпадающее, конечно, с указанным неформальным представлением, но всё же схватывающее его существенные черты. В качестве уточнения психологического представления о неформальном доказательстве математика и предлагает понятие формального доказательства. Именно о них, а не о содержательных доказательствах говорит Теорема Гёделя, точно так же, как теоремы геометрии — об идеальных математических точках, прямых, окружностях и треугольниках. (Вряд ли кто-нибудь сможет указать «физическую точку», лежащую на «физической прямой», да таких объектов и не бывает в реальности; однако несомненно, что математические точка и прямая отражают некоторые существенные черты реального мира.) Я потому так долго говорю на эти философские темы, что хочу убедить уважаемого читателя, что по своей логической природе понятие формального доказательства не отличается от других идеальных понятий математики.

§2. Формальные языки

Формальные доказательства составляют предмет специального раздела математики — теории доказательств. Теория доказательств, в свою очередь, является частью более крупного раздела математики — математической логики. Математическую логику можно определить как учение о формальных языках. (Здесь используется узкое понимание термина «математическая логика». В широком понимании математическая логика включает в себя также основания математики и теорию алгоритмов.) Формальные языки, в отличие от естественных языков, это такие языки, синтаксические системы которых точно и полностью описаны.

Синтаксическая система, или просто *синтаксис*, какого-либо языка указывает, какие выражения этого языка являются правильными, а какие нет. Так, синтаксическая система русского языка указывает, что выражения «он бежал» и «глядя из окна, у меня слетела шляпа»[1] являются правильными, а выражения «она бежал» и «глядя окна, мною слетела шляпа» являются неправильными. Но для русского языка, как и для любого другого естественного языка, так описать его синтаксическую систему, чтобы это описание было одновременно и точным, и полным, весьма затруднительно, если вообще возможно; во всяком случае, это никому ещё не удавалось.

Напротив, формальные языки потому и называются формальными, что в них среди всех возможных сочетаний знаков совершенно точными, формальными инструкциями выделены правильные выражения. Простейший пример — язык для записи шашечных позиций. Как известно, шашки бывают белыми и чёрными и, помимо того, подразделяются на простые и дамки, а расставляются исключительно на чёрных полях шашечной доски. Из чувства патриотизма будем подразумевать русские шашки, а посему наша доска будет 64-клеточной, а шашек будет 24 — по дюжине каждого цвета. Для простоты под *позицией* будем понимать любое расположение шашек на доске, невзирая ни на количество этих шашек (лишь бы на каждом поле стояло не более одной шашки), ни на цвет полей на которых они стоят. Под *правильным выражением* будем понимать запись позиции в шашечной нотации.

Пример выражения:
Белые: дамки a2, a3, a4, g8; простые a8, e1. Чёрные: дамки b1; простые f5.

Слегка более сложный язык — язык записи шахматных позиций. Предоставляем читателю записать позицию, где на доске стоят 64 чёрных короля.

---

[1] Это выражение означает, что шляпа слетела, когда она глядела из окна.

Читатель, при желании, легко составит синтаксические правила для образования правильных выражений двух наших языков, «шашечного» и «шахматного».

Среди позиций выделим *допустимые* — так мы будем называть те позиции, которые могут реально встретиться в какой-либо партии, пусть очень «глупой», но сыгранной без нарушения правил. Как описать допустимые позиции или, что то же самое, соответствующие им допустимые записи? Для этого удобно прибегнуть к формальным доказательствам. Нет ничего страшного в том, чтобы говорить о формальном доказательстве позиций или их записей. Возможно, некоторым читателям такие обороты речи покажутся слишком странными и они скажут, что доказательствами, хотя бы и формальными, могут обладать лишь утверждения. Извольте, будем говорить о формальных доказательствах утверждений вида «данная позиция (запись) является допустимой».

Говорить вообще о формальных доказательствах можно лишь в условиях, при которых и подлежащие доказательству утверждения, и формальные доказательства представляют собою тексты, организованные по совершенно точным синтаксическим правилам. Иначе говоря, и те, и другие должны быть элементами тех или иных формальных языков. Формальный язык интересующих нас утверждений описан (в двух вариантах — для шашек и для шахмат). Рассмотрим теперь языки записи шашечных и шахматных партий. Оба этих языка имеют совершенно точный синтаксис, определяемый правилами соответствующей игры. Любая запись партии, сделанная с соблюдением синтаксиса, может считаться формальным доказательством допустимости той позиции, которая возникла в конце записанной партии.

Более детальная, и притом «математическая», иллюстрация следует ниже.

Любой конечный список знаков[2)] называется *алфавитом*. Эти знаки называются *буквами* алфавита. Всякая конечная цепочка букв данного алфавита называется *словом* в этом алфавите. Ограничимся «линейными» языками, в которых все выражения языка суть слова.

Всякий формальный язык начинается со своего алфавита. Синтаксическая система выделяет среди всех слов в алфавите языка *правильные выражения*. (Заметим, что для русского языка его традиционный алфавит придётся расширить, добавив в него знаки препинания, включая знак пробела, и ещё кое-что.)

Построим язык L_- с алфавитом из 7 знаков: $)\quad(\quad=\quad x\quad y\quad z\quad -$

[2)] Не имеется в виду, что эти знаки что-либо означают. Лучше бы говорить поэтому не «знаки», а «значки».

Знаки x, y, z называются *переменными*.

Правильные выражения формализованных языков обычно подразделяются на *термы* и *формулы*. На содержательном уровне *формулами* называются такие сочетания знаков, которые предназначены для выражения утверждений (например: «три чётно») или утверждений, зависящих от параметра или параметров (например, «x чётно», «$x + 2y$ чётно»). *Термы* выступают в качестве имён (например: «3+5») или имён, зависящих от параметра или параметров (например, «x» или «$x + 2y$»). На формальном же уровне и термы, и формулы определяются чисто синтаксически, без апелляции к их смыслу.

Термы языка L_- определяются индуктивно:

1. Всякая переменная есть терм;
2. Если s и t есть терм, то и $(s - t)$ есть терм.

Пример терма:

$$((z - y) - (x - (x - (z - x))))$$

Формулы языка L_- имеют вид $(s = t)$, где s и t суть термы.

На синтаксическом уровне построение языка закончено. Как правило, однако, формальный язык наделяется ещё и семантической системой или дедуктивной системой или и той, и другой.

Семантическая система, или просто *семантика*, какого-либо языка выделяет среди всех формул этого языка те, которые *объявляются истинными*; говорят также, что им *приписывается значение истина*.

Вот один из способов такого выделения (приписывания) для построенного формального языка. Вместо переменных будем подставлять натуральные числа, а знак «$-$» понимать как умножение. При каждой подстановке каждый терм приобретает *значение* в виде натурального числа. Например, если вместо x, y, z подставить, соответственно, 2, 5 и 3, то переменная y получит значение 5, а выписанный выше терм получит значение $((3 \times 5) \times (2 \times (2 \times (3 \times 2))))$, то есть 360. Разумно считать — и мы примем эту точку зрения, — что *знак «=» всегда понимается как обычное равенство*. Припишем формуле значение *истина* (объявим её *истинной*), коль скоро она превращается в истинное утверждение при любой подстановке, другими словами — коль скоро значения левой и правой части совпадают при любых значениях переменных. Тогда формула $(z - x) = (x - z)$ [3] окажется истинной. Мы нарочно выбрали столь необычную интерпретацию

[3] Читатель скажет, что это выражение не является формулой, и будет прав. Здесь использовано традиционное разрешение опускать внешние скобки. Получающееся при этом выражение рассматривается как всего лишь сокращение для подлинной формулы, каковая в нашем случае выглядит так: $((z - x) = (x - z))$. Указанное разрешение будет действовать и в дальнейшем изложении.

знака «$-$» чтобы подчеркнуть произвол в наделении языка семантикой (за исключением семантики знака «$=$»).

Приведём теперь пример другой, более естественной семантики, которую будем именовать *стандартной*. В этой семантике знак «$-$» понимается как разность, а вместо переменных разрешается подставлять любые действительные числа. Условия истинности формулы остаётся прежним: формула истинна, если при любой подстановке она превращается в истинное утверждение. Формула $(z-x)=(x-z)$ теперь уже не будет истинной. Зато истинными окажутся формулы

$$\text{I.}\quad x=(x-(y-y))$$
$$\text{II.}\quad (x-(y-z))=(z-(y-x))$$

Дедуктивная система какого-либо языка выделяет среди всех формул те, которые *объявляются доказуемыми*. В наших шашечном и шахматном языках вместо «доказуемый» говорилось «допустимый». В этом параграфе мы ограничимся такими системами, в которых доказуемость задаётся индуктивно при помощи *аксиом* и *правил вывода*. Это делается так. Некоторые формулы объявляются аксиомами. Каждое правило вывода применяется к одной или нескольким формулам и указывает, как из этих формул можно получить новую формулу. Далее, делаются два заявления. Во-первых, каждая аксиома объявляется доказуемой формулой. Во-вторых, объявляется, что если правило вывода применено к доказуемым формулам, то и полученная формула также считается доказуемой. В шашечном и шахматном языках в роли аксиом выступали записи начальных позиций, говоря точнее — их записи; правилами вывода служили правила игры.

Проиллюстрируем сказанное на примере языка L_{-}. Аксиома одна — это формула $y=y$. Правило вывода также одно, и оно таково. В любом месте любой формулы разрешается заменить y любым термом. Легко видеть, что доказуемыми окажутся все формулы. Этот пример — искусственный и неинтересный. Содержательный пример будет ниже.

Говоря формально, доказуемость и истинность никак не связаны. Целесообразно, однако, наделить язык такой дедуктивной системой, при которой все доказуемые формулы оказываются истинными; если этот эффект имеет место, дедуктивная система называется *корректной* относительно данной семантической системы. (Желательно при этом, что бы как можно больше истинных формул оказалось доказуемыми).

Теперь — обещанный содержательный пример. Аксиомами объявляются выписанные выше формулы I и II. Правил вывода два:

Правило подстановки. Даны: формула, переменная, терм. Разрешается подставить этот терм вместо этой переменной в эту формулу.

Правило замены. Даны две формулы: A и $(s = t)$. Разрешается в любом месте формулы A заменить s на t или t на s.

Проверим, что формула $(x - x) = (y - y)$ является доказуемой. Для этого выпишем следующую цепочку из 9 формул:

$$
\begin{aligned}
&(1)\quad x = (x - (y - y));\\
&(2)\quad (x - (y - z)) = (z - (y - x));\\
&(3)\quad z = (z - (y - y));\\
&(4)\quad z = (z - (x - x));\\
&(5)\quad (y - y) = ((y - y) - (x - x));\\
&(6)\quad (x - (x - z)) = (z - (x - x));\\
&(7)\quad (x - (x - (y - y))) = (z - ((y - y) - (y - y));\\
&(8)\quad (x - x) = ((y - y) - (x - x));\\
&(9)\quad (x - x) = (y - y).
\end{aligned}
$$

Двигаясь по этой цепочке, последовательно убедимся в том, что все входящие в неё формулы доказуемы. В самом деле, формулы (1) и (2) суть аксиомы и потому доказуемы. Формула (3) доказуема потому, что получается из доказуемой (1) подстановкой z вместо x. Аналогично, (4) получается из (3) подстановкой x вместо y. Далее, (5) получается из (4) подстановкой $(y-y)$ вместо z, а (6) из (2) — подстановкой x вместо y. Если в (6) подставим $(y - y)$ вместо z, получим (7). Правило замены позволяет, используя доказуемую формулу (1), поменять в доказуемой (7) $(x-(y-y))$ на x. Это же правило, применённое к (8) и (5), даёт формулу (9), которая тем самым оказывается доказуемой.

Введённая дедуктивная система корректна относительно стандартной семантики. В самом деле, обе аксиомы истинны в этой семантике, а правила вывода, будучи применены к истинным формулам, дают формулу, являющуюся истиной. Отсюда следует, что если формула не является истинной, то она не может быть доказуемой. В частности, недоказуемы формулы $x = y$, $(z - x) = (x - z)$ и т. п.

Цепочка формул (1)–(9) представляет собой пример формального доказательства. В общем случае *формальным доказательством* называется такая цепочка формул, в которой каждый член либо является аксиомой, либо получается из предыдущих формул той же цепочки согласно одному из правил вывода. Формальное доказательство считается формальным доказательством своей последней формулы.

Вот ещё один пример дедуктивной системы и формального доказательства в ней. Пусть заданы какой-то формальный язык и какая-то дедуктивная система. Большие латинские буквы означают в этом примере произвольные формулы нашего языка. Предполагается, что алфавит языка содержит букву $\Rightarrow$, а правила образования позволяют из любых двух формул X и Y построить формулу $(X \Rightarrow Y)$.

Предполагается, что среди правил вывода имеется такое

правило (*MP*): если доказуемы обе формулы $X \Rightarrow Y$ и X, то доказуема и формула Y.

Предполагается также, что среди аксиом имеется бесконечное число аксиом нижеследующего вида:

Группа I: $X \Rightarrow (Y \Rightarrow X)$

Группа II: $((X \Rightarrow (Y \Rightarrow Z)) \Rightarrow ((X \Rightarrow Y) \Rightarrow (X \Rightarrow Z))$

Требуется убедиться, что любая формула вида $A \Rightarrow A$ является доказуемой. Убеждаемся в требуемом, предъявляя цепочку формул, образующую формальное доказательство нашей формулы. Для наглядности рядом с каждой формулой цепочки указываем в квадратных скобках причину, позволяющую этой формуле быть членом формального доказательства.

Итак, вот формальное доказательство формулы $A \Rightarrow A$:

1. $A \Rightarrow ((A \Rightarrow A) \Rightarrow A))$ [Гр. I]
2. $(A \Rightarrow ((A \Rightarrow A) \Rightarrow A)) \Rightarrow ((A \Rightarrow (A \Rightarrow A)) \Rightarrow (A \Rightarrow A))$ [Гр. II]
3. $(A \Rightarrow (A \Rightarrow A)) \Rightarrow (A \Rightarrow A)$ [(MP) из 1 и 2]
4. $A \Rightarrow (A \Rightarrow A)$ [Гр. I]
5. $A \Rightarrow A$ [(MP) из 3 и 4]

В §1 было изложено содержательное доказательство некоторого утверждения арифметики, опирающееся на аксиомы Пеано. Опыт формализации изложенного рассуждения в виде формального доказательства дан в разделе «Формальный аксиоматический метод» моей книжки «Простейшие примеры математических доказательств» (М.: МЦНМО, 2009).[4)]

Фиксируем какой-либо формальный язык и какую-либо его дедуктивную систему. Если отделить члены формального доказательства один от другого посредством какого-либо разделительного знака, например точки с запятой, то оно предстанет в виде слова некоторого алфавита — *алфавита языка формальных доказательств*; обозначим этот алфавит буквой Д. Дальнейшее изложение опирается на следующие два обстоятельства. Во-первых, существует алгоритм, позволяющий по любому слову в алфавите Д эффективно распознать, является оно формальным доказательством или нет (ясно, что для этого правила вывода должны быть эффективными). Во-вторых, существует алгоритм, который по каждому формальному доказательству d даёт ту формулу, для которой d является формальным доказательством.

[4)] Названный раздел воспроизведён на с. 381–388 книги: *В. А. Успенский*. Апология математики. СПб: Амфора, 2010.

Часть вторая: последние приготовления

§ 3. Понятия и термины

Термины и обозначения. Начинать ли *натуральный ряд* с нуля или с единицы, иными словами, считать ли ноль *натуральным числом* или нет — это вопрос соглашения. Каждому решению присущи свои преимущества и недостатки. Мы примем, что ноль — натуральное число, тем самым включая его в натуральный ряд $\mathbb{N}$. Натуральное число как абстрактную сущность следует отличать от его записи в какой-либо системе обозначений. Принято, тем не менее, такие записи также именовать натуральными числами.

Термин *кортеж*, к большому сожалению, ещё недостаточно популярен. Этим термином обозначается конечная строка $(a_1, a_2, \ldots, a_n)$; при этом n называется *длиной* кортежа. Чаще, увы, такую строку называют «энкой», не задумываясь, что делать в случае $n = k$. Ясно, что каждую функцию от n аргументов можно трактовать как функцию одного аргумента — кортежа длины n.

Множество всех числовых (т. е. составленных из натуральных чисел) кортежей длины k обозначается через $\mathbb{N}^k$; при этом вместо $\mathbb{N}^1$ обычно пишут просто $\mathbb{N}$.

Говоря о *функции из X в Y*, мы не будем предполагать, что она непременно определена во всех точках множества X. Более того, среди таких функций встречается и функция, которая не определена ни в одной точке. Она так и называется *нигде не определённой функцией*; её называют также *пустой*. Если функция f не определена в точке a, то выражение $f(a)$ считается бессмысленным. Утверждение, что функция f определена в a записывается так: $!f(a)$.

Мы не решились менять термин «отображение», поэтому вынуждены заявить, что функция из X в Y задаёт, вообще говоря, не отображение, а частичное отображение. Понятия образа и полного прообраза при этом сохраняются: для функции f *образ* множества A есть множество

$$f(A) \ = \ \{y \mid \exists x \in A \ f(x) = y\},$$

а *полный прообраз* множества A есть множество

$$f^{-1}(A) \ = \ \{x \mid \exists y \in A \ f(x) = y\}.$$

В применении к выражениям, которые могут быть и не определены, вместо знака равенства $=$ используется *знак условного равенства* $\simeq$. Запись $A \simeq B$ означает, что выражения A и B определены или не определены одновременно и, если определены, то их значения совпадают. Таким образом, утверждение $x - x \simeq y - y$ верно, а утверждение $\frac{x}{x} \simeq \frac{y}{y}$ неверно. Поскольку мы не требуем, чтобы рассматриваемые функции были определены для всех значений аргумента, то совпадение функций φ и ψ надо выразить так: $\varphi(x) \simeq \psi(x)$. Множество $\{(a,\ b) \mid f(a) = b\}$ называется *графиком* функции f; аналогично — для функции нескольких аргументов.

Термины *алфавит*, *буква* и *слово в данном алфавите* были приведены в § 2. Длина слова x обозначается так: $|x|$. Совокупность всех слов в каком-либо алфавите Б называется *словарным пространством* и обозначается Б*. Подмножества словарного пространства называются *словарными множествами.* Кортеж слов в алфавите Б легко записывается в виде слова в большем алфавите, полученного из Б добавлением новой буквы в качестве разделительного знака. Поэтому прямое произведение словарных множеств уместно считать словарным множеством.

Натуральное число как запись (см. начало раздела) есть слово в подходящем алфавите.

Алгоритмы, вычислимые функции, перечислимые и разрешимые множества. Элементы словарного пространства нетрудно расположить в последовательность. Любой очевидный способ такого расположения приведёт к последовательности, которая окажется вычислимой. *Вычислимой* же называется, вообще, всякая последовательность, для которой существует алгоритм, дающий по номеру члена последовательности сам этот член.

Что касается понятия «*алгоритм*», то оно принимается как первичное, неопределяемое понятие. При применении к какому-либо своему *входу* алгоритм может либо давать *выход*, он же *результат*, либо не давать ничего. Таким образом, в области возможных входов данного алгоритма возникает как её часть *область результативности*, состоящая из тех входов, для которых применение алгоритма приводит к какому-либо результату.

Функция называется *вычислимой*, если её область определения совпадает с областью результативности какого-либо алгоритма и для любого аргумента значение функции совпадает с результатом применения этого алгоритма к этому аргументу. В частности, пустая функция вычислима: её область определения совпадает с областью результативности алгоритма, не дающего результата ни при каком входе. Всякая последовательность есть функция, и потому приведённое выше определение вычислимой последовательности согласуется с определением вычислимой функции.

Если данное множество есть область значений некоторой вычислимой функции, определённой на натуральном ряду (то есть последовательности), то говорят, что эта функция *перечисляет* это множество. Словарное множество называется *перечислимым*, если оно либо пусто, либо перечисляется некоторой вычислимой функцией. Объединение и пересечение перечислимых множеств перечислимы. Всякое конечное множество перечислимо. Если перечислимое множество бесконечно, то его можно перечислить (то есть расположить в вычислимую последовательность) без повторений.

Теорема о графике. *Функция тогда и только тогда вычислима, когда её график перечислим.*

Следствия. 1. *Область определения и область значений вычислимой функции перечислимы.* 2. *Образ и полный прообраз перечислимого множества при частичном отображении, задаваемом вычислимой функцией, перечислимы.*

Словарное множество называется *разрешимым*, коль скоро его характеристическая функция относительно объемлющего словарного пространства вычислима. Дополнение к разрешимому множеству разрешимо. Всякое разрешимое множество перечислимо.

Теорема Чёрча – Поста. *Множество разрешимо тогда и только тогда, когда оно само и его дополнение перечислимы.*

Основной контрпример теории алгоритмов. *В каждом словарном пространстве существует перечислимое подмножество с неперечислимым дополнением.*

Следствие. *Существует перечислимое множество натуральных чисел, дополнение коего до всего натурального ряда неперечислимо.*

В этом разделе мы привели лишь самые общие факты из теории алгоритмов. Более специальные факты будут приведены в тех параграфах, где они понадобятся.

Все утверждения теории алгоритмов, приводимые в настоящем очерке, приводятся в нём без доказательства. Доказательства можно найти в книге: Н. К. Верещагин, А. Х. Шень. *Вычислимые функции*. Изд. 2-е, испр. М.: МЦНМО, 2002. Что же касается Теоремы Соломонова – Колмогорова, без доказательства приводимой в § 7, то она, к сожалению, в названной книге отсутствует; однако её доказательство может легко быть извлечено из доказательства «Основной теоремы» из § 3 статьи: А. Н. Колмогоров. *Три подхода к определению понятия «Количество информации»* // А. Н. Колмогоров. Избранные труды. Т. 3. Теория информации и теория алгоритмов. М.: Наука, 2005. С. 187–196.

§ 4. Выразимость в языках и окончательная формулировка Теоремы

Возвращаясь к Теореме Гёделя, сформулируем её так:

Пусть дан достаточно богатый формальный язык. Невозможно предъявить такое понятие формального доказательства, чтобы все доказуемые формулы были бы истинными, а все истинные — доказуемыми.

Разумеется, эта формулировка не может считаться окончательной. И термин «понятие формального доказательства» и термин «достаточно богатый язык» слишком расплывчаты. Да и понятие формального языка требует уточнения.

Что касается общего понятия формального языка, то нам достаточно понимать, что каждый язык имеет свой алфавит Б, свой синтаксис (то есть правила образования правильных выражений) и свою семантику (то есть способность выражать какие-то содержательные утверждения). Некоторые слова над Б опознаются как формулы, причём множество формул предполагается разрешимым. В словарном пространстве Б* образуется множество T *истин* (то есть таких формул, которые выражают истинные утверждения), $T \subset$ Б*. Говоря о Теореме Гёделя, традиционно имеют в виду не произвольный формальный язык, а какой-либо язык арифметики, то есть такой язык, в котором есть средства для обозначения натуральных чисел и операций над ними. Мы не будем уклоняться от указанной традиции. Подробнее о выразительных средствах, наличие которых предполагается в языке, будет сказано ниже.

С понятием формального доказательства мы расправимся следующим образом. Сейчас мы заменим его строго определённым математическим объектом, который назовём *дедуктикой*.

Каждое формальное доказательство есть цепочка знаков, то есть слово, в некотором алфавите Д (алфавите доказательств), который может совпадать или не совпадать с Б. Таким образом, в словарном пространстве Д*, где Д — алфавит доказательств, выделяется некоторое множество D формальных доказательств, $D \subset$ Д*. К формальным доказательствам предъявляются два требования — во-первых, чтобы можно было эффективно отличить формальное доказательство от слова из Д*, не являющегося таковым, и, во-вторых, чтобы по формальному доказательству можно было узнать, какое слово из Б* в нём доказывается. Иначе говоря, во-первых, D есть разрешимое множество, и, во-вторых, имеется вычислимая функция доказывания $\delta\colon$ Д$^* \to$ Б*; $\delta(d) = b$ означает, что d есть доказательство для b. Вот всякую такую тройку (Д, D, δ) мы и будем называть *дедуктикой*.

Для каждой дедуктики в Б* возникает множество P *доказуемых* слов:

$$P = \{b \mid \exists d\ \delta(d) = b\}.$$

Теперь приведённую выше формулировку Теоремы Гёделя можно изложить так:

Пусть дан достаточно богатый формализованный язык. Невозможно предъявить такую дедуктику, чтобы все доказуемые формулы были бы истинными, а все истинные — доказуемыми.

Остаётся сообщить, какие языки считаются достаточно богатыми. Точное определение мы отложим до конца параграфа, а на содержательном уровне скажем, что богатство языка понимается нами как его способность *выражать* определённые числовые множества и числовые функции.

К разъяснению того, что значит *выражать*, мы сейчас и перейдём.

Будем предполагать, что в алфавит Б входят буквы 0 и $'$, и это обеспечивает обозначения для натуральных чисел. Выражения $0, 0', 0'', 0'''$ и т. д., служащие именами натуральных чисел, называются *нумералами*. Также будем предполагать, что в алфавит входят знаки сложения и умножения, а синтаксис позволяет образовывать стандартные термы, такие как, например, $(((0'' + 0'''') \cdot 0) + (0''' + 0'''))$.

Предполагается также, что правила образования позволяют конструировать переменные в счётном количестве, например, так: $\bigcirc$, $\bigcirc_{|}$, $\bigcirc_{||}$, $\bigcirc_{|||}$, … — включив предварительно в Б буквы $\bigcirc$ и $_{|}$. Будем, однако, писать, как обычно пишут, x, y, x_{25}, … , подразумевая, что это — всего лишь сокращения для подлинных переменных, каковые суть слова в Б. Областью значений каждой из переменных служит натуральный ряд $\mathbb{N}$.

Педантизм или занудство — как хотите — требует, чтобы было оговорено требование, чтобы в языке имелись способы выражения обычных логических средств: отношения равенства, отрицания (то есть частицы НЕ), союзов И, ИЛИ, ЕСЛИ… ТО, кванторов $\exists$ и $\forall$; вряд ли кто-нибудь сочтёт это требование обременительным, скорее скажут: «А как же иначе?».

Итак, предполагается, что алфавит языка содержит скобки) и (, а также все логические знаки:

$$\neg \quad \& \quad \vee \quad \Rightarrow \quad \exists \quad \forall \quad =$$

Предполагается, что сочетания этих знаков управляются стандартными синтаксическими правилами, которые мы формулировать не будем. Ограничимся примером: если $\mathbf{A}$ формула, а z переменная, то разрешается образовать формулу $\forall z\, \mathbf{A}$.

Формулы бывают *закрытые* и *открытые*. Закрытые формулы (чаще их называют *замкнутыми* формулами) не содержат переменных как параметров:

$$0 = 0', \quad \neg\exists x\, (x' = 0).$$

Семантика предполагается стандартной, поэтому первая из приведённых формул ложна, а вторая истинна. Ещё два примера. Закрытая формула

$$0''''' > 1 \;\&\; \forall y\, \forall z (y \cdot z = 0''''' \;\Rightarrow\; y = 1 \vee z = 1) \qquad (1)$$

выражает утверждение «число 5 — простое» и потому истинна; если же в подслове $0'''''$ (в каждом из двух вхождений этого подслова) прибавить или убавить один штрих, формула превратится в ложную. Закрытая формула

$$\exists u\,(u \cdot 0'' = 0''') \tag{2}$$

выражает утверждение «3 делится на 2».

Естественно требовать, чтобы любая доказуемая формула была истинной. Мы будем из этого исходить.

Примеры открытых формул:

$$x = 0', \quad \neg\exists x\,(z' = y).$$

В первой формуле один параметр, во второй два. Открытые формулы сами по себе не истинны и не ложны, но становятся таковыми, если вместо параметров подставить числа. (Разумеется, это вольность речи. На самом деле число есть абстрактный объект, и подставляем мы не числа, а их имена, то есть нумералы.)

Если формула с параметрами x_1, x_2, …, x_k обозначена через $\mathbf{A}(x_1, x_2, \ldots, x_k)$, то выражение $\mathbf{A}(m_1, m_2, \ldots, m_k)$ обозначает ту формулу, которая получается из первоначальной подстановками каждого из чисел m_i вместо соответствующего параметра x_i.

Открытую формулу с одним параметром x

$$x > 1 \mathbin{\&} \forall y\, \forall z (y \cdot z = x \;\Rightarrow\; y = 1 \vee z = 1) \tag{3}$$

будем сокращённо писать так: $\mathbf{Pr}(x)$. Эта формула *выражает* множество простых чисел. Сказанное означает следующее. Будем вместо x подставлять в неё различные натуральные числа. При каждой такой подстановке она превращается в закрытую формулу, которая оказывается истинной ровно в тех случаях, когда подставляемое число является простым. Например, формулы $\mathbf{Pr}(0'')$ и $\mathbf{Pr}(0''')$ истинны, а формула $\mathbf{Pr}(0')$ ложна.

Как уже отмечалось, закрытая формула $\exists u\ (u \cdot 0'' = 0''')$ выражает утверждение «3 делится на 2». Открытая формула с двумя параметрами — первым y и вторым z

$$\exists u\,(u \cdot z = y) \tag{4}$$

выражает множество таких пар чисел, в которых первый член делится на второй. Это значит, что для таких и только для таких пар чисел (4) становится истинной при подстановке первого члена пары вместо y и второго члена пары вместо z.

Открытая формула с двумя параметрами — первым x и вторым y

$$x \cdot x = y \tag{5}$$

выражает некоторое множество пар чисел. Это множество является графиком функции возведения в квадрат.

Для читателя не составит труда придать точный смысл выражению: «данная формула выражает данное множество троек, четвёрок и так далее натуральных чисел». Например, формула $x = y + z$ при одном из шести упорядочений своих параметров выражает множество таких троек, в которых первый член равен сумме двух других (а при одном из других упорядочений — множество троек, в которых второй член равен сумме двух других). Если параметр x считать третьим, то выражаемое формулой множество окажется графиком функции сложения.

Множество натуральных чисел, или пар натуральных чисел, или троек натуральных чисел, или четвёрок, пятёрок и т. д. — вообще, множество кортежей натуральных чисел фиксированной длины — называется *выразимым* в данном языке, если существует выражающая его формула. Поскольку формул счётное число, а множеств описанного вида несчётно, то большинство множеств невыразимо.

Когда говорят, что формула *выражает* функцию, подразумевают, что формула выражает график этой функции; таким образом, формула, выражающая функцию n переменных содержит $n+1$ параметр. Таким образом, *выразимость функции* означает выразимость её графика.

Вот теперь мы готовы обсудить с уважаемым читателем, что разумно понимать под достаточным богатством языка. Свойство достаточного богатства нужно нам для того, чтобы в рассматриваемом языке проявлялся *эффект Гёделя*, заключающийся в невозможности дедуктики, при котором множество всех доказуемых формул совпало бы с множеством всех истинных формул. В Третьей части будут предложены четыре различные пути доказательства теоремы Гёделя; каждый из этих путей-дорог исходит из своего понимания богатства языка, достаточного для доказательства. Эти понимания весьма близки, потому что каждое из них интерпретирует богатство как выразимость перечислимых множеств и вычислимых функций. Дороги Гёделя и Чейтина опираются на предположение о выразимости любой вычислимой функций одного аргумента. Дорога Шеня опирается на предположение о выразимости любой вычислимой функции двух аргументов. Дорога Колмогорова опирается на предположение о выразимости любого перечислимого множества натуральных чисел.

Замечание. *Если в языке выразима любая вычислимая функция одного аргумента, то выразимо и любое перечислимое множество кортежей натуральных чисел фиксированной длины* (в том числе любое перечислимое множество натуральных чисел), *а, следовательно, и любая вычислимая функция одного или нескольких аргументов*. В самом деле, пустое множество выражается любой формулой, принимающей значение ложь при любом значении своих параметров. Непустое же перечислимое $G \subset \mathbb{N}^k$ располагаем в вычислимую последовательность. Пусть $(\pi_1(n), \pi_2(n), \ldots, \pi_k(n))$ есть n-й член этой последовательности. Функции π_1, π_2, $\ldots$, π_k очевидным образом вычислимы и потому выражаются, соответственно, формулами $\mathbf{P}_1$ с параметрами x и y_1, $\mathbf{P}_2$ с параметрами x и y_2, ..., $\mathbf{P}_k$ с параметрами x и y_k. Тогда множество G выражается следующей формулой с параметрами $y_1, y_2, \ldots, y_k$:

$$\exists x\,(\mathbf{P}_1 \,\&\, \mathbf{P}_2 \,\&\, \ldots \,\&\, \mathbf{P}_k).$$

Мы видим, что используемое в дороге Колмогорова требование выразимости в языке любого перечислимого множества натуральных чисел вытекает из требований, предъявляемых остальными дорогами. Таким образом, оно является наименее ограничительным, и притом его уже достаточно, чтобы Теорема Гёделя имела место! Естественно поэтому дать такое определение: **язык называется достаточно богатым, если в нём выразимо любое перечислимое множество натуральных чисел.** Три другие дороги требуют чуть-чуть более сильного предположения о выразимости любых вычислимых функций одного аргумента; языки, в которых выполняется это предположение, можно было бы назвать *языками с выразимыми функциями.*

Остаётся, однако, главный вопрос — что это за языки, в которых имеет место выразимость перечислимых множеств и вычислимых функций? Может быть, таких нет вовсе или же они крайне неестественны? Они есть и очень естественны. Достаточно, чтобы были выразимы две главные функции: сложение и умножение. А тогда уже все перечислимые множества и все вычислимые функции непременно будут выразимы. Это есть одна из замечательных теорем математической логики, так называемая **Теорема об арифметичности перечислимых множеств**. Доказательство её достаточно длинно, и мы его приводить не будем.

Таким образом, главную техническую часть Теоремы Гёделя составляет Теорема об арифметичности перечислимых множеств. Одно из её доказательств основано на тщательном изучении работы алгоритмов: рассматривается идеальный компьютер и доказывается, что его работа может быть описана в терминах сложения, умножения и логических операций. Другое доказательство — его применял сам Гёдель — состоит в замене вычислимых функций *рекурсивными.* Понятие рекурсивной функции, в отличие от понятия вычислимой функции, не основано на алгоритмической интуиции, а имеет строгое математическое определение. Объёмы обоих понятий тем не менее совпадают.

Исторический комментарий. Первые математические уточнения интуитивных понятий алгоритма и вычислимой функции появились в 1936 г. «А как же рекурсивные функции, которыми пользовался Гёдель шестью годами ранее? — спросит читатель. — Разве их нельзя считать математическим уточнением понятия вычислимой функции?». Всё дело в том, что Гёдель не использовал весь класс рекурсивных функций, который тогда ещё не был открыт; для его целей была достаточна часть этого класса, состоящая из так называемых *примитивно-рекурсивных* функций.

Часть третья: четыре дороги

У всех дорог одна и та же схема: предполагается, что предъявлен формальный язык и предъявлена дедуктика, а затем доказывается невозможность совпадения двух множеств: множества истинных формул и множества доказуемых формул. Отправные пункты дорог различаются: дорога Колмогорова отправляется от достаточного богатства языка (то есть от возможности выразить любое перечислимое множество чисел), остальные дороги отправляются от выразимости функций (то есть от возможности выразить любую вычислимую функцию с натуральными аргументами и значениями).

§ 5. Дорога Гёделя[5)]

Эта дорога основана на парадоксе лжеца, называемого также парадоксом Эпименида. Апостол Павел в послании к Титу (1:12), имея в виду критян, говорит: «Из них же самих один ⟨...⟩ сказал: „Критяне всегда лжецы ⟨...⟩“». По некоторым источникам, сказавшего критянина звали Эпименид. Если считать его слова верными, то получается, что он и сам лжец. Здесь парадокс ещё не возникает. Парадокс возникнет, если предположить, что кроме него критян больше не было, а сам Эпименид ничего другого не сказал. Примерно в такой форме парадокс возник в IV веке до нашей эры у древнегреческого софиста Эвбулида, которому приписывают такое высказывание: «Высказывание, которое я сейчас произношу, ложно». Попытка выяснить, истинно оно или ложно, приводит к противоречию; поэтому, с точки зрения логики, оно вообще не есть высказывание. Основываясь на этой идее, Гёдель нашёл формулу, выражающую недоказуемость самой себя. Легко понять, что эта формула истинна, но не доказуема. В самом деле, если бы она была доказуема, то утверждение, которое она выражает, было бы истинным. Утверждение же это состоит в её недоказуемости. Следовательно, она недоказуема. Значит, формула, выражающая эту недоказуемость (то есть она сама) истинна. Вот мы и

[5)] Проложена в 1930 г. Статья Гёделя с доказательством его великой Теоремы появилась в 1931 г.: Kurt Goedel. *Ueber formal unentscheidbare Saetze der Principia Mathematica und verwandter Systeme* I // Monatshefte fuer Mathematik und Physik, 1931, **38**, S. 173–198 (приблизительный перевод названия: «О формально неразрешимых предложениях системы Principia Mathematica и близких к ней систем»). Однако годом ранее была опубликована её краткая аннотация: Kurt Goedel. *Einige metamathematische Resultate ueber Entscheidungsdefinitheit und Widerspruchsfreiheit* // Anzeiger der Akademie der Wissenschaften in Wien, Mathematisch-naturwissenschaftliche Klasse, 1930, **67**, S. 214–215.

нашли истинную, но недоказуемую формулу, существование которой провозглашает Теорема Гёделя.

Посмотрим, как Гёдель для каждой дедуктики построил такую формулу.

Пусть $\mathbf{A}$ — некоторая формула. Через $\mathbf{A}(r)$ обозначаем результат подстановки в формулу $\mathbf{A}$ числа r вместо параметра x, через $\mathbf{A}(r,s)$ — результат подстановки в $\mathbf{A}$ чисел r и s вместо параметров x и y, через $\mathbf{A}(r,s,t)$ — результат подстановки в $\mathbf{A}$ чисел r, s и t вместо вместо параметров x, y и z. Напомним, что множество всех формул и множество всех формальных доказательств разрешимы, а, следовательно, перечислимы. Поэтому каждое из них можно без повторений расположить в вычислимую последовательность или, как говорят, *перечислить без повторений*.

1. Перечисляем без повторений все формулы:

$$C_1,\ C_2,\ C_3,\ \ldots$$

2. Перечисляем без повторений все формальные доказательства:

$$\Gamma_1,\ \Gamma_2,\ \Gamma_3,\ \ldots$$

3. Рассматриваем функцию g:

$$g(s) = t$$

тогда и только тогда, когда

Γ_s есть доказательство формулы C_t.

Ясно, что функция g вычислима. Поэтому её выражает некоторая формула G с параметрами x и y. Когда, для каких пар чисел формула $G(s,t)$ истинна? Она истинна тогда только тогда, когда доказательство с номером s является доказательством формулы с номером t.

4. Для каждого числа t формула $\exists w\ G(w,t)$ означает доказуемость формулы C_t.

5. Для каждого числа t формула $\neg\exists w\ G(w,t)$ означает недоказуемость формулы C_t.

6. Перечисляем без повторений все открытые формулы с параметром x:

$$A_1,\ A_2,\ A_3,\ \ldots\ .$$

7. Рассматриваем функцию f:

$f(m,n)$ есть номер формулы $A_m(n)$.

Ясно, что f — вычислимая функция. Поэтому её выражает некоторая формула F с параметрами x, y и z. Когда, для каких троек чисел формула $F(m,n,p)$ истинна? Она истинна тогда и только тогда, когда p есть номер формулы $A_m(n)$.

8. Для чисел e, m, n формула

$$\exists u\ [G(e,u)\&F(m,n,u)]$$

означает, что число e есть номер доказательства формулы $A_m(n)$, имеющей номер $f(m,n)$.

9. Для чисел m, n формула

$$\neg\exists w\ \exists u\ [G(w,u)\&F(m,n,u)]$$

означает недоказуемость формулы $A_m(n)$, имеющей номер $f(m,n)$.

10. Для каждого числа n формула

$$\neg\exists w\ \exists u\ [G(w,u)\&F(n,n,u)]$$

означает недоказуемость формулы $A_n(n)$, имеющей номер $f(n,n)$.

11. Теперь рассматриваем формулу

$$\neg\exists w\ \exists u\ [G(w,u)\&F(x,x,u)].$$

с параметром x. Эта формула есть A_q при некотором q.

12. В силу п. 7 формула $A_q(q)$, то есть формула

$$\neg\exists w\ \exists u\ [G(w,u)\&F(q,q,u)],$$

имеет номер $f(q,q)$.

13. В силу п. 10 формула

$$\neg\exists w\ \exists u\ [G(w,u)\&F(q,q,u)],$$

означает недоказуемость формулы с номером $f(q,q)$.

14. Сравнивая пп. 12 и 13, убеждаемся, что присутствующая в них формула (одна и та же!) означает недоказуемость самой себя. Построение закончено.

§ 6. Дорога Колмогорова[6)]

Теорема 1. *Множество P всех доказуемых формул перечислимо.*

Доказательство. Множество P есть $\delta(D)$, где D — множество всех формальных доказательств, а δ — функция выделения доказанного. Множество D разрешимо, а, значит, и перечислимо. Функция δ вычислима. Следовательно, P перечислимо как образ перечислимого множества при частичном отображении, задаваемом вычислимой функцией.

Теорема 2. *Существует выразимое неперечислимое множество натуральных чисел.*

Доказательство. Достаточно взять какое-либо перечислимое множество натуральных чисел с неперечислимым дополнением и рассмотреть

6) Указана в 1952 г.

этот дополнение. Оно и будет искомым множеством, поскольку выражается формулой, служащей отрицанием той формулы, которое выражает выбранное перечислимое множество.

Следствие теоремы 2. *Множество T всех истинных формул неперечислимо.*

Доказательство. Обозначим через K выразимое неперечислимое множество $K \subset \mathbb{N}$, а через $\mathbf{B}$ ту формулу с параметром x, которая его выражает. Обозначим через f вычислимую функцию, относящую каждому числу $n \in \mathbb{N}$ формулу $\mathbf{B}(n)$ (напомним, что это есть результат подстановки в формулу $\mathbf{B}$ числа n вместо параметра x). Очевидно, что

$$K = f^{-1}(T).$$

Если бы множество T было перечислимым, то по перечислимым было бы и K как его полный прообраз.

Сопоставляя теорему 1 и следствие теоремы 2, получаем, что множества P и T не могут совпадать, что и утверждается теоремой Гёделя. В предположении корректности дедуктики получаем существование недоказуемой истинной формулы.

§ 7. Дорога Чейтина[7)]

Эта дорога основана на так называемом *парадоксе Берри*, известном из статьи Бертрана Рассела 1908 г.[8)].

[7)] Проложена в 1974 г.; на сайте `http://www.cs.auckland.ac.nz/CDMTCS/chaitin/unm2.html` Чейтин излагает историю о том, как в начале названного года он уговорил Гёделя назначить ему встречу, чтобы рассказать о своём доказательстве, и как эта встреча не состоялась.

[8)] Bertrand Arthur William Russell (18.05.1872 – 02.02.1970), третий граф Рассел, — знаменитый английский логик, математик, философ и общественный деятель, один из двух авторов того самого трёхтомника *Principia Mathematica*, на который ссылается Гёдель в названии своей статьи 1931 г. Трёхтомник содержал титаническую попытку вывести все математические истины средствами математической логики, то есть опираясь только на точно сформулированные аксиомы и точно сформулированнные правила вывода. Что до статьи Рассела, излагающей парадокс Берри, то вот её библиоописание: Bertrand Russell. *Mathematical Logic as Based on the Theory of Types* // American Journal of Mathematics. Vol. 30, No. 3 (July, 1908), p. 222–262. На с. 223 автор указывает, что парадокс сообщил ему Бéрри (G. G. Berry) [его годы жизни: 1867–1928], один из библиотекарей Бодлианской библиотеки. (Бодлианская библиотека (Bodleian Library) библиотека Оксфордского университета, оспаривающая у Ватиканской право называться старейшей в Европе, а у Британской звание самого крупного книжного собрания Великобритании.) В той же статье, кстати, публикуется и «главный» парадокс математики, *парадокс Рассела*: «**Рассмотрим множество всех таких множеств, которые не содержат самих себя в качестве своего элемента; является ли это множество своим элементом?**».

Парадокс Берри возникает из следующего описания некоторого натурального числа: «Наименьшее натуральное число, которое нельзя описать посредством менее ста слов»[9]. Попытка выяснить, допускает или не допускает само это число описание посредством менее ста слов, приводит к противоречию (поскольку приведённое выше его описание содержит менее ста слов).

Для того, чтобы из этого парадокса возникла теорема Гёделя, надо воспользоваться понятием *колмогоровской сложности* слова. Понятие колмогоровской сложности в начале 60-х годов ввели независимо Рэй Соломонов (Ray Solomonoff, 1926–2009), Андрей Николаевич Колмогоров (1903–1987) и Грегори Чейтин (Gregory Chaitin, 1947–)[10].

Оказывается возможным предъявить число с парадоксальными свойствами: оно имеет сложность гораздо большую, чем та, наличие которой у него можно доказать. Это и есть сложностная версия парадокса Берри. Она допускает и такую формулировку, более близкую к нашему изложению: про числа, имеющие очень большую сложность, это их свойство не может быть доказано.

Переходим к математическому воплощению идеи.

Через Ξ обозначается множество всех двоичных слов, то есть $\{0,1\}^*$. Каждое двоичное слово, начинающееся с единицы, а также однобуквенное слово 0, служит записью натурального числа в двоичной системе счисления; остальные слова, то есть пустое слово и все слова (кроме упомянутого однобуквенного), начинающиеся с ноля, таковыми записями не являются. Для каждого слова $\xi \in \Xi$ обозначаем через $\nu(\xi)$ то натуральное число, для которого ξ служит двоичной записью; если такого числа нет, обозначение не определено.

Напомним, что длина слова ξ обозначается так: $|\xi|$. Легко проверить, что для всех слов ξ, кроме слова 0, выполняется неравенство:

$$|\xi| \;\leqslant\; \log \nu(\xi) + 1, \tag{1}$$

где логарифм берётся по основанию 2.

Напомним также, что все функции у нас не обязательно всюду определены. Пусть E — произвольное множество, f — произвольная функция из Ξ в E. **Сложность** элемента $e \in E$ **относительно** функции f есть, по

---

[9] В статье Рассела речь шла о девятнадцати слогах.

[10] В своей статье «Теорема Гёделя о неполноте в элементарном изложении», опубликованнной в журнале «Успехи математических наук» в 1974 г., автор этих строк ссылался на Григория Цейтина. Предотвращая возможное (и естественное) заблуждение, поспешим указать, что Грегори Чейтин и Григорий Самуилович Цейтин (1936–) являются разными лицами.

определению, число

$$K_f(e) = \min\{|\xi| \mid f(\xi) = e\}. \qquad (2)$$

Как всегда, минимум по пустому множеству есть ∞. (Психологическое пояснение: f рассматривается как функция, дающая объект e по его описанию ξ, и сложность — это наименьшая возможная длина описания. Сложность неописуемого объекта бесконечна.)

Замечание. *Количество элементов, сложность которых не превосходит заданного числа, конечно.*[11)]

Скажем, что функция f **не хуже** функции g, если с точностью до аддитивной константы, не зависящей от e, сложность любого e относительно f не превосходит сложности этого же e относительно g:

$$\exists C\ \forall e\ K_f(e) \leqslant K_g(e) + C. \qquad (3)$$

Теперь в качестве E берём натуральный ряд $\mathbb{N}$. Тогда можно говорить о вычислимых функциях из Ξ в E.

Определение. Вычислимая функция называется **оптимальной**, если она не хуже любой другой вычислимой функции.

Теорема Соломонова – Колмогорова. *Оптимальные функции существуют.*

(Можно показать, что оптимальная функция не может быть всюду определённой.) Фиксируем какую-нибудь оптимальную функцию f и для каждых чисел e и n рассмотрим утверждение

$$K_f(e) > n.$$

В формальном языке это утверждение выражается формулой $\mathbf{I}(n,e)$, получающейся из некоторой формулы $\mathbf{I}$ с двумя параметрами x и y при подстановке вместо этих параметров соответственно чисел n и e. Как найти такую формулу $\mathbf{I}$, скажем позже. Пока что мы исходим из того, что такая формула есть. Выпишем её определяющее свойство:

$$\mathbf{I}(n,e) \text{ истинна тогда и только тогда, когда } K_f(e) > n. \qquad (\clubsuit)$$

Располагаем все доказуемые формулы языка в вычислимую последовательность и фиксируем эту последовательность. Строим вычислимую функцию g из Ξ в E со следующим алгоритмом вычисления. Если ξ не есть двоичная запись натурального числа, оставляем функцию g не определённой на этом ξ. Если же ξ есть запись числа $\nu(\xi)$, то ищем в

[11)] Уместно привести фразу из упомянутой статьи Рассела: “only a finite number of names can be made with a given finite number of syllables”.

последовательности доказуемых формул первую формулу вида

$$\mathbf{I}(\nu(\xi), e) \tag{4}$$

с этим ξ и каким-то e. Если и когда мы такую формулу найдём (чего может и не произойти), то берём e в качестве значения функции g на аргументе ξ.

Таким образом, все натуральные числа разбиваются на два класса — класс тех чисел, для двоичных записей которых функция g определена, и класс тех чисел, для двоичных записей которых функция g не определена. Число n относится к первому классу в том и только в том случае, если для него существует такое e, что формула $\mathbf{I}(n, e)$ доказуема. Наша ближайшая цель — убедиться, что все числа первого класса не превосходят некоторой константы.

Приступаем к построению требуемой оценки. Пусть n принадлежит первому классу. Берём число e, входящее в ту первую — в нашей последовательности доказуемых формул — формулу, которая имеет вид (4). Раз формула $\mathbf{I}$ доказуема, то она истинна. В силу свойства (♣) формулы $\mathbf{I}$ справедливо неравенство

$$K_f(e) \ > \ n. \tag{5}$$

Пусть $n = \nu(\xi)$. Тогда

$$g(\xi) = e. \tag{6}$$

Оценим сверху сложность числа e относительно функции g. Ввиду (6),

$$K_g(e) \ \leqslant \ |\xi|. \tag{7}$$

В сочетании с (1) это даёт

$$K_g(e) \ \leqslant \ \log \nu(\xi) + \mathbf{1}, \tag{8}$$

или, что то же самое

$$K_g(e) \ \leqslant \ \log n + 1, \tag{9}$$

Оценим сверху сложность того же e относительно функции f. Сочетая (3) и (9), получаем:

$$K_f(e) \ \leqslant \ \log n + 1 + C, \tag{10}$$

где C — присутствующая в (3) константа, определяемая сочетанием функций f и g. Соединяя (5) и (9), получаем:

$$n \ < \ \log n + 1 + C. \tag{11}$$

Но начиная с некоторого N (зависящего, разумеется, от C), неравенство (11) не имеет места. Поэтому все числа первого класса меньше этого N.

Теперь мы можем предъявить целую серию истинных, но не доказуемых формул — что и даст теорему Гёделя. Берём произвольное n, такое что $n \geqslant N$. В силу сделанного Замечания, существует такое e, для которого $K_f(e) \ > \ n$. В силу (♣) будет истинной формула $\mathbf{I}(n, e)$. Однако эта формула будет недоказуемой, так как n — число второго класса. Таким образом, хотя при $n \geqslant N$ числа сложности большей, чем n, существуют, ни про одно из них нельзя доказать, что его сложность превосходит n.

Осталось выполнить обещание и показать, как строится формула **I**.

Построение нужной формулы осуществляем в два шага. На первом шагу раскрываем смысл утверждения $K_f(e) > n$ и получаем более подробную его запись. На втором шагу преобразуем эту запись в формулу рассматриваемого языка.

Первый шаг. Неравенство $K_f(e) > n$ равносильно такому утверждению: для всякого ξ, для которого $f(\xi) = e$, справедливо неравенство $|\xi| > n$. Записываем это:

$$\forall\xi\ [f(\xi) = e \Rightarrow\ |\xi| > n]. \tag{12}$$

Второй шаг. Запись (12) ещё не является формулой рассматриваемого формального языка. В ней участвуют двоичные слова, а в формулах языка речь может идти только о числах. Легко перейти от слов к числам. Для этого устроим взаимно однозначное и вычислимое соответствие между Ξ и натуральным рядом и вместо того, чтобы говорить о двоичных словах, будем говорить об их номерах, то есть о соответствующих числах. Соответствие задаётся двумя взаимно обратными вычислимыми функциями: π из $\mathbb{N}$ в Ξ и ρ из Ξ в $\mathbb{N}$:

$$\pi(m) = \xi \quad \text{равносильно} \quad \rho(\xi) = m. \tag{13}$$

Вычислимой функции f из Ξ в $\mathbb{N}$ соответствует вычислимая же функция $\widehat{f}$ из $\mathbb{N}$ в $\mathbb{N}$, аргументами которой вместо слов служат их номера:

$$\widehat{f}(m) = f(\pi(m)); \quad f(\xi) = \widehat{f}(\rho(\xi)). \tag{14}$$

График функции $\widehat{f}$ является перечислимым множеством, а потому выражается некоторой формулой F с параметрами x и y, так что $F(m, e)$ истинна тогда и только тогда, когда $\widehat{f}(m) = e$, т. е. $f(\pi(m)) = e$.

Упражнение. Докажите, что
(а) множество всех пар (ξ, n), для которых $|\xi| > n$, перечислимо;
(б) множество всех пар (k, n), для которых $|\pi(k)| > n$, перечислимо.

Множество, о котором идёт речь в части (б) Упражнения, в силу своей перечислимости, выражается некоторой формулой G с двумя параметрами. Переходя в записи (12) от двоичных слов к их номерам, переписываем эту запись в виде формулы языка:

$$\forall w\ [F(w, e)\ \Rightarrow\ G(w, n)]. \tag{15}$$

В качестве искомой формулы **I** можно взять формулу

$$\forall w\ [F(w, y)\ \Rightarrow\ G(w, x)]. \tag{16}$$

Тогда формула $\mathbf{I}(n, e)$ как раз и будет истинной в точности для тех пар чисел n и e, для которых выполнено неравенство $K_f(e) > n$, что и требовалось.

§ 8. Дорога Шеня[12)]

Эта дорога основана на понятии программы вычислимой функции и на связанной с этим понятием одной из наиболее замечательных теорем теории алгоритмов — Теореме о неподвижной точке. В этой теореме рассматриваются алгоритмы, преобразующие программы вычислимых функций в программы вычислимых функций; предполагается, что алгоритм приводит к результату в применении к любой программе. Теорема о неподвижной точке гласит, что каков бы ни был такой алгоритм, одна из программ непременно перейдёт в эквивалентную ей программу, то есть в программу той же самой функции. Теорема Гёделя получается из теоремы о неподвижной точке следующим образом. Строим такой алгоритм преобразования программ: для каждой программы P алгоритм выдаёт в качестве результата $\mathbf{a}(P)$ первую из таких программ, неэквивалентность которых программе P доказуема в рассматриваемом формальном языке. Если бы этот алгоритм давал результат на всех программах, то в силу Теоремы о неподвижной точки нашлась бы программа P_0 с противоречивыми свойствами: P_0 и $\mathbf{a}(P_0)$ были бы эквивалентны и в то же самое время их неэквивалентность была бы доказуема. Поэтому наш алгоритм не для всех программ приводит к результату. А, значит, найдётся такая замечательная программа Q, что ни для какой программы нельзя доказать её неэквивалентность программе Q. (Легко видеть, между прочим, что всякая столь замечательная программа вычисляет нигде не определённую функцию.) Достаточно теперь взять любую программу, не эквивалентную программе Q, чтобы получить истинное, но не доказуемое утверждение.

Строгое изложение указанной идеи состоит из двух частей. В первой части интуитивное понятие программы заменяется некоторым математическим понятием и Теорема о неподвижной точке приобретает тем самым более точную формулировку. Во второй части уточняется переход от Теоремы о неподвижной точке к Теореме Гёделя.

Первая часть

Говоря о программах вычислимых функций, мы имеем здесь в виду вычислимые функции одной переменной с натуральными аргументами и значениями. Чтобы перейти от интуитивного понятия программы к его точно описанному аналогу, укажем, какие свойства интуитивного понятия нам требуются. Прежде всего, мы предполагаем, что все программы

[12)] Сообщена в 2007 г.

записаны в виде слов в некотором *алфавите программ* и что в соответствующем словарном пространстве они образуют разрешимое множество. Раз множество программ разрешимо, то оно перечислимо и, будучи бесконечным, может быть без повторений расположено в вычислимую последовательность. Другими словами, это множество может быть таким образом занумеровано натуральными числами, что существуют как алгоритм перехода от номера к программе с этим номером, так и алгоритм перехода от программы к её номеру. Обозначим через p_n программу с номером n. Переходя от программ к их номерам, можно так переформулировать Теорему о неподвижной точке: *для всякой всюду определённой вычислимой функции χ из $\mathbb{N}$ в $\mathbb{N}$ найдётся такое число q, что p_q и $p_{\chi(q)}$ эквивалентны, то есть суть программы одной и той же функции.*

Принимается за очевидное наличие алгоритма, который для каждого n и для каждого x даёт результат применения программы p_n ко входу x и ничего не даёт, если результата применения p_n к x не существует. Этот алгоритм задаёт вычислимую функцию F от двух аргументов: n и x. Таким образом, если закрепить в этой функции F некоторое фиксированное число n в качестве первого аргумента, то возникшая функция от второго (переменного) аргумента будет не что иное, как функция, вычисляемая программой p_n. Поэтому для функции F выполняется такое

СВОЙСТВО УНИВЕРСАЛЬНОСТИ: для всякой вычислимой функции f из $\mathbb{N}$ в $\mathbb{N}$ существует такое число n, что для всякого x имеет место условное равенство

$$f(x) \simeq F(n, x).$$

(В качестве n достаточно взять номер программы, вычисляющей f.)

Для функции F выполняется также

СВОЙСТВО ГЛАВНОСТИ: для всякой вычислимой функции G из $\mathbb{N} \times \mathbb{N}$ в $\mathbb{N}$ существует такая всюду определённая вычислимая функция π из $\mathbb{N}$ в $\mathbb{N}$, что для любых n и x имеет место условное равенство

$$G(n, x) \simeq F(\pi(n), x).$$

(В качестве $\pi(n)$ надо взять номер программы, которая осуществляет вычисление функции от x по следующей схеме: «к паре (n, x) примени алгоритм вычисления функции G».)

Заметим, что из второго свойства вытекает первое, но для наглядности мы их разделили. Всякая вычислимая функция, обладающая свойством универсальности, называется *универсальной*. Всякая вычислимая универсальная функция, обладающая к тому же свойством главности, называется *главной универсальной*. Таким образом, описанная выше функция F — одна из главных универсальных функций. Определение главной универсальной функции и отражает нужные нам свойства программ.

Пусть Φ — произвольная числовая функция двух переменных. Число n называется **номером** функции f одной переменной **относительно** функции Φ, коль скоро

$$\forall x\ [f(x) \simeq \Phi(n, x)].$$

Теорема о неподвижной точке приобретает теперь такую точную формулировку:
Пусть Φ — главная универсальная функция. Тогда для всякой всюду определённой вычислимой функции χ из $\mathbb{N}$ в $\mathbb{N}$ найдётся такое число q, что q и $\chi(q)$ суть номера одной и той же функции относительно Φ.

Вторая часть

Фиксируем какую-то функцию Φ двух переменных. Два числа назовём *эквивалентными относительно* Φ, если они являются номерами одной и той же функции. В случае вычислимости функции Φ утверждение об эквивалентности чисел m и n можно выразить в языке, то есть можно указать такую формулу с двумя параметрами, которая тогда и только тогда обращается в истину при подстановке чисел m и n вместо этих параметров, когда числа m и n эквивалентны. Укажем такую формулу. Поскольку Φ вычислима, её график перечислим и выражается некоторой формулой A с параметрами x, y и z, так что для любой тройки чисел $i,\ k,\ l$

$$A(i,k,l) \text{ истинна тогда и только тогда, когда } \Phi(i,k) = l.$$

Эквивалентность чисел m и n означает, что

$$\forall x\ [\Phi(m, x) \simeq \Phi(n, x)],$$

то есть, что для любой пары чисел k и l

$$\Phi(m,k) = l \text{ тогда и только тогда, когда } \Phi(n,k) = l.$$

Последняя равносильность выражается такой формулой:

$$[A(m,k,l) \Rightarrow A(n,k,l)]\ \&\ [A(n,k,l) \Rightarrow A(m,k,l)].$$

А эквивалентность чисел m и n выражается формулой

$$\forall y\ \forall z\ \{[A(m,y,z) \Rightarrow A(n,y,z)]\ \&\ [A(n,y,z) \Rightarrow A(m,y,z)]\}.$$

Обозначим эту формулу через $B(m,n)$.

Итак, формула $B(m,n)$ выражает эквивалентность чисел m и n относительно Φ. Соответственно, неэквивалентность чисел m и n выражается формулой $\neg B(m,n)$.

Выберем теперь в качестве Φ какую-нибудь главную универсальную функцию. Расположим все доказуемые формулы в перечислимую последовательность. Рассмотрим алгоритм, который для каждого натурального

m сперва находит в последовательности доказуемых формул первую формулу вида $\neg B(m, n)$ с этим m и каким-то n, а затем выдаёт это n в качестве результата. Указанный алгоритм преобразования m в n вычисляет некоторую функцию χ. График этой функции перечислим и потому выражается некоторой формулой G.

Покажем, что функция χ определена не для всех натуральных чисел. Убедимся в этом способом от противного. Предположим, что функция χ определена для всех натуральных чисел. Тогда к ней можно применить теорему о неподвижной точке. Значит, существует такое число m_0, что числа m_0 и $\chi(m_0)$ эквивалентны и, стало быть, формула $B(m_0, \chi(m_0))$ истинна. По построению функции χ, формула $\neg B(m_0, \chi(m_0))$ доказуема в формальном языки, значит, тоже истинна. Но формула не может быть истинной вместе со своим отрицанием.

Итак, χ не определена для некоторого натурального q. Это значит, что среди формул вида $\neg B(q, n)$ с этим q и произвольным n нет доказуемых формул. В силу универсальности функции Φ имеется бесконечно много чисел, не эквивалентных числу q. Для любого такого n формула $\neg B(q, n)$ окажется истинной, но не доказуемой. Что и даёт теорему Гёделя.

Упражнение. Как в формальном языке выразить теорему о неподвижной точке? (Проблема состоит в том, чтобы как-то выразить оборот «для всякой всюду определённой вычислимой функции χ одной переменной».)

Добавление

Синтаксическая, или дедуктивная, версия теоремы Гёделя

§ Д-1. Синтаксическая версия теоремы о неполноте

И семантическая, и синтаксическая формулировки теоремы Гёделя начинаются с презумпции, что формализованный язык рассматривается вместе с произвольной, но фиксированной системой формальных доказательств, или, в нашей терминологии, с произвольной, но фиксированной дедуктикой. И когда говорится о доказуемости, подразумевается доказуемость относительно этой дедуктики.

Семантическая версия теоремы Гёделя о неполноте состоит в том, что существует истинная, но не доказуемая формула языка. Этот эффект называется *семантической* неполнотой. Таким образом, семантическая версия опирается на выделение из множества всех формул подмножества истинных формул. Синтаксическая версия не опирается на понятие истины в применении к формулам рассматриваемого языка. Она состоит в том, что существует закрытая формула, которую нельзя ни доказать, ни опровергнуть. (*Опровергнуть* означает доказать отрицание.) Этот эффект называется *синтаксической*, или *дедуктивной*, *неполнотой*. Разумеется, дедуктивная неполнота может иметь место лишь при условии *синтаксической непротиворечивости* (она же *дедуктивная непротиворечивость*) языка, означающей невозможность того, чтобы какая-либо формула была доказуема вместе со своим отрицанием: в противоречивом языке любая формула является доказуемой (в силу законов логики высказываний, каковые предполагаются действующими — см. ниже). Эффект, противоположный синтаксической неполноте и состоящий в том, что любая закрытая формула либо доказуема сама, либо имеет доказуемое отрицание, называется *синтаксической*, или *дедуктивной*, *полнотой*.

В этом Добавлении нас будет интересовать именно синтаксическая неполнота. Поэтому эпитеты «синтаксический» и «дедуктивный» будут опускаться, и мы будем говорить просто «непротиворечивость», «полнота» и т. п.

Окончательный вариант Добавления сложился под влиянием устных и письменных бесед автора с крупнейшим специалистом нашей страны по теории доказательств — Львом Дмитриевичем Беклемишевым.

Используя понятия «непротиворечивость», «полнота» и, вообще, любые понятия, опирающиеся на понятие доказуемости, в применении к *языку*, мы допускаем то, что Бурбаки называет вольностью речи (abus de language). Дело в том, что в логико-математическом обиходе *языком* принято называть чисто синтаксическую конструкцию, позволяющую из всех слов в заданном алфавите выделять правильно построенные выражения (имена, переменные, термы, формулы и т. п.) и указывать правила обращения с ними. Язык, наделённый семантикой, — это уже объект другого рода, который, за неимением лучшего термина, можно называть *интерпретированным языком*. Язык плюс дедуктика — это объект ещё одного рода, который принято называть *теорией*. Только к теориям, строго говоря, применимы понятия, связанные с доказуемостью. Обычно при этом термин «теория» применяют лишь в тех случаях, когда применяемая в рассматриваемом языке система формальных доказательств такова, что вмещает в себя все законы предикатной логики (в том числе законы логики высказываний); именно в таком узком понимании используется указанный термин в настоящем Добавлении. Наконец, если язык снабжён и семантикой, и дедуктикой, его следует, по-видимому, называть *интерпретированной теорией*.

Итак, следует различать четыре категории объектов: языки, интерпретированные языки, теории, интерпретированные теории. Однако, если нет опасения путаницы, всех их можно называть просто *языками*, имея в виду, что точное значение термина *язык* в каждом конкретном случае понятно из контекста. Так, когда говорят о семантической версии теоремы Гёделя, термином *язык* обозначается интерпретированная теория, а когда говорят о синтаксической версии этим термином обозначается теория.

Доказуемость какой-либо формулы **F** принято обозначать так:

$$\vdash \mathbf{F}.$$

Предметом нашего изложения будет язык формальной арифметики. Его синтаксические средства должны быть таковы, чтобы при наделении его стандартной семантикой он мог выражать факты арифметики. Таким образом, среди его выражений имеются нумералы (имена натуральных чисел), аргументами и значениями всех рассматриваемых функций служат натуральные числа, а областью изменения каждой переменной служит натуральный ряд. Мы исходим из того, что язык (как теория) непротиворечив.

В силу законов предикатной логики

$$\text{если нумералы } m \text{ и } n \text{ совпадают, то } \vdash (m = n). \tag{15}$$

Для наглядности при переходе от числа к нумералу, являющегося именем этого числа, мы не будем менять обозначения. Таким образом, если буквы m и n обозначают в тексте какие-нибудь числа, то в составе формул формального языка те же буквы будут обозначать соответствующие этим числам нумералы, то есть символы 0 с m и n штрихами соответственно. Поэтому (15) можно переформулировать так:

$$\text{если числа } m \text{ и } n \text{ совпадают, то } \vdash (m = n). \tag{15$'$}$$

Мы предполагаем, что система формальных доказательств обеспечивает доказуемость некоторых простейших формул. Некоторые из них отметим особо — см. ниже (16)–(18).

$$\text{Если числа } m \text{ и } n \text{ различны, то } \vdash \neg(m = n). \tag{16}$$

$$\text{Следствие: если } \vdash (m = n), \text{ то } m = n. \tag{16$'$}$$

(Неравенство невозможно в силу непротиворечивости.)

Утверждения (15′) и (16′) можно объединить:

$$\vdash (m = n), \text{ тогда и только тогда, когда } m = n. \tag{17}$$

В силу законов формальной арифметики справедливы следующие факты (18)–(20):

$$\vdash (x = y \vee x < y \vee y < x). \tag{18}$$

А поскольку $x \leqslant y$ и $x \geqslant y$ суть не что иное, как сокращения для, соответственно, $x < y \vee x = y$ и $x > y \vee x = y$, то

$$\vdash (x \leqslant y \vee y \geqslant x). \tag{19}$$

Для любой формулы $\mathbf{P}(u)$ и любого числа s,

$$\begin{gathered}\text{если } \vdash \mathbf{P}(0),\ \vdash \mathbf{P}(1), \ldots,\ \vdash \mathbf{P}(s),\\ \text{то } \vdash \forall u[u \leqslant s \Rightarrow \mathbf{P}(u)].\end{gathered} \tag{20}$$

При рассмотрении семантической версии теоремы Гёделя фундаментальную роль играло понятие выразимости. При рассмотрении синтаксической версии оно заменяется понятием верифицируемости. Будем говорить, что теория *верифицирует* функцию $f\colon \mathbb{N} \to \mathbb{N}$ посредством формулы $F(x, y)$ с двумя параметрами, коль скоро выполнены два условия:

(1) для любых чисел m и n если $f(m) = n$, то $\vdash F(m, n)$;

(2) для любых чисел m, n_1 и n_2 если $\vdash F(m, n_1)$ и $\vdash F(m, n_2)$, то $\vdash (n_1 = n_2)$.

Поскольку во всех наших рассуждениях теория предполагается фиксированной, то позволим себе её не упоминать и говорить для краткости, что при выполнении условий (1) и (2) формула $F(x, y)$ *верифицирует* функцию $f\colon \mathbb{N} \to \mathbb{N}$.

Наконец, говорят, что функция f *верифицируема*, если для неё существует верифицирующая её формула.

Вместо центрального для семантической версии требования достаточного богатства языка, состоящего в выразимости перечислимых множеств, для синтаксической версии выдвигается *требование верифицируемости вычислимых функций*. Это есть то основное ограничение на язык, при котором мы собираемся доказать синтаксическую версию.

Естественность указанного требования обосновывается следующими соображениями. Если функция f вычислима, то алгоритмический процесс перехода от аргумента m к результату n может быть прослежен и запротоколирован, и полученный протокол можно трактовать как формальное доказательство равенства $f(m) = n$. Поскольку алгоритм един для всех аргументов, то и полученному протоколу можно придать единообразную для всех аргументов форму, с пустыми местами, заполняемыми конкретными нумералами m и n. Требование верифицируемости функции f состоит по существу в том, чтобы рассматриваемый формализованный язык был в состоянии оформить указанную форму в виде доказуемой формулы с двумя параметрами. В стандартном формальном языке арифметики это оказывается возможным. Сказанное касалось части (1) определения представимости. Что касается части (2), то ясно, что коль скоро предъявлены два протокола одного и того же алгоритма, первый из которых ведёт от аргумента m к результату n_1, а второй — от того же самого аргумента к результату n_2, само это предъявление является наглядным доказательством равенства указанных результатов. Разумеется, указанные соображения имеют всего лишь эвристический характер и должны быть подтверждены строгим содержательным доказательством, каковое оказывается возможным в случае, когда в качестве рассматриваемого формального языка выступает стандартный язык формальной арифметики.

Полезная лемма. *Пусть формула $F(x, y)$ верифицирует функцию f, и пусть числа m и n таковы, что f определена на m и $f(m) \neq n$. Тогда, если язык полон, то $\vdash \neg F(m, n)$.*

Доказательство. Поскольку f определена на m, то для некоторого числа p, отличного от n, будет $f(m) = p$. В силу первой части определения верифицируемости тогда $\vdash F(m, p)$. Если бы было $\vdash F(m, n)$, то в силу второй части определения верифицируемости было бы $\vdash (n = p)$. Но такое невозможно в силу $n \neq p$. Итак, формула $F(m, n)$ недоказуема, а тогда, согласно предположению о полноте, доказуема формула $\neg F(m, n)$.

§ Д-2. Условие омега-непротиворечивости

Сам Гёдель доказал свою теорему в форме синтаксической версии. При этом относительно рассматриваемого формального языка он использовал предположение, более сильное, нежели просто непротиворечивость, а именно предположение об ω-непротиворечивости языка. Гёдель назвал язык *ω-непротиворечивым*, коль скоро ни для какой формулы вида $\exists x\mathbf{F}(x)$, где x — произвольная переменная, не могут одновременно выполняться два утверждения:

(1) $\vdash \exists x\mathbf{F}(x)$;

(2) для любого числа m имеет место $\vdash \neg\mathbf{F}(m)$.

Заметим, что если язык ω-непротиворечив, то он и (просто) непротиворечив, поскольку в противоречивом языке любая формула является доказуемой.

В указанном предположении мы докажем сейчас синтаксическую версию. При этом мы будем использовать не дорогу самого Гёделя, а дорогу Колмогорова. Более точно следовало бы говорить не о самих дорогах полностью, а о началах этих дорог. Именно, мы будем использовать не нумерацию формальных доказательств, с которой начинается дорога Гёделя, а предъявление неразрешимого перечислимого множества, с которого начинается дорога Колмогорова.

Итак, пусть перечислимое множество $R \subset \mathbb{N}$ неразрешимо, что равносильно тому, что его дополнение $\mathbb{N} \setminus R$ неперечислимо. Пусть f — какая-либо всюду определённая на $\mathbb{N}$ вычислимая функция, образом которой служит R. Пусть F — какая-либо верифицирующая её формула.

Доказательство неполноты проводим от противного. Сделаем предположение, что имеет место синтаксическая полнота, то есть, что любая закрытая формула либо доказуема сама, либо имеет доказуемое отрицание. Покажем, что такое предположение приводит к противоречию.

Рассмотрим множество E всех доказуемых формул вида $\neg\exists x F(x,n)$. Оно перечислимо как пересечение двух перечислимых множеств: перечислимого множества всех теорем и разрешимого множества всех формул вида $\neg\exists x F(x,n)$. Множество C, определяемое равенством

$$C = \{n \mid \neg\exists x F(x,n) \in E\},$$

также перечислимо как полный прообраз перечислимого множества E при вычислимом отображении, относящем каждому числу $n \in \mathbb{N}$ формулу $\neg\exists x F(x,n)$. Поэтому C не может совпадать с неперечислимым множеством $\mathbb{N} \setminus R$. А мы сейчас, опираясь на предположение о полноте, как раз и докажем совпадение, что и даст требуемое противоречие.

Для доказательства равенства $C = \mathbb{N} \setminus R$ достаточно доказать два утверждения (1) $C \cap R = \varnothing$ и (2) $\mathbb{N} \setminus R \subset C$.

Докажем (1). Если $n \in R$, то для некоторого числа m имеет место $f(m) = n$. Тогда, в силу определения верифицируемости, $\vdash F(m,n)$. В силу законов логики предикатов, если $\vdash F(m,n)$, то $\vdash \exists x F(x,n)$ (утверждение 1-a). Если же $n \in C$, то $\vdash \neg\exists x F(x,n)$ (утверждение 1-b). Ввиду непротиворечивости языка утверждения (1-a) и (1-b) несовместимы.

Докажем (2). Пусть $n \notin R$. Тогда для каждого числа m справедливо неравенство $f(m) \neq n$. В силу Полезной леммы, $\vdash \neg F(m,n)$. Итак, для каждого числа m имеет место $\vdash \neg F(m,n)$. Условие ω-непротиворечивости запрещает, чтобы при этом было $\vdash \exists x F(x,n)$. Следовательно, в силу предположения полноты, будет $\vdash \neg\exists x F(x,n)$, то есть $n \in C$.

§ Д-3. Неотделимые множества

Говорят, что множество U *отделяет* множество A от множества B, если $A \subset U$ и $B \cap U = \varnothing$. В случае, если все рассматриваемые множества являются подмножеством некоторого универсума, то дополнение к U будет отделять множество B от множества A.

Мы будем рассматривать подмножества натурального ряда $\mathbb{N}$. Два множества называются *отделимыми*, коль скоро существует отделяющее одно от другого разрешимое множество. Если множества не пересекаются и не являются отделимыми, они называются *неотделимыми*. Из теории алгоритмов известен следующий факт: *существуют неотделимые перечислимые множества натуральных чисел.*

Пара неотделимых перечислимых множеств может служить инструментом для доказательства синтаксической версии теоремы Гёделя.

Общая идея такова. Доказательство проводим методом от противного, то есть выдвигаем гипотезу, что язык синтаксически полон. Фиксировав произвольную пару неотделимых перечислимых множеств и предположив синтаксическую полноту языка, для каждого из указанных неотделимых множеств пытаемся найти такое перечислимое надмножество, чтобы эти надмножества оказались взаимно дополнительны. Тогда каждое из них будет, во-первых, разрешимым, а во-вторых, будет отделять исходные неотделимые множества, что невозможно.

Теперь реализуем изложенную идею. Язык по-прежнему предполагается ω-непротиворечивым.

Пусть множества A и B натуральных чисел перечислимы и неотделимы. Пусть f — какая-нибудь определённая на $\mathbb{N}$ вычислимая функция, множеством значений которых служит A. Пусть $F(x, y)$ — формула, верифицирующая функцию f. Таким образом,

$$\text{если } f(m) = n, \text{ то} \vdash F(m, n).$$

В качестве надмножеств, покрывающих, соответственно, A и B, берём такие два множества:

$$F_A := \{n \mid \vdash \exists x F(x, n)\};$$
$$F_B := \{n \mid \vdash \neg\exists x F(x, n)\}.$$

Про эти множества надлежит установить, что они перечислимы, взаимно дополнительны и покрывают, соответственно, A и B.

Перечислимость и — ввиду предположенных непротиворечивости и полноты — взаимная дополнительность очевидны.

Обнаружить, что $A \subset F_A$, очень просто. Пусть $n \in A$. Это означает, что для некоторого числа m имеет место $f(m) = n$. Тогда $\vdash F(m, n)$, откуда, по законам логики предикатов, $\vdash \exists x F(x, n)$.

Убедимся , что $B \subset F_B$. Пусть $n \in B$. Надо показать, что $n \in F_B$. Предположим противное: $n \notin F_B$. Тогда $n \in F_A$. Это значит, что $\vdash \exists x F(x, n)$. В силу ω-непротиворечивости языка невозможно, чтобы для каждого натурального m было $\vdash \neg F(m, n)$. Значит, найдётся такое число m, для которого формула $\neg F(m, n)$ недоказуема. Предположенная полнота языка позволяет применить Полезную лемму и заключить, что в таком случае $f(m) = n$. Но тогда $n \in A$, что невозможно, так как A и B не пересекаются.

§ Д-4. Устранение условия омега-непротиворечивости: дорога Россера.

Через несколько лет после того, как Гёдель опубликовал свою статью, Россер устранил условие омега-непротиворечивости. Он усовершенствовал ту дорогу, идя по которой Гёдель пришёл к своей синтаксической версии, и доказал теорему о неполноте в предположении лишь простой непротиворечивости теории. Способ Россера состоял в построении формулы, выражающей не только — как формула Гёделя — свою недоказуемость, но и ещё кое-что. Слегка огрубляя ситуацию, можно сказать так: Россер нашёл формулу, утверждающую, что если она доказуема, то доказуемо и её отрицание.

В § Д-2 мы действовали колмогоровскими методами, но нам, вслед за Гёделем, пришлось предположить омега-непротиворечивость теории. Основанное на том же предположении другое, слегка более сложное (но тоже колмогоровского типа) доказательство синтаксической неполноты было изложено в § Д-3; это доказательство ровно настолько сложнее доказательства из § Д-2, насколько понятие пары неотделимых перечислимых множеств сложнее понятия неразрешимого перечислимого множества. Сейчас мы приложим идеи Россера к усовершенствованию доказательства из § Д-3 в надежде избавиться от условия омега-непротиворечивости. Таким образом, § Д-3 можно рассматривать как подготовительный по отношению к настоящему § Д-4.

Проанализируем, как мы действовали в § Д-3.

Мы фиксировали произвольную пару $(A,\ B)$ неотделимых перечислимых числовых множеств. Затем мы нашли семейство $\{F_n\}$ закрытых формул со следующими свойствами:

(1) если $n \in A$, то $\vdash F_n$;

(2) если $n \in B$, то $\vdash \neg F_n$ (в демонстрации этого свойства и была использована ω-непротиворечивость).

Далее доказательство синтаксической неполноты проводилось от противного. Предположив полноту, мы получили, что для любого n либо $\vdash F_n$, либо $\vdash \neg F_n$; и то и другое вместе невозможно ввиду простой непротиворечивости. Таким образом, множества $\{n \mid\ \vdash F_n\}$ и $\{n \mid\ \vdash \neg F_n\}$ оказались взаимно дополнительными, а также — поскольку они очевидным образом перечислимы — разрешимыми. Каждое из них отделяет одно из множеств A и B от другого, что невозможно ввиду неотделимости последних.

В § Д-3 в роли F_n выступила формула $\exists x F(x, n)$, где $F(x, y)$ — любая из формул, верифицирующих произвольную вычислимую функции f, перечисляющую множество A. Теперь, поскольку мы хотим довольствоваться не омега-, а простой непротиворечивостью, нам потребуется усложнить формулу F_n.

Как и в предыдущем параграфе, доказательство ведём от противного, то есть предполагаем, что язык синтаксически полон. За обозначениями функции f и формулы $F(x, y)$ сохраняем их прежний смысл — как в § Д-3. Аналогично, g — какая-нибудь определённая на $\mathbb{N}$ вычислимая функция, множеством значений которых служит B, а формула $G(x, y)$ верифицирует функцию g.

Полагаем

$$F_n := \exists x\{F(x, n) \& \forall u[u \leqslant x \Rightarrow \neg G(u, n)]\}.$$

Пусть $n \in A$. Тогда для некоторого числа s будут справедливы два утверждения: во-первых, $f(s) = n$ и, во-вторых, $\forall u(u \leqslant s \Rightarrow g(u) \neq n)$. Из первого утверждения вытекает, что $\vdash F(s, n)$.

Посмотрим, что вытекает из второго. Так как A и B не пересекаются, то для каждого числа q имеет место $g(q) \neq n$ и потому, ввиду предположенной полноты и Полезной леммы, $\vdash \neg(g(q) = n)$. В частности,

$$\vdash \neg G(q,0),\ \vdash \neg G(q,1),\ \ldots,\ \vdash \neg G(q,s).$$

А тогда (см. (6) из § Д-1)

$$\vdash \forall u[u \leqslant s \Rightarrow \neg G(u,n)].$$

В сочетании с ранее доказанным утверждением $\vdash F(s,n)$ получаем

$$\vdash \{F(s,n) \& \forall u[u \leqslant s \Rightarrow \neg G(u,n)]\}.$$

По законам логики предикатов отсюда вытекает, что $\vdash F_n$. Доказательство свойства (1) завершено.

Переходим к доказательству свойства (2). Второй раз используем метод *от противного* (не забудем, что мы находимся внутри большого рассуждения от противного, начавшегося с гипотезы о полноте теории, каковую мы и стремимся опровергнуть). Пусть $n \in B$. Совершенно аналогично только что проведённому рассуждению убеждаемся, что в этом случае $\vdash G_n$, где G_n есть сокращение для формулы

$$\exists y\{G(y,n) \& \forall v[v \leqslant y \Rightarrow \neg F(v,n)]\}.$$

Итак, $\vdash F_n$ и $\vdash G_n$. Для завершения доказательства пункта (2) достаточно убедиться, что это невозможно. К этому мы и приступаем.

Третий раз прибегаем к методу *от противного* и допускаем, что $\vdash F_n$ и $\vdash G_n$. Тогда $\vdash (F_n \& G_n)$. Ложность формулы $(F_n \& G_n)$ на содержательном уровне очевидна. Однако синтаксическая версия потому и называется синтаксической, что не использует семантику языка. Должно продемонстрировать не её ложность, а её недоказуемость. Доказуема же она не может быть ввиду изначально предположенной простой непротиворечивости языка. Дело в том, что доказуемо её отрицание, а именно формула $\neg(F_n \& G_n)$. Демонстрация этого факта дана в Приложении к настоящему параграфу.

Приложение к § Д-4: доказуемость формулы $\neg(F_n \& G_n)$.

Подкванторные части формул F_n и G_n обозначаем, соответственно, $\overline{F_n}$ и $\overline{G_n}$. Таким образом, нам предстоит убедиться, что

$$\vdash \neg(\exists x \overline{F_n} \& \exists y \overline{G_n}). \tag{21}$$

Две формулы $\mathbf{A}$ и $\mathbf{B}$ называются *эквивалентными*, коль скоро доказуемы обе импликации: $\mathbf{A} \Rightarrow \mathbf{B}$ и $\mathbf{B} \Rightarrow \mathbf{A}$, так что эквивалентность сохраняется при переходе от формул к их отрицаниям. Если одна из двух эквивалентных формул доказуема, то, очевидным образом, доказуема и другая.

Один из законов логики предикатов гласит, что если формула $\mathbf{A}$ не содержит параметра y, а формула $\mathbf{B}$ не содержит параметра x, то формулы $(\exists x\mathbf{A} \& \exists y\mathbf{B})$ и

$\exists x \exists y(\mathbf{A} \& \mathbf{B})$ эквивалентны. Поэтому достаточно убедиться в доказуемости формулы

$$\neg \exists x \exists y(\overline{F_n} \& \overline{G_n}), \tag{22}$$

эквивалентной формуле

$$\forall x \forall y \neg(\overline{F_n} \& \overline{G_n}). \tag{23}$$

А чтобы проверить доказуемость формулы (23), достаточно убедиться, что доказуема формула

$$\neg(\overline{F_n} \& \overline{G_n}). \tag{24}$$

Этого потому окажется достаточным, потому что к открытой формуле (22) с параметрами x и y останется лишь дважды применить одно из правил вывода логики предикатов, а именно *правило обобщения*, и получить требуемую доказуемость формулы (23).

Итак, начнём *формально* доказывать формулу (24).

С этой целью подвергнем эту формулу эквивалентным, то есть сохраняющим эквивалентность, преобразованиям. Число n является фиксированным, поэтому для упрощения восприятия формул, отбросим n в выражениях $F(x,n)$, $G(y,n)$, $F(v,n)$, $G(u,n)$ и будем писать сокращённо $F(x)$, $G(y)$, $F(v)$, $G(u)$. Формулу (24) переписываем так:

$$\neg\{F(x) \& \forall u[u \leqslant x \Rightarrow \neg G(u)] \& G(y) \& \forall v[v \leqslant y \Rightarrow \neg F(v)]\}. \tag{25}$$

В (25) перегруппировываем члены:

$$\neg\{(F(x) \& \forall v[v \leqslant y \Rightarrow \neg F(v)]) \& (G(y) \& \forall u[u \leqslant x \Rightarrow \neg G(u)])\}. \tag{26}$$

Занося в (26) отрицание внутрь фигурных скобок, получаем по законам логики высказываний:

$$\{\neg(F(x) \& \forall v[v \leqslant y \Rightarrow \neg F(v)]) \vee \neg(G(y) \& \forall u[u \leqslant x \Rightarrow \neg G(u)])]\}. \tag{27}$$

Если мы формально докажем (27), то тем самым будет формально доказана и (21).

Формула (27) имеет вид $\{K \vee L\}$. По законам исчисления высказываний, если $\vdash \{C \vee D\}$, $\vdash \{C \Rightarrow K\}$, $\vdash \{D \Rightarrow L\}$, то $\vdash \{K \vee L\}$. В нашем случае K это $\neg(F(x) \& \forall v[v \leqslant y \Rightarrow \neg F(v)])$, L это $\neg(G(y) \& \forall u[u \leqslant x \Rightarrow \neg G(u)])$. В качестве C берём формулу $x \leqslant y$, а в качестве D — формулу $y \leqslant x$. Тогда утверждение (19) из § Д-1 можно переписать в виде $\vdash C \vee D$. Таким образом, для наших целей достаточно убедиться в том, что $\vdash \{C \Rightarrow K\}$ и $\vdash \{D \Rightarrow L\}$.

Проверяем утверждение $\vdash \{C \Rightarrow K\}$. (Утверждение $\vdash \{D \Rightarrow L\}$ проверяется совершенно аналогично.) Выписываем цепочку из 8 утверждений о доказуемости формул. Первое утверждение вытекает из законов формальной арифметики, каждое их остальных — из предыдущего, а последнее утверждение есть утверждение о доказуемости формулы $\{C \Rightarrow K\}$, что нам и требуется.

$$\begin{gathered}
\vdash\ [x \leqslant y \& F(x)] \Rightarrow \exists v[v \leqslant y \& F(v))];\\
\vdash\ \neg[x \leqslant y \& F(x)] \vee \exists v[v \leqslant y \& F(v)];\\
\vdash\ \neg(x \leqslant y) \vee \neg F(x) \vee \exists v[v \leqslant y \& F(v)];\\
\vdash\ \neg(x \leqslant y) \vee \neg F(x) \vee \exists v \neg[\neg(v \leqslant y) \vee \neg F(v)];\\
\vdash\ \neg(x \leqslant y) \vee \neg F(x) \vee \exists v \neg[v \leqslant y \Rightarrow \neg F(v)];\\
\vdash\ \neg(x \leqslant y) \vee \neg F(x) \vee \neg\forall v[v \leqslant y \Rightarrow \neg F(v)];\\
\vdash\ \neg(x \leqslant y) \vee \neg\{F(x) \& \forall v[v \leqslant y \Rightarrow \neg F(v)]\};\\
\vdash\ x \leqslant y \Rightarrow \neg\{F(x) \& \forall v[v \leqslant y \Rightarrow \neg F(v)]\}.
\end{gathered}$$

§ Д-5. Устранение условия омега-непротиворечивости: дорога Беклемишева

Отправной точкой по-прежнему служит существование пары неотделимых перечислимых множеств натуральных чисел. Пусть $(A,\ B)$ — такая пара. Рассмотрим функцию f, принимающую значение 0 на A, значение 1 на B и не определённую на остальных натуральных числах. Её график, как легко видеть, перечислим, поэтому она вычислима. Пусть $F(x, y)$ — какая-либо верифицирующая её формула. Положим

$$\begin{gathered}
\overline{A} = \{n \mid \vdash F(n, 0)\},\\
\overline{B} = \{n \mid \vdash \neg F(n, 0)\}.
\end{gathered}$$

Теорию предполагаем непротиворечивой. Доказательство неполноты теории проводим от противного. Предположим, что она полна, и придём к противоречию с неотделимостью множеств A и B.

Множества $\overline{A}$ и $\overline{B}$ очевидным образом перечислимы. В силу непротиворечивости и полноты они взаимно дополнительны, а потому разрешимы. Если мы обнаружим, что (1) $A \subset \overline{A}$ и (2) $B \subset \overline{B}$, то A и B окажутся отделимыми. Утверждение (1) вытекает из первой части определения верифицируемости. Убедимся в справедливости утверждения (2). Если $m \in B$, то $f(m) = 1$. Беря 0 в качестве числа n в Полезной лемме, получаем $\vdash \neg F(m, 0)$, что приводит к (2).

Существенное сокращение доказательства (по сравнению с предыдущим параграфом) достигнуто за счёт того, что была использована не всюду определённая функция.

Наш семинар: математические сюжеты

Проблема тринадцати шаров (элементарный подход)

Х. Маехара

Сколько единичных шаров могут одновременно касаться единичного шара в трехмерном пространстве, не пересекаясь друг с другом? Эта проблема, известная как проблема Ньютона – Грегори или проблема тринадцати шаров, стала предметом дискуссии между И. Ньютоном и Д. Грегори в 1694 г. Ответ следующий.

Теорема 1. *Максимальное число попарно непересекающихся единичных шаров, касающихся данного единичного шара, равно двенадцати.*

Первое доказательство этой теоремы было получено в 1953 г. Шутом и ван дер Варденом [17]. В 1956 г. Лич опубликовал двухстраничную работу [8] с наброском доказательства. Но проследить доказательство по этой работе трудно из-за значительных пробелов. За последние десять лет появились посвященные проблеме тринадцати шаров работы [1–5, 7, 11–16]. Мусин [13] решил аналогичную задачу в четырехмерном пространстве: максимальное число единичных шаров, касающихся четырехмерного единичного шара, равно 24. Он же, используя схожие методы, получил новое решение проблемы 13 шаров [14]. Кроме того, в работе [15] Мусин и Тарасов нашли максимально возможное значение d_{13} минимума сферических расстояний между 13 точками единичной сферы ($d_{13} \approx 57.1367°$) и доказали, что конфигурация, на которой достигается значение d_{13}, единственна с точностью до изометрии.

Перевод А. А. Заславского.

Мы приведем элементарное доказательство теоремы 1, полученное на основе работ [11, 12] и доступное старшеклассникам. Доказательства всех необходимых лемм и формул также будут представлены.

1. Терминология и обозначения

Обозначим через S^2 единичную сферу с центром в начале координат O трехмерного евклидова пространства. *Большим кругом* называется сечение S^2 плоскостью, проходящей через O. *Ребром* (или *отрезком*) называется дуга большого круга, меньшая π. Ребро с концами A, B (а также его длину) будем обозначать AB. Подмножество $W \subset S^2$ называется *выпуклым* если для любых двух точек из W соединяющее их ребро лежит в W. В дальнейшем под любой фигурой (дугой, треугольником, четырехугольником, кругом, окружностью и т. д.) подразумевается фигура на сфере S^2, если не оговорено иное. Треугольник ABC на S^2 — это выпуклая область на S^2, ограниченная тремя отрезками AB, BC, CA. Символ $\triangle(x, y, z)$ обозначает треугольник с длинами сторон x, y, z. *Шапкой* называется область S^2, ограниченная окружностью. если дан треугольник ABC, то $cap(ABC)$ обозначает шапку, ограниченную описанной окружностью ABC и содержащую ABC. *Центр* шапки лежит внутри нее, и сферическое расстояние от него до любой граничной точки называется *радиусом* шапки. Треугольник ABC называется *существенным*, если центр $cap(ABC)$ лежит внутри треугольника ABC. Дуга окружности с концами A, C, проходящая через B обозначается $\widehat{ABC}$. *Четырехугольником* $ABCD$ называется (не обязательно выпуклая) область S^2, ограниченная четырьмя отрезками AB, BC, CD, DA. Площадь треугольника или четырехугольника обозначается $|ABC|$, $|\triangle(x, y, z)|$, $|ABCD|$.

Пусть четырехугольник $ABCD$ является объединением треугольников ABC и ACD, как на рисунке 1. Если D не лежит внутри $cap(ABC)$, то диагональ AC называется *собственной диагональю* $ABCD$.

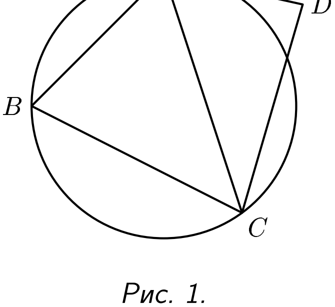


Рис. 1.

2. Основные формулы и леммы

Доказательства следующих формул и лемм приведены в разделе 4.

(2.1) Формула площади. $|ABC| = \angle A + \angle B + \angle C - \pi$.

(2.2) Сферическая теорема косинусов. В $\triangle(x, y, z)$

$$\cos z = \cos x \cos y + \sin x \sin y \cos\theta,$$

где θ — угол $\triangle(x, y, z)$, противоположный z.

Так как $\sin x \sin y > 0$, то $\cos z$ возрастает с ростом $\cos\theta$ при фиксированных x, y. Отсюда следует.

(2.3) При фиксированных x, y угол θ является монотонно возрастающей функцией z. $\square$

(2.4) Лемма Тота [6]. Пусть d — наименьшая сторона ABC. Если радиус $cap(ABC)$ меньше, чем d, то $|ABC| \geqslant |\triangle(d, d, d)|$.

(2.5) Лемма о собственной диагонали [3, 9]. Пусть AC — собственная диагональ четырехугольника $ABCD$. Если $ABCD$ деформируется так, что длины его сторон остаются неизменными, а длина AC уменьшается, то $|ABCD|$ уменьшается.

Пусть ABC — *существенный треугольник*, а треугольник $AB'C$ симметричен ABC относительно плоскости ACO. Тогда AC — существенная диагональ четырехугольника $ABCB'$, являющегося объединением треугольников ABC, $AB'C$. Так как $|ABC| = |ABCB'|/2$, из леммы о собственной диагонали следует, что:

(2.6) При уменьшении стороны x существенного треугольника $\triangle(x, y, z)$ $|\triangle(x, y, z)|$ уменьшается. $\square$

Если P — центр $cap(ABC)$, то общая точка луча $\overrightarrow{OP}$ с плоскостью ABC является центром описанной окружности плоского треугольника ABC. Следовательно, *треугольник ABC существенный тогда и только тогда, когда плоский треугольник ABC не тупоугольный*. Отсюда вытекает

(2.7) Для любых $x, y, z \in [\pi/3, \pi/2]$, $\triangle(x, y, z)$ существенный. $\square$

(2.8) Для любых $x, y \in [\pi/3, 2\pi/3]$, $\triangle(x, y, \pi/2)$ существенный. $\square$

3. Доказательство теоремы 1

Подмножество $X \subset S^2$ назовем $\frac{\pi}{3}$-*отделимым*, если сферическое расстояние между любыми двумя точками X не меньше $\frac{\pi}{3}$. Пусть два единичных шара касаются S^2 в точках P и Q. Из рисунка 2 видно, что эти шары не пересекаются тогда и только тогда, когда $\angle POQ \geqslant \frac{\pi}{3}$. Следовательно теорема 1 равносильна следующей.

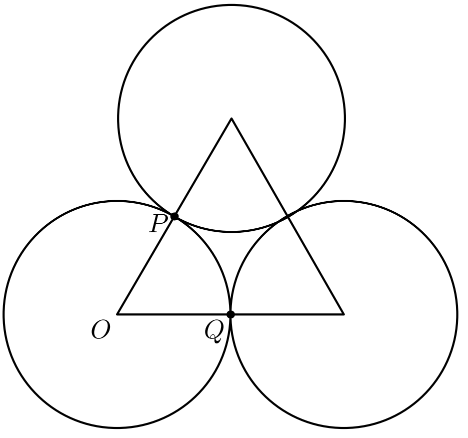


Рис. 2.

Теорема 2. *Максимальная мощность $\frac{\pi}{3}$-отделимого подмножества S^2 равна двенадцати.*

Мы будем доказывать теорему именно в этой формулировке. Обозначим через n максимальную мощность $\frac{\pi}{3}$-отделимого подмножества S^2. Надо доказать, что $n = 12$. Введем следующие символы a, b, δ.

$$a := \pi/3, \ b := \arccos(1/7) \approx 1.427, \ \delta := |\triangle(a, a, a)|.$$

Применяя (2.1) и (2.2), получаем:

$$\delta \approx 0.551, \qquad |\triangle(a, a, b)| \approx 0.667, \qquad |\triangle(a, b, b)| \approx 0.892,$$
$$|\triangle(b, b, b)| \approx 1.194, \quad |\triangle(a, \pi/2, \pi/2)| \approx 1.047, \quad |\triangle(a, a, \pi/2)| \approx 0.679.$$

Впишем в S^2 правильный икосаэдр. Проекции его ребер на S^2 из O разбивают S^2 на 20 равносторонних треугольников площади $4\pi/20 \approx 0.628$. Так как $0.628 > \triangle(\frac{\pi}{3}, \frac{\pi}{3}, \frac{\pi}{3}) \approx 0.551$, стороны этих треугольников больше, чем $\frac{\pi}{3}$. Значит вершины икосаэдра образуют $\frac{\pi}{3}$-отделимое множество, т. е. $n \geqslant 12$.

Пусть теперь $\mathcal{X} \subset S^2$ — $\frac{\pi}{3}$-отделимое множество максимальной мощности n. Выпуклая оболочка $\Gamma(\mathcal{X})$ множества $\mathcal{X}$ является выпуклым многогранником, вершины которого принадлежат $\mathcal{X}$. Точка O лежит внутри $\Gamma(\mathcal{X})$, так как иначе можно было бы добавить к $\mathcal{X}$ точку, не нарушая условия $\frac{\pi}{3}$-отделимости. Проекции ребер $\Gamma(\mathcal{X})$ на S^2 из O, разбивают S^2 на сферические многоугольники. Проведя, если необходимо, диагонали этих многоугольников, мы получим граф G на S^2, разбивающий S^2 на треугольники. Обозначим через t число получившихся треугольников G. Суммарная площадь этих t треугольников равна 4π. Применяя (2.1), получаем

$$4\pi = \text{сумма внутренних углов } t \text{ треугольников} - \pi t = 2\pi n - \pi t.$$

Следовательно

1° $t = 2n - 4$.

Так как O лежит внутри $\Gamma(\mathcal{X})$, а каждая грань $\Gamma(\mathcal{X})$ задает опорную плоскость к $\Gamma(\mathcal{X})$, получаем, что

- 2° Ни одна вершина G не лежит внутри описанной окружности треугольника G.
- 3° Поэтому каждое ребро G является собственной диагональю четырехугольника, полученного объединением двух треугольников, прилежащих к этому ребру.
- 4° Радиус описанной окружности каждого треугольника G меньше a (иначе центр этой окружности можно было бы добавить к $\mathcal{X}$, не нарушая $\frac{\pi}{3}$-отделимости).

По лемме Тота (2.4) и 4°, площадь каждого треугольника G не меньше, чем $\delta = |\triangle(a,a,a)|$. Так как $2n - 4 \leqslant 4\pi/\delta \approx 22.8$, то $n \leqslant 13$. Таким образом, чтобы доказать, что $n = 12$, достаточно доказать, что $n \neq 13$. Будем доказывать это от противного.

Предположим $n = 13$.

Лемма 1. *В этом случае в G есть не больше одного ребра длины $\geqslant b$.*

Доказательство. Из 1° и $n = 13$ следует, что $t = 22$. Пусть наибольшее ребро G — это общее ребро AC треугольников ABC и ACD, а e — следующее по длине ребро (а также его длина) G. Покажем, что $e < b$.

(i) Предположим сначала, что $e > \pi/2$. Деформируем четырехугольник $ABCD$, сохраняя длины его сторон так, чтобы AC стало равно $\frac{\pi}{2}$. Тогда $|ABCD|$ уменьшится по лемме о собственной диагонали (2.5), а так как длины всех ребер не меньше $\frac{\pi}{3}$, то по свойству 4° оба треугольника ABC, ACD будут существенными в силу (2.7) и (2.8). Если e является стороной $ABCD$, то

$$|ABCD| > |\triangle(a,\pi/2,\pi/2)| + \delta \approx 1.0472 + 0.5513$$

и $4\pi > 21\delta + 1.0472 \approx 12.624 > 4\pi$ — противоречие.

Если e не является стороной $ABCD$, то $|ABCD| > 2|\triangle(a,a,\pi/2)| \approx \approx 1.359$. Аналогично сумма площадей двух треугольников с общим ребром e не меньше, чем $2|\triangle(a,a,\pi/2)|$. Тогда $4\pi > (22-4)\delta + 2\times 1.359 \approx \approx 12.642 > 4\pi$ — противоречие. Таким образом, $e \leqslant \pi/2$.

(ii) Теперь предположим, что $b \leqslant e \leqslant \pi/2$. Треугольники, отличные от ABC, ACD, существенные по (2.7). Если e — сторона $ABCD$, то $|ABCD| > |\triangle(a,b,b)| + |\triangle(a,a,b)|$, и, так как существует другой треугольник со стороной e:

$$4\pi \geqslant (22-3)\delta + 2|\triangle(a,a,b)| + |\triangle(a,b,b)| \approx 12.701 > 4\pi,$$

противоречие. Если e не является стороной $ABCD$, то сумма площадей двух треугольников с общим ребром e не меньше, чем $|\triangle(a,a,b)|$, и, значит, $4\pi \geqslant (22-4)\delta + 4|\triangle(a,a,b)| \approx 12.59 > 4\pi$ — противоречие. Следовательно $e < b$. Таким образом в G есть не больше одного ребра с длиной $\geqslant b$.

Лемма 2. *Пусть $\theta = \theta(x,y,z)$ — угол треугольника $\triangle(x,y,z)$, противоположный z. Если $a \leqslant x \leqslant y < b$ и $a \leqslant z$, то $\theta > \pi/3$.*

Доказательство. По (2.3) $\theta(x,y,z) \geqslant \theta(x,y,a)$. Положим $f(x,y,z) = = \cos\theta$. Так как $f_y(x,y,z) = (\cos x - \cos y \cos z)/(\sin^2 y \sin x) > 0$ при $a \leqslant \leqslant x \leqslant y < b$, то $f(x,y,a) < f(x,b,a)$. Поскольку

$$f_x(x,b,a) = \sqrt{3}(2 - 7\cos x)/(24\sin^2 x),$$

максимум $f(x,b,a)$ на интервале $a \leqslant x \leqslant b$ достигается при $x = a$ или $x = b$. Так как $f(a,b,a) = \dfrac{1}{2} > \dfrac{47}{96} = f(b,b,a)$, то $f(x,y,z) < f(a,b,a) = \dfrac{1}{2}$. Значит, $\cos\theta < \dfrac{1}{2}$, т. е. $\theta > \dfrac{\pi}{3}$.

Теперь, если в G нет ребер длины $\geqslant b$, то из леммы 2 следует, что степень каждой вершины G не больше 5. (Величина b была выбрана так, чтобы гарантировать это условие.) Однако G имеет $(22\times 3)/2 = 33$ ребра, так что средняя степень равна $66/13 > 5$ — противоречие. Поэтому в G ровно одно ребро длины $\geqslant b$. Пусть граф G_1 получается из G удалением этого ребра. Тогда степень каждой вершины G_1 не больше 5.

(1) G_1 — плоский граф с 32 ребрами и 21 гранью, одна из которых — четырехугольник, а остальные 20 — треугольники.

Поскольку сумма степеней вершин G_1 равна 64, то

(2) G_1 имеет одну вершину степени 4 и 12 вершин степени 5.

Так как $\triangle(x,y,z)$ является существенным при $x,y,z \in [a,b]$, площадь любого треугольника, образованного ребрами G_1 меньше $|\triangle(b,b,b)|$. Так как $3\delta > |\triangle(b,b,b)|$, внутри такого треугольника не может быть вершин. Значит,

(3) любой треугольник, образованный ребрами G_1, является гранью G_1.

Упражнение. Докажите, что не существует графа, удовлетворяющего условиям (1–3).

Указание. Начните анализ структуры такого графа с четырехугольной грани.

Отсюда следует, что $n \neq 13$, и теорема 2 доказана.

4. Некоторые факты сферической геометрии

4.1. Формула площади и сферическая теорема косинусов

Пусть дана точка $P \in S^2$. Точка $Q \in S^2$ называется *противоположной* P, если PQ — диаметр S^2. Точку, противоположную P, будем обозначать через P^*. Область S^2, ограниченную двумя большими кругами, проходящими через $P \in S^2$ и противоположную точку P^*, будем называть *лункой*. Если внутренний угол при вершине лунки равен θ, то ее площадь равна $4\pi \times \dfrac{\theta}{2\pi} = 2\theta$.

Формула площади. $|ABC| = \angle A + \angle B + \angle C - \pi$.

Доказательство. Пусть A^*, B^*, C^* — точки, противоположные A, B, C, соответственно. Обозначим площади четырех образовавшихся треугольников через Δ, u, v, w (см. рисунок 3). Так как треугольники CA^*B^* и C^*AB

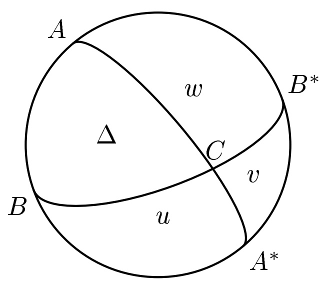


Рис. 3.

симметричны относительно центра O сферы S^2, $\Delta + v$ равно площади $2\angle C$ лунки CAC^*B. Значит $(\Delta + u) + (\Delta + v) + (\Delta + w) = 3\Delta + u + v + w$ равно суммарной площади трех лунок $2\angle A + 2\angle C + 2\angle B$. Поскольку $\Delta + u + v + w$ равно площади 2π полусферы, то $2\Delta + 2\pi = 2(\angle A + \angle B + \angle C)$, откуда и следует искомая формула. $\square$

Сферическая теорема косинусов. Пусть x, y, z — стороны ABC, противоположные вершинам A, B, C, соответственно. Тогда $\cos z = \cos x \cos y + \sin x \sin y \cos \angle C$.

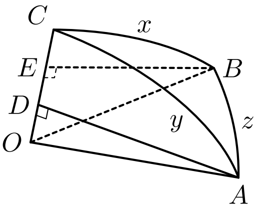


Рис. 4.

Доказательство. Пусть D, E — проекции точек A, B на прямую OC (см. рисунок 4). Угол между векторами $\overrightarrow{DA}$, $\overrightarrow{EB}$ равен $\angle C$. Так как $DA = \sin y$ и $EB = \sin x$, то $\overrightarrow{DA} \cdot \overrightarrow{EB} = \sin x \sin y \cos \angle C$. С другой стороны, поскольку $\overrightarrow{DA} = \overrightarrow{OA} - \cos y\, \overrightarrow{OC}$ и $\overrightarrow{EB} = \overrightarrow{OB} - \cos x\, \overrightarrow{OC}$, получаем, что

$$\overrightarrow{DA} \cdot \overrightarrow{EB} = (\overrightarrow{OA} - \cos y\, \overrightarrow{OC}) \cdot (\overrightarrow{OB} - \cos x\, \overrightarrow{OC}).$$

Так как правая часть равна $\cos z - \cos x \cos y$, то

$$\cos z - \cos x \cos y = \sin x \sin y \cos \angle C. \qquad \square$$

4.2. Теорема Лекселя и лемма Тота

Теорема 3 (Сферическая теорема о вписанном угле [7, 9, 10]). *Рассмотрим треугольник ABC. Пусть P — центр* cap(ABC). *Тогда $\angle C - (\angle A + \angle B) = \pm 2\angle PAB$, со знаком $(-)$ если $\widehat{ACB}$ большая дуга, и $(+)$ в противном случае. Значит,*

$$\angle AXB - (\angle XAB + \angle XBA) \text{ постоянно при } X \in \widehat{ACB}.$$

Доказательство на рисунке 5. (Заметим, что надо еще рассмотреть случай несущественного треугольника.)

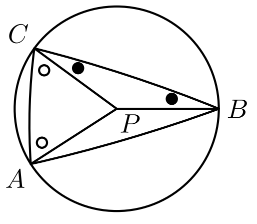


Рис. 5.

Следствие 1. *$\widehat{ACB}$ большая дуга (полуокружность) тогда и только тогда, когда $\angle C < \angle A + \angle B$ ($\angle C = \angle A + \angle B$). В частности, если $\angle C < \frac{\pi}{2}$, то $\widehat{ACB}$ большая дуга.* $\square$

Теорема 4 (Теорема Лекселя). *Дан треугольник ABC. Пусть X — внутренняя точка полусферы, ограниченной большим кругом ABA^*B^* и содержащей C. Если $X \in \widehat{A^*CB^*}$, то $|ABX| = |ABC|$. Если X лежит внутри (вне) cap(A^*CB^*), то $|ABX| > |ABC|$ ($|ABX| < |ABC|$).*

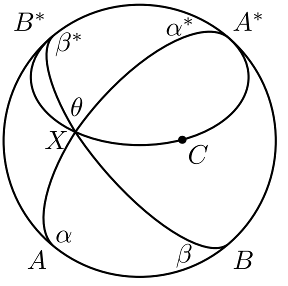


Рис. 6.

Доказательство [10]. На рисунке 6 видно, что, если $X \in \widehat{A^*CB^*}$, то $\theta - -\alpha^* - \beta^*$ постоянно по теореме 3, т. е. $\theta - (\pi - \alpha) - (\pi - \beta)$ постоянно. Значит, $\theta + \alpha + \beta - \pi$ постоянно, что влечет $|ABX| = |ABC|$. Если X лежит на CA^* и $X \neq C$, то, очевидно, $|ABX| > |ABC|$. Поэтому, если X лежит внутри $cap(A^*CB^*)$, то $|ABX| > |ABC|$. Аналогично, если X лежит вне $cap(A^*CB^*)$, то $|ABX| < |ABC|$. □

Лемма 3 (Лемма Тота). *Пусть $d = AB$ — кратчайшая сторона треугольника ABC. Если радиус $cap(ABC)$ меньше d, то*

$$|ABC| \geqslant |\triangle(d, d, d)|.$$

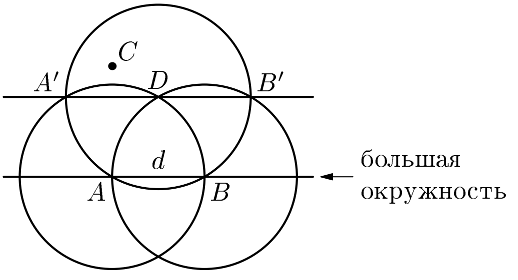


Рис. 7.

Доказательство. Проведем окружности радиуса d с центрами A, B (рисунок 7). Пусть D — точка пересечения этих окружностей, лежащая по ту же сторону от AB, что и C. Проведем окружность радиуса d с центром D. Пусть A', B' — вторые точки ее пересечения с окружностями с центрами B

и A, соответственно. Так как радиус $cap(ABC)$ меньше d, C лежит внутри $cap(ABA')$. Поскольку $|ABD| = |ABA'D|/2 = |ABA'| = |ABB'|$, получаем, что по теореме Лекселя $\widehat{A'DB'} \subset \widehat{A^*DB^*}$. Следовательно, C лежит внутри $cap(A^*DB^*)$, и по теореме Лекселя $|ABC| \geqslant |ABD|$. □

4.3. Изопериметрическая теорема и лемма о собственной диагонали

Лемма 4. *Пусть A, B, C — вершины $\triangle(x, y, z)$, противоположные x, y, z, соответственно. Если $\widehat{ABC}$ большая дуга, то $|\triangle(x, y, z)|$ уменьшается при уменьшении y. Если $\widehat{ABC}$ — полуокружность, то $|\triangle(x, y, z)|$ уменьшается при любом изменении y.*

Доказательство. (i) Предположим, что $\widehat{ABC}$ большая дуга. Тогда, по следствию 1, $\widehat{B^*A^*C}$ тоже большая дуга. Значит, центр $cap(B^*CA^*)$ и A^* лежат по одну сторону от B^*C. Тогда, если двигать C, не меняя A, B и сохраняя x постоянным, то $y = AC$ уменьшается, при выходе C из исходной $cap(A^*CB^*)$ (см. рисунок 8). По теореме Лекселя $|ABC|$ уменьшается.

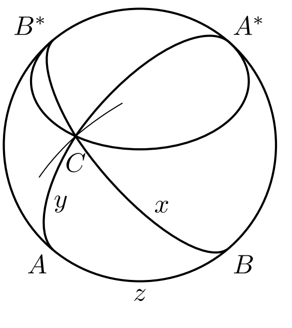


Рис. 8.

(ii) Пусть $\widehat{ABC}$ полуокружность. Тогда $\widehat{B^*A^*C}$ тоже полуокружность, и B^*C — диаметр $cap(B^*CA^*)$. Значит окружность с центром B и радиусом x касается дуги $\widehat{B^*CA^*}$ в точке C. Тогда, если двигать C, не меняя A, B, x, то C выйдет из исходной $cap(A^*CB^*)$. Следовательно, $|ABC|$ уменьшится. □

Следствие 2. *Площадь треугольника ABC с заданными $AB = z$, $BC = x$ (при $x + z < \pi$) максимальна, если $\widehat{ABC}$ полуокружность.* □

Следствие 3. *Площадь выпуклого четырехугольника $ABCD$ с заданными $AB = x$, $BC = y$, $CD = z$ (при $x + y + z < \pi$) максимальна, когда $\widehat{ABD}$, $\widehat{ACD}$ полуокружности.*

Доказательство. Если $\widehat{ABD}$ ($\widehat{ACD}$) не является полуокружностью, то по лемме 4 можно изменить длину AD, оставляя неизменными x, y, z и BD ($|AC|$), так, что $|ABCD|$ увеличится.

Теорема 5 (Изопериметрическая теорема). *Если деформировать вписанный многоугольник, сохраняя длины сторон, то его площадь уменьшится.*

Доказательство. Пусть $ABCD$ выпуклый четырехугольник, вписанный в окружность, x, y, z, w — длины его сторон (см. рисунок 9). Будем считать, что z — наибольшая сторона, а $w \geqslant y$. Пусть AP — диаметр шапки. Тогда

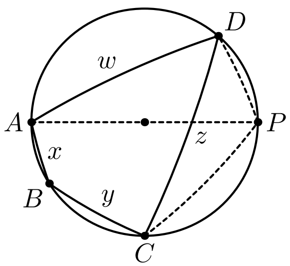


Рис. 9.

AP пересекает CD (возможно, $P = C$), и $w + DP < \pi$, $x + y + CP < \pi$ (так как описанная окружность не является большим кругом). Зафиксировав C, D, P и длины сторон x, y, z, w, деформируем четырехугольник в $A'B'CD$. По следствиям 2 и 3 $|A'PD| < |APD|$ и $|A'B'CP| < |ABCP|$. Следовательно,

$$|A'B'CD| \leqslant |A'B'CP| + |A'DP| - |CPD| < \\ < |ABCP| + |PAD| - |CPD| = |ABCD|.$$

□

Лемма 5. *Пусть $ABCD$ выпуклый четырехугольник, вписанный в окружность, и $AD \leqslant CD$. Если уменьшать AD, не меняя AB, BC, CD, то $|ABCD|$ уменьшается.*

Доказательство. Пусть $A'B'C'D'$ — выпуклый четырехугольник, полученный из $ABCD$ уменьшением AD при неизменных длинах остальных сторон. Зафиксировав вершины A', B', C' и длину $C'D'$, восстановим исходную длину $A'D'$, получив четырехугольник $A'B'C'D''$ (см. рисунок 10). Так как $A'D'' = AD \leqslant CD = C'D''$, $\widehat{A'C'D''}$ — большая дуга. По лемме 4 $|A'C'D''| > |A'C'D'|$, и, значит, $|A'B'C'D''| > |A'B'C'D'|$. Деформировав $A'B'C'D''$ во вписанный четырехугольник с сохранением длин сторон,

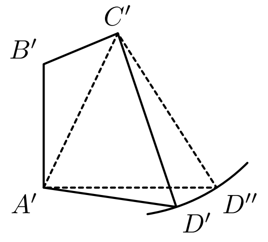


Рис. 10.

получим четырехугольник, равный $ABCD$. По изопериметрической теореме $|ABCD| \geqslant |A'B'C'D''|$, следовательно, $|ABCD| > |A'B'C'D'|$. $\square$

Лемма 6 (Лемма о собственной диагонали). *Пусть AC — собственная диагональ четырехугольника $ABCD$. При деформации, сохраняющей длины сторон $ABCD$ и уменьшающей длину AC, $|ABCD|$ уменьшается.*

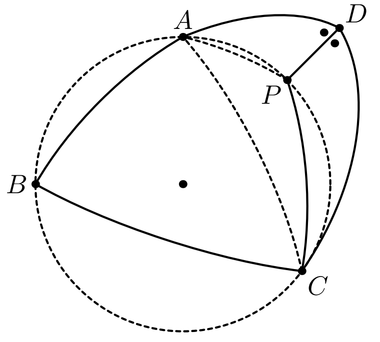


Рис. 11.

Доказательство. Если D лежит на границе $cap(ABC)$, то $|ABCD|$ уменьшается с уменьшением AC по изопериметрической теореме. Поэтому можно считать, что D лежит вне $cap(ABC)$. Пусть P — точка внутри треугольника ACD, лежащая на $\widehat{ACB}$, и такая, что DP — биссектриса $\angle D$. Тогда дуги $\widehat{ADP}$ и $\widehat{CDP}$ большие по следствию 1. Предположим, что $AP \leqslant CP$. Теперь, сохраняя CPD внутри $ABCD$, уменьшим AC. Тогда $\angle ADC$ уменьшается по (2.3). Значит, $\angle ADP$ и AP уменьшаются. Так как $\widehat{ADP}$ большая дуга, то $|ADP|$ уменьшается. С другой стороны, поскольку $AP \leqslant CP$, $|ABCD|$ уменьшается при уменьшении AP по лемме 5. Следовательно, при уменьшении AC $|ABCD|$ уменьшается. $\square$

Список литературы

[1] Anstreicher K. M. *The thirteen spheres: A new proof* // Discrete Comput. Geom. Vol. 31, 2004. P. 613–625.

[2] Bezdek K. *Sphere packing revisited* // European J. Combin. Vol. 27, 2006. P. 864–883.

[3] Böröczky K. *The Newton-Gregory problem revisited* // Discrete geometry (Ed. by A. Bezdek) New York: Marcel Dekker, 2003. P. 103–110.

[4] Böröczky K., Szabó L. *Arrangement of* 13 *points on a sphere* // Discrete geometry (Ed. by A. Bezdek) New York: Marcel Dekker 2003. P. 111–184.

[5] Casselman B. *The difficulty of kissing in three dimensions* // Notices Amer. Math. Soc. Vol. 51, 2004. P. 884–885.

[6] Fejes Tóth L. *On the densest packing of spherical caps* // Amer. Math. Monthly Vol. 56, 1949. P. 330–331.

[7] Hsiang Wu-Yi. *Least action principle of crystal formation of dense packing type and Kepler' conjecture* // Singapore: World Scientific, 2001.

[8] Leech J. *The problem of thirteen spheres* // Math. Gazette. Vol. 40, 1956. P. 22–23.

[9] Maehara H. *Geometry of circles and spheres*. [На японском] Tokyo: Asakura-shoten, 1998.

[10] Maehara H. *Lexell's theorem via an inscribed angle theorem* // Amer. Math. Monthly. Vol. 106, 1999. P. 352–353.

[11] Maehara H. *Isoperimetric theorem for spherical polygons and the problem of* 13 *spheres* // Ryukyu Math. J. Vol. 14, 2001. P. 41–57.

[12] Maehara H. *The problem of thirteen spheres — a proof for undergraduates* // European J. Combin. Vol. 28, 2007. P. 1770–1778.

[13] Musin O. R. *The kissing number in four dimensions* // Ann of Math. Vol. 168, no 1, 2008. P. 1–32.

[14] Musin O. R. *The kissing problem in three dimensions* // Discrete Comput Geom. Vol. 35, 2006. P. 375–384.

[15] Musin O. R., Tarasov A. S. *The strong thirteen sphere problem*. Preprint.

[16] Pfender F., Ziegler G. M. *Kissing numbers, sphere packings, and some unexpected proofs* // Notices Amer. Math. Soc. Vol. 51, 2004. P. 873–883.

[17] Schütte K., Waerden, van der, B. L. *Das Problem der dreizehn Kugeln* // Math. Ann. Bd. 125, 1953. S. 325–334.

Х. Маехара, Research Institute of Educational Development, Tokai University

Взаимное отталкивание примитивных вычетов

В. И. Арнольд

Исследование распределения остатков от деления на натуральное число n, которые взаимно просты с n, среди всех n остатков $\{0, 1, \ldots, n-1\}$, показывает, что взаимно простые с n остатки распределены вдоль конечной окружности $\mathbb{Z}_n$ длины n совсем не так, как были бы распределены независимые случайные точки, расположенные на этой окружности в таком же количестве. А именно, взаимно простые с n остатки отталкивают своих соседей.

Ключевые слова: Стохастичность, число Рейнольдса, распределение Колмогорова, равномерная распределенность, дзета-функция Эйлера, группы Эйлера.

Примитивными вычетами по модулю n я называю взаимно простые с натуральным числом n остатки от деления на n. Например, при $n = 100$ имеется $k = 40$ примитивных вычетов, а именно:

$$x = 1,\ 3,\ 7,\ 9,\ 11,\ 13,\ 17,\ 19,\ 21,\ 23,\ 27,\ 29,\ 31,\ 33,\ 37,\ 39,\ 41,\ 43,\ 47,\ 49,\ 51,\ 53,\ 57,\ 59,\ 61,\ 63,\ 67,\ 69,\ 71,\ 73,\ 77,\ 79,\ 81,\ 83,\ 87,\ 89,\ 91,\ 93,\ 97,\ 99.$$

Ниже обсуждается вопрос, *случайно ли распределены среди n точек конечной окружности длины n, $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, состоящей из всех остатков от деления на n, эти $k(n)$ примитивных вычетов (взаимно простых с n).*

Примеры показывают, что, хотя примитивные вычеты расположены среди всех вычетов довольно равномерно, их распределение статистически сильно отличается от распределения такого же числа $k(n)$ независимых случайных точек на целочисленной окружности $\mathbb{Z}_n$ длины n.

А именно, примитивные вычеты *взаимно отталкиваются*, малые расстояния между соседними примитивными вычетами встречаются гораздо реже, чем они встречаются для случайных выборок такого же числа $k(n)$ независимых точек в $\mathbb{Z}_n$.

Рукопись В. И. Арнольда отредактирована С. В. Конягиным. Редколлегия сборника «Математическое просвещение» выражает глубокую благодарность С. В. Конягину за помощь в подготовке к печати этой работы.

§ 1. Вычисление безразмерного параметра случайности β

Для k точек на целочисленной окружности длины n обозначим через $(a_1, a_2, \ldots, a_k)$ длины дуг, на которые эти точки делят окружность, т. е. расстояния между соседними точками. Таким образом

$$a_1 + a_2 + \ldots + a_k = n.$$

Пример. Для сорока примитивных вычетов по модулю $n = 100$ предыдущий список доставляет сорок последовательных дуг с длинами

$$2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,$$
$$4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2,\ 2,\ 4,\ 2,\ 2$$

(расстояние между остатками 99 и 1 на окружности $\mathbb{Z}_{100}$ равно 2).

Рассмотрим сумму квадратов длин всех этих k дуг, на которые делят целочисленную окружность длины n примитивные вычеты по модулю n:

$$B = a_1^2 + a_2^2 + \ldots + a_k^2.$$

Эта сумма квадратов сильно зависит от распределения изучаемых k точек на окружности $\mathbb{Z}_n$.

В предыдущем примере (где $n = 100$) мы замечаем десять троек последовательных дуг длины 2, дающие в сумму квадратов длин дуг вклад

$$10(4 + 4 + 4) = 120,$$

и еще десять дуг длины 4, дающие вклад

$$10 \cdot 16 = 160.$$

Таким образом, при $n = 100$ мы вычислили значение

$$B(100) = 120 + 160 = 280.$$

Можно доказать (см., например, [1]), что значение B всегда заключено между наименьшим значением

$$B_0 = \frac{n^2}{k}$$

и наибольшим значением (превосходящим наименьшее в k раз),

$$B_1 = n^2.$$

Чтобы определить безразмерный аналог «числа Рейнольдса» (позволяющий сравнивать распределения k точек на окружностях $\mathbb{Z}_n$ разных длин n), я предложил в [1] рассматривать «безразмерный параметр стохастичности»

$$\beta = B/B_0 = B/(n^2/k).$$

В рассмотренном выше примере ($n = 100$, $k = 40$) получаются такие значения:

$$B_0 = \frac{100^2}{40} = 250, \quad \beta = \frac{B}{250} = 1{,}120.$$

Для случайных k точек, независимо набросанных на окружность длины n, в [1] вычислены средние значения параметров стохастичности,

$$\widehat{B} = \hat{\beta} B_0$$

(оказалось, что $\hat{\beta} \approx 2$ при больших k).

А именно, там доказано, что математическое ожидание значения безразмерного параметра стохастичности составляет

$$\hat{\beta} = \frac{2k}{k+1}\,.$$

В приведенном выше примере значение параметра $\beta = 1{,}12$ гораздо меньше среднего значения $\hat{\beta} \approx 1{,}951$ (из-за того, что примитивные вычеты далеко не независимы, а «отталкиваются» от своих соседей).

Если для набора k точек в $\mathbb{Z}_n$ наблюдается значительно меньшее, чем $\hat{\beta}$ (т. е. чем 2) значение безразмерного параметра стохастичности β, то это свидетельствует об отталкивании соседних точек. А если наблюдается значительно большее, чем $\hat{\beta}$ (т. е. чем 2) значение β, то это свидетельствует об их взаимном притяжении.

Мы увидим сейчас, что для примитивных вычетов получаются значительно меньшее, чем 2, значения безразмерного параметра стохастичности β.

В следующих таблицах приведены значения параметров (n, k, B, β) для примитивных остатков от деления на n целых чисел, $2 \leqslant n \leqslant 31$.

Число α_a в этих таблицах означает число дуг длины a (при $n = 100$ вышеприведенный список доставляет

$$\alpha_2 = 30, \quad \alpha_4 = 10.)$$

Всегда выполняются соотношения

$$\sum \alpha_a = k, \quad \sum a\alpha_a = n, \quad \sum a^2\alpha_a = B.$$

Еще Эйлер заметил (в своей работе о дзета-функции), что в среднем

$$\frac{k}{n} \approx \frac{6}{\pi^2} = \frac{1}{\zeta(2)} \approx 0{,}608$$

(при больших n), а также, что для простого $n = p$ выполнены соотношения

$$k(p) = p - 1, \quad k(p^\omega) = (p-1)p^{\omega-1},$$

и что «φ-функция» $k = k(n)$ мультипликативна:

$$k(uv) = k(u)k(v),$$

если u и v взаимно просты (обычное в теории чисел обозначение этой мультипликативной функции k аргумента n есть $\varphi(n)$).

Вычисление указанных в таблицах значений повторяет приведенное выше в случае $n = 100$ прямое перечисление примитивных вычетов, но оно проще, так как значения k меньше. Например, примитивные вычеты по модулям $n = 2, \ldots, 12$ образуют такие картинки

$$
\begin{array}{ll}
n=2\colon & |-\overset{1}{\circ}-|;\\
n=3\colon & |-\overset{1}{\circ}-\overset{2}{\circ}-|;\\
n=4\colon & |-\overset{1}{\circ}--\overset{3}{\circ}-|;\\
n=5\colon & |-\overset{1}{\circ}-\overset{2}{\circ}-\overset{3}{\circ}-\overset{4}{\circ}-|;\\
n=6\colon & |-\overset{1}{\circ}----\overset{5}{\circ}-|;\\
n=7\colon & |-\overset{1}{\circ}-\overset{2}{\circ}-\overset{3}{\circ}-\overset{4}{\circ}-\overset{5}{\circ}-\overset{6}{\circ}-|;\\
n=8\colon & |-\overset{1}{\circ}--\overset{3}{\circ}--\overset{5}{\circ}--\overset{7}{\circ}-|;\\
n=9\colon & |-\overset{1}{\circ}-\overset{2}{\circ}--\overset{4}{\circ}-\overset{5}{\circ}--\overset{7}{\circ}-\overset{8}{\circ}-|;\\
n=10\colon & |-\overset{1}{\circ}--\overset{3}{\circ}----\overset{7}{\circ}--\overset{9}{\circ}-|;\\
n=11\colon & |-\overset{1}{\circ}-\overset{2}{\circ}-\overset{3}{\circ}-\overset{4}{\circ}-\overset{5}{\circ}-\overset{6}{\circ}-\overset{7}{\circ}-\overset{8}{\circ}-\overset{9}{\circ}-\overset{10}{\circ}-|;\\
n=12\colon & |-\overset{1}{\circ}----\overset{5}{\circ}--\overset{7}{\circ}----\overset{11}{\circ}-|.
\end{array}
$$

Подобные рисунки доставляют следующие значения параметров стохастичности систем примитивных вычетов:

n	2	3	4	5	6	7	8	9	10	11
k	1	2	2	4	2	6	4	6	4	10
B	4	5	8	7	20	9	16	15	28	13
B_0	4	4,5	8	6,25	18	8,167	16	13,5	25	12,1
β	1	1,111	1	1,120	1,111	1,102	1	1,111	1,120	1,074
α_1	0	1	0	3	0	5	0	3	0	9
α_2	1	1	2	1	1	1	4	3	3	1
α_3	0	0	0	0	0	0	0	0	0	0
α_4	0	0	0	0	1	0	0	0	1	0
Δ	2	2	2	2	4	2	2	2	4	2

Величина Δ в этой таблице — длина наиболее длинной из дуг, на которые примитивные вычеты делят окружность $\mathbb{Z}_n$.

При бóльших значениях n встречаются и более длинные дуги.

n	12	13	14	15	16	17	18	19	20	21
k	4	12	6	8	8	16	6	18	8	12
B	40	15	36	33	32	19	60	21	56	43
B_0	36	14,1	32,67	28,125	32	18,06	54	20,05	50	36,75
β	1,111	1,065	1,102	1,173	1	1,052	1,111	1,047	1,12	1,17
α_1	0	11	0	3	0	15	0	17	0	5
α_2	2	1	5	3	8	1	3	1	6	5
α_3	0	0	0	2	0	0	0	0	0	2
α_4	0	0	1	0	0	0	3	0	2	0
Δ	4	2	4	3	2	2	4	2	4	3

Третий десяток значений n приводит к следующим значениям параметров стохастичности:

n	22	23	24	25	26	27	28	29	30	31
k	10	22	8	20	12	18	12	28	8	30
B	52	25	80	35	60	45	72	31	132	33
B_0	48,4	24,04	72	31,25	56,333	40,5	65,333	30,036	112,5	32,033
β	1,074	1,059	1,111	1,12	1,065	1,111	1,102	1,032	1,173	1,030
α_1	0	21	0	15	0	9	0	27	0	29
α_2	9	1	4	5	11	9	10	1	3	1
α_3	0	0	0	0	0	0	0	0	0	0
α_4	1	0	4	0	1	0	2	0	3	0
α_5	0	0	0	0	0	0	0	0	0	0
α_6	0	0	0	0	0	0	0	0	2	0
Δ	4	2	4	2	4	2	4	2	6	2

Сопоставляя найденные выше значения безразмерного параметра стохастичности β, мы замечаем, что, хотя все они гораздо меньше критического значения 2, они заметно осциллируют при изменении числа n.

Чтобы сгладить эти осцилляции и выявить систематическое поведение чисел $\beta(m)$, составим их чезаровские средние

$$\hat{\beta}(n) = \left(\sum_{m=2}^{n} \beta(m)\right) \Big/ \left(\sum_{m=2}^{n} 1\right) = \sum_{m=2}^{n} \beta(m)/(n-1).$$

В следующих таблицах приведены значения сумм

$$\Sigma(n) = \sum_{m=2}^{n} \beta(m)$$

и значения чезаровских средних

$$\hat{\beta}(n) = \Sigma(n)/(n-1) :$$

n	2	3	4	5	6	7	8	9	10	11
$\Sigma(n)$	1,00	2,111	3,111	4,231	5,342	6,444	7,444	8,565	9,675	10,749
$\hat{\beta}(n)$	1,00	1,056	1,037	1,058	1,068	1,074	1,063	1,069	1,075	1,075

Второй десяток значений n доставляет еще более медленный рост чезаровских средних $\hat{\beta}(n)$:

n	12	13	14	15	16	17	18	19	20	21
$\Sigma(n)$	11,860	12,925	14,027	15,20	16,20	17,252	18,363	19,5	20,62	21,79
$\hat{\beta}(n)$	1,075	1,077	1,079	1,086	1,080	1,078	1,080	1,083	1,085	1,089

Третий десяток значений n почти прекращает рост значений чезаровских средних $\hat{\beta}(n)$:

n	22	23	24	25	26	27	28	29	30	31
$\Sigma(n)$	22,864	23,929	25,034	26,154	27,219	28,33	29,432	30,464	31,537	32,567
$\hat{\beta}(n)$	1,089	1,087	1,088	1,089	1,089	1,090	1,090	1,088	1,089	1,086

В пределах первого десятка значений ($2 \leqslant n \leqslant 11$) величина чезаровского среднего $\hat{\beta}(n)$ возрастает от 1 до 1,075, т. е. на 0,075. В пределах второго десятка значений ($12 \leqslant n \leqslant 21$) величина чезаровского среднего $\hat{\beta}(n)$ возрастает на 0,011, т. е. примерно в семь раз меньше, чем в предыдущем десятке. При дальнейшем росте n (в пределах $22 \leqslant n \leqslant 31$) прирост значения чезаровского среднего $\hat{\beta}(n)$ мало заметен.

Все это позволяет предположить, что чезаровское среднее $\hat{\beta}(n)$ остается и при больших значениях n ограниченным (и даже недалеким от 1).

Чтобы эмпирически проверить эту гипотезу, я вычислил значения параметров стохастичности для $n = 100$ и для окрестности этого значения, $96 \leqslant n \leqslant 104$. Вот полученные 9 значений безразмерного параметра стохастичности β:

n	96	97	98	99	100	101	102	103	104
k	32	96	42	60	40	100	32	102	48
B	320	99	252	189	280	103	372	105	290
B_0	288	$96\frac{1}{96}$	$228\frac{2}{3}$	163,35	250	102,01	$325\frac{1}{8}$	104,09	$225\frac{1}{3}$
β	1,111	1,010	1,102	1,157	1,120	1,0097	1,144	1,0087	1,065
Δ	4	2	4	3	4	2	6	2	4
σ	1,111	2,121	3,223	4,380	5,500	6,510	7,654	8,663	9,728

Здесь $\sigma(n) = \sum_{m=96}^{n} \beta(m)$. Поэтому среднее значение $\hat{\hat{\beta}}$ параметра β по указанной девятке значений n (вблизи $n = 100$) составляет

$$\hat{\hat{\beta}}(100) = \frac{\sigma(104)}{9} \approx 1{,}081.$$

Мы видим, что средние значения безразмерного показателя стохастичности $\beta(n)$ (системы примитивных вычетов по модулю n) мало отличаются друг от друга при $n \approx 20$ и при $n \approx 100$. Это подтверждает высказанную выше гипотезу о стабилизации $\hat{\beta}(n)$ при $n \to \infty$ (утверждающую, что значения $\hat{\beta}(n)$ не только остаются ограниченными при $n \to \infty$, но что они приближаются к стабильному значению $\beta^* \approx 1{,}08$).

Однако даже такая стабилизация чезаровских средних не означает еще стремления к пределу самих значений показателя стохастичности $\beta(n)$: они могут сильно отличаться от средних $\hat{\beta}(n)$ столь редко, что эти различия не разрушат стабилизации $\hat{\beta}(n) \to \beta^*$.

В [7] показано, что величины $\beta(n)$ ограничены и, более того, $\beta(n) \to 2$, где предел берется по любой подпоследовательности значений n, для которых $k(n)/n \to 0$. По-видимому, используя технику статьи [7], можно доказать, что $\beta(n) < 2$ и $\hat{\beta}(n)$ близко к 1 при всех n, но это потребует больших компьютерных вычислений.

§ 2. Параметр стохастичности Колмогорова λ

А. Н. Колмогоров [6] ввел свой параметр стохастичности λ в работе 1933 г., опубликованной по-итальянски в журнале страховой статистики. Он использовал этот параметр для решения практически важного вопроса, правдоподобно ли утверждение о случайном происхождении данной числовой последовательности $\{x_1, x_2, \ldots\}$.

Для k вещественных чисел x_i $(1 \leqslant i \leqslant k)$ Колмогоров определил значение своего параметра λ следующей конструкцией. Обозначим через f_k «эмпирическую считающую функцию» данного набора k точек, определенную условием

$$f_k(x) = (\text{число не превосходящих } x \text{ значений } x_i).$$

Колмогоров обсуждает вопрос о правдоподобности гипотезы, что $\{x_i\}$ суть k независимых значений случайной величины x, распределенной на вещественной оси с данной непрерывной функцией распределения F:

$$F(X) = (\text{вероятность события } x \leqslant X).$$

Составим «теоретическую считающую функцию»

$$F_k(X) = kF(X)$$

(это значение есть математическое ожидание числа меньших или равных X значений x_i в выборке из k независимых значений с данным распределением F).

Для вычисления своей характеристики λ множества k вещественных чисел x_i Колмогоров упорядочивает их (так, чтобы $x_1 \leqslant x_2 \leqslant \ldots x_k$).

Если все k значений x_i различны, то ступенчатая считающая функция f_k равна 0 в точках $x < x_1$, 1 в точках полуинтервала $x_1 \leqslant x < x_2$, 2 в точках полуинтервала $x_2 \leqslant x < x_3$ и т. д., до значения k в точках $x \geqslant x_k$. Считающая функция f_k множества $k = 4$ взаимно простых с числом $n = 10$ вычетов по модулю 10 имеет график, изображенный на рис. 1.

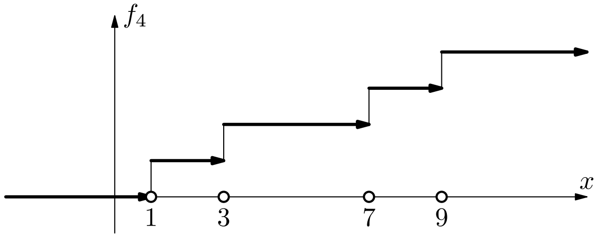


Рис. 1.

Если же в наборе $\{x_i\}$ есть повторения, то величина $f_k(z+0)-f_k(z-0)$ скачка в точке z равна числу совпавших с z точек изучаемого набора (т. е. числу тех i, для которых $x_i = z$).

Основной шаг в конструкции Колмогорова — исследование отличия эмпирической считающей функции f_k от теоретической считающей функции F_k.

Определение 1. *Отклонение Колмогорова* S_k определяется (для k точек $\{x_i\}$) как

$$S_k = \sup_{x \in \mathbb{R}} |f_k(x) - F_k(x)|.$$

Определение 2. *Параметр Колмогорова* λ_k имеет (для данного набора k точек $\{x_i\}$) значение

$$\lambda_k = S_k/\sqrt{k}.$$

Замечание. Деление на квадратный корень из k позволяет сравнивать между собою (по значению параметра стохастичности λ) наборы из разного числа точек $\{x_i\}$.

Дело в том, что разность $f_k - F_k$ можно представить как сумму k независимых слагаемых порядка 1 с математическим ожиданием ноль для каждого слагаемого (считающих по 1 точке в каждом слагаемом). По закону больших чисел типичные значения этой суммы имеют порядок величины $\sqrt{k}$ (и они велики при больших k).

Деление на $\sqrt{k}$ уничтожает это масштабирующее влияние числа k изучаемых точек — велико или мало значение λ_k определяется не тем, велико или мало число k точек x_i, а тем, насколько они случайны.

Колмогоров в [6] доказал, что для k независимых случайных величин x_i с одинаковой непрерывной функцией распределения F получаемые из разных выборок k точек $\{x_i\}$ значения параметра стохастичности λ_k имеют (на положительной полуоси $\Lambda \geqslant 0$) универсальную функцию распределения Φ_k (не зависящую от F_k), а при $k \to \infty$ эти функции стремятся к универсальной функции распределения Колмогорова

$$\Phi(\Lambda) = \sum_{s=-\infty}^{+\infty} \Big((-1)^s e^{-2s^2\Lambda^2} \Big).$$

Здесь функции распределения Φ_k определяются условиями

$$\Phi_k(\Lambda) = (\text{вероятность события } \ \lambda_k \leqslant \Lambda).$$

Сходимость $\Phi_k \to \Phi$ (при $k \to \infty$) — равномерная (на полуоси $\Lambda \geqslant 0$).

Функция Φ монотонно растет от $\Phi(0) = 0$ (где равны нулю и все ее производные) до $\Phi(\infty) = 1$.

Среднее значение так распределенной величины Λ равно

$$\widehat{\Lambda} = \sqrt{\pi/2}\ln 2 \approx 0{,}869.$$

Вдали от среднего значения $\widehat{\Lambda}$ функция Φ быстро стремится к 0 (слева) и к 1 (справа). Например,

$$\Phi(0{,}4) \approx 0{,}003, \quad \Phi(1{,}8) \approx 0{,}997.$$

Поэтому как слишком малые значения параметра Колмогорова λ_k данного набора $\{x_i\}$, так и слишком большие значения указывают на малое правдоподобие стохастичности изучаемого набора k чисел $\{x_i\}$.

Вот значения функции Φ в некоторых точках Λ:

Λ	0,4	0,6	0,8	1,0	1,2	1,4	1,6	1,8	2,0
$10^4\Phi$	28	1357	4558	7300	8877	9603	9888	9969	9993

Медианное значение Λ^* параметра стохастичности Колмогорова (для которого как меньшие, так и большие его значения параметра стохастичности в распределении Колмогорова Φ имеют одинаковую вероятность 1/2) составляет $\Lambda^* \approx 0{,}83$.

Описанная выше теорема Колмогорова [6] (об универсальном распределении Φ) доказана им в предположении непрерывности функции распределения F независимых вещественных величин x_i (релятивистское соображение состоит в том, что величины с разными такими распределениями превращаются друг в друга заменами координат на оси x). Поэтому не зависящие от выбора координаты на оси x величины (вроде колмогоровского отклонения S и параметра Колмогорова λ) ведут себя для любого непрерывного распределения так же, как для равномерного распределения вдоль некоторого отрезка (где все вычисления сводятся к суммированию объемов симплексов, биномиальным коэффициентам и формуле Стирлинга для их асимптотик).

Я же собираюсь (незаконно) применить эту теорию к наборам точек на конечных окружностях $\mathbb{Z}_n$, к которым она формально не относится.

Прежде всего, вычеты образуют дискретное множество, так что функция распределения никак не может быть непрерывной. Я надеюсь, однако, что при больших n конечная окружность $\mathbb{Z}_n$ приближается к вещественной окружности S^1, и что естественно-научные факты, доказанные для предельного случая непрерывного множества значений x_i, дают хорошее приближение к положению дел для значений, принадлежащих к дискретным множествам значений, аппроксимирующим непрерывные.

Другая трудность состоит в том, что конечная окружность $\mathbb{Z}_n$ аппроксимирует не прямую $\mathbb{R}$, к которой относится теория Колмогорова, а гладкую окружность S^1. Беда здесь в топологическом различии: одномерная группа когомологий окружности нетривиальна и заданные на ней распределения обычно не имеют поэтому однозначных функций распределения.

Ведь функция распределения — это первообразная его плотности, так что это многозначная функция на накрывающей окружность прямой при полном обходе окружности получает приращение (равное массе изучаемого распределения на окружности).

Чтобы преодолеть эту топологическую трудность, я использовал следующую конструкцию.

Рассмотрим на окружности два распределения одинаковой суммарной массы (в нашем случае одно из распределений дискретное, сосредоточенное в точках x_i, а другое — непрерывное).

Рассматривая окружность $S^1 = \mathbb{R}/\mathbb{Z}$ при помощи ее универсальной накрывающей $\mathbb{R}$, мы поднимем оба распределения на окружности до $\mathbb{Z}$-периодических распределений на накрывающей прямой.

На прямой распределение с плотностью ρ уже имеет функцию распределения

$$f(X) = \int_{X_0}^{X} \rho(n)\, dx.$$

Выбор этой первообразной (зависящей от «начальной точки» X_0) неоднозначен — она определена лишь с точностью до «постоянной интегрирования», первообразными являются также и функции $f + C$ при любой постоянной c.

Чтобы определить аналог отклонения Колмогорова S для пары распределений одинаковой массы на окружности, я поступлю следующим образом.

Обозначим через $\tilde{f}_k$ и $\widetilde{F}_k$ функции распределения на накрывающей прямой, полученные поднятием исходного дискретного распределения k точек на окружности и «теоретического» непрерывного распределения такой же суммарной массы k (в нашем случае это будет равномерное распределение:

$$\widetilde{F}_k(y) = py \;\text{ на накрывающей прямой}).$$

Коэффициент p выбирается так, чтобы приращение величины $\widetilde{F}_k$ при увеличении y на длину рассматриваемой окружности составляло число k точек (массу) изучаемого распределения.

В нашем случае конечной окружности $\mathbb{Z}_n = \mathbb{Z}/(n\mathbb{Z})$ длины n мы получаем для коэффициента p значение k/n.

Чтобы определить отклонение S_k, остается лишь выбрать постоянные интегрирования (c и C) для интегралов f_k и F_k вдоль накрывающей прямой.

Определение 3. *Отклонение* S_k между эмпирическим распределением k точек на окружности длины n и непрерывным распределением массы k на ней определяется как

$$S_k = \inf_{c\in\mathbb{R}}\Bigl(\sup_{x\in\mathbb{R}} |\tilde{f}_k(x) - (\widetilde{F}_k(x) + c)|\Bigr),$$

где $\tilde{f}_k$ и $\widetilde{F}_k$ — какие угодно две первообразные функции поднятий на накрывающую прямую (для эмпирического распределения k точек на окружности и для непрерывного распределения массы k вдоль нее).

Замечание. Величина отклонения S_k не зависит от того, какие именно первообразные взяты: при взятии нижней грани по c вторая первообразная согласуется с первой.

Пример 1. Рассмотрим величины

$$M = \sup_x(\tilde{f} - \widetilde{F}), \quad m = \inf_x(\tilde{f} - \widetilde{F}).$$

Величина отклонения есть

$$S = \frac{M - m}{2}\,,$$

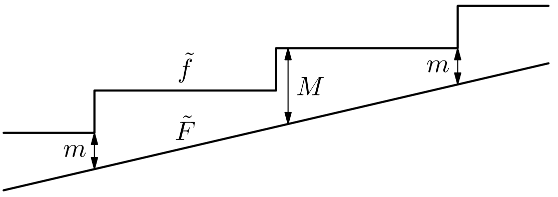


Рис. 2.

так как к $\widetilde{F}$ можно прибавить оптимизирующую постоянную $c = \dfrac{M+m}{2}$ (для которой отклонения верх и вниз одинаковы).

Пример 2. Для четырех взаимно простых с $n = 10$ остатков $\{1, 3, 7, 9\}$ от деления на 10 получается такая оптимальная первообразная $\widetilde{F}_4 + c$:

$$p = \frac{k}{n} = \frac{2}{5}\,, \quad S_4 = \frac{4}{5}\,, \quad \lambda_4 = \frac{4}{5\sqrt{4}} = \frac{2}{5}\,.$$

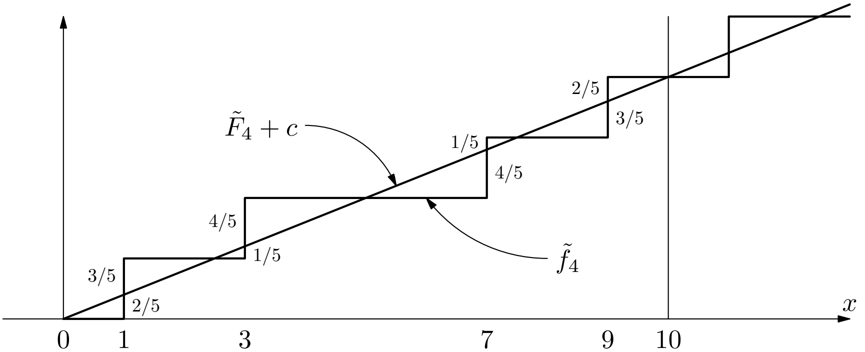


Рис. 3.

Приведенное выше определение 3 отклонения S_k позволяет определить значение $\lambda_k = S_k/\sqrt{k}$ параметра Колмогорова для распределения k точек на окружности. Теорема Колмогорова об универсальном распределении Φ обобщается и на этот вариант его теории (только универсальное предельное распределение Φ параметра Колмогорова λ в этом случае несколько изменится). Каждому распределению на окружности $S^1 = \mathbb{R}/\mathbb{Z}$ соответствует распределение на $[0, 1)$. Сопоставляя этому распределению случайные величины

$$S_k^+ = \sup_{x\in\mathbb{R}} (f_k(x) - F_k(x)),$$
$$S_k^- = \inf_{x\in\mathbb{R}} (f_k(x) - F_k(x)),$$

мы видим, что вышеопределенное отклонение S_k для распределения S^1 удовлетворяет соотношению

$$S_k = (S_k^+ - S_k^-)/2.$$

Поскольку предельное совместное распределение $(S_k^+/\sqrt{k}, S_k^-/\sqrt{k})$ известно ([4, с. 403]), то можно найти и предельное распределение для $S_k/\sqrt{k}$. Вычисления показывают, что для непрерывной функции распределения предел математических ожиданий величины $S_k/\sqrt{k}$ равен $\sqrt{\pi/8} \approx 0{,}63$.

Переход от точек на непрерывной окружности S^1 к точкам на конечных окружностях $\mathbb{Z}_n$ (с большими значениями n) строго не обоснован. Но это, я предполагаю, лишь временное техническое затруднение — асимптотика поведения ответов для $\mathbb{Z}_n$ при $n \to \infty$ согласуется, вероятно, с обобщением на случай случайных величин на окружности S^1 результатов Колмогорова о вещественных случайных величинах.

§ 3. Вычисление значений параметра стохастичности Колмогорова для групп Эйлера примитивных вычетов по модулю n

Приведенные в § 1 сведения о семействах примитивных вычетов позволяют легко нарисовать картинки типа примера 2 из § 2 выше для рассмотренных в § 1 значений n. Эти вычисления дают следующие значения величины отклонения S_k и параметра Колмогорова λ_k:

n	4	5	6	7	8	9	10	11	12	13
k	2	4	2	6	4	6	4	10	4	12
S_k	$\frac{1}{2}$	$\frac{4}{5}$	$\frac{2}{3}$	$\frac{11}{12}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{4}{5}$	$\frac{10}{11}$	$\frac{2}{3}$	$\frac{12}{13}$
λ_k	0,354	0,400	0,471	0,374	0,250	0,272	0,400	0,287	$\frac{1}{3}$	0,266

Второй десяток значений параметров λ_k делает столь же маловероятной гипотезу о случайности примитивных вычетов по этим модулям (при $\lambda \leqslant 0{,}4$ вероятность случайности меньше $3/1000$ по таблице функции Колмогорова Φ, приведенной выше):

n	14	15	16	17	18	19	20	21	22	23	24
k	6	8	8	16	6	18	8	12	10	22	8
S_k	$\frac{6}{7}$	$\frac{14}{15}$	$\frac{1}{2}$	$\frac{16}{17}$	$\frac{2}{3}$	$\frac{18}{19}$	$\frac{4}{5}$	$\frac{8}{7}$	$\frac{10}{11}$	$\frac{22}{23}$	$\frac{2}{3}$
λ_k	0,350	0,330	0,177	0,235	0,272	0,223	0,283	0,330	0,287	0,204	0,236

Оставшиеся 7 значений ($25 \leqslant n \leqslant 31$) доставляют следующие значения параметров Колмогорова λ_k:

n	25	26	27	28	29	30	31
k	20	12	18	12	28	8	30
S_k	$\frac{4}{5}$	$\frac{12}{13}$	$\frac{2}{3}$	$\frac{6}{7}$	$\frac{28}{29}$	$\frac{14}{15}$	$\frac{30}{31}$
λ_k	0,179	0,289	0,157	0,247	0,182	0,330	0,177

Распределение получившихся 28 значений параметров Колмогорова λ_k совершенно непохоже на распределение Колмогорова Ф. Разделим ось значений λ на 6 частей:

$$\lambda \leqslant 0{,}2; \quad 0{,}2 < \lambda \leqslant 0{,}25; \quad 0{,}25 < \lambda \leqslant 0{,}3;$$
$$0{,}3 < \lambda \leqslant 0{,}35; \quad 0{,}35 < \lambda \leqslant 0{,}4; \quad \lambda > 0{,}4$$

мы получим в этих частях, соответственно,

$$5,\ 6,\ 7,\ 5,\ 4,\ 1$$

значений. Суммы этих значений составляют, соответственно,

$$0{,}872;\ 1{,}385;\ 1{,}956;\ 1{,}673;\ 1{,}528;\ 0{,}471.$$

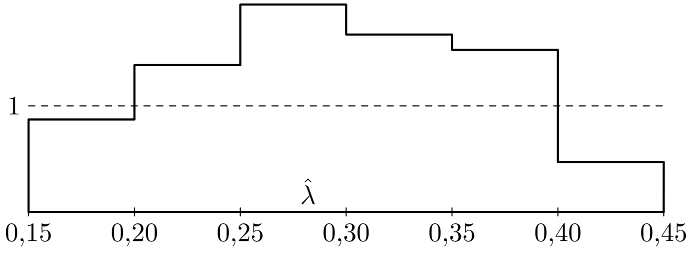


Рис. 4.

Среднее значение $\hat\lambda$ по всем 28 значениям параметра λ составляют примерно 0,282. В распределении Колмогорова суммарная вероятность всех 28 таких значений составляет менее 1/300, так они малы.

Мы заключаем, что изучаемые распределения k примитивных вычетов на окружности $\mathbb{Z}_n$ весьма сильно отличаются от случайных распределений k независимых точек. Это еще одно подтверждение вывода § 1 о зависимости между примитивными вычетами (которые ведь отталкиваются от своих соседей).

Ясно также, что вычисленные значения $\lambda_k(n)$ имеют тенденцию убывать с ростом n. Это можно было бы объяснить ростом с n знаменателя в формуле $\lambda_k = S_k/\sqrt{k}$, если бы отклонения S_k не слишком росли с ростом n (и, следовательно, с ростом $k \sim (6/\pi^2)n$).

Приведенные выше значения наводят на мысль, что величины отклонений S_k остаются ограниченными (и даже величинами порядка 1) при росте k.

Однако, в действительности дело обстоит иначе: при некоторых натуральных значениях n отклонения S_k сколь угодно велики.

А именно, возьмем в качестве n произведение всех j простых чисел, меньших некоторого p_{j+1}:

$$n = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot \ldots \cdot p_j.$$

Лемма 1. *Ни одно из целых чисел x в интервале $2 \leqslant x \leqslant p_{j+1}$ не взаимно просто с n.*

Действительно, все простые множители числа x, меньшего p_{j+1}, являются делителями числа n, а значит — его общими делителями с x.

Лемма 2. *Имеет место следующее неравенство для отклонения D, построенного по $k(n)$ примитивным вычетам по модулю n:*

$$D \geqslant \frac{k}{n} \, \frac{p_{j+1} - 1}{2} \, .$$

Доказательство. График ступенчатой считающей функции f примитивных вычетов по модулю n содержит горизонтальный участок $\{1, \ldots, p_{j+1}\}$ длины $p_{j+1} - 1$. Функция распределения для равномерного распределения массы k вдоль окружности длины n (определенная на накрывающей окружность прямой) линейна с наклоном $p = k/n$.

Отклонение никакой линейной функции наклона p от функции, постоянной на отрезке длины L, не может быть меньшим половины произведения pL. Это и доставляет утверждение леммы 2 (при $p = k/n$, $L = p_{j+1} - 1$). □

Лемма 3. *Произведение*

$$\frac{k(n)}{n} \, \frac{p_{j+1} - 1}{2}$$

принимает сколь угодно большие значения для чисел n, являющихся произведениями достаточно большого числа j последовательных простых чисел

$$n = 2 \cdot 3 \cdot 5 \cdot \ldots \cdot p_j.$$

Запишем число $k(n)$ взаимно простых с n остатков от деления на n по формуле мультипликативности Эйлера

$$k = \prod_{i=1}^{j} (p_i - 1).$$

Мы получим тогда

$$\frac{k}{n} = \prod_{i=1}^{j} \frac{p_i - 1}{p_i} = \prod_{i=1}^{j} \Big(1 - \frac{1}{p_i}\Big).$$

Следовательно, в силу известных оценок для последнего произведения [2, с. 94] справедливо асимптотическое равенство

$$\frac{k}{n} = \frac{e^{-\gamma}}{\ln L}(1 + o(1))$$

при $j \to \infty$, $L = p_{j+1} - 1$, где γ — постоянная Эйлера. Таким образом, утверждение леммы 3 вытекает из леммы 2.

Заметим, что для больших j утверждение леммы 2, основанное на том, что существует L последовательных чисел, не взаимно простых с n, можно усилить. Теорема Эрдеша и Ранкина показывает, что при $L > 20$ найдется не менее

$$\frac{cL \ln L \ln\ln\ln L}{(\ln\ln L)^2}$$

последовательных чисел с таким свойством, где $c > 0$ — некоторая константа. Более того, в [8] показано, что при $L \to \infty$ можно взять $c = 2e^{\gamma} + o(1)$.

Пример. Рассмотрим значение

$$n = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 43 \cdot 47$$

(соответствующее $j = 15$). В этом случае $p_{j+1} = 53$,

$$n = 1 \cdot 2 \cdot 4 \cdot 6 \cdot 10 \cdot 12 \cdot 16 \cdot 18 \cdot 22 \cdot 28 \cdot 30 \cdot 36 \cdot 40 \cdot 42 \cdot 46.$$

Значение k/n можно сосчитать, умножая выписанные сомножители, но проще заменить это умножение сомножителей сложением их логарифмов.

Таблица логарифмов дает (десятичные) логарифмы

$$\lg n \approx 17{,}78879, \quad \lg k \approx 16{,}93088.$$

Поэтому $\lg(n/k) \approx 0{,}85791$, так что

$$\frac{n}{k} \approx 7{,}21 \quad \text{и} \quad \frac{k}{n} \approx 0{,}139.$$

Зная длину $L = 52$ горизонтального участка $[1, 53]$ графика эмпирической функции распределения, мы получаем (лемма 2) для отклонения S оценку снизу

$$S \geqslant \frac{k}{n}\frac{L}{2} \approx 0{,}139 \cdot 26 \approx 3{,}61.$$

Этот пример показывает, что предположение $S \leqslant 2$, доставляемое приведенными выше таблицами (где $n \leqslant 31$) ошибочно: отклонения S бывают сколь угодно большими.

Правда, нужные для этого значения n очень велики (так что величина показателя стохастичности Колмогорова $\lambda = S/\sqrt{k}$ в приведенном выше примере и других, подобных ему, мала). Можно было предположить, что большие отклонения $S_{k(m)}$ встречаются столь редко, что чезаровские средние

$$\hat{S}_n = \frac{\sum\limits_{m=1}^{n} S_{k(m)}}{n}$$

остаются при $n \to \infty$ ограниченными (или даже имеют конечный предел). Оказывается, однако, что это не так! В [3] доказано, что

$$S_{k(m)} \geqslant c 2^{\omega(m)/2}/\omega(m),$$

где $c > 0$ — абсолютная константа, а $\omega(m)$ — число различных простых делителей m. Далее, известно, что для почти всех натуральных m величина $\omega(m)$ близка к $\ln\ln m$. Точнее, для любого $b \in (1/2, 1)$ справедливо асимптотическое соотношение

$$\sum_{m=1}^{n}{}' 1 = o(n) \quad (n \to \infty),$$

где $\sum'$ означает, что суммирование проводится по тем $m \leqslant n$, для которых $|\omega(m) - \ln\ln m| > (\ln\ln m)^b$ ([2, задача 20 к главе 1]). Если теперь через $\sum''$ обозначить суммирование по тем $m \leqslant n$, для которых $|\omega(m) - \ln\ln m| \leqslant$ $\leqslant (\ln\ln m)^b$, то мы получим

$$\sum_{m=1}^{n}{}'' S_{k(m)} \geqslant n(\ln n)^{(\ln 2/2)+o(1)} \quad (n \to \infty).$$

Следовательно,

$$\hat{S}_n \geqslant (\ln n)^{(\ln 2/2)+o(1)} \quad (n \to \infty).$$

Неслучайность распределения примитивных вычетов по модулю n может быть продемонстрирована следующим фактом: соответствующие значения показателей стохастичности Колмогорова $\lambda_{k(n)} = \dfrac{S_{k(n)}}{\sqrt{k(n)}}$ сходятся к нулю при $n \to \infty$, причем довольно быстро. В [5] показано, что

$$S_{k(n)} \leqslant 2^{\omega(n)}.$$

Далее, поскольку при $k = k(n)$

$$\frac{k}{n} \geqslant \prod_{i=1}^{\omega(n)} \frac{p_i - 1}{p_i} \geqslant \prod_{i=1}^{\omega(n)} \frac{i}{i+1} = \frac{1}{\omega(n)+1}\,,$$

из вышеприведенной оценки для произведения в последнем неравенстве следует, что

$$\frac{S_k}{\sqrt{k}} \leqslant \frac{2^{\omega(n)}(\sqrt{\omega(n)+1})}{\sqrt{n}}\,.$$

Так как $2^{\omega(n)} \leqslant d(n)$, где $d(n)$ — количество делителей числа n, то из известной оценки для $d(n)$ (см. [2, с. 32]) следует, что

$$\frac{S_k}{\sqrt{k}} \leqslant \frac{\exp((\ln 2 + o(1))\ln n)/\ln\ln n)}{\sqrt{n}} \to 0$$

при $n \to \infty$.

Список литературы

[1] Арнольд В. И. *Группы Эйлера и арифметика геометрических прогрессий.* — М.: МЦНМО, 2003. — 43 с.

[2] Прахар К. *Распределение простых чисел.* — М.: Мир, 1967. — 512 с.

[3] Bell J. P., Bober J. W. *Bounded step functions and factorial ratio sequences* // Intern. J. Number Theory — 2009. — V. 5 — P. 1419–1431.

[4] Doob J. L. *Heuristic approach to the Kolmogorov – Smirnov theorems* // Ann. Math. Statist. — 1949. — V. 20. — P. 393–403.

[5] Friedlander J. B., Shparlinski I. E. *On the distribution of the power generator* // Math. Comp. — 2001. — V. 70 — P. 1575–1589.

[6] Kolmogorov A. N. *Sulla determinazione empirica di una legge di distribuzione* // Giorn. Ist. Ital. Attuari. — 1933. — V. 4, № 1. — P. 83–91.

[7] Montgomery H. L., Vaughan R. C. *On the distribution of reduced residues* // Ann. Math. — 1986. — V. 123. — P. 311–333.

[8] Pintz J. *Very large gaps between consecutive primes* // J. Number Theory. — 1997. — V. 63. — P. 286–301.

Окружность, описанная вокруг многочлена

Ю. В. Вязовецкий А. С. Тихонов

1. Введение

Надеемся, название этой работы вас заинтересовало и заинтриговало. «Как такое возможно?» — воскликнет удивлённый читатель и наша ближайшая цель — представить соответствующие разъяснения. Хорошо известно (см. любой учебник по высшей алгебре [1–4,6,9,10]), что произвольный многочлен n-й степени имеет (с учётом их кратности) ровно n корней в комплексной плоскости. И для случая кубического многочлена без кратных корней мы всегда можем провести единственную окружность[1)], проходящую через три его корня[2)]. «Ну и что?» — воскликнет упрямый читатель, — «Подумаешь! Вокруг треугольника всегда можно описать окружность». И будет прав. Спорить не станем. Конечно, надо представить на ваш суд что-то более содержательное и об этом сейчас пойдёт речь.

С давних пор в математике большую пользу приносило взаимодействие и взаимопроникновение различных её ветвей: алгебраическими методами решались сложные геометрические задачи (начиная с изобретателя аналитической геометрии Р. Декарта) и, наоборот, геометрические методы с успехом использовались в алгебре, анализе и даже в физике. Одному примеру такого взаимодействия и посвящена данная работа.

Нашей целью является решение кубических уравнений в радикалах. Задача эта классическая и была решена ещё в 16-м веке: формулы Кардано дают полное решение.[3)] Мы хотим предложить геометрическую интерпретацию для них. Наши формулы не будут полностью совпадать с формулами Кардано, однако величины, участвующие в этих формулах, легко выражаются друг через друга. Таким образом, можно говорить о

[1)] Мы будем считать прямые предельным случаем окружностей, т. е. окружностью с бесконечно большим радиусом, как это принято в комплексном анализе.

[2)] С геометрической интерпретацией комплексных чисел можно ознакомиться в книге [5].

[3)] Однако сам Дж. Кардано, впервые опубликовавший эти формулы в 1545 г., указывает на авторство Н. Тартальи (1535). Исторически же первоизобретателем этого метода решения считается С. Ферро (1515).

некоторой модификации способа Кардано, которая имеет свои преимущества/недостатки.

Мы совершим наше путешествие по маршруту: Алгебра – Геометрия – Алгебра. В разделе 2 кратко излагаются необходимые сведения из теории дробно-линейных преобразований и обсуждаются их геометрические свойства. Во третьем разделе мы представляем основные геометрические соображения, лежащие в основе нашего решения, а затем в разделе 4 мы переводим их на язык алгебры.

2. Дробно-линейные преобразования и их свойства

Преобразования комплексной плоскости вида

$$w = \frac{az+b}{cz+d}\,, \quad ad-bc \neq 0\,, \quad a,b,c,d \in \mathbb{C}$$

называются дробно-линейными преобразованиями. Дробно-линейные преобразования обладают следующими свойствами:

1. Круговое свойство: окружности при дробно-линейных преобразованиях переходят в окружности.

2. Консерватизм углов: при дробно-линейных преобразованиях углы между окружностями сохраняются.

3. Сохранение сопряжённости: точки, сопряжённые относительно окружности[4)], переходят в точки, сопряжённые относительно образа этой окружности.

4. Единственным дробно-линейным преобразованием, переводящим три различные точки z_1, z_2 и z_3 в три различные точки w_1, w_2 и w_3, соответственно, является преобразование, задаваемое формулой $\dfrac{(w-w_1)(w_3-w_2)}{(w-w_2)(w_3-w_1)} = \dfrac{(z-z_1)(z_3-z_2)}{(z-z_2)(z_3-z_1)}$.

Мы не будем здесь останавливаться на обосновании этих свойств и отошлем заинтересованного читателя к [5] или [8]. Чтобы читатель понимал все геометрическое богатство этих свойств, хотелось бы попутно отметить, что дальнейшая их разработка приводит к модели геометрии Лобачевского на евклидовой плоскости [8]. Однако, нам — в другом направлении.

4) Две точки называются сопряжёнными относительно окружности, если они лежат на одном и том же луче, выходящем из центра окружности, и произведение их расстояний от центра окружности равно квадрату радиуса этой окружности. В предельном случае сопряжёнными точками являются центр окружности и бесконечно удалённая точка ∞.

3. Геометрический этап решения

В этом разделе мы изложим геометрические идеи, лежащие в основе предлагаемого в работе способа решения кубических уравнений. Наша цель — найти корни уравнения третьей степени, записанного в общем виде

$$w^3 + aw^2 + bw + c = 0\,, \tag{1}$$

используя при этом дробно-линейные преобразования. Для этого рассмотрим дробно-линейное преобразование, которое переводит корни $z_1 = 1$, $z_2 = \epsilon$ и $z_3 = \epsilon^2$, где $\epsilon = -1/2 + i\sqrt{3}/2$, простейшего уравнения $z^3 = 1$ в корни данного уравнения w_1, w_2 и w_3, соответственно. Такое (единственное) преобразование существует. Более того, для него существует явная формула (см. предыдущий раздел). Обозначим через w_0 и w_∞ точки, в которые наше преобразование переводит точки 0 и ∞ (очевидно, они сопряжены относительно окружности, проходящей через корни уравнения). Назовём w_0 и w_∞ точками разрешимости для нашего уравнения[5)].

Для уравнения $z^3 = 1$ точки разрешимости — это 0 и ∞. Как легко видеть, они (и только они) являются неподвижными точками дробно-линейных преобразований, циклически переставляющих корни z_1, z_2, z_3, т. е. (в этом частном случае) поворотов на $120°$ вокруг начала координат, и переставляются при транспозициях корней (например, преобразование $w = 1/z$, которое сохраняет z_1 и переставляет z_2 и z_3, переставляет также 0 и ∞).

Для любой другой тройки точек w_1, w_2, w_3 точки разрешимости w_0 и w_∞ удовлетворяют тем же свойствам, поскольку эти свойства сохраняются при дробно-линейных преобразованиях[6)].

Точки разрешимости допускают и чисто геометрическое описание. Именно, для любых трёх различных точек w_1, w_2, w_3 точки разрешимости w_0 и w_∞ лежат на пересечении трёх окружностей (прямых) ℓ_1, ℓ_2, ℓ_3. Эти окружности задаются следующими свойствами:

- ▷ окружность ℓ_i проходит через w_i;
- ▷ окружности ℓ_i ортогональны к окружности, описанной вокруг w_1, w_2, w_3;

[5)] Мы будем также называть w_0 и w_∞ точками разрешимости и для треугольника, вершинами которого являются корни нашего уравнения w_1, w_2, w_3. Точки разрешимости пополняют богатое семейство замечательных точек в треугольнике (ортоцентр, барицентр, центры вписанной и описанной окружностей и т. д.)

[6)] Замечание для тех, кто знаком с теорией Галуа. Раз точки разрешимости, которые очевидным образом рационально выражаются через корни уравнения, сохраняются при циклических перестановках корней, то существует квадратное уравнение с коэффициентами, которые рационально выражаются через коэффициенты исходного уравнения, корнями которого являются точки разрешимости. В следующем разделе мы найдём это уравнение явно.

▷ окружности ℓ_i и ℓ_j образуют между собой в точках пересечения (т. е. в точках w_0 и w_∞) углы по 120° (см. рисунок 1).

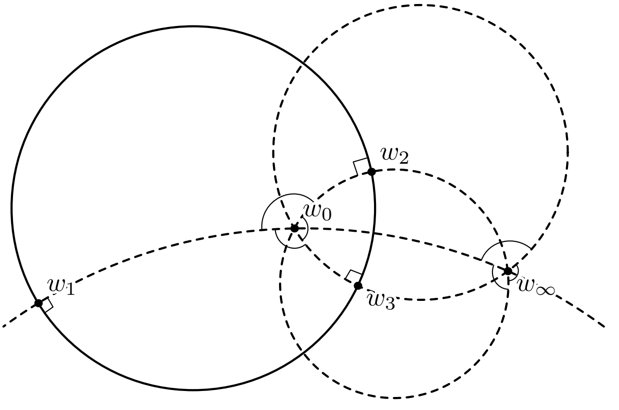


Рис. 1.

Ввиду свойств дробно-линейных преобразований, эти свойства достаточно проверить для точек разрешимости 0 и ∞ уравнения $z^3 = 1$. В этом случае окружности являются прямыми, проходящими через точки 1, ϵ и ϵ^2 и ортогональными единичной окружности. Очевидно, что эти прямые пересекаются в точках 0 и ∞ под углами 120°.

Можно также показать, что пара точек w_0 и w_∞ указанными тремя условиями определяется однозначно. Однако мы опустим обоснование этого факта, так как он нам не понадобится в дальнейшем.

Итак, мы получили геометрическую характеризацию точек разрешимости и доказали их существование для трёх произвольных различных точек комплексной плоскости (случай с кратными корнями требует отдельного рассмотрения).

Теперь осуществим преобразование

$$\omega = \frac{w - w_0}{w - w_\infty}\,, \tag{2}$$

которое переведёт точку w_0 в 0, а точку w_∞ — в ∞. Согласно свойствам дробно-линейных преобразований, оно переведёт корни данного уравнения в точки, лежащие в вершинах правильного треугольника с центром в точке 0. Можно показать, что полученные три точки удовлетворяют уравнению вида $\omega^3 = A$. Тогда задача решения кубического уравнения сведётся к вычислению величины A и нахождению точек разрешимости w_0 и w_∞. Этому посвящён следующий (алгебраический) раздел данной статьи.

4. Алгебраический этап решения

Заранее зная из геометрических соображений (см. предыдущий раздел), что преобразование (2) переведёт наше исходное кубическое уравнение (1) в уравнение $\omega^3 = A$, мы подставим $w = \dfrac{w_\infty \omega - w_0}{\omega - 1}$ (это дробно-линейное преобразование является обратным к преобразованию (2)) в исходное уравнение (1) и потребуем, чтобы коэффициенты в числителе при ω^2 и ω были равны 0. Тогда получаем систему

$$\left\{ \begin{array}{lcl} 3w_\infty^2 w_0 + aw_\infty^2 + 2aw_\infty w_0 + bw_0 + 2bw_\infty + 3c & = & 0 \\ 3w_\infty w_0^2 + aw_0^2 + 2aw_\infty w_0 + bw_\infty + 2bw_0 + 3c & = & 0 \end{array} \right. .$$

Группируя, получаем

$$\left\{ \begin{array}{lcl} w_\infty(3w_\infty w_0 + a(w_\infty + w_0) + b) + (aw_\infty w_0 + b(w_\infty + w_0) + 3c) & = & 0 \\ w_0(3w_\infty w_0 + a(w_\infty + w_0) + b) + (aw_\infty w_0 + b(w_\infty + w_0) + 3c) & = & 0 \end{array} \right. .$$

Учитывая, что $w_\infty \neq w_0$, имеем

$$\left\{ \begin{array}{lcl} 3w_\infty w_0 + a(w_\infty + w_0) + b & = & 0 \\ aw_\infty w_0 + b(w_\infty + w_0) + 3c & = & 0 \end{array} \right. .$$

Решая последнюю систему относительно переменных $w_\infty w_0$ и $w_\infty + w_0$, находим

$$w_\infty + w_0 = \frac{9c - ab}{a^2 - 3b}, \qquad w_\infty w_0 = \frac{b^2 - 3ac}{a^2 - 3b}. \tag{3}$$

Таким образом, w_0 и w_∞ являются корнями квадратного уравнения

$$(a^2 - 3b)w^2 - (9c - ab)w + (b^2 - 3ac) = 0. \tag{4}$$

Заметим, что к успеху привело то, что мы точно знали вид преобразования (2). Если бы мы использовали дробно-линейное преобразование общего вида $\omega = \dfrac{\alpha w + \beta}{\gamma w + \delta}$, то получили бы систему из двух уравнений, но с четырьмя неизвестными. Анализ решений такой недоопределённой системы весьма затруднителен.

Величина A легко вычисляется $A = \dfrac{w_0^3 + aw_0^2 + bw_0 + c}{w_\infty^3 + aw_\infty^2 + bw_\infty + c}$ и, учитывая, что w_0 и w_∞ являются корнями квадратного уравнения (4), получаем

$$\omega^3 = \frac{3w_0 + a}{3w_\infty + a}. \tag{5}$$

Извлекая кубический корень и делая преобразование $w = \dfrac{w_\infty \omega - w_0}{\omega - 1}$, находим корни исходного уравнения (1).

В наших рассуждениях мы предполагали, что точки w_0 и w_∞ являются различными конечными точками комплексной плоскости. Для полноты решения следует рассмотреть вырожденные случаи:

1. Одна из точек w_0 или w_∞ является бесконечно удалённой, что алгебраически означает, что коэффициент при w^2 в (4) равен 0.

   Упражнение 1. Доказать, что условие $a^2 = 3b$ является необходимым и достаточным для того, чтобы корни уравнения (1) лежали в вершинах правильного треугольника и решения (1) задавались бы формулой $w_{1,2,3} = -a/3 + \sqrt[3]{a^3/27 - c}$.

2. Случай $w_0 = w_\infty$, что эквивалентно тому, что уравнение (1) имеет кратные корни, то есть

$$D^2 = [(w_1 - w_2)(w_2 - w_3)(w_3 - w_1)]^2 = \\ = a^2b^2 - 4b^3 - 4a^3c - 27c^2 + 18abc = 0\,.$$

   Упражнение 2. Доказать, что дискриминант уравнения (4) равен $-3D^2$ и в случае, когда $D = 0$, корни уравнения (1) находятся по формулам $w_{1,2} = \dfrac{9c - ab}{2(a^2 - 3b)}$ и $w_3 = \dfrac{2ab + 9c - a^3}{a^2 - 3b}$.

Осталось обсудить связь нашего подхода с формулами Кардано. Как известно [4], для приведённого кубического уравнения (то есть, когда в (1) $a = 0$), корни находятся по формулам Кардано

$$w_{1,2,3} = \sqrt[3]{\widetilde{w}_0} + \sqrt[3]{\widetilde{w}_\infty}\,,$$

где $\widetilde{w}_0$ и $\widetilde{w}_\infty$ являются корнями квадратного уравнения

$$w^2 + cw - \frac{b^3}{27} = 0\,.$$

В случае $a = 0$ уравнение (4) имеет вид $-3bw^2 - 9cw + b^2 = 0$, которое можно переписать в виде $\dfrac{b^2w^2}{9} + c\dfrac{bw}{3} - \dfrac{b^3}{27} = 0$. Откуда моментально следует, что $\widetilde{w}_0 = \dfrac{bw_0}{3}$ и $\widetilde{w}_\infty = \dfrac{bw_\infty}{3}$.

5. Заключение

Итак, привлекая дробно-линейные преобразования комплексной плоскости, мы можем дать геометрическую интерпретацию формул Кардано, а также имеем ещё один способ нахождения корней кубического уравнения. При этом точки разрешимости, вводимые в работе, а также окружность, проходящая через корни уравнения, играют решающую роль в нашем подходе.

Список литературы

[1] Варден, Ван дер, Б. Л. *Алгебра*. М.: Наука, 1976.

[2] Калужнин Л. А., Сущанский В. И. *Преобразования и подстановки*. М.: Наука 1979.

[3] Кострикин А. Н. *Введение в алгебру*. М.: Наука, 1977.

[4] Курош А. Г. *Курс высшей алгебры*. М.: Наука, 1968.

[5] Понарин Я. П. *Алгебра комплексных чисел в геометрических задачах*. М.: МЦНМО, 2004.

[6] Прасолов В. В. *Многочлены*. М.: МЦНМО, 2001.

[7] Прасолов В. В., Соловьев Ю. П. *Эллиптические функции и алгебраические уравнения*. М.: Факториал, 1997.

[8] Привалов И. И. *Введение в теорию функций комплексного переменного*. М.: Наука, 1977.

[9] Фаддеев Д. К. *Лекции по алгебре*. М.: Наука, 1984.

[10] Чезаро Э. *Элементарный учебник алгебраического анализа и исчисления бесконечно малых. Часть первая*. Одесса, 1913.

Ю. В. Вязовецкий, ученик 11-го класса гимназии им. А. П. Чехова, г. Ялта
А. С. Тихонов, доцент кафедры математического анализа Таврического национального университета им. В. И. Вернадского, г. Симферополь
Email: `tikh@mail.ru`

Простое доказательство теоремы Абеля о неразрешимости уравнений в радикалах*

А. Б. Скопенков

Теорема Абеля. *Калькулятор имеет кнопки*

$$1, \quad +, \quad -, \quad \times, \quad : \quad \text{\textit{и}} \quad \sqrt[n]{} \quad \text{\textit{для любого } } n.$$

Он оперирует с комплексными числами и при нажатии кнопки $\sqrt[n]{}$ *выдает все значения корня (и как-то их случайно нумерует). Калькулятор вычисляет числа с абсолютной точностью и имеет неограниченную память. При делении на 0 он выдает ошибку и прекращает работу.*

Тогда ни при каком $n \geqslant 5$ *не существует программы для этого калькулятора, которая по коэффициентам многочлена* n*-й степени выдает конечное множество чисел, содержащее все его корни.*

Комментарии по поводу понятия программы и другая формулировка приведены в пункте «путёвые перестановки для программы и коммутаторы».

В этой заметке мы покажем, как можно было бы придумать простое доказательство теоремы Абеля. Мы следуем изложению [11], которое упрощено по сравнению с [1] (не используется понятия римановой поверхности и гомотопии). В отличие от [11], здесь излагается способ *придумать* доказательство.[1)]

Для понимания самого доказательства достаточно прочитать определение путёвой перестановки из следующего пункта, пункт «план простого доказательства теоремы Абеля» и решения задач, на которые есть ссылки в этом пункте.

*Обновляемая версия: `www.mccme.ru/circles/oim/abel.pdf`

[1)]Заметка основана на занятиях кружка «Олимпиады и математика» и выездных школ команды Москвы на Всероссийскую олимпиаду (`http://www.mccme.ru/circles/oim/mat.htm`). В одном месте (задача 7b, аналогично решаются задачи 8b и 9b) вместе с эвристическим рассуждением из [11] приводится и более строгое (которое мне сообщил А. Канель со ссылкой на А. Ногина). В отличие от [11], здесь используется понятие осторожного пути вместо рассуждений о перестановках башни значений радикальной формулы при обходе параметром замкнутого пути, не проходящего через особые точки радикальной формулы.

Заметим, что теорема (Галуа) о неразрешимости в радикалах *одного конкретного* уравнения [4, 5, 8] — более сильная.

Приводимое доказательство не использует терминов «группа Галуа» и «расширение поля». Несмотря на отсутствие этих *терминов*, *идеи* приводимого доказательства являются *отправными* для теории Галуа (более подробно см. [3, 7, 10]). Материал преподносится в виде задач, к которым даются указания. (И отсутствие терминов, и присутствие задач, характерно не только для дзенских монастырей, но и для серьезного изучения математики.) В конце приводятся задачи для исследования. Если условие задачи является формулировкой утверждения, то это утверждение и надо доказать.

Благодарю М. Вялого, А. Канеля-Белова, Г. Челнокова и участников моих занятий за полезные обсуждения.

1. (a) Теорему Абеля достаточно доказать для уравнений 5-й степени.

(b) Теорему Абеля достаточно доказать для уравнений $z^5 - z + a = 0$.

(На самом деле, и необходимо, поскольку любое уравнение пятой степени сводится к написанному при помощи некоторой программы для нашего калькулятора.)

(c) (это не задача, а *загадка* [2]) Сформулируйте вещественный аналог теоремы Абеля. Вытекает ли он из комплексного? А комплексный из вещественного?

2. (a) У калькулятора оторвали кнопки $\sqrt[n]{}$. Теперь не существует программы для решения квадратного уравнения.

(b)* Не существует программы для решения кубического уравнения, использующей извлечение корня только один раз.

(c)* Не существует программы для решения уравнения 4-й степени, использующей извлечение корня не более двух раз.

Вряд ли у вас получится решить задачу 2c без знания дальнейшего материала. К ней нужно возвращаться по мере его изучения.

Путёвые перестановки

В этом и следующих двух пунктах калькулятор не используется (поэтому даже читатель, у которого возникли вопросы о работе калькулятора, может смело решать задачи).

3. (a) Руководитель кружка двигался по единичной окружности на комплексной плоскости, сделал один оборот и вернулся в исходную точку. Он велел ученику двигаться на комплексной плоскости так, чтобы координата z ученика в любой момент равнялась бы квадрату координаты a руководителя. Как пришлось двигаться ученику?

(b) На следующее занятие кружка на комплексную плоскость пришло два ученика. Руководитель встал в точке 1, а учеников расставил в точки 1 и -1. Потом он велел каждому ученику двигаться так, чтобы координата a руководителя в любой момент равнялась бы квадрату координаты z ученика. А сам пошел по единичной окружности на комплексной плоскости против часовой стрелки, сделал один оборот и вернулся в исходную точку 1. Как пришлось двигаться ученикам? Где они оказались в конце занятия?

4. На следующее занятие кружка на комплексную плоскость пришло уже n учеников. Руководитель не растерялся, встал в точке 1, а учеников расставил в точки $\cos(2\pi k/n) + i\sin(2\pi k/n)$, $k = 1, 2, \ldots, n$. Потом он сказал: «Моя координата равна n-й степени координаты любого из вас! Двигайтесь так, чтобы сохранить это замечательное свойство».

(a) Руководитель сам пошел по единичной окружности на комплексной плоскости против часовой стрелки, сделал один оборот и вернулся в исходную точку 1. Как пришлось двигаться ученикам? Где они оказались в конце занятия?

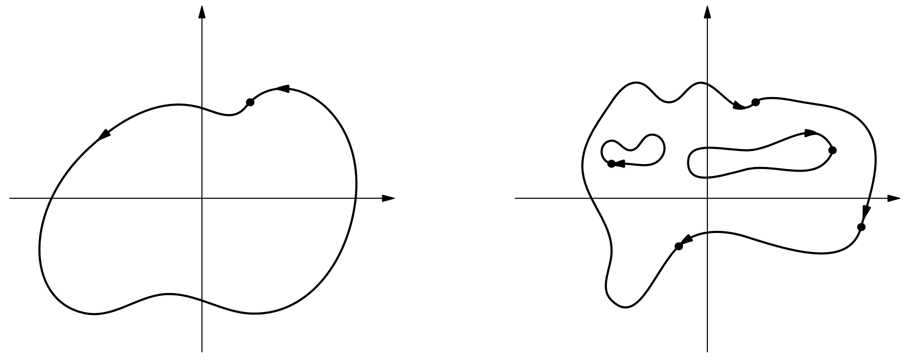
(b) Руководитель называется *добрым*, если он не проходит через 0. Для любого замкнутого маршрута доброго руководителя в конце занятия ученики переставятся.

(c) Для произвольного замкнутого маршрута доброго руководителя если в конце ученик 1 оказался в точке $\cos(2\pi k/n)+i\sin(2\pi k/n)$, то ученик $\cos(2\pi/n)+i\sin(2\pi/n)$ оказался в точке $\cos(2\pi(k+1)/n)+i\sin(2\pi(k+1)/n)$.

(d) Для произвольного замкнутого маршрута доброго руководителя в конце занятия ученики переставятся по степени некоторого цикла.

Определение путёвой перестановки. Пусть $p_a(z)$ — семейство многочленов степени n, коэффициенты которого (но не степень) непрерывно зависят от параметра a. Пусть уравнение $p_{a_0}(z) = 0$ имеет n различных корней, которые обозначены $z_1, \ldots, z_n$. Будем изменять параметр a (руководитель) вдоль некоторого непрерывного замкнутого пути с началом и концом в a_0. Пусть для любой точки a этого пути уравнение $p_a(z) = 0$ имеет n различных корней. Будем двигать i-го ученика, начиная в z_i, и так, чтобы в каждый момент времени его координата была корнем уравнения $p_a(z) = 0$. Тогда в конце движения ученики переставятся.[2)] Назовем полученную перестановку n-элементного множества **путёвой** для семейства уравнений $p_a(z) = 0$ и данной нумерации корней уравнения $p_{a_0}(z) = 0$.

[2)] Для первого знакомства с приводимыми идеями читателю полезно воспользоваться без доказательства существованием такого движения учеников (ср. задачи 4bcd), а также использовать аналогичные наглядные соображения при решении задач 7b, 8b и 9b ниже. Строгое обоснование вытекает из *комплексной теоремы о неявной функции*.



изменение параметра a изменение корней

Рис. 1.

Это действительно перестановка, т. е. два ученика не могут в конце оказаться в одной точке (задача 4b). Цикл $(12\dots n)$ является путёвым для $p_a(z) = z^n - a$ и «естественной» нумерации его корней (задача 4a); все путёвые перестановки являются степенями этого цикла (задача 4d).

5. (a) Тождественная перестановка путёвая для любого семейства многочленов и любой нумерации корней.

(b) Как перестановка, отвечающая вышеописанному замкнутому пути, зависит от начальной нумерации корней?

С этого места мы пропускаем слова «для некоторой нумерации корней», говоря о путёвых перестановках.

Ответы и указания

4. (c) Если $x(t)$ — закон движения первого ученика, то $x(t)(\cos(2\pi/n) + i\sin(2\pi/n))$ — закон движения второго.

(d) Следует из (c).

5. (a) Если параметр a в процессе движения стоит на месте, то каждый корень остается на месте.

(b) *Ответ*: при перенумерации, заданной перестановкой α, перестановка σ, отвечающая вышеописанному замкнутому пути, перейдет в перестановку $\alpha\sigma\alpha^{-1}$.

Зачем нужны путёвые перестановки?

Если бы удалось доказать, что любая путёвая перестановка для семейства уравнений, разрешимого в радикалах, циклическая, то для доказательства теоремы Абеля достаточно было бы привести пример семейства уравнений и нециклической путевой для него перестановки. Однако перестановка $(13)(24) = (1234)^2$ путёвая для $p_a(z) = z^4 - a$.

Хорошо было бы найти другое свойство путёвых перестановок для уравнений, разрешимых в радикалах, которое не выполняется для путёвых перестановок произвольных уравнений. Этого сделать не получится, ибо *любая перестановка является путёвой для некоторого семейства уравнений, разрешимого в радикалах, и некоторой нумерации корней.* (Действительно, если перестановка является произведением k циклов длин $n_1, \dots, n_k$, то искомое семейство уравнений — $\Pi_{s=1}^k((z - n_s)^{n_s} - a)$.)

И всё-таки мы докажем теорему Абеля. Мы найдем свойство *множества* всех путёвых перестановок для уравнения, разрешимого в радикалах, не выполненное для *множества* всех путёвых перестановок произвольного уравнения.

Покажем отправную идею на примере решения задачи 2a. (Это решение сложнее придуманного вами ранее, но зато обобщается до доказательства теоремы Абеля.)

6. (a) Для каких a уравнение $z^2 - 2z + a = 0$ имеет ровно два корня? (Здесь и далее имеются в виду комплексные корни.)

Ответ. Для $a \neq 1$.

(b) Как переставляются корни уравнения $z^2 - 2z + a = 0$ при следующем изменении параметра a?

Сначала от 0 до 0.99 по отрезку,

потом по окружности радиуса 0.01, обходящей вокруг точки 1 один раз

наконец, обратно от 0.99 до 0 по отрезку.

(c) Если $q(a)$ является отношением многочленов от a, то только тождественная перестановка является путёвой для $p_a(z) = z - q(a)$ (здесь путь параметра a не проходит через нули числителя и знаменателя).

(d) Выведите из (b,c) решение задачи 2a.

Ответы и указания

Обозначим $e^{it} := \cos t + i \sin t$. (В настоящем тексте это нужно воспринимать именно как обозначение. Свойство $e^{i(t_1+t_2)} = e^{it_1}e^{it_2}$ следует из формул для синуса и косинуса суммы.)

6. (b) Сначала корни приближаются к 1 с разных сторон. Чтобы корень двигался по закону $z(t) = 1 + \varepsilon e^{it}$, параметр a должен двигаться по закону $a(t) = 2z(t) - z^2(t) = 1 - \varepsilon^2 e^{2it}$. Поэтому возьмем $\varepsilon = 0.1$. Когда a совершит один оборот, каждый из корней совершит пол-оборота, т. е. они поменяются местами. Далее корни снова удалятся от 1.

Ответ: корни меняются местами.

Путёвые перестановки для уравнений 3-й, 4-й и 5-й степени

7. (a) Для каких a уравнение $z^3 - 3z + a = 0$ имеет ровно три корня?

Указание. Проще всего решать эту задачу при помощи следующих фактов:
– любой многочлен степени n имеет ровно n комплексных корней, и
– любой кратный корень многочлена является также корнем его производной.

Ответ. Для $a \neq \pm 2$.

(b) Как переставляются корни уравнения $z^3 - 3z + a = 0$ при следующем изменении параметра a?

Сначала от 0 до $2 - \delta_1$ по отрезку,
потом по кривой, «обходящей вокруг точки 2 один раз»,
наконец, обратно от $2 - \delta_2$ до 0 по отрезку.

(*В задачах 7bc, 8bc и 9bc выберите сами кривую и малые числа δ_1, δ_2, чтобы было удобно находить перестановку.*)

(c) Как переставляются корни уравнения $z^3 - 3z + a = 0$ при следующем изменении параметра a?

Сначала от 0 до $-2 - \delta_1$ по отрезку,
потом по кривой, «обходящей вокруг точки -2 один раз»,
наконец, обратно от $-2 - \delta_2$ до 0 по отрезку.

(d) Для $p_a(z) = z^3 - 3z + a$ все перестановки путёвые.

8. (a) Для каких a уравнение $z^4 - 4z + a = 0$ имеет ровно четыре корня?

Ответ. Для $a \neq 3, 3\alpha, 3\alpha^2$, где $\alpha = (1 + i\sqrt{3})/2$.

(b) Как переставляются корни уравнения $z^4 - 4z + a = 0$ при следующем изменении параметра a?

Сначала от 0 до $3 - \delta_1$ по отрезку,
потом по кривой, «обходящей вокруг точки 3 один раз»,
наконец, обратно от $3 - \delta_2$ до 0 по отрезку.

(c) Как переставляются корни уравнения $z^4 - 4z + a = 0$ при следующем изменении параметра a?

Сначала от 0 до $(3 - \delta_1)\alpha$ по отрезку,
потом по кривой, «обходящей вокруг точки 3α один раз»,
наконец, обратно от $(3 - \delta_2)\alpha$ до 0 по отрезку.

(d) Для $p_a(z) = z^4 - 4z + a$ все перестановки путёвые.

9. (a) Для каких a уравнение $z^5 - 5z + a = 0$ имеет ровно пять корней?

Ответ. Для $a \neq 4, 4i, -4, -4i$.

(b) Как переставляются корни уравнения $z^5 - 5z + a = 0$ при следующем изменении параметра a?

Сначала от 0 до $4 - \delta_1$ по отрезку,
потом по кривой, «обходящей вокруг точки 4 один раз»,
наконец, обратно от $4 - \delta_2$ до 0 по отрезку.

(c) Как переставляются корни уравнения $z^5-5z+a=0$ при следующем изменении параметра a?

Сначала от 0 до $(4-\delta_1)i$ по отрезку,
потом по кривой, «обходящей вокруг точки $4i$ один раз»,
наконец, обратно от $(4-\delta_2)i$ до 0 по отрезку.

(d) Для $p_a(z)=z^5-5z+a$ все перестановки путёвые.

10. Найдите все путёвые перестановки для
(a) $p_a(z)=z^4+2(1-2a)z^2+1$; (b) $p_a(z)=(z^3-a)^3-a(a-1)$.

Ответы и указания

7. (b) Ответ: корни $0, \sqrt{3}$ поменяются местами, корень $-\sqrt{3}$ остается на месте.

Зададим движение корня 0, по нему восстановим движение параметра a, а затем движение остальных корней. См. рис. 2.

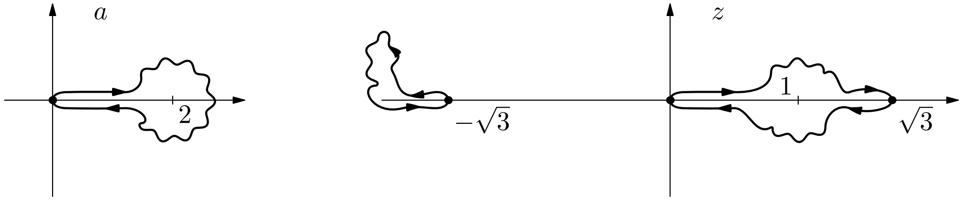


Рис. 2.

Сначала первый корень $z=0$ приближается к 1 слева и превращается в $1-\delta_1$. При этом $a=3z-z^3$ приближается к 2 слева. Значит, второй корень $\sqrt{3}$ приближается к 1 справа и превращается в $1+\delta_2$. А третий корень $-\sqrt{3}$ остается отрицательным. См. рис. 3.

Затем пусть первый корень идет в $1+\delta_2$ по некоторой кривой, близкой к 1 и не пересекающей вещественной оси нигде, кроме своих концов. Тогда второй корень остается близким к 1, а третий корень остается отрицательным. Значит, корень второй корень придет в $1-\delta_1$.[3)]

Далее первый корень приближается к $\sqrt{3}$ слева, примерно повторяя начальную часть движения второго корня в противоположном направлении. Значит, второй корень приближается к 0 справа, примерно повторяя начальную часть движения первого корня в противоположном направлении. А третий корень остается отрицательным.

[3)] Вот неформальная иллюстрация этого движения. Пусть первый корень движется по закону $z(t)=1+\delta_1 e^{it}$. Тогда параметр a движется по закону $a(t)=3z(t)-z^3(t)=$ $=2-3\delta_1^2e^{2it}-\delta_1^3e^{3it}$. Так как δ_1 мало, эта кривая близка к окружности $2-3\delta_1^2e^{2it}$. А значит, второй корень придет примерно в точку $1-\delta_1$ (ибо третий корень остается далеко).

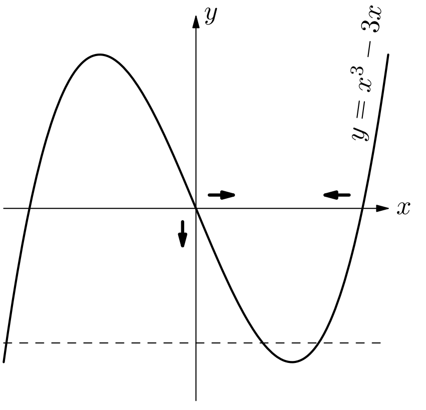


Рис. 3.

(c) *Указание*. Используйте нечетность.

Ответ: корни 0, $-\sqrt{3}$ поменяются местами, корень $\sqrt{3}$ остается на месте.

(d) Докажите, что любая перестановка 3-элементного множества представляется в виде композиции транспозиций (12) и (13).

8. (b) Аналогично задаче 7b корни 0, $\sqrt[3]{4}$ поменяются местами. Остальные два корня в процессе движения не пересекают ось Ox и поэтому в конце движения будут на своих прежних местах.

Ответ: корни 0, $\sqrt[3]{4}$ поменяются местами, остальные два корня остаются на месте.

(c) $p_{\alpha a}(\alpha z) = \alpha p_a(z)$.

Ответ: корни 0, $\sqrt[3]{4}\alpha$ поменяются местами, остальные два корня остаются на месте.

(d) Докажите, что любая перестановка 4-элементного множества представляется в виде композиции транспозиций (12), (13) и (14).

9. (b) Аналогично задачам 7b и 8b. Корни 0, $\sqrt[4]{5}$ поменяются местами, а остальные три корня вернутся на свои прежние места. (Ибо корни $i\sqrt[4]{5}$ и $-i\sqrt[4]{5}$ в процессе движения не пересекают вещественной оси, а корень $-\sqrt[4]{5}$ остается вещественным отрицательным.)

(d) Докажите, что любая перестановка 5-элементного множества представляется в виде композиции транспозиций (12), (13), (14) и (15). Для этого докажите, что

– любая перестановка 5-элементного множества является композицией циклов,

– любой цикл является композицией транспозиций,

– любая транспозиция является композицией транспозиций (12), (13), (14) и (15).

Осторожные пути

11. (a) Пусть $q(a)$ — отношение многочленов от a. Докажите, что все путёвые перестановки для $p_a(z) = z^n - q(a)$ являются степенью некоторого одного цикла.

(b)* Пусть существует программа для решения уравнения $p_a(z) = 0$, использующая извлечение корня только один раз. Обязательно ли все путёвые перестановки являются степенью некоторого одного цикла?

Наш калькулятор имеет неприятную особенность: результат вычислений не всегда однозначно определяется вводимыми данными (например, программа, выдающая *первое* значение $\sqrt{1}$, будет случайно выдавать 1 или -1). Эту неприятность можно преодолеть, осознав, что теорему Абеля достаточно доказать для «симметричных» программ для нашего калькулятора (или, эквивалентно, для похожего калькулятора, оперирующего с *множествами* комплексных чисел). Но всё равно будет непросто дать определение путёвой перестановки *для программы*, которое необходимо для решения задачи 11b. Мы поступим по-другому.

Радикальной формулой относительно $a_0, \ldots, a_n$ называется (упорядоченный) набор рациональных функций (т. е. отношений многочленов) $p_1, \ldots, p_s$ от $n+1, \ldots, n+s$ переменных, соответственно, и целых положительных чисел $k_1, \ldots, k_s$. Для радикальной формулы определим выражения

$$z_1, \ldots, z_s \quad \text{формулой} \quad z_j^{k_j} = p_j(a_0, \ldots, a_n, z_1, \ldots, z_{j-1}), \quad j = 1, 2, \ldots, s.$$

Уравнение $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$ с комплексными переменными коэффициентами называется *разрешимым в радикалах*, если существует радикальная формула относительно $a_0, \ldots, a_n$, для которой любой корень уравнения является одним из значений одного из выражений $z_1, \ldots, z_s$. Ср. [11, Chapter 5].

Мы докажем теорему Абеля в следующей эквивалентной форме: *ни при каком* $n \geqslant 5$ *уравнение* $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$ *с комплексными переменными коэффициентами не является разрешимым в радикалах*.

Для доказательства теоремы Абеля полезно следующее понятие. Назовем замкнутый путь на плоскости **осторожным** для данной радикальной формулы, в которой $a_0, \ldots, a_n$ зависят от параметра a, если при изменении параметра a вдоль этого пути каждое значение каждого выражения z_s возвращается на место.

12. (a) Если путь является осторожным относительно каждой из двух радикальных формул $p_1, \ldots, p_s; k_1, \ldots, k_s$ и $q_1, \ldots, q_t, l_1, \ldots, l_t$, то он является осторожным относительно их *суммы* $p_1, \ldots, p_s, q_1, \ldots, q_t, p_s+q_t; k_1, \ldots, k_s, l_1, \ldots, l_t, 1$.

(b) Определите разность, произведение и частное радикальных формул. Докажите для них аналог предыдущего пункта.

(c) *n-й степенью* замкнутого пути называется новый замкнутый путь, который получается прохождением исходного пути n раз. Если путь является осторожным относительно радикальной формулы $p_1, \ldots, p_s; k_1, \ldots, k_s$, то его n-я степень является осторожной относительно радикальной формулы $p_1, \ldots, p_s, z_s; k_1, \ldots, k_s, n$.

(d) $k_1 \cdot \ldots \cdot k_s$-я степень любого пути является осторожной относительно радикальной формулы $p_1, \ldots, p_s; k_1, \ldots, k_s$.

(e) Если семейство уравнений $p_a(z) = 0$ разрешимо при помощи радикальной формулы $p_1, \ldots, p_s; k_1, \ldots, k_s$, то $k_1 \cdot \ldots \cdot k_s$-я степень любой путёвой перестановки тождественна.

Последним утверждением нельзя воспользоваться для доказательства теоремы Абеля, ибо любая перестановка в некоторой степени равна тождественной.

Коммутатором двух замкнутых путей называется новый замкнутый путь, который получается последовательным прохождением
– первого пути,
– второго пути,
– первого пути в обратную сторону,
– второго пути в обратную сторону.

13. (a) Если оба пути — осторожные относительно радикальной формулы $p_1, \ldots, p_s; k_1, \ldots, k_s$, то их коммутатор — осторожный относительно радикальной формулы $p_1, \ldots, p_s, z_s; k_1, \ldots, k_s, n$.

(b) Если существует программа для решения уравнения $p_a(z) = 0$, использующая одноэтажные извлечения корней, то любые две путёвые перестановки коммутируют (т. е. $\sigma\tau = \tau\sigma$).

(c) Не существует программы для решения кубического уравнения, использующей одноэтажные извлечения корней.

(Сравните с задачей 2b и вашим решением ее.)

14. (a) Какое условие на множество путёвых перестановок следует из существования программы для решения уравнения $p_a(z) = 0$, использующей не более, чем двухэтажное извлечение корня?

(b) Выведите из вашего решения пункта (a) и задачи 8d решение задачи 2c.

15. (a) Какое условие на множество путёвых перестановок следует из существования программы для решения уравнения $p_a(z)=0$, использующей не более, чем трехэтажное извлечение корня?

(b) Существуют две не коммутирующие перестановки 5-элементного множества, каждая из которых является коммутатором некоторых двух произведений коммутаторов.

(c) Выведите из (a,b) и задачи 9 необходимость наличия хотя бы трех этажей в программе, якобы решающей семейство уравнений $z^5-z+a=0$.

(d) Докажите теорему Абеля.

Ответы и указания

11. (a) Аналогично задачам 4cd.

13. (a) Значение выражения z_s возвращается на место в результате обхода числом a каждого из двух данных замкнутых путей L_1 и L_2. Поэтому n значений выражения z_{s+1} имеют вид $x, x\varepsilon, x\varepsilon^2, x\varepsilon^3, \dots, x\varepsilon^{n-1}$ для некоторого x и $\varepsilon := \cos(2\pi/n) + i\sin(2\pi/n)$. Аналогично задаче 4d для любого замкнутого пути L найдется такое $k(L)$, что в результате изменения параметра a вдоль этого пути число $x\varepsilon^s$ переходит в число $x\varepsilon^{s+k(L)}$. Поэтому в результате прохождения параметром a коммутатора путей L_1 и L_2 число $x\varepsilon^s$ переходит в число $x\varepsilon^{s+k(L_1)+k(L_2)-k(L_1)-k(L_2)}$.

(b) Следует из (a).

14. (a) *Подсказка.* Условие $\sigma\tau = \tau\sigma$ (на перестановки) равносильно тождественности перестановки $\sigma\tau\sigma^{-1}\tau^{-1}$. Эта перестановка называется *коммутатором* перестановок σ, τ. Если имеется двухэтажная формула, то для путёвых перестановок σ, τ коммутатор (т. е. перестановка $\sigma\tau\sigma^{-1}\tau^{-1}$) может не быть тождественным. Однако для коммутаторов выполняется некоторое условие. Найдите его!

Ответ. Если существует программа для решения уравнения $p_a(z)=0$, использующая не более, чем двухэтажное извлечение корня, то коммутаторы путёвых перестановок коммутируют. (Даже произведения коммутаторов путёвых перестановок коммутируют.)

Доказательство аналогично решению задачи 13a.

15. (a) Если существует программа для решения уравнения $p_a(z)=0$, использующая не более, чем трехэтажное извлечение корня, то коммутаторы коммутаторов путёвых перестановок коммутируют. (Даже произведения коммутаторов произведений коммутаторов путёвых перестановок коммутируют.) Доказательство аналогично задачам 13ab и 14a.

(b) Пример можно придумать напрямую или доказав, что любая четная перестановка является произведением коммутаторов. (Определение четной перестановки напомнено ниже.)

План простого доказательства теоремы Абеля

Мы довольно долго *придумывали* доказательство теоремы Абеля. *Изложить* же доказательство можно совсем коротко. (Освобождение доказательства от деталей, возникших при его придумывании и не нужных для него самого — важная часть его проверки.) Приведем план такого изложения.

Теорема Абеля вытекает из следующих трех лемм. Для их формулировки введем следующее определение (которое поможет нам коротко произносить громоздкие конструкции, возникшие в задачах 13b, 14a и 15a). Для данного семейства уравнений $p_a(z) = 0$ назовем 0-*путёвыми* путёвые перестановки. Если уже определены n-путёвые перестановки, то назовем перестановку $(n+1)$-*путёвой*, если она представляется в виде композиции коммутаторов некоторых n-путёвых перестановок.

Лемма о коммутаторах. *Если существует программа для решения семейства уравнений* $p_a(z) = 0$, *использующая не более, чем* n*-кратное извлечение корня, то лишь тождественная перестановка является* n*-путёвой для этого семейства.*

Лемма о примере уравнения. *Существует такое семейство* $p_a(z)$ *уравнений* 5*-й степени, множество всех путёвых перестановок которого совпадает с множеством всех перестановок* 5*-элементного множества.*

Лемма о четных перестановках. *Любая четная перестановка 5-элементного множества является произведением коммутаторов четных перестановок.*

Напомним, что перестановка называется *четной*, если она представляется в виде произведения циклов длины 3. (Это определение равносильно общепринятому.)

Лемма о коммутаторах следует из задачи 13a (аналогично задачам 13b, 14a, 15a). Лемма о примере уравнения следует из задачи 9d. Доказательство леммы о четных перестановках — задача (*указание*: достаточно доказать, что таковым является цикл длины 3).

Задачи для исследования

В математике имеется много результатов, непосредственно связанных с теоремой Абеля. См., например, [6, 9].

Следующие задачи показывают, что метод Феррари для решения уравнения 4-й степени «самый простой», а формула дель Ферро – Кардано – Тартальи в виде

$$x = \frac{1}{2}\left(\sqrt[3]{-a+\sqrt{a^2-4}} - \sqrt[3]{a+\sqrt{a^2-4}}\right)$$

для решения кубического уравнения $x^3 - 3x + a = 0$ — не «самая простая».

16. Существует программа для калькулятора, строящая по числу a конечное множество, содержащее все корни уравнения $x^3 - 3x + a = 0$, и содержащая только одно извлечения корня из выражения, содержащего корни.

17. Не существует

(a) формулы вида $z = \sqrt[k]{p + \sqrt[l]{q + \sqrt[m]{r}}} + \sqrt[n]{s + \sqrt[o]{t}}$ для решения уравнения $x^4 - 4x + a = 0$ ни для каких целых положительных k, l, m, n, o и рациональных функций p, q, r, s, t от a.

(b) программы, строящей по числу a конечное множество, содержащее все корни уравнения $x^4 - 4x + a = 0$, и содержащей только одно «трехэтажное» извлечения корня.

18. Существует ли программа для *вещественного* аналога калькулятора, определенного в начале заметки, находящая все *вещественные* корни

(a) уравнения $x^3 + px + q = 0$ по его коэффициентам p, q?

(b) уравнения $x^4 + px^2 + qx + r = 0$ по его коэффициентам p, q, r?

(c) уравнения $x^5 + px^3 + qx^2 + rx + s = 0$ по его коэффициентам p, q, r, s?

19. Добавим к калькулятору кнопку, выдающую по числу $\cos\alpha$ все значения числа $\cos(\alpha/5)$. Появится ли программа для решения уравнения 5-й степени?

Задачи 16 и 17 несложны. (Задачу 16 можно решить, не используя изложенных выше идей, а задачу 17 — вряд ли.) К сожалению, автору не удалось найти в литературе ответы на естественные вопросы задач 18 и 19 (хотя, видимо, ответы известны специалистам).

Ответы и указания

16. $b := \frac{1}{2}\sqrt[3]{-a + \sqrt{a^2 - 4}}$ и $x := b - \dfrac{1}{b}$.

17. Если бы такая формула/программа существовала, то все коммутаторы произведений коммутаторов (перестановок 4-элементного множества) были бы степенью некоторой одной перестановки. А это не так.

Список литературы

[1] Алексеев В. Б. *Теорема Абеля.* М.: Наука, 1976.

[2] Виро О. Я., Иванов О. А., Нецветаев Н. Ю., Харламов В. М. *Элементарная топология.* М.: МЦНМО, 2010.

[3] Козлов П., Скопенков А. *В поисках утраченной алгебры: в направлении Гаусса (подборка задач)* // Мат. Просвещение, сер. 3, вып. 12, 2008. С. 127–144. Эл. версия: `http://arxiv.org/abs/0804.4357`

[4] Колосов В. А. *Теоремы и задачи алгебры, теории чисел и комбинаторики*. М.: Гелиос, 2001.

[5] Прасолов В. В. *Многочлены*. М.: МЦНМО, 2003. Эл. версия: http://www.mccme.ru/prasolov

[6] Прасолов В. В., Соловьев Ю. П. *Эллиптические функции и алгебраические уравнения*. М.: Факториал, 1997.

[7] Скопенков А. *Философски-методическое отступление* // Сборник материалов московских выездных математических школ. Под ред. А. Заславского, Д. Пермякова, А. Скопенкова, М. Скопенкова и А. Шаповалова. М.: МЦНМО, 2009. Эл. версия: `http://www.mccme.ru/circles/oim/mvz.pdf`

[8] Тихомиров В. М. *Абель и его великая теорема* // Квант, №1, 2003. С. 11–15.

[9] Хованский А. Г. *Топологическая теория Галуа*. Москва, МЦНМО, 2008.

[10] Челноков Г. Р. *Основы теории Галуа в интересных задачах*. `http://www.mccme.ru/circles/oim/materials/grishalois.pdf`.

[11] Fuchs D., Tabachnikov S. *Mathematical Omnibus*. AMS, 2007.

А. Б. Скопенков, механико-математический факультет Московского государственного университета им. М. В. Ломоносова, Независимый московский университет, Московский институт открытого образования
Инфо: `http://dfgm.math.msu.su/files/skopenkov/PAPERSCI.pdf`
e-mail: `skopenko@mccme.ru`

Четырехвалентные графы с крестовой структурой

В. О. Мантуров

Одним из важнейших классов графов является класс *крестовых графов*: таковыми мы будем называть четырехвалентные графы, у которых в каждой вершине указывается разбиение четырех полуребер, инцидентных каждой вершине графа, на две пары «противоположных». Понятие противоположности будет важно при вложениях крестовых графов в поверхности: будем говорить, что вложение графа в двумерную поверхность *согласовано с крестовой структурой*, если полурёбра, противоположные в вершине, являются противоположными на поверхности. Любой четырехвалентный граф, вложенный в поверхность, естественным образом наследует из этой поверхности крестовую структуру. Будем называть не противоположные полурёбра, инцидентные одной вершине, *соседними*.

На протяжении всей статьи графы подразумеваются связными и конечными; петли и кратные ребра допускаются (иногда графы с петлями и кратными ребрами называют *мультиграфами*).

Замечание 1. Для определения валентности графа в вершине удобно ввести понятие *полуребра*. Добавляя на каждом ребре графа по новой вершине, мы получим разбиение ребер на *полурёбра*. Полурёбра удобны для подсчета валентности вершин графа: если вершина исходного графа инцидентна некоторому ребру дважды, то это значит, что на преобразованном графе эта вершина инцидентна двум разным полурёбрам.

Крестовые графы могут возникнуть следующим образом. Пусть S — двумерная поверхность, а Γ — вложенный в поверхность S произвольный граф. Построим *срединный* (*медиальный*) граф $M(S,\Gamma)$ следующим образом. Вершинами нового графа будут середины ребер исходного графа, а рёбра нового графа будут строиться следующим образом: каждое ребро будет соответствовать *углу*, т. е. паре соседствующих ребер и вершине между ними. Вообще говоря, если два ребра графа Γ имели две общие концевые вершины, то на графе $M(S,\Gamma)$ две соответствующие этим ребрам вершины будут соединены двумя различными ребрами, см. рис. 1.

Полученный граф $M(S,\Gamma)$ будет четырехвалентным: у каждого ребра имеются два соседних ребра с одного конца и два соседних ребра с другого

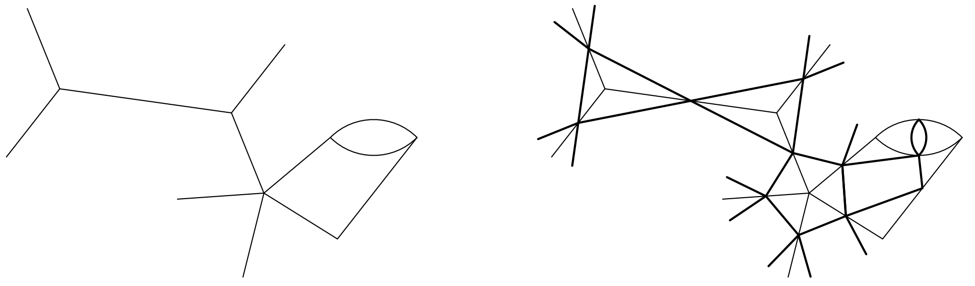


Рис. 1. Построение медиального графа

конца; более того, на данном графе естественным образом определяется наследуемая из поверхности крестовая структура.

Рассмотрим простейший четырехвалентный граф с одной вершиной и двумя ребрами, ей инцидентными. Отметим, что на этом графе имеется две разные крестовые структуры. В одной крестовой структуре два полуребра одного ребра являются противоположными, а в другой крестовой структуре они являются соседними. Отметим, что первый крестовый граф Γ_1 не вложи́м в плоскость, а второй крестовый граф вложи́м в плоскость, см. рис. 2.

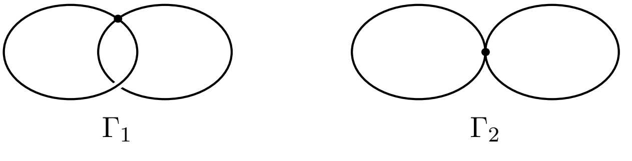


Рис. 2. Простейший граф и две его крестовые структуры

Как оказывается, граф Γ_1 естественным образом вложи́м в тор с сохранением крестовой структуры: представляя тор в виде квадрата с отождествленными противоположными сторонами, мы можем изобразить на торе граф Γ_1 в виде параллели и меридиана с одной точкой пересечения, см. рис. 3.

Граф Γ_1 не вложи́м в плоскость: действительно, предположив, что одна окружность вложена в плоскость, мы получим (по теореме Жордана) две области, на которые она делит плоскость. Вторая окружность, проходя через точку перекрестья, должна переходить из одной области в другую, откуда она не может возвратиться.

Как оказывается, граф Γ_1 в некотором смысле является единственным препятствием планарности крестового графа. А именно, имеет место следующая

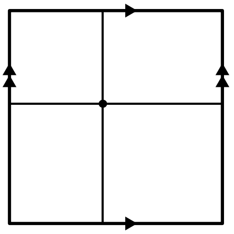


Рис. 3. Вложение графа Γ_1 в тор

Теорема 1. *Граф не вложи́м в плоскость тогда и только тогда, когда у него имеются два цикла без общих ребер, обладающие единственной точкой перекрестья.*

Дадим некоторые пояснения. *Циклом* мы называем непрерывное отображение окружности в граф, взаимно однозначное вне прообразов вершин графов. Этим мы подчеркиваем, что в цикле каждое ребро встречается не более одного раза. Мы говорим, что два цикла без общих ребер имеют *перекрестье* в некоторой вершине v, если один из них содержит одну пару противоположных полуребер в этой вершине, а другой — вторую. При этом в формулировке теоремы не накладывается никаких ограничений на количество (нетрансверсальных) пересечений двух циклов. Далее мы будем называть *препятствием Васильева* два цикла без общих ребер, имеющие единственную точку перекрестья.

Ясно, что на плоскости нельзя изобразить две несамопересекающиеся замкнутые кривые, которые бы имели перекрестье в одной точке. С другой стороны, если две кривые «касаются» в нескольких точках, то эти касания можно «развести» в точках нетрансверсального пересечения так, как показано на рис. 4.

Таким образом, предполагая, что граф вложен в плоскость с сохранением крестовой структуры, мы можем «развести» два цикла во всех

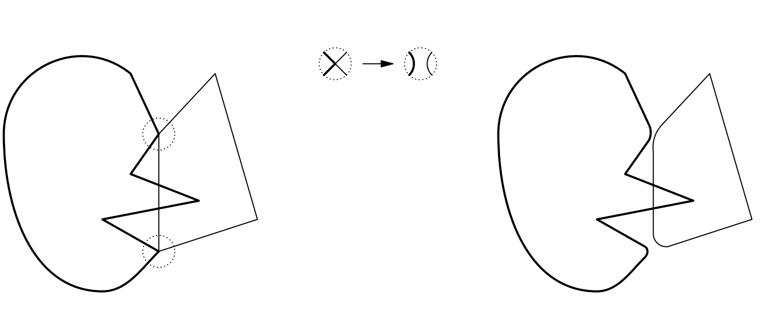


Рис. 4. Устранение нетрансверсальных пересечений

точках, где они пересекаются не трансверсально; если окажется, что при этом останется ровно одна точка перекрестья, то это будет противоречить планарности графа.

Приведенный выше результат был выдвинут в качестве гипотезы В. А. Васильевым [1] и доказан автором настоящей статьи в [3].

Простая часть гипотезы следует из приведенных выше рассуждений; доказательству сложной части (о том, что если граф не вложи́м в плоскость, то два таких цикла обязательно найдутся) посвящена оставшаяся часть настоящей статьи.

1. Поворачивающие обходы. Хордовые диаграммы

Назовем *хордовой диаграммой* конечный трехвалентный граф, состоящий из цикла, проходящего через все вершины (*окружности* хордовой диаграммы) по одному разу и оставшихся ребер, которые мы будем называть *хордами* хордовой диаграммы. Такие хордовые диаграммы рассматриваются с точностью до эквивалентности — изоморфизма графов, переводящего окружность в окружность. Хордовую диаграмму будем называть *ориентированной*, если ориентирована ее окружность; для ориентированных хордовых диаграмм изоморфизм предполагает сохранение ориентации.

Назовем *поворачивающим обходом* такой способ прохождения всех ребер четырехвалентного оснащенного графа с заходом в каждую вершину дважды, при котором в каждой вершине мы переходим с ребра на *не противоположное* ему (его же можно считать отображением f окружности в граф). Хордовая диаграмма $C(\Gamma)$, соответствующая поворачивающему обходу Γ строится следующим образом. Окружностью хордовой диаграммы является отображаемая окружность S^1, а хордами соединяются пары точек, имеющие один и тот же образ относительно отображения f.

Упражнение 1. Покажите, что у каждого связного крестового графа существует поворачивающий обход.

При этом хорды хордовой диаграммы, соответствующей поворачивающему обходу, естественно делить на два типа. Хорды *первого типа* соответствуют перекресткам первого типа, которые выглядят следующим образом. В данном перекрестке по двум противоположным (полу)ребрам ориентация обхода направлена к вершине, а по двум другим противоположным (полу)ребрам она ориентирована от вершины. В ином случае, если некоторое (полу)ребро ориентировано по направлению к вершине, а противоположное ему (полу)ребро ориентировано по направлению от вершины, будем говорить, что хорда *имеет второй тип*.

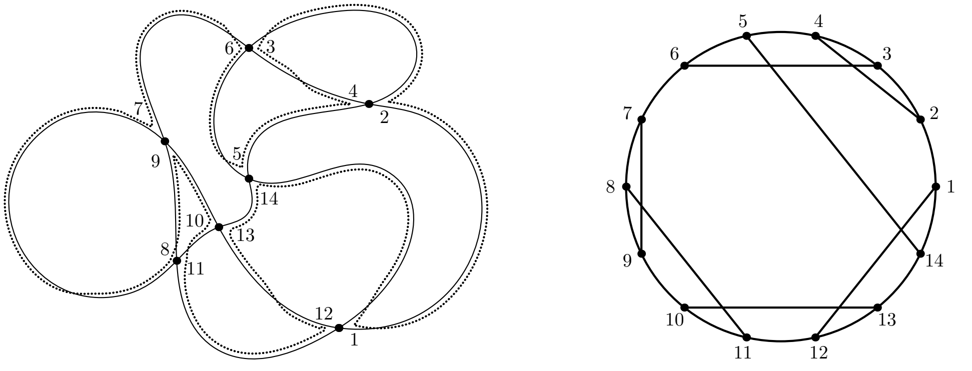


Рис. 5. Поворачивающий обход и поворачивающая хордовая диаграмма

На рис. 5 каждая вершина пронумерована два раза согласно двум моментам прохождения поворачивающего обхода через вершину (на этом примере все хорды имеют первый тип).

Как оказывается, поворачивающие обходы играют решающую роль в определении планарности крестовых графов.

2. d-диаграммы

Скажем, что две хорды p, q хордовой диаграммы D *зацеплены*, если их концы расположены на окружности в чередующем порядке.

Таким образом, при изображении хордовой диаграммы на плоскости посредством евклидовой окружности и прямолинейных хорд, зацепленные хорды изображаются пересекающимися.

Хордовая диаграмма называется *d-диаграммой*, если множество ее хорд можно разбить на два семейства таким образом, что любые две хорды из одного семейства оказались незацеплены.

d-Диаграммы играют ключевую роль в различных вопросах теории узлов, см., напр., [5].

Читатель может в качестве упражнения проверить следующий факт: *хордовая диаграмма вложима в плоскость как граф тогда и только тогда, когда она является d-диаграммой.*

Подсказка изображена на рис. 6.

3. Доказательство гипотезы Васильева

Скажем, что крестовый граф обладает *седловой ориентацией*, если можно ориентировать его рёбра таким образом, чтобы в каждой вершине V некоторые два противоположных ребра были ориентированы по направлению к вершине V, а другие два — по направлению от вершины V. На

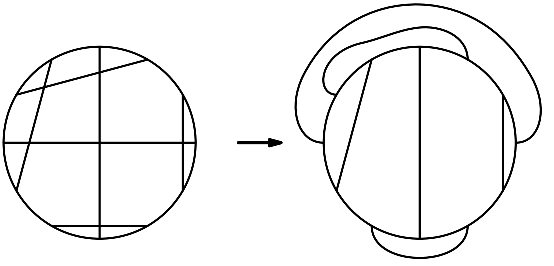


Рис. 6. d-диаграмма и вложение в плоскость

рисунке 7 слева изображен планарный крестовый граф и его седловая ориентация, а на рисунке справа — крестовый граф (с двумя вершинами), седловой ориентацией не обладающий.

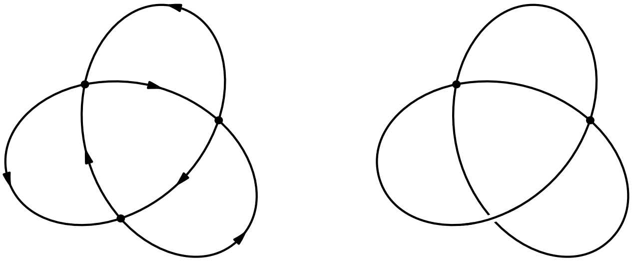


Рис. 7. Граф с седловой ориентацией и граф без седловой ориентации

Ясно, что если у связного крестового графа имеется седловая ориентация, то она единственная с точностью до одновременного обращения всех стрелок: стартуя с любой стрелки и двигаясь вдоль любого поворачивающего обхода, мы этот самый поворачивающий обход ориентируем в ту или другую сторону, и если в каждой вершине обход подходит сам к себе «правильно», то получаем седловую ориентацию. Иными словами, из определения седловой ориентации следует, что хордовая диаграмма, соответствующая некоторому (а следовательно, и любому) поворачивающему обходу крестового графа имеет все хорды первого типа.

Упражнение 2. У всякого плоского крестового графа имеется седловая ориентация.

Лемма 1. *Предположим, что крестовый граф* Γ *не обладает седловой ориентацией. Тогда на графе* Γ *найдется препятствие Васильева.*

Эту лемму мы оставляем читателю в качестве упражнения. Подсказка: попытайтесь из поворачивающего обхода построить седловую ориентацию и рассмотрите какую-либо из вершин, где этого не удается сделать.

Эта лемма сводит нашу задачу к случаю графов, обладающих седловой ориентацией.

Лемма 2. *Пусть крестовый граф* Γ *обладает седловой ориентацией. Пусть для некоторого его поворачивающего обхода соответствующая хордовая диаграмма является d-диаграммой. Тогда* Γ *планарен.*

Доказательство следует из рис. 8. Идея состоит в том, что каждая хорда вместе с двумя маленькими дугами в окрестности ее концов удаляется и заменяется на пару пересекающихся линий. Получающийся граф и будет искомым крестовым графом.

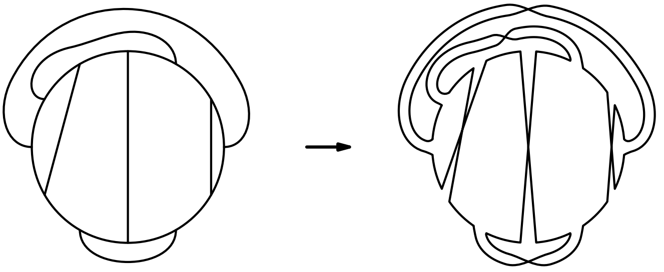


Рис. 8. От диаграммы к крестовому графу

Замечание 2. В лемме 2 можно заменить выражение «для некоторого обхода» на «для любого обхода». Действительно, планарность графа будет гарантировать, что хордовая диаграмма, соответствующая каждому поворачивающему обходу этого графа, является d-диаграммой.

Таким образом, для доказательства основной теоремы нам достаточно доказать следующее утверждение.

Лемма 3. *Пусть крестовый граф* Γ *обладает седловой ориентацией и пусть при этом для некоторого (и, следовательно, для любого) поворачивающего обхода этого графа соответствующая хордовая диаграмма не является d-диаграммой. Тогда граф* Γ *обладает препятствием Васильева.*

Заметим сначала, что если хордовая диаграмма не является d-диаграммой, то у нее найдется набор хорд, образующий «нечетноугольник», см. рис. 9. Это утверждение остается читателю в качестве упражнения.

Заметим, что если эта хордовая диаграмма ($(2n+1)$-угольник) соответствует поворачивающему обходу некоторого крестового графа Γ, у которого все вершины имеют первый тип, то на этом графе Γ легко выделяется препятствие Васильева: один цикл состоит из тех ребер графа, соответствующих дугам хордовой диаграммы, помеченным на рис. 9 буквой a, а

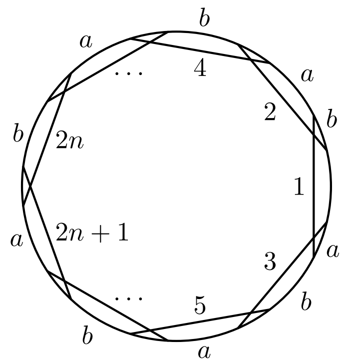


Рис. 9. Хордовая диаграмма $(2n+1)$-угольника

другой цикл (симметричный первому) состоит из ребер (дуг), помеченных буквой b. Эти два цикла имеют ровно одну точку пересечения, которая соответствует вершине 1 и является точкой перекрестья.

Рассмотрим теперь произвольный граф Γ', обладающий седловой ориентацией, хордовая диаграмма некоторого поворачивающего обхода которого содержит $(2n+1)$-угольник. Тогда два цикла, имеющие ровно одно перекрестье, легко переносятся на Γ': в вершинах графа Γ', которым соответствуют хорды, не принадлежащие $(2n+1)$-угольнику, каждый из этих двух циклов будет поворачивать, следовательно, других перекрестий не появится.

Лемма 3 доказана, что завершает доказательство основной теоремы.

4. Куда вложи́мы произвольные графы — постановка задачи

Предположим теперь, что крестовый граф не является планарным. Зададимся вопросом, в поверхность какого рода этот четырехвалентный граф может быть вложен с сохранением крестовой структуры. Этому посвящена работа [4].

5. Дальнейшие направления деятельности и нерешенные задачи

Следующая задача является переформулировкой задачи о вложениях крестовых графов.

1. Пусть дана матрица симметричная матрица M размера $n \times n$ над полем из двух элементов. Каково минимальное значение суммы рангов двух матриц $\operatorname{rank}(M_I) + \operatorname{rank}(M_J)$, где два подмножества I, J образуют

разбиение множества индексов исходной матрицы: $I \sqcup J = \{1, \ldots, n\}$, а квадратные матрицы M_I и M_J получаются из матрицы M взятием соответствующих множествам I и J наборов строк и столбцов?

Разумеется, эта задача может быть решена прямым перебором 2^{n-1} вариантов, однако, конечно, имеется в виду возможность нахождения ее быстрого (например, полиномиального по n решения).

Как именно эта задача связана с задачей о вложении крестовых графов, мы расскажем в следующей публикации. Скажем лишь, что сумма рангов 0 возможна лишь тогда, когда обе матрицы M_I и M_J нулевые. А это очень похоже на d-диаграммы — два семейства хорд, таких что хорды из одного семейства попарно не пересекаются.

2. Как мы видели, вопрос о планарности крестового графа удобным образом решается в терминах поворачивающего обхода. В терминах трансверсального обхода[1)] он также решается, см., напр., [7, 8]. Однако можно предложить следующий способ действий: имея трансверсальный обход, выбираем какой-нибудь поворачивающий обход и по нему определяем, является ли граф планарным.

Но здесь возникает вопрос о том, как представить себе процесс перерисовывания в виде удобной формулы. Оказывается, это можно легко сделать на языке матриц пересечения хордовых диаграмм, чему посвящена статья Д. П. Ильютко [2]. Благодаря формуле Ильютко можно, например, легко указать, как связаны хордовые диаграммы *разных* поворачивающих обходов одной и той же хордовой диаграммы, а также получить ряд важных комбинаторных и чисто алгебраических следствий.

Эти и смежные вопросы мы планируем обсудить в дальнейших публикациях. Продолжение следует.

Автор выражает благодарность В. А. Васильеву и М. Н. Вялому за полезные обсуждения.

Список литературы

[1] Васильев В. А. *Инварианты первого порядка и когомологии пространств вложений самопересекающихся кривых в* $\mathbb{R}^n$ // Известия РАН. Сер. Мат. Т. 69, №5, 2005. С. 3–52.

[2] Ильютко Д. П. *Оснащенные 4-графы: эйлеровы циклы, гауссовы циклы и поворачивающие обходы* // Матем. сбор., 2010. В печати.

[1)] Трансверсальный обход определяется так же, как и поворачивающий, с той лишь разницей, что в каждой вершине мы переходим на противоположное ребро. Трансверсальные обходы существуют лишь у уникурсальных графов (тех, которые можно нарисовать, не отрывая карандаша от бумаги и двигаясь гладко).

[3] Мантуров В. О. *Доказательство гипотезы Васильева о планарности сингулярных зацеплений* // Известия РАН. Т. 69, 2005. С. 169–178.

[4] Мантуров В. О. *Вложения оснащенных четырехвалентных графов в двумерные поверхности* // Доклады РАН. Т. 494, №3, 2009. С. 308–310.

[5] Мантуров В. О. *Теория узлов*. Москва-Ижевск: РХД, 2005. — 512 с.

[6] Прасолов В. В. *Элементы комбинаторной и дифференциальной топологии*. М.: МЦНМО, 2004. — 352 с.

[7] Cairns G., Elton D. *The planarity problem for signed Gauss words* // Journal of Knot Theory and Its Ramifications. Vol. 2, 1993. P. 359–367.

[8] Cairns G., Elton D. *The planarity problem*. II // Journal of Knot Theory and Its Ramifications. Vol. 5, 1996. P. 137–144.

В. О. Мантуров, механико-математический факультет МГУ

Задача об аддитивных цепочках и ее обобщения

С. Б. Гашков

В последние несколько десятков лет возрос интерес к алгоритмической стороне математики. Он проявляется, в частности, в постановке задач теоретико-сложностного характера в разнообразных областях, начиная с классического анализа и алгебры (например, [2, 7, 9, 10, 13, 16, 17, 20–23, 25]). Эти задачи отличаются друг от друга как по постановке, так и по методам, однако их большой круг объединяется возможностью формулировки в терминах схемной сложности вычисления функций.

1. Аддитивные цепочки и возведение в степень

Самой старой из относящихся к этой тематике задач является задача об аддитивных цепочках, рассматривавшаяся в тридцатые годы А. Шольцем и А. Брауэром. Идеи, ведущие к аддитивным цепочкам, в [9] прослежены до времен древних Египта и Индии. В последние двадцать лет аддитивные цепочки нашли применение в криптографических алгоритмах (см., например, [15]) и число публикаций на эту тему возросло настолько, что требует отдельного обзора (такой обзор уже написан [18]). Мы ограничимся некоторыми примерами.

Аддитивные цепочки возникают, например, в следующей олимпиадной задаче.

За какое наименьшее количество взвешиваний на чашечных весах можно отвесить один килограмм сахарного песку, если имеется лишь одна одниграммовая гирька?

На первый взгляд кажется, что единственный способ решения этой задачи — отвесить один грамм, положить в эту же чашку гирьку, отвесить в другой чашке два грамма, переложить гирьку в нее, и так далее, добавляя по одному грамму, после тысячного взвешивания отмерить наконец-то килограмм.

Но есть и более быстрый способ. Нужно лишь заметить, что если мы научились отвешивать за k взвешиваний m грамм, то, сделав еще одно взвешивание, можно, даже не используя гирьку, отвесить еще m грамм и, ссыпав обе порции вместе, получить $2m$ грамм за $k+1$ взвешивание. А если при этом взвешивании положить на одну из чашек гирьку, то за $k+1$ взвешивание можно отвесить $2m \pm 1$ грамм.

Если нужно отмерить n грамм, то можно записать n в двоичном виде $(a_m \dots a_1)$, где $2^{m-1} \leqslant n < 2^m$, $a_m = 1$ и воспользоваться формулой

$$n = a_m 2^{m-1} + \dots + a_2 2 + a_1 = (\dots((2a_m + a_{m-1})2 + a_{m-2})\dots)2 + a_1,$$

последовательно отвешивая по

$$b_1 = a_m,\ b_2 = 2b_1 + a_{m-1},\ b_3 = 2b_2 + a_{m-2},\ \dots,\ b_m = b_{m-1}2 + a_1 = n$$

граммов (можно использовать и двоичную запись с отрицательными цифрами).

В используемой формуле читатели увидят схему Горнера. Она будет встречаться у нас и далее.

Идея, лежащая в основе этого метода взвешивания, стара, как сама математика. Ее применяли и древние египтяне, и древние индусы, но конечно, не для взвешивания, а для умножения — алгоритм умножения столбиком был придуман не сразу. А до этого умножение сводилось к сложению и удвоению. Такой метод умножения дожил почти до нашего времени, он удобен при вычислениях на счетах. Сейчас он никому не нужен — счеты вытеснены калькуляторами. Но как возвести на калькуляторе число a в тысячную степень, если у него нет специальной операции возведения в степень? Умножать 999 раз не нужно, а можно применить совершенно тот же прием, последовательно вычисляя

$$a^3 = a^2 a,\ a^7 = (a^3)^2 a,\ a^{15} = (a^7)^2 a,\ a^{31} = (a^{15})^2 a,$$
$$a^{62} = (a^{31})^2,\ a^{125} = (a^{62})^2 a, a^{250} = (a^{125})^2,\ a^{500} = (a^{250})^2,\ a^{1000} = (a^{500})^2.$$

Если вспомнить, что 1000 имеет двоичную запись 1111101000, то можно заметить, что если отбросить старший бит (равный единице), то каждому следующему биту соответствует операция возведения в квадрат, если он нулевой, или возведение в квадрат с последующим умножением на число a— основание степени. Число a не нужно каждый раз заново набирать на клавиатуре. Можно в начале вычислений занести его в память, и когда нужно, после нажатия кнопки для умножения, вызывать его из памяти и потом нажимать кнопку «равно». Посчитаем общее число операций умножения в рассмотренном вычислении. Число возведений в квадрат на единицу меньше длины двоичной записи показателя степени, а число умножений общего вида на единицу меньше суммы цифр двоичной записи.

Для любого n обозначим $\lambda(n)$ уменьшенную на единицу длину двоичной записи числа n, а $\nu(n)$ — ее сумму цифр (т. е. число единиц в ней). Тогда в общем случае число операций умножения, использованных в этом методе[1)] возведения в степень n, равно $\lambda(n) + \nu(n) - 1$.

[1)] Он называется *бинарным.*

Покажем, что меньшим числом операций обойтись нельзя, если только не обновлять содержимое ячейки памяти. Для удобства будем рассматривать изменение не самих степеней, а их показателей, которые после каждой операции будут складываться. Обозначим показатель степени у числа, находящегося в регистре, через a_i (в начальный момент $a_0 = 1$), тогда изменяться он может одним из двух следующих способов: $a_{i+1} = 2a_i$ или $a_{i+1} = a_i + 1$ (в первом случае содержимое регистра возводится в квадрат, во втором случае оно умножается на содержимое ячейки памяти, а результат опять записывается в регистр). Очевидно, что в первом случае $\nu(a_{i+1}) = \nu(a_i)$, $\lambda(a_{i+1}) = \lambda(a_i) + 1$, откуда

$$\nu(a_{i+1}) + \lambda(a_{i+1}) = \nu(a_i) + \lambda(a_i) + 1.$$

Во втором случае $\nu(a_{i+1}) \leqslant \nu(a_i) + 1$ (равенство возможно, только если a_i четно), $\lambda(a_{i+1}) \leqslant \lambda(a_i) + 1$ (равенство возможно, только если a_i нечетно, точнее, когда $a_i = 2^{\lambda(a_i)} - 1$), поэтому одновременно эти равенства не выполняются; значит, во втором случае

$$\nu(a_{i+1}) + \lambda(a_{i+1}) \leqslant \nu(a_i) + \lambda(a_i) + 1,$$

причем равенство возможно лишь когда a_i четно или $a_i = 1$. Так как $\lambda(1) + \nu(1) = 1$, то отсюда с помощью индукции выводится, что $\nu(a_l) + \lambda(a_l) \leqslant l+1$, причем равенство возможно только когда прибавление единицы всегда выполняется после одного или нескольких удвоений (можно считать, что на первом шаге всегда происходит удвоение), при этом после каждого прибавления единицы $\nu(a_i)$ увеличивается на единицу, значит $\nu(a_l) - 1$ равно числу выполненных при вычислении a_l прибавлений единицы. Значит,

$$\nu(n) + \lambda(n) = \nu(a_{l(n)}) + \lambda(a_{l(n)}) \leqslant l(n) + 1,$$

где $l(n)$ число операций умножения для возведения в n-ю степень. Поэтому $l(n) \geqslant \nu(n) + \lambda(n) - 1$. Аналогично можно показать, что число операций умножения регистра на постоянное число из памяти не меньше $\nu(n) - 1$. Далее, равенство $l(n) = \nu(n) + \lambda(n) - 1$ возможно лишь когда прибавление единицы всегда выполняется после одного или нескольких удвоений, при этом число прибавлений единицы равно $\nu(n) - 1$ и n представимо в виде

$$n = 2^{\lambda(n)} + 2^{\alpha_{\nu(n)-1}} + \ldots + 2^{\alpha_1},$$

где α_i строго возрастают. Такое представление возможно только одно, и поэтому существует только один способ возведения в n-ю степень в указанных условиях — указанный выше бинарный метод.

Но если разрешается обновлять содержимое ячейки памяти (используя для этого пересылки из других ячеек), то бинарный метод вычисления x^n в некоторых случаях можно улучшить. Для этого, например, можно применить *метод множителей.* Его идея заключается в следующем. Если мы умеем возводить в степень n за $l(n)$ операций, и возводить в степень m за $l(m)$ операций, то можно, после того как закончено вычисление x^n, занести его в ячейку памяти, и далее вычислить $x^{nm} = (x^n)^m$ за $l(m)$ операций, используя тот же метод, что и для вычисления x^m. Тогда общее число операций будет равно $l(nm) = l(n) + l(m)$.

Например, вычисляя x^5 бинарным методом за 3 операции и применяя два раза метод множителей, получаем, что $l(125) = 3l(5) = 9$. Выполняя еще три возведения в квадрат, имеем $l(1000) = l(125) + 3 = 12$. Бинарный метод требовал $\lambda(1000) + \nu(1000) - 1 = 9 + 6 - 1 = 14$ операций.

Заметим еще, что существует и другой вариант бинарного метода, с тем же числом возведений в квадрат и с тем же общим числом умножений, но для его выполнения требуется большое число ячеек памяти.

Что же такое *аддитивная цепочка*? Это любая, начинающаяся с 1, последовательность натуральных чисел $a_0 = 1, a_1, \ldots, a_m$, в которой каждое число является суммой каких-то двух предыдущих чисел (или удвоением какого-то предыдущего числа). Обозначим $l(n)$ наименьшую длину аддитивной цепочки, заканчивающейся числом n. Длиной цепочки $a_0 = 1, a_1, \ldots, a_m$ называем число m. Например, $1, 2, 3, 5, 7, 14$ — минимальная цепочка для 14, т. е. $l(14) = 5$.

Аддитивные цепочки можно изображать в виде ориентированного графа, в котором в вершину a_i идут рёбра от вершин a_j, a_k, если $a_i = a_j + a_k$ (если такое представление неоднозначно, выбираем любое из них и рисуем только два ребра). Можно считать, что все числа в цепочке разные, просто удаляя из нее повторяющиеся числа, и располагать числа в цепочке в порядке возрастания.

Граф для предыдущего примера см. на рис. 1.

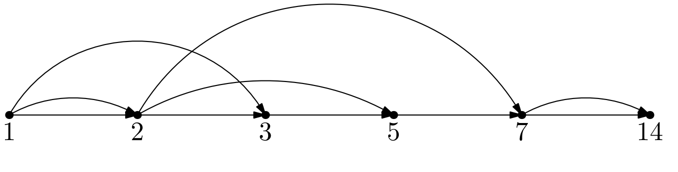


Рис. 1.

Очевидно, что наименьшее число умножений, необходимое для возведения в n-ю степень, равно $l(n)$.

Бинарный метод дает оценку $l(n) \leqslant \lambda(n) + \nu(n) - 1$. Методом множителей — оценку $l(nm) \leqslant l(n) + l(m)$. Справедлива и нижняя оценка $l(n) \geqslant \lambda(n)$. Из нее следует, что $l(2^n) = n$. Для доказательства оценки $l(n) \geqslant \lambda(n)$ достаточно заметить, что для любой аддитивной цепочки справедливы неравенства $a_i \leqslant 2^i$, в чем можно убедится, проведя индукцию.

База индукции очевидна, а индукционный шаг обосновывается неравенством
$$a_i = a_j + a_k \leqslant 2^j + 2^k \leqslant 2^{i-1} + 2^{i-1} = 2^i.$$
Из неравенства $n = a_{l(n)} \leqslant 2^{l(n)}$ следует, что $l(n) \geqslant \lceil \log_2 n \rceil = \lambda(n)$, символ $\lceil x \rceil$ в в последней формуле означает наименьшее целое число, не меньшее x. Более тонкие нижние оценки читатель может найти [9] (вместе с массой другой информации об аддитивных цепочках и не только о них), но они доказываются непросто, а по своей точности ненамного превосходят доказанную почти очевидную оценку.

Интересно, что бинарный метод был по существу известен древним индусам, потом был переоткрыт арабскими математиками, но задача о точном вычислении функции $l(n)$ появилась (согласно [9]) в одном французском журнале в 1894 г., потом заново была переоткрыта в тридцатые годы в Германии, и неоднократно переоткрывалась в дальнейшем[2)], но до сих пор в общем случае не решена. Не известно даже, существует ли алгоритм полиномиальной сложности[3)] для вычисления функции $l(n)$. Не решены также многие другие задачи об аддитивных цепочках. Например, неизвестно, верно ли равенство
$$l(2^n - 1) = n + l(n) - 1$$
(гипотеза Шольца). Некоторые естественные гипотезы об аддитивных цепочках оказались неверны. За информацией об всем этом мы отсылаем читателя к [9] — единственной пока монографии на русском языке, где есть раздел, посвященный аддитивным цепочкам.

Наилучшая из общих верхних оценок была доказана в тридцатые годы А. Брауэром и имеет вид
$$\lambda(n)\left(1 + \frac{1}{\lambda(\lambda(n))} + \frac{O(\lambda(\lambda(\lambda(n))))}{(\lambda(\lambda(n)))^2}\right).$$
Она вытекает из следующей теоремы, если в ней положить $k = \lambda(\lambda(n)) - 2\lambda(\lambda(\lambda(n)))$.

Теорема 1 (А. Брауэр). *При $k < \log_2 \log_2 n$ справедливо неравенство*
$$l\,(n) < (1 + 1/k)\,\lceil \log_2 n \rceil + 2^{k-1} - k + 2.$$

Доказательство. Представим n в двоичной записи
$$n = \sum_{i=0}^{m} \alpha_i 2^i,$$

[2)] Например, на рубеже 50-х и 60-х годов двадцатого века в Москве Р. Вальским.

[3)] Это означает, что время работы алгоритма ограничено некоторым полиномом от $\log n$.

Рис. 2. Альфред Брауэр

где $\alpha_i = 0$ или 1, $m = \lfloor \log_2 n \rfloor$. Выделим в наборе $(\alpha_0, \ldots, \alpha_m)$ не более чем $\lceil \frac{m+1}{k} \rceil$ непересекающихся блоков $A_0, \ldots, A_s$, $s < \lceil \frac{m+1}{k} \rceil$, следующим образом. Каждый блок состоит из не более чем из k подряд идущих цифр и заканчивается единицей, кроме блока A_s, состоящего из старших k цифр, а вне блоков стоят только нули. Числа $a_0, \ldots, a_s$, двоичными записями которых являются эти блоки, не превосходят $2^k - 1$ и все нечетны, кроме возможно a_s. Тогда n можно представить в виде

$$n = 2^{l_0} \left(2^{l_1} \ldots \left(2^{l_{s-2}} \left(2^{l_{s-1}} \left(2^{l_s} a_s + a_{s-1} \right) + a_{s-2} \right) + \ldots + a_1 \right) + a_0 \right),$$

где $l_s + l_{s-1} + \ldots + l_0 = m + 1 - k$. Все числа $a_0, \ldots, a_{s-1}$ содержатся в аддитивной цепочке $1, 2, 3, 5, 7, \ldots, 2^k - 1$ длины $2^{k-1} + 1$. Если число a_s не содержится в этой цепочке, то его можно поставить в ее конец, вычислив как $a_s = (a_s - 1) + 1$. Поэтому для вычисления n достаточно добавить к этой цепочке последовательность

$$a_s, 2a_s, 4a_s, \ldots, 2^{l_s} a_s, 2^{l_s} a_s + a_{s-1},$$

$$2 \left(2^{l_s} a_s + a_{s-1} \right), 4 \left(2^{l_s} a_s + a_{s-1} \right), \ldots,$$

$$2^{l_{s-1}} \left(2^{l_s} a_s + a_{s-1} \right), 2^{l_{s-1}} \left(2^{l_s} a_s + a_{s-1} \right) + a_{s-2}, \ldots,$$

$$2^{l_{s-2}} \left(2^{l_{s-1}} \left(2^{l_s} a_s + a_{s-1} \right) + a_{s-2} \right),$$

. . . .

$$2^{l_1} \left(\ldots 2^{l_{s-2}} \left(2^{l_{s-1}} \left(2^{l_s} a_s + a_{s-1} \right) + a_{s-2} \right) + \ldots + a_1 \right) + a_0, \ldots,$$

$$2^{l_0} \left(2^{l_1} \left(\ldots 2^{l_{s-2}} \left(2^{l_{s-1}} \left(2^{l_s} a_s + a_{s-1} \right) + a_{s-2} \right) + \ldots + a_1 \right) + a_0 \right),$$

длина которой равна

$$l_s + l_{s-1} + \ldots + l_0 + s + 1 = m + 2 + s - k.$$

Поэтому

$$l(n) < 2^{k-1} + 1 + m + 2 + s - k < m + 2 + \left\lceil \frac{m+1}{k} \right\rceil + 2^{k-1} - k.$$

Можно считать, что $n \neq 2^m$, тогда $m + 1 = \lceil \log_2 n \rceil$ и

$$l(n) < \lceil \log_2 n \rceil (1 + 1/k) + 2^{k-1} - k + 2.$$

Понятие аддитивной цепочки имеет некоторые естественные обобщения. Например, изучались цепочки с вычитаниями. Можно рассматривать цепочки с различными ограничениями, например цепочки, в которых запрещены удвоения, или цепочки, в которых разрешается сложения только следующего типа $a_{i+1} = a_i + a_k$, но удвоения разрешены. Такие цепочки называются линейными, о них в [9] доказаны интересные теоремы. Можно рассматривать также различные меры сложности аддитивных цепочек, отличные от их длины. Одно из естественных обобщений аддитивных цепочек будет рассмотрено в следующем разделе.

2. Сложность вычисления сумм

Введем следующее определение. *Векторная аддитивная цепочка* — это последовательность векторов, начинающаяся с единичных базисных векторов $(1, 0, \ldots, 0)$, $\ldots$, $(0, \ldots, 0, 1)$, в которой каждый следующий вектор получается сложением двух предыдущих векторов (не обязательно непосредственно предшествующих), или удвоением какого-то предыдущего вектора (т. е. прибавлением его к самому себе). Сложностью системы векторов называется длина (без учета базисных векторов) кратчайшей цепочки, содержащей эту систему (если в системе некоторые вектора совпадают, то требуется,чтобы в цепочке содержались все различные вектора, встречающиеся в системе). Обозначим $L(p, q, N)$ наибольшую сложность системы, состоящих из p векторов размерности q, компоненты которых принадлежат множеству $0, 1, \ldots, N - 1$. Эта величина совпадает с наибольшей *аддитивной сложностью* систем p линейных форм от q переменных с натуральными коэффициентами, меньшими N, и с наибольшей *мультипликативной сложностью* систем из p одночленов от q переменных,входящих в них в степенях, меньших N. Под аддитивной и мультипликативной сложностью здесь понимается сложность реализации соответствующей вектор-функции схемами из функциональных элементов в базисах $\{x + y\}$ и $\{x \cdot y\}$ соответственно (см. соответствующие определения в следующем разделе).

Н. Пиппенджер [24] показал, что при $w < N^{o(v)}$, где $w = \max(p, q)$, $v = \min(p, q)$,

$$L(p, q, N) = v \log N + \frac{H}{\log H} \left(1 + O\left(\frac{\log\log H}{\log H} \right)^{1/2} \right) + O(w),$$

где $H = pq \log N$. Отсюда следует, в частности, довольно неожиданная теорема Яо о том, что длина кратчайшей аддитивной цепочки, содержащей несколько заданных чисел, асимптотически равна длине кратчайшей цепочки, вычисляющей наибольшее из них, если только этих чисел не слишком много (см. [9]).

Оценки Пиппенджера в [4] дополнены и уточнены следующим образом: $L(p, q, N) + p = L(q, p, N) + q$, если $w \geqslant N^v - v - 1$, то $L(p, q, N) = N^v - v - 1 + (w - p)$, если $w < N^v - v - 1$, то

$$L(p, q, N) = v \log N + \frac{H}{\log H} \left(1 + O\left(\frac{\log\log H}{\log H} \right)^{1/2} \right) + \left(2 + \frac{1}{\log H} \right) w - p.$$

причем при $v \log N \geqslant \dfrac{H \log\log H}{\log^2 H}$ справедливо равенство

$$L(p, q, N) = v \log N + \frac{H}{\log H} \left(1 + O\left(\frac{\log\log H}{\log H} \right) \right),$$

если $w < N^v$, то

$$p + L(p, q, N) \leqslant \frac{H}{\log H} \left(1 + O\left(\frac{\log\log H}{\log H} \right)^{1/2} \right) + (2w + v \log N) \left(2 + \frac{1}{\log H} \right).$$

Частным случаем рассмотренной задачи является задача вычисления системы степеней одной переменной $x^{n_1}, \ldots, x^{n_p}$, поставленная Д. Кнутом [9]. Обозначим $l(n_1, \ldots, n_p)$ соответствующую сложность вычисления системы линейных форм от одной переменной. Положим $N = \max n_i$, $H = \log(n_1 \ldots n_p)$. Тогда приведенные выше результаты можно уточнить следующим образом [4]:

$$l(n_1, \ldots, n_p) \leqslant \log N + \frac{H}{\log H} \left(1 + O\left(\frac{\log\log H}{\log H} \right)^{1/2} \right) + O(p).$$

В книге для подготовки к математическим олимпиадам [12] в 13-й главе имеется подборка задач «Сложность суммирования». В ней идет речь о сложности вычисления линейных преобразований с булевыми матрицами над полем $\{0, 1\}$ в базисе, состоящем из одной операции $x \oplus y$ — сложения по модулю два. Повторять формулировки этих задач вряд ли здесь уместно, но имеет смысл привести формулировки некоторых более общих утверждений из [4]. Величина $L(p,q,N)$ далее та же, что и выше.

Справедливы неравенства

$$L(p,q,N) \leqslant L(p,q\lceil\log_2 N\rceil,2),$$

и для любого $s \leqslant q$

$$L(p,q,N) \leqslant (s-1)p + sN^{\lceil q/s\rceil}.$$

Второе неравенство доказывается применением вентильной конструкции Лупанова [11]. Формально этого неравенства в [11] нет, там есть подобные неравенства для сложности вентильных схем, но из оценки для вентильных схем вытекает оценка для рассматриваемых схем в силу почти очевидного неравенства $C(p,q,N) \geqslant L(p,q,N)+p$, и конструкция для вентильных схем без существенных изменений переносится и на рассматриваемые нами схемы. Чтобы не отвлекаться в сторону, мы не будем здесь рассматривать вентильные схемы. В чем состоит эта конструкция, можно понять, посмотрев в [4] или [11], но можно и прочитать решения соответствующих задач из [12]. Строго говоря в [11] рассматривался случай $N=2$ и рассматривались, так сказать, дизъюнктивные вентильные схемы, а не вентильные схемы по модулю два, которые более подходят к рассматриваемой ситуации, но принципиального значения это не имеет. В случае $N=2$ из указанного неравенства при выборе $s = q/(\log_2 p - \log_2\log_2 p) + O(1)$ имеем

$$L(p,q,2) \leqslant \frac{pq}{\log_2 p}\left(1+\frac{O(\log_2\log_2 p)}{\log_2 p}\right).$$

Если $q>p$, то в силу равенства $L(p,q,2)+p = L(q,p,2)+q$ это неравенство можно усилить до

$$L(p,q,2) \leqslant \frac{pq}{\log_2 q}\left(1+\frac{O(\log_2\log_2 q)}{\log_2 q}\right).$$

В общем случае, если положить $w=\max(p,q), v=\min(p,q)$, то

$$L(p,q,2) \leqslant \frac{wv}{\log_2 w}\left(1+\frac{O(\log_2\log_2 w)}{\log_2 w}\right).$$

Заметим, что из упомянутой выше теоремы Пиппенджера следует асимптотически точный результат

$$L(p,q,2) = v + \frac{H}{\log H}\left(1+O\left(\frac{\log\log H}{\log H}\right)^{1/2}\right) + O(w),$$

где $H=pq$.

2.1. Лемма о транспонировании матриц

Простейший вариант этой леммы содержится в формулировке задачи 16 из цикла «Сложность суммирования» 13-й главы книги [12]. Сформулировать ее можно следующим образом.

Пусть A — матрица из нулей и единиц, все m строк и n столбцов которой ненулевые. Обозначим $L(A)$ сложность вычисления определяемого этой матрицей линейного преобразования AX в базисе $\{x+y\}$. Тогда $L(A) + m = L(A^T) + n$, где A^T — транспонированная матрица.

История многократно переоткрывавшейся леммы о транспонировании[4)] восходит, как утверждают французские авторы, к работам Теллегена об электрических цепях. Исторический обзор имеется в [19]. В нем выражается сомнение в правомерности приписывания чести открытия этой леммы Теллегену, и указывает, что вероятно, она принадлежит Фидуччиа (1973 г.), который ее распространил и на схемы для билинейных преобразований. Разные варианты леммы о транспозиции и некоторые их применения были указаны в 1980 – 1981 годах в работах Оливоса и Кнута – Пападимитриу, изложение которых можно найти в [9]. В 1981 году лемму о транспозиции получил также в [14] А. Ф. Сидоренко.

Укажем некоторые применения леммы о транспозиции (имеющиеся в [4, 12]). Первое заключается в выводе равенства $L(p, q, N) + p = L(q, p, N) + q$ с помощью которого в [4] получены некоторые уточнения результатов Пиппенждера [24], причем так называемый «трудный случай» теоремы Пиппенджера сводится к несложно доказываемой методом Брауэра оценке Страуса (см. [4] или [9])

$$L(p, q, N) \leqslant p\left(1 + \frac{q}{t}\right)\lceil \log_2 N\rceil + 2^t q + pq.$$

Та же идея позволяет легко доказать упоминавшуюся выше теорему Яо. Фактически так же, но без явного применения леммы о транспозиции, теорема Яо доказана в [9].

Второе применение леммы о транспозиции, указанное в [4], связано с выводом равенства $L(B_n) = 2^{n+1} - 2(n+1)$, где B_n есть $(n, 2^n - 1)$-матрица, столбцы которой образованы всевозможными различными наборами 0 и 1 длины n. Для транспонированной матрицы B_n^T равенство $L(B_n^T) = 2^n - n - 1$ очевидно. Нужное нам равенство следует из него с помощью леммы о транспозиции. Матрица B_n интересна тем, что она определяет матрицу оператора поиска ошибки в коде Хэмминга (если ее рассматривать над полем из двух элементов). Если из из матрицы B_n выбросить все столбцы с одной единицей, то получится матрица кодирования для кода Хэмминга. Ее сложность находится аналогично, и она равна $2^{n+1} - 3n - 2$.

[4)] И автор этой статьи тоже ее переоткрыл в 1985 г.

3. Схемная сложность вычисления функций

Здесь будут дано общее определение понятия схемной сложности и указаны некоторые конкретные примеры.

Базисом называется произвольное множество B функций (операций) $\omega : E^n \to E$, чьи аргументы и значения принадлежат любому заданному множеству E (содержащему хотя бы два элемента).

Схемой в базисе B называется произвольная последовательность функций $f_1(X), \ldots, f_{L+n}(X)$, $X = (x_1, \ldots, x_n)$, в которой первые n функций определяются равенствами $f_i(X) = x_i$, а каждая следующая функция $f_l(X)$ вычисляется через некоторые из предыдущих с помощью одной из базисных операций w_k, $k = k(l)$. Схемы иногда называют *неветвящимися программами*. *Сложность* схемы — это число L. Функция f *реализуется* схемой S, если f равна какой-то из функций f_i схемы S. *Сложностью* (*схемной реализации*) функции f назовем число $L_B(f)$, равное наименьшей из сложностей схем, реализующих f. Все введенные определения естественно распространяются на случай реализации *вектор-функций* (*операторов* $E^n \to E^m$). Можно также перенести эти определения на случай базисов с произвольными неотрицательными весами элементов. Сложностью схемы тогда называется сумма весов, входящих в схему базисных элементов.

Случай $E = \{0, 1\}$ относится к алгебре логики, где впервые в массовом количестве появились задачи о сложности вычисления функций (см., например, [10]). Функции, о которых в этом случае идет речь, называются булевыми. В случае $E = \{0, 1, \ldots, k-1\}$ при $k > 2$ речь идет о сложности вычисления функций многозначной логики.

В случае $E = [0, 1]$ или $E = \mathbb{R}$ можно рассматривать и базисы, состоящие из непрерывных функций. Если в качестве B взять $\{x-y, x+y, xy\} \cup \mathbb{R}$, то в терминах сложности схем в этом базисе можно сформулировать многие результаты алгебраической теории сложности, например результаты о сложности вычисления многочленов, в том числе и об аддитивной и мультипликативной сложности, о сложности так называемых параллельных вычислений, о сложности умножения матриц и т. д. (см., например, [2, 9, 20, 22, 23]).

Очевидно, что понятие схемы в случае базиса, состоящего из операции сложения и единицы, превращается в понятие аддитивной цепочки.

Приведем примеры задач, на первый взгляд не связанных с рассматриваемой постановкой, но которые допускают переформулировку в ее терминах. Некоторые из этих задач фактически были известны довольно давно.

Таковой является известная в фольклоре задача о построении последовательно-параллельных схем (так называемых П-схем) из единичных резисторов, имеющих заданное сопротивление и содержащих минимально возможное их число. Эта задача сводится к задаче о вычислении меры сложности $L_B(r)$ для $r \in \mathbb{Q}_+$ и $B = \{x + y, 1/(1/x + 1/y)\}$. Распространенное мнение, будто она легко решается разложением r в цепную дробь, неверно; надо использовать так называемые *ветвящиеся цепные дроби*, и все же явно вычислить эту меру сложности не удается. Интересные результаты по этой задаче получил О. М. Касим-Заде [8] (см. также [3, 5]).

Другим примером является задача о нахождении минимального числа прямых и окружностей, которые надо провести циркулем и линейкой (или только одним циркулем, как в построениях Мора – Маскерони), чтобы выполнить данное планиметрическое построение. Эта задача была поставлена в 19-м веке Лемуаном, но в определенном смысле ее история начинается в Древней Греции. При некоторых естественных ограничениях на проводимые построения она сводится к вычислению (с точностью до порядка) меры сложности $L_B(F)$ вектор-функции F в базисе $B = \{x \pm y, xy, x/y, \sqrt{x}\}$. Например, если задан единичный отрезок, и надо построить одну точку, то в качестве F можно взять вектор, составленный из двух положительных констант. Если положить веса рациональных операций в базисе B равными нулю, то получим задачу о так называемой иррациональной сложности[5)].

В заключение рассмотрим вопрос о сложности геометрических построений более подробно.

4. Сложность геометрических построений

Допустим, что на плоскости дано некоторое множество точек, прямых и окружностей, которое обозначим M_0. Назовем *построением циркулем и линейкой* при заданном M_0 любую последовательность множеств M_0, $M_1, \ldots, M_L$, начинающуюся с M_0, и такую, что каждое следующее множество M_{i+1} получается из предыдущего множества M_i добавлением либо некоторой прямой, проходящей через какие-то две точки из множества M_i, либо окружности с центром в какой-то из точек множества M_i и радиусом, равным длине некоторого отрезка с концами в точках из M_i, а также всех точек пересечения добавленной линии со всеми линиями из множества M_i. Число L назовем *сложностью этого построения. Сложностью построения множества* M точек, отрезков и окружностей и прямых при заданном M_0 назовем минимальную сложность такого построения M_0, $M_1, \ldots, M_L$, для которого множество M_L содержит все прямые и окружности из M,

[5)]Однако высказанное в [6] утверждение о тождестве иррациональной сложности и наименьшего числа применений циркуля в планиметрическом построении неверно, так как согласно теореме Штейнера любое построение можно выполнить одной линейкой, если задана окружность с центром.

все точки из M и концы всех отрезков из M. Аналогично определяется *сложность построения одним циркулем.*

Большая подборка задач о сложности геометрических построений приведена в [5]. Далее приводятся некоторые задачи из [5].

Задача 1. Пусть дан единичный отрезок (точнее, даны только его концы). Для любого натурального n можно построить одним циркулем отрезки длины n и $1/n$ сложностью не более

$$3\log_3 n + \frac{\log_3 n}{\log_3\log_3 n}\left(1 + \frac{2\log_3\log_3\log_3 n + O(1)}{\log_3\log_3 n}\right).$$

Сложность построения отрезка длины n одним только циркулем не меньше $\lfloor \log_\varphi (n + 1/2)\rfloor$, где $\varphi = \dfrac{\sqrt{5}+1}{2}$. Можно разделить отрезок на n равных частей одним циркулем со сложностью не больше

$$n + 3\log_3 n + \frac{\log_3 n}{\log_3\log_3 n}\left(1 + \frac{2\log_3\log_3\log_3 n + O(1)}{\log_3\log_3 n}\right).$$

Сложность разделения отрезка одним циркулем на n равных частей не меньше $n-1$.

Указание. Доказательство основано на использовании метода аддитивных цепочек, троичной системы и следующей леммы:

можно построить со сложностью не более $4n$ одном циркулем такое множество точек, что среди отрезков с концами в этом множестве найдутся отрезки с длинами $1, 2, \ldots, n^2$.

Заметим, что если отрезок дан полностью (а не только его концы), то в оценке сложности задачи разделения его на n частей можно первое слагаемое n заменить на $n/2$.

В случае n, равного степени двойки, утверждение задачи 1 можно существенно усилить.

Задача 2. Пусть задан единичный отрезок. Можно построить циркулем отрезок длины 2^{-n} со сложностью не более

$$2\log_2 n + \frac{\log_2 n}{\log_2\log_2 n} + \frac{(2\log_2 n)\,(\log_2\log_2\log_2 n + O(1))}{(\log_2\log_2 n)^2}.$$

При $n = 2^m$ этот отрезок можно построить со сложностью $2\log_2 n + 4$. Можно разделить данный отрезок на 2^n равных частей одним циркулем со сложностью не более $2^{n-1} + 9$, если дан весь отрезок, а не только его концы.

Указание. Доказательство основано на использовании аддитивных цепочек и следующей леммы:

если дан отрезок длины $x < 1$, то можно построить со сложностью не более $4n$ одним циркулем такое множество точек, что среди отрезков с концами в этом множестве найдутся отрезки длиной x, x^2, …, x^{n^2}.

При использовании линейки сложность построения, указанная в задаче 1, может быть уменьшена.

Задача 3. Можно построить циркулем и линейкой отрезки длины n и $1/n$ со сложностью не более

$$\frac{\log_2 n}{2\log_2\log_2 n}\left(1+\frac{3\log_2\log_2\log_2 n+O(1)}{\log_2\log_2 n}\right)$$

и разделить отрезок циркулем и линейкой на n равных частей со сложностью не более

$$\frac{n}{2}+\frac{\log_2 n}{2\log_2\log_2 n}\left(1+\frac{3\log_2\log_2\log_2 n+O(1)}{\log_2\log_2 n}\right).$$

Сложность разделения отрезка циркулем и линейкой на n равных частей не меньше $\lceil (n-1)/2\rceil$. Для некоторой бесконечной последовательности чисел n сложность построения отрезка длины n циркулем и линейкой больше

$$\frac{1}{6}\frac{\log_2 n}{\log_2\log_2 n}\,.$$

Указание. Доказательство верхней оценки основано на использовании того факта что сложность вычисления произвольного числа n в базисе $\{x+y, xy, 1\}$

$$\tau(n)\leqslant\frac{\log_2 n}{\log_2\log_2 n}\left(1+\frac{3\log_2\log_2\log_2 n+O(1)}{\log_2\log_2 n}\right),$$

и следующей леммы:

можно построить со сложностью не более $4n$ *одном циркулем такое множество точек, что среди отрезков с концами в этом множестве найдутся отрезки с длинами* $1, 2, \ldots, n^2$.

Доказательство нижней оценки основано на мощностных соображениях.

В [1] часть приведенных выше результатов была независимо получена (в несколько более слабом виде), причем один из них был предложен в качестве задачи[6)] на Всероссийской олимпиаде школьников. Но в [1] получены и несколько новых интересных результатов, из которых, в частности, следует, что оценка, указанная в задаче 2, по порядку неулучшаема. А именно, там показано существование такой константы c, что если отрезок длины a построен со сложностью l, то $a < 2^{c^l}$. Если добавить к рассуждениям [1] использование леммы Ландау – Миньотта, то можно получить оценку $L(a) < 2^{c^l}$, где $L(a)$ — сумма модулей коэффициентов минимального многочлена с целыми коэффициентами, корнем которого является число a. Фактически, это утверждение в [1] есть, но доказательство

[6)] Второй автор — А. Я. Белов — задачу на олимпиаду предлагал, а первый автор — М. В. Алехнович — будучи во время олимпиады школьником, ее решал.

не приведено (но доказано, что степень этого многочлена не больше 2^l). Если далее воспользоваться известными в теории трансцендентных чисел теоремами о точности аппроксимации числа π алгебраическими числами данной степени и высоты, то можно доказать, что если $|\sqrt{\pi} - a| < \epsilon$, то $l > \Theta(\log_2 \log_2 1/\epsilon)$, т. е. сложность приближенного решения квадратуры круга по порядку не меньше двойного логарифма от $1/\epsilon$, где ϵ — относительная погрешность построения стороны квадрата, равновеликого данному кругу. Используя алгоритм вычисления π из [21], можно выполнить приближенное решение квадратуры круга со сложностью $O(\log_2 \log_2 1/\epsilon)$ (см. [5]). Значит, эта оценка по порядку неулучшаема.

Список литературы

[1] Алехнович М. В., Белов А. Я. *Сложность алгоритмов при построении циркулем и линейкой* // Фундаментальная и прикладная математика. Т. 7, №2, 2001. С. 597–614.

[2] Ахо А., Хопкрофт Д., Ульман Д. *Построение и анализ вычислительных алгоритмов*. М.: Мир, 1979.

[3] Гашков С. Б. *Алгоритм Евклида, цепные дроби, числа Фибоначчи и квадрирование прямоугольников* // Математическое просвещение. Сер. 2, вып. 6, 2002. С. 93-115.

[4] Гашков С. Б., Кочергин В. В. *Об аддитивных цепочках векторов, вентильных схемах и сложности вычисления степеней* // Методы дискретного анализа в теории графов и сложности. Т. 52, 1992. С. 22–40.

[Анг. пер.: Gashkov S., Kochergin V. On addition chains of vectors, gate circuits, and the complexity of computation of power, Syberian Advances in Mathematics, 1994, v.4, no 4, 1–16.]

[5] Гашков С. Б., Чубариков В. Н. *Арифметика. Алгоритмы. Сложность вычислений.* 1е изд. М.: Наука, 1996, 2е изд. М.: Высшая школа, 2000, 3е изд. М.:Дрофа, 2005.

[6] Григорьев Д. Ю. *Нижние оценки в алгебраической сложности вычислений* // Теория сложности вычислений. I., Записки научных семинаров ЛОМИ. Т. 118, 1982. С. 25–82.

[7] Карацуба А. А. *Сложность вычислений* // Труды Математического ин-та РАН. Т. 211, 1995. С. 1–17.

[8] Касим-Заде О. М. *О сложности схем из единичных сопротивлений и о некоторых свойствах чисел Фибоначчи* // Труды матем. института им. Стеклова. Т.218. М.: Наука, 1997. С. 233–248.

[9] Кнут Д. *Искусство программирования* Т. 2. Изд. Вильямс, 2000.

[10] Лупанов О. Б. *Асимптотические оценки сложности управляющих систем*. М.: Изд. МГУ, 1984.

[11] Лупанов О. Б. *О вентильных и контактно-вентильных схемах* // ДАН СССР, т. 111, №6, 1956. С. 1171–1174.

[12] *Математика в задачах*. М.: МЦНМО, 2009.

[13] Ноден П., Китте К. *Алгебраическая алгоритмика*. М.: Мир, 1999.

[14] Сидоренко А. Ф. *Сложность аддитивных вычислений семейств целочисленных линейных форм* // Записки научных семинаров ЛОМИ. Т. 105, 1981. С. 53–61.

[15] Смарт Н. *Криптография*. М.: Техносфера, 2005.

[16] Трауб Дж., Вожьняковский Х. *Общая теория оптимальных алгоритмов*. М.: Мир, 1983.

[17] Трауб Дж., Вожьняковский Х., Васильковский Г. *Информация, неопределенность, сложность*. М.: Мир, 1988.

[18] Bernstein D. J. *Pippenger's exponentiation algorithm.* `http://cr.yp.to/papers.html#pippenger`

[19] Bernstein D. J. *The transposition principle.* `http://cr.yp.to/transposition.html`.

[20] Blum L., Cucker F., Shub M., Smale S. *Complexity and real computation.* Springer-Verlag, 1998.

[21] Borwein J. M., Borwein R. P. *Pi and the AGM*. Wiley, New York, 1987.

[22] *Handbook of theoretical computer science. Algorithms and Complexity.* Elsevier-MIT Press, 1990.

[23] Gathen, von zur, J., Gerhard J. *Modern computer algebra.* Cambridge University Press, 1999.

[24] Pippenger N. *On the evaluation of powers and monomials* // SIAM J. Comput. Vol. 980, № 9, 1980. P. 230–250.

[25] `www.ccas.ru/personal/karatsuba/algen.htm`

С. Б. Гашков, механико-математический факультет МГУ им. М. В. Ломоносова
E-mail: `sergey.gashkov@lsili.ru`, `sbgashkov@gmail.com`

Строго равнобедренные множества

Ю. И. Ионин

§1. Вступление

На рис. 1 изображены шесть точек — вершины и центр правильного пятиугольника. Любые три из них образуют равнобедренный треугольник. В 1947 году знаменитый венгерский математик Пол Эрдёш задал следующие два вопроса.

1. Существует ли множество из семи точек на плоскости, любые три из которых являются вершинами равнобедренного треугольника?

2. Каково максимальное число точек в пространстве, любые три из которых являются вершинами равнобедренного треугольника?

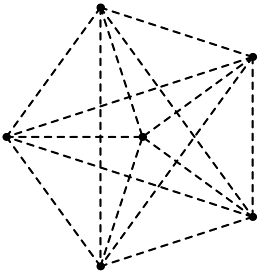


Рис. 1.

Отрицательный ответ на первый вопрос был получен Келли [7] в 1948 году. Келли также привел следующий пример множества из восьми точек в трехмерном пространстве: пять вершин и центр правильного пятиугольника и две точки на прямой, перпендикулярной плоскости пятиугольника и проходящей через его центр; расстояние от обеих точек до центра пятиугольника должно быть равно радиусу описанной окружности пятиугольника. Для любых трех точек из этого множества среди трех попарных расстояний между ними есть два равных. В 1962 году Крофт [4] доказал, что никакое множество из девяти точек в трехмерном пространстве не обладает этим свойством, а в 2006 году Кидо [8] доказал, что любое множество из восьми точек в трехмерном пространстве с двумя различными расстояниями между ними подобно примеру Келли (короткое доказательство утверждений Крофта и Кидо дано в [6]). С тех пор пример Келли считается решением задачи Эрдёша для трехмерного евклидова пространства. Однако, в примере Келли среди восьми точек есть три, лежащие на одной прямой, и потому не образующие треугольник, так что, строго говоря, семь точек в $\mathbb{E}^3$ — ответ задачи Эрдёша. В настоящей статье мы рассмотрим задачу Эрдёша в строгом смысле для евклидова пространства $\mathbb{E}^n$. Кроме определения максимального числа точек, любые три из которых образуют равнобедренный треугольник, мы попытаемся, где возможно, найти все конфигурации с максимальным числом точек.

Мы рассматриваем $\mathbb{E}^n$ как множество всех точек (векторов) $\mathbf{u} = (u_1, u_2, \dots, u_n)$, где u_i — вещественные числа. Если X — непустое подмножество пространства $\mathbb{E}^n$, то аффинное подпространство $\langle X \rangle$, порожденное множеством X, — это множество всех точек вида $\sum_{\mathbf{u}\in X} a_{\mathbf{u}}\mathbf{u}$, где среди коэффициентов $a_{\mathbf{u}}$ все, кроме конечного числа, равны 0 и сумма всех ненулевых коэффициентов равна 1. Если аффинное подпространство $\langle X \rangle$ порождается каким-нибудь множеством из $m+1$ точек и не порождается никаким множеством из m точек, то мы называем $m = \dim X$ размерностью множества X. В частности, $\dim X = 0$ означает, что X состоит из одной точки; $\dim X = 1$ означает, что $|X| \geqslant 2$ и все точки множества X лежат на одной прямой.

Мы используем обозначение $\|\cdot\|$ для евклидовой нормы и $\langle \cdot, \cdot \rangle$ для стандартного скалярного произведения, так что $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$.

Определение. Непустое множество S в евклидовом пространстве $\mathbb{E}^n$ называется *равнобедренным множеством*, если для любых трех точек множества S среди трех попарных расстояний между ними есть хотя бы два равных. Если, кроме того, никакие три точки множества S не лежат на одной прямой, то множество S называется *строго равнобедренным*.

После подготовительной работы в §§2, 3 мы приведем доказательство теоремы Блокхауса [2], устанавливающей верхнюю границу $(n+1)(n+2)/2$ для числа точек равнобедренного множества в $\mathbb{E}^n$. Эта граница достигается в размерностях 1, 2, 6 и 8, а для строго равнобедренных множеств — в размерностях 2, 6 и 8.

В §4 мы найдем точную верхнюю границу для числа точек строго равнобедренного множества в $\mathbb{E}^n$ для всех $n \leqslant 8$, а для $n \leqslant 6$ мы найдем все строго равнобедренные множества, реализующие эти границы. Аналогичные результаты для равнобедренных множеств (при $n \leqslant 8$ и $n \leqslant 7$ соответственно) получены автором в работе [6], свободно доступной в интернете. В §5 мы рассмотрим обобщения задачи Эрдёша на другие метрические пространства.

§2. Множества с двумя расстояниями

Для любого множества S в евклидовом пространстве обозначим через $\operatorname{dist}(S)$ число различных ненулевых расстояний между точками множества S.

Упражнение 1. Доказать: если множество $S \subset \mathbb{E}^n$ таково, что $\operatorname{dist}(S) = 1$, то $|S| \leqslant n+1$; если при этом $|S| = n+1$, то множество S — правильный n-мерный симплекс — определено однозначно с точностью до подобия.

Упражнение 2. Пусть $n \geqslant 3$ и пусть множество Λ_n образовано всеми точками пространства $\mathbb{E}^{n+1}$, у которых две координаты равны 1, а остальные $n-1$ координат равны 0. Доказать: множество Λ_n образовано серединами ребер правильного n-мерного симплекса, $\operatorname{dist}(\Lambda_n) = 2$, $\dim \Lambda_n = n$.

Упражнение 3. Доказать: если множество $S \subset \mathbb{E}^2$ таково, что $\operatorname{dist}(S) = 2$, то $|S| \leqslant 5$; если $|S| = 5$, то S — множество вершин правильного пятиугольника. Доказать, что существует ровно шесть (с точностью до подобия) различных четырехточечных множеств на плоскости с двумя ненулевыми расстояниями.

Очевидно, если $\operatorname{dist}(S) \leqslant 2$, то S — равнобедренное множество. Следующий результат, полученный в работе [9], показывает, что любое достаточно большое множество с двумя расстояниями является на самом деле строго равнобедренным.

Теорема 1. Пусть $S \subset \mathbb{E}^n$ и пусть $\operatorname{dist}(S) = 2$. Пусть $d_1 < d_2$ — ненулевые расстояния между точками множества S. Если $S > 2n+3$, то существует натуральное число $t \leqslant (-1+\sqrt{2n})/2$ такое, что

$$\left(\frac{d_2}{d_1}\right)^2 = \frac{t+1}{t}\,.$$

Следствие 1. Пусть $S \subset \mathbb{E}^n$ и пусть $\operatorname{dist}(S) = 2$. Если $|S| > 2n+3$, то S — строго равнобедренное множество.

Доказательство. Так как $(t+1)/t < 4$, то $d_2 < 2d_1$, и потому никакие три точки множества S не лежат на одной прямой.

Следствие 2. Пусть $S \subset \mathbb{E}^n$ и пусть $\operatorname{dist}(S) = 2$. Если $n \leqslant 4$, то $|S| \leqslant 2n+3$.

Доказательство. Если $n \leqslant 4$, то $(-1+\sqrt{2n})/2 < 1$, в то время как $(t+1)/t > 1$.

Следствие 3. Пусть $S \subset \mathbb{E}^n$ и пусть $\operatorname{dist}(S) = 2$. Пусть $d_1 < d_2$ — ненулевые расстояния между точками множества S. Если $5 \leqslant n \leqslant 12$ и $|S| > 2n+3$, то $d_2 = d_1\sqrt{2}$.

Доказательство. Если $n \leqslant 12$, то $(-1+\sqrt{2n})/2 < 2$, так что $t = 1$, $(t+1)/t = 2$.

В 1966 году Эйнгорн и Шёнберг [5] доказали, что множество $S \subset \mathbb{E}^3$ такое, что $\operatorname{dist}(S) = 2$, состоит не более чем из шести точек. При этом имеется ровно шесть (с точностью до подобия) конфигураций из шести точек с двумя различными ненулевыми расстояниями между точками. Две из них образованы вершинами правильного октаэдра и вершинами правильной треугольной призмы с квадратными боковыми гранями. Шестерки вершин правильного икосаэдра, не содержащие противоположных вершин, дают еще четыре конфигурации. Пусть A — вершина правильного икосаэдра. Соседи вершины A образуют правильный пятиугольник $BCDEF$. Оставшиеся шесть вершин — A', B', C', D', E', F' — диаметрально противоположны (по отношению к описанной сфере) вершинам A, B, C, D, E, F, соответственно. Если две вершины не противоположны, то расстояние между ними равно либо стороне, либо диагонали правильного пятиугольника $BCDEF$. Поэтому, если S — любое множество вершин правильного икосаэдра, не содержащее противоположных вершин, то $\operatorname{dist}(S) < 2$. Следующие четыре множества из шести вершин попарно не подобны:

$$\begin{gathered}\{A,B,C,D,E,F\}, \quad \{A,B',C',D',E',F'\},\\ \{A,B,C,D,E,F'\}, \quad \{A,B',C,D',E,F\}.\end{gathered} \tag{1}$$

Первые два множества состоят из вершин правильных пятиугольных пирамид, у которых боковые рёбра равны либо стороне, либо диагонали основания. Последнее множество образовано вершинами усеченной правильной треугольной пирамиды, у которой боковое ребро равно стороне меньшего основания, а острые углы боковых граней равны 72°.

Упражнение 4. Доказать: любое множество из шести вершин правильного икосаэдра, никакие две из которых не противоположны, конгруэнтно одному из множеств (1).

Упражнение 5. Доказать: в пространстве $\mathbb{E}^n$ существует множество из $2n$ точек с двумя ненулевыми расстояниями d_1 и d_2, такими, что $(d_1/d_2)^2$ иррационально. (*Указание*: найдите n-мерный аналог последнего из множеств (1).)

В работах [1] и [2] доказано, что если множество $S \subset \mathbb{E}^n$ таково, что $\operatorname{dist}(S) = s$, то $|S| \leqslant C^s_{n+s}$. Мы приведем доказательство этой теоремы только для интересующего нас случая $s = 2$.

Теорема 2. Если множество $S \subset \mathbb{E}^n$ таково, что $\operatorname{dist}(S) = s$, то $|S| \leqslant C^2_{n+2}$. Если, кроме того, все точки множества S лежат на сфере, то $|S| \leqslant C^2_{n+2} - 1 = n(n+3)/2$.

Доказательство. Пусть $\mathrm{dist}(S) = 2$. Если теорема справедлива для конечных множеств, то S не может быть бесконечным. Поэтому мы можем предположить, что S — конечное множество в $\mathbb{E}^n$. Пусть d_1, d_2 — различные ненулевые расстояния между точками множества S. Для каждой точки $\mathbf{u} = (u_1, u_2, \ldots, u_n) \in S$ рассмотрим следующий многочлен $F_{\mathbf{u}}$ от n переменных $\mathbf{x} = (x_1, x_2, \ldots, x_n)$: $F_{\mathbf{u}}(\mathbf{x}) = (\|\mathbf{x} - \mathbf{u}\|^2 - d_1^2)(\|\mathbf{x} - \mathbf{u}\|^2 - d_2^2)$. Если $\mathbf{u}, \mathbf{v} \in S$, $\mathbf{u} \neq \mathbf{v}$, то $\|\mathbf{u} - \mathbf{v}\| = d_1$ или d_2, и потому $F_{\mathbf{u}}(\mathbf{v}) = 0$. Кроме того, $F_{\mathbf{u}}(\mathbf{u}) = (d_1 d_2)^2 \neq 0$. Эти свойства многочленов $F_{\mathbf{u}}$ гарантируют их линейную независимость над полем вещественных чисел. В самом деле, если $\sum_{\mathbf{u}\in S} a_{\mathbf{u}} F_{\mathbf{u}} = 0$, где $a_{\mathbf{u}} \in \mathbb{R}$, то, для любой точки $\mathbf{v} \in S$, $\sum_{\mathbf{u}\in S} a_{\mathbf{u}} F_{\mathbf{u}}(\mathbf{v}) = a_{\mathbf{v}}(d_1 d_2)^2 = 0$, откуда следует, что $a_{\mathbf{v}} = 0$.

Раскроем скобки в выражении для $F_{\mathbf{u}}$:

$$F_{\mathbf{u}}(\mathbf{x}) = (\|\mathbf{x}\|^2 - 2\langle \mathbf{u}, \mathbf{x}\rangle + \|u\|^2 - d_1^2)(\|\mathbf{x}\|^2 - 2\langle \mathbf{u}, \mathbf{x}\rangle + \|u\|^2 - d_2^2) =$$
$$= \|\mathbf{x}\|^4 - 4\sum_{i=1}^{n} u_i(\|\mathbf{x}\|^2 x_i) + \sum_{1\leqslant i\leqslant j\leqslant n} \alpha_{ij} x_i x_j + \sum_{i=1}^{n} \beta_i x_i + \gamma,$$

где α_{ij}, β_i, γ — вещественные числа, конкретные значения которых нас не интересуют.

Таким образом, каждый из линейно независимых многочленов $F_{\mathbf{u}}$, $u \in S$, является линейной комбинацией одних и тех же $(n+1)(n+4)/2$ многочленов:

$$\|\mathbf{x}\|^4, \quad \|\mathbf{x}\|^2 x_i \ (1 \leqslant i \leqslant n), \quad x_i x_j \ (1 \leqslant i \leqslant j \leqslant n), \quad x_i \ (1 \leqslant i \leqslant n), \quad 1. \tag{2}$$

Мы предоставляем читателю проверить, что многочлены (2) линейно независимы.

Мы теперь покажем, что никакая линейная функция, кроме тождественного нуля, не является линейной комбинацией многочленов $F_{\mathbf{u}}$. Отсюда будет следовать, что многочлены $F_{\mathbf{u}}$ ($\mathbf{u} \in S$), x_i ($1 \leqslant i \leqslant n$) и константа 1 линейно независимы. Поскольку каждый из них является линейной комбинацией $(n+1)(n+4)/2$ многочленов (2), мы получим, что $|S| + n + 1 \leqslant (n+1)(n+4)/2$, откуда $|S| \leqslant C_{n+2}^2$.

Итак, предположим, что

$$\sum_{\mathbf{u}\in S} a_{\mathbf{u}} F_{\mathbf{u}}(\mathbf{x}) = c + \sum_{i=1}^{n} b_i x_i,$$

где $a_{\mathbf{u}}$, c и b_i — вещественные числа. Положим $\mathbf{b} = (b_1, b_2, \ldots, b_n)$. Тогда

$$\sum_{\mathbf{u}\in S} a_{\mathbf{u}} F_{\mathbf{u}}(\mathbf{x}) = c + \langle \mathbf{b}, \mathbf{x}\rangle.$$

Положив $\mathbf{x} = \mathbf{v}$, где $\mathbf{v} \in S$, мы получим $a_{\mathbf{v}}(d_1 d_2)^2 = c + \langle \mathbf{b}, \mathbf{v} \rangle$. Следовательно,

$$\sum_{\mathbf{u} \in S} (c + \langle \mathbf{b}, \mathbf{u} \rangle) F_{\mathbf{u}}(\mathbf{x}) = (d_1 d_2)^2 (c + \langle \mathbf{b}, \mathbf{x} \rangle). \tag{3}$$

Обе части этого равенства являются линейными комбинациями линейно независимых многочленов (2), и потому многочлены (2) должны входить в эти линейные комбинации с одним и тем же коэффициентом. Равенство коэффициентов при $\|x\|^4$ означает, что $c|S| + \sum_{\mathbf{u} \in S} \langle \mathbf{b}, \mathbf{u} \rangle = 0$, откуда

$$\frac{1}{|S|} \sum_{\mathbf{u} \in S} \langle \mathbf{b}, \mathbf{u} \rangle = -c. \tag{4}$$

Для $i = 1, 2, \ldots, n$ равенство коэффициентов при $\|x\| 2x_i$ означает, что

$$\sum_{\mathbf{u} \in S} (c + \langle \mathbf{b}, \mathbf{u} \rangle) u_i = 0$$

откуда

$$\sum_{\mathbf{u} \in S} (c + \langle \mathbf{b}, \mathbf{u} \rangle) b_i u_i = 0,$$

$$\sum_{i=1}^{n} \sum_{\mathbf{u} \in S} (c + \langle \mathbf{b}, \mathbf{u} \rangle) b_i u_i = 0,$$

$$\sum_{\mathbf{u} \in S} \sum_{i=1}^{n} (c + \langle \mathbf{b}, \mathbf{u} \rangle) b_i u_i = 0,$$

$$\sum_{\mathbf{u} \in S} (c + \langle \mathbf{b}, \mathbf{u} \rangle) \langle \mathbf{b}, \mathbf{u} \rangle = 0.$$

Используя (4), мы получаем

$$\frac{1}{|S|} \sum_{\mathbf{u} \in S} \langle \mathbf{b}, \mathbf{u} \rangle^2 = c^2. \tag{5}$$

Равенства (4) и (5) означают, что квадрат среднего арифметического чисел $\langle \mathbf{b}, \mathbf{u} \rangle$, $\mathbf{u} \in S$, равен среднему арифметическому их квадратов. Поэтому все числа $\langle \mathbf{b}, \mathbf{u} \rangle$ равны друг другу и, следовательно, равны $-c$. Но тогда равенство (3) означает, что многочлен $c + \langle \mathbf{b}, \mathbf{x} \rangle = c + \sum_{i=1}^{n} b_i x_i$ является тождественным нулем, и первая часть теоремы доказана.

Пусть множество S лежит на некоторой сфере. Не теряя общности, мы предположим, что это сфера радиуса 1 с центром в начале координат. Будем рассматривать многочлены $F_{\mathbf{u}}$ и многочлены (2) не на всём пространстве $\mathbb{E}^n$, а только на этой сфере. Тогда $x_n^2 = 1 - x_1^2 - x_2^2 - \cdots - x_{n-1}^2$, и потому мы можем исключить многочлен x_n^2 из списка (2). Список (2) теперь состоит из $(n+1)(n+4)/2-1$ многочленов. Сохраняя оставшуюся часть проведенного выше доказательства неизменной, мы получим неравенство $|S| \leqslant C_{n+2}^2 - 1$. $\square$

В работе [10] Лисонек разработал алгоритм для нахождения множеств с двумя расстояниями. Не вдаваясь в детали, отметим, что математической основой алгоритма служат теорема 1 и следующая теорема, доказанная Шёнбергом в 1935 году. Доказательство этой теоремы приведено в классической книге Л. Блюменталя [3].

Теорема 3. *Пусть все диагональные элементы симметрической матрицы $D = [d_{ij}]$ порядка $m \geqslant 2$ равны 0, а все внедиагональные элементы положительны. Зададим симметрическую матрицу $C = [c_{ij}]$ порядка $m-1$ равенствами $c_{ij} = d_{im}^2 + d_{jm}^2 - d_{ij}^2$. Пусть n — число положительных собственных чисел матрицы C с учетом кратностей. Следующие утверждения эквивалентны:*

(а) все собственные числа матрицы C неотрицательны;

(б) n — минимальная размерность евклидова пространства, содержащего такие m точек $\mathbf{u}_1$, $\mathbf{u}_2$, $\ldots$, $\mathbf{u}_m$, что $\|\mathbf{u}_i - \mathbf{u}_j\| = d_{ij}$ для всех $i, j = 1, 2, \ldots, m$.

В заключение этого параграфа мы приведем три таблицы (см. с. 161). В первой строке каждой таблицы указана размерность евклидова пространства. Во второй строке для пространства данной размерности дается точная верхняя граница для числа точек в множестве с двумя ненулевыми расстояниями, в равнобедренном множестве и в строго равнобедренном множестве, соответственно. В третьей строке дается число множеств, реализующих эту границу (с точностью до подобия).

Для $n \geqslant 4$ таблица 1 получена в [10], таблица 2 получена автором этой статьи в [6]. Обоснование таблицы 3 дано в §4 настоящей статьи.

§3. Как раскрасить рёбра графа

Напомним, что *полный граф* K_n — это граф с n вершинами, в котором каждая пара вершин соединена одним ребром. Раскраска ребер полного графа — это присвоение каждому ребру некоторого цвета. Раскраска ребер графа в цвета $c_1, c_2, \ldots, c_k$ называется *связной*, если для каждого цвета c_i можно пройти по ребрам этого цвета из любой вершины графа в любую другую вершину. Будем говорить, что вершины u, v, w образуют разноцветный треугольник, если рёбра uv, uw, vw окрашены в три разных цвета. Следующая лемма была доказана в [7]. Мы следуем доказательству, приведенному в [8].

Табл. 1. Множества с двумя ненулевыми расстояниями в $\mathbb{E}^n$

n	1	2	3	4	5	6	7	8
Максимальное число точек в множестве	3	5	6	10	16	27	29	45
Число множеств с максимальным числом точек	1	1	6	1	1	1	1	$\geqslant 1$

Табл. 2. Равнобедренные множества в $\mathbb{E}^n$

n	1	2	3	4	5	6	7	8
Максимальное число точек в множестве	3	6	8	11	17	28	30	45
Число множеств с максимальным числом точек	1	1	1	2	1	1	1	$\geqslant 1$

Табл. 3. Строго равнобедренные множества в $\mathbb{E}^n$

n	1	2	3	4	5	6	7	8
Максимальное число точек в множестве	2	6	7	11	17	28	29	45
Число множеств с максимальным числом точек	1	1	∞	2	1	1	∞	$\geqslant 1$

Лемма 1. Пусть рёбра графа K_n раскрашены в k цветов без разноцветных треугольников. Тогда справедливы следующие два утверждения.

(1) Пусть красный — один из данных k цветов и пусть Γ — граф с теми же вершинами, что и K_n, ребрами которого являются в точности все красные рёбра графа K_n. Пусть C_1 и C_2 — различные связные компоненты графа Γ. Тогда все рёбра графа K_n, соединяющие вершины из C_1 с вершинами из C_2, окрашены в один и тот же цвет.

(2) Если $k \geqslant 3$, то данная раскраска в k цветов не является связной.

Доказательство. (1) Пусть z — вершина из C_1 и пусть u, v — две вершины из C_2. В C_2 есть путь, ведущий из u в v, состоящий из красных ребер v_0v_1, v_1v_2, $\ldots$, $v_{m-1}v_m$, где $v_0 = u$, $v_m = v$. Так как вершина z не лежит в C_2, рёбра zv_0, zv_1, $\ldots$, zv_m не могут быть красными. Рассмотрим последовательно треугольники zv_0v_1, zv_1v_2, $\ldots$, $zv_{m-1}v_m$. Так как ни один из них не является разноцветным, мы получим, что рёбра zv_0, zv_1, $\ldots$, zv_m должны быть одного и того же цвета, так что zu и zv — рёбра одного и того же цвета.

(2) При $n \leqslant 3$ это утверждение верно, поскольку K_n либо нельзя раскрасить в $k \geqslant 3$ цветов, либо такая раскраска содержит разноцветный треугольник. Мы продолжим доказательство индукцией по n.

Предположим, что $n \geqslant 4$ и что любая раскраска ребер графа K_{n-1} в $k \geqslant 3$ цветов либо допускает разноцветный треугольник, либо не является связной. Пусть дана связная раскраска ребер графа K_n в $k \geqslant 3$ цветов. Мы докажем, что в графе K_n найдется разноцветный треугольник. Пусть красный, желтый и зеленый — три данных цвета. Перекрасим все рёбра других цветов в зеленый цвет. Мы получим связную раскраску в три цвета, и любой разноцветный треугольник при этой раскраске был разноцветным до перекраски.

Выберем вершину x и рассмотрим граф K_{n-1}, полученный удалением из K_n вершины x и всех выходящих из нее ребер. Если бы в графе K_{n-1} не было ребер какого-нибудь цвета, скажем, красного, то поскольку в K_n можно пройти из x в любую другую вершину по красным ребрам, все рёбра, выходящие из x, были бы красными. Но тогда из x нельзя никуда пройти по желтым (как и по зеленым) ребрам. Следовательно, в раскраске графа K_{n-1} встречаются все три цвета. Любой разноцветный треугольник в графе K_{n-1} является таковым и в K_n. Поэтому предположим, что в графе K_{n-1} нет разноцветных треугольников. В силу индукционного предположения имеющаяся раскраска графа K_{n-1} не является связной.

Пусть, для определенности, не из всякой вершины графа K_{n-1} можно пройти в любую другую вершину по красным ребрам. Обозначим через Γ граф с теми же вершинами, что и K_{n-1}, и с теми и только теми ребрами, которые принадлежат графу K_{n-1} и окрашены в красный цвет. Граф Γ не является связным, так что пусть $C_1, C_2, \ldots, C_\ell$ — его связные компоненты, $\ell \geqslant 2$. В силу утверждения (1) при $i \neq j$ все рёбра, соединяющие вершины компоненты C_i с вершинами компоненты c_j, окрашены в один и тот же (не красный) цвет. Поскольку в графе K_n можно пройти из x в любую другую вершину по красным ребрам, в каждой компоненте C_i есть вершина y_i, соединенная с x красным ребром. Кроме того, в графе K_n есть желтое ребро xu и зеленое ребро xv.

Случай 1. Вершины u и v принадлежат одной и той же компоненте, скажем, C_1.

Так как рёбра uy_2 и vy_2 одного и того же (не красного) цвета, то один из треугольников xuy_2, xvy_2 — разноцветный.

Случай 2. Вершина u принадлежит C_1, вершина v принадлежит C_2.

Если все рёбра между C_1 и C_2 — желтые, то xy_1v — разноцветный треугольник; если все рёбра между C_1 и C_2 — зеленые, то xy_2u — разноцветный треугольник. Лемма доказана. $\square$

В работе [2] Блокхаус применил лемму 1 и получил результат, позволяющий в значительной степени свести поиск максимальных равнобедренных множеств к поиску максимальных множеств с двумя ненулевыми расстояниями.

Теорема 4. Пусть S — конечное равнобедренное множество в $\mathbb{E}^n$ и пусть $\operatorname{dist}(S) \geqslant 3$. Тогда S можно представить в виде объединения непустых непересекающихся подмножеств X и Y таких, что $\dim X \geqslant 1$, $\operatorname{dist}(X) \leqslant 2$ и каждая точка множества Y является центром сферы, содержащей множество X. Более того, аффинные подпространства $\langle X\rangle$ и $\langle Y\rangle$ ортогональны, и потому $\dim X + \dim Y \leqslant \dim S$.

Доказательство. Рассмотрим полный граф $K_{|S|}$, множество вершин которого совпадает с множеством S. Пусть $d_1, d_2, \ldots, d_k$ — все различные ненулевые расстояния между точками множества S. Выберем k различных цветов $c_1, c_2, \ldots, c_k$ и окрасим ребро uv графа $K_{|S|}$ в цвет c_i в том и только в том случае, если расстояние между точками u и v множества S равно d_i. Так как S — равнобедренное множество, полученная раскраска ребер графа $K_{|S|}$ не содержит разноцветных треугольников. В силу леммы 1 граф Γ, образованный всеми вершинами графа $K_{|S|}$ и всеми ребрами какого-то одного цвета, скажем, красного, не является связным. Так как в графе $K_{|S|}$ есть хотя бы одно красное ребро, в графе Γ есть связная компонента C_1, содержащая не менее двух вершин. Пусть X — подмножество множества S, образованное всеми вершинами компоненты C_1 и пусть $Y = S \setminus X$. Пусть $\mathbf{y} \in Y$. Тогда вершина y принадлежит некоторой связной компоненте $C_2 \neq C_1$ графа Γ, так что в силу леммы 1 все рёбра $\mathbf{xy}$, где $\mathbf{x} \in X$, окрашены в один и тот же цвет. Это означает, что все точки $\mathbf{x} \in X$ находятся на одном и том же расстоянии от точки y, т. е. y — центр сферы, содержащей множество X.

Если $\operatorname{dist}(X) \geqslant 3$, то мы применим к X доказанную часть теоремы и представим X в виде $X = X_1 \cup Y_1$, где $X_i \cap Y_1 = \varnothing$, $|X_1| \geqslant 2$, $Y_1 \neq \varnothing$ и каждая точка $\mathbf{y} \in Y_1$ — центр сферы, содержащей X_1. Тогда $S = X_1 \cup (Y \cup Y_1)$ и каждая точка $\mathbf{y} \in Y \cup Y_1$ — центр сферы, содержащей X_1. При этом $|X_1| < |X|$, так что если с самого начала подмножество X было выбрано с минимальным возможным числом элементов, то $\operatorname{dist}(X) \leqslant 2$.

Так как каждая точка $\mathbf{y} \in Y$ равноудалена от всех точек множества X, то и ортогональная проекция $\mathbf{z}$ точки $\mathbf{y}$ на аффинное подпространство $\langle X\rangle$ равноудалена от всех точек множества X. Так как множество X порождает подпространство $\langle X\rangle$, то в $\langle X\rangle$ есть единственная сфера, содержащая X, и точка $\mathbf{z}$ — центр этой сферы. Так как множество Y порождает подпространство $\langle Y\rangle$, то ортогональная проекция всего подпространства $\langle Y\rangle$ на подпространство $\langle X\rangle$ состоит из единственной точки $\mathbf{z}$, т. е., подпространства $\langle X\rangle$ и $\langle Y\rangle$ ортогональны. $\square$

Применив теоремы 2 и 4, Блокхаус [2] получил оценку числа точек равнобедренного множества, лежащего в евклидовом пространстве $\mathbb{E}^n$.

Теорема 5. Пусть $S \subset \mathbb{E}^n$ — равнобедренное множество. Тогда S конечно и $|S| \leqslant C_{n+2}^2$. Более того, если $|S| = C_{n+2}^2$, то либо $\operatorname{dist}(S) \leqslant 2$, либо $S = X \cup \{\mathbf{y}\}$, где $\operatorname{dist}(X) \leqslant 2$ и $\mathbf{y}$ — центр сферы, содержащей X.

Доказательство. Как и в доказательстве теоремы 2, мы можем с самого начала предположить, что S — конечное множество. Если $|S| \leqslant$ $\leqslant C_{n+2}^2$, то $|S| \leqslant C_{n+2}^2$ в силу теоремы 2. Предположим, $\operatorname{dist}(S) \geqslant 3$, и пусть X и Y — подмножества множества S, удовлетворяющие теореме 4. Мы продолжим доказательство индукцией по n. Теорема выполняется для $n = 1$, так что пусть $n \geqslant 2$ и пусть теорема выполняется для евклидовых пространств размерности меньше n.

Пусть $\dim X = r$, $\dim Y = s$. Тогда $r + s \leqslant \dim S \leqslant n$. Так как $r \geqslant 1$, то $s < n$. Применяя теорему 2 к множеству X, лежащему на сфере (с центром в любой точке множества Y), мы получаем неравенство $|X| \leqslant C_{r+2}^2 - 1$. Если $s = 0$, то пусть $Y = \{\mathbf{y}\}$. Тогда $|S| \leqslant C_{n+2}^2$ и X лежит на сфере с центром $\mathbf{y}$. Если $s \geqslant 1$, мы применим индукционное предположение к множеству Y:

$$|S| = |X| + |Y| \leqslant C_{r+2}^2 + C_{s+2}^2 - 1 < C_{n+2}^2.$$

Действительно, так как $(r+1) + s \leqslant n + 1$, то

$$\begin{aligned} C_{n+2}^2 &= (1 + 2 + \cdots + (n+1)) \geqslant \\ &\geqslant (1 + 2 + \cdots + (r+1)) + s(r+1) + (1 + 2 + \cdots + s) \geqslant \\ &\geqslant (1 + 2 + \cdots + (r+1)) + (1 + 2 + \cdots + (s+1)) = \\ &= C_{r+2}^2 + C_{s+2}^2 > C_{r+2}^2 + C_{s+2}^2 - 1. \end{aligned}$$

□

§4. Максимальные строго равнобедренные множества

В этом параграфе мы найдем точную верхнюю границу для числа точек в строго равнобедренном подмножестве пространства $\mathbb{E}^n$ при $n \leqslant 8$ и найдем все строго равнобедренные множества, реализующие эту границу, при $n \leqslant 6$. Изложение следует, в основном, схеме, развитой в статье [9]. Но прежде нам понадобится еще одно понятие, связанное с раскраской полного графа — на этот раз в два цвета.

Определение. Раскраска полного графа в два цвета, скажем, красный и зеленый, называется *сильно регулярной с параметрами* (v, k, λ, μ), если (1) граф имеет v вершин, (2) из каждой вершины выходит ровно k красных ребер, (3) если вершины x и y соединены красным ребром, то в графе есть ровно λ вершин z таких, что xz и yz — красные рёбра, (4) если вершины x и y соединены зеленым ребром, то в графе есть ровно μ вершин z таких, что xz и yz — красные рёбра. При этом граф Γ, образованный всеми вершинами и всеми красными ребрами полного графа, называют *сильно регулярным графом с параметрами* (v, k, λ, μ).

Разумеется, граф Γ', образованный всеми вершинами и всеми зелеными ребрами полного графа, тоже сильно регулярен. Его параметры — $(v, v-k-1, v-2k+\lambda, v-2k+\mu-2)$. Граф Γ' называется *дополнением графа* Γ.

Пример. Граф, образованный вершинами и сторонами правильного пятиугольника, — сильно регулярный с параметрами $(5, 2, 0, 1)$.

Упражнение 6. Проверьте, что каждый из следующих графов сильно регулярен и найдите его параметры.

(а) Граф Петерсена, изображенный на рис. 2.

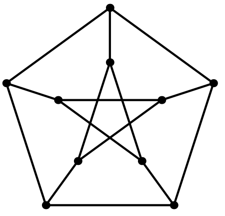


Рис. 2.

(б) Дано натуральное число n. Вершины графа — все точки координатной плоскости с целыми координатами (x, y) такими, что $1 \leqslant x \leqslant n$, $1 \leqslant y \leqslant n$. Две вершины, (x, y) и (u, v), соединены ребром в том и только в том случае, если $x = u$ или $y = v$.

(в) Дано натуральное число $n \geqslant 2$. Вершины *треугольного графа* $T(n)$ — все двухэлементные подмножества множества $\{1, 2, \ldots, n\}$. Две вершины соединены ребром в том и только в том случае, если соответствующие подмножества имеют общий элемент.

(г) Вершины *графа Клебша* — все подмножества множества $\{1, 2, 3, 4, 5\}$ с четным числом элементов. Две вершины соединены ребром в том и только в том случае, если симметрическая разность соответствующих подмножеств состоит из четырех элементов.

Следующая лемма будет применяться при решении задачи Эрдёша в пространствах $\mathbb{E}^3$, $\mathbb{E}^4$ и $\mathbb{E}^7$.

Лемма 2. Пусть множество $S \subset \mathbb{E}^n$ таково, что $\operatorname{dist}(S) = 2$ и пусть d — одно из ненулевых расстояний между точками множества S. Образуем полный граф, вершинами которого являются все точки множества S.

Пусть две вершины соединены красным ребром, если расстояние между ними (как точками множества S) равно d, и зеленым ребром в противном случае. Предположим, что полученная раскраска сильно регулярна с параметрами (v, k, λ, n), причем $v \geqslant 2k + 1$. Предположим также, что $\dim S = n - 1$, $\dim(S \setminus \{\mathbf{x}\}) = n - 1$ для любой точки $\mathbf{x} \in S$ и, кроме того, выполняется по крайней мере одно из следующих трех условий:

(1) $v \leqslant 3k - 2\lambda$ и $v \leqslant 3k - 2\mu + 2$;
(2) $v \leqslant 3k - 2\lambda$ и $2k \geqslant 2\mu + n - 1$;
(3) $v \leqslant 3k - 2\mu + 2$ и $2k \geqslant 2\lambda + n + 1$.

Пусть точка $\mathbf{p} \in \mathbb{E}^n$ такова, что $S \cup \{\mathbf{p}\}$ — равнобедренное множество размерности n. Тогда либо $\operatorname{dist}(S \cup \{\mathbf{p}\}) = 2$, либо $\mathbf{p}$ — центр сферы, содержащей множество S.

Доказательство. Неравенство $v \geqslant 2k + 1$ означает, что число зеленых ребер, выходящих из любой вершины данного графа, не меньше числа красных ребер.

Предположим, что $\operatorname{dist}(S \cup \{\mathbf{p}\}) \geqslant 3$. Тогда $S \cup \{\mathbf{p}\} = X \cup Y$, где X и Y удовлетворяют теореме 2. Если $Y = \{\mathbf{p}\}$, то $\mathbf{p}$ — центр сферы, содержащей множество S, так что предположим, что $Y \neq \{\mathbf{p}\}$, т. е., $S \cap Y \neq \varnothing$. Пусть $\mathbf{y} \in S \cap Y$. Определение множества Y означает, что все рёбра, соединяющие $\mathbf{y}$ с точками множества $S \cap X$, — одного цвета, и потому $|S \cap X|$ не превосходит числа зеленых ребер, выходящих из любой вершины графа: $|S \cap X| \leqslant v - k - 1$.

Пусть $\mathbf{a}, \mathbf{b} \in S \cap X$, $\mathbf{a} \neq \mathbf{b}$. Предположим сначала, что $\mathbf{a}$ и $\mathbf{b}$ соединены красным ребром. Тогда имеется λ вершин, соединенных как с $\mathbf{a}$, так и с $\mathbf{b}$ красными ребрами, и $v - 2k + \lambda$ вершин, соединенных с $\mathbf{a}$ и с $\mathbf{b}$ зелеными ребрами. Так как каждая вершина $\mathbf{y} \in S \cap Y$ соединена с $\mathbf{a}$ и $\mathbf{b}$ ребрами одного цвета, мы получаем, что $|S \cap Y| \leqslant v - 2k + 2\lambda$. Следовательно, в этом случае $v = |S| = |S \cap X| + |S \cap Y| \leqslant 2v - 3k + 2\lambda - 1$, т. е., $v > 3k - 2\lambda + 1$. Если $\mathbf{a}$ и $\mathbf{b}$ соединены зеленым ребром, мы получаем аналогично, что $|S \cap Y| \leqslant v - 2k + 2\mu - 2$, $v = |S| \leqslant 2v - 3k + 2\mu - 3$, так что $v \geqslant 3k - 2\mu + 3$. Поэтому, если выполняется условие (1), то между точками множества $S \cap X$ нет ни красных, ни зеленых ребер. Так как $|X| \geqslant 2$, то $|S \cap X| = 1$.

Кроме того, полученные неравенства показывают, что если выполняется условие (2), то никакие две точки множества $S \cap X$ не соединены зеленым ребром, а если выполняется условие (3), то никакие две точки множества $S \cap X$ не соединены красным ребром. В обоих случаях все рёбра между точками множества $S \cap X$ одного и того же цвета, т. е., $\operatorname{dist}(S \cap X) = 1$. Отсюда $|S \cap X| \leqslant \dim(S \cap X) + 1 \leqslant n$ и $|S \cap Y| = v - |S \cap X| \geqslant v - n$. Однако, если точки множества $S \cap X$ соединены красными ребрами, то $v - n \leqslant |S \cap Y| \leqslant v - 2k + 2\lambda$, откуда $2k \leqslant 2\lambda + n$, что противоречит условию (3). Если же точки множества $S \cap X$ соединены зелеными ребрами, то $v - n \leqslant |S \cap Y| \leqslant v - 2k + 2\mu - 2$, и мы получаем противоречие с условием (2).

Итак, множество $S \cap X$ состоит из единственной точки $\mathbf{a}$, т. е., $X = \{\mathbf{a}, \mathbf{p}\}$. Так как каждая точка множества Y равноудалена от $\mathbf{a}$ и $\mathbf{p}$, то множество Y лежит в гиперплоскости π, проходящей через середину отрезка $\mathbf{ap}$ перпендикулярно к этому отрезку. Так как $\dim Y = \dim(S \setminus \{\mathbf{a}\}) = n - 1$, то множество Y порождает π. Так как $\mathbf{a} \notin \pi$, мы получаем $\dim S = n$, что противоречит условию леммы. $\square$

Следствие. Если основание пирамиды — правильный пятиугольник и любые три вершины пирамиды образуют равнобедренный треугольник, то пирамида — правильная.

Доказательство. Пусть S — множество всех вершин основания пирамиды и пусть $\mathbf{p}$ — вершина пирамиды. Тогда $\operatorname{dist}(S) = 2$. Пусть стороны основания окрашены в красный цвет, а диагонали — в зеленый. Полученная раскраска сильно регулярна с параметрами $(5, 2, 0, 1)$, так что мы можем применить лемму 2. Если $\mathbf{p}$ — центр сферы, содержащей основание пирамиды, то ортогональная проекция вершины пирамиды на плоскость основания — центр основания. Если $\operatorname{dist}(S \cup \{\mathbf{p}\} = 2$, то среди пяти боковых ребер пирамиды есть по крайней мере три равных, и мы приходим к тому же заключению.

Упражнение 7. Найдите доказательство этого следствия, не опирающееся на лемму 2 и понятие сильно регулярного графа. Верно ли аналогичное утверждение для треугольной, четырехугольной, шестиугольной пирамиды?

Пусть S — строго равнобедренное множество с максимально возможным числом точек в пространстве $\mathbb{E}^n$, $n \leqslant 7$. Если $\operatorname{dist}(S) = 2$, то S найдено в [10] и число точек множества S указано в таблице 1. Если $\operatorname{dist}(S) \geqslant 3$, то пусть X и Y — подмножества множества S, удовлетворяющие теореме 3.

$$n = 2$$

Множество, состоящее из вершин и центра правильного пятиугольника, — строго равнобедренное, так что $|S| \geqslant 6$. Тогда $\operatorname{dist}(S) \geqslant 3$ (упражнение 3). Если $\dim X = 2$, то $\dim Y = 0$ и $|X| \leqslant 5$. Следовательно, $|X| = 5$, т. е., X — множество всех вершин правильного пятиугольника, а $Y = \{by\}$, где $\mathbf{y}$ — центр пятиугольника. Если $\dim X = 1$, то $\dim Y \leqslant 1$, так что $|X| = 2$, $|Y| \leqslant 2$, $|S| \leqslant 4$.

$n = 3$

Присоединив центр сферы, описанной около правильного икосаэдра, к любой из фигур (1), мы получим четыре строго равнобедренных множества из 7 точек в $\mathbb{E}^3$, так что $|S| \geqslant 7$. Тогда $\operatorname{dist}(S) \geqslant 3$ (см. абзац после следствия 3) и, так как $\operatorname{dist}(X) \leqslant 2$, то $|X| \leqslant 6$. Если $Y = \{\mathbf{y}\}$, то $|X| = 6$, а $\mathbf{y}$ — центр сферы, содержащей X. Если X — множество вершин правильного октаэдра или правильной треугольной призмы с квадратными гранями, то S не является строго равнобедренным, так что S одно из упомянутых выше четырех множеств.

Пусть $\dim Y \neq 0$. Предположим, $\dim X = 2$. Тогда $\dim Y = 1$, $|X| \leqslant 5$, $|Y| = 2$. Следовательно, $|X| = 5$, т. е., X — множество всех вершин правильного пятиугольника. Так как каждая точка множества $Y = \{\mathbf{u}, \mathbf{v}\}$ равноудалена от вершин пятиугольника, то прямая ℓ, содержащая Y, перпендикулярна плоскости пятиугольника и проходит через его центр $\mathbf{c}$. Пусть $\mathbf{w}$ — середина отрезка $\mathbf{uv}$. Пусть $\mathbf{x} \in X$. Если $\mathbf{w} = \mathbf{c}$, то S — строго равнобедренное множество. Если $\mathbf{w} \neq \mathbf{c}$, то $\|\mathbf{u}-\mathbf{x}\| \neq \|\mathbf{v}-\mathbf{x}\|$, так что расстояние между $\mathbf{u}$ и $\mathbf{v}$ должно быть равно либо расстоянию от $\mathbf{u}$ до точек множества X, либо расстоянию от $\mathbf{v}$ до точек множества X. Таким образом, мы получаем бесконечно много семиточечных строго равнобедренных множеств, выбрав произвольную точку $\mathbf{u}$ на прямой ℓ, произвольную точку $\mathbf{x} \in X$ и затем определив возможные положения точки $\mathbf{v} \in \ell$ так, чтобы $\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u} - \mathbf{x}\|$ или $\|\mathbf{u} - \mathbf{v}\| = \|\mathbf{v} - \mathbf{x}\|$ (рис. 3).

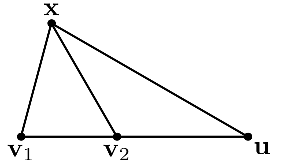


Рис. 3.

Предположим, что $\dim X = 1$. Тогда $|X| = 2$ и $\dim Y \leqslant 2$, так что $|Y| \leqslant 6$. Если $|Y| = 6$, то Y состоит из вершин и центра правильного пятиугольника, а X лежит на прямой, перпендикулярной плоскости пятиугольника. Точка множества X, не лежащая в плоскости пятиугольника, и вершины пятиугольника образуют пирамиду, удовлетворяющую следствию леммы 2. Следовательно, обе точки множества X и центр пятиугольника лежат на одной пряной, т. е., множество S не является строго равнобедренным. Пусть $|Y| = 5$. Если $\operatorname{dist}(Y) = 2$, то Y — множество вершин правильного пятиугольника, так что мы опять применяем следствие леммы 2, откуда S — ранее полученное множество из семи точек. Если $\operatorname{dist}(Y) \geqslant 3$, то мы применяем теорему 3 к множеству Y. Однако, если

$Y = X_1 \cup Y_1$ и $\dim X_1 + \dim Y_1 \leqslant 2$, то $|Y| \leqslant 4$. Таким образом, мы нашли все строго равнобедренные множества из семи точек в $\mathbb{E}^3$.

$$n = 4$$

Как установлено в работе [10], если $S \subset \mathbb{E}^4$ и $\operatorname{dist}(S) = 2$, то $|S| \leqslant 10$, и множество Λ_4 (см. упражнение 2) — это единственное (с точностью до подобия) множество из 10 точек с двумя ненулевыми расстояниями в $\mathbb{E}^4$. Присоединив к этому множеству центр описанной сферы $(\frac{2}{5}, \frac{2}{5}, \frac{2}{5}, \frac{2}{5}, \frac{2}{5})$, мы получим строго равнобедренное множество из 11 точек.

Пусть S — произвольное строго равнобедренное множество в $\mathbb{E}^4$ и пусть $|S| \geqslant 11$. Тогда $\operatorname{dist}(S) \geqslant 3$ и мы применяем теорему 2. Если $\dim X = 4$, то $\dim Y = 0$. Следовательно, $|X| = 10$ и $|Y| = 1$. Так как $\operatorname{dist}(X) = 2$, то S — множество из 11 точек, которое мы только что описали.

Если $\dim X = 3$, то $\dim Y \leqslant 1$, так что $|X| \leqslant 5$, $|Y| \leqslant 2$, $|S| \leqslant 7$.

Если $\dim X = 1$, то $\dim Y \leqslant 3$, $|X| = 2$, $|Y| \leqslant 7$, $|S| \leqslant 9$.

Пусть $\dim X = 2$. Тогда $\dim Y \leqslant 2$, $|X| \leqslant 5$, $|Y| \leqslant 6$. Так как $|S| \geqslant 11$, мы получаем, что X состоит из вершин правильного пятиугольника, а Y состоит из вершин и центра другого правильного пятиугольника, причем плоскости пятиугольников ортогональны. Применяя следствие леммы 2 к произвольной вершине первого пятиугольника и к множеству всех вершин второго пятиугольника, мы получаем, что ортогональная проекция всех вершин и, следовательно, плоскости первого пятиугольника на плоскость второго пятиугольника — центр второго пятиугольника. Отсюда следует, что пятиугольники имеют общий центр. Так как треугольник, образованный вершиной первого пятиугольника, вершиной второго пятиугольника и их общим центром, должен быть равнобедренным, то пятиугольники конгруэнтны. Таким образом, мы получили второй пример строго равнобедренного множества из 11 точек в $\mathbb{E}^4$ — вершины и общий центр двух конгруэнтных правильных пятиугольников, лежащих в ортогональных плоскостях.

$$n = 5$$

В работе [10] установлено, что если $S \subset \mathbb{E}^5$ и $\operatorname{dist}(S) = 2$, то $|S| \leqslant 16$; при этом $|S| = 16$ в том и только в том случае, если S — множество из 16 вершин пятимерного куба, никакие две из которых не являются концами одного ребра куба. С точностью до подобия множество S может быть представлено как множество всех точек, у которых каждая координата равна 0 или 1 и сумма всех координат четна. Присоединив к множеству S точку $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ — центр куба, мы получим строго равнобедренное множество из 17 точек. Заметим, что граф Γ, вершинами которого являются точки множества S и две вершины соединены ребром в том и только в том случае, если расстояние между ними (как точками множества S) равно 2, — это сильно регулярный граф Клебша (см. упражнение 6(г)).

Если $S \subset \mathbb{E}^5$ — строго равнобедренное множество и $\operatorname{dist}(S) \geqslant 3$, мы применяем теорему 2. Если $\dim X = 5$, мы получаем описанную выше конфигурацию. Если $\dim X = 4, 3, 2, 1$, то $|S|$ не превосходит $10+2$, $6+6$, $5+7$, $2+11$, соответственно, так что $|S| \leqslant 17$.

$$n = 6$$

Согласно [10], наибольшее множество с двумя расстояниями в $\mathbb{E}^6$ состоит из 27 точек и единственно. Это множество может быть представлено как следующее подмножество пространства $\mathbb{E}^8$:

$$T = \{\mathbf{a}_i, \mathbf{b}_i : 1 \leqslant i \leqslant 6\} \cup \{\mathbf{c}_{ij} : 1 \leqslant i < j \leqslant 6\},$$

где у точки $\mathbf{a}_i$ i-тая и седьмая координаты равны 2, а остальные шесть координат равны 0, у точки $\mathbf{b}_i$ i-тая и восьмая координаты равны 2, а остальные шесть координат равны 0, у точки c_{ij} i-тая и j-тая координаты равны -1, а остальные шесть координат равны 1. Поскольку все 27 точек лежат в гиперплоскостях $x_1 + x_2 + \dots + x_6 = 2$ и $x_7 + x_8 = 2$, множество T лежит в шестимерном евклидовом пространстве.

Упражнение 8. (а) Проверьте, что расстояние между любыми двумя точками множества T равно 4 или $\sqrt{8}$.

(б) Проверьте, что T лежит на сфере радиуса $4/\sqrt{3}$ с центром $\mathbf{q} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1, 1\right)$ и что $\mathbf{q} \in \langle T \rangle$.

(в) Пусть вершинами полного графа являются точки множества T, причем две вершины соединены красным ребром, если расстояние между ними равно 4, и зеленым ребром, если расстояние между ними равно $\sqrt{8}$. Проверьте, что эта раскраска — сильно регулярная с параметрами $(27, 10, 1, 5)$. Соответствующий сильно регулярный граф называется *графом Шлефли.*

Нетрудно проверить, что множество $T \cup \{\mathbf{q}\}$ является строго равнобедренным. Если S — произвольное строго равнобедренное множество в $\mathbb{E}^6$ и $|S| \geqslant 28$, то $\operatorname{dist}(S) \geqslant 3$, и мы применяем теорему 2. Если $\dim X = 6$, мы получаем описанное множество из 28 точек. Если $\dim X = 5, 4, 3, 2, 1$, то $|S|$ не превосходит $16+2$, $10+6$, $6+7$, $5+11$, $2+17$, соответственно, так что $|S| \leqslant 28$.

$$n = 7$$

Наибольшее (и единственное с точностью до подобия) множество с двумя расстояниями в $\mathbb{E}^7$, состоит, согласно [10], из следующих 29 точек пространства $\mathbb{E}^8$:

семь точек, у которых восьмая и еще какая-нибудь координата равны -1, а остальные шесть координат равны 1; 21 точка, у которых две из первых семи координат равны 2, а остальные шесть координат равны 0; точка $(1,1,1,1,1,1,1,-3)$.

Все эти точки лежат в гиперплоскости $x_1+x_2+\dots+x_8=4$, т.е., в семимерном евклидовом пространстве.

Упражнение 9. Проверьте, что расстояние между любыми двумя из указанных 29 точек равно 4 или $2\sqrt{2}$. Докажите, что данные 29 точек не лежат на одной сфере.

Данное множество из 29 точек является строго равнобедренным, но оно не пригодно для получения большего равнобедренного множества. Если мы присоединим к множеству Λ_7 (см. упражнение 2) центр описанной сферы, мы получим другое строго равнобедренное множество из 29 точек. Кроме того, если мы выберем в какой-нибудь гиперплоскости π единственное множество T из 27 точек с двумя расстояниями (см. случай $n=6$), то это множество лежит на сфере с центром в точке $\mathbf{q}\in\pi$, и потому, присоединив к $\mathbf{q}$ одну из точек пересечения этой сферы и прямой ℓ, проходящей через $\mathbf{q}$ и ортогональной гиперплоскости π, мы снова получим строго равнобедренное множество из 29 точек в $\mathbb{E}^7$. Более того, мы можем получить бесконечно много строго равнобедренных множеств из 29 точек вида $S\cup\{\mathbf{u},\mathbf{v}\}$, где $\mathbf{u},\mathbf{v}\in\ell$ выбраны так, что расстояние между $\mathbf{u}$ и $\mathbf{v}$ равно расстоянию между одной из этих точек и любой точкой множества T (рис. 3). Мы теперь докажем, что в $\mathbb{E}^7$ не существует строго равнобедренного множества из 30 точек.

Пусть $S\subset\mathbb{E}^7$ — строго равнобедренное множество и пусть $|S|=30$. Тогда $\operatorname{dist}(S)\geqslant 3$, и мы, как всегда, применяем теорему 2. Если $\dim X=7$, то $|Y|=1$, так что $|X|=29$. Однако, единственное множество с двумя расстояниями из 29 точек в $\mathbb{E}^7$ не лежит на сфере. Если $\dim X=6,5,4,3,2$, то $\dim Y\leqslant 1,2,3,4,5$, соответственно, и тогда $|S|\leqslant 27+2$, $16+6$, $10+8$, $6+11$, $5+17$, так что $|S|<30$.

Пусть $\dim X=1$, $\dim Y=6$, так что $|X|=2$, $|Y|=28$. Мы можем предположить, что множество S лежит в $\mathbb{E}^8$ и при этом $Y=T\cup\{\mathbf{q}\}$, где множество T и точка $\mathbf{q}$ те же, что и в случае $n=6$. Пусть $\pi=\langle T\rangle=\langle Y\rangle$. Тогда $\dim\pi=6$. Так как множество с двумя расстояниями в пятимерном евклидовом пространстве состоит не более чем из 16 точек, любые 17 точек множества T порождают π. Рассмотрим полный граф, вершины которого — точки множества T, а рёбра раскрашены, как указано в упражнении 8(в). Мы можем применить лемму 2, где $\mathbf{p}$ — точка множества X, не лежащая в π.

Если $\mathbf{p}$ — центр сферы, содержащий множество T, то ортогональная проекция точки $\mathbf{p}$ на подпространство π — точка $\mathbf{q}$, и тогда две точки множества X и точка $\mathbf{q}$ лежат на одной прямой.

Предположим теперь, что $\operatorname{dist}(T \cup \mathbf{p}) = 2$. Если мы покажем, что точка $\mathbf{p}$ равноудалена от 17 точек множества S, то ортогональная проекция $\mathbf{p}$ на подпространство π — точка $\mathbf{q}$, и мы опять получим что три точки множества T лежат на одной прямой.

Пусть $\mathbf{p} = (p_1, p_2, \ldots, p_8)$. Применяя описание точек $\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_{ij}$, данное в случае $n = 6$, мы получим, что для любых четырех различных индексов $i, j, k, l \in \{1, 2, 3, 4, 5, 6\}$

$$\begin{aligned} &\|\mathbf{p} - \mathbf{a}_i\| = \|\mathbf{p} - \mathbf{a}_j\| \quad \Leftrightarrow \quad \|\mathbf{p} - \mathbf{b}_i\| = \|\mathbf{p} = \mathbf{b}_j\| \quad \Leftrightarrow \quad p_i = p_j, \\ &\|\mathbf{p} - \mathbf{a}_i\| < \|\mathbf{p} - \mathbf{a}_j\| \quad \Leftrightarrow \quad \|\mathbf{p} - \mathbf{b}_i\| < \|\mathbf{p} = \mathbf{b}_j\| \quad \Leftrightarrow \quad p_i = p_j + 2; \end{aligned} \tag{6}$$

$$\begin{aligned} &\|\mathbf{p} - \mathbf{a}_i\| = \|\mathbf{p} - \mathbf{a}_i\| \quad \Leftrightarrow \quad p_7 = p_8, \\ &\|\mathbf{p} - \mathbf{a}_i\| < \|\mathbf{p} - \mathbf{a}_i\| \quad \Leftrightarrow \quad p_7 = p_8 + 2, \\ &\|\mathbf{p} - \mathbf{a}_i\| > \|\mathbf{p} - \mathbf{a}_i\| \quad \Leftrightarrow \quad p_7 = p_8 - 2; \end{aligned} \tag{7}$$

$$\begin{aligned} &\|\mathbf{p} - \mathbf{c}_{ij}\| = \|\mathbf{p} - \mathbf{c}_{ik}\| \quad \Leftrightarrow \quad p_j = p_k, \\ &\|\mathbf{p} - \mathbf{c}_{ij}\| > \|\mathbf{p} - \mathbf{c}_{ik}\| \quad \Leftrightarrow \quad p_j = p_k + 2; \end{aligned} \tag{8}$$

$$\begin{aligned} &\|\mathbf{p} - \mathbf{c}_{ij}\| = \|\mathbf{p} - \mathbf{c}_{kl}\| \quad \Leftrightarrow \quad p_i + p_j = p_k + p_l, \\ &\|\mathbf{p} - \mathbf{c}_{ij}\| > \|\mathbf{p} - \mathbf{c}_{kl}\| \quad \Leftrightarrow \quad p_i + p_j = p_k + p_l + 2. \end{aligned} \tag{9}$$

Так как среди расстояний $\|\mathbf{p} - \mathbf{a}_i\|$ есть не более двух различных, из (6) следует, что среди первых шести координат точки $\mathbf{p}$ есть не более двух различных. Если первые шесть координат точки $\mathbf{p}$ равны друг другу, то из (6) и (8) следует, что точка $\mathbf{p}$ равноудалена от по крайней мере 21 точки множества S. Поэтому предположим, что среди первых шести координат точки $\mathbf{p}$ встречаются ровно два разных значения. Тогда из (9) следует, что одно из этих значений встречается только один раз, а из (6) следует, что $p_7 = p_8$. Пусть $p_2 = p_3 = p_4 = P_5 = p_6$. Тогда точка $\mathbf{p}$ равноудалена от 10 точек $\mathbf{a}_i$, $\mathbf{b}_i$, $2 \leqslant i \leqslant 6$. Нетрудно проверить, что подпространство $\langle \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5, \mathbf{a}_6 \rangle$ является пересечением гиперплоскостей $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 2$, $x_7 + x_8 = 2$ и $x_1 = 0$. Так как ни одна из точек $\mathbf{c}_{ij}$ не лежит в гиперплоскости $x_1 = 0$, то 10 точек $\mathbf{a}_i$, $\mathbf{b}_i$, $2 \leqslant i \leqslant 6$, и любая из точек $\mathbf{c}_{ij}$ порождают π.

Таким образом, $\mathbb{E}^7$ не содержит строго равнобедренных множеств из 30 точек.

$n = 8$

В работе [10] найдено множество $S \subset \mathbb{E}^9$, состоящее из 9 точек, у которых восемь координат равны $\frac{1}{3}$ и одна координата равна $-\frac{2}{3}$, и 36 точек с двумя координатами, равными 1, и семью координатами, равными 0. Расстояние между любыми двумя из этих 45 точек равно 2 или $\sqrt{2}$, и все они лежат в гиперплоскости $x_1 + x_2 + \cdots + x_9 = 2$, т. е. в восьмимерном евклидовом пространстве. Поскольку $|S| = C_{10}^2$, это наибольшее возможное равнобедренное и строго равнобедренное множество в $\mathbb{E}^8$.

§5. Задача Эрдёша в других пространствах

Определение равнобедренного множества имеет смысл в любом метрическом пространстве. Теорема 4 (в части, не касающейся ортогональности и размерности) тоже справедлива в любом метрическом пространстве. Мы рассмотрим три примера. В первом примере задача Эрдёша рассматривается во всей полноте, в двух других примерах получена оценка числа точек в множестве с двумя расстояниями, что открывает возможности для исследования равнобедренных множеств в этих пространствах.

1. Бинарное пространство Хэмминга

Бинарное пространство Хэмминга размерности n — это множество H_n всех последовательностей $\mathbf{a} = (a_1, a_2, \ldots, a_n)$, где $a_i \in \{0, 1\}$. Расстояние $h(\mathbf{a}, \mathbf{b})$ между точками $\mathbf{a}$ и $\mathbf{b} = (b_1, b_2, \ldots, b_n$ пространства H_n определяется как число индексов i, для которых $a_i \neq b_i$. Пространство H_n естественно интерпретировать как множество вершин единичного куба в $\mathbb{E}^n$. Поскольку расстояние между двумя точками пространства H_n равно квадрату евклидова расстояния между этими точками как вершинами куба, равнобедренное множество в H_n является строго равнобедренным множеством в $\mathbb{E}^n$ (никакие три вершины куба не лежат на одной прямой). Перечислим кратко основные результаты, касающиеся задачи Эрдёша в пространстве H_n. Доказательства и ссылки даны в работе [6].

Пусть S — равнобедренное множество в H_n.

(a) Если $\operatorname{dist}(S) = 1$, то $|S| \leqslant n+1$. При этом равенство возможно в том и только в том случае, если существует матрица Адамара порядка $n + 1$. Матрица Адамара — это квадратная матрица с попарно ортогональными строками, все элементы которой равны ± 1. Легко доказать, что порядок матрицы Адамара делится на 4 (или равен 1 или 2). В 1893 году Жак Адамар предположил, что для любого натурального n существует матрица Адамара порядка $4n$. Предположение Адамара до сих пор не доказано и не опровергнуто. Наименьшее значение n, для которого существование матрицы Адамара порядка n остается открытым — 668.

(б) Если $\operatorname{dist}(S) = 2$, то $|S| \leqslant 1 + C_{n+1}^2$. При этом равенство имеет место в том и только в том случае, если (1) $n = 2$ и $S = H_2$ или (2) $n = 5$ и S — множество из 16 вершин пятимерного куба, никакие две из которых не принадлежат одному ребру.

(в) Если $\operatorname{dist}(S) = 2$ и множество S, рассматриваемое как подмножество пространства $\mathbb{E}^n$, лежит в какой-нибудь гиперплоскости, то $|S| \leqslant C_n^2$. Пусть $S = \{\mathbf{x} \in H_n : h(\mathbf{0}, \mathbf{x}) = 2\}$. Тогда $\operatorname{dist}(S) \leqslant 2$, $|S| = C_n^2$ и S лежит в гиперплоскости $x_1 + x_2 + \cdots + x_n = 2$.

(г) Если $\operatorname{dist}(S) \geqslant 3$, то $|S| \leqslant 1 + C_n^2$. Пусть $\mathbf{u} = (1, 1, \ldots, 1)$ и пусть $S = \{\mathbf{u}\} \cup \{\mathbf{x} \in H_n : h(\mathbf{0}, \mathbf{x}) = 2\}$. Тогда S — равнобедренное множество и $|S| = 1 + C_n^2$; если $n \geqslant 5$ и $n \neq 6$, то $\operatorname{dist}(S) \geqslant 3$.

(д) Если $n = 6$ и $\operatorname{dist}(S) \geqslant 3$, то $|S| \leqslant 12$; если $n \leqslant 4$, то $\operatorname{dist}(S) \leqslant 2$.

2. Сферическое пространство

Сферическое пространство $\mathbb{S}^n$ — это множество всех точек $\mathbf{u} \in \mathbb{E}^{n+1}$ таких, что $\|\mathbf{u}\| = 1$. Расстояние между точками $\mathbf{u}, \mathbf{v} \in \mathbb{S}^n$ определяется как $\arccos(\langle \mathbf{u}, \mathbf{v} \rangle)$.

Пусть S — множество в $\mathbb{S}^n$ с двумя ненулевыми расстояниями, α и β. Мусин [11] доказал, что если $\cos\alpha + \cos\beta \geqslant 0$, то $|S| \leqslant C_{n+1}^2$. Если $n \geqslant 7$ и $S = \dfrac{1}{\sqrt{2}}\Lambda_n$ (см. упражнение 2), то $\cos\alpha + \cos\beta \geqslant 0$ и $|S| = C_{n+1}^2$. Если $\cos\alpha + \cos\beta < 0$ то, как показано в [11], $|S| < C_{n+1}^2$ при $7 \leqslant n \leqslant 39$, исключая $n = 22$ и $n = 23$.

3. Гиперболическое пространство

Гиперболическое пространство $\mathbb{H}^n$ — это множество всех точек $\mathbf{u} = (u_0, u_1, \ldots, u_n) \in \mathbb{E}^{n+1}$ таких, что $u_0^2 - u_1^2 - \cdots - u_n^2 = 1$. Расстояние между точками $\mathbf{u}$ и $v = (v_0, v_1, \ldots, v_n)$ определяется как $\operatorname{arcch}(u_0 v_0 - u_1 v_1 - \ldots - u_n v_n)$. Как показано в [2], если $S \subset \mathbb{H}^n$, $\operatorname{dist}(S) = s$, то $|S| \leqslant C_{n+s}^s$. Для каждого $n \geqslant 10$ Лисонек [10] приводит пример множества $s \subset \mathbb{H}^n$, такого, что $\operatorname{dist}(S) = 2$ и $|S| = C_{n+2}^2$.

Список литературы

[1] Bannai Ei., Bannai Et., Stanton D. *An upper bound for the cardinality of an s-distance set in real Euclidean space*, II // Combinatorica. Vol. 3, 1983. P. 147–152.

[2] Blokhuis A. *Few-distance sets.* CWI Tract. 7, 1984. P. 1–70.

[3] Blumenthal L. M. *Theory and Applications of Distance Geometry.* Oxford, 1953.

[4] Croft H. T. 9-*point and* 7-*point configurations in* 3-*space* // Proc. London Math. Soc. Vol. s3–12, issue 1, 1962. P. 400–424.

[5] Einhorn S. J., Schoenberg I. J. *On Euclidean sets having only two distances between points* // Indagationes Mathematicae. Vol. 28, 1966. P. 479–504.

[6] Ionin Y. J. *Isosceles sets* // The Electronic Journal of Combinatorics. Vol. 16, 2009, R 141. `www.combinatorics.org`.

[7] Kelly L. M. *E 735* // The American Mathematical Monthly. Vol. 54, 1947. P. 227–229.

[8] Kido H. *Classification of isosceles eight-point sets in three-dimensional Euclidean space* // Europ. J. Combin. Vol. 27, 2006. P. 329–341.

[9] Larman D. G., Rogers C. A., Seidel J. J. *On two-distance sets in Euclidean space* // Bull. London Math. Soc. Vol. 9, 1977. P. 261–267.

[10] Lisonek P. *New maximal two-distance sets* // J. Combin. Theory Ser. A. Vol. 77, 1997. P. 318–338.

[11] Musin O. R. *Spherical two-distance sets* // J. Combin. Theory Ser. A. Vol. 116, 2009. P. 988–995.

Ю. И. Ионин, Central Michigan University, USA

Одна геометрическая задача, приводящая к биллиардному закону отражения

Г. А. Гальперин А. Ю. Плахов

Следующее геометрическое утверждение играет важную роль в задачах ньютоновской аэродинамики [2, 3]. Оно позволяет строить «невидимые объекты» типа криволинейного треугольника ABC, изображенного на рис. 4 в конце настоящей статьи. Цель настоящей заметки — доказательство этого утверждения.

Теорема. *Пусть F_1F_2C — прямоугольный треугольник с вершиной прямого угла в F_2, $\mathcal{E}$ и $\mathcal{H}$ — софокусные, с фокусами в точках F_1 и F_2, эллипс и гипербола, проходящие через точку C. (Мы рассматриваем только ту ветвь гиперболы $\mathcal{H}$, которая содержит C.) Рассмотрим луч с вершиной в точке F_1, пересекающий эллипс $\mathcal{E}$ и ветвь гиперболы $\mathcal{H}$ в некоторых точках A и B. Тогда отрезок F_2C образует равные углы с отрезками F_2A и F_2B: $\alpha = \beta$ (см. рис. 1).*

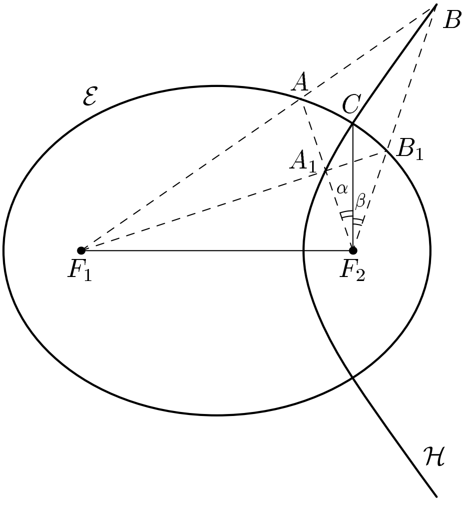


Рис. 1. $\alpha = \beta$

Отметим следующее свойство, вытекающее непосредственно из нашей теоремы.

Следствие. *Пусть A_1 — точка пересечения луча F_2A с ветвью гиперболы $\mathcal{H}$ и пусть луч F_1A_1 пересекает эллипс в точке B_1* (*рис.* 1). *Тогда, согласно теореме, точки B, B_1 и F_2 лежат на одной прямой. Иными словами, тройки точек F_1AB, $F_1A_1B_1$, F_2A_1A и F_2B_1B коллинеарны.*

Доказательство теоремы использует следующее характеристическое свойство биссектрисы треугольника.

Лемма. *Рассмотрим треугольник ABC и отрезок BD, соединяющий вершину B с точкой D, лежащей на противолежащей стороне AC. Обозначим $a_1 = AB$, $a_2 = BC$, $b_1 = AD$, $b_2 = DC$ и $f = BD$* (*см. рис.* 2). *Отрезок BD является биссектрисой угла B тогда и только тогда, когда $(a_1 + b_1)(a_2 - b_2) = f^2$.*

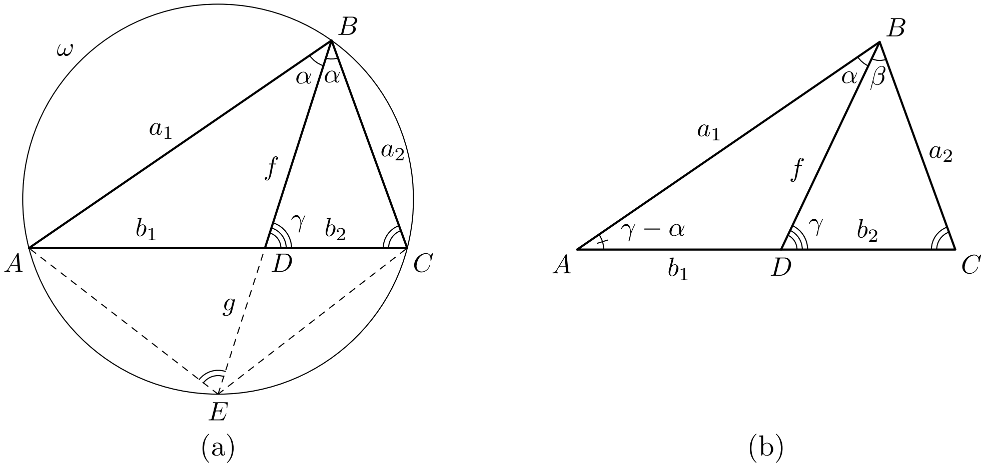


Рис. 2. Доказательство прямого (a) и обратного (b) утверждения о биссектрисе

Доказательство. Пусть $f = BD$ есть биссектриса угла B, проведенная к стороне AC. Докажем следующие три соотношения на величины a_1, a_2, b_1, b_2 и f:

1. $a_1/a_2 = b_1/b_2$;

2. $a_1a_2 - b_1b_2 = f^2$;

3. $(a_1 + b_1)(a_2 - b_2) = f^2$.

Равенства 1 и 2 хорошо известны; каждое из них также является характеристическим свойством биссектрисы треугольника.

Первое свойство вытекает из сравнения площадей:

$$\frac{a_1}{a_2} = \frac{\frac{1}{2}a_1 f \sin\alpha}{\frac{1}{2}a_2 f \sin\alpha} = \frac{S_{ABD}}{S_{BCD}} = \frac{\frac{1}{2}b_1 h}{\frac{1}{2}b_2 h} = \frac{b_1}{b_2}\,, \tag{1}$$

где $\alpha = \measuredangle ABD = \measuredangle CBD$, а h — высота, опущенная из вершины B на сторону AC.

Второе свойство биссектрисы основывается на понятии «степени» точки относительно окружности. Опишем окружность ω вокруг треугольника ABC. *Степенью точки* D внутри окружности ω называется произведение длин отрезков любой хорды, проходящей через D, на которую D делит эту хорду (все такие произведения одинаковы для данной точки D). Обозначив $DE = g$, получаем для точки D: $b_1 b_2 = fg$ (рис. 2).

Заметим, что $\triangle ABE$ подобен $\triangle DBC$ по двум углам:

$$\measuredangle ABE = \measuredangle DBC = \alpha \quad \text{и} \quad \measuredangle AEB = \measuredangle ACB = \frac{1}{2}\,\overset{\frown}{AB}\,.$$

Поэтому

$$\frac{a_1}{f+g} = \frac{f}{a_2}\,,$$

откуда

$$a_1 a_2 = f^2 + fg \quad \Rightarrow \quad f^2 = a_1 a_2 - fg = a_1 a_2 - b_1 b_2,$$

что и требовалось.

Приступим к доказательству того, что для биссектрисы f выполняется равенство 3, и обратно, отрезок BD, удовлетворяющий этому равенству, является биссектрисой. Отметим, что нам неизвестны какие-либо упоминания в литературе об этом свойстве.

Легко видеть, что алгебраические соотношения 1, 2 и 3 являются «линейно зависимыми»: из любых двух следует третье. Тем самым из свойств биссектрисы 1, 2 немедленно вытекает прямое утверждение: биссектриса f удовлетворяет условию 3.

Вывод обратного утверждения требует применения теоремы синусов и некоторой тригонометрии. Обозначим $\alpha = \measuredangle ABD$, $\beta = \measuredangle CBD$ и $\gamma = \measuredangle BDC$ (см. рис. 2(b)). Наша цель — доказать равенство $\alpha = \beta$. Применяя теорему синусов к $\triangle ABD$, имеем

$$\frac{a_1}{\sin\gamma} = \frac{b_1}{\sin\alpha} = \frac{f}{\sin(\gamma-\alpha)}\,,$$

а применяя теорему синусов к $\triangle BDC$, имеем

$$\frac{a_2}{\sin\gamma} = \frac{b_2}{\sin\beta} = \frac{f}{\sin(\gamma+\beta)}\,.$$

Отсюда находим

$$a_1 + b_1 = \frac{f}{\sin(\gamma-\alpha)}(\sin\gamma + \sin\alpha) = f\frac{\sin\dfrac{\gamma+\alpha}{2}}{\sin\dfrac{\gamma-\alpha}{2}}\,,$$

$$a_2 - b_2 = \frac{f}{\sin(\gamma+\beta)}(\sin\gamma - \sin\beta) = f\frac{\sin\dfrac{\gamma-\beta}{2}}{\sin\dfrac{\gamma+\beta}{2}}\,,$$

и, применяя условие 3, получаем

$$f^2\,\frac{\sin\dfrac{\gamma+\alpha}{2}\sin\dfrac{\gamma-\beta}{2}}{\sin\dfrac{\gamma-\alpha}{2}\sin\dfrac{\gamma+\beta}{2}} = f^2,$$

откуда

$$\sin\frac{\gamma+\alpha}{2}\sin\frac{\gamma-\beta}{2} = \sin\frac{\gamma-\alpha}{2}\sin\frac{\gamma+\beta}{2}\,,$$

$$\cos\frac{\alpha+\beta}{2} - \cos\left(\gamma + \frac{\alpha-\beta}{2}\right) = \cos\frac{\alpha+\beta}{2} - \cos\left(\gamma - \frac{\alpha-\beta}{2}\right),$$

$$\cos\left(\gamma + \frac{\alpha-\beta}{2}\right) = \cos\left(\gamma - \frac{\alpha-\beta}{2}\right).$$

Из последнего равенства и условий $0 < \alpha,\ \beta,\ \gamma < \pi$ следует, что $\alpha = \beta$, что и требовалось доказать.

Теперь приступим к доказательству теоремы.

Продлим отрезок BF_2 до второго пересечения с эллипсом в некоторой точке A'. Обозначим

$$f = F_1F_2,\ c = F_2C,\ a_1 = F_1A',\ b_1 = F_2A',\ a_2 = F_1B \text{ и } b_2 = F_2B$$

(см. рис. 3). Обозначим C' вторую точку пересечения эллипса с ветвью гиперболы $\mathcal{H}$. В силу фокального свойства эллипса выполнено равенство

$$F_1A' + F_2A' = F_1C' + F_2C',$$

то есть

$$a_1 + b_1 = \sqrt{f^2+c^2} + c. \tag{2}$$

Далее, в силу фокального свойства гиперболы имеем

$$F_1B - F_2B = F_1C - F_2C,$$

то есть

$$a_2 - b_2 = \sqrt{f^2+c^2} - c. \tag{3}$$

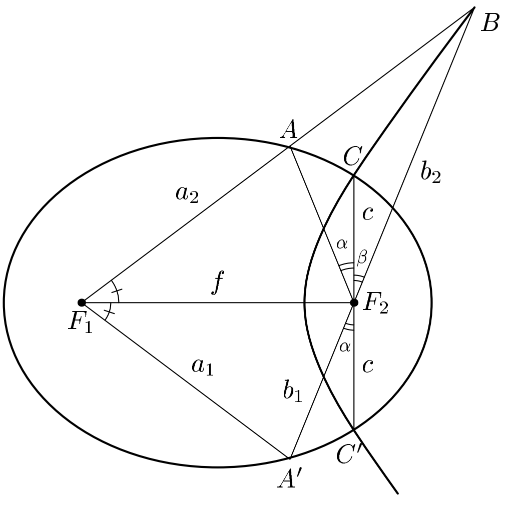


Рис. 3. Вспомогательное построение

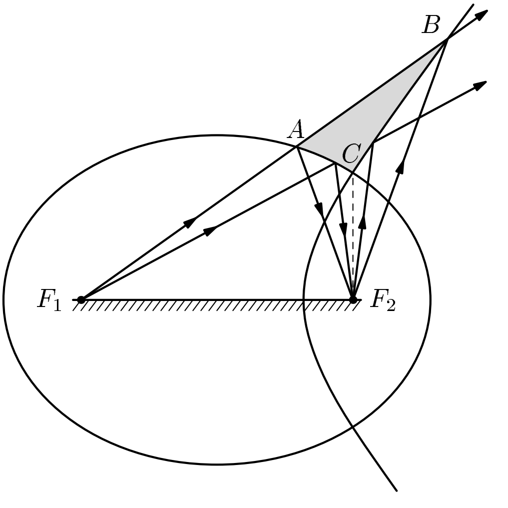


Рис. 4. Криволинейный треугольник ABC невидим из фокуса F_1: все выходящие из F_1 световые лучи огибают препятствие ACB так, как если бы его не было вовсе

Перемножая левые и правые части (2) и (3),получаем

$$(a_1 + b_1)(a_2 - b_2) = f^2,$$

и с учетом леммы заключаем, что F_1F_2 — биссектриса угла F_1 треугольника $A'F_1B$. Это, в свою очередь, означает, что точка A' симметрична A относительно прямой F_1F_2, и в силу симметрии имеем

$$\measuredangle AF_2C = \measuredangle A'F_2C'. \tag{4}$$

С другой стороны, углы $\measuredangle BF_2C$ и $\measuredangle A'F_2C'$ вертикальны, а значит, равны:

$$\measuredangle BF_2C = \measuredangle A'F_2C'. \tag{5}$$

Из (4) и (5) следует, что

$$\measuredangle AF_2C = \measuredangle BF_2C,$$

то есть $\alpha = \beta$. Теорема доказана.

Список литературы

[1] Гальперин Г. А. *Биллиардная формула, измеряющая расстояния в многомерном пространстве Лобачевского* // Математическое просвещение. Серия 3, вып. 8, 2004. С. 93–112.

[2] Плахов А. Ю. *Рассеяние в биллиардах и задачи ньютоновской аэродинамики* // Успехи математических наук. Т. 64, 2009. С. 97–166.

[3] Aleksenko A., Plakhov A. *Bodies of zero resistance and bodies invisible in one direction* // Nonlinearity. Vol. 22, 2009. P. 1247–1258.

Г. А. Гальперин, Eastern Illinois University, USA
E-mail: `ggalperin@eiu.edu`
А. Ю. Плахов, Aveiro University, Portugal
E-mail: `plakhov@ua.pt`

О минимуме максимума гладких функций

Г. Г. Магарил-Ильяев

В заметке приводятся необходимые и достаточные условия минимума функции, являющейся максимумом конечного числа гладких функций многих переменных. Для функций одного переменного «из картинок» сразу видно, что если $\widehat{x}$ — локальный минимум максимума функций f_1 и f_2, где $f_1(\widehat{x}) = f_2(\widehat{x})$ и, скажем, $f_1'(\widehat{x}) \leqslant f_2'(\widehat{x})$, то справедливо включение $0 \in [f_1'(\widehat{x}), f_2'(\widehat{x})]$. Достаточным это условие не является, как показывает пример функции $f(x) = \max(-x, \sin x)$. Но оно становится достаточным, если в некоторой окрестности $\widehat{x}$ графики функций f_1 и f_2 лежат выше своих касательных в точке $\widehat{x}$.

Все эти наблюдения mutatis mutandis имеют место и в более общей ситуации. Основной результат заключается в том, что если функция f есть максимум функций $f_1, \ldots, f_m$, определенных в окрестности некоторой точки $\widehat{x} \in \mathbb{R}^n$ и $f_1(\widehat{x}) = \ldots = f_m(\widehat{x})$, то принадлежность нуля выпуклой оболочке производных (градиентов) функций $f_1, \ldots, f_m$ в $\widehat{x}$ является необходимым условием локального минимума f в точке $\widehat{x}$, и достаточным, если график каждой из функций f_i, $1 \leqslant i \leqslant m$, лежит выше некоторой гиперплоскости, проходящей через $\widehat{x}$.

Для точной формулировки результата нужны некоторые начальные сведения о пространстве $\mathbb{R}^n$, понятие производной функции на $\mathbb{R}^n$, понятие выпуклого множества, выпуклой оболочки множества, а также теорема Минковского об отделимости точки от выпуклого замкнутого множества в $\mathbb{R}^n$. Ниже все эти факты (которые вполне стандартны) напоминаются, но если они известны читателю, то следующий параграф можно пропустить.

Предварительные сведения

Пространство $\mathbb{R}^n$ это совокупность всех упорядоченных наборов $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ из n действительных чисел (если $n = 1$, то это просто множество действительных чисел, и мы пишем $\mathbb{R}$ вместо $\mathbb{R}^1$), называемых *векторами* или *вектор-столбцами*, а числа x_i, $1 \leqslant i \leqslant n$ — *координатами вектора* x. Ради экономии места, элементы $\mathbb{R}^n$ будем записывать так

$x = (x_1, \ldots, x_n)^T$, где ^{T} обозначает *транспонирование* (перевод строки в столбец). В $\mathbb{R}^n$ естественным образом вводится операция (покоординатного) сложения векторов и операция (покоординатного) умножения вектора на число.

Пусть $a = (a_1, \ldots, a_n)$ — вектор-строка из n действительных чисел. Для каждого $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ обозначим $a \cdot x = \sum_{i=1}^n a_i x_i$. Ясно, что отображение $x \mapsto a \cdot x$ есть линейная функция (или, говорят, линейный функционал) на $\mathbb{R}^n$, т. е. $a \cdot (x_1 + x_2) = a \cdot x_1 + a \cdot x_1$ и $a \cdot \alpha x = \alpha a \cdot x$ для любых $x, x_1, x_2 \in \mathbb{R}^n$ и $\alpha \in \mathbb{R}$. Легко видеть, что и любой линейный функционал l на $\mathbb{R}^n$ задается подобным образом с $a = (l(e_1), \ldots, l(e_n))$, где $e_1 = (1, 0, \ldots, 0)^T, \ldots, e_n = (0, \ldots, 0, 1)^T$. Таким образом, если обозначить через $(\mathbb{R}^n)^*$ множество, элементы которого суть те же наборы из n действительных чисел, но расположенные в строку (с аналогичными операциями сложения и умножения на числа), то его можно отождествить с пространством всех линейных функционалов на $\mathbb{R}^n$, которое называют *сопряженным* или *двойственным* пространством к $\mathbb{R}^n$.

Далее, каждому $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ можно сопоставить линейный функционал $a \mapsto a \cdot x$ на $(\mathbb{R}^n)^*$ и тогда совершенно аналогично устанавливается, что сопряженное (двойственное) пространство к $(\mathbb{R}^n)^*$ можно отождествить с $\mathbb{R}^n$.

Пусть $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$. Величина $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$ называется *длиной* или *евклидовой нормой* вектора x, а величина $d(x, y) = |x - y|$ — *расстоянием* между векторами x и y.

Пусть $\widehat{x} \in \mathbb{R}^n$ и $\delta > 0$. Множество $U_{\mathbb{R}^n}(\widehat{x}, \delta) = \{\, x \in \mathbb{R}^n \mid |x - \widehat{x}| < \delta \,\}$ называется *открытым шаром* с центром в точке $\widehat{x}$ радиуса δ.

Множество $G \subset \mathbb{R}^n$ называется *открытым*, если с каждой своей точкой оно содержит и некоторый шар с центром в этой точке.

Любое открытое множество, содержащее данную точку, называется *окрестностью* этой точки.

Множество $F \subset \mathbb{R}^n$ называется *замкнутым*, если его дополнение $\mathbb{R}^n \setminus F$ — открытое множество.

Пусть $x_1, x_2 \in \mathbb{R}^n$. Множество $[x_1, x_2] = \{\, x \in \mathbb{R}^n \mid (1 - \alpha)x_1 + \alpha x_2,\ 0 \leqslant \alpha \leqslant 1 \,\}$ называется *отрезком* (соединяющим точки x_1 и x_2). В случае, когда $n = 1, 2$ и 3 — это обычный отрезок.

Непустое множество $A \subset \mathbb{R}^n$ называется *выпуклым*, если с любыми двумя своими точками оно содержит и отрезок, соединяющий эти точки.

Если A — произвольное непустое подмножество $\mathbb{R}^n$, то наименьшее выпуклое множество, содержащее A называется выпуклой оболочкой A и обозначается $\operatorname{co} A$.

Упражнение. Доказать, что $\operatorname{co} A$ — множество всех выпуклых комбинаций элементов из A, т. е. векторов вида $x = \alpha_1 x_1 + \ldots + \alpha_m x_m$, где $x_i \in A$, $\alpha_i \geqslant 0$, $i = 1, \ldots, m$, и $\sum_{i=1}^m \alpha_i = 1$.

Пусть $x^* \in (\mathbb{R}^n)^*$, $x^* \neq 0$ и $\gamma \in \mathbb{R}$. Множество $H = H(x^*, \gamma) = \{x \in \mathbb{R}^n \mid x^* \cdot x = \gamma\}$ называется *гиперплоскостью*. Гиперплоскость порождает два полупространства $H_+(x^*, \gamma) = \{x \in \mathbb{R}^n \mid x^* \cdot x \leqslant \gamma\}$ и $H_-(x^*, \gamma) = \{x \in \mathbb{R}^n \mid x^* \cdot x \geqslant \gamma\}$.

Пусть A и B — непустые подмножества $\mathbb{R}^n$. Говорят, что эти множества *отделимы*, если существует такая гиперплоскость, что A и B принадлежат различным полупространствам, порожденным этой гиперплоскостью.

Данное геометрическое определение отделимости равносильно, очевидно, следующему алгебраическому: множества A и B отделимы, если существует такой ненулевой элемент $x^* \in (\mathbb{R}^n)^*$, что

$$\sup_{a \in A} x^* \cdot a \leqslant \inf_{b \in B} x^* \cdot b.$$

Если неравенство строгое, то говорят, что множества A и B *строго отделимы*.

Теорема (Минковского об отделимости). *Пусть A — непустое выпуклое замкнутое подмножество $\mathbb{R}^n$ и $b \notin A$. Тогда множество A и точка b строго отделимы.*

Все перечисленные выше понятия (начиная с длины вектора и кончая теоремой Минковского) дословно переносятся на пространство $(\mathbb{R}^n)^*$ (двойственным к которому является пространство $\mathbb{R}^n$).

Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$. Говорят, что функция $f : U \to \mathbb{R}$ *дифференцируема в точке* $\widehat{x}$, если существует линейный функционал на $\mathbb{R}^n$, т. е. вектор $a = (a_1, \ldots, a_n) \in (\mathbb{R}^n)^*$ такой, что для всех $h \in \mathbb{R}^n$, для которых $\widehat{x} + h \in U$ справедливо представление

$$f(\widehat{x} + h) = f(\widehat{x}) + a \cdot h + r(h),$$

где $|r(h)|/|h| \to 0$ при $h \to 0$, или пишут $r(h) = o(h)$. Вектор a, определяемый этим представлением однозначно, называется *производной функции f в точке* $\widehat{x}$ и обозначается $f'(\widehat{x})$.

Беря в качестве h векторы $(h_1, 0, \ldots, 0)^T, \ldots, (0, \ldots, 0, h_n)^T$, получим, что $a_i = \partial f(\widehat{x})/\partial x_i$ — частная производная функции f по x_i в точке $\widehat{x}$, $i = 1, \ldots, n$, и таким образом, $f'(\widehat{x}) = (\partial f(\widehat{x})/\partial x_1, \ldots, \partial f(\widehat{x})/\partial x_n)$.

Основной результат и приложение

Пусть U — открытое подмножество $\mathbb{R}^n$, $f_i \colon U \to \mathbb{R}$, $i = 1, \ldots, m$. Рассмотрим задачу

$$f(x) = \max_{1 \leqslant i \leqslant m} f_i(x) \to \min, \quad x \in U, \qquad (P)$$

заключающуюся в нахождении таких точек $\widehat{x} \in U$, где функция f достигает минимума. Нас будут интересовать локальные минимумы f. Точка

$\widehat{x}$ — локальный минимум функции f, если существует такая окрестность $U_0 \subset U$ точки $\widehat{x}$, что $f(x) \geqslant f(\widehat{x})$ для всех $x \in U_0$.

Перед формулировкой теоремы дадим одно определение. Пусть U — окрестность точки $\widehat{x} \in \mathbb{R}^n$ и $f\colon U \to \mathbb{R}$. Скажем, что f *выпукла в точке* $\widehat{x}$, если найдутся такие $x^* \in (\mathbb{R}^n)^*$ и окрестность $U_0 \subset U$ точки $\widehat{x}$, что $f(x) - f(\widehat{x}) \geqslant x^* \cdot (x - \widehat{x})$ для всех $x \in U_0$. Другими словами, в окрестности U_0 график функции f лежит выше гиперплоскости $y = x^* \cdot (x - \widehat{x}) + f(\widehat{x})$.

Теорема (основная). *Пусть в задаче* (P) *функции* f_i, $1 \leqslant i \leqslant m$, *дифференцируемы в* $\widehat{x} \in U$ *и* $f_1(\widehat{x}) = \ldots = f_m(\widehat{x})$. *Тогда условие* $0 \in \operatorname{co}\{\, f_1'(\widehat{x}), \ldots, f_m'(\widehat{x}) \,\}$ *необходимо, а если функции* f_i, $1 \leqslant i \leqslant m$, *выпуклы в* $\widehat{x}$, *то и достаточно для того, чтобы* $\widehat{x}$ *было локальным минимумом функции* f.

Доказательство. Необходимость. Пусть $\widehat{x}$ — локальный минимум функции f. Обозначим $A = \operatorname{co}\{\, f_1'(\widehat{x}), \ldots, f_m'(\widehat{x}) \,\}$ и допустим, что $0 \notin A$. Тогда по теореме отделимости найдется такой вектор $\overline{x} \in \mathbb{R}^m$, что $\sup_{y \in A} y \cdot \overline{x} < 0$ или что тоже $\sup\{\, \sum_{i=1}^m \alpha_i f_i'(\widehat{x}) \cdot \overline{x} \mid \alpha_i \geqslant 0,\ 1 \leqslant i \leqslant m,\ \sum_{i=1}^m \alpha_i = 1 \,\} = \max_{1 \leqslant i \leqslant m}\{\, f_i'(\widehat{x}) \cdot \overline{x} \,\} < 0$. Отсюда и в силу дифференцируемости функций f_i, $i = 1, \ldots, m$, в точке $\widehat{x}$ следует, что для каждого $1 \leqslant i \leqslant m$ и всех достаточно малых $t > 0$ выполняется неравенство $f_i(\widehat{x} + t\overline{x}) = f_i(\widehat{x}) + t(f_i'(\widehat{x}) \cdot \overline{x} + o(t)/t) < f_i(\widehat{x})$, из которого вытекает, что $f(\widehat{x} + t\overline{x}) < f(\widehat{x})$ в противоречие с тем, что $\widehat{x}$ — локальный минимум f.

Достаточность. Пусть $0 \in A$. Тогда, очевидно, $\sup_{y \in A} y \cdot x \geqslant 0$ для всех $x \in \mathbb{R}^m$, что, как показано выше, равносильно неравенству $\max_{1 \leqslant i \leqslant m} \{\, f_i'(\widehat{x}) \cdot x \,\} \geqslant 0$. Для каждого $1 \leqslant i \leqslant m$ по условию существует такой элемент $x_i^* \in (\mathbb{R}^n)^*$, что $f_i(x) - f_i(\widehat{x}) \geqslant x_i^* \cdot (x - \widehat{x})$ для x, близких к $\widehat{x}$. Так как функция f_i, дифференцируема в $\widehat{x}$, то $x_i^* = f_i'(\widehat{x})$. Действительно, для любого $h \in \mathbb{R}^n$ и достаточно малых $t > 0$ имеем $f_i(\widehat{x} + th) - f_i(\widehat{x}) \geqslant t x_i^* \cdot h$, или $t^{-1}(f_i(\widehat{x} + th) - f_i(\widehat{x})) \geqslant x_i^* \cdot h$. Переходя к пределу при $t \to 0$, получаем, что $f_i'(\widehat{x}) \cdot h \geqslant x_i^* \cdot h$. Но h произвольно и поэтому $x_i^* = f_i'(\widehat{x})$. Тогда для x, достаточно близких к $\widehat{x}$, учитывая, что $f_1(\widehat{x}) = \ldots = f_m(\widehat{x})$, будем иметь $f(x) - f(\widehat{x}) = \max_{1 \leqslant i \leqslant m}\{\, f_i(x) - f_i(\widehat{x}) \,\} \geqslant \max_{1 \leqslant i \leqslant m}\{\, f_i'(\widehat{x}) \cdot (x - \widehat{x}) \,\} \geqslant 0$, т. е. $\widehat{x}$ — локальный минимум функции f.

Отметим, что если функции f_i, $1 \leqslant i \leqslant m$, дополнительно, выпуклы на $\mathbb{R}^n$, то необходимость может быть выведена из (весьма непростой) теоремы Дубовицкого – Милютина о субдифференциале максимума выпуклых функций (см. [1]). Постановка, которая рассмотрена здесь, представляется вполне естественной, но автор затрудняется дать какие-либо ссылки по этому поводу.

Приведем здесь одно приложение данного результата к экстремальным задачам. Пусть U — открытое подмножество $\mathbb{R}^n$, $f_i\colon U \to \mathbb{R}$, $i = 0, 1, \ldots, m$.

Рассмотрим задачу

$$f_0(x) \to \min, \quad f_i(x) \leqslant 0, \quad i = 1, \ldots, m. \tag{P_1}$$

Функция $\mathcal{L}\colon U \times \mathbb{R}^{m+1} \to \mathbb{R}$, $\mathcal{L}(x, \overline{\lambda}) = \sum_{i=0}^m \lambda_i f_i(x)$, где $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m)$, называется *функцией Лагранжа задачи* (P_1), а вектор $\overline{\lambda}$ — набором *множителей Лагранжа.* Справедлива следующая

Теорема (правило множителей Лагранжа для задачи (P_1)). *Пусть в задаче (P_1) все функции дифференцируемы в точке $\widehat{x} \in U$. Тогда, если $\widehat{x}$ является локальным минимумом в этой задаче, то найдется такой ненулевой набор множителей Лагранжа $\overline{\lambda} = (\lambda_0, \ldots, \lambda_m)$, что $\lambda_i \geqslant 0$, $i = 0, 1, \ldots, m$, и*

$$\mathcal{L}_x(\widehat{x}, \overline{\lambda}) = 0 \ \Leftrightarrow \ \sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0.$$

Доказательство. Заметим сначала, что можно считать, что $f_1(\widehat{x}) = \ldots = f_m(\widehat{x}) = 0$. В самом деле, отбросим те ограничения, для которых $f_i(\widehat{x}) < 0$. Тогда $\widehat{x}$ будет локальным минимумом и в новой задаче. Если для нее доказано утверждение, то дополнив найденный набор множителей Лагранжа нулевыми компонентами, соответствующие тем номерам, где $f_i(\widehat{x}) < 0$, получим утверждение в общем случае.

Если $\widehat{x}$ — локальный минимум в задаче (P_1) , то $\widehat{x}$ — локальный минимум и в такой задаче

$$f(x) = \max\{\, f_0(x) - f_0(\widehat{x}), f_1(x), \ldots, f_m(x)\,\} \to \min, \quad x \in U. \tag{i}$$

Действительно, если это не так, то в любой окрестности $\widehat{x}$ найдется точка $\widetilde{x}$ такая, что $f(\widetilde{x}) < 0$ (поскольку $f(\widehat{x}) = 0$). Это означает, что $\widetilde{x}$ — допустимая точка в задаче (P_1) и $f(\widetilde{x}) < f(\widehat{x})$, в противоречие с тем, что $\widehat{x}$ — локальный минимум. Итак, $\widehat{x}$ — локальный минимум в задаче (i). Тогда по основной теореме найдутся такие $\lambda_i \geqslant 0$, $i = 0, 1, \ldots, m$, $\sum_{i=0}^m \lambda_i = 1$, что $\sum_{i=0}^m \lambda_i f_i'(\widehat{x}) = 0$, а это и есть утверждение теоремы.

Список литературы

[1] Магарил-Ильяев Г. Г., Тихомиров В. М. *Выпуклый анализ и его приложения.* М.: Книжный дом «ЛИБРИКОМ», 2011. 3-е изд.

Г. Г. Магарил-Ильяев, Московский государственный институт радиотехники, электроники и автоматики (технический университет)

Конкурсы и олимпиады

Олимпиады: дверь в математику или спорт?

А. Я. Белов

1. Введение

По всему миру проводятся математические конкурсы и олимпиады. Появились специалисты по их проведению, возникла олимпиадная математика со своей методикой работы и своей литературой. Олимпиадный мир стал жить собственной жизнью, но его кажущаяся самодостаточность породила ряд проблем.

О пользе и вреде олимпиад, о том, как их проводить, постоянно ведутся кулуарные дискуссии. Во время заседаний методических комиссий, когда следует принимать решения, эти дискуссии превращаются в довольно трудные разговоры. Поэтому необходимо их провести открыто на страницах журнала. Сожалея о неизбежной субъективности, с благодарностью выслушаю замечания.

Ситуация вокруг олимпиад парадоксальна. С одной стороны, на олимпиады тратятся значительные ресурсы. Талантливых школьников и их учителей сводят вместе, прежде всего, математические олимпиады. У истоков олимпиадного движения стояли великие ученые. Многие школьники, особенно на периферии, получают математическое образование, нацеленное в первую очередь на подготовку к олимпиадам, со всеми его плюсами и минусами. (С этим надо считаться научным руководителям и организаторам учебного процесса.)

С другой стороны, распространены суждения о вреде олимпиад, зачастую весьма странные. Например, бытует мнение о том, что успехи на олимпиадах мало связаны с научной карьерой. Доходит до курьезов — до утверждений о том, что «среди крупных ученых нет победителей олимпиад», хотя среди бывших победителей олимпиад известных математиков

Математическое просвещение, сер. 3, вып. 15, 2011(187–203)

во много раз больше, чем среди неолимпиадников. Автор сталкивался с попытками одного преподавателя забрать своих учеников из его кружка. Один из аргументов, высказанных им школьникам, звучал так: «Концевич в олимпиадах не участвовал».

Такого рода суждения очень легко критиковать, а иногда — и высмеивать. Владимир Соловьев высказал мысль о том, что любая ложная социальная теория базируется на искажении некоторой правды, и для победы над «теорией» необходимо эту внутреннюю правду осознать и выявить. Мне представляется, критики олимпиад чувствуют серьезные проблемы, зачастую не умея их сформулировать.[1)]

Хотя олимпиадное движение играет значительную роль в математическом образовании школьников, особенно на периферии, олимпиадные деятели зачастую заявляют, что цель олимпиад ограничивается выявлением талантливых учащихся и формированием первоначального интереса к предмету. При этом декларируют, что «олимпиада — это не математика» (а то, что преподают в школе или колледже, это математика?), и любят подчеркивать, что далеко не все становятся математиками (хотя содержание олимпиадных задач должно быть полноценным, вне зависимости от будущей профессии).

В адрес некоторых задач со стороны некоторых активных деятелей олимпиадного движения высказывалась и такая критика, пусть и в полемическом запале: «*Плохо, что эта задача „из науки“*». Постоянно ведутся разговоры о спортивных достижениях. Победа школьника на олимпиаде иногда ценится выше его публикации в академическом журнале.

Против тезиса о важности научного содержания олимпиад высказывается и такой странный аргумент: «Олимпиады — это только небольшой жизненный эпизод». Даже если считать прямое и косвенное действие олимпиадного мира на подростка незначительным (что неправда), этот аргумент выглядит странно. Следуя такой логике, можно оправдать плохое качество любого отдельно взятого урока, ибо один урок мало что решает. Эту аргументацию не следует анализировать как научную, но выявлять скрытые мотивы. Я ни разу не слышал ничего подобного со стороны действующих ученых. Уход от ответственности за научное будущее ученика узким олимпиадным деятелям необходим, чтобы оправдать автономное существование олимпиадного мира.

Работа над этой статьей началась еще в 1996 году. В дальнейшем стало понятно, что проблемы носят не только внутрироссийский, но и международный характер. Так, участники команды одной страны, не набравшие

[1)] Есть математики-спортсмены, нацеленные на решение конкретных проблем, и математики-«домостроители», занимающиеся построением теорий, они чаще всего и становятся критиками олимпиад.

достаточного количества баллов, не были допущены до фотографирования. На международных олимпиадах по математике неоднократно наблюдались случаи коррупции (сообщение задач членам своей команды или «пробивание» в вариант задач, которые команда знает, пристрастная проверка работ, сбор компромата на участников других команд и т. д.), от чего сильно страдают и содержание вариантов, и результаты участников. Разговоры педагогов о школьниках напоминают разговоры о скаковых лошадях. Другой побудительной причиной написания этой статьи послужили внутренние проблемы, связанные с самим олимпиадно-педагогическим сообществом.

Благодарности. Автор признателен А. И. Буфетову, Э. Б. Винбергу, М. Н. Вялому, А. Домошницкому, А. К. Ковальджи, В. Н. Латышеву, Н. Х. Розову, В. М. Тихомирову, Б. Р. Френкину за полезные обсуждения и поддержку. Автор признателен своим коллегам по проведению олимпиад, благодаря которым эта статья была написана.

Зачем нужны олимпиады? Предваряя обсуждение, автор считает необходимым обозначить личное отношение к олимпиадам и олимпиадной математике. К сожалению, обоснование этой точки зрения выходит за рамки настоящей статьи. Автору близки позиции, изложенные в работах [4, 11].

Решение олимпиадных задач (разного уровня сложности) служит основой для почти всех математических кружков. Подготовка к олимпиадам оказывает значительное влияние на первоначальные занятия школьников математикой. Именно в решении трудной задачи может состоять достижение подростка. Более того, он зачастую оказывается в равном положении со взрослым. На трудных задачах вырабатывается интеллектуальная техника и соответствующие волевые качества. Но главное — сам факт достижения серьезной, но посильной цели в подростковом возрасте.

Часто утверждается, что олимпиадные умения не связаны с большой наукой, а зависимость между олимпиадными успехами и научной карьерой весьма слабая. Автор с этим категорически не согласен, поскольку препятствий для развития таланта множество. Прежде всего, человек даже очень талантливый встречается с теми или иными житейскими обстоятельствами. Они его могут надломить или даже сломать. Он может уйти в зарабатывание денег, столкнуться с семейными проблемами. Поэтому как бы мы ни выявляли таланты в юном возрасте, какая-то часть (и, увы, очень большая) из них в зрелом возрасте погаснет. (Небольшая популяция в Древней Греции поставила много великих ученых — больше, чем та же Греция произвела за последние две тысячи лет). Не исключено, что олимпиадные успехи больше говорят об изначальном таланте, чем будущее научное творчество.

Деятельность вокруг олимпиад стала заметным явлением в области современного математического образования. Их роль далеко не ограничивается обнаружением талантливых учащихся. Благодаря олимпиадной математике удается увидеть роль стандартных идей и рассуждений. Появились подборки олимпиадных задач по темам «принцип Дирихле», «правило крайнего», «инварианты» и др., объединенные единством метода.

Математики-непедагоги, как правило, не уделяют должное внимание «тривиальным» вещам, которые между тем играют исключительную роль как в мышлении математика, так и в подготовке к олимпиадам.

Благодаря олимпиадам возникли знаменитые книги Д. Пойа [8–10]. Да и облик так называемой «венгерской математики» (вспомним Пола Эрдёша) сформировался во многом под влиянием олимпиад.

Возможно, что выделение стандартных рассуждений может привести к революции и в самой математике, а в дальнейшем — и физике. Сама работа над изложением, казалось бы, известных результатов, часто приводила к открытиям. Так было со схемами Дынкина и с уравнением Гейзенберга. Олимпиадная математика с ее систематизацией идей и методов может послужить детонатором. Процесс детонации может начаться в комбинаторике. Мне представляется, что эта наука должна быть организована не так, как классическая область математики, а подобно некоторым тематическим подборкам олимпиадных задач. В основе ее организации должно лежать единство метода. Правильный учебник по комбинаторике должен быть чем-то вроде учебника шахматной игры или олимпиадного самоучителя.

Из истории олимпиад. Олимпиадное движение возникло свыше ста лет тому назад. Первые олимпиады состоялись в 1884 году в Австро-Венгрии. Они возникли из конкурсных экзаменов. Затем олимпиады появились в Венгрии. В дальнейшем олимпиадное движение ширилось. Олимпиады вышли за рамки конкурсных задач. В 30-е годы по инициативе Б. Н. Делоне возникли Ленинградские, а затем — Московские городские олимпиады (впервые в СССР математическая олимпиада для школьников была проведена в Тбилиси в декабре 1933 г.). У истоков олимпиад в СССР стояли ведущие ученые — А. Н. Колмогоров, И. М. Гельфанд, П. С. Александров, С. Л. Соболев, Л. Г. Шнирельман и другие. В дальнейшем возникла целая система национальных и международных олимпиад. Олимпиадный мир стал жить собственной жизнью. До недавнего времени его лидерами были хорошие математики (в том числе и создатель питерской олимпиадной школы, которую сейчас некоторые критикуют за излишне спортивную направленность). Не все знают, что создатель Турнира городов Н. Н. Константинов имеет красивые математические результаты.

Подробнее об истории математических олимпиад см. предисловия к книгам [2, 6], а также размышления В. М. Тихомирова в книге [12].

Кто сейчас делает олимпиады? Хотя у истоков олимпиад стояли великие ученые, в последующем уровень олимпиадных деятелей постепенно снижался. (Затем — раскол научного сообщества, вызванный проблемами 70-х – начала 80-х годов, события 90-х годов усилили эту тенденцию.) Ослабла связь олимпиад с научным сообществом. Появились так называемые «олимпиадные функционеры», т. е. специалисты по организации и проведению олимпиад. В жюри многих турниров высокого уровня почти не осталось профессиональных математиков даже среднего уровня. В последние годы в некоторых странах появились олимпиадные деятели, представляющие собой «нематематиков» и в то же время пытающиеся доминировать в олимпиадном мире, иногда откровенно противопоставляя себя научному сообществу.

Возникшая проблема является относительно новой. Как мне кажется, она в значительно меньшей мере наблюдается в олимпиадах по другим предметам, в силу меньшей развитости соответствующих субкультур. Чтобы «олимпиадная математика» смогла сложиться, был необходим первоначальный приток идей из большой науки. Сейчас, однако, такая необходимость многими олимпиадными деятелями не только не ощущается, но ими оказывается сопротивление переносу идей из этого источника.

1.1. О вкусах спорят!

На это мне указал замечательный математик и человек, ныне покойный, Ромен Васильевич Плыкин (он был организатором и председателем жюри Всероссийских конференций школьников). Очевидно, что есть «прекрасное» и «безобразное» в жизни, есть понятие плохого и хорошего вкуса в живописи, в одежде и в еде. Можно иметь предпочтения в живописи или предпочитать китайскую кухню французской. Но если вкус человека, разбирающегося во французской кухне, вине или китайском чае, заслуживает уважения, то вкус любителя фастфуда — нет (потребитель в обществе потребления и потреблять-то не умеет!).

Любая сильная идея является в обличии красоты (beauty is power itself). Значение математика определяется не только его «пробивной силой» (т. е. возможностью «пробить» трудную задачу), но и вкусом, которые, впрочем, тесно связаны. За свой вкус математик отвечает своей судьбой. Плохая эстетика задач наносит ущерб учащимся.

На мой взгляд, имеется группировка вкусовых предпочтений членов жюри олимпиад в зависимости от того, являются ли они действующими учеными или только функционерами. Если этот факт получит дополнительное подтверждение, то о плохих или хороших вкусах можно будет говорить более объективно.

В свое время за счет математического профессионализма и, как следствие, лучших эстетических критериев подбора задач, в одном областном центре удалось добиться результатов лучше, чем в мегаполисах. (В дальнейшем финансирование талантливых детей в этом регионе пошло по сомнительному направлению, в том числе математические лагеря перестали проводиться.)

2. Два подхода к олимпиадам

В олимпиадном мире сложились две ценностные ориентации. Они проявляются во всем: в подборе задач, выработке критериев оценок, и — что немаловажно — отражаются на роли и авторитете тех или иных личностей и, как следствие, на кадровых вопросах. Проявляются эти подходы и в организации математических лагерей. В значительной степени люди привержены одному из этих подходов, причем не всегда осознанно. Поэтому автор пытается дать описание, условно говоря, «научного» и «спортивного» стиля олимпиад.

2.1. Перерождение в большой спорт

В последнее время в олимпиадном мире усилился и доминирует «спортивный» подход. Я получил публичный упрек от одного олимпиадного деятеля, что для меня «*олимпиада лишь средство обучения математике, а для него — СПОРТИВНОЕ СОРЕВНОВАНИЕ*».

Распространение подобной точки зрения связано с тремя группами причин. Во-первых, с общим спортивным духом нашего времени. Во-вторых, с недостатком действующих математиков в олимпиадном мире и, как следствие, с ухудшением качества кадров. И в-третьих, это следствие логики развития олимпиадного движения без обратных связей, которой надо противостоять. (Иногда в предметах с менее долгой олимпиадной традицией научный уровень жюри бывает выше, поскольку первоначальный импульс в них задают крупные ученые, по той же причине бывает выше научный уровень жюри в странах с более молодой олимпиадной традицией). Ситуация усугубляется влиянием бизнеса, который в ряде случаев извращает творческие конкурсы (не только по математике).

Спортивный подход выражается фразой: «олимпиада — это спорт по решению головоломок». Этот лозунг влечет за собой многое: усиливается тренерство, ужесточаются формальные требования и, соответственно, критерии оценок. Так, если спортсмен переступит черту на 10 см, а прыгнет на 10 метров, ему не зачтут прыжок 9,9 м, его прыжок не зачтут вовсе.

Показательны слова одного из олимпиадных деятелей, сказанные во время обсуждения описки учащегося: «А если тебе зарплату не так подсчитали?» Налицо непонимание цели и смысла олимпиад.

К придумыванию новых задач относятся как к составлению шахматных этюдов и головоломок, только вместо фигур комбинируют объекты из школьной программы и стандартные олимпиадные темы. А в шахматах вопроса «откуда такое расположение фигур» просто не возникает. Комбинировать пытаются всё со всем. Особенно ценится внешняя обертка. Отношение к задачам как к головоломкам ведет к возникновению химер — когда комбинируется несовместимое по своей внутренней природе. Вот типичные примеры такого рода задач:

1. *Можно ли расставить числа от* 1 *до* 100 *в ряд так, чтобы сумма любых трех, идущих подряд, была простым числом?*

2. *Стороны треугольника — простые числа. Может ли его площадь быть целым числом?*

В этих двух задачах простота числа притянута искусственно.

Или еще —

3. *Найдите все целые числа, равные сумме факториалов своих цифр.*

Это мертвая математика.

Примеры отторжения содержательных задач (разных авторов). Речь пойдет о вкусе жюри, а не об аргументации, связанной с известностью или уровнем сложности задач.

Следующие две задачи были сочтены методической комиссией «неестественными».

1. *Поезд ехал один час от пункта A в пункт B, проехав* 60 *км. Доказать, что в какой-то момент его ускорение было не менее* 240 *км/ч*^{2}.

2. *Плоскость покрыта единичными кругами. Докажите, что некоторая точка покрыта не менее трех раз.*

Однако вторая задача отражает в простейшей форме фундаментальное понятие *топологической размерности*. Пространство имеет *размерность* n, когда имеются сколь угодно мелкие покрытия без перекрытий по $n+2$, но нельзя избавиться от перекрытий по $n+1$. У того, кому она кажется неестественной, скорее всего, плохой вкус.

На одном фестивале была предложена задача по нахождению угла между некоторыми диагоналями правильного додекаэдра. Идея решения состояла в рассмотрении вписанного куба. Эта задача была отвергнута как «неолимпиадная». Однако в задачных конкурсах до недавнего времени использовались разного рода стереометрические «монстры».

Заметим, что способный от природы школьник всё же может увидеть куб, вписанный в додекаэдр, а натасканный на «стандартные» олимпиадные темы «спортсмен», скорее всего, не увидит.

С другой стороны, многие задачи, например, на построение инвариантов, могут быть решены только учащимися, хорошо владеющими этой идеей (которая, к сожалению, даже намеками не входит в школьный курс). Здесь возникает проблема «джентльменского набора» идей и методов, без владения которыми «самородок» не достигнет больших успехов на олимпиаде.

Еще пример «неолимпиадной задачи».

Ломаная делит круг на две равные части. Доказать, что она проходит через его центр. Стиль решения этой задачи непривычен для олимпиадных деятелей. Тут дело вовсе не в трюке. Надо осознать, что значит *две равные части.* Это значит, что *есть движение, переводящее одну часть в другую.* А все типы движений плоскости описаны в теореме Шаля. Далее следует небольшой перебор. (Подробнее — см. «Математическое просвещение», сер. 3, вып. 6, 2002. С. 139–140.)

Или еще пример стереометрической задачи, отторжение которой говорит о дурном вкусе. *Можно ли разбить пространство на усеченные октаэдры*?

Разговор о конкретных вариантах олимпиад, плохих и хороших задачах давно назрел. К сожалению, здесь мы имеем возможность только поставить вопрос о его необходимости.

Причины появления задач-химер и отторжения содержательных задач. Дело в том, что поучительная задача чему-то учит. Но тогда и обратно — решение такой задачи непредсказуемо зависит от особенностей решателя и его культуры. Если процесс решения задачи оказывает воздействие на культуру решателя, то его результаты, в свою очередь, должны от этой культуры зависеть. Эта зависимость тем менее предсказуема, чем более глубокой оказывается задача. Следовательно, такая задача неудобна в плане оценки ее сложности.

Спортивный принцип предполагает стандартизацию. Одна из его основ — использование относительно стандартных приемов решения задач. При решении искусственных задач участники более равны, а самые «равные» должны получать премии.

Большой спорт тяготеет к ограничению поля деятельности и четкой формализации правил. Поэтому не случайна узость тематики задач, отсюда опасность вырождения олимпиад. Кроме того, стиль решения содержательной задачи (за которой стоит целое поле идей и сюжетов) непривычен для нематематика, а следовательно, не соответствует его вкусам.

Манипуляторство при проведении занятий. Разница между содержательным и спортивным стилем олимпиад примерно такая же, как между задачами придумать способ сборки кубика Рубика и соревнованиями на скорость его сборки. Спортивный стиль, «головоломочность» оказывают влияние и на ведение занятий. Внимание смещается с внутренней сущности на формальные манипуляции материалом.

Применительно к преподаванию это приводит не только к упору на натаскивание к олимпиадам. Парадоксальным образом зачастую наблюдается любовь тренеров к ученым словам и манипулирование ими. (Примеры курьезов такого рода — мини-курс на тему: «три определения комплексного числа с доказательством их равносильности»[2)], абстрактно изучаются «n-арные операции», обсуждаются «результаты» типа такого: группа S_3 вкладывается в группу автоморфизмов свободной группы с тремя образующими и т. д. и т. п.)

На наш взгляд, важна связь учителя с живым источником, которая сама по себе служит опорой и дисциплинирующим началом, что позволяет быть менее формальным. Жесткость и формальность в преподавании связана и с узостью кругозора (чем уже, тем жестче). Отсутствие или слабость живой связи с наукой приводит к возрастанию роли внешней обертки, а сами математические понятия становятся чем-то вроде заклинаний.

Причины распространения формально-спортивного подхода. Помимо уже обсуждавшейся логики организации соревнований в большом спорте есть и другая столь же важная причина распространения сверхспортивного стиля. Дело в том, что взгляд на математику как на науку о решении занимательных задач и головоломок самый доступный. Более того, он необходим при первоначальном знакомстве математикой, а следовательно, и в преподавании. Естественно, что этот самый доступный, безусловно, ценный и живой взгляд на математику, при всей его узости, получил распространение. Но при доведении его до крайности возникает сверхспортивный стиль.

Глубина понимания без узости объекта изучения сразу не достигается (с этим связан подростковый экстремизм типа «ничего мне не нужно, кроме геометрии»), но бывает лучше вначале достичь глубины, чем широты. В доступности спортивного стиля есть и позитивная сторона. Олимпиадный тренер даже спортивного толка может много сделать для развития образования в своем регионе.

[2)] Вполне осмысленно анализировать разные определения выпуклости фигур, поскольку это дается сравнительно легко и помогает решать задачи. А в «игре» с комплексными числами, с одной стороны, имеется стремление подражать «большой науке», а с другой — мало содержания.

Талантливый школьный учитель или местный деятель образования (а то обстоятельство, что он смог возвыситься над рутиной, говорит о многом), совершив усилие, иногда даже сверхусилие, входит в олимпиадный мир. Но чтобы ему понять, что мотивировки задач принципиально важны, требуется еще одно усилие, которое редко когда совершается. Помимо всего прочего, человек горд собой — у него уже есть результаты, а они зачастую ослепляют (вплоть до снобизма).

Социальные опасности. Олимпиадный мир — это только ветка, а не дерево с собственными корнями, поэтому терять связь с научным миром никак нельзя. Каковы бы ни были олимпиадные деятели, они светятся хотя бы отраженным светом. Нынешняя эволюция олимпиад предвещает мало хорошего. Если возобладает чисто спортивный подход, то математическая олимпиада, очевидно, не сможет конкурировать с иными соревнованиями — ни по зрелищности, ни по популярности. Возникнут новые деятели, специалисты по проведению игры типа «завоюй красный флажок» и они вытеснят старых.

Если олимпиадный деятель может отстранить ученого от участия в подготовке олимпиады, то, наверное, и чиновнику можно заменить олимпиадного деятеля?

Последняя олимпиадная реформа, связанная с отменой зонального этапа, невозможностью принимать участие в городской олимпиаде, не став победителем районной, и т. д., оказалась возможной в том числе из-за отсутствия в олимпиадном движении крупных ученых. Отчуждение между научным и олимпиадным сообществами может приводить к нежеланию последнего, чтобы в обсуждении вопросов образования участвовали ученые и, как следствие, к сдаче позиций перед чиновниками. В письме одного регионального олимпиадного деятеля к президенту РФ говорится о привлечении энтузиастов, работающих на местах, к обсуждению организационно-педагогических вопросов, но не говорится о привлечении ученых.

Когда на это обстоятельство было указано автору обращения, то он ответил, что ученые всё равно в олимпиадной деятельности не участвуют. Такое чувство, что у сообщества пропадает инстинкт самосохранения. Однако, благодаря связям московских олимпиадных деятелей с научным сообществом удалось существенно уменьшить вред от прошедшей олимпиадной реформы.

В последнее время к функциям олимпиады добавился способ поступления, альтернативный ЕГЭ, что создает дополнительные проблемы, в частности, с содержанием задач и секретностью подготовки вариантов, т. е. сужением круга лиц, готовящих варианты.

2.2. Подход к олимпиадам, имеющий источником науку

Девиз другой олимпиадной идеологии: *преподавание и олимпиады должны отражать науку*. Этой идеологии следовали всесоюзные олимпиады и старые олимпиады во многих странах (с единственной оговоркой — изначально олимпиады были близки по духу к вступительным экзаменам).

Впоследствии выяснилось, что олимпиада — это, в том числе, полигон для отработки новых тем и сюжетов (так был устроен задачник «Кванта»). В современных олимпиадах эта идеология присутствует не в чистом виде (в наиболее чистом виде — в Турнире городов, особенно на его летних конференциях [5], и в Московской олимпиаде, при всех их недостатках).

Этот девиз о связи с математикой имеет конкретное преломление.

Прежде всего, химерам в олимпиадах отказано в праве на существование. Ведь постановка задачи столь же важна, как и умение ее решать. Сила математика, как уже говорилось, во многом зависит от его вкуса. Совершенно необходимо иметь чутье на естественность. Без этого человек находится вне науки. Неестественная трудная задача на олимпиаде портит вкус и наносит огромный вред участникам.

Хорошая олимпиадная задача получается путем оформления идей и сюжетов из науки. Реже возникает новый сюжет в элементарной математике.

К спорту — отношение утилитарное, как к средству заставить подростка выложиться, достичь глубины, изучить технику. Приоритет содержательных соображений над спортивными, когда это возможно.

Математик, занимающийся большой проблемой, ищет связанные с ней задачи, где предполагаемые идеи решения работают в более простой ситуации. Невозможно также придумать несколько идей сразу — нужны промежуточные этапы, поэтому ценятся *продвижения*. Под наличием решения, как правило, понимается наличие «каркаса», т. е. основных идей. Такой подход в миниатюре можно переносить на олимпиадное творчество. Академические ценности по своей природе неформальны, но имеют огромное значение для воспитания будущих ученых.

К научному стилю, так или иначе, тяготеют практически все профессиональные математики, занимавшиеся олимпиадами.

2.3. Два подхода и человеческий фактор

В человеческих сообществах наличие разных ценностных ориентаций преломляется через субъективный фактор, амбиции и приводит к «политическим» последствиям. Понятно, что человек стремится создать среду обитания для себя, стремится установить такую структуру ценностей, чтобы его рейтинг поднялся. При этом от шкалы ценностей сильно зависит выбор авторитетов и ценность тех или иных «заслуг». С этим связаны дополнительные проблемы при организации олимпиадного сообщества.

В наше время происходят неприятные явления, друг друга усиливающие. Во-первых, повторим, ослабляется влияние научного сообщества.

Во-вторых, ослабляется внимание ведущих ученых к вопросам образования вообще и олимпиад в частности. Проблемы олимпиадного движения часто аргументируют именно этим явлением. Однако эта аргументация бывает лицемерна, поскольку главный фактор — вытеснение ученых. Традиции математического образования живы, и можно указать довольно много ученых докторского уровня, которых можно привлечь к организации и проведению олимпиад — не как «генералов», а для практической работы.

Для создавшегося положения характерно именно сочетание этих двух обстоятельств. Если бы ведущие ученые принимали деятельное участие в олимпиадах, как это было 20–30 лет назад, создавшееся положение не возникло бы. С другой стороны, если бы не было процесса вытеснения (об этом — ниже), то ряд людей докторского уровня участвовал бы в турнирах и олимпиадах, и уровень мероприятий был бы совсем иным.

Разумеется, без определенной преподавательской и некоторой олимпиадной квалификации сама по себе научная квалификация для составителя варианта мало полезна. Действительно, существует определенный набор вопросов, существенных при подготовке олимпиады (наличие утешительной задачи, балансировка варианта по сложности, наличие разных тем в варианте олимпиады). Но грамотному математику, интересующемуся олимпиадами (а на олимпиадах воспиталось много математиков), всему этому научиться несложно.

В учебном процессе по некоторым предметам ставится «зачет», выражающий только наличие определенных умений, при этом неважно, насколько эти умения доведены до совершенства, а по некоторым предметам — «экзамен с оценкой». За соблюдение формата варианта жюри можно присудить только «зачет» или «незачет». История ставит оценку только за научное содержание. У создателей олимпиадного движения не было стремления ни к тонкой балансировке варианта, ни к близости распределения решивших задачи с желаемым. Эти стремления появились и стали приводить к засилью искусственных задач, которые (см. выше) проще придумывать и удобнее оценивать. Такого рода «профессионализация» есть паразитическая часть олимпиадной культуры. Она позволяет рационализировать отторжение хороших задач и создает авторитет некоторым олимпиадным деятелям, ограничивает круг лиц, занимающихся олимпиадами.

Когда за олимпиаду отвечал математик приемлемого уровня, то даже при отсутствии олимпиадного опыта научное содержание олимпиад выправлялось. Что касается соблюдения организационных принципов и форматов, оно быстро приходило в норму. Так было, в частности, с Всесоюзными и Московскими городскими олимпиадами в середине 80-х годов. Роль сильного математика может быть нетривиальной. Бывает очень непросто понять, как наши сильные качества или слабости отражаются на нашем окружении. Так, только проанализировав свои разговоры с учениками, я обнаружил, что почти все время я говорил о технике атаки на тот или иной открытый вопрос, но не о красоте той или иной математической теории, и осознал необходимость дополнительных душевных усилий в этом направлении. О силе руководителя прежде всего говорит его окружение.

Об олимпиадном сообществе. Плохо, когда педагогическое сообщество отделено от научного и предоставлено самому себе. Часть возникающих нежелательных явлений связана с внутренними особенностями самого элитного (по содержанию, а не престижности) образования, которое с неизбежностью базируется на культе успеха. Более способные учащиеся получают больше внимания и ресурсов. (Противникам элитного образования следует указать на их лицемерие — ведь в жизни более успешный человек больше получает. В конце концов, «ибо кто имеет, тому дано будет и приумножится, а кто не имеет, у того отнимется и то, что имеет» (Мф. 13:12).) Применительно к математическому образованию, больше внимания получает тот, кто лучше решает задачи.

Многие замечали, что великие тренеры часто возникают из неудавшихся олимпиадников. Психологические комплексы, о которых так любят говорить, несут в себе не только плохое, но и хорошее. Зачастую, излечив человека от комплексов, мы его излечим и от талантов и будущих достижений. Не понимая этого, гуманистические психологи[3)] делают подобные суждения некомпетентными.

Тем не менее, при этом возникают психологические проблемы, которые следует учитывать и, по возможности, смягчать. В сообществе, в котором доминируют неудавшиеся ученые, возникает весьма специфическая атмосфера. Если человек уделяет большее внимание учащемуся, более успешному в решении трудных задач, в то время как у него самого есть проблемы с творческой реализацией (особенно если он был вундеркиндом), он тем самым усиливает свои комплексы, порой это становится опасным. Кроме того, зачастую психика преподавателя приближается к подростковой.

Разумеется, отнюдь не все преподаватели и педагогические деятели таковы. Есть вполне успешные люди, пришедшие в эту область (как учителя, методисты или организаторы) по зову сердца, а не потому, что наука не получилась.

[3)] Это их самоназвание.

Есть механизм, выдвигающий наиболее амбициозных деятелей. Влияние человека сильно зависит от энергии, которую он вкладывает, и, к сожалению, от уровня его агрессии. (В этом отношении человеческое сообщество не отличается от иной популяции животных.) Овладение культурными ценностями, равно как и научная работа, требует очень много энергии. Если же всего этого нет, то при том же природном энергетическом потенциале человек будет выглядеть более ярко, чем тот, у кого есть другая большая работа. Бурлит мелкая вода. Есть люди с блестящим, но бесплодным интеллектом (быстрота мыслительных операций — еще не интеллект, интеллект — это не ум, а ум — не мудрость). Однако выдвижение человека и его роль сильно зависит от этого внешнего блеска.[4] Этим объясняется парадоксальная ситуация, сложившаяся в жюри некоторых олимпиад и турниров: ответственными за варианты для относительно слабых учащихся (работа, связанная с меньшими амбициями) являются более сильные математики и методисты.[5]

Если олимпиада — единственное поле реализации человека, то он будет энергично добиваться должностей, занимаясь даже неприятной работой, да и локтями поработает. А человека, который может реализоваться в науке, гораздо легче прогнать — у него есть другое поле деятельности. Происходит отрицательный отбор. Подобные механизмы действия инстинкта агрессии изложены в книгах [3, 7].

Далеко не все педагогические деятели прошлого были сильными математиками, но они были членами научного сообщества. Есть и сейчас деятели, высказывающие глубокие методические идеи, авторы журнальных статей. Но зачастую задают тон люди с узкой специализацией — придумывание олимпиадной задачи или создание варианта олимпиады, профессионального с точки зрения формата мероприятия, но с пренебрежением к содержанию.

Научный мир (особенно в точных науках) при всех своих нынешних недостатках (см. например, книгу А. Гротендика «Урожаи и посевы») более здоровый, чем олимпиадно-педагогический. В нем есть объективные критерии.

Но даже в более здоровом научном сообществе необходима определенная формальная самозащита (ученые степени, звания и т. д.), часто с издержками. К западному опыту организации науки надо относиться критично, но не игнорировать. Следует осознать причины тех или иных решений. Там человек меряется, прежде всего, по публикациям (с учетом

[4] Внешний блеск очень полезен в преподавании. Но в преподавании сильным учащимся важнее научная квалификация.

[5] Кроме того, следует отметить, что студенту, пусть даже талантливому математику и педагогу, входящему в жюри мероприятия, иногда трудно спорить со старшим по возрасту, в особенности с тем, с кем он взаимодействовал, будучи школьником.

их количества и индекса цитируемости соответствующих журналов), а уже затем — по преподаванию. Этот способ имеет много недостатков. Зачастую очень хорошие специалисты имеют более низкий индекс цитирования, чем некоторые ученые среднего уровня. Тем не менее, нижний и средний уровень эти формальные параметры измеряют хорошо. Репутация организации зависит от публикаций сотрудников.[6)]

3. Заключение

В создавшихся условиях особенно важны традиции, заложенные создателями олимпиадного движения. В частности, они способны затруднить, а иногда и остановить вредное реформаторство. То обстоятельство, что у истоков олимпиад стояли великие люди, имеет важное значение и для привлечения общественного внимания и ресурсов (включая финансовые). Но, в то же время, это и ответственность — по меньшей мере, моральная. Например, если мы объявляем, что наша олимпиада основана таким-то ученым, или называем мероприятие его именем, то общественность это воспринимает как заявление в плане верности традиции, и мы тем самым ставим вопрос: как бы относился имярек к его проведению и стилю?[7)]

Автору близка мысль члена жюри Всероссийской олимпиады по математике А. С. Голованова о роли гласности, высказанная им в интернет-дискуссиях. Об этом в отношении уровня жюри должны заботиться организаторы турниров.[8)]

Избегая несправедливых обид, мы не считаем нужным называть приверженцев того или иного подхода. Мы не оцениваем отдельных личностей, равно как и конкретные мероприятия, но подходы и тенденции. Борьба разных стилей происходит внутри человека. Автор опирается и на свой внутренний опыт. Он не может осуждать деятеля, прервавшего свою заграничную поездку, чтобы подготовить международника, или человека, для которого олимпиадный диплом стоит больше публикации в научном

[6)] За рейтингами стоит нетривиальное управление. Сравнение успехов студентов на олимпиадах (и в научной работе) между российскими и западными университетами невыгодно Западу. Поэтому соответствующие параметры не учитываются в подсчете рейтингов. О системе PISA подробно рассказано в статье Д. Малати в журнале «Математическое образование». Понижение рейтинга отечественных журналов вызвало снижение потока хороших статей и, как следствие, цепную реакцию понижения.

[7)] В некоторых случаях вопрос надо согласовать с наследниками. На Западе к этому чувствительны.

[8)] Некоторое представление как о научных, так и о популярных статьях, можно получить на платных сайтах `http://www.zentralblatt-math.org`, `http://www.ams.org/mathscinet`, и бесплатных `http://scholar.google.com/`, `http://kvant.mirror1.mccme.ru/`, `http://www.turgor.ru/lktg/index.php`, `http://www.mathnet.ru`).

журнале. Он сам испытывал бóльшую радость, когда его задача попадала в вариант олимпиады, чем от публикации, скажем, в Journal of Algebra, а олимпиадная премия могла обрадовать больше, чем защита диссертации. К нашим чувствам надо относиться критично. Иногда попадаются математики, жестко следующие спортивному подходу, а с другой стороны — есть педагоги, преданные науке.[9)] Кроме того, сами формальные показатели нуждаются в изучении и уточнении. Да и ситуация в каждом из конкурсов должна быть изучена подробно.

Олимпиадные функционеры тоже нужны — прежде всего, в технических вопросах, а также вопросах, связанных с формированием простой части вариантов.

Разнообразие интересов членов методкомисии улучшает атмосферу и результат работы. Очень плохо, когда свет клином сошелся на мероприятиях. Очень важна диверсификация интересов деятеля. Говоря о рейтингах, кроме научных статей есть еще статьи в достаточно престижные популярные журналы, например, «Квант».

Чтобы возникли новые авторы, важен малый жанр. Для написания большой статьи надо потратить много ресурсов, да ее еще могут не принять. Пусть и школьники, и учителя из провинции, общаясь с редакторами научных журналов, получат навык и вкус к научным миниатюрам.[10)]

Нельзя жить прошлыми заслугами. Уровень популярных журналов в России пока выше, чем за рубежом, но это не может долго продолжаться по инерции. Об авторитете отечественных олимпиад также нужно заботиться, а он зависит, прежде всего, от их содержания.

И самое главное. Поднять научный уровень олимпиад и одновременно спортивные достижения МОЖНО. Это проще, чем кажется. Не так сложно научиться технике составления вариантов, да и среди действующих математиков есть достаточное число людей, прошедших олимпиадную школу. Организация постоянно действующего семинара по олимпиадным задачам в конце 80-х – начале 90-х годов на мехмате привела к появлению новых олимпиадных деятелей и достаточного запаса новых олимпиадных задач. Важно привлекать действующих математиков, причем разнообразных специальностей, чтобы были покрыты основные разделы науки. Тех,

---

[9)] Поведение действующего математика, получившего чрезмерно спортивное воспитание, зачастую противоречиво. Взгляды людей, переставших заниматься наукой и оторванных от научного сообщества, эволюционируют в сторону первого стиля. Иногда поведение математика в отношении задач своей узкой специальности отвечает второму подходу, а остальных задач — первому.

[10)] Автор еще в 90-е годы критиковал (часто с ущербом для себя) некоторых журнальных деятелей, ревниво относящихся к своему делу и подавлявших других, особенно провинциальных авторов. Многие люди, испытавшие на себе это подавление, в последующем вели себя не лучше — если не хуже. Дело не в конкретных персоналиях, а в проблемах самого педагогического сообщества.

кого можно привлечь к олимпиадному движению, довольно много, как в столицах, так и на периферии.

Обучение живой математике даже с точки зрения спортивных достижений гораздо эффективнее, чем натаскивание на мертвые задачи. Этому есть примеры. Сила приводит к успехам, но проявляется она в обличии красоты.

Данная статья призвана начать обсуждение проблем олимпиадного движения, которое, как утверждали практически все математики и методисты, прочитавшие эту статью, давно назрело. Необходимо начать работу по осмыслению накопленного опыта, и это осмысление принесет плоды. Призываем всех к дальнейшему обсуждению.

Список литературы

[1] Бончковский Р. Н. *Вторая Московская математическая олимпиада* // Успехи матем. наук, 1936. Вып. 2. С. 275–278. (См. также `www.mathnet.ru/rm8891`.)

[2] Гальперин Г. А., Толпыго А. К. *Московские математические олимпиады*. М.: Просвещение, 1986.

[3] Дольник В. Р. *Непослушное дитя биосферы. Беседы о поведении человека в компании птиц, зверей и детей.* М: МЦНМО, 2009. (См. также `http://www.ethology.ru/library/?id=321`)

[4] Константинов Н. Н. *Турнир Городов и Математическая олимпиада* // Математическое Просвещение. Сер. 3, вып. 1, 1997. С. 164–174.

[5] Константинов Н. Н., Френкин Б. Р. *Летние конференции Турнира городов: Избранные материалы (Вып.* 1). М.: МЦНМО, 2009.

[6] Леман А. А. *Сборник задач московских математических олимпиад*. М.: Просвещение, 1965.

[7] Лоренц К. *Агрессия (так называемое зло)*. М: Римис, 2009.

[8] Пойа Д. *Как решать задачу*. М.: Учпедгиз, 1961.

[9] Пойа Д. *Математика и правдоподобные рассуждения*. М.: Наука, 1975.

[10] Пойа Д. *Математическое открытие*. М.: Наука, 1976.

[11] Скопенков А. Б. *Олимпиады и математика* // Математическое просвещение. Сер. 3, вып. 10, 2006. С. 57–63.

[12] Федоров Р. М., Канель-Белов А. Я., Ковальджи А. К., Ященко И. В. *Московские математические олимпиады.* 1993–2005. М: МЦНМО, 2006.

Традиции математической олимпиады в Грузии

Н. Х. Розов

Грузинская земля в XX веке явила миру замечательную когорту блестящих математиков и механиков, снискавших славу грузинской математической школе. Ее родоначальником по праву считается А. Размадзе, в честь которого назван Математический институт в Тбилиси. В математическом мире хорошо известны фундаментальные работы Н. И. Мусхелишвили, А. К. Харадзе, Г. Н. Николадзе, В. Д. Купрадзе, И. Н. Векуа, А. В. Бицадзе Р. В. Гамкрелидзе, И. Т. Кигурадзе и многих других.

Все эти люди вошли в науку и внесли в нее свой выдающийся вклад в неразрывном единстве и при тесном сотрудничестве со всей той математической корпорацией, которую принято называть «советской математической школой». Достаточно вспомнить, например, жизненный путь Ильи Нестеровича Векуа. Воспитанник и преподаватель Тбилисского университета, он работал профессором механико-математического факультета Московского университета и сотрудником Математического института им. В. А. Стеклова в Москве, был ректором Новосибирского университета, ректором Тбилисского университета, Президентом АН Грузинской ССР, являлся академиком АН СССР и академиком АН ГССР.

Такая «математическая атмосфера» Грузии просто не могла не найти реализации в особом внимании к математическому просвещению молодежи, к поиску и поддержке молодых людей, проявляющих склонность и интерес к математике. Это вполне закономерно. Но — что самое удивительное — Грузия оказалась у истоков олимпиадного математического движения в Советском Союзе.

Первая математическая олимпиада для школьников в СССР была проведена в Тбилиси. 3 ноября 1933 года по инициативе двух энтузиастов, заслуженных учителей Грузинской ССР С. Е. Вашакмадзе и Т. Д. Петраковой, в 26 опытно-показательной столичной школе собрались юные любители математики, чтобы помериться своими математическими способностями. А чуть позже, 28 ноября, прошла и районная олимпиада школьников по математике. Интересно, что общественность и пресса широко и заинтересованно откликнулись на это начинание. Газеты «Ахалгазрда комунисти», «Тифлисский рабочий» и др. напечатали большие статьи об олимпиаде. Она нашла энергичную поддержку в постановлении

Бюро организаций коммунистического воспитания детей ЦК ЛКСМ Грузии, опубликованном в газете «Ахалгазрда комунисти» 25 декабря того же года.

Очень жаль, что нам не известны ни задачи, предлагавшиеся на этой первой олимпиаде, ни фамилии первых победителей и их судьбы. Быть может, грузинские коллеги помогут пополнить историю математических олимпиад этими бесценными свидетельствами.

Как известно, весной 1934 года городская олимпиада юных математиков прошла в Ленинграде. Именно это событие традиционно считается началом «движения школьных математических олимпиад» в СССР. Осенью 1935 года городская математическая олимпиада для школьников впервые была проведена в Москве.

В первый раз городская школьная математическая олимпиада в Тбилиси прошла 18 июня 1934 года. А ее победители были приглашены на олимпиаду в Тбилисский университет, активными организаторами которой явились видные грузинские математики А. Харадзе, Л. Гокиели, Д. Долидзе, М. Кониашвили, К. Сулаквелидзе.

Республиканская олимпиада юных математиков Грузии родилась в 1957 году — чести возглавлять ее оргкомитет удостоился чл.-корр. АН ГССР В. Челидзе.

История распорядилась так, что Грузия еще раз оказалась колыбелью важного математического олимпиадного события. В 1967 году состоялась Первая Всесоюзная олимпиада школьников по математике. И ее гостеприимной хозяйкой была именно грузинская земля. В 1981 году Грузия снова принимала Всесоюзную математическую олимпиаду школьников, председателем ее оргкомитета был чл.-корр. АН ГССР Т. Гегелия. Активную роль в ее проведении тогда сыграли известные грузинские математики И. Т. Кигурадзе, братья Т. и З. Чантурия и Л. Гоголадзе (последний, кстати, и до сих пор занимается проведением республиканских математических олимпиад в Грузии).

По мотивам задачника «Математического просвещения»

Решение задачи про вписывание пятиугольника

Р. Н. Карасёв*

Теорема. *Во всякое выпуклое тело K в $\mathbb{R}^3$ можно вписать плоский правильный пятиугольник, плоскость которого проходит через любую наперёд заданную точку $a \in \operatorname{int} K$.*

Из этой теоремы следует решение задачи 3.9 задачника «Математического просвещения» (формулировку см. в вып. 8, с. 247).

План доказательства взят из статьи В. В. Макеева [1]. Оно основано на применении некоторого обобщения теоремы Борсука – Улама [4]. Однако для понимания доказательства достаточно владеть программой первого семестра университета, не нужно знать ни формулировки этой теоремы, ни слов «степень отображения», «эквивариантные отображения». Новые объекты появляются только в конце (при доказательстве факта 2) и в тех простейших частных случаях, в которых они используются (ср. [2]).

В создании текста также принимали участие Илья Богданов и Аркадий Скопенков.

Доказательство теоремы для гладкого строго выпуклого[1] *тела K.* Будем считать, что точка $a \in K$ (через которую должна пройти точка пятиугольника по условию задачи) — это начало координат. Рассмотрим в плоскости $x0y$ правильный пятиугольник, составленный из упорядоченных против часовой стрелки единичных векторов $e_1, \ldots, e_5$, первый из которых является направляющим вектором оси $0x$.

*Работа выполнена при частичной поддержке фонда «Династия», гранта Президента РФ МК-113.2010.1, грантов РФФИ 10–01–00096 и 10–01–00139, АВЦП «Развитие научного потенциала высшей школы», проект 2.1.1/500, ФЦП «Научные и научно-педагогические кадры иннновационной России» на 2009–2013 годы

[1] Это ограничение техническое, почти всё доказательство можно понять и проверить, не зная, что такое гладкость и строгая выпуклость.

Доказательство проведём от противного — предположим что требуемых вписанных пятиугольников нет. Неформально говоря, начало доказательства — попытка вписать в K пятиугольник с центром в начале координат. Для $i = 1, 2, 3, 4, 5$ и вращения ρ трёхмерного пространства $\mathbb{R}^3$, ось которого проходит через начало координат, обозначим

$$h_i(\rho) := \max\{s : s\rho(e_i) \in K\}.$$

Иначе говоря, $h_i(\rho)$ — длина отрезка, высекаемого телом K на луче, выходящем из начала координат в направлении $\rho(e_i)$. Поскольку мы предполагаем отсутствие вписанных пятиугольников, то числа $h_1(\rho), \ldots, h_5(\rho)$ никогда не равны. Положим $\sigma(\rho) := \dfrac{1}{5}(h_1(\rho) + \cdots + h_5(\rho))$. Обозначим через $SO(3)$ множество вращений трёхмерного пространства, оси которых проходят через начало координат. Определим отображение

$$g\colon SO(3) \to \mathbb{R}^4 \quad \text{формулой } g(\rho) := (h_1(\rho) - \sigma(\rho), \ldots, h_4(\rho) - \sigma(\rho)).$$

Можно было бы добавить к этому отображению пятую координату $h_5(\rho) - \sigma(\rho)$, но она зависит от первых четырёх — ведь сумма всех пяти координат будет равна нулю.

Ясно, что отсутствие вписанного пятиугольника равносильно тому, что $0 \notin g(SO(3))$. Поэтому можно определить отображение

$$f\colon SO(3) \to S^3 \quad \text{формулой} \quad f(\rho) := \frac{g(\rho)}{|g(\rho)|}\,.$$

Напомним определение трёхмерной сферы:

$$S^3 := \{(x, y, z, t) \in \mathbb{R}^4 \;:\; x^2 + y^2 + z^2 + t^2 = 1\}.$$

Непрерывные отображения[2] $f_0, f_1\colon SO(3) \to S^3$ называются *гомотопными*, если они получаются друг из друга непрерывной деформацией, т. е. включаются в семейство непрерывных отображений $f_t\colon SO(3) \to S^3$, $t \in [0, 1]$, непрерывно зависящих от параметра t.[3]

Мы получим противоречие, если докажем следующие два факта.

1. f гомотопно отображению $f_1\colon SO(3) \to S^3$, для которого $f_1(SO(3)) \neq S^3$.

2. f не гомотопно отображению $f_1\colon SO(3) \to S^3$, для которого $f_1(SO(3)) \neq S^3$.

Доказательство факта 1. Говоря неформально, мы сдвигаем предполагаемый центр пятиугольника вдоль ρe_1 и определяем h_{ti} аналогично h_i,

[2] Читатель легко определит *расстояния* между точками пространства $SO(3)$ и между точками сферы S^3. Тогда обычное ε-δ определение непрерывности переносится на случай отображений $SO(3) \to S^3$.

[3] Непрерывная зависимость от параметра t в этом конкретном случае означает непрерывность отображения $F\colon [0, 1] \times S^3 \to S^3$, заданного формулой $F(t, x) = f_t(x)$.

но с новым центром. Формально, положим

$$h_{ti}(\rho) := \max\{s : s\rho(e_i) + th_1(\rho)\rho(e_1) \in K\}.$$

Для каждого t можно по приведённой выше схеме построить[4] отображения $f_t\colon SO(3) \to S^3$. Это семейство отображений f_t непрерывно зависит от $t \in [0,1]$.[5] Заметим, что $h_{11}(\rho) = 0$ для любого вращения ρ. Кроме того, $h_{1i}(\rho) > 0$ при некоторых $i = 2, \ldots, 5$ и ρ. Поэтому первая координата вектора $f_1(\rho)$ отрицательна для любого ρ. Значит, $f_1(SO(3)) \neq S^3$. $\square$

Доказательство факта 2. Обозначим через p вращение относительно оси z на угол $-2\pi/5$. Ясно, что $h_i(\rho\circ p) = h_{i+1}(\rho)$. Здесь и далее нумерация индексов циклическая по модулю 5. Определим отображение

$$s\colon \mathbb{R}^4 \to \mathbb{R}^4 \quad \text{формулой } s(x_1, x_2, x_3, x_4) := (x_2, x_3, x_4, -x_1 - x_2 - x_3 - x_4).$$

Ясно, что s непрерывно,

$$g(\rho \circ p) = s(g(\rho)) \text{ и } s^5 = \mathrm{id}_{\mathbb{R}^4}$$

(ибо, вспомнив про пятую координату $x_5 := -x_1 - x_2 - x_3 - x_4$, получим $s^k(x_1, x_2, x_3, x_4) = (x_{1+k}, x_{2+k}, x_{3+k}, x_{4+k}, x_k))$. Определим отображение

$$r\colon S^3 \to S^3 \text{ формулой } r(v) := \frac{s(v)}{|s(v)|}\,.$$

Ясно, что r непрерывно,

$$f(\rho \circ p) = r(f(\rho)) \quad \text{и} \quad r^5 = \mathrm{id}_{S^3}\,.$$

Хочется применить следующую теорему (обобщение теоремы Борсука – Улама).

Пусть q — простое число, $F, P, R\colon S^n \to S^n$ — непрерывные отображения, для которых $F \circ P = R \circ F$, $P^q = R^q = \mathrm{id}_{S^n}$ и для любого $x \in S^n$ имеем $P(x) \neq x$, $Q(x) \neq x$. При этих условиях F не может быть гомотопно отображению $F_1\colon S^n \to S^n$, для которого $F_1(S^n) \neq S^n$.

Попробуйте доказать эту теорему для $n = 1$! Для произвольного n доказательство более сложное, подробное изложение можно найти в книге [5]. По сути надо доказать, что степень отображения F ненулевая.

Чтобы применить эту теорему, определим отображение

$$\pi\colon S^3 \to SO(3) \text{ формулой: } \pi(v) \text{ действует на } x \in \mathbb{R}^3 \text{ как } \pi(v)(x) = vxv^{-1}.$$

Здесь S^3 отождествлена с множеством кватернионов единичной нормы, а $\mathbb{R}^3$ с множеством кватернионов вида $xi + yj + zk$.

---

[4] Естественно, в предположении отсутствия вписанного пятиугольника с центром $t\rho e_1$, лежащего в плоскости $\rho\langle e_1, \ldots, e_5\rangle$.

[5] Доказательство — упражнение для читателя. Для непрерывности при $t = 1$ необходима гладкость и строгая выпуклость тела K.

Ясно, что это отображение непрерывно и что у любого вращения из $SO(3)$ ровно два прообраза при отображении π (эти прообразы отличаются знаком). Возьмём некоторый элемент $\bar{p}$ в прообразе p относительно π. При этом $\bar{p}^5$ может оказаться либо $+1$, либо -1; в последнем случае возьмём вместо $\bar{p}$ элемент $-\bar{p}$, тогда $\bar{p}^5$ станет $+1$. Тогда

$$f(\pi(v\bar{p})) = f(\pi(v)p) = rf(\pi(v)) \quad \text{для любого } v \in S^3.$$

Поэтому можно применить сформулированную теорему для $n = 3$, $q = 5$, $F = f \circ \pi$, $R = r$ и P, определённого как умножение на $\bar{p}$ справа. Получим, что $f \circ \pi$ не гомотопно отображению $F_1 \colon S^3 \to S^3$, для которого $F_1(S^3) \neq \neq S^3$. Значит, исходное f также не гомотопно отображению $f_1 \colon SO(3) \to S^3$, для которого $f(SO(3)) \neq S^3$. □

Сведение случая произвольного выпуклого тела K к случаю гладкого строго выпуклого K. Заметим, что существует последовательность гладких и строго выпуклых тел K_n, которые приближают (скажем, в метрике Хаусдорфа[6)]) тело K. Каждое K_n имеет вписанный пятиугольник P_n, плоскость которого α_n проходит через начало координат. Если размеры P_n не стремятся к нулю, то из соображений компактности можно выбрать предельный пятиугольник P, вписанный в K. Далее предполагаем, что размеры P_n стремятся к нулю. Заметим такой геометрический факт: для правильного пятиугольника $A_1A_2A_3A_4A_5$ обозначим (индексы считаются по модулю 5)

$$B_i = A_{i-1}A_i \cap A_{i+1}A_{i+2}$$

и рассмотрим звезду $A_1B_1A_2B_2 \dots A_5B_5$. Для P_n обозначим соответствующую звезду Q_n и заметим, что с одной стороны размеры Q_n стремятся к нулю, а с другой стороны Q_n содержит в себе $K_n \cap \alpha_n$. Но размеры $K_n \cap \alpha_n$ не могут стремиться к нулю, так как можно выбрать некоторый шар радиуса $\varepsilon > 0$ с центром в начале координат, содержащийся во всех K_n. □

Замечание. Аналогично (см. статью В. В. Макеева [1]) доказывается утверждение: пусть p нечётное простое, а $d = \dfrac{p+1}{2}$, тогда во всякое выпуклое тело K в $\mathbb{R}^d$ можно вписать двумерный правильный p-угольник, плоскость которого проходит через любую наперёд заданную точку $a \in \operatorname{int} K$. Правда, в этом утверждении вместо теоремы Борсука – Улама нужно применять более сильное утверждение, в котором пространство прообраза (в доказательстве оно равно многообразию Штифеля V_d^2 ортонормированных 2-реперов в $\mathbb{R}^d$) не обязано быть сферой. Однако можно вычислить некоторый инвариант (когомологический индекс) этого пространства, с

[6)] Расстояние Хаусдорфа между непустыми компактами определяется как

$$\operatorname{dist}(X, Y) = \max\{\max_{x \in X} \operatorname{dist}(x, Y), \max_{y \in Y} \operatorname{dist}(X, y)\}.$$

точки зрения которого оно «похоже» на сферу соответствующей размерности.

Обобщённая теорема Борсука – Улама. Очень подробная информация по теоремам типа Борсука – Улама и их применениям в геометрии и комбинаторике содержится в книге И. Матушека [5], которая, в принципе, не требует предварительных знаний от читателя. Приведём для полноты формулировки этих теорем. Определения понятий «полиэдр» и «n-связность» читатель может найти в стандартном учебнике по топологии, например [3].

Определения.

1) Полиэдр X, на котором действует конечная группа G непрерывными преобразованиями, называется *G-пространством.* Если для любых $g \in G \setminus \{e\}$ и $x \in X$ имеем $gx \neq x$, то действие называется *свободным.*

2) Непрерывное отображение $f : X \to Y$ между G-пространствами называется *G-эквивариантным* (или G-отображением), если оно перестановочно с действием G, то есть для любых $g \in G$ и $x \in X$ имеем $f(gx) = gf(x))$.

Теорема. *Пусть G — конечная группа, состоящая более чем из одного элемента, а $f\colon X \to Y$ — G-отображение из $(n-1)$-связного в n-мерный полиэдр со свободным действием G. Тогда f не может быть гомотопным постоянному отображению, т. е. не может быть непрерывно продеформировано в отображение, образ которого состоит из одной точки.*

В качестве упражнения читателю предлагается доказать, что если для непрерывного отображения $f\colon X \to S^n$ образ $f(X)$ не равен S^n, то оно гомотопно постоянному.

Существует и другая формулировка теоремы Борсука – Улама. Она эквивалентна первой, но доказательство эквивалентности не тривиально.

Теорема. *Пусть G — конечная группа, состоящая более чем из одного элемента. Тогда не существует G-отображения из n-связного в n-мерный полиэдр со свободным действием G.*

Список литературы

[1] Макеев В. В. *Вписанные и описанные многогранники выпуклого тела* // Матем. заметки. Т. 55, вып. 4, 1994. С. 128–130.
Англ. пер. Math. Notes. Vol. 55, no 4, 1994. P. 423–425.

[2] Скопенков А. *Философски-методическое отступление* // Математика в задачах. Сборник материалов выездных школ команды Москвы

на Всероссийскую математическую олимпиаду. Под ред. А. Заславского, Д. Пермякова, А. Скопенкова, М. Скопенкова, А. Шаповалова. М.: МЦНМО, 2009. С. 11–14.
Эл. версия `http://www.mccme.ru/circles/oim/mvz.pdf`

[3] Фоменко А. Т., Фукс Д. Б. *Курс гомотопической топологии*. М.: Наука, 1989.

[4] Borsuk K. *Drei Sätze über die n-dimensionale euklidische Sphäre* // Fund. Math. B. 20, 1933. S. 177–190.

[5] Matoušek J. *Using the Borsuk – Ulam theorem*. Berlin – Heidelberg: Springer Verlag, 2003.

Роман Карасёв, каф. высшей математики, Московский физико-технический институт, Институтский пер. 9, Долгопрудный, РФ 141700
Email: `r_n_karasev@mail.ru`

Задача о фишках и потоки на кубической решетке

М. Л. Матдинов

Введение. Пусть в некоторых клетках плоской решетки стоят фишки (не более одной в клетке), а некоторые клетки — пустые. Требуется сдвинуть все фишки на ограниченное расстояние так, чтобы все клетки были заполнены. Когда это можно сделать?

Например, пусть все клетки, кроме одной, заполнены. Тогда фишка, стоящая справа от этой клетки, сдвигается на шаг влево, фишка, стоящая справа от нее, тоже сдвигается на шаг влево и т. д. В результате весь ряд фишек справа от пустой клетки сдвигается на одну клетку влево и клетка заполняется. С другой стороны, если половина плоскости пустая, то ее заполнить невозможно. Невозможно осуществить требуемое и если имеются сколь угодно большие пустые квадраты.

Не очень сложно понять, что следующее условие необходимо:

Существует такая константа $C > 0$, что количество пустых клеток в каждом квадрате с произвольной стороной a не превосходит $C \cdot a$. $(*)$

Упражнение. Покажите, что данное условие необходимо.

На самом деле это условие является и достаточным. В [1, с. 217], была предложена следующая *задача* 4.11:

Задача (А. Я. Белов). На некоторых клетках бесконечной доски стоят фишки (не более одной на каждой клетке), некоторые клетки — пустые. Назовем расстановку *почти полной*, если найдется такое число C, что можно так сдвинуть каждую фишку на расстояние, не превышающее C (иногда нулевое), чтобы пустых клеток не осталось. Назовем расстановку *не слишком пустой*, если найдется такое число D, что количество пустых клеток в любом квадрате не превосходит DP, где P — периметр квадрата. Докажите, что почти полные расстановки — это в точности не слишком пустые.

Под ред. Б. Р. Френкина.

В 1998 г. эта задача предлагалась при отборе команды Москвы на Всероссийскую олимпиаду. Участник олимпиады Д. Ершов предложил идею, связанную с использованием теоремы Менгера. Решение задачи, основанное на этой идее см. [2, с. 249–252].

Пространственное обобщение этой задачи до сих пор оставалось нерешенным.

Отметим, что родственная задача (автор неизвестен) была предложена на V Международном математическом Турнире городов осенью 1983 г. (см. [3, с. 31, задача 10 и с. 247–248, решение задачи 10]).

Условие $(*)$ для n-мерного случая формулируется так:

Существует такая константа $C > 0$, что в произвольном кубе со стороной l содержится не более $C \cdot l^{n-1}$ пустых клеток. $(*_n)$

Упражнение. Докажите, что данное условие необходимо.

Как будет показано ниже, это условие в n-мерном случае также является достаточным.[1)]

Исходная задача о фишках связана со свойствами фазовых потоков на кубических решетках. Задача облегчается, если условие $(*_n)$ заменить на более сильное:

Существует такая константа $C > 0$, что в произвольном выпуклом теле с площадью поверхности S содержится не более $C \cdot S$ пустых клеток.

Доказательство соответствующего утверждения содержится в работе [8]. Его непрерывные аналоги исследовались в работах [10, 11].

Интересен аналог нашего результата для *билипшицевых* отображений дискретных множеств. Пусть даны два дискретных множества D_1 и D_2 на плоскости, для которых существуют константы $C_1 > 0$ и $C_2 > 0$ такие, что в каждом круге площади $S > 1$ содержится не менее чем C_1S и не более C_2S точек каждого из множеств D_i. Если же $S \leqslant 1$, то не более одной. Возникает естественный

Вопрос: Существуют ли константы $M_1 > 0$ и $M_2 > 0$ и биекция $F\colon D_1 \to D_2$ такие, что $M_1|x-y| < |F(x)-F(y)| < M_2|x-y|$?

Этому вопросу и его обобщениям был посвящен проект М. Вялого на IV конференции Международного математического Турнира городов в 1992 г. [9].

В дальнейшем на этот вопрос был получен отрицательный ответ [5]. Тем не менее в ряде важных случаев, относящихся к арифметическим

[1)] Как нам стало известно, аналогичный результат независимо получен А. Магазиновым.

объектам в теории групп, в частности решеткам, ответ всё же положительный [5, 7].

Благодарности. Автор благодарен А. Белову, В. Шаричу, С. Горбаню, М. Содину за полезные обсуждения, М. Громову за указание связи с проблемой билипшицевых отображений, А. Райгородскому и всем участникам его семинара за внимание к работе, а также С. Дориченко и участникам его кружка за проявленный интерес.

Теорема 1. *Пусть некоторые клетки N-мерного клетчатого пространства $\mathbb{Z}^N$ заполнены фишками так, что в каждой клетке — не более одной фишки.*

Предположим, что существует такая константа $C \in \mathbb{R}$, что при любом натуральном l в каждом N-мерном гиперкубе со стороной l содержится не более Cl^{N-1} свободных клеток. Тогда существует такое натуральное число k, что все фишки могут одновременно сдвинуться, каждая на расстояние, не превосходящее k, так, чтобы свободных клеток не осталось (под *расстоянием* между клетками подразумевается максимум из расстояний между их проекциями на оси координат).

Доказательство теоремы 1. Пусть Ω — множество всех клетчатых кубов в $\mathbb{Z}^N$. Далее, пусть $m(L)$ — длина стороны произвольного куба $L \in \Omega$, и пусть L содержит $h(L)$ свободных клеток.

Для каждого куба $L \in \Omega$ и числа k построим следующий граф $G(L, k)$. Его вершины — клетки куба L и еще две особые вершины a и b («источник» и «сток»). Клетки куба L (т. е. остальные вершины $G(L, k)$) являются *смежными* тогда и только тогда, когда расстояние между ними не превосходит k. Кроме того, с вершиной a смежны все свободные клетки в L и только они, а с вершиной b смежны все клетки в L, из которых можно сделать шаг длины не более k, выводящий за пределы L.

Нам потребуется следующая

Лемма 1. *При условии $(*_n)$ существует такое натуральное k, что для любого куба $L \in \Omega$ в графе $G(L, k)$ между вершинами a и b существует $h(L)$ непересекающихся путей.*

Покажем, как из леммы 1 *вытекает доказательство основной теоремы.*

Утверждение этой леммы означает следующее: *существует такое k, что при любом натуральном m существуют непересекающиеся пути с длиной шага, не превосходящей k, начинающиеся во всех свободных клетках куба с центром в начале координат и стороной $2m$ и выходящие за его пределы.*

Пусть $a_{m,i,1}$ — первая клетка пути, ведущего из a_i (она определена только если a_i лежит внутри соответствующего куба). Заметим, что величина $a_{m,i,1}$ может принимать лишь конечное множество значений при данном i, поскольку эта клетка лежит на расстоянии не более k от a_i. Поэтому мы можем выбрать последовательность значений m так, чтобы клетка $a_{m,i,1}$ была одной и той же. Затем выберем подпоследовательность так, чтобы следующая клетка пути была одной и той же, и так далее. Аналогичные пути построим для всех a_i. Пусть $\{a_i = a_{i,0}, a_{i,1}, a_{i,2}, \dots\}$ — путь, начинающийся в a_i. Свободна в нем клетка a_i и только она. Сдвинем фишки из $a_{i,1}$ в a_i, из $a_{i,2}$ в $a_{i,1}$ и т. д. Аналогично передвинем фишки на остальных путях. Прочие фишки оставим на месте. Тогда все фишки сдвинутся на ограниченное расстояние, все свободные клетки будут заполнены, а новых свободных клеток не появится, что и требуется в теореме.

Итак, остается доказать лемму 1. Вначале приведем формулировку теоремы Менгера (см., например, [4, с. 64]):

Теорема Менгера. *Между двумя несмежными вершинами графа x и y существует не менее t непересекающихся путей тогда и только тогда, когда при выкидывании любых $t-1$ вершин, отличных от x и y, эти две вершины остаются в одной компоненте связности.*

Доказательство леммы 1. В силу теоремы Менгера нам достаточно установить следующий факт:

При условии $(*_n)$ существует такое k, что в любом кубе $L \in \Omega$ после выкидывания $h(L)-1$ клеток можно из какой-то оставшейся свободной клетки выбраться за пределы L, идя по оставшимся клеткам шагами длины, не превосходящей k.

Пусть $L \in \Omega$, A — множество свободных клеток в L, B — множество выкинутых клеток и из оставшихся свободных клеток нельзя выйти из L указанным образом. Покажем, что $|A| < |B|$ при всех $k > k_0$, где константа k_0 не зависит от L. Тем самым требуемый факт будет доказан, поскольку $|A| = h(L)$.

Можно считать, что $A \backslash B$ не пусто. Выберем достаточно малое положительное ε. Пусть X — множество клеток, до которых можно добраться из $A \backslash B$. Заключим X в куб со стороной 2^t, объем которого не менее $\dfrac{|X|}{\varepsilon}$. Определим множество кубов M с помощью следующей процедуры, которую мы вначале применим к нашему кубу со стороной 2^t:

Если пересечение X с кубом составляет хотя бы долю ε от объема куба, причисляем куб к M. В противном случае, если пересечение не пусто, то делим куб на 2^N равных кубов и применяем процедуру к каждому из них. Если пересечение пусто, не делаем ничего.

Таким образом, X будет покрыто объединением нескольких непересекающихся кубов, причем для любого куба $U \in M$ имеем

$$\varepsilon \leqslant \frac{|U \cup X|}{|U|} < 2^N \varepsilon. \qquad (**)$$

(Правое неравенство выполнено, так как в противном случае мы не делили бы на 2^N частей куб, из которого получился куб U.)

Пусть T — объединение всех кубов из M. Достаточно показать, что $|T \cap A| < |T \cap B|$ и $|A \backslash T| \leqslant |B \backslash T|$. Но второе неравенство следует из того, что если $\alpha \in A$ и $\alpha \notin T$, то $\alpha \notin X$ и потому $\alpha \in B$. Первое неравенство будет доказано, если мы установим, что $|U \cap A| < |U \cap B|$ для любого $U \in M$. В силу условия $(*_n)$ $U \cap A$ содержит не более $Cm(U)^{N-1}$ клеток. Поэтому достаточно показать, что $|U \cap B| > Cm(U)^{N-1}$.

Пусть Y — произвольное подмножество в кубе $L \in \Omega$, а $B_k(Y, L)$ — множество всех клеток из L, удаленных не более чем на k хотя бы от одной клетки из Y.

Лемма 2. *При условии $(*_n)$ для любого положительного $C \in \mathbb{R}$ и любых $\alpha, \beta \in (0,1)$ существует такое натуральное k, что для любого достаточно большого куба $L \in \Omega$ и подмножества его клеток $Y \subset L$ выполняется следующее утверждение: если $\alpha|L| < |Y| < \beta|L|$, то $|B_k(Y, L)| > Cm(L)^{N-1}$.*

С помощью леммы 2 можно завершить доказательство леммы 1. Заметим, что $B_k(X \cap U, U)$ содержится в $B \cap U$. Положим $\alpha = \varepsilon$, $\beta = 2^N \varepsilon$. Из леммы 2 получаем, что *при некотором k в любом $U \in M$ достаточно большого размера, удовлетворяющем условию $(**)$, будет более чем $Cm(U)^{N-1}$ клеток из $B \cap U$, где C — константа из условия основной теоремы.*

При достаточно больших k эта оценка верна и для маленьких $|U|$. Действительно, возьмем $\varepsilon < 2^{-N-1}$, тогда $|U \cap X|$ составляет менее половины от $|U|$. При большом k остальная часть U содержится в $B_k(X \cap U, U)$, а тогда и в $B \cap U$, откуда $|B \cap U| > \frac{1}{2}|U| = \frac{1}{2}m(U)^N$. Осталось показать, что $\frac{1}{2}m(U)^N > Cm(U)^{N-1}$. Для этого достаточно, чтобы выполнялось условие $m(U) > 2C$. Но так как $1 \leqslant |U \cap X| < 2^N \varepsilon |U|$, то нужно лишь выбрать ε достаточно малым.

Доказательство леммы 2. Зафиксируем некоторое δ, $0 < \delta < 1$. Выберем некоторое k. Поскольку мы хотим доказать наше утверждение лишь для достаточно больших кубов, можно считать, что $|B_k(Y, L)| < \delta|L|$. Пусть $Y' = L \backslash Y \backslash B_k(Y, L)$. Заметим, что $|Y'| > (1 - \beta - \delta)|L|$.

Пусть $x, y \in L$. Назовем *линейным путем* между x и y путь наименьшей длины между ними с шагом 1, причем вначале движение происходит параллельно первой координатной оси, затем параллельно второй и т. д. Ясно, что между двумя ячейками из L существует единственный линейный путь.

Введем следующие обозначения: $w(x) = 1$ при $x \in B_k(Y, L)$; $w(x) = 0$ в противном случае; $W(x, y)$ — число точек из $B_k(Y, L)$ на линейном пути из x в y; $N(x)$ — число линейных путей, проходящих через x.

Если линейный путь проходит через x, то x принадлежит его компоненте, параллельной i-й координатной оси, где $1 \leqslant i \leqslant N$. Начало этой компоненты — одна из m клеток, откуда можно попасть в x, двигаясь параллельно i-й оси. Конец i-й компоненты — одна из тех же m клеток. При каждом из вариантов начала i-й компоненты имеется m вариантов начала предыдущей компоненты (при $i > 1$) и т. д., а при каждом из вариантов конца i-й компоненты имеется m вариантов конца следующей компоненты (при $i < N$) и т. д. Поэтому существует не более m^{N+1} линейных путей, i-я компонента которых проходит через x, откуда $N(x) \leqslant Nm^{N+1}$.

Заметим, что если $x \in Y$, $y \in Y'$, то $W(x, y) = W(y, x) \geqslant k$, поэтому выполняется неравенство:

$$\sum_{x,y \in L} W(x, y) \geqslant 2k|Y||Y'| \geqslant 2k\alpha(1 - \beta - \delta)|L|^2 = 2k\alpha(1 - \beta - \delta)m(L)^{2N}.$$

Заметим также, что выполняется очевидное равенство

$$\sum_{x,y \in L} W(x, y) = \sum_{x \in L} N(x)w(x)$$

Таким образом, мы имеем следующую цепочку соотношений:

$$2k\alpha(1 - \beta - \delta)m(L)^{2N} \leqslant \sum_{x,y \in L} W(x, y) = \sum_{x \in L} N(x)w(x) \leqslant$$

$$\leqslant Nm(L)^{N+1} \sum_{x \in L} w(x) = Nm(L)^{N+1}|B_k(Y, L)|$$

Следовательно,

$$|B_k(Y, L)| \geqslant \frac{2k}{N}\alpha(1 - \beta - \delta)m(L)^{N-1}.$$

Остается лишь выбрать k достаточно большим, чтобы правая часть была больше $Cm(L)^{N-1}$. Таким образом, лемма 2, а вместе с ней лемма 1 и основная теорема 1, доказаны.

Список литературы

[1] Математическое Просвещение. Сер. 3, вып. 4, 2000.

[2] Математическое Просвещение. Сер. 3, вып. 8, 2004.

[3] Толпыго А. К. *Тысяча задач Международного математического Турнира городов*. М.: МЦНМО, 2010.

[4] Харари Ф. *Теория графов*. М.: УРСС, 2003.

[5] Burago D., Kleiner B. *Separated nets in Euclidean space and Jacobians of bi-Lipschitz maps* // Geom. Funct. Anal. Vol. 8, 1998. P. 273–282.

[6] Burago D., Kleiner B. *Rectifying separated nets* // Geom. Funct. Anal. Vol. 12, 2002. P. 80–92.

[7] Gromov M. *Asymptotic invariants of infinite groups* // Geometric Group Theory. Vol. 2 (Sussex, 1991). Cambridge: Cambridge Univ. Press, 1993. P. 1–295.

[8] Gromov M. *Metric structures for Riemannian and Non-Riemannian spaces*. Boston: Birkhäuser, 1999.

[9] IV *Летняя конференция Международного математического Турнира Городов. Задачи и решения.* Москва, 1993. С. 11–13, 33–46.

[10] Sodin M., Tsirelson B. *Uniformly spread measures and vector fields* // Записки научн. сем. ПОМИ. Т. 366, 2009. С. 116–127.

[11] Sodin M., Tsirelson B. *Random complex zeroes, II. Perturbed lattice* // Israel Journal of Mathematics. Vol. 152, 2006. P. 105–124.

М. Л. Матдинов, СУНЦ МГУ
Email: `marmarselsel@mail.ru`

Несколько прямых, проходящих через точку Фейербаха

Ф. Ивлев

В этой заметке содержится, в частности, решение задачи 14.8 из задачника «Математического просвещения»[1)]

Всем известно, что в любом треугольнике существует вписанная окружность. Также хорошо известен факт, что во всяком треугольнике середины его сторон и основания высот треугольника лежат на одной окружности, называемой *окружностью девяти точек* или *окружностью Эйлера* треугольника. Если рассматривать разносторонний треугольник (а в нашем исследовании случай равнобедренного и равностороннего треугольника является вырожденным), то две вышеупомянутые замечательные окружности касаются. Этот факт в 1822 году доказал немецкий математик К. Фейербах, именем которого часто называют как окружность девяти точек, так и саму точку касания окружностей.

В данной работе рассмотрены несколько новых замечательных прямых, проходящих через точку Фейербаха. Сам факт, что эти прямые проходят через точку Фейербаха был замечен Львом Емельяновым и Татьяной Емельяновой в 2001 году. На Летней конференции Турнира городов они включили его как недоказанный пункт в задачу «Семейство Фейербаха»[2)]. Первое решение, видимо, было получено Д. Гринбергом (Darij Grinberg)[3)].

Приведенное в данной заметке доказательство отличается от найденного Гринбергом. Оно затрагивает несколько новых свойств точек, являющихся точками пересечения несоответственных сторон серединного треугольника и треугольника из точек касания вписанной окружности со

[1)] Условие задачи в 14 выпуске «Математического просвещения» содержало опечатки. Правильное условие см. в этом номере, с. 235.

[2)] Ссылка на страницу конференции: `http://www.turgor.ru/lktg/2001/index.php`. Ссылка на условия задачи (см. п. 22) `http://www.turgor.ru/lktg/2001/feerbach/ru/probl_r.zip`.

[3)] Ссылка на его работы `http://www.cip.ifi.lmu.de/~grinberg/geometry2.html`, на статью, о которой идет речь, — `http://www.cip.ifi.lmu.de/~grinberg/GenFeuerPDF.zip`.

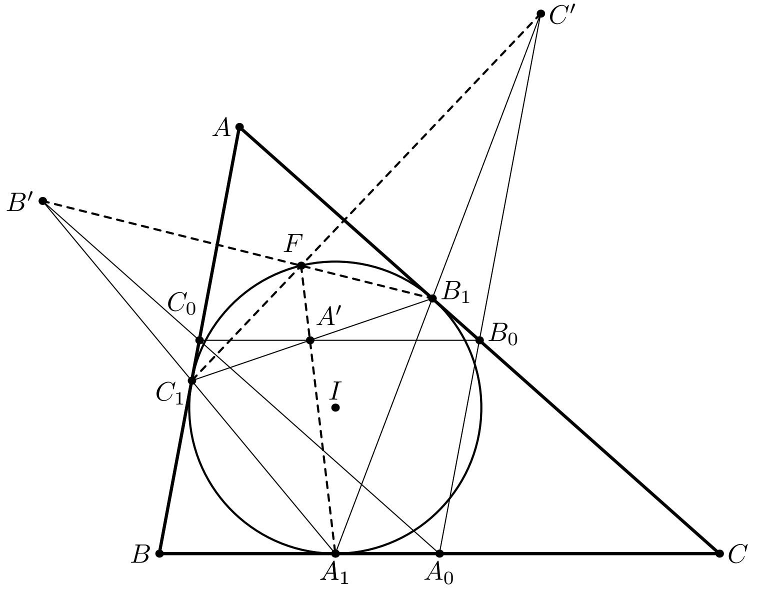


Рис. 1. Основная теорема: прямые A_1A', B_1B', C_1C' проходят через точку Фейербаха

сторонами. Приводятся новые свойства центров вневписанных окружностей серединного треугольника. Также указываются еще три замечательные прямые, проходящие через точку Фейербаха (помимо, найденных Емельяновыми) и даже целый класс таких прямых. (См. теорему 2 и лемму 5 ниже.)

В конце статьи приведено, пожалуй, самое короткое и простое доказательство поставленной задачи, найденное Александром Скутиным[4)]. В решении Скутина приводятся еще несколько интересных свойств рассматриваемой конструкции.

Будем называть треугольник, образованный точками касания вписанной окружности со сторонами, *треугольником Жергонна*.

Обозначим вершины исходного треугольника A, B, C, вершины треугольника Жергонна — A_1, B_1, C_1, а середины сторон треугольника ABC через A_0, B_0 и C_0 (см. рис. 1). Рассмотрим точки пересечения соответственных сторон серединного треугольника $A_0B_0C_0$ и треугольника Жергонна $A_1B_1C_1$. Обозначим их через A', B' и C' соответственно (A' соответствует точке пересечения прямых B_1C_1 и B_0C_0). Наша цель — изучить свойства прямых A_1A', B_1B', C_1C' и доказать, что они проходят через точку Фейербаха. Назовем это утверждение Основной теоремой.

[4)]Решение приводится с согласия А. Скутина.

Наметим план доказательства. Если читатель знает некоторые леммы, то он может сразу перейти к следующим шагам решения. Также сюда можно заглядывать, чтобы вспомнить обозначения.

Обозначим центр вписанной окружности $\triangle ABC$ через I.

Обозначим точку пересечения прямых CI и C_1A_1 через C_A, CI и C_1B_1 — через C_B. Аналогично определяются точки A_B, A_C, B_A, B_C.

1. Лемма 1. Точка C_A лежит на C_0B_0. Аналогично для остальных точек.
2. Лемма 2. Точки C_A, C_B, A_B, A_C лежат на одной окружности, ортогональной вписанной окружности с центром в точке O_A.
3. Лемма 3. O_A и два центра аналогичных окружностей (обозначим их через O_B и O_C) лежат на прямых $B'C'$, $A'C'$ и $A'B'$ соответственно.
4. Лемма 4. Обозначим за M_A середину стороны O_BO_C треугольника $O_AO_BO_C$.

   Тогда M_A лежит на прямой A_1A'.
5. Доказательство основной теоремы.
6. Лемма 5. Даны точки A_1, B_1, C_1 на сторонах треугольника. A_0, B_0, C_0 — середины сторон $\triangle ABC$. Точка пересечения A_1B_1 с A_0B_0 — A'_1. Аналогично определяются точки B'_1 и C'_1. Точка A_2 симметрична точке A_1 относительно A_0. Аналогично определяются точки B_2 и C_2. Точка пересечения A_2B_2 с A_0B_0 — A'_2. Аналогично определяются точки B'_2 и C'_2.

   Тогда прямые $A_1A'_1$, $B_1B'_1$, $C_1C'_1$, $A_2A'_2$, $B_2B'_2$, $C_2C'_2$ пересекаются в одной точке.
7. Теорема 2. Обозначим точки касания вневписанных окружностей со сторонами через A_2, B_2 и C_2. Пусть A'' — точка пересечения прямых B_0C_0 и B_2C_2. Аналогично определим точки B'' и C''. Тогда прямые A_2A'', B_2B'' и C_2C'' проходят через точку Фейербаха.

Перейдем непосредственно к доказательству. При доказательстве лемм используются некоторые факты из статьи [1]. А именно: вершины A, B и C лежат на сторонах треугольника $A'B'C'$, и каждая из вершин треугольника $A'B'C'$ является полюсом его противоположной стороны относительно вписанной окружности.

В дальнейшем удобно будет рассматривать вместо обычных углов ориентированные углы. Будем обозначать через $(l,\, r)$ ориентированный угол между прямыми l и r — угол, на который нужно повернуть прямую l, чтобы она стала параллельна прямой l. Подробнее об ориентированных углах можно прочитать в [2, глава 2].

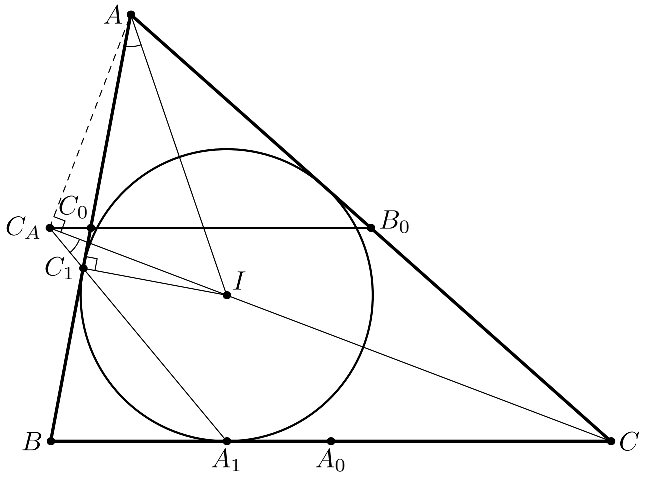


Рис. 2.

Лемма 1. Точка C_A лежит на средней линии $\triangle ABC$, параллельной BC. Аналогично для остальных точек.

Доказательство. Рассмотрим только случай точки C_A, для остальных точек рассуждения аналогичны.

$$\begin{aligned}
&\angle B = 180^\circ - \angle A - \angle C \quad \Rightarrow\\
&\qquad \angle BA_1C_1 = \frac{180^\circ - \angle B}{2} = \frac{\angle A + \angle C}{2} \quad \Rightarrow\\
&\qquad \angle(A_1B,\, A_1C_1) = \angle(A_1C,\, CI) + \angle(AI,\, AC_1) \quad \Rightarrow\\
&\qquad \angle(AI,\, AC_1) = \angle(A_1B,\, A_1C_1) - \angle(A_1C,\, CI) =\\
&\qquad = \angle(A_1C,\, C_1C_A) - \angle(A_1C,\, CC_A) =\\
&\qquad = \angle(IC_A,\, A_1C) + \angle(A_1C,\, C_1C_A) = \angle(IC_A,\, C_AC_1).
\end{aligned}$$

Следовательно, точки A, C_1, C_A и I лежат на одной окружности. Тогда $\angle AC_AI = \angle AC_1I = 90^\circ$ (см. рис. 2). Следовательно, B_0 — центр описанной окружности $\triangle AC_AC$, как середина гипотенузы прямоугольного треугольника. Получаем, что $\angle C_AB_0A = 2\angle C_ACA = \angle C$, т. е. $C_AB_0 \parallel BC$, что и означает, что C_AB_0 — средняя линия. □

Далее для доказательства мы будем пользоваться языком проективных и полярных преобразований. Подробнее о них можно прочитать в [1] и [2].

Лемма 2. Точки A_B, A_C, B_A, B_C лежат на одной окружности.

Доказательство. Мы получили, что $C_0C_A \parallel BC$, а следовательно, все стороны $\triangle C_1C_0C_A$ параллельны соответственным сторонам $\triangle C_1BA_1$.

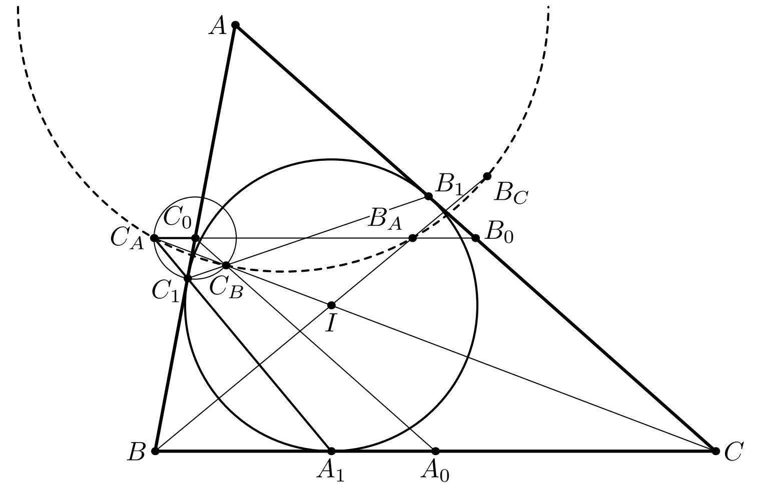


Рис. 3.

То есть он тоже равнобедренный, и потому $C_0C_1 = C_0C_A$. По аналогичным причинам $C_0C_1 = C_0C_B$. Следовательно, C_0 — центр описанной окружности треугольника $C_AC_1C_B$ (см. рис. 3). Заметим, что центр этой окружности лежит на касательной к вписанной окружности в их общей точке C_1. Значит сама окружность ортогональна вписанной окружности. Следовательно, точки C_A и C_B, лежащие с I на одной прямой, инверсны относительно вписанной окружности. Аналогично инверсны относительно вписанной окружности пары точек A_B с A_C и B_A с B_C. Следовательно, точки A_B, A_C, B_A, B_C лежат на одной окружности, ортогональной вписанной окружности. Назовем эту окружность ω_C. Аналогично определим окружности ω_A и ω_B. Центры этих окружностей обозначим за O_A, O_B и O_C соответственно. □

Так как окружности ω_A, ω_B и описанная окружность $\triangle C_AC_BC_1$ проходят через точки C_A и C_B, то их центры лежат на одной прямой. То есть C_0 лежит на прямой O_AO_B. Аналогично $B_0 \in O_AO_C$ и $A_0 \in O_BO_C$. Причем O_AO_B перпендикулярна C_AC_B, а следовательно перпендикулярна и биссектрисе угла C. Заметим, что биссектриса угла C_0 треугольника $A_0B_0C_0$ параллельна биссектрисе угла C. А значит, прямая O_AO_B перпендикулярна и биссектрисе угла C_0. Вспомнив, что эта прямая проходит через вершину C_0, получаем, что O_AO_B — внешняя биссектриса треугольника $A_0B_0C_0$. Заметим, что аналогичное верно и для прямых O_AO_C и O_BO_C, а следовательно, O_A, O_B и O_C — центры вневписанных окружностей треугольника $A_0B_0C_0$. Но тогда точки A_0, B_0 и C_0 являются основаниями высот треугольника $O_AO_BO_C$. То есть окружность Эйлера треугольника

ABC так же является окружностью Эйлера треугольника $O_AO_BO_C$, так как они обе проходят через точки A_0, B_0 и C_0.

Лемма 3. O_A лежит на $B'C'$.

Доказательство. Обозначим вписанную окружность треугольника ABC через ω. Напомним, что окружности ω и ω_A ортогональны. Обозначим за K и L точки их пересечения. Тогда O_AK и O_AL — касательные к ω, так как ω ортогональна ω_A. Следовательно, поляра O_A относительно ω — прямая KL. Также KL является полярой I относительно ω_A. Как мы знаем, A' — полюс $B'C'$ относительно ω. Следовательно, для того чтобы показать, что O_A лежит на $B'C'$, достаточно показать, что поляра O_A проходит через A'. То есть, что A' лежит на KL. Для этого достаточно заметить, что I — точка пересечения C_AC_B и B_AB_C, а A' — точка пересечения C_AB_A и C_BB_C. Следовательно, A' лежит на поляре I относительно ω_A, то есть на KL, ч. т. д. □

Обозначим середины сторон $\triangle O_AO_BO_C$ через M_A, M_B и M_C.

Лемма 4. Точки M_A, A_1, A' лежат на одной прямой.

Доказательство. Пусть Q — точка пересечения прямых B_1C_1 и BC. Тогда, так как прямые AA_1, BB_1, CC_1 пересекаются в одной точке, то по теореме о полном четырехвершиннике $(B, C, A_1, Q) = -1$. Спроектируем это двойное отношение из точки A' на прямую O_BO_C. При этом Q перейдет в бесконечно удаленную точку, точки B и C перейдут в O_B и O_C, а следовательно, A_1 перейдет в середину отрезка O_BO_C, то есть в точку M_A (см. рис. 4). Следовательно, точки M_A, A_1, A' лежат на одной прямой. □

Основная теорема. Прямые A_1A', B_1B' и C_1C' проходят через точку Фейербаха.

Доказательство. Рассмотрим треугольники $M_AM_BM_C$ и $A_1B_1C_1$. Соответственные стороны у них перпендикулярны соответствующим биссектрисам треугольника $\triangle ABC$, а значит параллельны между собой. Следовательно, эти треугольники гомотетичны. Так как описанная окружность треугольника $M_AM_BM_C$ является окружностью Эйлера треугольника ABC, а описанная окружность треугольника $A_1B_1C_1$ — вписанной окружностью треугольника ABC, то центром гомотетии будет точка Фейербаха треугольника ABC. Следовательно, прямые A_1A', B_1B', C_1C', проходящие через соответствующие вершины треугольника, пересекаются в этом центре гомотетии, то есть в точке Фейербаха, ч. т. д. □

Именно это факт заметили Емельяновы, и доказал Гринберг. Наши рассуждения можно продолжить и найти еще одну тройку замечательных прямых, проходящих через точку Фейербаха, и даже целый класс

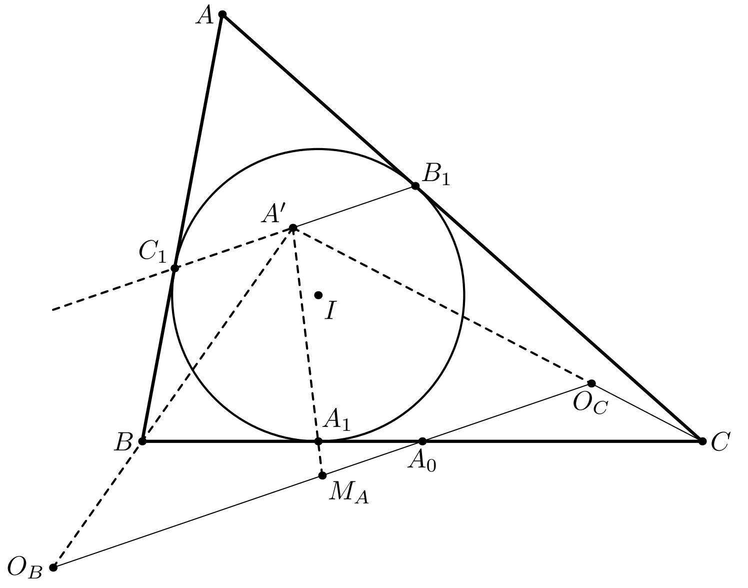


Рис. 4.

специальных троек прямых, проходящих через точку Фейербаха. Сейчас мы покажем, что в условии основной теоремы можно было заменить заменить треугольник Жергонна на треугольник из точек касания вневписанных окружностей со сторонами, часто называемый *треугольником Нагеля*. Обозначим его вершины через A_2, B_2 и C_2. Для них аналогичным образом определим точки $A''B''C''$. Тогда утверждение теоремы 2 состоит в том, что прямые $A''A_2$, $B''B_2$, $C''C_2$ также пересекутся в точке Фейербаха.

Докажем более общий факт, а именно:

Лемма 5. Даны точки A_1, B_1, C_1 на сторонах треугольника. A_0, B_0, C_0 — середины сторон треугольника ABC. Точка пересечения A_1B_1 с A_0B_0 — A'_1. Аналогично определяются B'_1 и C'_1. Точка A_2 — образ точки A_1 при гомотетии в A_0 с коэффициентом k. Аналогично определяются точки B_2 и C_2 (коэффициент k для всех трех точек одинаковый). Для них определим точки A'_2, B'_2, C'_2 аналогичным образом. Тогда прямые $A_1A'_1$, $B_1B'_1$, $C_1C'_1$, $A_2A'_2$, $B_2B'_2$, $C_2C'_2$ пересекаются в одной точке.

Доказательство. Сначала покажем, что прямые $A_1A'_1$, $B_1B'_1$, $C_1C'_1$ пересекаются в одной точке. Так как точки A_2, B_2, C_2 тоже лежат на сторонах треугольника ABC, то прямые $A_2A'_2$, $B_2B'_2$, $C_2C'_2$ тоже будут пересекаться в одной точке.

Сделаем аффинное преобразование, переводящее $\triangle ABC$ в правильный. Тогда точки A_0, B_0, C_0 перейдут в середины сторон нового треугольника, так как аффинное преобразование сохраняет отношения на прямой. В свою очередь точки A_1, B_1, C_1 перейдут в какие-то точки на сторонах. При этом точки A_2, B_2, C_2 будут по прежнему гомотетичны точкам A_1, B_1, C_1 с тем же коэффициентом k. То есть условие леммы можно считать сохранившимся.

Обозначим длину высоты из A_1 на стороны A_0B_0 и A_0C_0 через h_A. Аналогично определим h_B и h_C. Поместим в точки A_1, B_1, C_1 массы $\frac{1}{h_A}$, $\frac{1}{h_B}$ и $\frac{1}{h_C}$ соответственно. Заметим, что в этом случае массы точек A_1 и B_1 группируются в точку C_1', так как $A_1C_1' \cdot \frac{1}{h_A} = \sin\angle(A_0B_0,\, A_1B_1) = B_1C_1' \cdot \frac{1}{h_B}$. Следовательно, центр масс всей системы будет лежать на прямой C_1C_1'. Аналогично показывается, что он лежит и на прямых A_1A_1' и B_1B_1'. Следовательно, эти прямые пересекаются в одной точке. Пусть это точка M. Так как M — центр масс, то $\overrightarrow{MA_1} \cdot \frac{1}{h_A} + \overrightarrow{MB_1} \cdot \frac{1}{h_B} + \overrightarrow{MC_1} \cdot \frac{1}{h_C} = 0$. А следовательно, и $(\overrightarrow{MA_0} + \overrightarrow{A_0A_1}) \cdot \frac{1}{h_A} + (\overrightarrow{MB_0} + \overrightarrow{B_0B_1}) \cdot \frac{1}{h_B} + (\overrightarrow{MC_0} + \overrightarrow{C_0C_1}) \cdot \frac{1}{h_C} = \overrightarrow{MA_0} \cdot \frac{1}{h_A} + \overrightarrow{MB_0} \cdot \frac{1}{h_B} + \overrightarrow{MC_0} \cdot \frac{1}{h_C} + \overrightarrow{A_0A_1} \cdot \frac{1}{h_A} + \overrightarrow{B_0B_1} \cdot \frac{1}{h_B} + \overrightarrow{C_0C_1} \cdot \frac{1}{h_C} = 0$. Заметим, что $\overrightarrow{A_0A_1} \cdot \frac{1}{h_A}$ — вектор коллинеарный вектору $\overrightarrow{BC}$ и по длине равный $\frac{1}{\sin 60^\circ}$. А следовательно, сумма трех таких векторов равна нулю. Откуда получаем, что $\overrightarrow{MA_0} \cdot \frac{1}{h_A} + \overrightarrow{MB_0} \cdot \frac{1}{h_B} + \overrightarrow{MC_0} \cdot \frac{1}{h_C} = 0$. То есть M — еще и центр масс системы точек A_0, B_0, C_0 с массами $\frac{1}{h_A}$, $\frac{1}{h_B}$, $\frac{1}{h_C}$ соответственно. Заметим, что аналогичные рассуждения верны и для точек A_2, B_2, C_2. А так как высоты из них относятся к высотам из точек A_1, B_1, C_1 с коэффициентом k, то полученные массы будут пропорциональны массам, получившимся в рассуждении о точках A_1, B_1, C_1. А следовательно, точка пересечения прямых A_2A_2', B_2B_2', C_2C_2' есть также центр масс точек A_0, B_0, C_0 с массами $\frac{1}{h_A}$, $\frac{1}{h_B}$, $\frac{1}{h_C}$ соответственно. То есть совпадает с M, ч. т. д. □

Известный факт, что точки касания вписанной окружности и вневписанной окружности со стороной симметричны относительно середины этой стороны. Тогда утверждение теоремы 2 получается, если применить лемму 5 с $k = 1$ к утверждению основной теоремы.

Заметим так же, что условие пересечения трех прямых в одной точке сохраняется при проективных преобразованиях. Так что в условии леммы 5 можно заменить серединный треугольник на произвольный чевианный. А для сохранения условия равенства отношений после проективного преобразования можно потребовать, чтобы двойные отношения точек $(A,\, B_1,\, B_0,\, C)$ и $(C,\, B_2,\, B_0,\, A)$ относились так же, как и две аналогичные пары четверок точек. В частности, это соображение можно применить для следующей конструкции: взять за треугольник $A_0B_0C_0$ треугольник из оснований биссектрис треугольника, а за точки A_1, B_1, C_1, A_2, B_2, C_2 — чевианные треугольники двух изогонально сопряженных точек. Получается сам по себе интересный факт, пока не нашедший себе продолжения.

Теперь приведем доказательство Александра Скутина. В нем мы будем доказывать нужный нам факт про каждую из трех прямых отдельно. Будем доказывать для прямой A_1A'. Для остальных прямых доказательство аналогично.

Лемма 1′. $\angle(A_0F,\, FA_1) = (\angle(CA,\, CB) + \angle(BA,\, BC))/2$.

Доказательство. Покажем, что

$$\angle(A_0F,\, FH) = \angle(CA,\, CB) + \angle(BA,\, BC),$$

где H — основание высоты, опущенной из A на BC. Тогда нужный нам факт будет очевидно следовать из леммы Архимеда. Заметим, что

$$\angle(A_0F,\, FH) = \angle(A_0B_0,\, B_0H),$$

так как все эти точки лежат на окружности Эйлера. Так же заметим, что ввиду того, что B_0 — середина гипотенузы прямоугольного треугольника AHC, $\angle(CA,\, CB) = \angle(CB_0,\, CH) = \angle(HC,\, HB_0)$. Значит, $\angle(A_0B_0,\, B_0H) = \angle(A_0B_0,\, A_0H) + \angle(A_0H,\, HB_0) = \angle(BA,\, BC) + \angle(CA,\, CB)$, ч. т. д. □

Лемма 2′. Точки B_C, C_B, A_1, A_0 и F лежат на одной окружности.

Доказательство. Покажем, что $\angle(A_0C_B,\, C_BA_1) = \angle(A_0F,\, FA_1)$. Из вписанности C_BIA_1B получаем, что

$$\angle(C_BI,\, C_BA_1) = \angle(BI,\, BA_1) = \angle(BA,\, BC)/2.$$

Из того, что $C_BA_0 = A_0C$, получаем равенство

$$\angle(A_0C_B,\, C_BC) = \angle(CC_B,\, C_A0) = \angle(CA,\, CB)/2.$$

Откуда получаем, что

$$\begin{aligned}\angle(A_0C_B,\, C_BA_1) = \angle(A_0C_B,\, C_BC) + \angle(C_BC,\, C_BA_1) = \\ = (\angle(CA,\, CB) + \angle(BA,\, BC))/2 = \angle(A_0F,\, FA_1).\end{aligned}$$

То есть C_B лежит на описанной окружности $\triangle A_1A_0F$. Аналогично получаем, что и B_C лежит на ней же. Утверждение доказано. □

Мы получили, что A_1F — радикальная ось вписанной и описанной около треугольника $C_BB_CA_1$ окружностей. Так что нам осталось показать, что A' лежит на ней. Для этого достаточно заметить, что A' — центр гомотетии, переводящей треугольник $C_0C_1C_B$ в треугольник $B_0B_CB_1$, так как у этих треугольников соответственные стороны параллельны. А значит, $\frac{A'C_1}{A'B_C} = \frac{A'C_B}{A'B_1}$, откуда следует равенство степеней относительно нужных окружностей $A'C_1 \cdot A'B_1 = A'C_B \cdot A'B_C$. Основная теорема доказана.

Заметим также, что данный результат обобщается для вневписанной окружности, так как алгебраически вписанная окружность неотличима от вневписанной. То есть, если везде заменить слово «вписанная» на «вневписанная», то полученный факт тоже верен.

Приведем также несколько задач, которые предлагается решить самостоятельно.

Задача 1. Докажите, что прямые A_0A', B_0B' и C_0C' пересекаются в одной точке. Назовем эту точку F'. Докажите, что прямая FF' является общей касательной вписанной окружности и окружности Эйлера.

Задача 2. Покажите, что прямые AA', BB' и CC' параллельны.

Задача 3. В обозначениях Леммы 5 докажите, что точки A_1, B_1, C_1, A_0, B_0, C_0 и M лежат на одной конике. То есть точа M является четвертой точкой пересечения описанных коник соответствующих шестерок точек.

Задача 4. Из задачи 3 видно, что точки A_0, A_1, B_0, B_1, C_0, C_1, F и F' лежат на одной конике. Покажите, что она касается с соответственными кониками для других точек Фейербаха в точках A_0, B_0 и C_0 соответственно, причем касательными являются прямые O_CO_B и ей аналогичные.

Автор выражает благодарность за помощь в доказательстве некоторых лемм Матдинову Марселю и Мокину Василию, а так же Скопенкову Аркадию и Заславскому Алексею за помощь в оформлении работы.

Список литературы

[1] Емельянов Л. А., Емельянова Т. Л. *Семейство Фейербаха* // Математическое просвещение. Сер. 3, вып. 6, 2002. С. 78–92.

[2] Прасолов В. В. *Задачи по планиметрии*. М.: МЦНМО, 2007.

Сопротивление между узлами решетки

Д. Баранов М. Скопенков А. Устинов

Данная подборка задач составлена по мотивам проекта «Случайные блуждания и электрические цепи» XXII Летней конференции Турнира городов и задачи 14.12 из задачника «Математического Просвещения» (вып. 14, с. 274).

1. Предположим, что имеется граф Γ, у которого сопротивление каждого ребра равно 1. Возьмем в графе Γ два смежных ребра AB и AC. Эти рёбра назовем *эквивалентными*, если существует перестановка вершин графа, переводящая вершины, соединенные ребром, в вершины, соединенные ребром, при которой A переходит в A, а B — в C. Вершину графа назовем *центром симметрии* графа Γ, если все рёбра, содержащие ее, эквивалентны. Граф Γ называется *правильным*, если все его вершины — центры симметрии графа.
(A) (А. Ходулёв) Пусть правильный граф содержит n вершин, A и B — соседние вершины степеней a и b, соответственно. Докажите, что сопротивление между ними равно

$$\left(\frac{1}{a}+\frac{1}{b}\right)\left(1-\frac{1}{n}\right).$$

(B) Если же взять 2-мерную целочисленную решетку, то $1/n$ в последней формуле нужно заменить нулем.
(C) Приведите пример бесконечного правильного графа, для которого формула из пункта (A) (с заменой $1/n$ нулем) не дает правильного ответа.
(D) Докажите, что существует такая функция $f(m,n)\colon \mathbb{Z}^2 \to \mathbb{R}$, что $f(0,0)=1$, $f(0,1)=-1$,

$$f(m,n)=\frac{1}{4}\left(f(m-1,n)+f(m+1,n)+f(m,n-1)+f(m,n+1)\right)$$

для каждой пары $(m,n)\neq(0,0),(0,1)$, и $f(m,n)$ стремится к 0 при стремлении $r(m,n):=\sqrt{m^2+n^2}$ к бесконечности.

2. К двум соседним вершинам проволочного (A) икосаэдра; (B) додекаэдра; (C) правильного графа, полученного из центрально-симметричного многогранника с n вершинами; подвели напряжение так, что по

соединяющему их ребру потек ток I. Какой при этом будет течь ток по диаметрально противоположному ребру?

3. (A) Из резисторов спаяна цепь. Может ли сопротивление между какими-то двумя ее клеммами увеличиться, если припаять еще один?
(B) Электрическая цепь спаяна из нескольких резисторов и одной батарейки. В цепи выбрали две клеммы с потенциалами A и B, отличные от полюсов батарейки, и объединили (закоротили) их в одну клемму. Докажите, что потенциал новой клеммы в полученной цепи будет заключен между величинами A и B.

4. (A) Из металлической сетки вырезано кольцо с внутренним радиусом rn и внешним — Rn (центры обоих кругов — в начале координат). Сопротивление каждого ребра равно 1. Если некоторое ребро разрезано, то сопротивление оставшегося куска пропорционально его длине. Докажите, что сопротивление кольца равно

$$\frac{1}{2\pi}\ln\frac{R}{r}+O\left(\frac{\ln n}{n}\right).$$

(B) Докажите аналогичную формулу для сопротивления между и центром и границей круга радиуса R, вырезанного из металлической сетки: $\frac{1}{2\pi}\ln R+O(1)$.

Авторы благодарны И.В. Богданову, А.Я. Канелю, М.В. Прасолову и Г.Р. Челнокову за полезные обсуждения.

Дмитрий Баранов, мехмат МГУ
E-mail:dimbaranov@mail.ru
Михаил Скопенков, ИППИ РАН
E-mail:skopenkov@rambler.ru
Алексей Устинов, ХО ИПМ ДВО РАН
E-mail:ustinov.alexey@gmail.com

Нам пишут

Отклик на статью Е. Алексеевой

А. Б. Скопенков

В [2] приводятся четкая формулировка и короткое доказательство основного результата работы [1], а также проясняется его связь с изопериметрической теоремой для плоскости Лобачевского. Доказательство по сути не отличается от приведенного в [1]. Однако ввиду красоты и важности результата короткое доказательство, освобожденное от ненужных деталей, может быть интересно читателю.

Список литературы

[1] Алексеева Е. *Гиперболические треугольники максимальной площади с двумя заданными сторонами* // Мат. Просвещение. Сер. 3, вып. 14, 2010. С. 175–183. См. также J. I. Alekseeva. *Hyperbolic triangles of the maximum area with two fixed sides*. `http://arxiv.org/abs/0911.5319`.

[2] *A simple proof of the isoperimetric theorem for the hyperbolic plane* A. Skopenkov (editor). `http://arxiv.org/pdf/1009.0897v1`.

Задачный раздел

В этом разделе вниманию читателей предлагается подборка задач разной степени сложности, в основном трудных. Составителям этой подборки кажется, что предлагаемые ниже задачи окажутся интересными как для сильных школьников, интересующихся математикой, так и для студентов-математиков.

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.

В скобках после условия задачи приводится фамилия автора (уточнения со стороны читателей приветствуются). Если автор задачи нам неизвестен, то в скобках указывается «фольклор».

1. Задачи на «устный счет»:

   а) Найдите первую цифру числа 2^{400}. (*А. Я. Белов*)

   б) Найдите $[2^{\sqrt{15}}]$, не пользуясь калькулятором. (*А. В. Спивак*)

   в) Что больше: $\sqrt[3]{60}$ или $2+\sqrt[3]{7}$? (*В. А. Сендеров*)

   г) Оцените $\int_0^{2\pi} \sin^{100} x\,dx$ с 20% погрешностью. (*В. И. Арнольд*)

2. Может ли сумма двух периодических функций с наименьшими периодами 1 и $\sqrt{2}$ снова быть периодической функцией, отличной от константы?

3. Гипербола $H=\{(x,y) : xy=1\}$ повернута на угол α относительно начала координат $(0,0)$; получилась гипербола H_α. Найдите угол между их касательными в точках пересечения H и H_α.

   (*А. В. Акопян, D.Schleicher*)

4. Во все точки целочисленной решетки на плоскости вбиты гвозди. На плоскость положили отрезок длины 2011, не задевающий ни одного их этих гвоздей.

   а) Можно ли передвинуть отрезок, не задевая ни одного гвоздя, так, чтобы в результате он развернулся на 180°?

   б) Существует ли такое начальное положение отрезка, при котором его можно повернуть вокруг некоторой точки на 180° так, чтобы он не задел ни одного гвоздя?

   (Авторам неизвестно, верно ли утверждение пункта б) для произвольного начального положения отрезка.)

   (*В. А. Сендеров, А. Я. Белов*)

5. Даны две бесконечно дифференцируемые функции $f(x) = x + a_2x^2 + a_3x^3 + \ldots$, $g(x) = x + b_2x^2 + b_3x^3 + \ldots$. Известно, что $f(g(x)) = x$ и что все числа $k!a_k$ — целые. Докажите, что что все числа $k!b_k$ — тоже целые. (*А. В. Бунькова*)

6. На ленте записана бесконечная последовательность цифр. Докажите, что либо из нее можно вырезать 10 стозначных чисел, идущих в порядке убывания, либо какая-нибудь комбинация цифр повторяется 10 раз подряд. А если длина ленты 100^{100} или 100^{20}? Оцените достаточную длину ленты. (*А. Я. Белов*)

7. *Алгебраическая кривая порядка* m задана уравнением
$$\sum_{k+l\leqslant m} a_{kl}x^k y^l = 0.$$
Докажите, что количество ее *овалов* (т. е. компонент связности) не превосходит $(m-1)(m-2)/2+1$. (*Теорема Гарнака*)

8. а) Расстояние между промежуточными лагерями 1 день пути. Экспедиция хочет отнести 1 банку консервов на расстояние n дней пути от базового лагеря и вернуться обратно. При этом каждый член экспедиции может нести с собой не более трех банок консервов, а в день он съедает одну банку. Каково наименьшее число банок консервов должно быть употреблено для этой цели? (Оставлять банки можно только в промежуточных лагерях.) б) Рассмотрите случай, когда каждый член экспедиции может нести с собой k банок консервов, а также случай камикадзе. в) Непрерывный аналог этой задачи: самолеты могут заправляться в воздухе, а дальность полета составляет 1000 км. Количество бензина, который для этого тратится принимается за единицу. Сколько нужно бензина чтобы пролететь 10 тыс. км? (*А.Я.Белов*)

9. а) Дана 2×2 матрица A с вещественными коэффициентами. Докажите, что ее можно представить как сумму квадратов двух матриц второго порядка с вещественными коэффициентами.
(*SEEMOUS* 2010)

   б)* Можно ли матрицу размера $n\times n$ с вещественными коэффициентами представить в виде суммы квадратов нескольких матриц размера $n\times n$ с вещественными коэффициентами? Если «да», то каково минимальное число квадратов?
(*Охад Ливне Бар-Он, Шахар Кармиели*)

10. На грани правильного тетраэдра отмечена точка. Докажите, что тетраэдр можно разрезать на четыре равных выпуклых многогранника так, чтобы эта точка была вершиной одного из них.
(*И. И. Богданов*)

11. а) Пусть $M = (a, b)$ — интервал на положительной полупрямой $(0, +\infty)$. Доказать, что интервалы $nM = (na, nb)$ $(n = 1, 2, \dots)$ покрывают полупрямую $(c, +\infty)$ для достаточно большого c и, значит, дополнение к их объединению имеет конечную меру.

    б) Придумать пример подмножества $M \subset (0, \infty)$ положительной меры, не обладающего указанным выше свойством, т. е. такого, что дополнение к объединению подмножеств nM $(n = 1, 2, \dots)$ имеет бесконечную меру. (*Э. Б. Винберг*)

12. а) Множество X вершин правильного n-угольника таково, что равна нулю сумма векторов с началом в центре многоугольника и концами в вершинах множества X. Пусть $n = p^\alpha q^\beta$, где p, q — простые числа. Докажите, что для некоторого k множество X можно разбить на непересекающиеся множества, каждое из которых является множеством вершин правильного k-угольника. Докажите, что при любом n, делящемся хотя бы на три различных простых числа, это не всегда верно. (Пара противоположных вершин рассматривается как 2-угольник).

    б) Пусть $v_1, \dots, v_n$ — векторы из центра правильного n-угольника в его последовательные вершины, $a = (a_1, \dots, a_n)$ — такой целочисленный вектор, что $\sum_{i=1}^n a_i v_i = 0$. Докажите, что a есть целочисленная линейная комбинация векторов вида

$$e_{d,k} = (\underbrace{0, \dots, 0}_{d}, 1, \underbrace{0, \dots, 0}_{k-1}, 1, \underbrace{0, \dots, 0}_{k-1}, \dots, 1, \underbrace{0, \dots, 0}_{k-d-1})$$

    при k, делящих n. (По сути, $e_{k,d}$ соответствует сумме векторов в вершины некоторого правильного (n/k)-угольника.)
(*И. И. Богданов, Э. Б. Винберг, Г. А. Гальперин, Г. Р. Челноков*)

Исправленные условия задач

К сожалению, некоторые условия задач из задачника «Математического просвещения» в вып. 14 содержали ошибки.

Приводим исправленные формулировки задач.

14.7. Пусть d — нечетный делитель натурального числа $p-1$. Докажите, что в p-ичной системе счисления существует d-значное число, равное определителю $d \times d$ матрицы, составленной из цифр этого числа и их циклических перестановок. Например, в десятичной системе счисления $(p = 10)$

$$692 = \det\begin{pmatrix} 6 & 2 & 9 \\ 9 & 6 & 2 \\ 2 & 9 & 6 \end{pmatrix} \text{ или}$$

$$456790123 = \det\begin{pmatrix} 4 & 5 & 6 & 7 & 9 & 0 & 1 & 2 & 3 \\ 3 & 4 & 5 & 6 & 7 & 9 & 0 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 & 7 & 9 & 0 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 9 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 9 \\ 9 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 9 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 7 & 9 & 0 & 1 & 2 & 3 & 4 & 5 \\ 5 & 6 & 7 & 9 & 0 & 1 & 2 & 3 & 4 \end{pmatrix}.$$

(*Н. И. Белухов*)

14.7. Дан $\triangle ABC$. A_1, B_1, C_1 — точки касания сторон BC, AC, AB с вписанной окружностью соответственно. A_0, B_0, C_0 — середины сторон. Обозначим точку пересечения прямых A_0B_0 и A_1B_1 через C'. Аналогично определяются точки A' и B'. Докажите, что прямые A_1A', B_1B' и C_1C' пересекаются в точке Фейербаха. (*Ф. Ивлев*)

Решения задач из предыдущих выпусков

12.1. Условие. $\cos\alpha = 1/3$. Докажите, что градусная мера угла α иррациональна.

Решение. Достаточно убедиться, что $\cos(2^n\alpha) \neq \cos(2^m\alpha)$ при $m \neq n$. В самом деле: если $\alpha = p/q \cdot 2\pi$, p, q — целые числа, то найдутся такие $n \neq m$, что 2^n и 2^m сравнимы по модулю q. В этом случае углам $2^n\alpha$ и $2^m\alpha$ отвечает одна и та же точка на единичной окружности и соответствующие косинусы совпадают.

Покажем, что $\cos(2^n\alpha)$ имеет вид $p_n/3^{2^n}$, где p_n не делится на 3. Тогда $\cos(2^n\alpha) \neq \cos(2^m\alpha)$ при $m \neq n$.

В силу равенства $\cos(2\varphi) = 2\cos^2(\varphi) - 1$ имеем $p_{n+1} = 2p_n^2 - 3^{2^n}$ (не делится на 3), а знаменатель дроби для $\cos(2^{n+1}\alpha)$ равен $3^{2^{n+1}}$, так что дело завершает индукция. Задача решена.

Замечание. На окружности радиуса $5^{k/2}$ лежит ровно $k+1$ точка с натуральными координатами. Полезное упражнение — доказать это элементарными методами, без использования теории гауссовых целых чисел.

Комментарии. 1. Пусть β — градусная мера угла правильного тетраэдра. Легко убедиться, что $\cos(2\beta) = 1/3$. На результате задачи 12.1 основано решение Третьей проблемы Гильберта о неравносоставленности куба и правильного тетраэдра. Грубо говоря, поскольку двугранные углы правильного тетраэдра иррациональны, при разрезании и составлении куба образуются «щели», ведь если примыкание нескольких ребер многогранников выходит на поверхность, то сумма их двугранных углов равна π, если на ребро куба — то $\pi/2$, а если образуется ребро «в глубине» — то сумма соответствующих двугранных углов равна 2π. Подробнее см. книгу В. Г. Болтянского «Третья проблема Гильберта».

2. Возникает вопрос: *Какие углы у треугольников на целочисленной решетке выражаются целым числом градусов*? (Ответ — углы, кратные $45°$).

Этот вопрос сводится к такому: *Какие правильные многоугольники можно вписать в целочисленную решетку*? Ответ — только квадраты.

В правильную треугольную решетку можно вписать только правильный треугольник и шестиугольник. Прочие правильные многоугольники никуда не вписываются.

Подробнее см. А. Я. Канель-Белов, А. К. Ковальджи, «Как решают нестандартные задачи». (*А. Я. Канель-Белов*)

13.3. Условие. Известно, что для любой последовательности $\{a_i\}_{i=1}^{\infty}$ такой, что $\sum_{i=1}^{\infty} a_i^2 < \infty$, ряд $\sum_{i=1}^{\infty} a_i b_i < \infty$. Докажите, что тогда $\sum_{i=1}^{\infty} b_i^2 < \infty$.

Решение. Докажем, что $|b_n| \to 0$.

Для $C > 0$ рассмотрим подпоследовательность m_k такую, что $|b_{m_k}| > C$. Если эта подпоследовательность бесконечна, то условие задачи нарушается для последовательности

$$a_n = \begin{cases} \dfrac{\operatorname{sign} b_{m_k}}{k}, & \text{если } n = m_k, \\ 0, & \text{в противном случае.} \end{cases}$$

Значит, для любого C лишь конечное количество членов последовательности b_n по модулю превосходит C.

Теперь рассуждаем от противного. Предположим, что $\sum_{i=1}^{\infty} b_i^2 = \infty$. Рассмотрим последовательность

$$a_n = \frac{b_n}{f_n}, \quad \text{где } f_n = \sum_{k=1}^{n} b_k^2.$$

Достаточно доказать, что $\sum_n a_n b_n$ расходится, а $\sum_n a_n^2$ — сходится.

Для натурального числа N выделим отрезок натурального ряда $\ell_N, \ell_N + 1, \ldots, r_N$, состоящий из всех чисел n, для которых $N - 1 \leqslant f_n < N$. Имеем оценку

$$\sum_{n\in[\ell_N, r_N]} a_n b_n = \sum_{n\in[\ell_N, r_N]} \frac{b_n^2}{f_n} \geqslant \frac{\sum_{n\in[\ell_N, r_N]} b_n^2}{N} \geqslant \frac{1}{2N}$$

при достаточно больших N. Отсюда получаем расходимость ряда $\sum_n a_n b_n$.

С другой стороны, при достаточно больших N для $n \in [\ell_N, r_N]$ выполняется $b_n^2 < 1/2$. Поэтому

$$\sum_{n\in[\ell_N, r_N]} a_n^2 = \sum_{n\in[\ell_N, r_N]} \frac{b_n^2}{f_n^2} \leqslant \frac{1}{2(N-1)^2}\,.$$

Отсюда получаем сходимость ряда $\sum_n a_n^2$. (*М. Н. Вялый*)

14.3. Условие. Бесконечное множество точек на плоскости таково, что все попарные расстояния целые. Докажите, что все точки лежат на одной прямой. А что если все попарные расстояния рациональны?

Решение. а) Рассмотрим пару точек A, B и произвольную точку X нашего множества M. Заметим, что в силу неравенства треугольника $\big||XA|-|XB|\big| \leqslant |AB|$. Поскольку все попарные расстояния между точками целые, количество возможных значений величины $\big||XA|-|XB|\big|$ не превосходит $|AB|$. Таким образом, все точки лежат на конечном числе гипербол $\big||XA|-|XB|\big|=c$; $c=0,\ldots,|AB|$ с фокусами A и B.

При $c=0$ и $c=|AB|$ эти гиперболы вырождаются в прямые (серединный перпендикуляр к отрезку $[AB]$ и прямая (AB) соответственно). На одной из них (гиперболе Γ) лежит бесконечное множество отмеченных точек, ибо множество M бесконечно.

Теперь возьмем пару точек $C \neq D$ из множества M, лежащие на Γ, и построим семейство гипербол Γ_i, с ними связанные. Поскольку множество отмеченных точек на Γ бесконечно, бесконечное подмножество из них лежат на одной из Γ_i, т. е. на пересечении $\Gamma \cap \Gamma_i$.

С другой стороны, любые две гиперболы либо совпадают, либо их пересечение состоит не более чем из четырех точек. Поэтому $\Gamma=\Gamma_i$. Поскольку фокусы гиперболы Γ_i лежат на Γ это возможно только, если они обе прямые. Обозначим эту прямую за l.

Поскольку не все точки из M лежат на одной прямой, есть отмеченная точка $O \in M$ не принадлежащая l. Пусть $Q \in M \cap l$.

Остается рассмотреть конечное семейство гипербол с фокусами Q и O (на которых лежат отмеченные точки) и их пересечение с прямой l. Поскольку ни одна из них не может совпасть с l, пересечение каждой из них с l состоит не более чем из двух точек, т. е. конечно. Значит, отмеченных точек на l конечно.

Получили нужное противоречие. Задача допускает пространственное обобщение.

б) Эта задача является контрапунктом к предыдущей.

Покажем, что на единичной окружности можно выбрать бесконечное множество точек такое, что попарные расстояния между ними рациональны. Если точки A, B отвечают углам φ и ψ, то расстояние между ними равно $2\sin((\varphi-\psi)/2)$. Поэтому если $\sin(n\varphi)$ рационален при всех n, то все попарные расстояния между точками $A_k=(\sin(2k/\varphi),\cos(2k/\varphi))$ рациональны.

С помощью равенств

$$\sin(k+1)\varphi=\sin(k\varphi)\cos(\varphi)+\cos(k\varphi)\sin(\varphi)$$

легко проверить по индукции что если $\sin(\varphi)$ и $\cos(\varphi)$ рациональны, то и $\sin(n\varphi)$ рационален при всех n. Если при этом градусная мера угла φ иррациональна, то точки A_k попарно различны и образуют искомое бесконечное множество.

Возьмем $\varphi = \arcsin 3/5$. Тогда $\sin(\varphi) = 3/5$, $\cos(\varphi) = 4/5$ так что $\sin(n\varphi)$ рационален при всех n. Остается проверить только, что градусная мера угла φ иррациональна.

А это доказывается аналогично решению задачи 12.1 (см. выше, с. 236). (*А. Я. Канель-Белов*)

14.5. Условие. Бесконечно Мудрый Таракан живет на плоскости. Он близорук и потому видит Истину, только когда находится не более, чем в одном шаге от нее. Первоначально Таракан находится в n шагах от Истины. Когда Таракан делает шаг, друзья говорят говорят ему, приблизился он к Истине, или нет. а) Докажите, что, пользуясь этой и только этой информацией, Таракан может достичь Истины менее чем за $n+10\ln n$ шагов. б) Докажите, что не существует алгоритма, позволяющего достичь Истины менее чем за $n+0.1\ln n$ шагов.

Решение.

а) Опишем алгоритм. Он состоит из двух фаз.

1. Определение направления на Истину.

2. Путь по выделенному направлению.

Конусом возможных направлений K назовем множество возможных направлений на Истину из начальной точки. Это угол, величина которого *раствор* равен φ. Вначале $\varphi = 2\pi$. *Действие* состоит в паре шагов – шаг от начальной точки и шаг к ней обратно.

Лемма. *Существует действие, сокращающее раствор в* 2 *раза.*

Доказательство. Случай $\varphi = 2\pi$ проверяется непосредственно. Пусть $\vec{l}$ — вектор биссектрисы K, $\vec{m}$ — вектор перпендикулярного направления. Ясно, что действие, связанное с направлением m приводит к искомому сокращению: ответу «приблизился» отвечает половина K лежащая в угле между $\vec{l}$, $\vec{m}$, а ответу «удалился» отвечает половина K лежащая в угле между $\vec{l}$, $-\vec{m}$. Лемма доказана.

За $\log_2(n)+2$ действий (т. е. $2[\log_2(n)+3]$ шагов) можно определить направление на Истину с точностью до угла $\varphi = 2\pi/(4n)$, при этом уклонение сторон угла от биссектрисы составит $\pi/(4n) < \arcsin(1/n)$ так что после этого путь по биссектрисе приводит к успеху за $n+1$ шаг.

б) Назовем шаг Таракана *разведывательным*, если его направление с направлением на деталь образует угол больше $\pi/3$. C есть *дуга возможных положений Истины*. Это дуга окружности с центром в начальной точке и радиусом n, такая, что любое положение на ней согласуется с ответами, полученными Роботом. Вначале это вся окружность.

Лемма 1. *Количество l разведывательных шагов не превосходит* $\log_2(n)/5$.

Доказательство. Разведывательные шаги приводят к приближению на величину, не превосходящую $l\cos\pi/3$. Поэтому общее число шагов не меньше $l+(n-l\cos\pi/3)=n+l(1-\cos\pi/3)=n+l/2$ и если $l>\log_2(n)$ то величина $n+l/2$ окажется больше $n+\log_2(n)/10$, что противоречит условию задачи. Лемма доказана.

Назовем *начальным участком* первые $n/3$ шага.

В справедливости следующего утверждения легко убедиться непосредственно.

Лемма 2. *После первого шага только разведывательный шаг приводит к изменению угла раствора. При этом C разбивается на две дуги, отвечающие ответам «приблизился» и «удалился» (возможно, одна из них пустая). Длина $|C'|$ наибольшей из них C' не превосходит $|C|/2$.*

Следствие. *Для любого алгоритма с l проверочными шагами существует положение Истины такое, что после начального участка угловой раствор дуги C будет не менее чем $\pi n/2^l$.*

Следующая лемма также проверяется непосредственно:

Лемма 3. *Пусть A — произвольная точка, тогда расстояние от A до самой далекой точки дуги C не меньше чем $d+|C|/2\sqrt{n}$.*

В силу леммы 1 тогда $|C|<\pi\cdot n/2^{\log_2 n/5}=\pi\cdot n^{4/5}$. И после первых $n/3$ шагов расстояние от положения Таракана до самой удаленной от него точки дуги C будет не меньше $2n/3+|C|/\sqrt{n}=2n/3+n^{0.3}/2$. Общее число шагов в наихудшем случае будет тогда не меньше чем $n/3+2n/3+n^{0.3}/2>$ $>n+\log_2(n)/10$. Задача решена. (*А. Я. Канель-Белов*)

14.6. Условие. $f(x,y)$ — бесконечно дифференцируемая функция от двух переменных с локальным минимумом в нуле. Других критических точек у ней нет. Верно ли, что этот минимум глобальный? (Точка называется *критической*, если в ней обе частные производные $\partial f/\partial x$ и $\partial f/\partial y$ обращаются в нуль.)

Ответ: точка $(0,0)$ не обязательно является точкой глобального минимума!

Первое решение. Рассмотрим многочлен $P(x)=x^2+y^2(1-x)^3$. Тогда $\partial P/\partial x=2x-3y^2(1-x)^2$, $\partial P/\partial y=2y(1-x)^3$. Если $\partial P/\partial y=0$, то $y=0$ либо $x=1$. Если при этом $x=1$ и $\partial P/\partial x=0$ то $x=0$, что противоречит равенству $x=1$.

Итак, $y=0$. Тогда равенство $\partial P/\partial x=2x-3y^2(1-x)^2=0$ влечет равенство $x=0$. Итак, $(0,0)$ — единственная критическая точка P и легко проверить, что это точка локального минимума. Она не является точкой глобального минимума, поскольку P принимает сколь угодно большие по модулю отрицательные значения. В самом деле, пусть $y = 1$, тогда $P(x, 1) = x^2 + (1 - x)^3$. Ясно, что $P(x, 1) \to -\infty$ при $x \to \infty$.

(*В. О. Бугаенко*)

Второе решение. Рассмотрим функцию общего вида (т. е., ее критические точки образуют дискретное множество), имеющую изолированный локальный минимум, не являющийся глобальным. Рассмотрим связную односвязную область, содержащую этот локальный минимум, а также некоторую точку, в которой она принимает меньшее значение, и не содержащую никаких других критических точек. Голоморфным преобразованием переведем эту область во всю плоскость. Тогда заданная функция на области определит новую функцию на всей плоскости. Она и будет искомой.

(*В. Гальперин*)

14.7. Условие. Пусть d — нечетный делитель натурального числа $p - 1$. Докажите, что в p-ичной системе счисления существует d-значное число, равное определителю $d \times d$ матрицы, составленной из цифр этого числа и их циклических перестановок.

Решение. Обозначим

$$q = \frac{p^d - 1}{(p - 1)d}\,.$$

Рассмотрим те кратные tq числа q, для которых выполняются условия $\frac{d-1}{2}(p - 1) < t < \frac{d+1}{2}(p - 1)$ и $(t, d) = 1$. Докажем, что каждое такое число равно определителю циркулянтной матрицы из его p-ичных цифр.

Под циркулянтной матрицей $C(x_1, x_2, \dots, x_n)$ мы понимаем матрицу вида

$$\begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ x_n & x_1 & x_2 & \dots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \dots & x_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ x_2 & x_3 & x_4 & \dots & x_1 \end{pmatrix}$$

Для определителя $\det C(x_1, x_2, \dots, x_n)$ циркулянтной матрицы известна следующая формула:

$$\det C(x_1, x_2, \dots, x_n) = \prod_{j=1}^{n} (x_1 + x_2\omega_j + x_3\omega_j^2 + \dots + x_n\omega_j^{n-1}),$$

где ω_j — корни n-й степени из единицы.

Мы используем эту формулу в двух случаях (в обоих ответ можно также получить элементарными преобразованиями строк и столбцов циркулянтной матрицы). В первом случае числа $x_1, x_2, \ldots, x_n$ образуют арифметическую прогрессию с суммой s и разностью h. В этом случае

$$C(x_1, x_2, \ldots, x_n) = sn^{n-2}(-h)^{n-1}.$$

Во втором случае числа $x_1, x_2, \ldots, x_n$ образуют геометрическую прогрессию вида $x_i = p^{n-i}$ для некоторого действительного числа p. В этом случае

$$C(x_1, x_2, \ldots, x_n) = (p^n - 1)^{n-1}.$$

Итак, рассмотрим число описанного выше вида. Обозначим через $A_1 = tq = \overline{a_1 a_2 \ldots a_d}$ его p-ичную запись, а через $A_i = \overline{a_i a_{i+1} \ldots a_d a_1 \ldots a_{i-1}}$ — i-ю циклическую перестановку, $i = 2, 3, \ldots, d-1$. Непосредственным перемножением матриц можно проверить, что

$$C(a_1, a_2, \ldots, a_d) C^T(p^{d-1}, p^{d-2}, \ldots, p, 1) = C(A_1, A_2, \ldots, A_{d-1}, A_d). \qquad (*)$$

Здесь C^T обозначает транспонированную матрицу. Из предыдущего имеем $\det C^T(p^{d-1}, p^{d-2}, \ldots, p, 1) = (p^d - 1)^{d-1}$.

Наша следующая цель — доказать два утверждения: (1) применением четного числа транспозиций строк и столбцов матрицу $C(A_1, A_2, \ldots, A_d)$ можно привести к виду $C(B_0, B_1, \ldots, B_{d-1})$, где $B_0, B_1, \ldots, B_{d-1}$ — перестановка чисел $A_1, A_2, \ldots, A_d$ в возрастающем порядке; (2) числа B_0, $B_1, \ldots, B_{d-1}$ образуют арифметическую прогрессию.

Начнем с доказательства утверждения (2). Заметим, что

$$A_{i+1} = pA_i - (p^d - 1)a_i.$$

Поэтому $A_i \equiv p^{i-1} A_1 \pmod{p^d - 1}$. Учитывая, что $(t, d) = 1$, отсюда следует, что числа $A_1, A_2, \ldots, A_d$ попарно различны и все они имеют одинаковый остаток по модулю $(p-1)q$. С другой стороны, все эти числа содержатся в интервале $(0, p^d - 1)$ длины $d[(p-1)q]$. Поэтому $A_1, A_2, \ldots, A_d$ — в точности все числа из интервала $(0, p^d - 1)$, имеющие один и тот же остаток по модулю $(p-1)q$. Значит, упорядоченный набор этих чисел образует арифметическую прогрессию с разностью $(p-1)q$. Из ограничений на t заключаем, что средний член $B_{(d-1)/2}$ этой прогрессии равен A_1, а ее сумма равна dA_1.

Теперь докажем (1). Обозначим $r = (d-1)/2$, так что $A_1 = B_r$.

Перестановками строк и столбцов за исключением первого преобразуем матрицу $C(A_1, A_2, \ldots, A_d)$ к виду

$$M = \begin{pmatrix} B_r & B_{r+1} & \dots & B_{d-1} & B_0 & \dots & B_{r-1} \\ B_{r-1} & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots \\ B_0 & \dots & \dots & \dots & \dots & \dots & \dots \\ B_d & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots \\ B_{r+1} & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Этого можно добиться, сделав одинаковое количество транспозиций строк и транспозиций столбцов. Действительно, у циркулянтной матрицы первая строка, прочитанная слева направо, совпадает с первым столбцом, прочитанным снизу вверх. Аналогичное свойство выполняется и для матрицы M. Значит, упорядочить первую строку и первый столбец в некотором порядке можно, выполнив равное количество транспозиций строк и столбцов.

Мы докажем, что матрица M — циркулянтная. Из этого легко следует утверждение (1). Действительно, циклическая перестановка строк (столбцов) сохраняет циркулянтный вид матрицы и является четной перестановкой (так как d нечетно). После r таких перестановок матрица M (в предположении, что она циркулянтная) станет равной $C(B_0, \dots, B_{d-1})$.

Рассмотрим элемент матрицы M, который стоит на пересечении столбца, в первой строке которого стоит $B_i = A_x$, и строки, в первом столбце которой стоит $B_j = A_y$. Этот элемент равен B_k для некоторого k. Условие циркулянтности матрицы M равносильно $i + j \equiv k + r \pmod d$.

С другой стороны, исходная матрица $C(A_1, \dots, A_d)$ циркулянтная. Поэтому $B_k = A_m$ и $m + 1 \equiv x + y \pmod d$.

Поскольку A_i образуют геометрическую прогрессию по модулю $p^d - 1$, из последнего сравнения получаем $A_1 A_m \equiv A_x A_y \pmod{p^d - 1}$. Из равенств $A_1 = B_r$, $A_m = B_k$, $A_x = B_i$, $A_y = B_j$ получаем

$$\begin{aligned} A_x A_y \equiv B_i B_j \equiv (B_0 + iq(p-1))(B_0 + jq(p-1)) \equiv \\ \equiv B_0^2 + B_0(i+j)q(p-1) \pmod{p^d - 1} \end{aligned}$$

и

$$A_1 A_m \equiv B_0^2 + B_0(r+k)q(p-1) \pmod{p^d - 1}.$$

Поэтому

$$B_0(i+j)q(p-1) \equiv B_0(r+k)q(p-1) \pmod{p^d - 1},$$

т. е.

$$B_0(i+j) \equiv B_0(r+k) \pmod d.$$

Но $(B_0, d) = 1$, так что выполняется сравнение

$$(i+j) \equiv (r+k) \pmod d$$

и матрица M циркулянтная.

Для завершения доказательства напишем равенство определителей, которое вытекает из $(*)$:

$$\det C(a_1, \ldots, a_d)(p^d-1)^{d-1} = \det C(A_1, \ldots, A_d) =$$
$$= \det C(B_0, \ldots, B_{d-1}) = sd^{d-2}(-h)^{d-1} =$$
$$= dA_1 d^{d-2}[(1-p)q]^{d-1} = A_1(p^d-1)^{d-1}.$$

Отсюда получаем искомое равенство

$$C(a_1, a_2, ... a_d) = A_1.$$

Замечание. Для десятичной системы это доказательство дает такие числа, равные определителям циркулянтных матриц, составленных из их цифр: 456 790 123, 469 135 802, 493 827 160, 506 172 839, 530 864 197 и 543 209 876.

Некоторые 3-значные числа указанного вида содержат нули в десятичной записи. В этом случае определитель упрощается и мы получаем два числа, которые равны сумме кубов их цифр: $370 = 3^3 + 7^3 + 0^3$ и $407 = 4^3+0^3+7^3$. Тот факт, что оба эти числа делятся на 37, послужил толчком для придумывания данной задачи. Одна из первых формулировок этого факта имела такой вид:

Пусть $\overline{abc}$ — кратно 37. Докажите, что число $a^3+b^3+c^3-3abc$ равно либо 0, либо еще одному кратному 37, составленному из десятичных цифр a, b и c.

(*Н. И. Белухов*)

Список решений задач из задачника «Математического просвещения»

Решения задач задачника «Математического просвещения» рассеяны по 15 выпускам нашего сборника. Поэтому нам показалось разумным привести список всех задач с указанием выпусков, в которых опубликованы их решения. Решения многих задач еще не опубликованы. Мы будем благодарны тем, кто пришлет решения задач.

В списке ссылки имеют вид ⟨номер выпуска⟩, ⟨страницы⟩.

Задача	Исправления	Решение(-я)	Задача	Исправления	Решение(-я)
1.1.		6, 137–139	1.2.		5, 218
1.3.		4, 218	1.4.		5, 218–221
1.5.		6, 139–140	1.6.		5, 221–223 10, 274
1.7.		4, 219	1.8.		4, 220
1.9.		5, 223–225	1.10.		5, 225–227
2.1.		4, 221	2.2.		5, 227
2.3.		4, 221	2.4.		4, 221–222
2.5.		4, 222	2.6.		4, 222–223
2.7.		6, 140–142	2.8.		6, 142–145
2.9.		6, 145–147	2.10.	6, 135	8, 248–249
3.1.		4, 223	3.2.		4, 223–224
3.3.		5, 227–228	3.4.		4, 224
3.5.		4, 225	3.6.		7, 190–193
3.7.		6, 147–148	3.8.		6, 148–149
3.9.	8, 247	15, 206–211	3.10.		8, 239–245
3.11.		8, 186–221			
4.1.		6, 149–150	4.2.		7, 193
4.3.		7, 193–194	4.4.		7, 194
4.5.		6, 150	4.6.		6, 150–151
4.7.		6, 151–152	4.8.		5, 228–229
4.9.		8, 237–238	4.10.		7, 194–195
4.11.		8, 249–252 15, 212–218	4.12.		
5.1.		6, 152–153	5.2.	6, 135	7, 195–196

Задача	Исправления	Решение(-я)	Задача	Исправления	Решение(-я)
5.3.		6, 153	5.4.		7, 196–198
5.5.		8, 252–254	5.6.		14, 275–276
5.7.		6, 153	5.8.		7, 198
5.9.		10, 232–242	5.10.		8, 254
		11, 145–148			
6.1.		8, 255	6.2.		9, 225
6.3.		9, 225–226	6.4.		9, 226–227
6.5.			6.6.		8, 255–256
6.7.		11, 165–166	6.8.		8, 256–257
6.9.		8, 222–228	6.10.		
		229–236			
6.11.		9, 227–229	6.12.		8, 258–259
7.1.		9, 229	7.2.		8, 259–260
7.3.		9, 229	7.4.		9, 229–230
		10, 274–275			
7.5.			7.6.		
7.7.			7.8.		13, 182–184
7.9.		9, 230–232	7.10.	11, 164	
7.11.		9, 232–233	7.12.		11, 166–168
8.1.		13, 184	8.2.		10, 281
8.3.		10, 281–282	8.4.		10, 282–284
8.5.		9, 215–217	8.6.		
8.7.			8.8.		9, 233
8.9.		11, 168–169	8.10.		
8.11.			8.12.		
9.1.		14, 276–277	9.2.		13, 184–185
9.3.			9.4.		10, 284–285
9.5.		14, 277	9.6.		12, 237
9.7.			9.8.		12, 238
9.9.		11, 169–171	9.10.		10, 265–272
9.11.					
10.1.		13, 185–186	10.2.		11, 172
10.3.			10.4.		11, 173–174
10.5.		14, 240–255	10.6.		12, 238–239
10.7.			10.8.		
10.9.		14, 277–278	10.10.		11, 149–158
					12, 229–231
10.11.			10.12.		
11.1.		13, 186			14, 278
11.2.		13, 186–189			
11.3.			11.4.		
11.5.			11.6.		14, 279
11.7.			11.8.		
11.9.			11.10.		
11.11.			11.12.		
12.1.		15, 236–237	12.2.		

Задача	Исправления	Решение(-я)	Задача	Исправления	Решение(-я)
12.3.			12.4.		14, 279–280
12.5.			12.6.		
12.7.			12.8.		14, 256–269
12.9.			12.10.		13, 189–190
12.11.			12.12.		
13.1.		14, 280–281	13.2.		14, 270–271
					14, 281
13.3.		15, 237	13.4.		
13.5.			13.6.		
13.7.			13.8.		
13.9.			13.10.		
13.11.			13.12.		
14.1.			14.2.		
14.3.		15, 237–239	14.4.		
14.5.		15, 239–240	14.6.		15, 240–241
14.7.	15, 235	15, 241–244	14.8.	15, 235	15, 219–228
14.9.			14.10.		
14.11.			14.12.		
15.1.			15.2.		
15.3.			15.4.		
15.5.			15.6.		
15.7.			15.8.		
15.9.			15.10.		
15.11.			15.12.		

Опечатки, замеченные в №14

Страница, Строка	Напечатано	Следует читать
273, 6 снизу	[4]	[1]
273, 10 снизу	[4]	[1]
285, 21 сверху	67	6–7

Подготовка оригинал-макета: LaTeX2ε,
METAPOST, М. Н. Вялый

Издательство Московского Центра
непрерывного математического
образования
119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241 74 83

Отпечатано с готовых диапозитивов в ППП «Типография „Наука“».
121099, Москва, Шубинский пер., д. 6.
Подписано в печать 07.02.11. Формат 70×100/16. Бумага офсетная. Печать офсетная. Печ. л. 15,5. Тираж 1000 экз. Заказ №