МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ

Третья серия

выпуск 23

Москва
Издательство МЦНМО
2019

УДК 51.009
ББК 22.1
М34

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

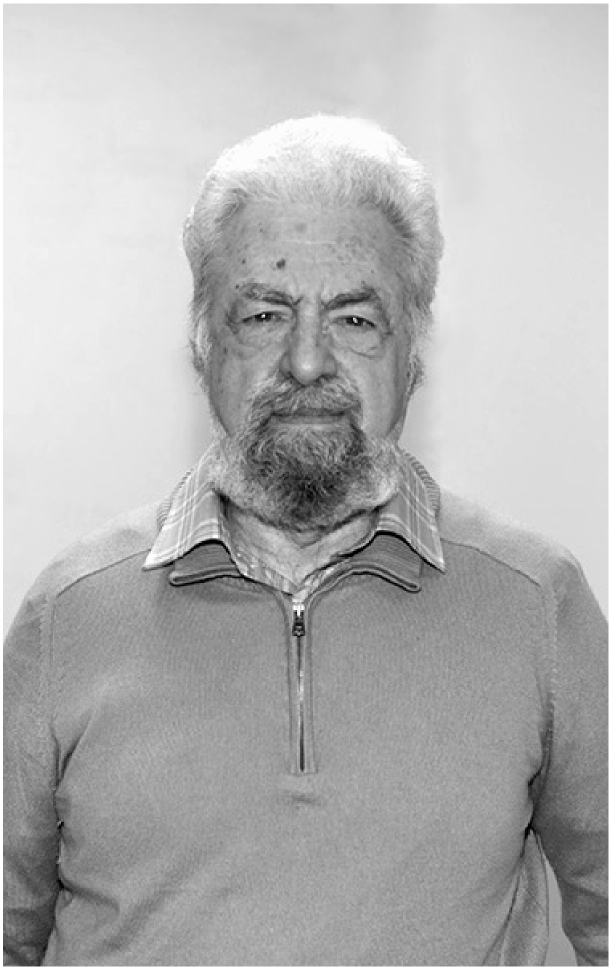
М34 **Математическое просвещение.** Третья серия, вып. 23. — М.: МЦНМО, 2019. — 240 с.
ISBN 978-5-4439-1368-1

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

УДК 51.009
ББК 22.1

12+

ISBN 978-5-4439-1368-1



Поздравляем
Юлия Сергеевича Ильяшенко
с 75-летием!

Содержание

Наш семинар: математические сюжеты

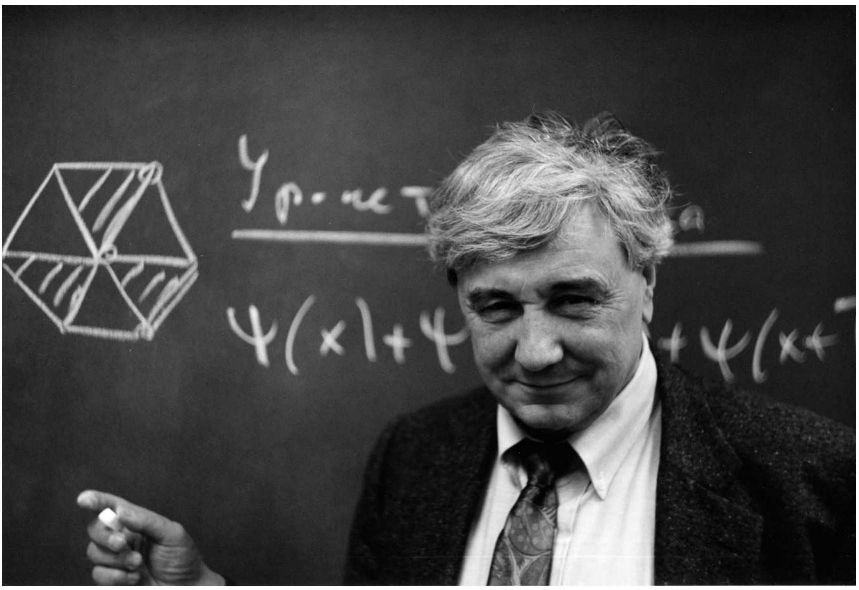
Нам пишут

Задачник (составитель А. Я. Канель-Белов)

Математический мир

Рассказ о Сергее Петровиче Новикове

И. А. Дынников, Д. В. Миллионщиков



20 марта 2018 года исполнилось 80 лет выдающемуся математику современности академику Сергею Петровичу Новикову. Он родился в семье знаменитых математиков: его отец, академик Пётр Сергеевич Новиков, был крупнейшим специалистом в области теории множеств, математической логики, теории алгоритмов и теории групп, а его мать, профессор Людмила Всеволодовна Келдыш, известна в математическом мире своими работами по геометрической топологии. Не забудем также напомнить, что

дед Сергея Петровича по матери Всеволод Михайлович Келдыш всеми отечественными энциклопедиями отмечен как создатель русского железобетона, а дядя, родной брат матери, Мстислав Всеволодович Келдыш был крупным специалистом в области прикладной математики и долгие годы возглавлял Академию наук СССР.

Неудивительно, что Сергей Петрович Новиков с самого раннего детства был знаком со многими замечательными учёными. Среди них были не только математики. В своих воспоминаниях о детстве Сергей Петрович тепло рассказывает о Михаиле Александровиче Леонтовиче — выдающемся физике, академике АН СССР. Леонтович учился с родителями Сергея Петровича Новикова на одном курсе физико-математического факультета МГУ. Сергей Петрович пишет и о других друзьях семьи Новиковых — Келдышей, среди которых были академические семьи Андроновых, Таммов, Ландсбергов, Парийских... Взрослые и дети этого большого сообщества вместе ходили в лодочные походы, в музеи и, конечно же, ходили друг к другу в гости. Часто собирались у Новиковых. Еда у Новиковых, по воспоминаниям гостей, была обычно очень простой, чаще всего на стол перед гостями ставили огромную кастрюлю с гречневой кашей, а самым интересным и важным за столом был, конечно же, разговор.

У Новиковых говорили и о науке, и об искусстве, и о литературе. Серёжа Новиков был внимательным слушателем этих бесед с самого раннего детства (одному из авторов этой статьи доводилось слышать об этом сразу от нескольких общих знакомых). Весьма вероятно, что именно удивительная широта тем для разговоров, высочайший культурный уровень собеседников, их талант выделить среди научных результатов и открытий только самые значимые и интересные, а также удивительная собственная память — всё это могло определить в Сергее Петровиче Новикове и необычайную широту его собственных научных и культурных интересов, и ту особую, свойственную только ему, принципиальность в науке. Несмотря на яркое раннее проявление математического таланта, широта интересов Сергея Новикова вплоть до последнего класса школы задержала выбор математики как будущей профессии.

В 1955 году Сергей Новиков поступает на механико-математический факультет МГУ. От родителей он знает о замечательной мехматской традиции добровольных учебных семинаров, где студенты решают разные нестандартные задачи, разбирают интересные темы, которых нет в обязательной программе. С первого же курса Сергей Новиков начинает посещать такой семинар для младшекурсников, которым руководил тогда Владимир Андреевич Успенский, сам в те годы ещё совсем молодой сотрудник МГУ. На этом семинаре первокурсник Сергей Новиков изучает основы алгебры, логики, теории множеств, теории функций. В те годы уже на втором курсе каждый студент мехмата МГУ должен был выбрать своего первого научного руководителя. Только написав в конце второго курса свою пробную курсовую работу, студент окончательно выбирал себе специализацию, а значит, научного руководителя и кафедру.

Второкурсника Сергея Новикова привлекает объявление о семинаре М. М. Постникова, В. Г. Болтянского и А. С. Шварца. В объявлении говорилось, что появилась новая увлекательная наука — современная алгебраическая топология и что «её не надо путать с такой ерундой, как общая топология» (последняя фраза была явной фрондой в адрес Павла Сергеевича Александрова). Молодой профессор М. М. Постников был уже хорошо известен своими работами и энциклопедическими познаниями почти во всех разделах алгебры, геометрии и топологии. Выбрав себе в качестве научного руководителя Михаила Михайловича, Сергей Петрович жадно впитывает новые математические идеи, возникающие в быстро развивающейся тогда топологии. Спустя годы он скажет в интервью В. М. Бухштаберу: «Мне помогли блестящие лекции Альберта Соломоновича Шварца и прекрасно написанные работы Жан-Пьера Серра» [1]. Именно в то время появляются замечательно выполненные М. М. Постниковым и Е. Б. Дынкиным переводы статей математиков французской школы: Серра, Тома, А. Картана, А. Бореля. Огромное влияние на Новикова на том этапе оказали и статьи английского математика Фрэнка Адамса.

На четвёртом курсе Сергей Петрович неожиданно остаётся фактически без научного руководителя — Михаил Михайлович Постников уезжает на долгий срок в Китай. Возникает знаменитый студенческий семинар «без научных руководителей», среди активных участников которого, кроме Сергея Петровича, мы видим Д. В. Аносова, Д. Б. Фукса, Г. Н. Тюрину, А. М. Виноградова и др. Сергей Петрович сказал позднее об этом периоде: «Именно в тот год я и сделал свои первые научные работы» [1]. Первые статьи Новикова относились к приложениям алгебр Хопфа к вычислению гомотопических групп сфер и комплексов Тома. Какая интересная получилась «точка входа в науку» у Сергея Петровича Новикова, как рано он обретает творческую самостоятельность как учёный! Однако сам Сергей Петрович в воспоминаниях даёт абсолютно иную оценку тем событиям: «Свою первую статью я опубликовал в 21 год. В то время я не был „молодым“: у таких людей, как, например, Арнольд, публикации появлялись уже в 18–19 лет, и это не считалось чем-то необычным. Я вырос в семье математиков, и моя мать жаловалась, что у всех уже есть научные публикации, кроме её сына» [1]. Под влиянием выдающихся западных математиков, звёзд мировой топологии Милнора, Хирцебруха, Атьи, Смейла, приезжавших в то время в Советский Союз, Новиков начинает заниматься теорией дифференцируемых многообразий и характеристических классов. После открытия Милнором в 1956 году различных гладких структур на семимерной сфере, стало ясно, что между понятиями топологического и гладкого многообразия имеется существенная разница (которая, однако, ещё не проявляется в размерности три). Естественно возникает вопрос: какие из инвариантов гладких многообразий зависят на самом деле только от топологии или даже только от гомотопического типа? Например, характеристические классы Штифеля — Уитни, в определении которых исходно используется гладкая структура, оказались инвариантами лишь гомотопического типа.

С классами Понтрягина ситуация оказалась гораздо сложнее: они не являются, вообще говоря, топологическими инвариантами. Однако Хирцебрух построил многочлены, которые для многообразий, размерность которых делится на четыре, выражают их сигнатуру через классы Понтрягина. Но сигнатура — это не только топологический, но и гомотопический инвариант! Используя этот факт, Новиков доказал в 1965 году, что классы Понтрягина будут топологическими инвариантами, если отбросить кручение, т. е. ограничиться рассмотрением рациональных классов. Это было одним из крупнейших достижений топологии того времени. Новиков доказал также гомотопическую инвариантность специальных интегралов Понтрягина — Хирцебруха по циклам, возникающим из гомологической алгебры фундаментальной группы, что привело его к знаменитой гипотезе о высших сигнатурах. Эта гипотеза Новикова оказала и продолжает оказывать большое влияние на развитие этой важнейшей области топологии вплоть до наших дней.

Ряд крупных достижений Новикова в топологии относятся к тому же периоду. Среди них полное вычисление кольца комплексных кобордизмов, классификация замкнутых односвязных многообразий размерности $n > 4$ с данным касательным гомотопическим типом, доказательство нестягиваемости группы диффеоморфизмов сферы (начиная с размерности 7) на ортогональную группу, развитие мощных алгебраических методов с использованием экстраординарных теорий когомологий, с помощью которых удалось, в частности, значительно продвинуться в вычислении стабильных гомотопических групп сфер. За полным списком результатов Новикова, который мы не в силах здесь охватить, советуем читателю обратиться к книге [1].

Первое, что сразу обращает на себя внимание при знакомстве со списком научных трудов С. П. Новикова — это необычайная широта тематики. Это и уже упомянутые глубокие проблемы алгебраической топологии, и различные аспекты геометрии, теории интегрируемых систем и теоретической физики. Однако область научных интересов Сергея Петровича гораздо шире. Например, математическая логика далека от основных интересов Сергея Петровича, но он хорошо знаком с теми её разделами, которыми занимался его отец, Пётр Сергеевич, а также ученик его отца и близкий друг Сергея Петровича академик Сергей Иванович Адян.

Одну работу в области математической логики сделал и сам Сергей Петрович. В 1962 году он доказал алгоритмическую нераспознаваемость класса гомеоморфизма сферы размерности пять. Один только этот замечательный результат мог бы сделать имя иному математику, но ко вкладу Сергея Петровича в топологию это уже тогда было лишь небольшое дополнение (которое, к слову, даже не было своевременно опубликовано). Эта работа замечательна тем, что в ней отчётливо виден фирменный почерк Новикова: сам результат лежит на пересечении двух казалось бы далёких областей — топологии и математической логики, а в основе доказательства спрятан хитроумный алгебраический трюк. Этот пример замечательно отражает стиль работы Сергея Петровича как истинного учёного: занимаясь любой темой, он не только продвигается вглубь, но и всегда как бы осматривается по сторонам в поисках точек соприкосновения с другими областями.

Сергею Петровичу всегда тесно в рамках какой-то одной области; этим он ярко выделяется на фоне многих математиков, которые зачастую посвящают всю свою научную жизнь какой-то одной узкой теме и пишут по ней десятки работ, каждая из которых на эпсилон улучшает предыдущую. Результаты Сергея Петровича — это, как правило, не только решения крупных проблем, но и открытие совершенно неожиданных связей между ранее далёкими друг от друга областями и создание на этой основе абсолютно новых направлений современной математики. Это хорошо иллюстрируют работы Новикова 1970-х годов. Тогда, уже будучи членом-корреспондентом Академии наук СССР, лауреатом Ленинской премии, обладателем золотой Филдсовской медали за выдающиеся достижения в алгебраической топологии, он, неожиданно для многих, переходит на работу в Институт теоретической физики имени Л. Д. Ландау. Новикова стали интересовать задачи, имеющие физическое происхождение, в частности, нелинейные уравнения математической физики.

Истины ради мы должны тут отметить, что первой собственно физической работой Сергея Петровича стала статья по общей теории относительности, написанная им в соавторстве с О. И. Богоявленским. Новиков занимался теорией однородных космологических моделей, изучая пространство однородных решений уравнения Эйнштейна. В этих исследованиях крайне полезными оказались его знания в теории динамических систем и особенно в теории странных аттракторов.

Незадолго до перехода Новикова на работу в Институт им. Ландау у физиков стало набирать популярность изучение уравнения Кортевега —

де Фриза (КдФ), ставшее к сегодняшнему дню поистине знаменитым. Изначально оно использовалось только для описания волн на мелкой воде, но в 1960-х годах было обнаружено, что спектр его применения намного шире. В работе Крускала, Грина, Гарднера и Миуры было открыто замечательное свойство уравнения КдФ — возможность его интегрирования методом обратной задачи рассеяния, а Лакс открыл коммутаторную форму записи этого уравнения, за которой впоследствии закрепилось название пара Лакса. Вскоре появились работы Л. Д. Фаддеева, В. Е. Захарова и А. Б. Шабата, открывшие другие важные аспекты теории уравнения КдФ и его обобщений. Внимание специалистов было сосредоточено на физически наиболее интересном быстроубывающем случае, в котором метод обратной задачи приводит к линеаризации динамики после перехода к данным рассеяния.

В своей новаторской работе 1974 года Новиков рассмотрел периодическую задачу для уравнения КдФ и привнёс в эту теорию поистине революционную идею: спектр оператора Штурма — Лиувилля в периодическом случае правильно рассматривать как риманову поверхность, так называемую спектральную кривую, а собственную функцию оператора Штурма — Лиувилля задавать как мероморфную функцию на спектральной кривой, зависящую от параметра. Этот подход позволяет строить массу точных решений, если спектральная кривая имеет конечный род как двумерная ориентированная поверхность. Специалисты называют эту ситуацию случаем так называемых конечнозонных потенциалов. Динамика уравнения превращается в динамику полюсов так называемой функции Бейкера — Ахиезера на спектральной кривой, а динамика образа дивизора полюсов на якобиане спектральной кривой (комплексный тор, построенный специальным образом по спектральной кривой) становится линейной. Таким способом можно получить большое семейство условно периодических решений, явно выражающихся в тета-функциях, которые в пределе дают также и быстроубывающие решения.

Эти идеи явного взаимодействия алгебраической геометрии и спектральной теории дифференциальных операторов породили целую большую теорию «методов конечнозонного интегрирования», которая может быть применена для других уравнений, в частности для уравнения Кадомцева — Петвиашвили (сокращённо — уравнение КП). Над этой теорией Сергей Петрович стал работать вместе со своими учениками Борисом Дубровиным и Игорем Кричевером. Теория в данном контексте играет совершенно особую роль; рождённая как приложение алгебраической геометрии к нелинейным уравнениям, она привела к решению проблемы алгебраической геометрии, известной с конца XIX века, — проблемы Римана — Шоттки о характеризации якобианов комплексных кривых. Дело в том, что формулы для решения уравнений КП, получаемые методом конечнозонного интегрирования, включают лишь матрицу периодов спектральной кривой, но отнюдь не произвольная матрица даёт решение. Новиков выдвинул гипотезу, что матрицы, приводящие к решениям уравнений КП, — это в точности матрицы периодов голоморфных дифференциалов комплексных кривых. Эта гипотеза была окончательно доказана в 1986 году Шиотой, после серьёзных продвижений Бориса Дубровина в направлении её доказательства.

Упомянутая нами гипотеза, вторая по счёту в нашем рассказе (но исторически не последняя!), оказала большое влияние на развитие важной области математики во второй половине XX века. Напомним, что первая, более ранняя, гипотеза Новикова о высших сигнатурах до сих пор является открытой проблемой и также во многом определила развитие целой отрасли топологии.

В начале 80-х Новиков стал одним из создателей ещё одного большого научного направления — аналога теории Морса для многозначных функционалов. Параллельно на Западе родственные идеи развивались несколько в ином ключе в работах Виттена и Флоера и их последователей. Наблюдение Новикова, давшее начало этой теории, состояло в том, что даже несложная механическая система после гамильтоновой редукции может перестать быть лагранжевой в традиционном понимании, но остаётся таковой, если обобщить постановку вариационной задачи, разрешая функционалу действия быть многозначным, имея однозначной лишь первую вариацию. Иначе говоря, лагранжиан системы следует рассматривать как замкнутую, но не обязательно точную 1-форму. Как оказалось, такие функционалы естественным образом возникают в самых разных задачах, а обобщение на них идей обычной теории Морса приводит к новым глубоким результатам.

В этом снова проявился один из удивительных талантов Новикова — увидеть в простом и общеизвестном примере нечто просто формулируемое, из которого потом, как дерево из семени, вырастает быстро развивающаяся область науки с многочисленными ответвлениями. Одна из ветвей этого дерева, в развитии которой довелось участвовать одному из авторов этого рассказа, нашла своё применение в реальной физике, а именно, в теории проводимости в нормальных металлах, а математические её аспекты оказались теснейшим образом связаны с теорией динамических систем.

У Сергея Петровича есть удивительное качество: он в состоянии понять абсолютно любой математический результат, и не просто понять, а мгновенно оценить его «нетривиальность», его уровень сложности и важности, связать его с другими областями.

Сергей Петрович — яркий пример «неудобного слушателя»: на его семинаре докладчику невозможно предугадать его вопросы. Новиков способен своими неожиданными и глубокими репликами раскрыть новые связи обсуждаемой темы, но также способен и полностью поменять исходный план докладчика.

Стремление к полной ясности в понимании и изложении даже самых сложных теорем — одна из самых ярких и привлекательных черт Сергея Петровича как математика. «Я мыслю конкретными примерами, мне не нужно пустых абстрактных теорем» — любимая фраза Новикова.

Перейдём к последней части нашего рассказа о Сергее Петровиче Новикове. Речь пойдёт о Новикове-педагоге.

Проблемы образования вообще и в первую очередь математического образования в стране и в мире всегда волновали и волнуют Сергея Петровича. В 1973 году в журнале «Природа» вышла статья совсем молодого члена-корреспондента АН СССР С. П. Новикова «Необходима перестройка математического образования». Статья написана в рамках полемики со статьёй американского математика Н. Спона (мл.), опубликованной в том же номере журнала. В своей статье Новиков лишь обозначает своё видение места, которое должна занимать математика в структуре человеческого знания. Он делает особый акцент именно на модернизации существовавшего тогда математического образования, направленной на взаимодействие с другими естественными науками.

Пройдёт совсем немного времени, и молодой профессор Сергей Новиков вместе со своими друзьями и единомышленниками создаст экспериментальный поток на мехмате МГУ. История, опыт этого экспериментального потока важны и интересны и в наши дни. Многие математики сейчас стали говорить о необходимости его воссоздания на мехмате. В ответ на это предложение звучат такие критические слова: «нельзя создавать элитарность и замкнутость внутри математического образования». Любопытно посмотреть, что говорится в письме, написанном С. П. Новиковым и В. А. Зоричем декану мехмата. В этом письме излагается основная идея и программа экспериментального потока. В самом начале мы встречаем фразу: «Отбор студентов в этот поток мог бы производиться с середины первого курса (или с начала второго) на нормальной объективной основе, исключающей отбор „элиты“, состоящей лишь из наиболее способных». С. П. Новиков и В. А. Зорич специально подчёркивали, что «организация учебной работы должна носить межкафедральный характер; детализированные программы должны быть выработаны всем коллективом совместно, учитывая максимальное взаимодействие предметов — не только математических, но и естественнонаучных, которые должны начинаться не позднее начала второго курса».

Конечно же, созданием экспериментального потока на мехмате МГУ педагогическая и научно-просветительская деятельность Сергея Петровича не ограничивалась. За год до создания экспериментального потока он в корне модернизировал курс дифференциальной геометрии для студентов-механиков на мехмате МГУ. Именно по его программе, с минимальными вариациями, профессора кафедры высшей геометрии и топологии мехмата МГУ читают этот курс вплоть до наших дней.

Одной из вершин научно-педагогической деятельности Новикова стало создание учебника-энциклопедии «Современная геометрия» (в соавторстве с Б. А. Дубровиным и А. Т. Фоменко). Этот трёхтомник переведён на множество иностранных языков, выдержал множество переизданий, и авторы могут лично свидетельствовать, что встречали «Современную геометрию» на книжной полке в рабочих кабинетах у всех знакомых геометров и топологов, у математиков других специальностей, а также у многих физиков по всему миру. Большой популярностью пользуется изданная в более поздние годы монография-учебник «Геометрические структуры и поля» С. П. Новикова и И. А. Тайманова.

Сергей Петрович Новиков — создатель крупнейшей научной школы. Под его руководством более 40 учеников защитили кандидатские диссертации. Затем стали и докторами наук: И. К. Бабенко, Ф. А. Богомолов, О. И. Богоявленский, М. А. Бродский, В. М. Бухштабер (член-корреспондент РАН), В. В. Веденяпин, А. П. Веселов, И. А. Володин, В. Л. Голо, П. Г. Гриневич, С. М. Гусейн-Заде, Б. А. Дубровин, И. А. Дынников, Г. Г. Каспаров, И. М. Кричевер, А. Я. Мальцев, А. С. Мищенко, О. И. Мохов, О. Р. Мусин, Р. Г. Надирадзе, Р. Г. Новиков, И. А. Тайманов (академик РАН), С. П. Царев. Под руководством Новикова защитили кандидатские диссертации также Л. А. Алания, А. Л. Брахман, Ф. Ф. Воронов, Г. С. Григорян, А. Джиакоббе, Р. Делео, А. В. Зорич, А. Ю. Лазарев, Т. К. Т. Ле, А. С. Лыскова, Д. В. Миллионщиков, М. В. Павлов, Н. В. Панов, Г. В. Потёмин, С. В. Пиунихин, А. А. Пересецкий, В. А. Садов.

Авторы этого рассказа не могут не поделиться интересными деталями, относящимися к тому периоду, когда они сами стали заниматься математикой под руководством Сергея Петровича. В те годы студент мехмата МГУ, записавшийся весной второго курса к Сергею Петровичу Новикову, обычно получал от него задание на лето: прочитать и освоить книгу Милнора «Теория Морса». Вся дальнейшая подготовка ученика строилась вокруг «Современной геометрии». В какой-то момент каждый ученик Новикова получал первую книгу из серии «Ландау — Лифшиц» — это была «Теоретическая механика». В аспирантуре все ученики Новикова должны были сдать ему экзамен по ещё одной книге Л. Д. Ландау и Е. М. Лифшица — по «Теории поля».

Обычно экзамен происходил в старом здании Института им. В. А. Стеклова. Сергей Петрович сидел в кабинете, в котором всегда было много народу, а аспиранты покорно ждали свободной минуты в коридоре. Дверь была всегда открыта. «Что там у Вас?» — спрашивал неожиданно Сергей Петрович, заметив в глубине коридора прижавшегося к стене аспиранта. «Отчёт... Теория поля», — бормотал аспирант. «Очень хорошо, — быстро говорил Сергей Петрович. — Это очень правильно. Прекрасная книга! Все мои ученики должны её освоить! Вот давайте, идите-ка и выведите формулу для метрики Шварцшильда!» «Олег! — обращался он к сидевшему в его кабинете Олегу Игоревичу Богоявленскому (одному из учеников более старшего поколения) — Олег, посмотрите формулу, когда он напишет».

В 80-е годы прошлого века на каждом году обучения у Сергея Петровича было по несколько учеников. Всего, вместе с аспирантами, набиралось до пятнадцати и даже больше человек. Темы курсовых работ в те годы не сильно различались у разных студентов. Существовало всего два варианта: «Некоторые проблемы топологии» и «Некоторые проблемы геометрии». Но конкретную задачу, над которой надо было работать, студент получал далеко не сразу. «Вы должны освоить...» — далее шёл список параграфов из «Современной геометрии» и других книг, которые надо было освоить и сдать Сергею Петровичу, прежде чем он решал, что наступил момент для формулировки задачи для курсовой работы. День, когда Сергей Петрович формулировал задачу своему новому ученику, наступал зачастую неожиданно: «Вот послушайте! — говорил он после семинара, обращаясь по фамилии к новому ученику, — вот есть такая хорошая задача!» Дальше он быстро и чётко формулировал задание. Сотоварищи счастливчика, удостоившегося чести получить задачу от Шефа, в этот момент с завистью смотрели на него.

90-е годы прошлого века сильно изменили судьбы многих отечественных учёных. Сергей Петрович Новиков, как и многие ведущие математики из бывшего Советского Союза, стал работать за границей — в университете Мэриленда (США). Но при этом Новиков ни в какой момент не терял связь с российской математикой: его семинар как активно работал, так и работает до сих пор 4–5 месяцев в году. Сергей Петрович руководит журналом «Успехи математических наук», имеющим самый высокий импакт-фактор среди российских математических журналов, в чём мы видим его прямую заслугу. Кафедра высшей геометрии и топологии, которой заведует Сергей Петрович Новиков, по-прежнему одна из ведущих кафедр механико-математического факультета МГУ, на ней по-прежнему учатся многие из сильнейших студентов мехмата. Руководимый им отдел геометрии и топологии Математического института им. В. А. Стеклова РАН имеет репутацию мирового уровня.

Мэриленд, конечно же, оказал влияние на Сергея Петровича, расширил его понимание современной мировой математики, американский опыт добавил новые грани его педагогическому таланту. Здесь мы хотим привести один любопытный рассказ Карстена Грова, коллеги Сергея Петровича по университету Мэриленда.

> Когда Новиков впервые появился на семинаре нашего отдела в Мэриленде, он сразу удивил участников своими вопросами: их было всегда много, и он задавал их сразу по ходу доклада. Это было удивительно, ведь обычно у нас задают вопросы после доклада, да и самих вопросов бывает не так много. А тут Новиков на семинарах стал спрашивать докладчиков буквально с самого начала. Многие удивлялись, ведь вопросы Новикова были часто простыми, и некоторым из нас стало казаться, что Новиков просто не понимает содержание доклада, раз задаёт столь элементарные вопросы. Но прошло какое-то время, прежде чем все мы осознали, что Новиков спрашивает много именно потому, что он ВСЁ понимает и стремится к полному пониманию каждого доклада! Многие из нас привыкли просто вежливо сидеть во время докладов, далёких от собственной научной темы, и не задавать вопросов, а тут мы увидели яркий живой ум математика, который всегда стремится дойти до самой сути.

Вернёмся к педагогической деятельности Сергея Петровича Новикова в России последних десятилетий. Сергей Петрович прочитал много лекций на летних школах в Дубне, а также на летних школах по геометрическим методам математической физики, которые проводила его кафедра высшей геометрии и топологии мехмата МГУ. Мы хотим предложить читателям пару небольших цитат из лекций и выступлений последнего десятилетия, обращённых к совсем юным математикам — студентам и школьникам.

> Геометрия и топология — это замечательные дисциплины, у них есть только одна сложность: когда вы занимаетесь естественными науками, те приложения геометрии и топологии, о которых вам говорили профессора в процессе вашего обучения, эти приложения не проявляются так ярко, как бы вам думалось, и приложения, как правило, бывают скрыты. Для того чтобы раскрыть эти приложения, вам требуются дополнительные знания, большие знания. Например, законы природы мы привыкли описывать дифференциальными уравнениями, а где же в этом месте искать геометрию и топологию? Чаще мы обнаруживаем следы алгебры.
>
> Чтобы осознать красоту и важность таких областей математики, как теория динамических систем, топология, алгебраическая геометрия, вам нужно очень много знать, особенно, если вы хотите осознать проявления этих новых областей в естественных науках.

Сергей Петрович Новиков является блестящим знатоком мировой истории, особенно впечатляет список первоисточников, которые он прочитал и с лёгкостью цитирует. Глубина его познаний в этой области поражает. Конечно, это во многом объясняется традициями семьи. Гармоничность развития и образования детей — одна из важнейших традиций в семье Новиковых-Келдышей. Сергей Петрович не очень ценит исторические романы. «Я предпочитаю изучать только подлинные исторические хроники и документы», — часто говорит он.

Сергей Петрович Новиков — гармонически развитая личность в самом античном смысле: к его математическим и гуманитарным талантам надо добавить и его разностороннее спортивное развитие: он прекрасно играет в настольный теннис, отличный пловец, друзья его молодости любят рассказывать о его сильных качествах как футболиста. Сергей Петрович любит длинные пешие прогулки.

Сергей Петрович — прекрасный семьянин. Несколько лет назад он и его супруга, Элеонора Викентьевна, отметили золотую свадьбу. У них трое детей. Все дети получили классическое физико-математическое образование, живут и работают в Москве: дочери Ирина и Мария — выпускницы МГУ и кандидаты физико-математических наук, сын Пётр закончил Московский физико-технический университет (Физтех). Подрастают многочисленные внуки Сергея Петровича.

Завершить наш рассказ мы хотим перечислением высших математических наград Сергея Петровича Новикова:

- премия Московского математического общества для молодых математиков (1964);
- Ленинская премия (1967) — за цикл работ по дифференцируемым многообразиям;
- Филдсовская медаль (1970) — за работы по топологии;
- премия имени Н. И. Лобачевского (1980) — за цикл работ по теории слоений;
- премия Вольфа (2005, Израиль) — за фундаментальный и новаторский вклад в алгебраическую и дифференциальную топологию и в математическую физику (в частности, за введение алгебро-геометрических методов);
- премия имени А. В. Погорелова НАН Украины (2008) — за цикл работ «Современные методы геометрии и топологии и их применение»;
- золотая медаль Н. Н. Боголюбова РАН (2009) — за выдающиеся результаты в области математики, теоретической физики и механики;
- золотая медаль имени Леонарда Эйлера РАН (2012) — за глубокий вклад в применение топологических методов в квантовой физике.

Список литературы

[1] Сергей Петрович Новиков. К семидесятилетию со дня рождения. Интервью, статьи, выступления / Под ред. В. М. Бухштабера. М.: МЦНМО, 2008.

[2] *Новиков С. П.* Необходима перестройка математического образования // Природа. 1973. № 2. С. 57.

[3] Сергей Петрович Новиков (к 75-летию со дня рождения) // УМН. 2013. Т. 68, вып. 3. С. 3.

Иван Алексеевич Дынников, МИАН им. В. А. Стеклова
dynnikov@mech.math.msu.su

Дмитрий Владимирович Миллионщиков, мехмат МГУ
mitia_m@hotmail.com

Елена Александровна Морозова

В. М. Тихомиров

15 мая 2018 года исполнилось 90 лет Елене Александровне Морозовой.

Она верно служит математике и преподаванию, своему механико-математическому факультету и Московскому университету, и это её служение продолжается свыше семидесяти лет.

Вот как всё для неё начиналось. Школьница Лёля Морозова пошла на лекцию по геометрии Лобачевского, которую читал в большой аудитории Политехнического музея профессор мехмата МГУ, тогда ещё не академик, а член-корреспондент АН СССР, Павел Сергеевич Александров. Аудитория была переполнена, сидели даже на всех ступеньках. Ассистировал Павлу Сергеевичу студент третьего курса мехмата МГУ Олег Локуциевский: он доходчиво демонстрировал геометрию Лобачевского на граммофонной трубе. От него слушатели узнали, что в университете возобновилось чтение воскресных лекций по математике для школьников и работа школьных математических кружков.

...Чтение лекций для школьников и организация школьных математических кружков на мехмате начались с 1934 года. Сначала кружки были не очень эффективны: школьникам давали прочитать литературу и сделать доклад по прочитанному. Это не слишком удавалось. Новый импульс к проведению кружков был получен от юноши, которого звали Додик Шклярский. Он первым стал на своём кружке решать задачи и составил список ярких и содержательных задач для своего школьного кружка. Всё дальнейшее восьмидесятилетнее развитие кружков, цель которых — поиск математических талантов, следует пути, намеченному Шклярским. На первой же олимпиаде, в которой участвовали школьники из кружка Шклярского, все первые премии были за ними. В начале войны Додик был направлен в отряд особого назначения, который действовал в тылу врага. Давид Оскарович Шклярский с войны не вернулся. Но его ученики продолжили дело своего учителя.

Лёля пошла в кружок, который вели участники кружка Шклярского — братья Ягломы Ися и Кика (Исаак Моисеевич и Акива Моисеевич) и ещё Оля (Ольга Александровна) Ладыженская, в будущем академик АН СССР. В кружке рассказывалось о различных геометрических преобразованиях и элементах топологии. Там решались интересные задачи.

Лёля решила поступать на мехмат. При поступлении возникло некоторое затруднение, связанное с тем, что она заболела во время экзаменов. Но все трудности помог преодолеть Андрей Николаевич Колмогоров. И в итоге Лёля Морозова оказалась на мехмате, который стал для неё не местом службы, а родным домом, где жили и работали не коллеги, а близкие люди, которые помогли ей выбрать путь в жизни, поддерживали её, а многие из них нуждались в помощи и защите, которую они получали от неё.

Лёля, поступив на мехмат, почти сразу же стала вести школьные кружки.

Автор этих строк впервые попал на воскресную лекцию на мехмате в четырнадцатилетнем возрасте. После лекции руководители стали агитировать за свои кружки. Сначала выступали какие-то взрослые люди, потом пошли студенты старших курсов, а завершали совсем ещё юные, почти дети, и в этой группе среди руководителей кружков появились девушки. Одна девушка выступала в паре с высоким худым мальчиком с несколько необычным высоким тембром голоса. Я запомнил, как звали мальчика: Коля Ченцов. Через некоторое время я узнал имя соруководительницы этого кружка. Её звали Лёля Морозова. Всё это произошло осенью 1948 года, семьдесят лет тому назад.

В кружке Лёли и Коли получали напутствие к будущей жизни многие выдающиеся математики. Среди них мои друзья и добрые знакомые — Владимир Михайлович Алексеев, Николай Сергеевич Бахвалов, Александр Дмитриевич Вентцель, Эрнест Борисович Винберг, Игорь Владимирович

Гирсанов, Алексей Валерьевич Забродин, Борис Теодорович Поляк, Юлиан Борисович Радвогин... Все они сохранили самые тёплые чувства к руководителям своих изначальных математических семинаров, а Елена Александровна всю последующую жизнь с гордостью вспоминает, как у неё с Колей занимались замечательные мальчики и девочки, ставшие её коллегами по профессии и возглавившие новые научные направления.

Более 25 лет Е. А. Морозова руководила школьными математическими кружками, организовывала юношеские математические школы, принимала участие в организации Всесоюзных математических олимпиад, была членом жюри Московских, Всесоюзных и Международных математических олимпиад. На IV–X Международных математических олимпиадах Е. А. Морозова была руководителем советской команды. По результатам этой деятельности написана и издана известная книга «Международные математические олимпиады» (первое издание 1967 г., затем переиздавалась в 1968, 1971 и 1976 гг.; I–III изд. — совместно с И. С. Петраковым, IV изд. — совместно с И. С. Петраковым и В. А. Скворцовым).

Ещё в школе Лёля полюбила геометрию, и узкой математической профессией Елены Александровны Морозовой стала геометрия в широком смысле этого слова. Закончив механико-математический факультет МГУ в 1951 г. и аспирантуру по кафедре высшей геометрии и топологии в 1954 г., Морозова защитила кандидатскую диссертацию в 1955 г. Её научным руководителем был профессор кафедры Сергей Владимирович Бахвалов. К областям научных интересов Е. А. Морозовой можно отнести геометрическую теорию тонких оболочек, вопросы изгибания поверхностей в целом.

Николай Николаевич Ченцов и Елена Александровна Морозова образовали замечательную творческую пару. Николай Николаевич был математиком очень широкого диапазона. Он, в частности, решил многое переосмыслить в теории вероятностей и математической статистике. Возникшая в конце сороковых годов теория информации позволила рассматривать статистику и некоммутативную теорию вероятностей как геометрические многообразия. Приложения различных геометрических идей и конструкций к статистике и некоммутативной теории вероятностей и стали предметом совместной деятельности Н. Н. Ченцова и Е. А. Морозовой.

Елена Александровна Морозова работает на геометрических кафедрах мехмата с 1951 года. Я не знаю никого, у кого педагогический стаж превышает стаж Морозовой. Все эти годы она ведёт семинарские занятия по курсам аналитической геометрии, линейной алгебры и геометрии, дифференциальной геометрии и топологии. Она читала лекции по аналитической геометрии на инженерном потоке, спецкурс по геометрической теории тонких оболочек, руководила курсовыми и дипломными работами.

Число бывших мехматян, которые с самыми тёплыми чувствами вспоминают семинары Е. А. Морозовой, огромно.

Однажды Елена Александровна выразила желание стать членом Дома учёных. В ту пору в Дом учёных принимали в основном лишь докторов наук, а кандидатов только в виде исключения. Нужна была рекомендация какого-то видного учёного. Я обратился к В. В. Козлову с просьбой подписать уже написанный мною текст. Но Валерий Васильевич сказал, что Елена Александровна была его преподавательницей и он напишет рекомендацию сам. И действительно, через пару дней мне передали текст, в котором В. В. Козлов нашёл замечательные возвышенные слова о Елене Александровне, и после получения этого текста директором Дома учёных вопрос о Е. А. Морозовой был немедленно решён.

Елена Александровна в течение многих лет (с 1960 по 1979 гг.) была заместителем П. С. Александрова как заведующего Отделением математики мехмата. Она стала доверенным лицом Павла Сергеевича в полном смысле этого слова. К ней замечательно относился и Андрей Николаевич Колмогоров. Это привело к тому, что в решении многих трудных проблем мехмата Елена Александровна играла очень значительную роль, осуществляя связь Колмогорова и Александрова с Иваном Георгиевичем Петровским. Елена Александровна всегда была на стороне правды и справедливости. После смерти Ивана Георгиевича Елена Александровна много раз в трудных случаях разрешала проблемы благодаря поддержке Виктора Антоновича Садовничего.

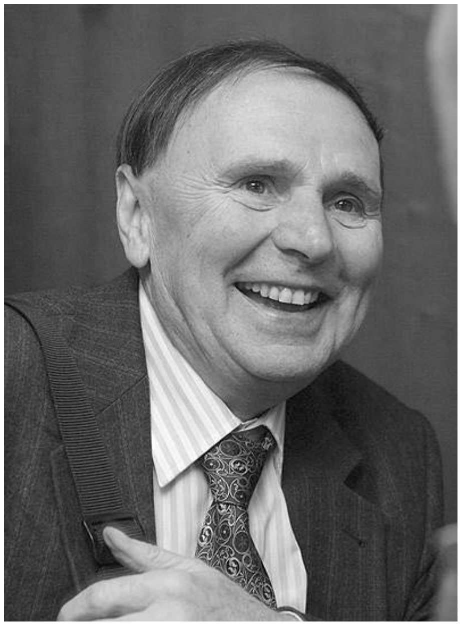
Вот два примера деятельности Елены Александровны по сохранению памяти о наших великих учителях. Павел Сергеевич Александров умер в 1982 году. Он завещал похоронить себя на поселковом кладбище невдалеке от Комаровки. Шли годы, но на могиле Александрова не было никакой мемориальной доски. И Елена Александровна сама поставила мемориальную доску! Конечно, факультет, усовестившись, оплатил её расходы. Несколько лет боролась Елена Александровна за то, чтобы вблизи Университета появилась улица Колмогорова. Здесь решающее содействие оказал ректор.

А сколько раз Е. А. Морозова вступала в борьбу за правое дело — не счесть!

Пожелаем Елене Александровне сохранить душевные силы для просвещения молодёжи, для творчества и благих дел.

Владимир Михайлович Тихомиров, мехмат МГУ
vmtikh@googlemail.com

Памяти
Андрея Анатольевича Зализняка



24 декабря 2017 года ушёл из жизни замечательный лингвист, академик РАН Андрей Анатольевич Зализняк. Его деятельность была тесно связана с математикой и информатикой. Ниже следуют статьи В. А. Успенского и А. А. Зализняка, посвящённые вручению ему литературной премии Александра Солженицына и ранее опубликованные в книге: *Зализняк А.* Из заметок о любительской лингвистике, М.: Русский мир, 2010, с. 194–212. Публикуем также статью А. Пиперски о научной деятельности А. А. Зализняка.

В. Успенский

О русском языке, о дешифровке древних текстов, о «Слове»

Три области названы в решении солженицынского жюри, три области, за достижения в коих Андрей Анатольевич Зализняк удостоен Солженицынской премии. Это русский язык. Это дешифровка древних текстов. Это «Слово о полку Игореве». Премия присуждена по совокупности. Однако достижения в каждой из этих областей столь значительны, что могли бы претендовать на самую высокую премию и по отдельности. Вот об этом позвольте поговорить.

Сперва о русском языке. Исследования Зализняка в этой области начались с его русско-французского словаря, вышедшего в 1961 г. Словарь предназначался для *франкоязычного* пользователя. Русский язык флективен, и это ставит перед составителем рассчитанного на иностранца русско-иностранного словаря непростую задачу: надлежит либо включить в русскую часть словаря все формы слова, что едва ли возможно практически, либо сопроводить словарь правилами русского словоизменения — что и сделал Зализняк. Он приложил к словарю свой первый шедевр — краткий очерк русского словоизменения, т. е. склонения и спряжения.

Очерк поражал отточенной логикой изложения. Зализняк установил основные схемы, по которым происходит русское словоизменение, придумал удобную индексацию этих схем и снабдил каждое словарное слово соответствующим индексом. Этот очерк затем помещался во многих русско-иностранных словарях, а из наработок, относящихся к склонению, родилась знаменитая монография 1967 г. «Русское именное словоизменение», вошедшая в золотой фонд русской и мировой лингвистики. Весной 1965 г. от одного из руководящих языковедов МГУ мне довелось услышать такой вопрос: «А что, до Зализняка не знали, как склонять русские слова?» Знали, конечно, но на уровне использования языка его носителями, а не на уровне исчерпывающего лингвистического описания. Полностью русское склонение было описано впервые именно Зализняком. Здесь существенно слово **«полностью»**: впервые было дано описание, не использующее слов «и так далее», многоточий и других подобных апелляций к аналогии. Параллельно шла титаническая работа над созданием «Грамматического словаря русского языка» — словаря, дающего для каждого (из более чем стотысячного списка!) русского слова все его формы. Работа продолжалась тринадцать лет и увенчалась выходом в 1977 г. первого издания словаря. Словарь сразу стал событием в русистике и сделался необходимым не только лингвистам, но и всем, использующим русский язык. В 2003 г. вышло его четвёртое издание. Сегодня пейзаж русистики немыслим без этого словаря и предшествующей ему монографии. «Посмотри у Зализняка» стало такой же формулой, как «посмотри у Даля».

Теперь о втором поприще Зализняка — о дешифровке древних текстов. Имеются в виду грамоты на берёсте, но не только они. Что здесь сделал Зализняк, если ограничиться главным?

Первое. Для большого числа грамот, ранее читавшихся неправильно, Зализняк установил их правильный смысл, что иногда имело важные последствия. Вот два примера в упрощённом пересказе. Фраза «посылаю щуку и клещи» давала основание для далеко идущих выводов о развитии кузнечного дела на Новгородчине и даже о близости рыбной и кузнечной слобод. Оказалось: ‘щуку и лещей’. Написание двьрикѣлѣ неправильно понималось как ‘двери кельи’. Зализняк разгадал правильный смысл: ‘двери целы’. Однако так называемая вторая палатализация требовала, чтобы в этой позиции звук *к* изменился в звук *ц*, чего, как показала грамота, не произошло. Предпринятое Зализняком тщательное исследование проблемы привело его к пониманию, что в диалекте древних новгородцев вообще не было второй палатализации, — а ведь наличие её во всём славянском мире считалось аксиомой!

Второе. Когда предшественники Зализняка встречали затруднение в осмыслении грамоты, они нередко объявляли писца неграмотным или описавшимся и заменяли одну букву другой. Зализняк первым догадался, что те, кто писал грамоты, писали именно то, что хотели, но только пользовались при этом особой, отличной от книжной, бытовой графической системой, имеющей свои собственные правила. Тем самым система письма берестяных грамот обрела своё законное место среди других древних письменностей — а ведь до исследований Зализняка совокупность записей на грамотах трактовалась как хаотическая и ни в какую систему не укладывающаяся. Фонд памятников, использующих указанную систему письма, позволил, далее, восстановить тот диалект, на котором общались между собой древние новгородцы.

Третье. Изучив живой, бытовой язык, на котором писались берестяные грамоты, Зализняк установил, что в древнерусском языке существовали два основных диалекта: северо-западный, на котором и говорили новгородцы, и юго-центро-восточный, на котором говорили все остальные восточные славяне.

Четвёртое. Согласно широко распространённому мнению различные языки и диалекты — и заведомо восточнославянские языки и диалекты —

образовались путём дивергенции, т. е. расхождения, расщепления, из некоего исходного языка или диалекта. Зализняк открыл, что в формировании того современного русского языка, на котором мы говорим, решающую роль сыграл процесс конвергенции, т. е. схождения, двух древних диалектов древнерусского языка — северо-западного и юго-центро-восточного.

Пятое. Зализняк явился основателем нового раздела палеографии, а именно палеографии берестяных грамот. Более того, он довёл этот раздел до большей разработанности, чем традиционная палеография древнерусских рукописей. Предпринятое Зализняком тщательное исследование всех найденных грамот позволило ему чётко выделить основные визуальные признаки грамоты, разделив эти признаки на пять групп:

1) алфавитные (Зализняк называет их графическими), определяемые тем набором графем (то есть букв алфавита и других знаков), который используется в грамоте;
2) палеографические, определяемые конкретным начертанием графем;
3) орфографические;
4) лингвистические — связанные с фонетическими, морфологическими, синтаксическими и лексическими особенностями текста;
5) формулярные — связанные с языковым этикетом.

Для каждого из признаков Зализняк провёл анализ встречаемости данного признака в тех грамотах, даты коих были установлены археологами методом так называемого *стратиграфического датирования*, даваемого глубиной залегания пластов, в которых грамоты были найдены. При этом им были выявлены две критические точки на хронологической шкале: точка, ранее которой признак не встречается, и точка, позже которой признак не встречается. Значимость этих точек недооценивается традиционной палеографией. Точки эти позволяют сразу отсечь некоторые признаки как нерелевантные — это те признаки, для которых критические точки находятся вне изучаемого исторического периода (для берестяных грамот это XI–XV века). А для каждого из релевантных признаков была определена частота его встречаемости в стратиграфически датированных грамотах. На основе критических точек и частот, установленных для различных признаков, Зализняком были составлены таблицы *внестратиграфического датирования*. Эти таблицы дают возможность датировать грамоту без апелляции к её археологической истории, на основе лишь её визуальных признаков — прежде всего алфавитных и палеографических.

Сама идея датировать предмет по его внешнему облику, опираясь на знание внешнего облика других, уже датированных, предметов, конечно, не нова — так искусствоведы датируют древнюю вазу по её орнаменту.

Однако Зализняк внёс здесь существенные новации. Во-первых, им выявлено и работает на датирование около 500 признаков, что во много раз больше, чем в любом учебнике палеографии. Во-вторых, им предложена схема действий по датированию, основанная на здравом смысле: например, из того, что некий признак бывает в документах XIII века, указанная схема ещё не позволяет делать вывод, что обладающий этим признаком исследуемый документ непременно относится к тому же веку; вместо этого схема предлагает достаточно надёжный временной интервал. В-третьих, им были составлены удобные в пользовании таблицы, позволяющие как бы автоматически извлекать из них тот временной интервал, к которому грамота с высокой вероятностью принадлежит.

Шестое. В 2000 г. в Новгороде была откопана деревянная книга первой четверти XI века — так называемый «Новгородский кодекс». Книга состояла из трёх дощечек с восковым покрытием и текстом, процарапанным по воску. Текст на воске читался сравнительно легко. Но были и тексты, нацарапанные на дереве, причём тексты двух видов: непосредственно нацарапанные на участках, не покрытых воском, и слабые следы, оставшиеся под воском от писания по воску. И те и другие читались с огромным трудом. К тому же в течение десятилетий тексты на дереве наслаивались друг на друга. Требовалось сверхчеловеческое искусство, чтобы в паутине царапин увидеть осмысленный текст, точнее — много таких скрытых текстов, наложенных друг на друга. Зализняк увидел и прочёл эти скрытые тексты. Сама догадка, что воск пронзался писалом насквозь, что оттого на деревянной подложке должны были остаться царапины и что эти еле заметные царапины допускают прочтение, сама эта догадка представляет собою отдельное замечательное достижение. Одним дано попадать в цели, в которые не могут попасть остальные. Другим дано видеть цели, которые не видят остальные.

Наконец, о «Слове о полку Игореве». Зализняк доказал его подлинность — при том понимании слова «доказал», какое вообще возможно в филологии. Доказательство опирается на анализ раскрытых им тончайших закономерностей древнерусского языка. Гипотетический фальсификатор должен был бы обладать немыслимыми качествами, а именно знать эти закономерности, иные из коих были обнаружены лишь недавно, — знать и скрывать своё знание от современников! Это при том, что, как известно, незнание можно скрыть, знание скрыть невозможно. Лет сорок назад я спросил Андрея Анатольевича, что он думает о подлинности «Слова». Он отослал меня к случившимся рядом Юрию Михайловичу Лотману и моему брату Борису Андреевичу Успенскому. «Разумеется, подлинное», — ответил Лотман. «Разумеется, подделка», — ответил брат. Сам же Зализняк ответа тогда не дал (сказал, что не знает). Вот теперь ответил. Скажу ещё, что если бы Зализняк с такой же убедительностью доказал поддельность «Слова», это было бы не меньшим достижением. Главное, что в споре поставлена точка.

В трудах Зализняка проявляются следующие характерные черты его творчества.

Первая черта. Абсолютное владение фактическим материалом.

Вторая черта. Безупречность логического анализа.

Третья черта. Небывалая для гуманитарных сочинений ясность изложения.

Четвёртая черта. Небывалая для гуманитарных сочинений аргументированность изложения.

Пятая черта. Уникальный сплав теории и практики. Практическая задача по созданию удобного в пользовании русско-французского словаря стимулировала теоретическое исследование русского словоизменения. Это исследование привело к уточнению базовых лингвистических понятий, таких как грамматическая категория вообще, как частные категории падежа, числа, рода, одушевлённости; при этом было открыто, что слова типа *сани* и *ножницы* суть слова особого, четвёртого рода, а не слова, употребляемые только во множественном числе, как всегда считалось. И уже на основе построенной теории возник, в свою очередь, «Грамматический словарь». Этот вручную составленный словарь служит сейчас, как утверждают специалисты, основой компьютерных программ анализа и синтеза словоформ для едва ли не всех систем информационного поиска, использующих русский язык.

Шестая черта. Способность к обозрению очень больших совокупностей текстов — причём совокупностей в известном смысле полных — с целью обнаружения закономерностей. Эта способность была необходимой и при создании «Грамматического словаря», и при создании методов внутренней датировки грамот.

Седьмая черта. Живое ощущение единства знания, отсутствие сектантского его разделения по ведомственным полочкам. Как разъяснил мне Валентин Лаврентьевич Янин, единство истории и филологии, характерное для энциклопедизма XVIII века, уже в XIX веке было в значительной мере утрачено, оставшись лишь в названии историко-филологических факультетов. Своим исследованием берестяных грамот и других древних текстов Зализняк возрождает это единство. Наряду с этим, в монографии «Русское именное словоизменение» можно встретить математическую теорему об устройстве ударения в русском языке.

Зализняк внёс решающий вклад и в проведение традиционных олимпиад по языковедению и математике, явившись, в частности, основателем нового, приближённого к математике, жанра лингвистических задач.

Восьмая черта. Научная честность и беспристрастность. Единственная цель — установление истины. С этим связано и тщательнейшее рассмотрение аргументов, выдвигаемых оппонентом.

И ещё кое-что, уже не относящееся, пожалуй, к научному творчеству. Большой учёный отнюдь не всегда хороший лектор. Зализняк — блестящий лектор. Неизменным успехом пользуются его занятия по структуре санскрита, древнеперсидского, арабского и других языков; слушателям раздаются краткие конспекты с необходимой грамматической информацией, а затем предлагаются домашние задания, состоящие в опирающихся на эту информацию чтении и анализе подлинных текстов. (Некогда один член-корреспондент сказал мне: «Зализняк не имеет лингвистического лица. То он преподаёт арабский, то старославянский...».) А в амфитеатре Московского университета, рассчитанном на сотни слушателей, на ежегодной традиционной лекции Зализняка о лингвистических итогах очередного сезона новгородских раскопок каждый раз не хватает сидячих мест.

И последнее. Надо отметить чрезвычайную, иногда даже чрезмерную скромность Андрея Анатольевича. 70 лет назад Пастернак высказал такой комплимент современному ему литератору: «Вы могли бы в гораздо большей степени навязать себя эпохе». Зализняк не навязывает себя эпохе. Скорей уж эпоха навязывает ему себя.

Владимир Андреевич Успенский (1930–2018)

А. Зализняк

Истина существует

Я благодарю Александра Исаевича Солженицына и всё жюри за великую честь, которой я удостоен.

В то же время не могу не признаться, что эта награда вызывает у меня не одни только приятные чувства, но и большое смущение.

В моей жизни получилось так, что моя самая прочная и долговременная дружеская компания сложилась в школе, — и с тех пор те, кто ещё жив, дружески встречаются несколько раз в год вот уже больше полувека. И вот теперь мне ясно, насколько едины мы были в своём внутреннем убеждении (настолько для нас очевидном, что мы сами его не формулировали и не обсуждали), что высокие чины и почести — это нечто несовместимое с нашими юношескими идеалами, нашим самоуважением и уважением друг к другу.

Разумеется, эпоха была виновата в том, что у нас сложилось ясное сознание: вознесённые к официальной славе — все или почти — получили её кривыми путями и не по заслугам. Мы понимали так: если лауреат Сталинской премии, то почти наверное угодливая бездарность; если академик, то нужны какие-то совершенно исключительные свидетельства, чтобы поверить, что не дутая величина и не проходимец. В нас это сидело крепко и в сущности сидит до сих пор. Поэтому никакие звания и почести не могут нам приносить того беспримесного счастья, о котором щебечут в таких случаях нынешние средства массовой информации. Если нам их всё-таки по каким-то причинам дают, нам их носить неловко.

«Устарело! — говорят нам. — Теперь уже всё по-другому, теперь есть возможность награждать достойных». Хотелось бы верить. И есть уже, конечно, немало случаев, когда это несомненно так. Но чтобы уже отжил и исчез сам фундаментальный принцип, свидетельств как-то ещё маловато...

А между тем наше восприятие российского мира не было пессимистическим. Мы ощущали так: наряду с насквозь фальшивой официальной иерархией существует подпольный гамбургский счёт. Существуют гонимые художники, которые, конечно, лучше официальных. Существует — в самиздате — настоящая литература, которая, конечно, выше публикуемой. Существуют не получающие никакого официального признания замечательные учёные. И для того, чтобы что-то заслужить по гамбургскому счёту, нужен только истинный талант, угодливости и пронырства не требуется.

Разумеется, материальные успехи определялись официальной иерархией, а не подпольной. Но мы же в соответствии с духом эпохи смотрели свысока на материальную сторону жизни. Западная формула: «Если ты умный, почему же ты бедный?» — была для нас очевидным свидетельством убогости такого типа мышления.

Ныне нам приходится расставаться с этим советским идеализмом. Для молодого поколения большой проблемы тут нет. Западная формула уже не кажется им убогой. Но нашему поколению полностью уже не перестроиться.

Мне хотелось бы сказать также несколько слов о моей упоминавшейся здесь книге про «Слово о полку Игореве». Мне иногда говорят про неё, что это патриотическое сочинение. В устах одних это похвала, в устах других — насмешка. И те и другие нередко меня называют сторонником (или даже защитником) подлинности «Слова о полку Игореве».

Я это решительно отрицаю.

Полагаю, что во мне есть некоторый патриотизм, но скорее всего такого рода, который тем, кто особенно много говорит о патриотизме, не очень понравился бы.

Мой опыт привёл меня к убеждению, что если книга по такому «горячему» вопросу, как происхождение «Слова о полку Игореве», пишется из патриотических побуждений, то её выводы на настоящих весах уже по одной этой причине весят меньше, чем хотелось бы.

Ведь у нас не математика — все аргументы не абсолютные. Так что если у исследователя имеется сильный глубинный стимул «тянуть» в определённую сторону, то специфика дела, увы, легко позволяет эту тягу реализовать — а именно, позволяет находить всё новые и новые аргументы в нужную пользу, незаметно для себя самого раздувать значимость аргументов своей стороны и минимизировать значимость противоположных аргументов.

В деле о «Слове о полку Игореве», к сожалению, львиная доля аргументации пронизана именно такими стремлениями — тем, у кого на знамени патриотизм, нужно, чтобы произведение было подлинным; тем, кто убеждён в безусловной и всегдашней российской отсталости, нужно, чтобы было поддельным. И то, что получается разговор глухих, в значительной мере определяется именно этим.

Скажу то, чему мои оппоненты (равно как и часть соглашающихся) скорее всего не поверят. Но это всё же не основание для того, чтобы этого вообще не говорить.

Действительным мотивом, побудившим меня ввязаться в это трудное и запутанное дело, был отнюдь не патриотизм. У меня нет чувства, что я был бы как-то особенно доволен от того, что «Слово о полку Игореве» написано в XII веке, или огорчён от того, что в XVIII. Если я и был чем-то недоволен и огорчён, то совсем другим — ощущением слабости и второсортности нашей лингвистической науки, если она за столько времени не может поставить обоснованный диагноз лежащему перед нами тексту.

У лингвистов, казалось мне, имеются гораздо бóльшие возможности, чем у других гуманитариев, опираться на объективные факты — на строго измеренные и расклассифицированные характеристики текста. Неужели текст не имеет совсем никаких объективных свойств, которые позволили бы отличить древность от её имитации?

Попытка раскопать истину из-под груды противоречивых суждений в вопросе о «Слове о полку Игореве» была также в значительной мере связана с более общими размышлениями о соотношении истины и предположений в гуманитарных науках — размышлениями, порождёнными моим участием в критическом обсуждении так называемой «новой хронологии» Фоменко, провозглашающей поддельность едва ли не большинства источников, на которые опирается наше знание всемирной истории.

Все мы понимаем, что в стране происходит великое моральное брожение.

Близ нас на Волоколамском шоссе, где годами нависали над людьми гигантские лозунги «Слава КПСС» и «Победа коммунизма неизбежна», недавно на рекламном щите можно было видеть исполненное столь же громадными буквами: «Всё можно купить!». Столь прицельного залпа по традиционным для России моральным ценностям я не встречал даже в самых циничных рекламах.

Вот Сцилла и Харибда, между которыми приходится искать себе моральную дорогу нынешнему российскому человеку.

Моральных, этических и интеллектуальных проблем здесь целый клубок.

По характеру моих занятий мне из них ближе всего тот аспект — пусть не самый драматичный, но всё же весьма существенный, — который касается отношения к знанию.

Вместе с яростно внушаемой нынешней рекламой агрессивно-гедонистической идеей «Возьми от жизни всё!» у множества людей, прежде всего молодёжи, произошёл также и заметный сдвиг в отношении к знанию и к истине.

Не хочу, однако, обобщать поспешно и чрезмерно. Всю жизнь, начиная с 25-летнего возраста (с одним не очень большим перерывом), я в той или иной мере имел дело со студентами. И это общение всегда было окрашено большим удовлетворением. Наблюдая сейчас за работой тех довольно многочисленных лингвистов, которых я в разное время видел перед собой на студенческой скамье, я чувствую, что их отношение к науке и способ действия в науке мне нравятся. И студенты, с которыми я имею дело теперь, по моему ощущению, относятся к своему делу с ничуть не меньшей отдачей и энтузиазмом, чем прежние.

Но за пределами этой близкой мне сферы я, к сожалению, ощущаю распространение взглядов и реакций, которые означают снижение в общественном сознании ценности науки вообще и гуманитарных наук в особенности.

Разумеется, в отношении гуманитарных наук губительную роль играла установка советской власти на прямую постановку этих наук на службу политической пропаганде. Результат: неверие и насмешка над официальными философами, официальными историками, официальными литературоведами. Теперь убедить общество, что в этих науках бывают выводы, не продиктованные властями предержащими или не подлаженные под их интересы, действительно очень трудно.

И напротив, всё время появляющиеся то тут, то там сенсационные заявления о том, что полностью ниспровергнуто то или иное считавшееся общепризнанным утверждение некоторой гуманитарной науки, чаще всего истории, подхватываются очень охотно, с большой готовностью. Психологической основой здесь служит мстительное удовлетворение в отношении всех лжецов и конъюнктурщиков, которые так долго навязывали нам свои заказные теории.

И надо ли говорить, сколь мало в этой ситуации люди склонны проверять эти сенсации логикой и здравым смыслом.

Мне хотелось бы высказаться в защиту двух простейших идей, которые прежде считались очевидными и даже просто банальными, а теперь звучат очень немодно:

1) Истина существует, и целью науки является её поиск.
2) В любом обсуждаемом вопросе профессионал (если он действительно профессионал, а не просто носитель казённых титулов) в нормальном случае более прав, чем дилетант.

Им противостоят положения, ныне гораздо более модные:

1) Истины не существует, существует лишь множество мнений (или, говоря языком постмодернизма, множество текстов).
2) По любому вопросу ничьё мнение не весит больше, чем мнение кого-то иного. Девочка-пятиклассница имеет мнение, что Дарвин неправ, и хороший тон состоит в том, чтобы подавать этот факт как серьёзный вызов биологической науке.

Это поветрие — уже не чисто российское, оно ощущается и во всём западном мире. Но в России оно заметно усилено ситуацией постсоветского идеологического вакуума.

Источники этих ныне модных положений ясны: действительно, существуют аспекты мироустройства, где истина скрыта и, быть может, недостижима; действительно, бывают случаи, когда непрофессионал оказывается прав, а все профессионалы заблуждаются.

Капитальный сдвиг состоит в том, что эти ситуации воспринимаются не как редкие и исключительные, каковы они в действительности, а как всеобщие и обычные.

И огромной силы стимулом к их принятию и уверованию в них служит их психологическая выгодность. Если все мнения равноправны, то я могу сесть и немедленно отправить и моё мнение в Интернет, не затрудняя себя многолетним учением и трудоёмким знакомством с тем, что уже знают по данному поводу те, кто посвятил этому долгие годы исследования.

Психологическая выгодность здесь не только для пишущего, но в не меньшей степени для значительной части читающих: сенсационное опровержение того, что ещё вчера считалось общепринятой истиной, освобождает их от ощущения собственной недостаточной образованности, в один ход ставит их выше тех, кто корпел над изучением соответствующей традиционной премудрости, которая, как они теперь узнали, ничего не стоит.

От признания того, что не существует истины в некоем глубоком философском вопросе, совершается переход к тому, что не существует истины ни в чём, скажем, в том, что в 1914 году началась Первая мировая война. И вот мы уже читаем, например, что никогда не было Ивана Грозного или что Батый — это Иван Калита. И что много страшнее, прискорбно большое количество людей принимает подобные новости охотно.

А нынешние средства массовой информации, увы, оказываются первыми союзниками в распространении подобной дилетантской чепухи, потому что они говорят и пишут в первую очередь то, что должно производить впечатление на массового зрителя и слушателя и импонировать ему, — следовательно, самое броское и сенсационное, а отнюдь не самое серьёзное и надёжное.

Я не испытываю особого оптимизма относительно того, что вектор этого движения каким-то образом переменится и положение само собой исправится. По-видимому, те, кто осознаёт ценность истины и разлагающую силу дилетантства и шарлатанства и пытается этой силе сопротивляться, будут и дальше оказываться в трудном положении плывущих против течения. Но надежда на то, что всегда будут находиться и те, кто всё-таки будет это делать.

Андрей Анатольевич Зализняк (1935–2017)

А. Ч. Пиперски

Андрей Анатольевич Зализняк (1935–2017)

24 декабря 2017 года в возрасте 82 лет ушёл из жизни великий российский лингвист, академик РАН Андрей Анатольевич Зализняк. На его счету — множество выдающихся открытий. Он изучал древненовгородские берестяные грамоты и описал грамматику древненовгородского диалекта; он убедительно доказал, что «Слово о полку Игореве» действительно было написано около 1200 года, а не является подделкой XVIII века, как многие считали раньше; он создал словарь русского словоизменения, который и по сей день лежит в основе компьютерных программ, работающих с русским языком; он развил и доработал определение падежа, данное А. Н. Колмогоровым, — и многое другое. Строгость и доказательность работ Андрея Анатольевича была и остаётся идеалом для всех его учеников — а таковыми считают себя почти все лингвисты Москвы и не только.

Отличительная черта научной деятельности Зализняка — интерес к математике и точным методам в языкознании. В 1950-е годы он учился на французском отделении филологического факультета Московского университета — казалось бы, что может быть более гуманитарным и далёким от точных наук? Однако он стал ходить к В. А. Успенскому на занятия по математике для филологов и оказался в числе лучших студентов. «Он меня абсолютно потряс. Он задавал какие-то вопросы, чрезвычайно глубокие, но совершенно перпендикулярные к вопросам всех остальных и к тому, чего я мог ожидать. Какой-то такой был поворот в его вопросах, словно он с другой стороны на всё смотрел. Сразу стало ясно, что это гений», — вспоминает Успенский[1)]. Вскоре Зализняк стал одним из руководителей семинара по математической лингвистике на механико-математическом факультете МГУ. О некоторых достижениях Зализняка, связанных с формализацией лингвистического знания, я и хочу рассказать поподробнее.

Русское словоизменение

В 1957–1958 гг. Зализняк был на стажировке в Париже. Там он преподавал французским студентам русский язык и обнаружил, что существующих описаний русского склонения и спряжения для иностранцев недостаточно: они слишком расплывчатые. Уже в 1961 году Зализняк

Печатается по брошюре: *Зализняк А. А.* Лингвистические задачи. М.: МЦНМО, 2018. С. 8–23.

1) Борьба за лингвистику. Беседа с математиком Владимиром Успенским. Часть 1. http://polit.ru/article/2009/07/09/linguist/

издал «Краткий русско-французский учебный словарь», приложив к нему абсолютно формальный 150-страничный очерк русского словоизменения. В 1965 году из этого очерка выросла кандидатская диссертация; работа оказалась настолько новаторской, что её сразу засчитали за докторскую. В 1967 году Зализняк издал книгу «Русское именное словоизменение», а в 1977 году — «Грамматический словарь русского языка».

В этом словаре при каждом слове стоят непривычные пометы: *дуб* м 1с, *порог* м 3а, *фонарь* м 2b. Ясно, что «м» обозначает мужской род, а зачем всё остальное? Ведь мы со школьных лет и так знаем, что *дуб*, *порог* и *фонарь* — это II склонение.

Однако для полного описания русского языка такого знания недостаточно. Чтобы убедиться в этом, образуем родительный падеж единственного числа и именительный падеж множественного числа:

нет ду́ба — есть дубы́,

нет поро́га — есть поро́ги,

нет фонаря́ — есть фонари́.

Видно, что в окончаниях этих слов появляются разные буквы (в одном — *а* и *ы*, в другом — *а* и *и*, в третьем — *я* и *и*), да и с ударением всё непросто: у слова *дуб* оно падает то на окончание, то на основу, у слова *порог* оно в обеих формах на основе, а у слова *фонарь* — на окончании. Именно эту информацию и несут в себе пометы.

Цифры обозначают, на что оканчивается основа и как будут выглядеть окончания:

1 — основа на твёрдый согласный (кроме заднеязычных *к*, *г* и *х*, шипящих *ш* и *ж* и свистящего *ц*);

2 — основа на мягкий согласный (кроме *ч* и *й*);

3 — основа на *к*, *г* или *х*; и т. д.

Таких типов у существительных насчитывается девять: 0 — несклоняемые слова (например, *кенгуру*), 1–7 — варианты традиционных I и II склонений (например, *акула*, *поезд*, *ведро*) и 8 — традиционное III склонение (например, *крепость*).

Буква обозначает схему ударения:

a — всегда на основе;

b — всегда на окончании;

c — подвижное ударение: в единственном числе на основе, а во множественном — на окончании; и т. д.

Всего у существительных насчитывается восемь схем, включающих в себя по крайней мере 10 слов. Больше 90 % существительных в «Грамматическом словаре» Зализняка относятся к схеме *a* — но это не значит, что остальными схемами можно пренебречь. Дело в том, что редкие схемы включают в себя очень частотные слова: например, многие названия частей тела. Посклоняйте слово *рука* и посмотрите, когда ударение будет на основе, а когда на окончании: распределение там очень прихотливое. И вот если считать не по словарю, а по текстам, то на все схемы, кроме *a*, придётся более четверти всех существительных (табл. 1).

Таблица 1

Схема ударения	Примеры	Число существительных в словаре	Доля в словаре (%)	Доля в текстах (%)
a	*карта*, *порог*	44 933	91,62	73,40
b	*фонарь*, *ступня*	3117	6,36	6,85
c	*дуб*, *море*	408	0,83	8,75
d	*изба*, *вино*	336	0,69	3,30
e	*ухо*, *доля*	185	0,38	5,04
f	*губа*, *червь*	35	0,07	0,38
d′	*спина*, *душа*	11	0,02	0,91
f′	*рука*, *гора*	15	0,03	1,37
Всего		49 040	100,00	100,00

Кстати, в этой таблице пары слов в каждой ячейке упорядочены так же, как и в самом «Грамматическом словаре» — в инверсионном (обратном) порядке, т. е. по алфавиту при чтении справа налево, а не слева направо. Если объединить эти 16 слов в один список и упорядочить его таким образом, получим: *изба*, *губа*, *рука*, *спина*, *гора*, *карта*, *душа*, *дуб*, *порог*, *море*, *вино*, *ухо*, *червь*, *фонарь*, *доля*, *ступня* (как и при обычной алфавитной сортировке, если последние буквы одинаковы, смотрим на предпоследние и т. д.).

Зализняк не зря выбрал такой порядок. Благодаря этому в словаре оказываются рядом слова с одинаковой правой частью, т. е. те, которые изменяются похоже: глаголы на *-нуть*, существительные на *-а* и т. п. В нашем примере видно: подряд расположились 7 слов традиционного I склонения, потом 7 слов II склонения и потом ещё 2 слова I склонения. Сейчас цель такой сортировки кажется надуманной: можно же взять оцифрованную версию словаря и за доли секунды выбрать все слова на *-нуть,* — но в 1977 году такой возможности не было. Однако с распространением компьютеров обнаружилось новое применение «Грамматического словаря»: именно на нём основаны программы, которые работают с русской морфологией. Без словаря гораздо хуже функционировали бы современные поисковые системы и голосовые помощники. Например, введя в поиск *Пушкин стихи*, вы получаете страницы, на которых написано *в стихах Пушкина.* Это не просто слова, которые отличаются на 1–2 буквы (иначе находились бы и страницы со словами *пушки* и *стихия*, а такого не происходит), а именно формы слов, построенные по моделям Зализняка. А голосовой помощник без словаря не построил бы родительный падеж от слова *фонарь*, а если бы он озвучивал текст, то не узнал бы, что надо читать *фонаря́*, а не *фона́ря*.

Легко понять, как на основе словаря Зализняка работает морфологический синтез. Если мы строим формы от слова *фонарь*, надо взять его помету (2b), выбрать из окончаний типа 2 то, которое выражает нужную комбинацию падежа и числа: например, -*я* для род. ед., — и узнать во вводной части словаря, что основа слов типа 2 получается отбрасыванием конечного *ь*. В результае выходит: *фонар-* + -*я* = *фонаря*. Чуть сложнее со словами, у которых несколько основ: так, слово *огурец* имеет помету 5*b, где * обозначает наличие беглой гласной (т. е. основ будет две: *огурец*- и *огурц-*). Опять-таки из вводной части словаря узнаём, что окончание род. ед. в типе 5 имеет вид -*а*; если есть знак *, то берётся основа без гласной в случае, если окончание не входит в множество {-∅ (пустая строка); -*й*; -*ь*; -*ью*}. Получаем *огурц-* + -*а* = *огурца*.

Противоположная задача — опираясь на «Грамматический словарь», понять по строке символов, является ли она формой какого-либо русского слова, и если да, то какие это могут быть формы и от каких слов — с вычислительной точки зрения решается сложнее. Требуется не только хранить все возможные наборы окончаний и помету для каждого слова, но и заранее построить все возможные основы и записать, в какой форме какая из основ используется. Дальше можно предложить такой алгоритм: разбиваем слово на две части всеми возможными способами и смотрим, является ли первая часть основой каких-либо слов. Если да, проверяем, бывает ли вторая часть окончанием у этих слов при такой основе, и если да, то выписываем те слова и формы, в которых это возможно. Проанализируем в качестве примера слово *паром*, для которого получается три разбора (табл. 2).

Аркадий Волож, руководитель «Яндекса», вспоминает, что в начале 1990-х его другу и коллеге Илье Сегаловичу на основе словаря Зализняка удалось создать морфологический анализатор, который работал настолько эффективно, что умещался в оперативную память тогдашнего персонального компьютера — 640 килобайт.

Можно, конечно, сказать: зачем все эти сложные обозначения, схемы ударения, чередования основ и т. д.? Давайте просто для каждого слова выпишем все формы с ударением, и готово. Но это очень неэкономно: на каждое существительное придётся по 12 форм, а на некоторые глаголы — больше 100 (считая четыре причастия во всех родах, числах и падежах). А «Грамматический словарь» — не такая уж и толстая книга, хотя и имеет

Таблица 2

Часть 1 (основа)	Часть 2 (окончание)	Возможная начальная форма	Есть у этого слова нужное окончание?	Результат
паром-	-∅	ПАРО́М	да, им. ед. и вин. ед.	ПАРО́М, им. ед. ПАРО́М, вин. ед.
паро-	*-м*	—		
пар-	*-ом*	ПАРИ́ТЬ	—	
		ПА́РИТЬ	—	
		ПА́РА	—	
		ПАР	да, твор. ед.	ПАР, твор. ед.
па-	*-ром*	ПА	—	
п-	*-аром*	—		
∅-	*-паром*	—		

большой формат: в 4-м издании описание русского словоизменения занимает 118 страниц, а за ними следуют 586 страниц собственно словаря.

Стремясь описать язык оптимальным образом, Зализняк ставил и решал в том числе и математические задачи. К одной из них подводит пример, с которого мы начали. Напомним, что слово *фонарь* имеет схему ударения *b*, т. е. ударение всегда на окончании. Прежде чем читать дальше, подумайте: нет ли здесь некоторого противоречия, которое надо разрешить?

Зададимся странным вопросом: куда падает ударение в форме *фонарь* — на основу или на окончание? На этот вопрос есть три варианта ответа:

1) на основу: ведь ясно, что гласный *а* входит в основу, а окончание здесь нулевое;
2) мы не знаем куда: основа и окончание здесь не соревнуются за ударение, поскольку окончание нулевое;
3) на окончание, но введём дополнительное правило: если ударение попало на нулевое окончание (и получается, что ему негде прозвучать), оно переносится на последний слог основы.

Обозначим ударение на основе через +, ударение на окончании через −, а «не знаем» через 0. Примем вариант 1 — самый, как кажется, естественный — и изобразим схему ударения во всех числах и падежах[2] для нескольких русских существительных (табл. 3). Нулей в полученной таблице не будет, поскольку этот вариант не предполагает ответа «не знаем».

[2] За исключением винительного падежа множественного числа, поскольку он всегда совпадает либо с родительным, либо с именительным в зависимости от одушевлённости (вин. *вижу столы* как им. *стоят столы*, вин. *вижу кошек* как род. *глаза кошек*).

Таблица 3

	ед. И.	Р.	Д.	В.	Т.	П.	мн. И.	Р.	Д.	Т.	П.	Примеры
I	+	+	+	+	+	+	+	+	+	+	+	*юноша*, *карта*, *тигр*, *порог*
II	−	−	−	−	−	−	−	−	−	−	−	*ступня*
III	+	−	−	+	−	−	−	−	−	−	−	*фонарь*
IV	+	−	−	+	−	−	−	−	−	−	−	*вратарь*
V	+	−	−	+	−	−	−	+	−	−	−	*сапог*

Видно, что вариант 1 не очень экономный: нам приходится вводить разные схемы ударения для неодушевлённого слова *фонарь* и одушевлённого слова *вратарь* (у них различается вин. ед.: *вижу фона́рь*, но *вижу вратаря́*), не говоря уже про очень похожее слово *сапог*, для которого понадобится схема V из-за того, что у него род. мн. имеет нулевое окончание: *нет сапо́г*. Всего таких схем, как посчитал Зализняк, окажется 22.

Ещё менее экономен вариант 2, который даёт нам таблицу с нулями (табл. 4). Если использовать его, схем выйдет целых 39. Слова *карта*, *тигр* и *порог* получат не ту же схему, что *юноша* (строка II не изменится, а в строках III–V вместо плюсов будут нули).

А самым экономным оказывается вариант 3 (табл. 5): если использовать его, выяснится, что *юноша*, *карта*, *тигр* и *порог* — это одна схема; *ступня*, *фонарь*, *вратарь* и *сапог* — ещё одна схема (и да, как это ни странно, в слове *порог* ударение на основе, а в слове *фонарь* — на окончании). Всего для

Таблица 4

	ед. И.	Р.	Д.	В.	Т.	П.	мн. И.	Р.	Д.	Т.	П.	Примеры
I	+	+	+	+	+	+	+	+	+	+	+	*юноша*
I′	+	+	+	+	+	+	+	0	+	+	+	*карта*
I″	0	+	+	+	+	+	+	+	+	+	+	*тигр*
I‴	0	+	+	0	+	+	+	+	+	+	+	*порог*

Таблица 5

	ед. И.	Р.	Д.	В.	Т.	П.	мн. И.	Р.	Д.	Т.	П.	Примеры
a	+	+	+	+	+	+	+	+	+	+	+	*юноша*, *карта*, *тигр*, *порог*
b	−	−	−	−	−	−	−	−	−	−	−	*ступня*, *фонарь*, *вратарь*, *сапог*

русского языка потребуется 11 таких схем. Выше были продемонстрированы только 8 из них, обозначенные как a, b, c, d, e, f, d′ и f′, поскольку остальные представлены совсем небольшим числом слов.

Зализняк не просто получил это решение интуитивно, но и доказал строго математически, что оно оптимально, если мы хотим минимизировать число строк в таблице. Вот как выглядит задача, которую он перед собой поставил:

> Имеется матрица из m строк и n столбцов, элементами которой являются знаки $+$, $-$ и 0. Заменить каждый из имеющихся в матрице нулей плюсом или минусом так, чтобы в полученной матрице без нулей (назовём её результирующей для данной матрицы) число s различных строк было минимальным.

Читателей, заинтересовавшихся полным решением, отсылаем к книге А. А. Зализняка «Русское именное словоизменение» (М.: Наука, 1967; переиздание: М.: Языки славянской культуры, 2002). Заметим, что для решения этой задачи автор сформулировал и доказал три теоремы; это нечасто встречается в лингвистических работах — но, оказывается, встречается!

Лингвистические задачи

В 1963 году Зализняк опубликовал статью «Лингвистические задачи», которая и составляет основную часть этой брошюры. В статье он продемонстрировал, что можно анализировать факты неизвестных языков, опираясь на строгую логику в сочетании с самыми базовыми представлениями о том, как устроены тексты на человеческом языке. Так, мы вправе ожидать, что порядок слов в языке не случаен, а подчиняется каким-то закономерностям: например, определение всегда стоит после определяемого слова или всегда перед ним, связанные друг с другом слова располагаются рядом, а не раскиданы в разные концы предложения и т. п. И тогда применение логических операций позволит нам понять, что есть что в непонятном тексте. Именно на этом основываются задачи 1–4, приводимые в статье.

Задачи Зализняка самодостаточны, т. е. никаких специальных знаний не нужно: например, в задаче 4 даются 14 фраз на баскском с переводами на венгерский, а дальше читателю предлагается самому перевести три фразы с венгерского на баскский. Автор специально подчёркивает, что это задача для тех, кто не знает ни того ни другого языка и, таким образом, даже не понимает, что эти фразы значат. Тем не менее, после тщательного анализа примеров становится ясно, что задача является разрешимой.

Другой тип задач, предложенных Зализняком, — это задачи на внутреннюю реконструкцию (№№ 5–7 в его статье). Это, конечно, не было так непривычно для лингвистов того времени, как перевод с неизвестного языка на неизвестный язык, но Зализняк в своих задачах довёл процедуру внутренней реконструкции до очень высокого уровня строгости. Видимое разнообразие форм того или иного языка сводится к глубинным представлениям, в которых одному значению соответствует одна форма и из которых реально наблюдаемая картина получается при помощи упорядоченных формул перехода. С конца 1960-х такой подход почти на 20 лет стал доминирующим в фонологии (это область лингвистики, которая изучает, как люди обрабатывают звуки в своём сознании, и стремится строго это описать); и в наши дни он продолжает использоваться наряду с другими моделями.

Самодостаточные лингвистические задачи впервые были предложены школьникам на Первой традиционной олимпиаде по языковедению и математике, которая прошла в 1965 году. Некоторые смеялись: «Что это такое — первая традиционная?». А она действительно стала традиционной, как того и хотел В. А. Успенский, включивший это слово в название. В 2017/18 учебном году олимпиада прошла в 48-й раз и была посвящена памяти Зализняка — создателя первых самодостаточных задач на стыке языка и логики. А с 2003 года проводится и Международная олимпиада по лингвистике, в которой участвуют уже три десятка стран: от Швеции до Бразилии.

Зализняк сочинил для олимпиады 26 задач и много лет активно участвовал в обсуждении и подготовке задач других авторов. И даже в те годы, когда обилие прочих дел уже не позволяло ему ходить на регулярные заседания задачной комиссии, все сочинённые задачи полагалось «озализнячивать» — показывать Зализняку перед тем, как давать школьникам.

История русского ударения

В 1970-е годы Зализняк обратился к истории русского языка и впоследствии плодотворно занимался ею более 40 лет. Изучение схем ударения в современном русском языке привело его к тому, что он решил разобраться, откуда они взялись, и построил строгую формальную модель, которая позволяет по составу древнерусского слова определить, какое в нём было ударение.

Рассмотрим современные русские слова *бы́ли* и *ры́ли*. В них обоих ударение падает на первый слог — но и в этом случае, как и в примере про *фонарь*, всё сложнее, чем кажется. Добавим отрицание и получим: *не́ были* и *не ры́ли*. У слова *ры́ли* — своё собственное, «настоящее» ударение, которое оно никому не отдаёт. А вот у слова *бы́ли* ударение «ненастоящее»; скажем, что у него как бы нет ударения, а просто автоматически усиливается первый слог; если же первым слогом становится частица *не*, она и получает это автоматическое усиление. Для нас это неожиданно: ведь ударение в *бы́ли* звучит так же, как в *ры́ли*, но Зализняк установил, что в древнерусском языке около 1000 года это различие было вполне живым. Имелось два типа ударения, различающихся по звучанию и совпавших только к XIV веку: автономное (самостоятельное, «настоящее») ударение, которое могло падать на любой слог слова, и автоматическое («ненастоящее»). Автоматическое ударение — это и есть то самое усиление на первом слоге. Условно будем обозначать автономное ударение в древнерусских словах привычным способом (*ры́ли*), а автоматическое — чёрточкой перед словом (¯*были*).

Возьмём три тройки существительных (возможно, с предлогом), которые до наших дней сохранили древнерусское ударение: *гру́ша — гру́шу — за гру́шу*, *жена́ — жену́ — за жену́* и *нога́ — но́гу — за́ ногу*. Вспоминается что-то знакомое: ударение всегда на основе, ударение всегда на окончании и подвижное ударение — это те же самые схемы, которые мы уже видели в работах Зализняка про современный русский язык. В своей книге «От праславянской акцентуации к русской» (1985) он наглядно объяснил, откуда и как эти схемы произошли.

Припишем каждой морфеме (корню, окончанию и т. п., а также предлогам, союзам и частицам) одну из трёх маркировок: ↓, → или –. Договоримся, что ударение определяется самой левой стрелкой. Если это ↓, то автономное («настоящее») ударение падает на отмеченный ею слог; если это →, то автономное («настоящее») ударение падает на слог правее неё; если же стрелок в слове нет, а есть только минусы, слово имеет автоматическое («ненастоящее») ударение. Для того чтобы описать все приведённые выше примеры, достаточно разметить морфемы так: *за* –, *груш* ↓, *жен* →, *ног* –, *а* ↓, *у* –. После этого получаем все реально наблюдаемые ударения:

гру́ш-а ↓ ↓	*гру́ш-у* ↓ –	*за гру́ш-у* – ↓ –
жен-а́ → ↓	*жен-у́* → –	*за жен-у́* – → –
ног-а́ – ↓	¯*ног-у* – –	¯*за ног-у* – – –

Этот простой алгоритм (с некоторыми добавлениями, которые, разумеется, изложены в работах Зализняка) позволяет полностью описать древнерусское ударение. А чтобы потренироваться, сделайте небольшое упражнение.

Упражнение 1

Даны пять глаголов в двух формах прошедшего времени: *могла́*, *могли́*; *мы́ла*, *мы́ли*; *жила́*, *жи́ли*; *вы́мыла*, *вы́мыли*; *вы́жила*, *вы́жили*. Известно, что место современного ударения во всех этих формах совпадает с древнерусским.

Задание. Для каждой морфемы, содержащей хотя бы один гласный (приставка *вы-*; корни *-мог-*, *-мы-*, *-жи-*; окончания *-а*, *-и*), укажите её древнерусскую маркировку (↓, → или −) так, чтобы получить правильное ударение во всех формах.

Примечание. Поскольку суффикс прошедшего времени *-л-* не содержит гласного, ему можно не приписывать маркировку или приписать маркировку −.

Как определить ударение по рукописи

Схема, представленная выше, описывает древнерусскую ситуацию, которая затем заметно изменилась. Мы и сами это замечаем: например, сейчас вполне можно сказать *за нóгу*. В своих работах Зализняк очень подробно описал, как изменялось древнерусское ударение. Для этого он проанализировал 89 старинных рукописей и печатных книг с проставленными ударениями. Ударения ставились в текстах далеко не всегда — но иногда получить информацию об ударении можно даже в случае, если ударение не проставлено.

Сделаем небольшой экскурс в историю русского языка. Наряду с «обычными» гласными, в древнерусском языке около 1000 года встречались два «особых» сверхкратких гласных: *ъ* (как *ы*, но короче) и *ь* (как *и*, но короче). Например, слова *столъ*, *годъ* и *мъхъ* 'мох' были двусложными (*сто-лъ*, *го-дъ*, *мъ-хъ*), слова *голубь*, *нужьно* 'нужно', *польза*, *тьмьно* 'темно', *шьвьць* 'швец' — трёхсложными (*го-лу-бь*, *ну-жь-но*, *по-ль-за*, *ть-мь-но*, *шь-вь-ць*). В XI–XIV веках с *ъ* и *ь* происходило изменение, которое можно описать так[3]:

1) в конце слова гласные *ъ* и *ь* выпадают[4];
2) далее движемся по слову справа налево; если мы находим *ъ* или *ь*, то смотрим на ближайший гласный справа от него:
   а) если это выпавший *ъ* или *ь*, то рассматриваемый на данном шаге *ъ* или *ь* проясняется, т. е. превращается в *о* или *е* соответственно;
   б) в любом другом случае (т. е. если справа стоит «обычный» гласный или прояснившийся *ъ* или *ь*) *ъ* и *ь* выпадают.

В табл. 6 продемонстрировано, как эти правила работают (каждую строку следует читать справа налево)[5].

[3] Эта закономерность была открыта ещё до Зализняка и часто называется законом Гавлика в честь чешского лингвиста Антонина Гавлика (1855–1925).

[4] Буква *ъ* продолжает сохраняться на письме до 1918 года, а *ь* во многих случаях и до сих пор, обозначая мягкость согласного.

[5] Согласный, который стоял перед выпавшим *ь*, иногда сохраняет мягкость (*голу<u>бь</u>*, *по<u>ль</u>за*), а иногда нет (*ну<u>ж</u>но*, *те<u>м</u>но*, *<u>ш</u>ве<u>ц</u>*). Для нашего изложения этот вопрос несущественен.

Таблица 6

Стало	3-й слог с конца	2-й слог с конца	1-й слог с конца	Было
стол		*сто*	*лъ* выпадает (1)	*столъ*
год		*го*	*дъ* выпадает (1)	*годъ*
мох		*мъ* *ъ* → *о* (2а)	*хъ* выпадает (1)	*мъхъ*
голубь	*го*	*лу*	*бь* выпадает (1)	*голубь*
нужно	*ну*	*жь* выпадает (2б)	*но*	*нужьно*
польза	*по*	*ль* выпадает (2б)	*за*	*польза*
темно	*ть* *ь* → *е* (2а)	*мь* выпадает (2б)	*но*	*тьмьно*
швец	*шь* выпадает (2б)	*вь* *ь* → *е* (2а)	*ць*	*шьвьць*

Видно, что в результате этого процесса появилось много слогов, которые оканчиваются на согласный: *мох*, первый слог в *поль-за* и т. п.

Упражнение 2

Проверьте себя: какие слова (не обязательно в начальной форме) получились из древнерусских слов *дьнь*, *львъ*, *сънъ*, *пьшено*, *лъжька*, *почьсть*, *льстимъ*, *льсть*, *чьтьць*, *Смольньску*?

А теперь, после этого экскурса, перенесёмся в XIV век. Именно тогда была написана рукопись, которая благодаря Зализняку стала известна как ценный источник знаний об ударении — справочник для судей под названием «Мерило праведное» (это значит «Весы правосудия»). Знаков ударения в рукописи нет, но зато различаются два типа *о*: «закрытое *о*», которое обозначалось буквой *ω* («омега») и звучало как нечто среднее между *о* и *у*, и «открытое *о*», которое обозначалось буквой *о* («он») и звучало как нечто среднее между *о* и *а*. Закрытое *о* возникло из раннедревнерусского *о* (причём только из «обычного» *о*, а не из *ъ*) в трёх случаях:

1) в начальном слоге, который ранее оканчивался на гласный, а после выпадения *ъ* и *ь* стал оканчиваться на согласный: *гωдъ* (было двусложное

¯го-дъ, которое превратилось в односложное *год*), *тѡч-но* (было *тó-чь-но*); назовём такой слог перестроенным;

2) под автономным ударением: *злѡ* (*зълó*), *поклѡнъ* (*поклóнъ*);
3) на два слога правее автономного ударения (по счёту слогов уже после выпадения *ъ* и *ь*): *обладаемѡе* (*обладáемое*) *вѡинствѡ* (*вóиньство*).

В остальных случаях в «Мериле» раннедревнерусское *о* даёт открытое *о*: *году*, *золото* (*¯году*, *¯золото* — автоматическое ударение, начальные слоги оканчиваются на гласный), *молода* (*молодá* — оба *о* безударные, начальный слог оканчивается на гласный). Если *ъ* превратился в *о*, это всегда открытое *о*: *мохъ* (*мъ́хъ*) — не помогает ни то, что начальный слог перестроенный, ни автоматическое ударение.

Зная всё это, как нам понять ударение по написанию в «Мериле праведном», при том что знаков ударения там нет? Например, мы видим в «Мериле» слово *теплѡтою* (из раннедревнерусского *теплотою*). Могло ли там быть автоматическое ударение (*¯теплотою*)? Нет, потому что тогда во втором слоге не появилась бы *ѡ*. А могло ли там быть автономное ударение на 1-м, 3-м или 4-м слоге: *тéплотою*, *теплотóю* или *теплотою́*? И опять-таки нет, потому что не появилась бы *ѡ* в нужном месте: *тéплотою* и *теплотóю* дали бы *теплотѡю* (проверьте, понимаете ли вы, почему в этих двух случаях написание совпадёт), а *теплотою́* вовсе не дало бы буквы *ѡ*. Значит, это слово имело ударение *теплóтою*, непривычное для нас сегодня.

А какое было ударение в слове *старость* (из раннедревнерусского *старость*)? Точно определить не удаётся, но у нас есть основания сказать, что там могло быть автоматическое ударение (*¯старость*) или автономное ударение на 1-м слоге (*стáрость*)[6], но только не автономное ударение на 2-м слоге: из *старóсть* в «Мериле» получилось бы *старѡсть*. В слове *золоту* (из *золоту*) могло быть автоматическое ударение (*¯золоту*) или автономное ударение на 3-м слоге (*золоту́*), но не автономное ударение на 1-м или 2-м слоге (*зóлоту* дало бы *зѡлоту*, а *золóту* — *золѡту*). В слове *пѡпъ* (из *попъ*) может быть любое ударение: если там было *¯попъ*, то *ѡ* получится по правилу про перестроенный слог, а если было *пóпъ*, то сразу по двум правилам — про перестроенный слог и про автономное ударение.

Зализняку было труднее, чем вам: ему пришлось выявить все эти правила, читая рукопись. Но и вы можете попробовать повторить заключительную часть его открытия и решить следующую задачу.

[6] Разумеется, если для какого-то слова у нас получилось несколько вариантов ударения, это не значит, что все они реально встречались: скорее всего, использовался только один из них, но мы не можем по написанию надёжно определить, какой именно.

Задача

Пусть вам известно написание слова в «Мериле праведном» (XIV век) и его раннедревнерусское написание, отражающее произношение около 1000 года. Постройте алгоритм, который позволяет определить, каким могло, а каким не могло быть древнерусское ударение в этом слове. Иначе говоря, для слова, которое в «Мериле праведном» является n-сложным, вам надо сказать, могло ли в нём быть: автоматическое ударение; автономное ударение на 1-м, 2-м, ..., n-м слоге.

Чтобы вы могли решать эту задачу не чисто умозрительно, а на реальном материале, вот несколько слов в обоих видах: *истиненъ* (*истиненъ*), *мѡжемъ* (*можемъ*), *живѡтъ* (*животъ*), *подѡбалѡ* (*подобало*), *свободенъ* (*свободьнъ*), *свободнѡму* (*свободьному*), *золото* (*золото*), *мѡрьскую* (*морьскую*), *погребенѡ* (*погребено*), *свитокъ* (*съвитъкъ*), *полезно* (*пользьно*), *золъ* (*зълъ*).

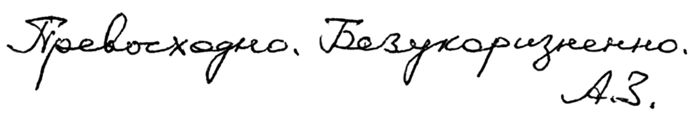
Указание. Помните, что ко времени написания «Мерила праведного» буквы ъ и ь уже не читаются как гласные и о них при определении ударения думать не надо.

Подсказка 1. Ровно для пяти из данных слов вы точно можете определить древнерусское ударение (т. е. доказать, что все варианты ударения, кроме одного, невозможны).

Подсказка 2. Ровно про три из данных слов вы точно можете сказать, что место ударения в них не такое же, как в современном русском языке.

* * *

Вот уже почти десять лет я бережно храню один листок бумаги. В 2008/09 году Зализняк читал в МГУ имени М. В. Ломоносова курс по истории русского ударения и дал на дом задание — то, что я только что предложил вам как задачу. Проверив мою работу, он исправил одну опечатку и написал:



С тех пор, читая книги и статьи самого Андрея Анатольевича и слушая его лекции, я всегда думал и продолжаю думать, что эти слова в первую очередь характеризуют не моё домашнее задание, а его самого и всё то, что он делал. Превосходно. Безукоризненно.

Март 2018 г. *А. Ч. Пиперски*

Ответы к упражнениям и решение задачи

Упражнение 1. *вы-* ↓, *-мог-* →, *-мы-* ↓, *-жи-* —, *-а* ↓, *-и* —.

Упражнение 2. *день*, *лев*, *сон*, *пшено*, *ложка*, *почесть*, *лъстим*, *лесть*, *чтец*, *Смоленску*.

Задача. Шаг 1 (классификация омег). Разделим все буквы ω на два класса: свободная ω и несвободная ω. Назовём несвободными такие ω, которые стоят в 1-м перестроенном слоге, и свободными — все остальные.

Шаг 2 (первичная генерация вариантов). Если в записи «Мерила» есть хотя бы одна свободная ω, слово может иметь только автономное ударение на слоге с самой левой свободной ω или на два слога левее (если такой слог есть). Если в слове нет ни одной свободной ω, слово может иметь любое ударение (автоматическое или автономное на любом слоге)

Шаг 3 (ограничение возможностей). Если в записи «Мерила» есть хотя бы одна буква *о* на месте раннедревнерусского *о*, для каждой такой буквы из возможностей, полученных в шаге 2, следует исключить автономное ударение на этом слоге и на два слога левее (если такой слог есть).

Рассмотрим несколько примеров:

Написание «Мерила»	Шаг 1 (свободные ω подчёркнуты)	Шаг 2 (отмечены все варианты)	Шаг 3 (отмечены все варианты)	Число вариантов
истиненъ	*истиненъ*	*¯и́сти́не́нъ*	*¯и́сти́не́нъ*	4
золото	*золото*	*¯зо́ло́то́*	*¯золото*	1
погребенω	*погребен*$\underline{\omega}$	*погре́бенώ*	*погре́бенώ*	2
подωбалω	*под*$\underline{\omega}$*бал*$\underline{\omega}$	*подώбалω*	*подώбалω*	1
свободенъ	*свободенъ*	*¯сво́бо́де́нъ*	*¯свободе́нъ*	2
свитокъ	*свитокъ*	*¯сви́то́къ*	*¯сви́то́къ*[7)]	3

Однозначно определяется ударение в словах *мώжемъ*, *живώтъ*, *подώбалω*, *свободнώму* и *¯золото*. В словах *подώбалω* и *свободнώму* ударение отличается от современного (*подоба́ло* и *свобо́дному*); в слове *¯свободе́нъ* мы не можем однозначно определить ударение, но оно в любом случае отлично от современного *свобо́ден*[8)].

[7)] Так как здесь *о* из *ъ*, исключить ударение на *о́* нельзя.

[8)] Из других источников известно, что на самом деле было ударение *¯свободенъ*.

Список литературы

[1] *Беликов В. И., Муравенко Е. В., Алексеев М. Е.* (ред.-сост.). Задачи лингвистических олимпиад. М.: МЦНМО, 2006.

[2] http://old.stsl.ru/manuscripts/medium.php?manuscript=015

[3] *Зализняк А. А.* Противопоставление букв о и ω в древнерусской рукописи XIV века «Мерило праведное» // Советское славяноведение. 1978. № 5, С. 41–68.

Александр Чедович Пиперски, НИУ ВШЭ

apiperski@gmail.com

Геометрия: классика и современность

Теорема о семи окружностях и другие новые теоремы

С. Дж. А. Ивлин, Г. Б. Мани-Каутс, Дж. А. Тиррелл

Вышедшая в 1974 г. книга C. J. A. Evelin, G. B. Money-Coutts, J. A. Tyrrell «The seven circles theorem and other new theorems» состоит из трёх независимых параграфов. Ниже публикуется перевод третьего параграфа. Сохранены авторские обозначения. Перевод второго параграфа опубликован в выпуске 21 «Математического просвещения». Редакция признательна С. В. Маркелову, обратившему внимание на эту книгу.

§ 3. Цепочки окружностей

Цепочкой окружностей в этом параграфе будет последовательность окружностей $S_1, S_2, \ldots, S_n$, каждая из которых касается соседних, т. е. S_1 касается S_2, S_2 касается $S_3, \ldots, S_{n-1}$ касается S_n. Цепочку, в которой S_n также касается S_1, назовём *замкнутой*. В противном случае будем называть цепочку *открытой*.

Все теоремы этого параграфа связаны с замкнутыми цепочками шести окружностей. Интересно, что каждый раз их формулировки требуют гораздо больше внимания, чем кажется на первый взгляд!

3.1. Теорема о семи окружностях

Наша первая цепочка, к которой относится теорема о семи окружностях, выглядит достаточно просто (рис. 1a). Однако, как мы уже отметили

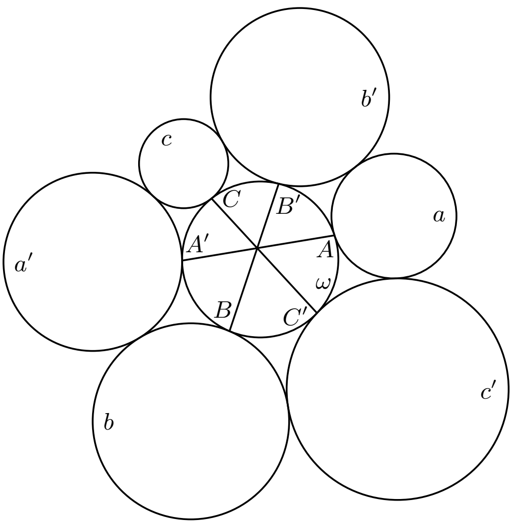
Перевод А. А. Заславского.

выше, точно изложить соответствующий результат не так легко. Начнём со следующей предварительной формулировки, которая верна лишь при некотором дополнительном условии.

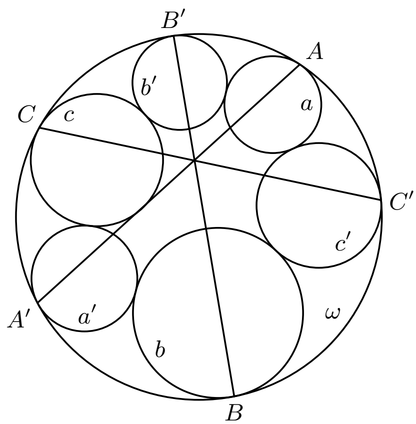
Утверждение о семи окружностях. *Пусть* a, b', c, a', b, c' — *замкнутая цепочка окружностей, касающихся окружности* ω *в шести различных точках* A, B', C, A', B, C' *соответственно. Тогда прямые* AA', BB', CC' *пересекаются в одной точке.*

Сразу добавим, что если точки A, B', C, A', B, C' лежат на ω именно в таком циклическом порядке, то утверждение верно без дополнительных условий. Рисунок 1 иллюстрирует именно этот случай. Однако точки могут лежать на ω и в другом порядке; на рис. 2 и 3 приведены примеры, когда утверждение верно, а на рис. 4 — когда не верно.

Чтобы разобраться в ситуации, посмотрим на все конфигурации, иллюстрирующие утверждение, *временно исключив из них последнюю окружность* c'. Можно заметить, что во всех случаях существуют две окружности, касающиеся a, b и ω, причём если в качестве c' взять одну из них, то без дополнительных условий утверждение верно, а если другую, то неверно. Хочется сказать, что утверждение верно в 50 % случаев! Везде, кроме рис. 4, окружность c' выбрана правильно, а на рис. 4 сделан неверный выбор. Для читателя не составит труда построить на рис. 4 другую окружность c', для



а)



б)

Рис. 1. *Примеры, когда утверждение о семи окружностях верно без дополнительных условий*

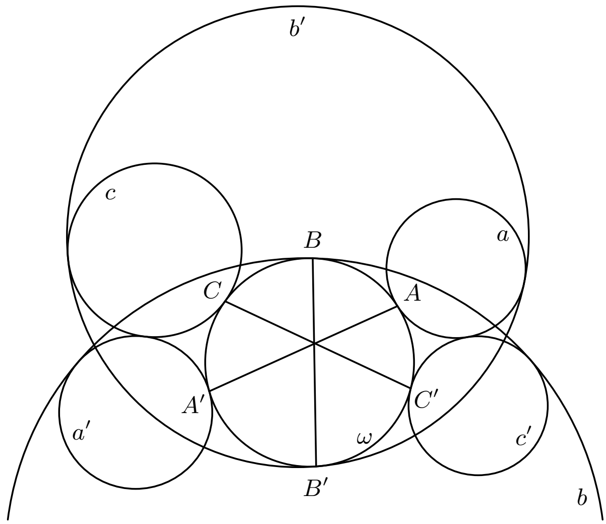


Рис. 2. Пример, когда утверждение о семи окружностях верно

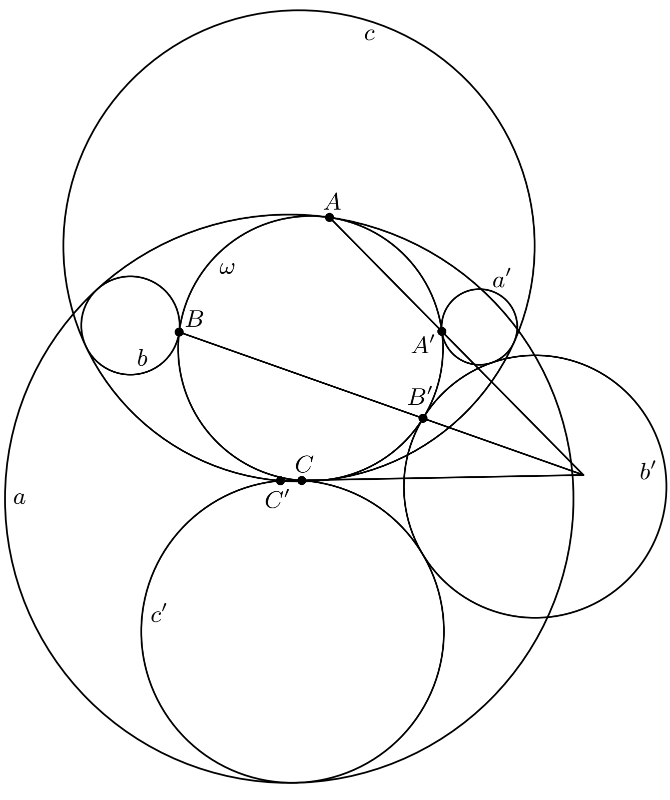


Рис. 3. Пример, когда утверждение о семи окружностях верно

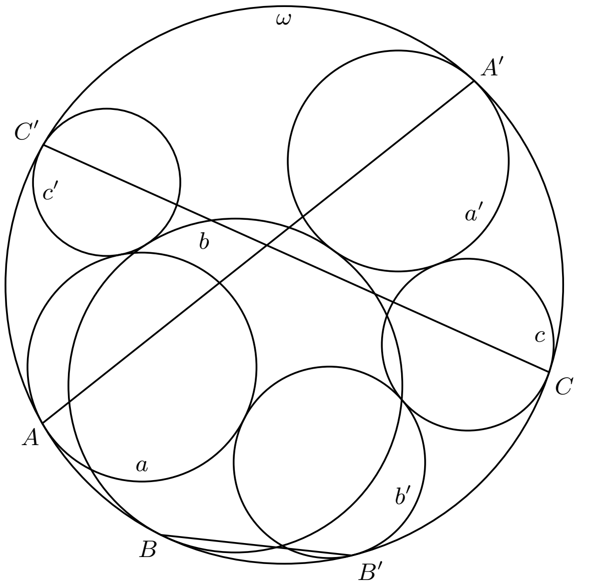


Рис. 4. Пример, когда утверждение о семи окружностях неверно

которой прямые AA', BB' и CC' будут пересекаться в одной точке. Если исследовать всевозможные расположения окружностей, станет ясно, что справедливость утверждения зависит от некоторого условия на порядок точек A, B', C, A', B, C' на окружности ω.

Для точной формулировки этого условия введём понятие *положительной дуги*: пусть P, Q, R — три точки окружности; дугу между P и R, содержащую Q, будем обозначать (PQR) и называть положительной дугой, если движение по ней от P к R происходит в положительном (против часовой стрелки) направлении. Теперь завершим формулировку теоремы о семи окружностях (используем обозначения из утверждения на с. 52).

Теорема 3.1. *Пусть среди четырёх дуг* $(BC'A')$, $(C'AA')$, $(AB'A')$, $(B'CA')$ *чётное число положительных. Тогда утверждение о семи окружностях верно.*

Доказательство. Сделаем инверсию с центром A'. Сохраняя для полученной фигуры исходные обозначения, получим:

(i) a' и ω являются параллельными прямыми (как образы окружностей, касающихся в точке A');

(ii) b и c, касаясь параллельных прямых a' и ω, являются *равными* окружностями;

(iii) b, c', a, b', c образуют открытую цепочку окружностей, касающихся ω в точках B, C', A, B', C соответственно.

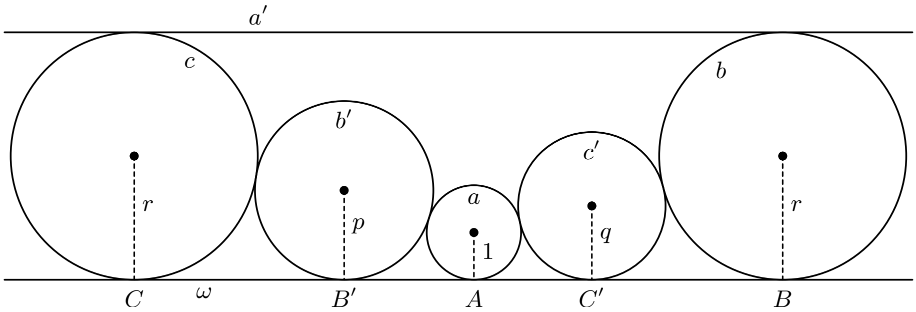


Рис. 5. Конфигурация после инверсии

Одна из возможных фигур изображена на рис. 5. Могут возникнуть и другие весьма различные конфигурации, однако все они должны удовлетворять следующим двум условиям.

(а) Все окружности в образе инверсии лежат по одну сторону от ω: иначе нашлись бы две касающиеся окружности, лежащие по разные стороны от ω, т. е. касающиеся ω в одной точке, что противоречит условиям теоремы. В терминах исходной конфигурации это означает, что окружности a, b', c, a', b, c' лежат либо все вне, либо все внутри ω.

(б) Все касания окружностей в открытой цепочке b, c', a, b', c являются внешними (хотя в исходной фигуре могло быть иначе). Действительно, окружности, касающиеся друг друга внутренним образом, не могли бы касаться ω в разных точках.

Читателю будет полезно в процессе доказательства обращаться к рис. 5, однако доказательство проходит для всех возможных конфигураций.

На рис. 5 не изображена точка, в которой находился центр инверсии, так как её положение несущественно для дальнейшего. Тем не менее, нам понадобится обозначение для этой точки, поэтому будем называть её O (обозначение A' не годится, так как точка A' при инверсии перешла в бесконечно удалённую). Образами прямых AA', BB' и CC' при инверсии будут соответственно прямая OA и окружности OBB', OCC'. А поскольку пучок прямых при инверсии (если её центр не совпадает с их общей точкой) переходит в пучок соосных окружностей, наша теорема означает, что прямая OA является радикальной осью окружностей OBB' и OCC'. Так как O лежит на этих окружностях, достаточно доказать, что A лежит на их радикальной оси, т. е.

$$(\overrightarrow{AB'}, \overrightarrow{AB}) = (\overrightarrow{AC'}, \overrightarrow{AC}) \qquad (1)$$

(эти выражения равны степеням точки A относительно обеих окружностей). Равенство (1) относится к *направленным отрезкам* на прямой ω, и его нужно доказать с учётом и абсолютных величин, и знаков.

Используя равенства

$$\overrightarrow{AB} = \overrightarrow{AC'} + \overrightarrow{C'B}, \quad \overrightarrow{AC} = \overrightarrow{AB'} + \overrightarrow{B'C},$$

приведём (1) к виду

$$(\overrightarrow{AB'}, \overrightarrow{C'B}) = (\overrightarrow{B'C}, \overrightarrow{AC'}). \tag{2}$$

Прежде всего покажем, что левая и правая части (2) имеют одинаковый знак. Согласно дополнительному условию теоремы, среди дуг $(BC'A')$, $(C'AA')$, $(AB'A')$, $(B'CA')$ в исходной фигуре чётное число положительных; значит, после инверсии, переводящей A' в бесконечно удалённую точку, чётное число из отрезков BC', $C'A$, AB', $B'C$ имеет направление, совпадающее с положительным направлением на прямой ω (независимо от того, какое из двух направлений мы считаем положительным). Следовательно, знаки частей в (2) совпадают.

Осталось показать, что правая и левая части (2) равны по абсолютной величине. Для этого будем считать, что радиусы окружностей a, b', c' на рис. 5 равны 1, p и q соответственно, а радиус окружностей b и c равен r. Элементарно доказывается, что если окружности с радиусами r_1, r_2 касаются внешним образом, то длина их общей касательной равна $2\sqrt{r_1 r_2}$. И так как любые соседние окружности в открытой цепочке b, c', a, b', c касаются внешним образом, то (без учёта знаков)

$$BC' = 2\sqrt{qr}, \quad C'A = 2\sqrt{q}, \quad AB' = 2\sqrt{p}, \quad B'C = 2\sqrt{pr}.$$

Итак, обе части (2) равны по модулю $4\sqrt{pqr}$, что и доказывает теорему. □

Попутно заметим, что левая и правая части (2) равны по модулю, даже если применить инверсию к «неправильной» конфигурации рис. 4, результат изображён на рис. 6. В этом случае теорема неверна из-за того, что знаки частей в (2) противоположны.

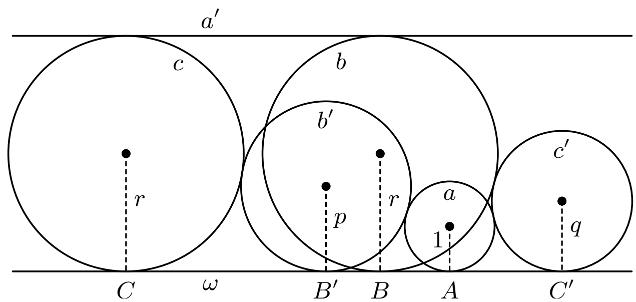


Рис. 6. Пример, когда утверждение о семи окружностях неверно. Конфигурация после инверсии

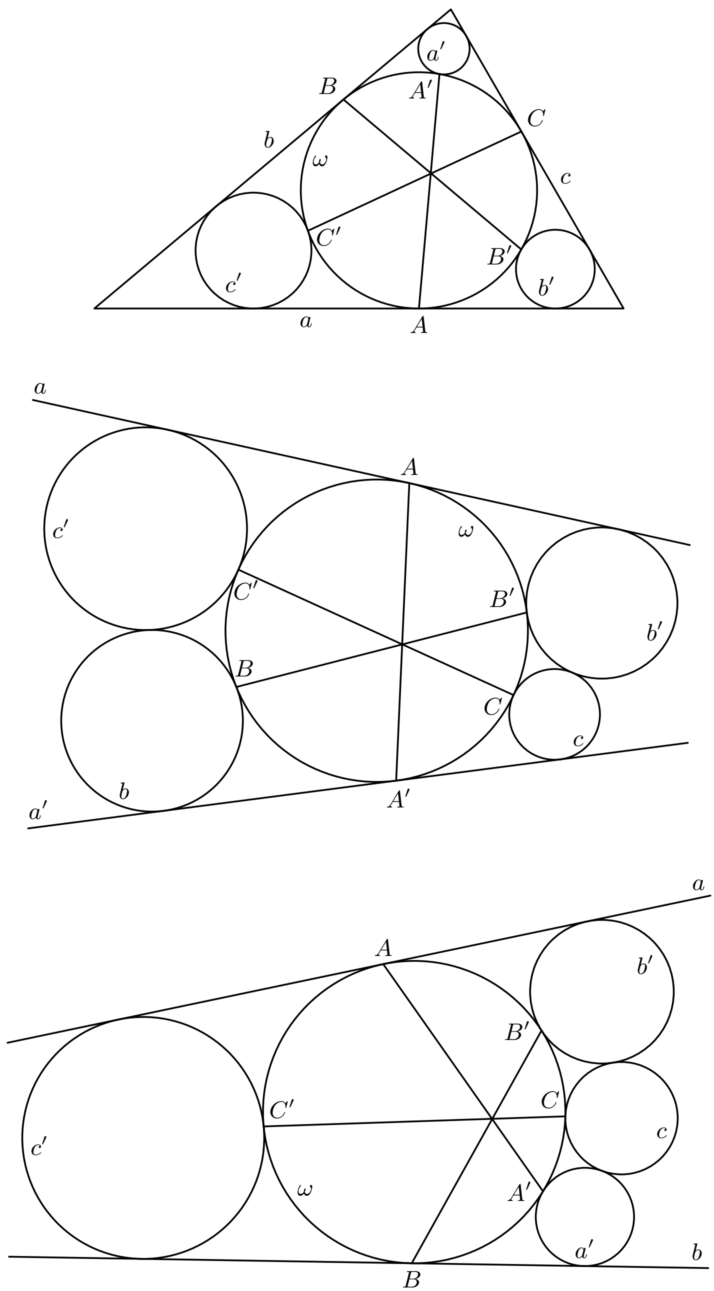


Рис. 7. *К теореме о семи окружностях. Вырожденные случаи*

Также можно заметить, что (в правильной конфигурации) точка пересечения прямых AA', BB' и CC' является точкой Брианшона шестиугольника, образованного касательными к ω в точках A, B', C, A', B, C'. Это непосредственно следует из известного факта, что если хорды AA' и BB' окружности (или любой коники) пересекаются в точке X, то X лежит на прямой, соединяющей полюса прямых AB' и $A'B$.

Наконец, заметим, что утверждение теоремы можно переформулировать следующим образом: (A, A'), (B, B'), (C, C') являются тремя парами соответственных точек некоторой инволюции окружности ω. Отсюда можно получить нетривиальное утверждение для случая, когда ω вырождается в прямую.

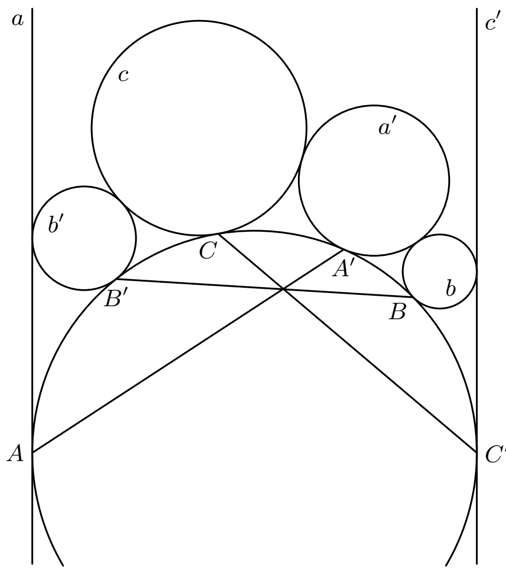


Рис. 8. К теореме о семи окружностях. Вырожденный случай

Другие специальные случаи теоремы получаются, когда в прямые вырождаются некоторые из окружностей a, b', c, a', b, c'; см. рис. 7–8 (на рис. 8 прямые a и c' параллельны).

3.2. Теорема о двух параллельных прямых и шести окружностях

Теорема, обсуждаемая в этом пункте, как мы увидим, родственна теореме о семи окружностях. Было бы интересно получить более общий результат, включающий обе эти теоремы как частные случаи, но нам пока не удалось доказать простую теорему такого рода.

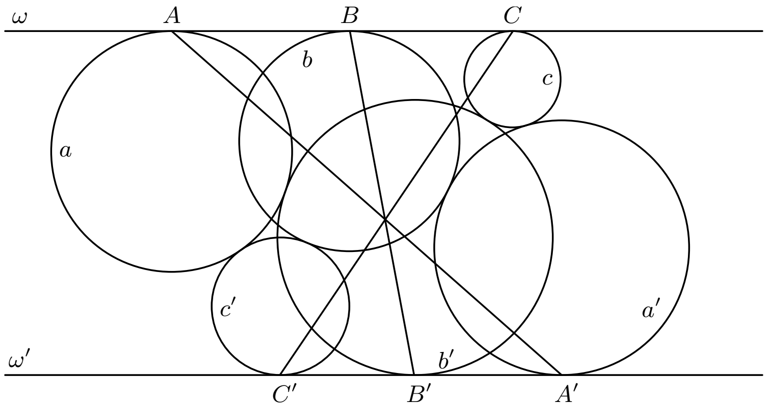


Рис. 9. К теореме 3.2 о двух параллельных прямых и шести окружностях

Идея настоящей теоремы заключается в замене окружности ω из теоремы о семи окружностях парой параллельных прямых ω и ω' и рассмотрении замкнутой цепочки окружностей a, b', c, a', b, c', поочерёдно касающихся ω и ω' (рис. 9). Тогда при выполнении некоторых условий верно заключение теоремы о семи окружностях, т. е. прямые AA', BB', CC', соединяющие «противоположные» точки касания, пересекаются в одной точке. Однако и здесь, чтобы точно сформулировать теорему, необходима аккуратность. Так, рассматривая рис. 9, можно заметить, что при данной окружности a окружность b' достаточно большого радиуса может пересекать прямую ω, и в этом случае окружность c может касаться b' как внешним, так и внутренним образом, а прямой ω — как «сверху», так и «снизу».

Чтобы ограничить разнообразие случаев, мы сосредоточимся на конфигурациях, в которых:

(а) все окружности касаются друг друга внешним образом;

(б) окружности a, b, c касаются ω снизу;

(в) окружности a', b', c' касаются ω' сверху.

При этих условиях теорема верна. На самом деле она верна при гораздо более общих условиях, но мы отложим обсуждение этого вопроса до конца этого пункта. Здесь же сформулируем доказываемое утверждение.

Теорема 3.2. *Пусть a, b', c, a', b, c' — замкнутая цепочка окружностей, поочерёдно касающихся параллельных прямых ω и ω' в точках A, B', C, A', B, C' соответственно, причём выполняются все условия* (а)–(в). *Тогда прямые AA', BB', CC' пересекаются в одной точке.*

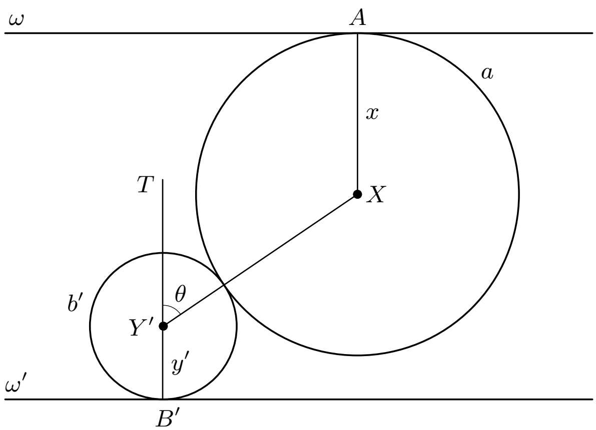
Доказательство. Обозначим расстояние между ω и ω' через $2k$. Наше доказательство будет основано на многократном применении следующей простой тригонометрической леммы.

Лемма. *Пусть a — окружность с центром X, касающаяся ω снизу, а b' — окружность с центром Y', касающаяся ω' сверху, причём касание a и b' внешнее. Если горизонтальная проекция отрезка XY' равна $2pk$, то его вертикальная проекция равна $k(1-p^2)$, причём эта величина положительна, если X лежит выше Y', и отрицательна, если X ниже Y'.*

Доказательство леммы. Обозначим через x и y' радиусы окружностей a и b' соответственно, а через A и B' — точки их касания с ω и ω'. Пусть $\theta = \angle TY'X$, где T — произвольная точка на продолжении $B'Y'$ за точку Y'. Два возможных случая изображены на рис. 10: на первом угол θ острый, Y' лежит ниже, чем X, и вертикальная проекция отрезка XY' положительна; на втором угол θ тупой, Y' лежит выше, чем X, и вертикальная проекция отрезка XY' отрицательна.

Рассмотрим ломаную $AXY'B'$, её вертикальная проекция равна $2k$ (расстоянию между ω и ω'), поэтому получаем $x + (x+y')\cos\theta + y' = 2k$, т. е.

$$x + y' = \frac{2k}{1+\cos\theta}. \qquad (1)$$



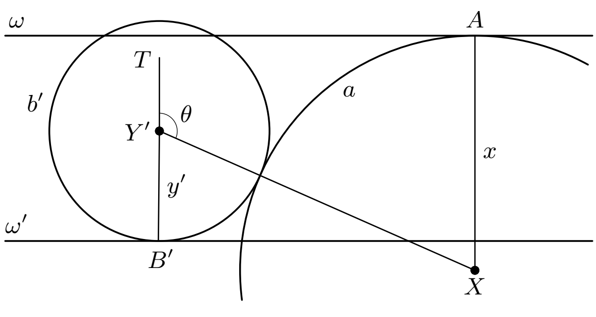


Рис. 10. К доказательству леммы

С другой стороны, горизонтальная h и вертикальная v проекции отрезка XY' равны

$$h = (x + y')\sin\theta, \quad v = (x + y')\cos\theta,$$

причём v положительна, если угол θ острый, и отрицательна, если θ тупой. С учётом (1) получаем

$$h = \frac{2k\sin\theta}{1+\cos\theta}, \quad v = \frac{2k\cos\theta}{1+\cos\theta}.$$

Теперь, поскольку

$$\frac{\sin\theta}{1+\cos\theta} = \operatorname{tg}\frac{\theta}{2}, \quad \frac{2\cos\theta}{1+\cos\theta} = 1 - \operatorname{tg}^2\frac{\theta}{2},$$

получаем $h = 2kp$, $v = k(1-p^2)$, где $p = \operatorname{tg}(\theta/2)$. Лемма доказана. □

Возвращаемся к доказательству теоремы 3.2. Пусть X, Y', Z, X', Y, Z' — центры окружностей a, b', c, a', b, c' соответственно. Обозначим через $2kp$, $2kq$, $2kr$, $2ks$, $2kt$, $2ku$ горизонтальные проекции отрезков XY', $Y'Z$, ZX', $X'Y$, YZ', $Z'X$ соответственно. В формулировке леммы знак горизонтальной проекции был несущественным; однако сейчас будем считать p положительным, если Y' лежит правее X, и отрицательным в противном случае, и аналогично зададим знаки q, r, s, t, u. Так как ломаная $XY'ZX'YZ'X$ замкнута, её горизонтальная проекция равна нулю, т. е.

$$p + q + r + s + t + u = 0. \tag{2}$$

Сумма вертикальных проекций также равна 0, что с учётом леммы даёт

$$k(1-p^2) - k(1-q^2) + k(1-r^2) - k(1-s^2) + k(1-t^2) - k(1-u^2) = 0. \tag{3}$$

Перепишем (3) в виде

$$r^2 - q^2 - s^2 = u^2 - p^2 - t^2. \tag{4}$$

Из (2) получаем

$$(r + q + s)^2 = (u + p + t)^2. \tag{5}$$

Сложив (4) и (5) и разделив на 2, получим

$$(r + s)(q + r) = (u + p)(t + u),$$

т. е.

$$\frac{r+s}{t+u} = \frac{u+p}{q+r}. \tag{6}$$

Но $k(r + s)$ — это сумма горизонтальных проекций отрезков ZX' и $X'Y$, т. е. горизонтальная проекция отрезка ZY, которая, очевидно, равна CB.

Аналогично $k(t+u)$, $k(u+p)$, $k(q+r)$ равны соответственно BA, $C'B'$ и $B'A'$. Поэтому (6) эквивалентно равенству

$$\frac{CB}{BA} = \frac{C'B'}{B'A'},$$

а так как A, B, C и A', B', C' лежат на параллельных прямых ω и ω' соответственно, то AA', BB', CC' проходят через одну точку. Теорема доказана. □

Заметим, что, за исключением леммы, доказательство базируется на прямых алгебраических выкладках, т. е. оно остаётся верным во всех случаях, когда верно утверждение леммы (независимо от того, верны ли условия о внешнем касании окружностей и т. п.). Но лемма оказывается верной и при значительном ослаблении сформулированных выше условий. А именно, если горизонтальная проекция отрезка XY' равна $2kp$, то вертикальная равна (с правильным знаком) $k(1-p^2)$ в следующих случаях:

(i) если одна окружность касается ω снизу, другая касается ω' сверху и окружности касаются внешним образом;

(ii) если обе окружности касаются соответственно ω и ω' сверху и касаются внутренним образом;

(iii) если обе окружности касаются соответственно ω и ω' снизу и касаются внутренним образом.

Проверку случаев (ii) и (iii) оставляем читателю; наше доказательство, разумеется, соответствовало случаю (i). Поскольку при условии справедливости леммы доказательство теоремы сводится к алгебраическим выкладкам, её можно сформулировать следующим образом.

Теорема 3.2$'$. *Если a, b', c, a', b, c' — замкнутая цепочка окружностей, поочерёдно касающихся параллельных прямых ω и ω' в точках A, B', C, A', B, C' соответственно, причём для каждой пары соседних окружностей выполнено какое-либо из условий* (i), (ii), (iii), *то прямые AA', BB', CC' пересекаются в одной точке.*

На рис. 11 пары (b, c') и (b', c) удовлетворяют условию (i), пары (a, b') и (a, c') — условию (ii), пары (a', b) и (a', c) — условию (iii).

В заключение этого пункта рассмотрим не лишённый изящества рис. 12. С одной стороны, это частный случай доказанной теоремы, поэтому прямые AA', BB', CC' пересекаются в некоторой точке X. С другой стороны, это частный случай теоремы о семи окружностях, поэтому прямые BB', DD', EE' пересекаются в некоторой точке Y. В качестве упражнения предлагаем читателю показать, что *точки X и Y совпадают*. (Подсказка: пусть K и K' — середины BC и $B'C'$, тогда KK' пересекается с BB' и в точке X, и в точке Y.)

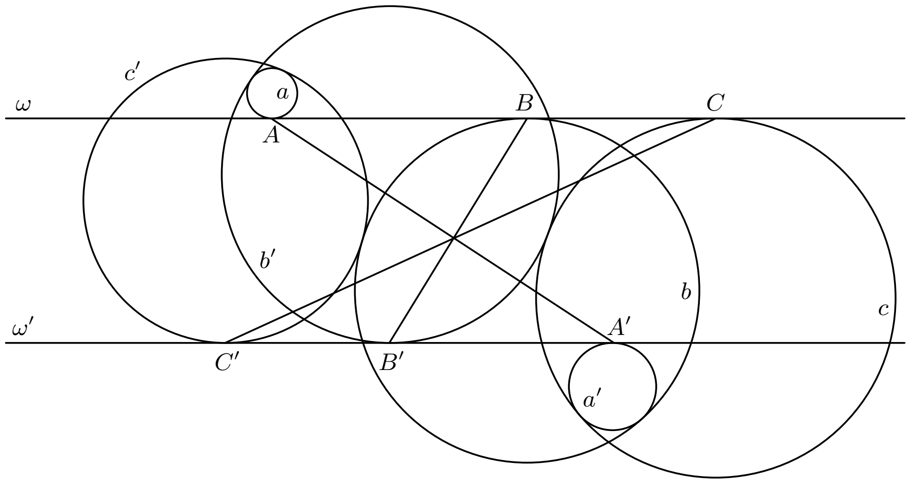


Рис. 11. Одна из конфигураций в теореме 3.2 о двух параллельных прямых и шести окружностях

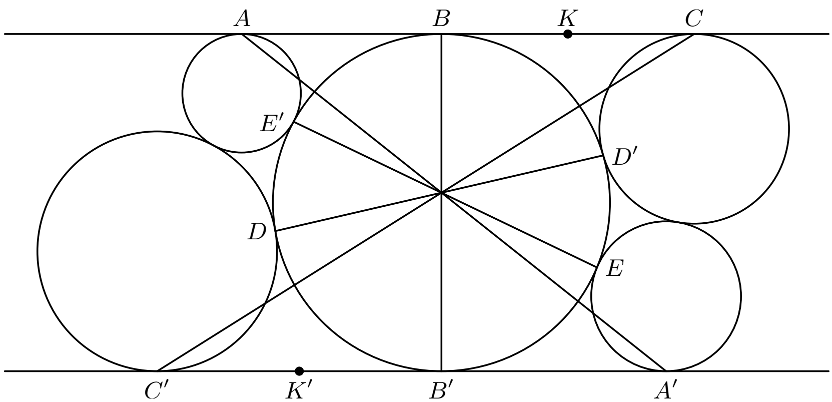


Рис. 12. Частный случай обеих теорем 3.1 и 3.2

3.3. Теорема о треугольнике и шести окружностях

Рассмотрим цепочку окружностей в некотором треугольнике ABC. Начнём с произвольной окружности x, касающейся сторон CA и AB (рис. 13). Построим цепочку, которую назовём *треугольной*, следующим образом:

окружность y касается AB, BC и x,
окружность z касается BC, CA и y,
окружность x' касается CA, AB и z,
окружность y' касается AB, BC и x',
окружность z' касается BC, CA и y',
окружность x'' касается CA, AB и z'.

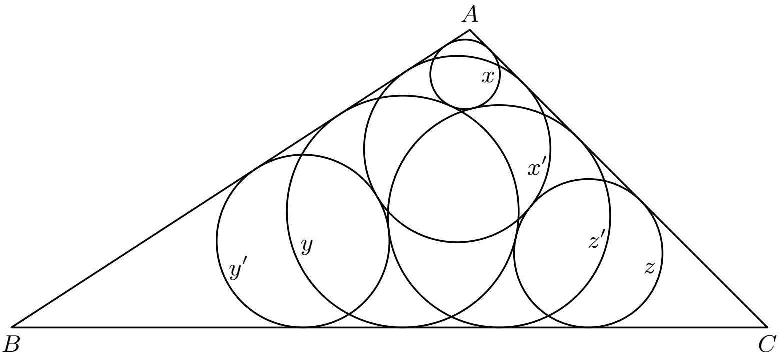


Рис. 13. Треугольная цепочка

Конечно, этих условий недостаточно, чтобы однозначно задать окружности $y, z, \ldots, x''$; на каждом шаге для этого есть несколько вариантов. Однако, как будет показано в этом пункте, *можно на каждом шаге осуществлять выбор так, что окружность x'' совпадёт с x, т. е. мы получим замкнутую треугольную цепочку из шести окружностей.*

Нетрудно заметить тесную связь этого утверждения с классической задачей Мальфатти, которую можно сформулировать так: *найти окружность x, которая совпадает с соответствующей окружностью x'*. Однако если задача Мальфатти имеет лишь конечное число решений [1)], то наша цепочка окружностей замыкается для любой окружности x.

Один из способов ограничить выбор на каждом шаге заключается в требовании, чтобы все окружности лежали *внутри треугольника ABC*. Легко понять, что тогда на каждом шаге существует *не более одной подходящей окружности* [2)], причём получается самая естественная конфигурация из возможных в теореме. Поэтому мы начнём с доказательства теоремы для этого случая, а другие разберём позже.

Теорема 3.3. *Если x, y, z, x', y', z', x'' — цепочка окружностей, вписанных в углы треугольника ABC и лежащих внутри него, то x'' совпадает с x.*

Доказательство. Окружности лежат внутри треугольника, поэтому (i) центры окружностей лежат на соответствующих *внутренних* биссектрисах треугольника; (ii) все касания в цепочке *внешние*.

[1)] Решение задачи Мальфатти методом, похожим на приведённый в этом пункте, см. в [3]. — *Прим. перев.*

[2)] Впрочем, такая цепочка может вообще не существовать. Например, если угол B тупой, а окружность x достаточно мала, то не существует окружности y, лежащей внутри треугольника.

Как обычно, обозначим через a, b, c длины сторон BC, CA, AB, и пусть l, m, n, l', m', n', l'' — длины касательных к x, y, z, x', y', z', x'' из A, B, C, A, B, C, A соответственно. Идея доказательства состоит в том, чтобы получить соотношения между соседними величинами l, m, n, l', m', n', l'' и, исключив из этих соотношений m, n, l', m', n', вывести равенство $l = l''$, из которого, очевидно, будет следовать утверждение теоремы.

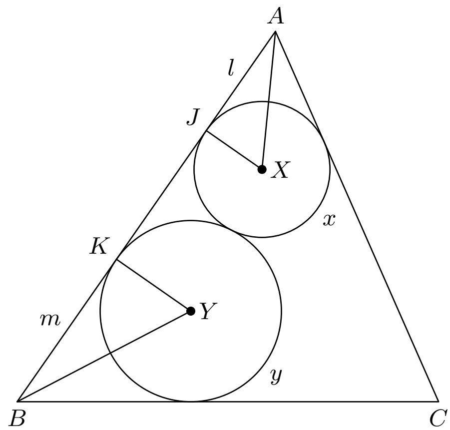


Рис. 14. *К теореме 3.3 о треугольнике и шести окружностях*

Выведем для начала соотношение между l и m. На рис. 14 изображены окружности x и y, их центры X и Y, а также точки J и K их касания со стороной AB. По определению $AJ = l$ и $BK = m$. Так как X лежит на внутренней биссектрисе угла A, радиус окружности x равен $JX = l\operatorname{tg}(A/2)$. Аналогично радиус окружности y равен $KY = m\operatorname{tg}(B/2)$. Как отмечено в п. 3.1, общая касательная окружностей с радиусами r_1, r_2, касающихся внешним образом, равна $2\sqrt{r_1 r_2}$; следовательно,

$$JK = 2\sqrt{lm\operatorname{tg}\frac{A}{2}\operatorname{tg}\frac{B}{2}}.$$

Поскольку $AB = AJ + JK + BK$, выполнено равенство

$$c = l + m + 2\sqrt{lm\operatorname{tg}\frac{A}{2}\operatorname{tg}\frac{B}{2}}. \tag{1}$$

Теперь, используя известные формулы, получаем, положив $s = (a+b+c)/2$:

$$\operatorname{tg}\frac{A}{2}\operatorname{tg}\frac{B}{2} = \sqrt{\frac{(s-b)(s-c)}{s(s-a)}}\sqrt{\frac{(s-a)(s-c)}{s(s-b)}} = \frac{s-c}{s}.$$

Поэтому (1) приводится к виду

$$c = l + m + 2\sqrt{lm}\sqrt{1 - \frac{c}{s}}, \tag{2}$$

что и даёт нужное соотношение между l и m.

Аналогично получаем, что m и n связаны равенством

$$a = m + n + 2\sqrt{mn}\sqrt{1 - \frac{a}{s}},$$

и т. д., всего имеем шесть равенств.

Чтобы исключить из этих равенств m, n, l', m', n', воспользуемся следующим любопытным приёмом. Обозначим через p, q, r, p', q', r', p'' значения квадратных корней из l, m, n, l', m', n', l'', а через f, g, h и t — значения корней из a, b, c и s. Определим также тупые углы θ, φ, ψ следующими равенствами:

$$\begin{aligned} \cos\theta &= -\sqrt{1 - \frac{a}{s}} = -\sqrt{1 - \frac{f^2}{t^2}}, \quad \sin\theta = \frac{f}{t}, \\ \cos\varphi &= -\sqrt{1 - \frac{b}{s}} = -\sqrt{1 - \frac{g^2}{t^2}}, \quad \sin\varphi = \frac{g}{t}, \\ \cos\psi &= -\sqrt{1 - \frac{c}{s}} = -\sqrt{1 - \frac{h^2}{t^2}}, \quad \sin\psi = \frac{h}{t}. \end{aligned} \tag{3}$$

Это возможно, поскольку подкоренные выражения лежат между 0 и 1. Теперь (2) можно записать в виде

$$h^2 = p^2 + q^2 - 2pq\cos\psi, \tag{4}$$

а пять других равенств приводятся к аналогичному виду с заменой h, p, q, ψ на соответствующие величины. Поскольку наша цель — доказать, что $l = l''$, исключим из (4) и пяти аналогичных равенств переменные q, r, p', q', r' и покажем, что $p = p''$.

Уравнение (4) выражает теорему косинусов для треугольника со сторонами p, q, h и углом ψ, противолежащим h; пять остальных соотношений выражают теорему косинусов для других пяти треугольников. Два из этих шести треугольников изображены на рис. 15а.

Заметим, что радиусы описанных окружностей всех этих треугольников равны. Действительно, по теореме синусов диаметр описанной окружности первого треугольника равен $h/\sin\psi$, что, согласно (3), равно t; значение t получается и для остальных пяти треугольников. Следовательно, если у двух треугольников на рис. 15а совместить стороны, равные q, то их описанные окружности тоже совместятся. На рис. 15б показаны три совмещённых таким образом треугольника. Продолжая этот процесс, можно впи-

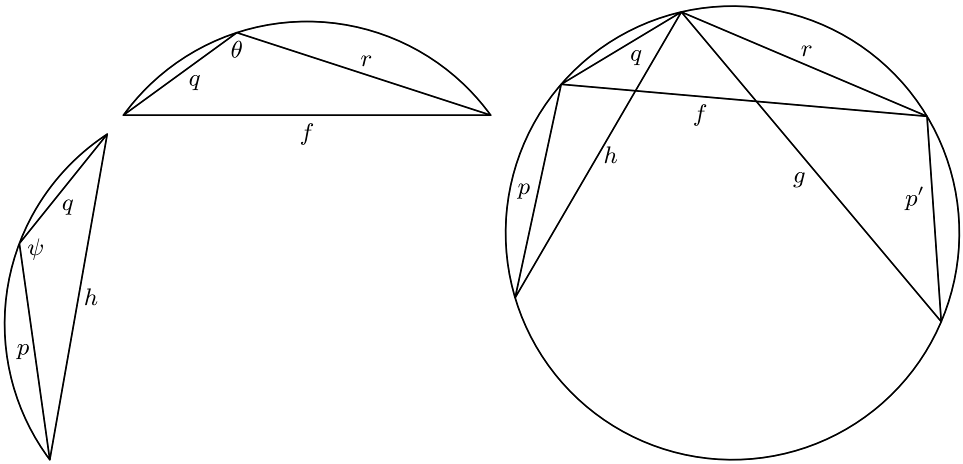


Рис. 15. Совмещение треугольников

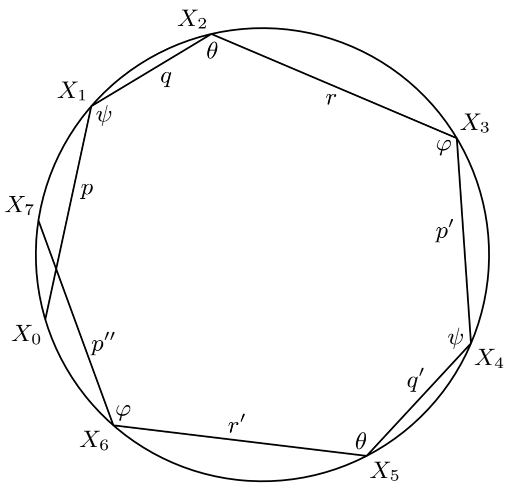


Рис. 16. Описанная окружность шести треугольников

сать в одну окружность все шесть треугольников, как показано на рис. 16 (некоторые отрезки для удобства удалены).

Итак, мы получили вписанную в окружность ломаную $X_0 \ldots X_7$ (рис. 16) с углами ψ в вершинах X_1, X_4, углами θ в вершинах X_2, X_5 и углами φ в вершинах X_3, X_6; наша цель — доказать, что $p = p''$, т. е. $X_0X_1 = X_6X_7$. Для этого заметим, что треугольники $X_0X_2X_4$ и $X_3X_5X_7$ равны, поскольку $X_0X_2 = X_3X_5 = h$, $X_2X_4 = X_5X_7 = g$ и углы, противоположные этим сторонам, опираются на равные дуги. Значит, равны углы $X_0X_2X_4$ и $X_3X_5X_7$. Вычитая эти углы[3)] из углов $X_1X_2X_3$ и $X_4X_5X_6$, равных θ, получим

$$\angle X_0X_2X_1 + \angle X_3X_2X_4 = \angle X_3X_5X_4 + \angle X_6X_5X_7. \tag{5}$$

Поскольку углы $X_3X_2X_4$ и $X_3X_5X_4$ опираются на одну дугу, из (5) имеем

$$\angle X_0X_2X_1 = \angle X_6X_5X_7.$$

Следовательно, равны и соответствующие этим углам хорды X_0X_1 и X_6X_7, т. е. $p = p''$. Теорема 3.3 доказана. □

Вернёмся теперь к описанию других случаев, когда треугольная цепочка замыкается. Ослабим условие, что окружности лежат внутри треугольника, потребовав лишь, чтобы центры окружностей лежали на внутренних биссектрисах его соответствующих углов. Тогда для данной окружности x окружность y можно выбрать тремя различными способами (два из которых при некоторых x приводят к мнимым окружностям).

а) Один из этих способов, который мы назовём *специальным* (он всегда даёт вещественную окружность), существенно отличается от двух остальных: при нём y касается AB в той же точке, что и x (и, следовательно, касается x внутренним образом). Если мы выбрали y специальным способом, то и окружность z надо выбрать специальным способом (касающейся BC в той же точке, что и y), и т. д. Такая цепочка обязательно замыкается (рис. 17).

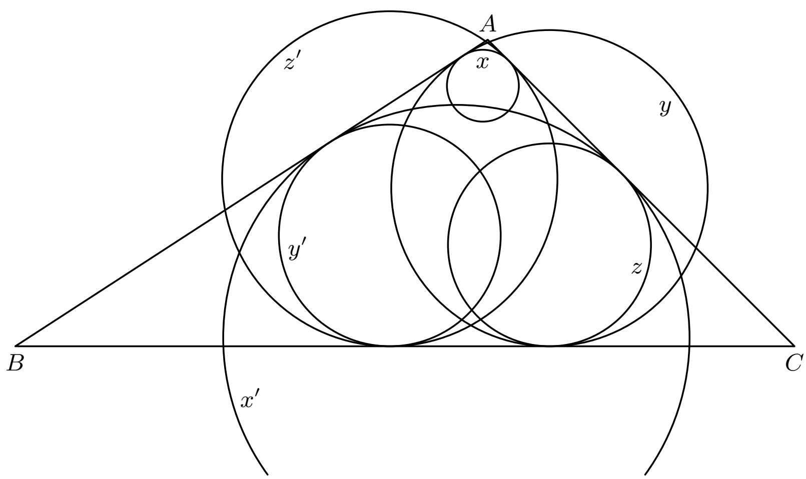


Рис. 17. Специальный выбор окружностей

[3)] В общем случае нужно рассматривать направленные углы. — *Прим. ред.*

Доказательство простое: если, как и раньше, обозначить через l, m, n, l', m', n', l'' длины соответствующих касательных, то

$$l + m = c, \quad m + n = a, \quad n + l' = b,$$
$$l' + m' = c, \quad m' + n' = a, \quad n' + l'' = b.$$

Отсюда $l = l''$, что и требовалось.

б) Два других способа дают вещественную окружность тогда и только тогда, когда x не лежит целиком вне треугольника; при этом, очевидно, y касается x внешним образом. Легко убедиться, что, при сохранении введённых ранее обозначений, m и l по-прежнему удовлетворяют соотношению (2), однако нужно использовать оба знака квадратного корня (соответственно двум возможностям выбора y). Поэтому будет удобно переписать (2) без иррациональностей:

$$s(c - l - m)^2 = 4(s - c)lm; \tag{6}$$

при данном l получаем квадратное уравнение относительно m, имеющее вещественные (лежащие между 0 и s) корни тогда и только тогда, когда l лежит между 0 и s. Это подтверждает сделанное выше утверждение об условиях вещественности y [4)].

Если теперь на каждом шаге построения треугольной цепочки выбирать один из двух неспециальных способов, то после шести шагов мы при данной окружности x получим 64 возможности для окружности x''; наша задача — определить, сколько из этих окружностей совпадают с x. С алгебраической точки зрения, мы имеем уравнение (6) и пять аналогичных соотношений для последующих шагов. Исключая m, n, l', m', n', получаем l'' как 64-значную функцию от l; нужно определить, сколько из этих 64 значений равны l.

Эту задачу можно решить методом, применённым выше для случая окружностей, лежащих внутри треугольника. Однако представляет интерес и альтернативный подход. Заменим l, m, n, l', m', n', l'' новыми переменными λ, μ, ν, λ', μ', ν', λ'':

$$l = s\sin^2\lambda, \quad m = s\sin^2\mu, \quad \ldots, \quad l'' = s\sin^2\lambda'', \tag{7}$$

новые переменные определяются с точностью до знака по модулю π. В новых переменных уравнение (6) приводится к виду $\cos 2\mu = \cos(2\lambda \pm 2\psi)$, где ψ определено [5)] в соответствии с (3). Следовательно,

$$\mu = \pm\lambda \pm \psi \pmod{\pi};$$

[4)] Если x совпадает с вневписанной окружностью, то $l = s$, откуда следует сказанное.

[5)] Уравнение (6) выражает теорему косинусов для треугольника с углами, по модулю π равными $\pm\psi$, $\pm\lambda$, $\pm\mu$, откуда следует нужное соотношение. — *Прим. перев.*

аналогично пять остальных соотношений приводятся к виду

$$\nu = \pm\mu \pm \theta \pmod{\pi}, \quad \ldots, \quad \lambda'' = \pm\nu' \pm \varphi \pmod{\pi}.$$

Таким образом,

$$\lambda'' = \pm\lambda + (\pm\psi \pm \theta \pm \varphi \pm \psi \pm \theta \pm \varphi) \pmod{\pi}, \tag{8}$$

где выбор знаков в правой части соответствует выбору одного из двух способов на каждом шаге построения цепочки. Если удастся выбрать знаки так, что $\lambda'' = \pm\lambda$, то в силу (7) получим, что $l'' = l$. Из (8) следует, что первые три знака можно выбрать произвольно, после чего остальные нужно выбрать так, чтобы слагаемые попарно уничтожились. Поэтому из 64 возможных при данной окружности x неспециальных цепочек замыкаются восемь.

Сформулируем окончательный результат.

Теорема 3.3′. *Пусть I — центр вписанной окружности треугольника ABC, а центры окружностей треугольной цепочки x, y, … лежат на биссектрисах AI, BI, CI. Тогда*:

а) *если на каждом шаге окружность выбирается специальным образом, то цепочка замыкается*;

б) *если на каждом шаге окружность выбирается неспециальным образом, то для любых x, y, z, x' можно выбрать y', z' так, чтобы цепочка замкнулась.*

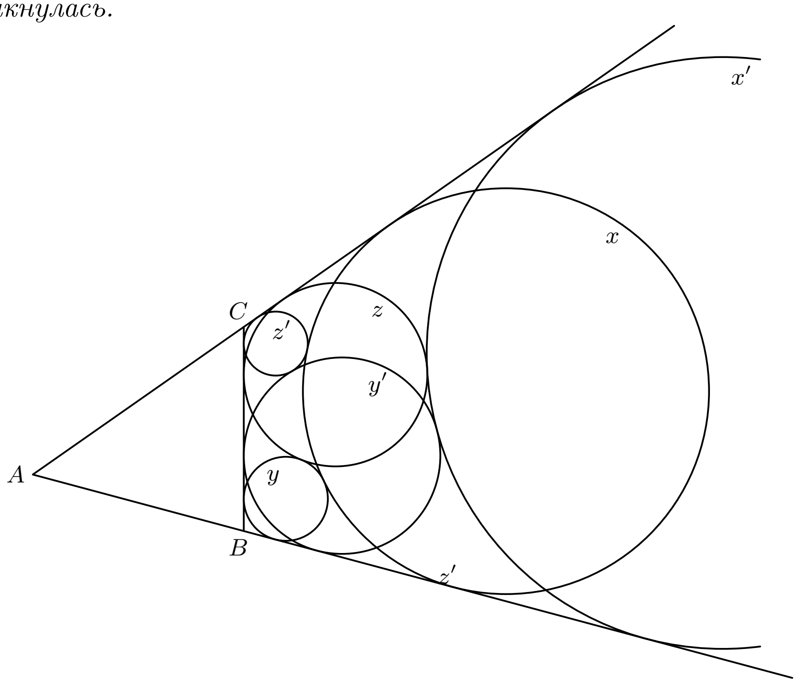


Рис. 18. Цепочка вне треугольника

В заключение отметим, что утверждение теоремы остаётся верным, если заменить I на центр одной из вневписанных окружностей. Пример такой конфигурации изображён на рис. 18.

3.4. Теорема о девяти окружностях

В этом пункте будет описана последняя цепочка. Она является обобщением цепочки из предыдущего пункта и получается из неё заменой сторон треугольника тремя окружностями. А именно, пусть a, b, c — три окружности на плоскости, а x — окружность, касающаяся b и c. Построим цепочку следующим образом:

окружность y касается c, a и x,
окружность z касается a, b и y,
окружность x' касается b, c и z,
окружность y' касается c, a и x',
окружность z' касается a, b и y',
окружность x'' касается b, c и z'.

Как и в случае треугольной цепочки, на каждом шаге есть несколько способов выбора окружности. Покажем, что при подходящем выборе (что это значит, объяснено ниже) окружность x'' совпадает с x.

На рис. 19 и 20 изображены два примера цепочек. Можно заметить, что девять окружностей a, b, c, x, y, z, x', y', z' образуют симметричную

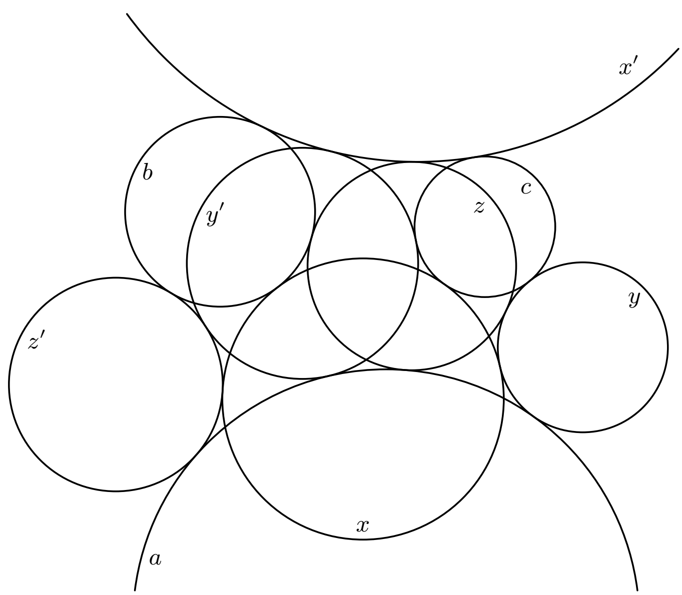


Рис. 19. Пример неспециального выбора

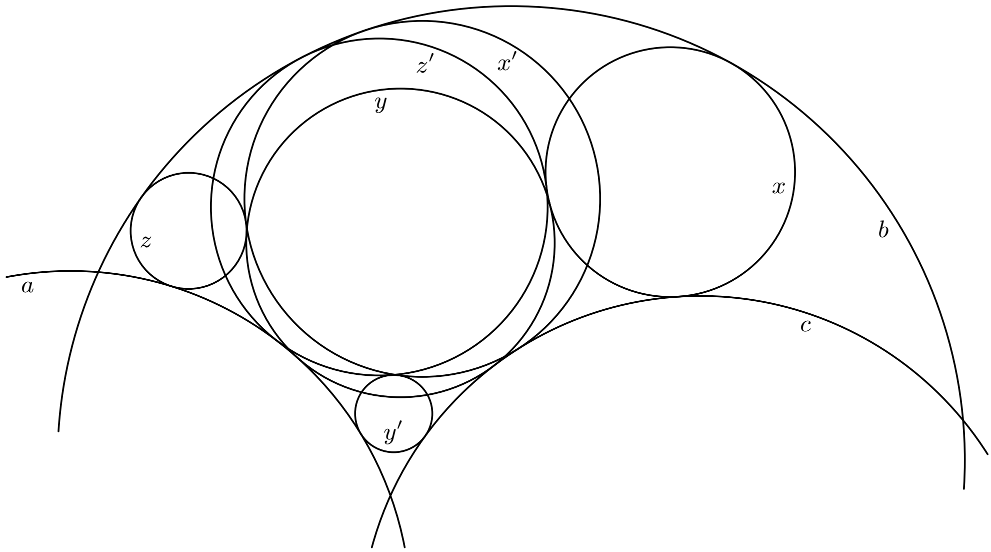


Рис. 20. Пример неспециального выбора

систему, каждая окружность касается четырёх других. Эту симметрию можно описать следующей таблицей:

$$\begin{array}{ccc} a & b & c \\ x & y' & z \\ x' & y & z' \end{array}$$

Две окружности касаются, если (и в общем случае — только если) они не находятся ни в одной строке, ни в одном столбце.

Опишем теперь упомянутый выше «подходящий выбор». Для начала напомним следующие факты.

(i) Для трёх окружностей общего положения на плоскости существует восемь касающихся их окружностей. Если две из исходных окружностей касаются (этот случай возникает на каждом шаге построения нашей цепочки), число таких окружностей уменьшается до шести. Однако две из них обычно засчитываются дважды, так что общее число остаётся равным восьми. В частности, в каждом семействе соосных окружностей, касающихся x и c, найдутся две окружности, которые также касаются a, и они задают два возможных выбора окружности y. Как и в предыдущем пункте, такой способ выбора y назовём *специальным* (таких окружностей y две, и они засчитываются дважды) четыре остальных выбора окружности y будем называть *неспециальными*. Аналогично будем различать специальные и неспециальные способы выбора окружности на остальных шагах построения цепочки.

(ii) Если окружность s касается двух данных окружностей, то прямая, соединяющая точки касания, проходит через один из двух центров гомотетии этих окружностей. Будем говорить, что s *принадлежит* этому центру. Из шести возможных способов выбора окружностей y три (один специальный и два неспециальных) принадлежат одному центру и три — другому.

(iii) Напомним, что центры гомотетии трёх окружностей образуют четыре тройки коллинеарных точек, т. е. являются шестью вершинами полного четырёхсторонника.

Приведём точную формулировку теоремы о девяти окружностях.

Теорема 3.4. *Пусть a, b, c — три окружности на плоскости. Зафиксируем три коллинеарных центра их гомотетии* (*по одному для каждой пары окружностей*). *Будем строить цепочку окружностей $x, y, \ldots,$ принадлежащих* (см. выше (ii)) *выбранным центрам. Тогда*:

а) *если на каждом шаге окружность выбирается специальным образом, то цепочка замыкается*;

б) *если на каждом шаге окружность выбирается неспециальным образом, то для любых окружностей x, y, z, x' существуют такие окружности y', z', что цепочка замыкается.*

Примеры неспециального выбора изображены на рис. 19 и 20, специального — на рис. 21.

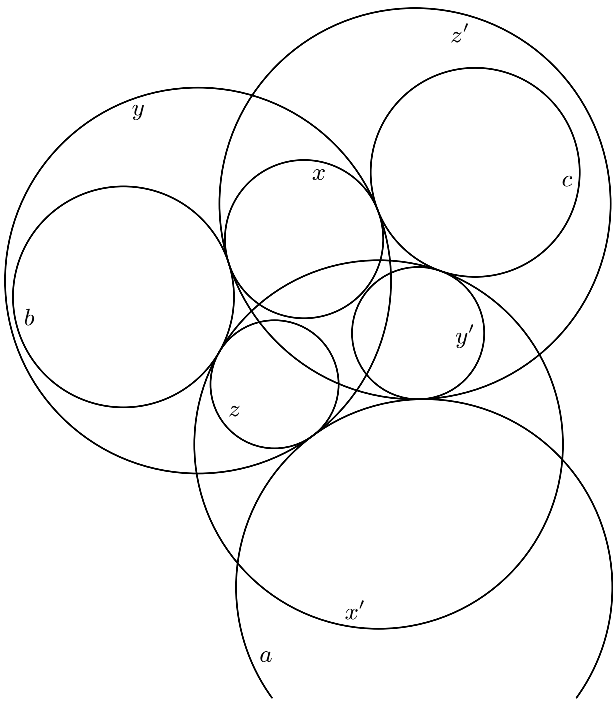


Рис. 21. Пример специального выбора

Доказательство. Единственное известное нам полное доказательство теоремы о девяти окружностях использует методы трёхмерной комплексной проективной геометрии и опубликовано в [1]. Здесь мы приведём доказательство, которое, хотя и не является полным, охватывает широкий класс случаев. А именно, мы докажем теорему: а) для случая трёх равных окружностей a, b, c и б) для случая, когда три выбранных центра гомотетии являются бесконечно удалёнными точками. Эти предположения не столь ограничительны, как может показаться, поскольку три данные окружности «как правило» можно инверсией перевести в равные окружности («как правило» означает «если некоторые кривые пересекаются в вещественных точках»).

Рассмотрим три равных окружности a, b, c с центрами A, B, C (рис. 22). Можно считать, что их радиусы равны 1. Пусть O — центр окружности, проходящей через точки A, B, C, а ρ — её радиус, так что $OA = OB = OC = \rho$. Положим также $\angle BOC = 2\alpha$, $\angle COA = 2\beta$, $\angle AOB = 2\gamma$, так что $\alpha + \beta + \gamma = \pi$.

Пусть теперь окружность x принадлежит бесконечно удалённому центру гомотетии b и c (это значит, что x касается b и c одинаковым образом,

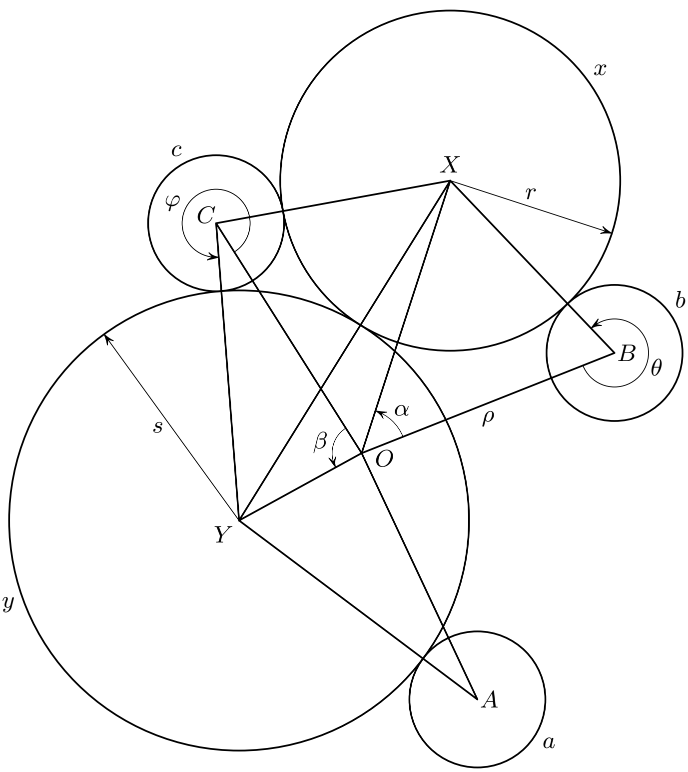


Рис. 22. К теореме о девяти окружностях. Случай равных окружностей

т. е. оба касания внешние или оба внутренние) и это правило соблюдается для всех окружностей цепочки. На рис. 22 изображены окружность x с центром X и радиусом r (с внешними касаниями) и выбранная неспециальным образом окружность y с центром Y и радиусом s. Обозначим через θ ориентированный угол между BO и BX и будем рассматривать θ как параметр, определяющий x. Аналогично будем задавать окружность y ориентированным углом φ между CO и CY, а окружности z, x', y', z', x'' — не показанными на рис. 22 углами ψ, θ', φ', ψ', θ''. Тогда нам надо показать, что $\theta'' = \theta \pmod{2\pi}$.

Сначала рассмотрим случай (не показанный на рис. 22), когда на каждом шаге делается специальный выбор. Тогда y касается c в той же точке, что и x. В этом случае, так как $\angle XBO = \angle XCO$ в силу симметрии, получаем (с учётом ориентации), что $\varphi = 2\pi - \theta$. Аналогично $\psi = 2\pi - \varphi$, $\theta' = 2\pi - \psi$, $\ldots$, $\theta'' = 2\pi - \psi'$ и, значит, $\theta'' = \theta$, что и требуется.

Оставшаяся часть доказательства посвящена исключительно неспециальному случаю, и первым её шагом будет вывод соотношения между θ и φ. Прежде всего заметим, что в силу симметрии X лежит на серединном перпендикуляре к BC и, значит, $\angle BOX = \alpha$. Поэтому

$$\angle OXB = \pi - \angle BOX - \angle OBX = \theta - \alpha - \pi,$$

и мы, применив к треугольнику BOX теорему синусов, получаем

$$\frac{r+1}{\sin\alpha} = \frac{\rho}{\sin(\theta-\alpha-\pi)}, \quad \text{т. е.} \quad \frac{1}{r+1} = \frac{\sin(\alpha-\theta)}{\rho\sin\alpha}. \tag{1}$$

Аналогично из треугольника COY находим

$$\frac{1}{s+1} = \frac{\sin(\beta-\varphi)}{\rho\sin\beta}. \tag{2}$$

Применяя к треугольнику XCY теорему косинусов, получаем, поскольку $\angle XCY = 4\pi - \theta - \varphi$:

$$(r+s)^2 = (r+1)^2 + (s+1)^2 - 2(r+1)(s+1)\cos(4\pi-\theta-\varphi),$$

или, после упрощения,

$$\sin^2\frac{\theta+\varphi}{2} = \Big(1 - \frac{1}{r+1}\Big)\Big(1 - \frac{1}{s+1}\Big).$$

Откуда с учётом (1) и (2) получаем

$$\sin^2\frac{\theta+\varphi}{2} = \Big(1 - \frac{\sin\xi\sin(\alpha-\theta)}{\sin\alpha}\Big)\Big(1 - \frac{\sin\xi\sin(\beta-\varphi)}{\sin\beta}\Big), \quad \text{где } \sin\xi = \rho^{-1}. \tag{3}$$

(Здесь 2ξ — угол между касательными из O к любой из окружностей a, b, c.)

Уравнение (3) является основным соотношением, связывающим[6)] параметры θ и φ окружностей x и y. Если θ задано, то существуют два значения φ, удовлетворяющие (3), что соответствует наличию двух неспециальных окружностей y для данной окружности x. Аналогично, заменив в (3) θ, φ, α и β на φ, ψ, β и γ соответственно, мы получим соотношение, связывающее параметры φ и ψ окружностей y и z; продолжая циклически менять буквы, мы получим аналогичные (3) соотношения, из которых нужно исключить φ, ψ, θ', φ', ψ', для вывода равенства $\theta'' = \theta \pmod{2\pi}$.

Детальное изложение этих выкладок было бы слишком длинным, поэтому ограничимся описанием основного приёма. Как уже было отмечено, уравнение (3) задаёт взаимно двузначное соответствие между θ и φ и, следовательно, может быть разрешено в эллиптических функциях. Соответствующее ему в п. 3.3 уравнение (6) удалось решить тригонометрическими и рациональными заменами, что позволило сравнительно просто доказать теорему о треугольнике. Однако данное уравнение не относится к такому классу и действительно требует применения эллиптических функций или чего-то эквивалентного. Тем не менее, попытаемся введением новых переменных вместо $\theta, \varphi, \ldots$ свести (3) к некоторой стандартной формуле теории эллиптических функций. Опуская детали, опишем результат такой замены.

Введём следующие обозначения:

$$\lambda = \sqrt{\frac{\sin(\alpha - \xi)}{\sin(\alpha + \xi)}}, \quad \mu = \sqrt{\frac{\sin(\beta - \xi)}{\sin(\beta + \xi)}}, \quad \nu = \sqrt{\frac{\sin(\gamma - \xi)}{\sin(\gamma + \xi)}}.$$

Отметим, что λ, μ, ν не обязаны быть вещественными. При этом несущественно, какое значение корня выбирается в каждом из трёх случаев. Положим также $k = \lambda\mu\nu$. Введём теперь параметр t посредством равенства

$$\frac{t - 5 - k^2}{t + 1 + 5k^2} = \frac{\lambda}{k} \cdot \frac{\sin\dfrac{\theta + \xi - \pi}{2}}{\sin\dfrac{\theta - \xi + \pi}{2}}, \tag{4}$$

а параметры $u, v, \ldots$, связанные с $\varphi, \psi, \ldots$, — аналогичными формулами, в которых λ заменяем на $\mu, \nu, \ldots$ соответственно. Уравнение (3) примет вид

$$(t + u + f)(4tuf - g_3) = \left(tu + tf + uf - \tfrac{1}{2}g_2\right)^2, \tag{5}$$

где $g_2 = 12(1 - 14k^2 + k^4)$, $g_3 = 8(1 + 33k^2 - 33k^4 - k^6)$ и

$$f = \frac{(k^2 + 5)\lambda\mu - (5k^2 + 1)}{\lambda\mu + 1}.$$

[6)] Следует заметить, что при других расположениях x и y доказательство требует некоторой модификации, но в итоге опять будет получено равенство (3).

Теперь уравнение (5) выражает хорошо известную форму теоремы сложения для эллиптической функции Вейерштрасса ρ (см., например, [2, с. 364]) и может быть записано в виде

$$\rho^{-1}(u) = \pm\rho^{-1}(t) \pm \rho^{-1}(f) \quad \text{(по модулю периодов)}, \tag{6}$$

где ρ^{-1} — функция, обратная ρ (при этом ρ^{-1} — двузначная функция по модулю периодов с двумя противоположными, равными по абсолютной величине значениями в каждой точке; поэтому для данного значения t в правой части (6) получаем четыре значения, разбивающиеся на две пары противоположных, т. е. два значения для u).

Уравнение (6) даёт простое выражение параметра u окружности y через параметр t окружности x. Аналогичная формула выражает параметр v окружности z через u; при этом g_2 и g_3 задаются теми же формулами, поскольку они зависят только от k, а функция k симметрична по α, β, γ. В итоге соотношение между u и v имеет вид

$$\rho^{-1}(v) = \pm\rho^{-1}(u) \pm \rho^{-1}(d) \quad \text{(по модулю периодов)}, \tag{7}$$

где ρ — та же эллиптическая функция, что и выше, а d — некоторая новая константа. Точно так же получаем, что параметр t' окружности x' выражается через v соотношением

$$\rho^{-1}(t') = \pm\rho^{-1}(v) \pm \rho^{-1}(e) \quad \text{(по модулю периодов)} \tag{8}$$

с некоторой новой константой e. Объединив (6), (7) и (8), получаем

$$\rho^{-1}(t') = \pm\rho^{-1}(t) \pm \rho^{-1}(f) \pm \rho^{-1}(d) \pm \rho^{-1}(e) \quad \text{(по модулю периодов)}, \tag{9}$$

что даёт выражение параметра t' через t. Шестнадцать возможных комбинаций знаков дают для данного t восемь значений t'. Аналогично, сделав три следующих шага, мы получаем выражение параметра t'' окружности x'' как восьмизначной функции от t':

$$\rho^{-1}(t'') = \pm\rho^{-1}(t') \pm \rho^{-1}(f) \pm \rho^{-1}(d) \pm \rho^{-1}(e) \quad \text{(по модулю периодов)}. \tag{10}$$

Очевидно, что при любом выборе знаков в (9) можно выбрать знаки в (10) так, чтобы выполнялось равенство $\rho^{-1}(t'') = \pm\rho^{-1}(t)$, а остальные слагаемые попарно уничтожились. Отсюда $t'' = t$, и из (4) получаем $\theta'' = \theta \pmod{2\pi}$. Теорема доказана. □

Читатель мог заметить, что конфигурации, иллюстрирующие теорему о девяти окружностях, весьма разнообразны и лишь малая часть их была представлена в данной книге. Можно добавить к ним ещё одну теорему, также имеющую множество специальных случаев, соответствующих вырождению некоторых окружностей. Как уже было отмечено, теорема п. 3.3 является вырожденным случаем, возникающим при замене окружностей a, b, c прямыми. Другой специальный случай возникает, когда a, b, c вырождаются в точки; его доказательство весьма элементарно и доступно даже школьнику. Также можно получить ряд специальных случаев, допуская совпадение некоторых окружностей. У нас нет возможности детально разобрать все эти возможности, но интересующийся читатель может сделать это самостоятельно.

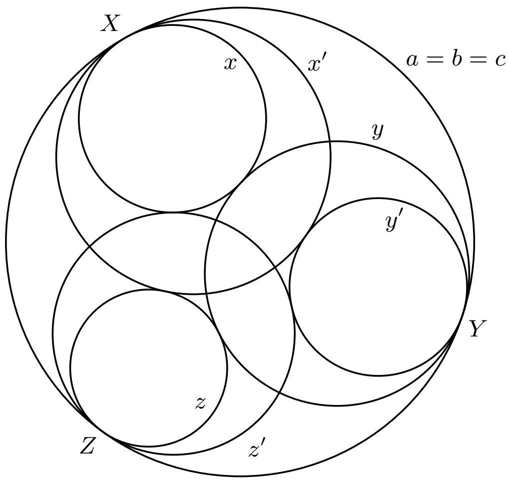


Рис. 23. Предельный случай теоремы о девяти окружностях

Однако мы не можем не остановиться в заключение на одном частном случае теоремы о девяти окружностях, возникающем при совпадении всех трёх окружностей a, b, c. Изобразив соответствующую конфигурацию, можно заметить, что в предельной ситуации верны две совершенно различные теоремы. Одна из них — обращение теоремы о семи окружностях (п. 3.1), а другая утверждает замкнутость цепочки, изображённой на рис. 23, где окружности цепочки касаются внешней окружности в трёх различных точках X, Y, Z. Впрочем, эти результаты значительно проще получаются с помощью инверсии!

Список литературы

[1] *Tyrrell J. A., Powell M. T.* A theorem in circle geometry // Bull. London Math. Soc. 1971. Vol. 3. P. 70–74.

[2] *Copson E. T.* Theory of functions of a complex variable. Oxford, 1935.

Литература, добавленная переводчиком

[3] *Беленький В. З., Заславский А. А.* Решение обобщённой задачи Мальфатти с помощью комплексной (гиперболической) тригонометрии // Математическое просвещение. Сер. 3. Вып. 2. М.: МЦНМО, 1998. С. 141–154.

[4] *Ивлин С. Дж. А., Мани-Каутс Г. Б., Тиррелл Дж. А.* Теорема о семи окружностях и другие новые теоремы: § 2. Обобщения теорем Паскаля и Брианшона // Математическое просвещение. Сер. 3. Вып. 21. М.: МЦНМО, 2017. С. 47–63.

Сесил Джон Алвин Ивлин (1904–1976)

Годфри Бёрдет Мани-Каутс (1905–1979)

Джон Алфред Тиррелл (1932–1992)

Обобщённая задача Аполлония

Е. А. Морозов

В работе исследуется обобщение известной задачи Аполлония о построении окружности, касающейся трёх данных. Рассматривается вопрос о максимальном возможном числе таких окружностей, в случае если исходных окружностей больше трёх. Доказано, что если не все исходные окружности касаются в одной точке, то в случае четырёх исходных окружностей имеется не более шести решений задачи Аполлония, а в случае пяти исходных окружностей — не более четырёх. Также дано описание всех четвёрок окружностей, для которых количество решений максимально.

§ 1. Введение

Задача (классическая задача Аполлония). Построить циркулем и линейкой окружность, касающуюся каждой из трёх данных.

Так сформулировал задачу известный древнегреческий геометр Аполлоний Пергский в III в. до н. э. С тех пор придумано много остроумных решений первоначальной задачи, однако нас будет интересовать лишь вопрос о максимально возможном числе искомых окружностей. Известно, что классическая задача имеет не более восьми решений (см., например, [1]). Мы рассмотрим обобщение этого вопроса на случаи четырёх и пяти окружностей.

Ключевую роль в наших рассуждениях играет инверсия. Это преобразование плоскости позволяет упростить конфигурацию исходных окружностей и осуществить перебор. Поэтому перед доказательством основных теорем мы подробнее рассмотрим конструкции, к которым нас приведёт это преобразование (см. § 3). Также мы используем один из результатов полного перебора всех случаев классической задачи Аполлония, описанного в [1].

Среди приёмов, которые мы используем для доказательства различных случаев основной теоремы, будут: подсчёт количества точек пересечения окружностей, использование их взаимного расположения на плоскости, а также факт об объединении решений классической задачи в пары, использующий алгебраическую интерпретацию окружности, — об этом подробнее в § 5.

§ 2. Формулировки основных результатов

Обобщённой окружностью (*объектом*) назовём окружность, прямую или точку на плоскости.

Две обобщённые окружности *обобщённо касаются*, если это одна из следующих конфигураций:

1) две касающиеся окружности или касающиеся окружность и прямая;
2) две параллельные прямые;
3) точка, лежащая на окружности или прямой.

Так как все касания у нас будут обобщёнными, то для краткости вместо «обобщённо касаются» будем говорить просто «касаются».

Заметим, что посредством стереографической проекции плоскость с добавленной к ней бесконечно удалённой точкой ∞ можно перевести в сферу. При этом преобразовании обобщённая окружность на плоскости перейдёт в окружность или точку на сфере, а касающиеся обобщённые окружности на плоскости — в касающиеся окружности или точку и проходящую через неё окружность на сфере. Поэтому фактически нами рассматривается сфера и её непустые пересечения с плоскостями. В дальнейшем мы будем считать, что бесконечно удалённая точка на плоскости присутствует — в частности, через эту точку проходят все прямые и сама она является обобщённой окружностью. Эту точку можно вернуть на видимую плоскость с помощью инверсии, которая меняет её местами со своим центром. Строгие определения и некоторые теоремы см. в [5].

Теперь можно сформулировать основные результаты.

Теорема 1. *На плоскости даны четыре различные обобщённые окружности, которые не все касаются в одной точке. Тогда существует не более шести обобщённых окружностей, касающихся каждой из данных.*

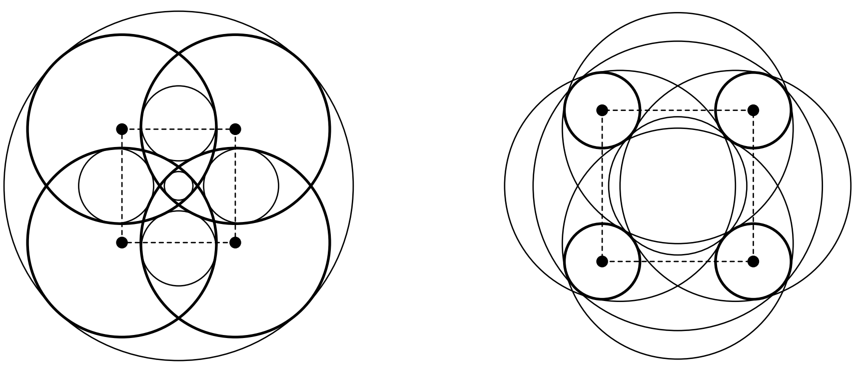


Рис. 1. Примеры к теореме 1

Теорема 2. *На плоскости дано пять различных обобщённых окружностей, которые не все касаются в одной точке. Тогда существует не более четырёх обобщённых окружностей, касающихся каждой из данных.*

Нетрудно найти примеры, когда количество окружностей, указанное в теоремах 1 и 2, действительно достигается. Соответствующие примеры для четырёх окружностей приведены на рис. 1. В этих примерах исходные окружности имеют равный радиус, а их центры лежат в вершинах квадрата. Обоснование примеров очевидно и не нуждается в комментариях.

Из шести окружностей на рис. 1, касающихся каждой из четырёх данных, выберем любые пять. Рассматривая эти пять окружностей в качестве исходных, получаем пример к теореме 2.

Итак, найти какие-то примеры, подтверждающие точность теоремы 1, нетрудно. Но *все* эти примеры допускают единое описание!

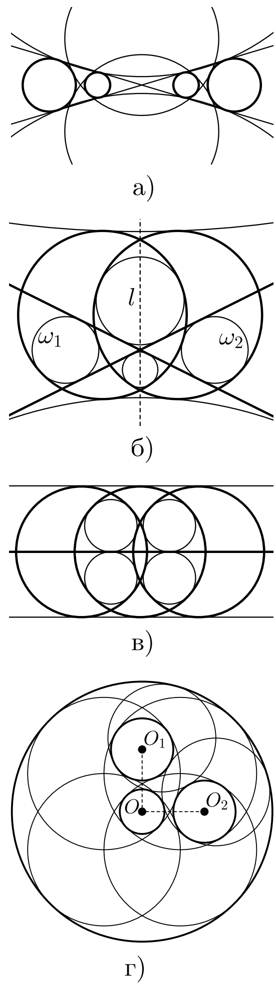


Рис. 2

Теорема 3. *Пусть конфигурация четырёх обобщённых окружностей на плоскости такова, что существует ровно шесть обобщённых окружностей, касающихся каждой из данных. Тогда несколькими инверсиями её можно свести к одной из следующих четырёх.*

1. *Исходные окружности — две непересекающиеся окружности, вписанные в угол, и две симметричные им относительно вершины этого угла* (рис. 2а).

2. *Возьмём непересекающиеся окружности ω_1 и ω_2, симметричные относительно прямой l. Исходные обобщённые окружности — две окружности, касающиеся ω_1 и ω_2 и симметричные друг другу относительно l, и две общие внутренние касательные к ω_1 и ω_2* (рис. 2б).

3. *Исходные обобщённые окружности — прямая и три окружности радиуса 1 с центрами на этой прямой, причём расстояние между соседними центрами равно $2/\sqrt{5}$* (рис. 2в).

4. *Возьмём две концентрические окружности с центром O и отношением радиусов $3+2\sqrt{2}$ — это две из исходных окружностей. Рассмотрим четыре окружности с центрами в вершинах квадрата, касающиеся обеих концентрических, причём так, что одна из концентрических содержит их, а другая*

в них содержится. Две оставшиеся исходные окружности с центрами O_1 *и* O_2 *касаются этих четырёх, причём* $\angle O_1OO_2 = 90^\circ$ (рис. 2г).

Примеры на рис. 1 соответствуют конфигурациям 1 и 2.

§ 3. Обозначения, договорённости и используемые конструкции

Договоримся о некоторых общих обозначениях. Данные в (обобщённой) задаче Аполлония объекты назовём *исходными*, а объект, касающийся всех данных, назовём *решением*. Назовём конфигурацию обобщённых окружностей *вырожденной*, если какие-то три из них касаются в одной точке. Вырожденная тройка объектов, очевидно, имеет бесконечно много решений.

Мы пользуемся следующими известными свойствами инверсии (о них см., например, [5]).

1. Пару непересекающихся окружностей или непересекающиеся окружность и прямую можно перевести инверсией в концентрические окружности.
2. Пару пересекающихся в двух точках окружностей или пересекающиеся в двух точках окружность и прямую можно перевести инверсией в пересекающиеся прямые.
3. Пару касающихся окружностей или касающиеся окружность и прямую можно перевести инверсией в параллельные прямые.
4. Если точка лежит вне окружности, то существует инверсия с центром в этой точке, оставляющая окружность на месте.
5. Если точка лежит внутри окружности, то существует композиция инверсии и центральной симметрии с общим центром в этой точке, оставляющая окружность на месте.

Если два объекта из трёх фиксированы, а третий α меняется, то соответствующие решения для всех трёх объектов мы называем *решениями для* α (или *решениями, порождёнными* α), а объект, касающийся только двух фиксированных, мы называем *допустимым*.

Рассмотрим тройку объектов, каждый из которых является окружностью или прямой. Для описания расположения такой тройки объектов на плоскости удобно использовать так называемые *метки Фитцджеральда*. Рассматриваемой тройке объектов сопоставляется набор букв, характеризующий их отношения: буква I используется для обозначения каждой пары пересекающихся объектов, буква T — для обозначения каждой пары касающихся объектов, а буква S — для обозначения того, что некоторая окружность или прямая разделяет два других объекта. Если ни одно из перечисленных выше отношений не имеет места, то мы говорим о пустой метке ($\varnothing$). Если все три объекта имеют общую точку, то запись заключается в квадратные скобки. Таким образом, например, метке $[TTT]$ соответствует вырожденная тройка объектов.

Известно (см. например [4]), что если три окружности или прямые попарно пересекаются, но не проходят через одну точку (это соответствует метке III), то возможны два случая: либо каждая обобщённая окружность разделяет точки пересечения двух других, либо нет. Обозначим эти случаи разными метками — III_1 и III_2 соответственно. Если три окружности попарно пересекаются и проходят через одну точку, то также возможно два случая: либо эти окружности имеют ровно одну общую точку, либо две. Эти случаи тоже обозначим разными метками $[III]_1$ и $[III]_2$ соответственно.

Также нам потребуются некоторые определения, связанные с тремя простейшими конструкциями из обобщённых окружностей. В каждой из них мы фиксируем два объекта и меняем третий.

3-1. Зафиксированы две пересекающиеся исходные прямые. Будем называть четыре части, на которые две пересекающиеся прямые разбивают плоскость, *секторами*. Секторы пронумеруем S_{I}, S_{II}, S_{III}, S_{IV} по традиции обозначения квадрантов — против часовой стрелки. Точку пересечения прямых обозначим через O.

Каждая окружность, касающаяся двух прямых, лежит в одном из секторов. Эти окружности — допустимые. Назовём *распределением* некоторого набора допустимых окружностей ненулевого радиуса упорядоченную четвёрку чисел x–y–z–t, где число на i-м месте обозначает, сколько окружностей вписано в i-й сектор. При необходимости мы указываем, относительно каких пересекающихся прямых записано распределение. Два распределения *изоморфны*, если они совпадают с точностью до переворотов и циклических сдвигов четвёрки x–y–z–t, т. е., например, разбиения 0–1–2–3 и 2–1–0–3 изоморфны. До первого упоминания о нумерации секторов распределения рассматриваются с точностью до изоморфизма.

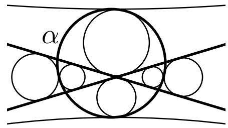
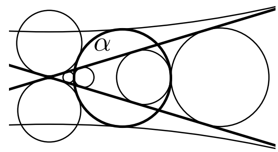
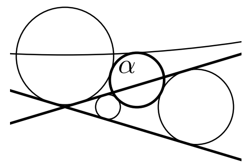
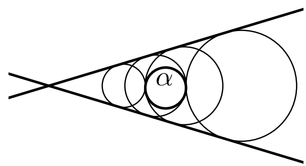
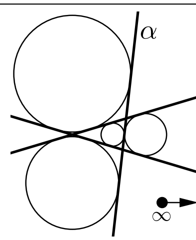
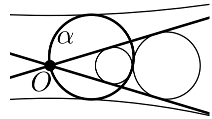
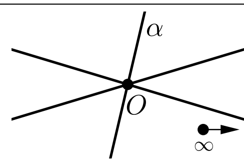
Также решениями могут являться точки O и ∞. Такие решения мы называем *дополнительными*.

Теперь добавим к двум зафиксированным прямым окружность или прямую α. В зависимости от метки этой тройки объектов припишем α некоторый *I-тип*. Посмотрим, каким может быть распределение решений для α.

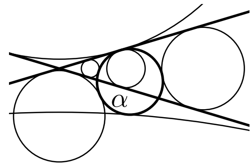
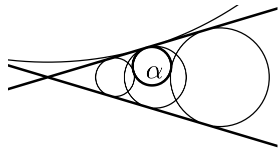
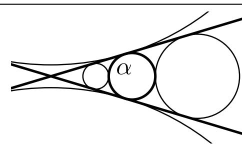
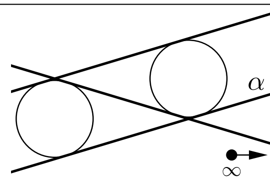
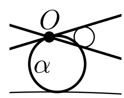
Лемма 3.1. *Даны две пересекающиеся прямые, а также окружность или прямая α, отличная от данных. Тогда, в зависимости от положения α на плоскости, вся тройка имеет одну из меток, указанных в табл.* 1. *Метка однозначно определяет распределение решений для α и количество дополнительных решений* (см. табл. 1).

Таблица 1

I-типы

Тип	Метка	Распределение решений для α	Доп. решения	Всего решений для α	Рисунки
I_1	III_1	2–2–2–2	$\varnothing$	8	
I_2	III_2	4–2–0–2	$\varnothing$	8	
I_3	II	2–2–0–0	$\varnothing$	4	
I_4	I	4–0–0–0	$\varnothing$	4	
I_5	$[III]_1$	2–1–0–1	O или ∞	5	 прямая  окружность
I_6	$[III]_2$	0–0–0–0	O и ∞	2	

Окончание таблицы 1

Тип	Метка	Распределение решений для α	Доп. решения	Всего решений для α	Рисунки
I_7	IIT	3–1–0–2	$\varnothing$	6	
I_8	IT	3–1–0–0	$\varnothing$	4	
I_9	ITT	2–1–0–1	$\varnothing$	4	
I_{10}	$[IIT]$	1–0–0–1	O или ∞	3	 прямая  окружность

Замечание 1. Количество решений указано без учёта кратности.

Замечание 2. Отметим, что окружность типа I_2 всегда пересекает обе стороны некоторого сектора.

Доказательство леммы 3.1 осуществляется одинаково для всех I-типов непрерывным увеличением допустимой окружности в одном из секторов до момента касания с α. Более подробное рассуждение см. в [1]. $\square$

3-2. Зафиксированы две концентрические окружности. Через ω обозначим окружность с меньшим радиусом, а через Ω — с бо́льшим. Общий центр ω и Ω обозначим через O. Допустимыми здесь являются все окружности, касающиеся Ω и ω. Они разбиваются на два рода, которые мы обозначим А и В: окружности *рода* А касаются ω внешним образом, а *рода* B — внутренним. Прямую, соединяющую O с центром окружности α, мы называем *диаметром* (а луч — *радиусом*), на котором лежит α. *Противоположными* называются окружности, лежащие на одном диаметре, но по разные стороны от O.

Теперь снова добавим ещё одну окружность α, но потребуем, чтобы она не имела общих точек с ω и Ω. Найдём количество допустимых окружностей, касающихся α.

Лемма 3.2. *Даны две концентрические окружности и окружность или прямая α, не имеющая общих точек с данными. Тогда полученная тройка либо не имеет решений, либо имеет восемь решений* (как на рис. 3).

Доказательство. Возможны следующие случаи:

1) α лежит внутри ω, тогда ω разделяет α и Ω;
2) α лежит снаружи Ω, тогда Ω разделяет α и ω;
3) α лежит внутри Ω, и ω лежит внутри α, тогда α разделяет ω и Ω;
4) α лежит внутри Ω, но ω не лежит внутри α, и α не лежит внутри ω.

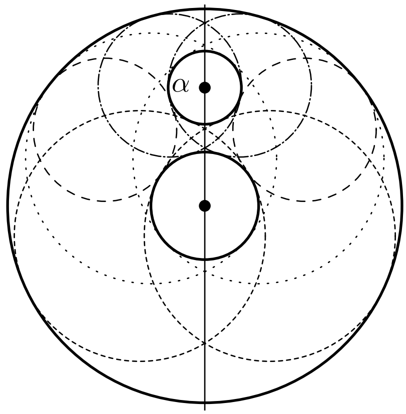


Рис. 3

Случаи 1–3 соответствуют метке S, и тогда решений нет. Случай 4 соответствует метке $\varnothing$, и тогда есть восемь решений, причём четыре из них имеют род A и ещё четыре — род B. Доказательство осуществляется посредством вращения потенциального решения до момента касания с α. □

Обратим внимание, что полученные восемь решений разбиваются на пары симметричных относительно радиуса, на котором лежит α. Следующее наблюдение оказывается весьма полезным в дальнейшем.

Лемма 3.3. *Среди любых пяти решений для α можно выбрать пару пересекающихся, а среди любых шести решений для α можно выбрать пару пересекающихся и симметричных относительно радиуса, на котором лежит α.*

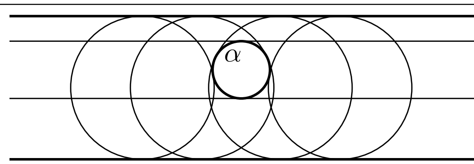
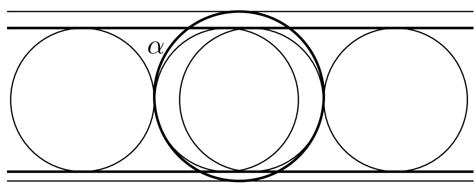
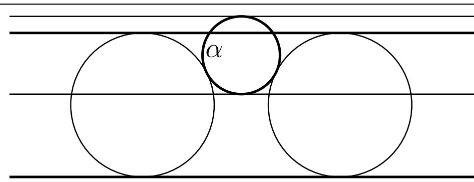
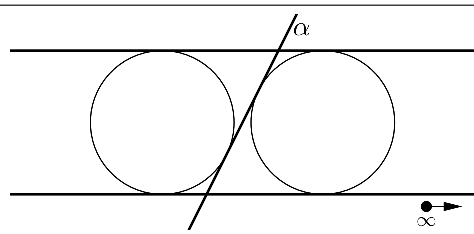
Доказательство. Заметим, что любые две окружности рода B пересекаются, и только в одной паре симметричных решений окружности могут не пересекаться — это пара решений рода A, которые касаются α внешним образом. Отсюда мгновенно следуют оба утверждения леммы. □

3-3. Зафиксированы две параллельные прямые. Будем считать направление фиксированных прямых *горизонтальным*, а перпендикулярное — *вертикальным.* Допустимыми являются окружности равного радиуса, касающиеся фиксированных прямых, а также все горизонтальные прямые и точка ∞. *Распределением* здесь называем упорядоченную пару чисел $x+y$, где x обозначает количество допустимых окружностей, а y — количество допустимых прямых. Решением может быть также точка ∞ — в этом случае она называется *дополнительным решением.*

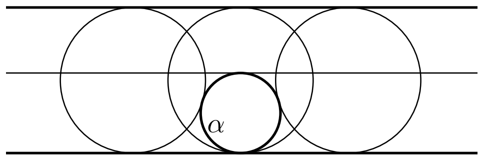
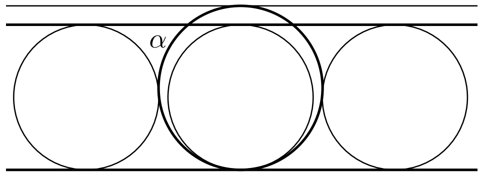
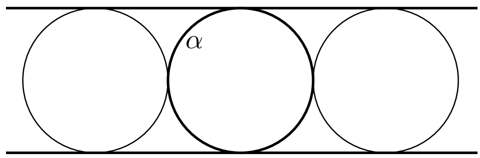
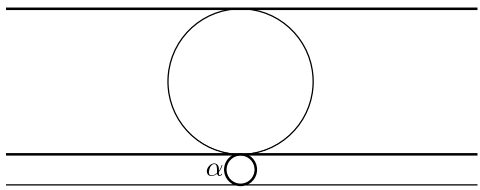
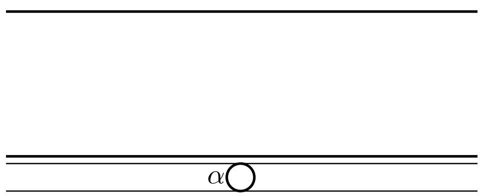
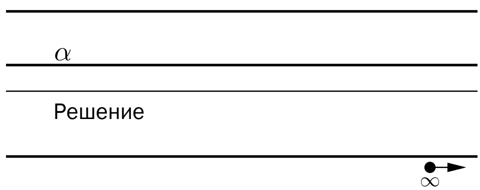
Лемма 3.4. *Даны две параллельные прямые, а также допустимая окружность или прямая α, отличная от данных. Тогда, в зависимости от положения α на плоскости, вся тройка имеет одну из меток, указанных в табл. 2. Кроме того, метка однозначно определяет распределение решений для α и количество дополнительных решений* (см. табл. 2).

Таблица 2

T-типы

Тип	Метка	Распределение решений для α	Доп. решения	Всего решений для α	Рисунки
T_1	T	4+2	$\varnothing$	6	
T_2	IIT	4+2	$\varnothing$	6	
T_3	IT	2+2	$\varnothing$	4	
T_4	$[IIT]$	2+0	∞	3	

Окончание таблицы 2

Тип	Метка	Распределение решений для α	Доп. решения	Всего решений для α	Рисунки
T_5	TT	3+1	$\varnothing$	4	
T_6	ITT	3+1	$\varnothing$	4	
T_7	TTT	2+0	$\varnothing$	2	
T_8	STT	1+1	$\varnothing$	2	
T_9	ST	0+2	$\varnothing$	2	
T_{10}	$[TTT]$	0+∞	∞	∞	

Определим T*-типы* аналогично I-типам. Как и в лемме 3.1, количество решений указано без учёта кратности.

Доказательство леммы 3.4 осуществляется одинаковым образом для всех T-типов посредством непрерывного перемещения допустимой окружности между параллельными прямыми до момента касания с α. □

§ 4. Леммы о пересечениях

В этом параграфе мы докажем несколько утверждений, полезных при доказательстве теорем 1 и 3 и особенно теоремы 2. В следующих четырёх леммах предполагается зафиксированной пара пересекающихся прямых.

Лемма 4.1 (первая лемма о пересечениях). *Если две окружности типа I_1 имеют два общих решения в одном секторе, то они имеют общую точку внутри этого сектора* (рис. 4 слева).

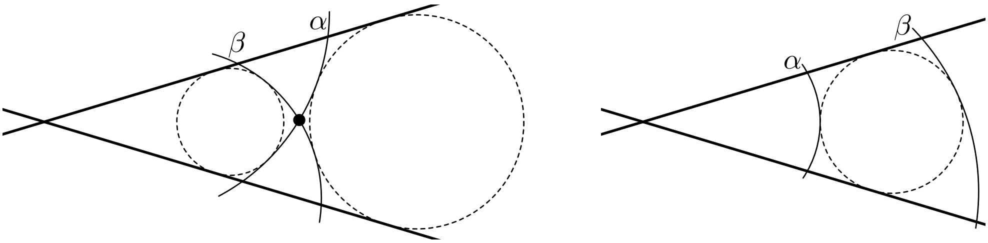


Рис. 4

Доказательство. Заметим, что если α и β не пересекаются в каком-то секторе, то они делят его на три области, из которых не более чем в одну можно вписать требуемую окружность-решение (рис. 4 справа). Так как в эту область, очевидно, можно вписать не более одной окружности, утверждение доказано. □

Лемма 4.2 (вторая лемма о пересечениях). *Пусть для окружностей α, β и γ, имеющих тип I_1, есть общее решение в некотором секторе. Тогда среди них какие-то две имеют общую точку внутри этого сектора.*

Доказательство. Предположим противное. Тогда дуги окружностей α, β и γ делят сектор на четыре части, каждая из которых граничит не более чем с двумя из этих окружностей (рис. 5). Решение должно быть вписано в одну из этих частей, т. е. оно не может касаться всех трёх окружностей α, β и γ — противоречие. □

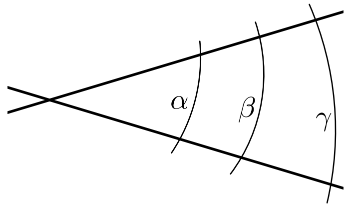
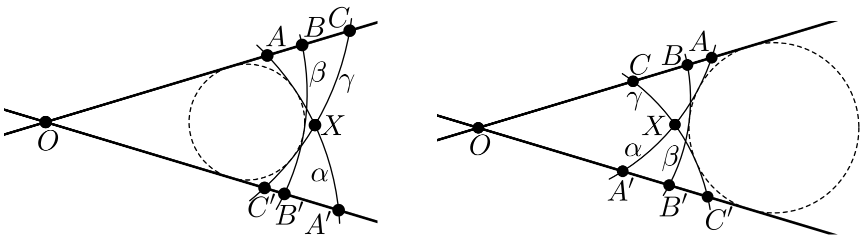


Рис. 5

Лемма 4.3 (третья лемма о пересечениях). *Пусть для окружностей α, β и γ, имеющих тип I_1, есть два общих решения в некотором секторе. Тогда среди них какие-то две имеют две общие точки внутри этого сектора.*

Доказательство. Предположим противное. Пусть дуги окружностей α, β и γ пересекают одну из сторон данного сектора в точках A, B и C, а вторую — в точках A', B' и C' соответственно. Пусть для определённости B лежит на отрезке AC (рис. 6).



β разделяет O и X *β не разделяет O и X*

Рис. 6

По лемме 4.1 дуги окружностей α и β имеют общую точку внутри рассматриваемого сектора, причём, согласно нашему предположению, ровно одну. Но тогда порядок троек точек O, A, B и O, A', B' должен отличаться, т. е. если A лежит на отрезке OB, то B' лежит на отрезке OA', и наоборот: если B лежит на отрезке OA, то A' лежит на отрезке OB'. То же верно и для двух других пар дуг окружностей (α и γ, β и γ). Отсюда ясно, что B' лежит на $A'C'$.

Пусть теперь X — точка пересечения дуг α и γ. Возможны два случая, в зависимости от того, разделяет ли β точки O и X (рис. 6). Нетрудно убедиться, что в обоих случаях α, β и γ делят сектор на семь частей, из которых только одна граничит со сторонами сектора и всеми тремя данными окружностями — только в эту часть может быть вписана искомая окружность. Получаем противоречие с условием леммы. □

Лемма 4.4 (четвёртая лемма о пересечениях). *Пусть даны окружность α типа I_1 и окружность β типа I_2, пересекающая обе стороны сектора S_{II}, причём α и β имеют в секторе S_{I} два общих решения. Тогда они имеют общую точку в секторе S_{I}.*

Доказательство. Разделим каждое общее решение для α и β в секторе S_{I} на две дуги точками касания с прямыми (рис. 7). Из полученных четырёх дуг рассмотрим две такие, которые имеют общую точку с α. Заметим, что α разделяет эти дуги, а β также имеет с каждой из них общую точку, поскольку β пересекает обе стороны сектора S_{II} и, значит, не пересекает вторую сторону сектора S_{I} (рис. 7). Следовательно, дуга β, лежащая в секторе S_{I}, должна пересекать дугу α в секторе S_{I} в силу непрерывности окружности. $\square$

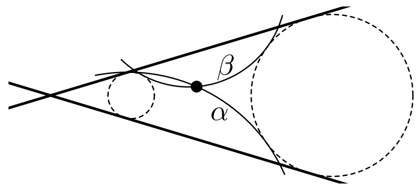


Рис. 7

§ 5. Геометрия окружностей

Для полного понимания этого раздела потребуется знание основ линейной алгебры. Однако для понимания формулировок необходимых нам результатов эти знания вовсе не обязательны. Читатель может ознакомиться только с формулировками (без доказательств) в разделе 5.2.

5.1. Основные понятия геометрии окружностей

Рассмотрим 5-мерное векторное пространство $\mathbb{R}^5$, снабжённое симметричной билинейной формой: $(\mathbf{x}, \mathbf{y}) := x_0y_3 + x_3y_0 + x_1y_1 + x_2y_2 - x_4y_4$ для $\mathbf{x} = (x_0, x_1, x_2, x_3, x_4)$, $\mathbf{y} = (y_0, y_1, y_2, y_3, y_4)$.

В этом пространстве *квадрика Ли* определяется условием:

$$\Omega = \{\mathbf{x}\colon (\mathbf{x}, \mathbf{x}) = 0\}.$$

Заметим, что для всякого $\mathbf{x}\in\mathbb{R}^5$ верно $\mathbf{x}\in\Omega \Leftrightarrow \langle\mathbf{x}\rangle\subset\Omega$, где $\langle\mathbf{x}\rangle=\{\lambda\mathbf{x}\colon \lambda\in\mathbb{R}\}$.

Каждой точке плоскости, а также каждой ориентированной окружности или прямой можно поставить в соответствие одномерное подпространство (прямую) $\langle\mathbf{x}\rangle \subset \Omega$ согласно следующим правилам.

1. Окружности с центром (p_1, p_2) и радиусом r отвечает $\langle(v, p_1, p_2, 1, \pm r)\rangle$. Здесь знак перед r зависит от ориентации окружности (см. подробности в [3]), а v подбирается так, чтобы полученное подпространство содержалось в Ω.
2. Прямой с нормальным вектором $\mathbf{p} = (p_1, p_2)$, где $\|\mathbf{p}\| = 1$, проходящей через точку $\mathbf{q}$, соответствует $\langle(-\mathbf{p}\cdot\mathbf{q}, p_1, p_2, 0, 1)\rangle$. Выбор одной из двух возможностей для p зависит от ориентации прямой (см. [3]).
3. Точке с координатами (p_1, p_2) соответствует $\langle(v, p_1, p_2, 1, 0)\rangle$, где v снова подбирается так, чтобы полученное подпространство содержалось в Ω. Точке ∞ соответствует $\langle(1, 0, 0, 0, 0)\rangle$.

Это соответствие взаимно однозначно.

Такое представление замечательно тем, что, как нетрудно проверить (см., например, [2]), объекты X и Y ориентированно касаются тогда и только тогда, когда ненулевые векторы $\mathbf{x}$ и $\mathbf{y}$, выбранные произвольно из соответствующих X и Y одномерных подпространств, удовлетворяют соотношению $(\mathbf{x}, \mathbf{y}) = 0$. Это позволяет переформулировать задачу Аполлония на линейно-алгебраическом языке.

Подробнее о геометрии окружностей см. [2] и [3].

5.2. Применение к задаче Аполлония

Пусть на плоскости дано несколько окружностей и окружность, касающаяся всех данных (решение). Сопоставим каждой данной исходной окружности знак «+», если она касается данного решения внешним образом, и знак «−», если внутренним. Мы говорим, что решение *соответствует* (*подходит*) данной *расстановке* (*набору, комбинации*) знаков.

Теперь поступим наоборот: сначала расставим знаки, а затем попробуем найти решение, которому этот набор знаков соответствовал бы. Сколько существует таких решений? Оказывается, если на конфигурацию данных окружностей наложены некоторые естественные ограничения, то таких решений не больше двух. Они поэтому называются *парными.*

Сперва нам потребуется следующая лемма.

Лемма 5.1. *Пусть тройка окружностей G, соответствующая прямым $\langle \mathbf{x} \rangle$, $\langle \mathbf{y} \rangle$ и $\langle \mathbf{z} \rangle$, невырожденна (ориентация здесь не учитывается). Тогда векторы $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{z}$ линейно независимы и $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{\perp} \not\subset \Omega$.*

Доказательство. Предположим противное, т. е. что $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{z}$ лежат в одной плоскости π. Тогда π имеет с квадрикой Ω три общие прямые. Но это возможно, лишь если $\pi \subset \Omega$. Пусть $\mathbf{t}$ — произвольный вектор из π. Тогда $\mathbf{x} - \mathbf{t} \in \pi$. Имеем $(\mathbf{t}, \mathbf{t}) = 0$ и $(\mathbf{x} - \mathbf{t}, \mathbf{x} - \mathbf{t}) = 0$, откуда

$$(\mathbf{x}, \mathbf{t}) = \frac{1}{2}\big((\mathbf{x}, \mathbf{x}) + (\mathbf{t}, \mathbf{t}) - (\mathbf{x} - \mathbf{t}, \mathbf{x} - \mathbf{t})\big) = 0.$$

Аналогично $(\mathbf{y}, \mathbf{t}) = (\mathbf{z}, \mathbf{t}) = 0$, т. е. $\langle \mathbf{t} \rangle$ соответствует некоторому объекту, который касается каждой окружности из G, и тройка G имеет бесконечно много решений. Но из лемм 3.1, 3.2, 3.4 следует, что бесконечным числом решений обладают лишь вырожденные тройки — противоречие.

Далее, если $\theta = \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{\perp} \subset \Omega$, то любой вектор из θ соответствует решению. Но $\dim \theta = 2 > 1$, поэтому тройка G снова имеет бесконечно много решений, т. е. вырождена. $\square$

Теперь можно сформулировать самый важный для нас факт.

Лемма 5.2. *Дано множество G, состоящее из окружностей на плоскости, причём $|G| \geqslant 3$ и G не вырождено. Каждой окружности из G сопоставлен знак. Тогда G имеет не более двух решений, подходящих для данной расстановки знаков.*

Доказательство. Выберем из G три произвольные окружности и ориентируем их согласно их знакам. Сопоставим им одномерные подпространства $\langle \mathbf{x} \rangle$, $\langle \mathbf{y} \rangle$ и $\langle \mathbf{z} \rangle$. Каждое подходящее решение соответствует окружности, ориентированно касающейся каждой окружности из G. Поэтому надо доказать, что существует не более двух векторов $\mathbf{t}$, таких что

$$(\mathbf{x}, \mathbf{t}) = (\mathbf{y}, \mathbf{t}) = (\mathbf{z}, \mathbf{t}) = (\mathbf{t}, \mathbf{t}) = 0.$$

По лемме 5.1 векторы $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{z}$ линейно независимы и $\theta = \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{\perp} \not\subset \Omega$. Но $\dim \theta = 2$. Значит, θ пересекает Ω не более чем по двум прямым, т. е. существует не более двух решений, что и требовалось. □

Следующая лемма достаточно естественна, причём она оказывается весьма полезной во многих ситуациях.

Лемма 5.3. *Пусть в условиях леммы* 5.2 *центры окружностей из* G *лежат на одной прямой. Тогда парные решения симметричны относительно этой прямой.*

Доказательство. Рассмотрим такую систему координат, в которой линия центров G является осью абсцисс. Снова выберем из G три окружности и сопоставим им векторы $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$. Обозначим $\theta = \langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle^{\perp}$. По лемме 5.1 имеем $\dim \theta = 2$ и $\theta \not\subset \Omega$.

Рассмотрим отображение $f \colon \mathbb{R}^5 \to \mathbb{R}^5$, определённое правилом:

$$f \colon (x_0, x_1, x_2, x_3, x_4) \mapsto (x_0, x_1, -x_2, x_3, x_4).$$

Отображение f, очевидно, линейно, и при этом оно сохраняет значение билинейной формы. Кроме того, оно отображает каждый из векторов $\mathbf{x}$, $\mathbf{y}$ и $\mathbf{z}$ в себя (поскольку для них $x_2 = 0$), поэтому в силу линейности f каждая точка из $\langle \mathbf{x}, \mathbf{y}, \mathbf{z} \rangle$ также отображается в себя. Поэтому и $f(\theta) = \theta$.

Если плоскость θ имеет с Ω меньше двух общих прямых, то доказательство не требуется. В противном случае θ пересекает Ω по двум прямым, причём векторы из этих прямых порождают θ. Это значит, что если каждая из этих прямых переходит в себя, то и каждая точка θ переходит в себя (в силу линейности f), т. е. f является тождественным преобразованием — противоречие. Следовательно, эти прямые переходят друг в друга под действием f. Это как раз и означает, что соответствующие им объекты симметричны относительно оси абсцисс, что и требовалось. □

§ 6. Доказательство теоремы 1

Предположим противное условию теоремы, т. е. что нашлись четыре объекта (не все из которых касаются в одной точке) с семью решениями.

Лемма 6.1. *Пусть среди четырёх или более исходных объектов есть точка, а количество решений больше четырёх. Тогда все исходные объекты касаются в одной точке.*

Доказательство. Известно (см., например, [1, с. 100]), что если в тройке объектов есть точка, то количество решений для этой тройки не больше четырёх, за исключением случая вырожденной тройки. Поэтому в нашем случае любая тройка объектов, содержащая точку, вырождена. Ясно, что тогда все объекты касаются в этой точке, что исключено условием. □

Из леммы 6.1 получаем, что все исходные объекты — окружности или прямые. Докажем, что среди них нет касающихся. Предположим противное и рассмотрим любую тройку исходных объектов, содержащую касающиеся. По лемме 3.4 (применённой после инверсии в точке касания) такая тройка либо имеет не более шести решений, либо она вырождена. Но у нас по предположению имеется хотя бы семь решений, поэтому любая тройка исходных объектов, содержащая касающиеся, вырождена. Тогда очевидно, что снова все объекты касаются в одной точке, что исключено формулировкой теоремы.

Таким образом, среди исходных объектов нет точек и пар касающихся. Далее рассмотрим два случая.

1. Среди исходных окружностей или прямых какие-то две пересекаются. Переведём инверсией любые две пересекающиеся обобщённые окружности в пересекающиеся прямые. Для краткости здесь и далее образы исходных объектов при одной или нескольких инверсиях называются просто исходными объектами, а образы решений — просто решениями.

Кроме двух пересекающихся прямых есть ещё две исходные обобщённые окружности. Обозначим их через α и β.

По лемме 3.1 в случае I-типов, отличных от I_1 и I_2, уже по отдельности существует не больше пяти решений, поэтому α и β могут быть только объектами этих типов. Если они имеют разные типы, то их распределения равны 2–2–2–2 и 0–2–4–2 и они пересекаются не больше чем по шести окружностям, т. е. всего есть не более шести решений. Поэтому α и β обязательно имеют одинаковые типы.

Пусть α и β имеют тип I_1. Тогда семь общих решений для α и β, очевидно, имеют распределение 2–2–2–1. Значит, согласно лемме 4.1 хотя бы в трёх секторах α и β должны пересекаться, что невозможно.

Пусть α и β имеют тип I_2. Тогда, чтобы решений было хотя бы 7, необходимо, чтобы эти окружности пересекали обе стороны одного и того же сектора, иначе уже распределения решений для них (4–2–0–2 и 2–4–2–0, либо 4–2–0–2 и 0–2–4–2 в одинаковой нумерации секторов) пересекаются не более, чем по четырём окружностям. С этого сектора и начнём нумерацию. Тогда в секторе S_{I} не более четырёх решений, а значит, в секторах S_{II} и S_{IV} в сумме не меньше трёх. Завершает доказательство первого случая следующая лемма.

Лемма 6.2. *Для двух окружностей α и β, имеющих тип I_2 и пересекающих обе стороны сектора S_{I}, есть не более двух общих решений в секторах S_{II} и S_{IV}.*

Доказательство. Предположим противное. Заметим, что α и β касаются своих решений в секторах S_{II} и S_{IV} внешним образом, т. е. соответствуют комбинации «+++». Центры этих решений лежат на одной прямой — биссектрисе секторов S_{II} и S_{IV}. Ясно, что α и β лежат по одну сторону от этой прямой, в частности, они не могут быть симметричны относительно неё. Противоречие с леммой 5.3. □

2. Никакие две окружности или прямые из исходных четырёх не пересекаются. Переведём инверсией любые два из данных объектов в концентрические окружности. Образы двух оставшихся объектов снова обозначим через α и β.

По лемме 3.2, для α и β отдельно существует четыре решения рода A и столько же рода B, причём каждое из этих множеств симметрично относительно радиуса, на котором расположена соответствующая окружность α или β. Так как у нас, согласно предположению, имеется не меньше семи решений, то хотя бы четыре из них одного рода. Далее нам потребуется несколько лемм.

Лемма 6.3. *Пусть окружности α и β порождают четыре общих решения одного рода. Тогда α и β лежат либо на одном диаметре, либо на перпендикулярных диаметрах. В последнем случае общие решения обязательно имеют род* B, *а их центры образуют прямоугольник* (рис. 8).

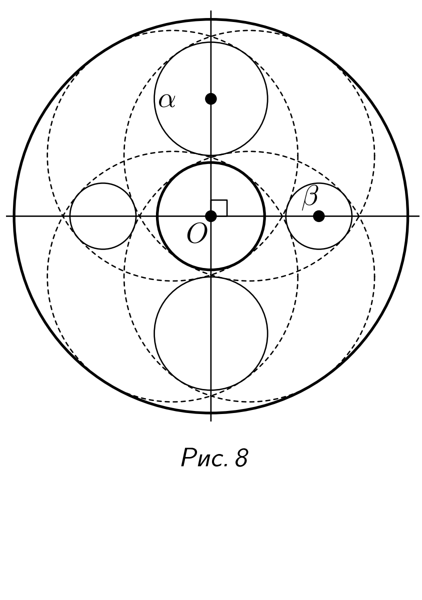


Рис. 8

Доказательство. Пусть α и β не лежат на одном диаметре (иначе доказательство не требуется). Обозначим множество четырёх общих решений одного рода для α и β через X. Поскольку окружности из X являются решениями и для α, и для β, то X симметрично относительно радиусов α и β. Однако по нашему предположению угол между диаметрами, на которых лежат α и β, ненулевой. Поэтому множество X переходит в себя при некотором повороте на $\varphi > 0$ относительно O. Можно считать $\varphi \leqslant 180°$.

Докажем, что $\varphi = 90°$ или $\varphi = 180°$. Рассмотрим произвольную окружность Γ из X. При поворотах на 0, φ, 2φ, и т. д. она переходит в окружность из X. Но в X ровно четыре окружности, поэтому Γ может иметь от двух до четырёх образов при таких поворотах (Γ не может перейти в себя, так как $0 < \varphi \leqslant 180°$). Эти случаи соответствуют поворотам на $180°$, $120°$ и $90°$ соответственно. Однако если образов три, то в X остаётся ещё одна окружность, не являющаяся образом Γ. У неё тоже три образа, которые не совпадают с образами Γ. Получается, что в X хотя бы шесть окружностей, а не четыре — противоречие. Остаются варианты $180°$ и $90°$, которые реализуются, только если центры окружностей из X образуют прямоугольник или квадрат соответственно.

Чтобы теперь доказать, что в X не могут быть окружности рода А, нам понадобится следующая достаточно очевидная лемма.

Лемма 6.4. *Если для α и β есть два общих противоположных решения P и Q рода* А, *то α и β лежат на одном диаметре.*

Доказательство. Заметим, что P и Q не пересекаются, так как они центрально симметричны относительно точки O, которую не содержат

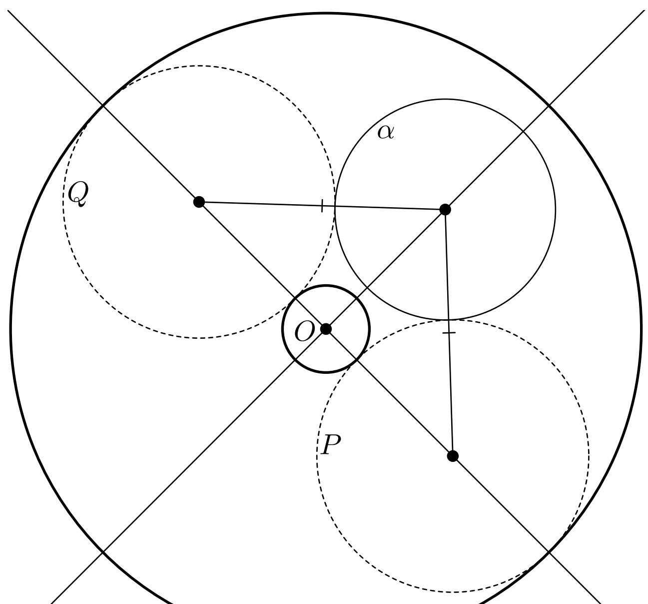


Рис. 9

внутри себя. Радиус окружности α меньше радиуса окружности P, поэтому если α и P касаются внутренним образом, то α лежит внутри P. Но так как P и Q не пересекаются, α не может касаться Q. Значит, α и аналогично β касаются P и Q внешним образом (рис. 9). Но радиусы P и Q равны, поэтому геометрическое место центров окружностей, касающихся внешним образом P и Q, является серединным перпендикуляром отрезка между центрами P и Q, который, очевидно, проходит через O, что и требовалось. □

Продолжим доказательство леммы 6.3. Из леммы 6.4 ясно, что в X не может быть окружностей рода А, так как окружности из X делятся на две пары противоположных, поэтому α и β лежали бы одновременно на двух диаметрах, что невозможно. Таким образом, в X только окружности рода В.

Так как центры окружностей из X образуют прямоугольник, то кроме концентрических существуют ещё четыре окружности, которые касаются всех окружностей из X — они вписаны в криволинейные симметричные четырёхугольники, образованные окружностями из X (рис. 8). Других окружностей, касающихся всех окружностей из X, быть не может по доказанному выше первому случаю основной теоремы. Действительно, в X четыре попарно пересекающихся окружности, и их уже касаются шесть других — две концентрические и четыре вписанные в криволинейные четырёхугольники. Центры этих четырёх окружностей лежат на двух перпендикулярных диаметрах; α и β должны быть двумя из этих четырёх окружностей, т. е. они должны лежать на двух перпендикулярных диаметрах — лемма 6.3 доказана. □

Лемма 6.5. *Пусть окружности α и β порождают четыре общих решения рода* В *и лежат на перпендикулярных диаметрах. Тогда они не могут порождать ещё три общих решения рода* А.

Доказательство. Пусть это не так. Обозначим общие решения рода А через P, Q, R. Докажем, что среди них есть пара противоположных. Так как все они касаются α, а всего среди решений для α четыре решения рода А, причём они разбиваются на две пары симметричных, то среди решений P, Q, R два симметричны относительно радиуса, на котором лежит α. Пусть это P и Q. Аналогично рассмотрим пару окружностей, симметричную относительно радиуса, на котором лежит β. Если это снова P и Q, то они переходят друг в друга при двух симметриях, угол между осями которых составляет 90°, а тогда P и Q совпадают, что невозможно. Пусть это Q и R (или P и R). Тогда, так как угол между радиусами, на которых лежат α и β, равен 90°, то P и R (соответственно Q и R) противоположны — отражение относительно двух перпендикулярных осей является центральной симметрией. Таким образом, среди P, Q и R есть пара противоположных. Следовательно, α и β имеют два противоположных решения рода А, из чего с помощью леммы 6.4 получаем, что они лежат на одном диаметре — противоречие. □

Продолжим доказательство теоремы 1. Так как α и β порождают четыре общих решения одного рода, то по лемме 6.3 они лежат либо на одном диаметре, либо на перпендикулярных диаметрах. Но во втором случае α и β порождают ещё три общих решения рода А, что невозможно по лемме 6.5. Таким образом, α и β лежат на одном диаметре, в частности, центры всех исходных окружностей лежат на одной прямой.

Окружности α и β порождают семь общих решений, поэтому по лемме 3.3 из них можно выбрать пару пересекающихся симметричных решений. Следующая лемма завершает доказательство теоремы 1.

Лемма 6.6. *Пусть центры всех четырёх исходных окружностей лежат на одной прямой, а среди решений есть два пересекающихся и симметричных относительно этой прямой. Тогда после инверсии в точке пересечения этих решений исходные окружности образуют одно из распределений* 4–0–0–0, 3–0–1–0 *или* 2–0–2–0 *относительно образов этих решений. Количество всех решений в этих случаях не превосходит* 4, 2 *и* 6 *соответственно.*

Доказательство. Центр инверсии лежит на линии центров исходных окружностей (так как точка пересечения двух симметричных окружностей лежит на их оси симметрии). Поэтому центры всех четырёх исходных окружностей до инверсии лежали на одной прямой, проходящей через центр инверсии. Значит, и после инверсии их центры лежат на одной прямой, а тогда и все исходные окружности должны лежать в двух противоположных секторах — первая часть леммы тем самым доказана.

С этого момента мы будем рассматривать решения в качестве исходных объектов (назовём их *новыми исходными объектами*), а четвёрку исходных объектов — наоборот — в качестве решений (назовём их *новыми решениями*). Обозначим последние через $\omega_1, \omega_2, \omega_3, \omega_4$ в том порядке, в котором они касаются любой из прямых.

Рассмотрим по отдельности все возможные распределения четырёх новых решений. В первом случае (распределение 4–0–0–0) по лемме 3.1 оставшиеся новые исходные объекты должны быть типа I_2 или I_4. Но среди четырёх решений для любой окружности типа I_4 заведомо есть пересекающиеся — те, что содержат окружность α на рисунке в табл. 1 внутри себя. Значит, все оставшиеся новые исходные объекты имеют тип I_2. Они касаются ω_1 и ω_4 внешним образом, а ω_2 и ω_3 — внутренним. По лемме 5.2, где в качестве G мы берём все 4 новых решения, получаем, что таких новых исходных объектов не более двух.

Для второго случая (3–0–1–0) вообще не существует подходящего нового исходного объекта никакого типа, в чём нетрудно убедиться с помощью таблицы 1.

В третьем же случае (2–0–2–0) оставшиеся новые исходные объекты могут быть типа I_1 или I_2.

Окружности типа I_2 должны касаться каждого из четырёх новых решений внешним образом. Применим для них лемму 5.2, взяв в качестве G четвёрку новых решений с комбинацией знаков «$++++$». Получаем, что есть не более двух окружностей типа I_2 (рис. 10 справа).

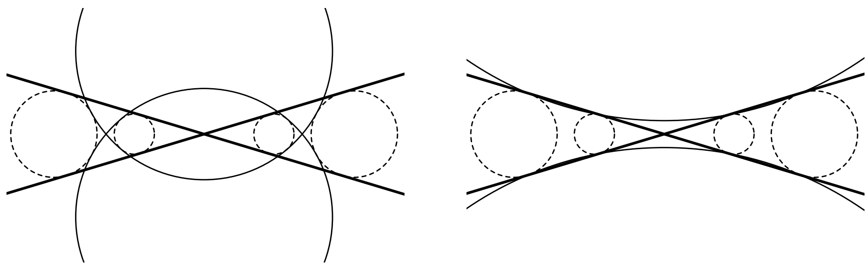


Рис. 10. Распределение новых решений 2–0–2–0

Окружности типа I_1 касаются ω_2 и ω_3 внутренним образом, а ω_1 и ω_4 — внешним. Применим ещё раз лемму 5.2, снова взяв в качестве G четвёрку новых решений с комбинацией знаков «$+--+$». Получаем, что есть не более двух окружностей типа I_1 (рис. 10 слева).

Таким образом, общее количество новых исходных объектов типов I_1 и I_2 в этом случае не больше четырёх, т. е. количество решений исходной задачи не больше шести. Доказательство леммы 6.6, а с ней и всей теоремы 1 завершено. $\square$

§ 7. Доказательство теоремы 2

Как и при доказательстве теоремы 1, предположим противное, т. е. что нашлась пятёрка объектов (не все из которых касаются в одной точке) с пятью решениями. По лемме 6.1 среди исходных объектов нет точек. Заметим, что количество решений равно количеству исходных объектов, а также что все решения не могут касаться в одной точке (иначе все исходные объекты касаются в этой же точке). Поэтому решения можно рассматривать в качестве исходных объектов (и наоборот), и среди решений тоже нет точек.

7-1. Ни среди исходных объектов, ни среди решений нет касающихся. Можно считать, что либо среди исходных объектов, либо среди решений есть пересекающиеся. Действительно, в противном случае переведём инверсией любые две исходные обобщённые окружности в концентрические и воспользуемся леммой 3.3.

Полученная пара пересекающихся объектов принадлежит некоторой пятёрке, которую мы и будем считать исходной. Переведём инверсией эти пересекающиеся объекты в прямые. Оставшиеся исходные объекты обозначим α, β и γ. Из леммы 3.1 получаем, что каждый из них должен иметь тип I_1, I_2 или I_5. Но если какой-то объект имеет тип I_5, то среди решений должна быть точка, что невозможно по лемме 6.1. Значит, каждый из объектов α, β и γ имеет тип I_1 или I_2. Прежде чем разбирать возможные случаи, заметим, что две окружности типа I_2 могут иметь больше четырёх общих решений только тогда, когда пересекают обе стороны одного и того же сектора. С этого сектора и начнём нумерацию.

7.1.1° Все окружности α, β и γ имеют тип I_1. Тогда в каждом секторе есть не более двух решений, а распределение всех пяти решений изоморфно 2–2–1–0, 2–1–2–0 или 1–1–1–2. Рассмотрим точки пересечения окружностей α и β, β и γ, α и γ и обозначим их множество через C. Для каждой пары окружностей таких точек не больше двух, поэтому всего их не больше шести.

Согласно лемме 4.2, в секторах с одним решением есть хотя бы одна точка из C. Далее, в секторе с двумя решениями каждая пара окружностей (α и β, β и γ, α и γ) имеет общую точку по лемме 4.1, а какая-то пара имеет в этом секторе даже две общие точки по лемме 4.3. Поэтому в таком секторе есть не меньше четырёх точек из C (все три окружности α, β и γ не могут иметь общую точку, поскольку тогда они имеют метку $[III]_1$ или $[III]_2$, и по лемме 3.1 общее количество решений, которые не являются точками, не больше четырёх).

Таким образом, в случае распределений 2–2–1–0 и 2–1–2–0 имеем $|C| \geqslant 2 \cdot 4 + 1 = 9$, а в случае распределения 1–1–1–2 имеем $|C| \geqslant 3 + 4 = 7$. Но $|C| \leqslant 6$ — противоречие.

7.1.2° Все окружности α, β и γ имеют тип I_2. Рассмотрим случаи в зависимости от количества решений в секторе S_{I}. Обозначим эти решения через ω_i, где $i = 1, 2, 3, 4$, в том порядке, в котором они касаются любой из прямых.

7.1.2А *В секторе S_{I} четыре решения.*Тогда α, β и γ касаются ω_1, ω_4 внешним образом, а ω_2, ω_3 — внутренним. Из леммы 5.2 для $G = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ и комбинации «$+--+$» получаем противоречие.

7.1.2Б *В секторе S_{I} три решения.*Рассмотрим окружность α. Окружности $\omega_1, \omega_2, \omega_3$ являются тремя из четырёх решений для α в секторе S_{I}. Простой перебор показывает, что если взять любые три решения из этих четырёх, то среднее из них касается α внутренним образом (см. рисунок в табл. 1). Поэтому α, β и γ касаются ω_2 внутренним образом. Однако всего решений пять, поэтому есть ещё два решения в секторах S_{II} и S_{IV}, которые касаются α, β и γ внешним образом. Обозначим их через σ_1 и σ_2. Применив лемму 5.2 для $G = \{\sigma_1, \sigma_2, \omega_2\}$ и комбинации «$++-$», получаем противоречие.

7.1.2В *В секторе S_{I} не более двух решений.* Тогда в секторах S_{II} и S_{IV} в сумме хотя бы три решения, что противоречит лемме 6.2.

7.1.3° Две окружности (например, α и β) типа I_2 и одна окружность (γ) типа I_1. За счёт окружности типа I_1 в секторе S_{I} не более двух решений, но тогда в секторах S_{II} и S_{IV} хотя бы три решения, что противоречит лемме 6.2.

7.1.4° Две окружности (α и β) типа I_1 и одна окружность (γ) типа I_2. В этом случае без ограничения общности возможны распределения решений 1–2–0–2 и 2–2–0–1. Разберём их отдельно.

7.1.4А *Распределение* 1–2–0–2. Так как в секторах S_{II} и S_{IV} по два решения, то из леммы 4.4 получаем (после перенумерации секторов), что все точки пересечения γ с α и β находятся в этих секторах. Далее, по лемме 4.1 окружности α и β также должны иметь общие точки в этих секторах. Таким образом, α, β и γ попарно не имеют общих точек в секторе S_{I} (рис. 11). В таком случае они разбивают его на пять частей, в каждую из которых можно вписать окружность, касающуюся лишь не более двух окружностей из α, β и γ — противоречие.

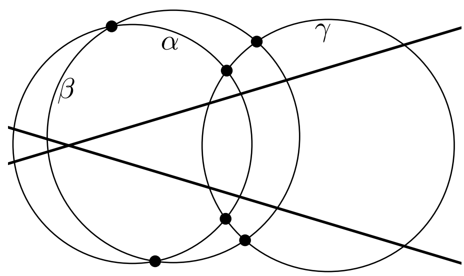


Рис. 11

7.1.4Б *Распределение* 2–2–0–1. Обозначим решения в секторе S_{I} через ω_1 и ω_2, а решения в секторе S_{II} через ω_3 и ω_4 так, чтобы решение с меньшим радиусом имело меньший индекс (рис. 12). Сделаем инверсию в точке пересечения α и β (она существует по лемме 4.1). Заметим, что ω_1 и ω_3 лежат внутри α и β, а ω_2 и ω_4 — вне. Но четыре области, на которые α и β делят плоскость, при инверсии перейдут в секторы соответствующих прямых. Поэтому распределение образов $\omega_1, \omega_2, \omega_3, \omega_4$ относительно прямых — образов α и β равно 2–0–2–0, и распределение *всех* решений относительно тех же прямых не может снова быть 2–2–0–1. Следовательно, задача сводится к одному из уже разобранных случаев 7.1.1°–7.1.3° или 7.1.4А.

7-2. Среди исходных объектов или среди решений есть касающиеся. Будем считать исходной пятёрку, в которой есть пара касающихся

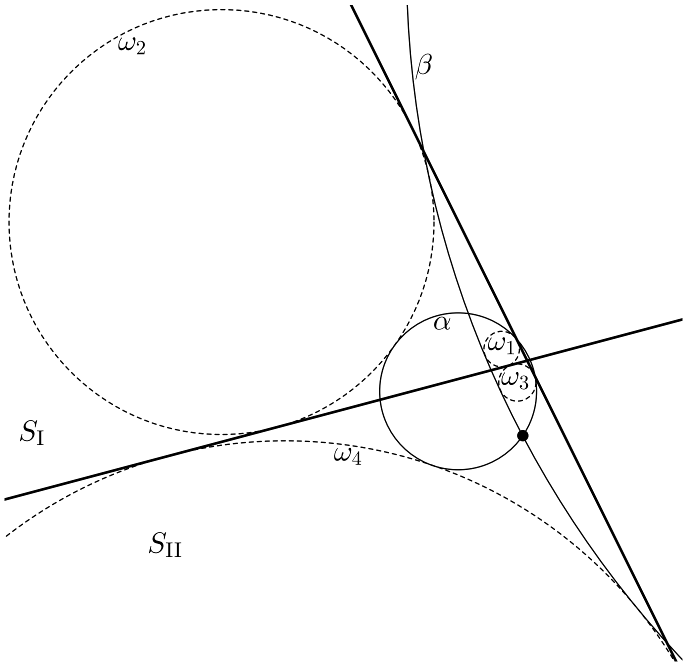


Рис. 12. *К доказательству случая* 4.2°

объектов. Переведём их инверсией в (горизонтальные) параллельные прямые. Докажем, что среди оставшихся трёх исходных объектов (мы их обозначим, как обычно, α, β и γ) не может быть горизонтальной прямой.

Лемма 7.1. *Пусть в конфигурацию нескольких исходных объектов входят три параллельные прямые, а количество решений больше двух. Тогда все исходные объекты — параллельные прямые или ∞.*

Доказательство. Заметим, что в условиях леммы все решения являются горизонтальными прямыми или ∞, т. е. все они касаются в точке ∞. Но если решений больше двух и все они касаются в одной точке, то, очевидно, и все исходные объекты касаются в той же точке, что и требовалось. □

Среди T-типов T_1–T_9 только два имеют 5 или более решений. Это типы T_1 и T_2, распределения решений для них одинаковы и равны 4+2. Значит, распределение пяти решений может иметь вид 3+2 или 4+1. Разберём эти случаи.

7.2.1° Распределение решений 4+1.

Лемма 7.2. *Пусть зафиксированы две горизонтальные прямые, и некоторые окружности α и β порождают четыре общих окружности-решения. Тогда линия центров окружностей α и β вертикальна.*

Доказательство. Предположим противное и обозначим множество из четырёх общих окружностей-решений для α и β через X. Это множество симметрично относительно вертикальных прямых, на которых лежат центры α и β. По нашему предположению эти прямые не совпадают, поэтому X переходит в себя при симметрии относительно каждой из некоторых двух параллельных прямых. Но тогда X переходит в себя и при некотором параллельном переносе, т. е. является бесконечным — противоречие. □

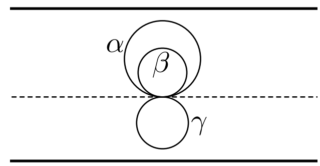


Рис. 13. Распределение 4+1

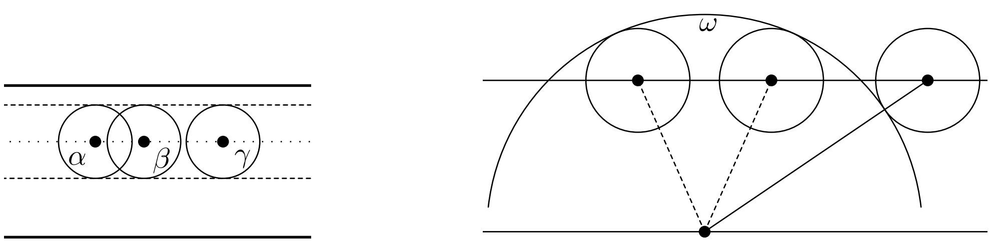
По этой лемме имеем, что центры α, β и γ лежат на одной вертикальной прямой, а тогда они касаются прямой-решения в одной и той же точке, т. е. образуют вырожденную тройку (рис. 13), что исключено леммой 7.1.

7.2.2° Распределение решений 3+2. 10 этом случае среди решений есть две параллельные прямые, поэтому радиусы окружностей α, β и γ равны половине расстояния между этими прямыми, а их центры лежат на одной горизонтальной прямой (рис. 14 слева). То же можно сказать и про три оставшихся решения-окружности (поскольку среди исходных объектов есть две параллельные прямые). Получаем два множества окружностей X и Y, которые удовлетворяют следующим условиям:

— в X и Y по три окружности;
— в каждом из множеств окружности имеют равный радиус;
— в каждом из множеств центры окружностей лежат на одной прямой, причём для X и для Y эти прямые параллельны (горизонтальны);
— любые две окружности из разных множеств касаются.

Мы докажем, что одновременное соблюдение этих условий невозможно.

Возьмём в $X \cup Y$ окружность ω с самым *левым* центром. Расстояние между центром окружности из X и центром окружности из Y может принимать одно из двух значений, в зависимости от того, внутреннее



Распределение $3+2$ *Финальное рассуждение*

Рис. 14

касание или внешнее. Поэтому из трёх таких расстояний для окружности ω найдутся два равных. Центры соответствующих окружностей образуют равнобедренный треугольник, одна из вершин которого окажется *левее* центра окружности ω (рис. 14 справа). Полученное противоречие завершает доказательство данного случая и всей теоремы 2. □

§ 8. Доказательство теоремы 3

Доказательство. Возьмём произвольную четвёрку объектов с ровно шестью решениями. По лемме 6.1 среди исходных объектов нет точек. Предположим, что среди исходных объектов есть касающиеся. Переведём их в горизонтальные прямые. Остаются ещё два исходных объекта α и β, и по лемме 7.1 ни один из них не может быть горизонтальной прямой. Но всего имеется шесть решений, поэтому α и β должны иметь тип T_1 или T_2 (см. табл. 2). Распределение решений для каждого из этих типов равно $4+2$, поэтому среди решений есть две параллельные горизонтальные прямые, и линия центров окружностей α и β горизонтальна, а их радиусы равны. Однако среди решений есть и четыре окружности, поэтому по лемме 7.2 линия центров α и β вертикальна — противоречие, и среди исходных объектов нет пар касающихся.

Рассмотрим те же случаи, что и в доказательстве теоремы 1.

8-1. Среди исходных окружностей или прямых какие-то две пересекаются. Переведём их инверсией в пересекающиеся прямые. Заметим, что среди I-типов, которые не содержат касающихся объектов, только типы I_1 и I_2 имеют не менее шести решений. Поэтому оставшиеся исходные окружности α и β имеют тип I_1 или I_2, и возможны три случая. Разберём их по отдельности.

8.1.1° Обе окружности α и β имеют тип I_2. Тогда α и β пересекают обе стороны одного и того же сектора (иначе распределения решений для них пересекаются не более чем по четырём окружностям). Пусть это сектор S_{I}. Тогда по лемме 6.2 в секторах S_{II} и S_{IV} не более двух решений, а значит, в секторе S_{I} есть четыре решения, причём α и β касаются их в одной и той же комбинации знаков «$+--+$», считая от O. По лемме 5.3 окружности α и β симметричны относительно биссектрисы секторов S_{I} и S_{III} (рис. 15).

Очевидно, что центр любой окружности, касающейся α и β внешним образом, должен лежать на их оси симметрии (так как радиусы окружностей α и β равны), т. е. на биссектрисе секторов S_{I} и S_{III}. Но центры решений в секторах S_{II} и S_{IV}, которые как раз касаются α и β внешним образом, лежат на биссектрисе секторов S_{II} и S_{IV} — противоречие. Значит, этот случай вообще невозможен.

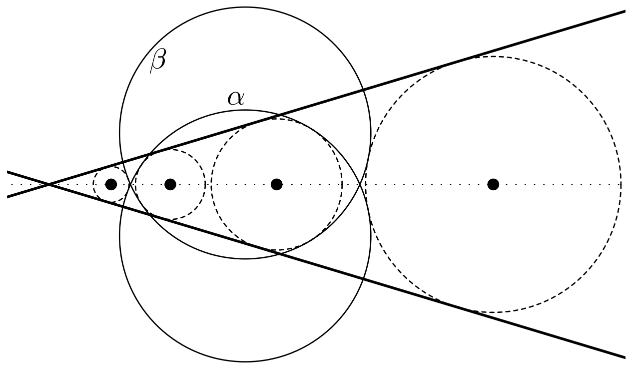


Рис. 15. Обе исходные окружности — типа I_2

8.1.2° Одна окружность (α) типа I_1 и одна окружность (β) типа I_2 (рис. 16 слева). Пусть β пересекает обе стороны сектора S_{II}. Тогда распределение решений равно 2–2–2–0. Сперва нам потребуется следующая (интересная и сама по себе) лемма.

Лемма 8.1. *Пусть окружность α типа I_1 и окружность β типа I_2, пересекающая обе стороны сектора S_{II} или S_{IV}, порождают четыре общих решения с распределением 2–0–2–0. Тогда множество этих решений симметрично относительно точки пересечения прямых O.*

Доказательство. Обозначим общие решения через ω_i (в том порядке, в котором они касаются любой из прямых), а через a_i обозначим длину касательной из O к ω_i для $i = 1, 2, 3, 4$.

Рассмотрим инверсию с центром в O, оставляющую β на месте (такая инверсия существует, так как O лежит вне любой окружности типа I_2). При такой инверсии прямые также остаются на месте, поэтому решения для β должны перейти в решения для β. Кроме того, после инверсии любое решение для β остаётся в своём секторе. Но решения для β в секторах S_{I} и S_{III} — это как раз окружности ω_i для $i = 1, 2, 3, 4$. Следовательно, ω_1 меняется местами с ω_2, а ω_3 — с ω_4. Отсюда $a_1a_2 = a_3a_4$.

Рассмотрим композицию инверсии и центральной симметрии с центром в O, оставляющую α на месте (такое преобразование существует, так как O лежит внутри любой окружности типа I_1). В этом случае любое решение для α переходит в противоположный сектор, поэтому ω_1 меняется с ω_3, а ω_2 меняется с ω_4. Отсюда $a_1a_3 = a_2a_4$.

Из полученных равенств следует, что $a_1{=}a_4$ и $a_2{=}a_3$, что и требовалось. □

Из этой леммы получаем, что α и β переходят в себя при симметрии относительно биссектрисы секторов S_{II} и S_{IV}. Поэтому оба решения в сек-

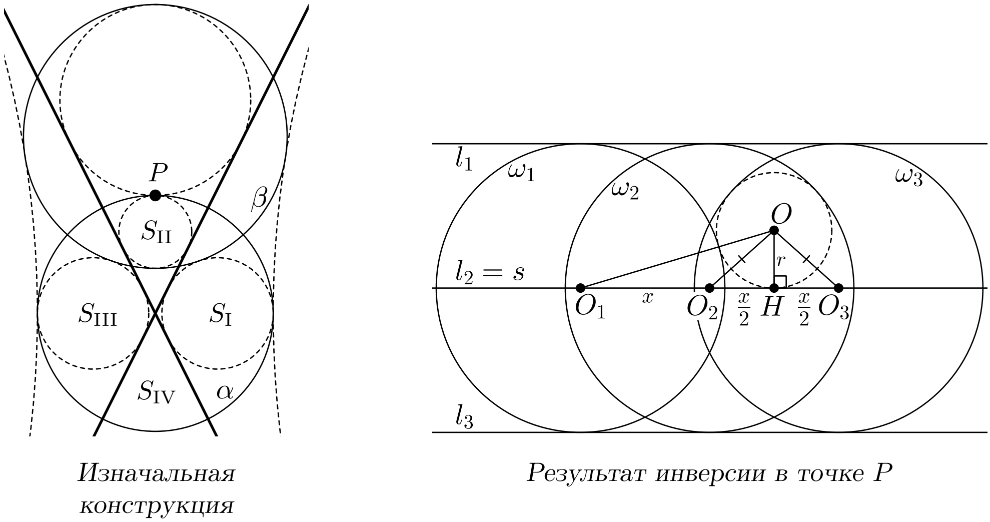


Рис. 16. Одна окружность типа I_1 и одна окружность типа I_2

торе S_{II} касаются α в одной и той же точке P. По лемме 4.4 получаем, что окружности α и β имеют общие точки в секторах S_{I} и S_{III}. Поэтому в секторе S_{II} они общих точек не имеют и $P \notin \beta$.

Докажем теперь, что исходная четвёрка соответствует конфигурации 3 из условия теоремы 3. Сделаем инверсию с центром в точке P. Тогда α и оба решения в секторе S_{II} переходят в некоторые три параллельные прямые l_1, l_2, l_3. Пусть l_2 — образ α, а расстояние между l_1 и l_3 равно 2. Так как $P \notin \beta$, то β переходит в некоторую окружность ω_2, касающуюся l_1 и l_3. Пересекающиеся прямые переходят в окружности ω_1 и ω_3, касающиеся l_1 и l_3, причём в силу симметрии центры этих окружностей равноудалены от центра ω_2 (рис. 16 справа). Обозначим центры ω_i через O_i при $i = 1, 2, 3$, и пусть $x = O_1O_2 = O_2O_3$.

Обозначим через s прямую, равноудалённую от l_1 и l_3. Четыре решения в секторах S_{I} и S_{III} при инверсии переходят в окружности, касающиеся ω_1, ω_2 и ω_3 в соответствии с комбинациями «$+--$» и «$--+$». Рассмотрим одну из этих комбинаций и обозначим подходящие для неё окружности через σ_1 и σ_2. По лемме 5.3 окружности σ_1 и σ_2 симметричны относительно s — линии центров ω_i. Поэтому радиусы окружностей σ_1 и σ_2 равны, а их линия центров вертикальна. Следовательно, σ_1 и σ_2 должны касаться l_2 в одной и той же точке, и l_2 совпадает с s.

Обозначим радиус образа любого из решений в секторе S_{I} и S_{III} через r, а его центр через O. Имеем $OO_1 = 1 + r$, $OO_2 = OO_3 = 1 - r$. Пусть OH — высота равнобедренного $\triangle O_2OO_3$. Имеем $O_2H = HO_3 = x/2$, $OH = r$.

По теореме Пифагора для прямоугольных $\triangle O_1HO$ и $\triangle O_2HO$:

$$\begin{cases} OH^2+HO_1^2=OO_1^2, \\ OH^2+HO_2^2=OO_2^2 \end{cases} \Leftrightarrow \begin{cases} \left(\frac{3x}{2}\right)^2=(1+r)^2-r^2, \\ \left(\frac{x}{2}\right)^2=(1-r)^2-r^2 \end{cases} \Leftrightarrow \begin{cases} \frac{9x^2}{4}=1+2r, \\ \frac{x^2}{4}=1-2r. \end{cases}$$

Отсюда $x=2/\sqrt{5}$, $r=2/5$, что соответствует конфигурации 3.

8.1.3° Обе окружности α и β имеют тип I_1. Тогда в каждом секторе не более двух решений и возможны распределения 1–2–1–2 и 1–1–2–2 (распределение 2–2–2–0 невозможно, так как иначе по лемме 4.1 α и β имеют три общие точки). Разберём их по отдельности.

8.1.3А *Распределение* 1–2–1–2. Тогда α и β касаются решений в секторах S_{II} и S_{IV} в комбинации «$+--+$» (в том порядке, в котором они касаются любой из прямых). По лемме 5.3 окружности α и β симметричны относительно биссектрисы секторов S_{II} и S_{IV}. Получаем конфигурацию 2.

8.1.3Б *Распределение* 1–1–2–2. Тогда по лемме 4.1 α и β имеют общие точки в секторах S_{I} и S_{II}. Сделаем инверсию в одной из них. Заметим, что одна из прямых разделяет точки пересечения α и β, а другая — нет (рис. 12). Поэтому после инверсии одна из прямых перейдёт в окружность типа I_1, а другая перейдёт в окружность типа I_2. Это соответствует случаю 8.1.2°, который уже рассмотрен.

8-2. Никакие две окружности или прямые из исходных четырёх не пересекаются.

Лемма 8.2. *Даны три попарно непересекающиеся окружности и шесть решений. Тогда можно выбрать две исходные окружности из данных трёх и четыре решения из данных шести так, что если перевести инверсией выбранные исходные окружности в концентрические, то выбранные решения будут одного рода* (А *или* В).

Доказательство. Рассмотрим таблицу из трёх строк и шести столбцов, где строки соответствуют окружностям, а столбцы — решениям. На пересечении i-й строки и j-го столбца стоит бит «1», если i-я окружность касается j-го решения внешним образом, и бит «0» — если внутренним. Докажем, что найдутся такие две строки и четыре столбца, что четыре пары битов в выбранных столбцах совпадают с точностью до замены обоих битов в паре на противоположные.

Обозначим строки таблицы через a, b и c. Рассмотрим побитовые суммы $a\oplus b$, $b\oplus c$ и $c\oplus a$. Если в какой-то из них найдутся четыре одинаковых бита, то соответствующая пара строк и четвёрка столбцов является искомой. В противном случае в каждой из рассматриваемых побитовых сумм ровно три бита «0» и три бита «1», т. е. общее количество единиц в этих суммах равно 9. Но сумма рассматриваемых сумм равна, очевидно, нулевой строке, т. е. общее количество единиц в этих суммах чётно — противоречие.

Заменим теперь бит «1» на знак «+» и бит «0» на знак «−». Рассмотрим две окружности и четыре решения, соответствующие найденным нами строкам и столбцам. Заметим, что для двух окружностей существует всего лишь две комбинации знаков с точностью до замены обоих на противоположные: «++» или «+−». В случае непересекающихся окружностей одна из этих комбинаций соответствует решению, разделяющему исходные окружности, а другая — решению, не разделяющему их. Поэтому и после инверсии, переводящей выбранные окружности в концентрические, выбранные решения будут одновременно разделять или не разделять концентрические окружности, т. е. все принадлежать к роду A или все к роду B. □

Продолжим разбор второго случая теоремы 3. Применим лемму 8.2 и переведём соответствующие окружности в концентрические. Остаются две исходные окружности α и β, порождающие шесть общих решений, из которых четыре имеют один и тот же тип. По лемме 6.3 окружности α и β лежат либо на одном диаметре, либо на перпендикулярных диаметрах. Разберём эти случаи по отдельности.

8.2.1° Окружности α и β лежат на одном диаметре. Сделаем инверсию в точке пересечения симметричных решений (они найдутся согласно лемме 3.3), переводя их в пересекающиеся прямые a и b. Рассмотрим шесть решений в качестве исходных объектов (назовём их *новыми исходными объектами*), а четыре исходных объекта в качестве решений (назовём их *новыми решениями*). Тогда по лемме 6.6 новые решения имеют распределение 2–0–2–0 относительно a и b. Новые исходные объекты при этом должны иметь тип I_1 или I_2, и по лемме 5.2 (применённой к четырём новым решениям) получаем, что объектов каждого из этих типов не больше двух. Поэтому имеется новый исходный объект каждого из этих типов. После применения леммы 8.1 получаем конфигурацию 1.

8.2.2° Окружности α и β лежат на перпендикулярных диаметрах (рис. 17). Тогда по лемме 6.3 четыре общих решения для α и β имеют род B, а их центры образуют прямоугольник. Также α и β порождают ещё два общих решения рода A.

Пусть R и r — радиусы концентрических окружностей, $R > r$. Обозначим решения рода A через σ_1 и σ_2. Далее, пусть A_1, B_1 — точки касания σ_1 с α и β соответственно, A_2, B_2 — точки касания σ_2 с α и β соответственно.

Докажем теперь, что исходная четвёрка соответствует конфигурации 4. Сделаем инверсию с центром в O и радиусом $\sqrt{Rr}$. При такой инверсии все

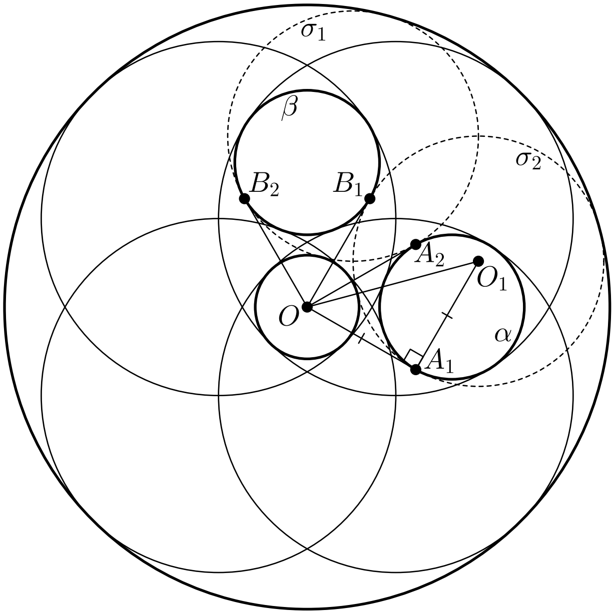


Рис. 17. Случай непересекающихся исходных окружностей

решения рода В перейдут в противоположные себе (так как степень точки O относительно всех решений рода В равна $-Rr$). Образы окружностей α и β должны касаться этих решений, поэтому α и β перейдут в себя. Это значит, что степень точки O относительно α и β равна Rr. Однако степень точки O относительно всех окружностей рода А также равна Rr, поэтому O лежит на радикальной оси σ_1 и α. Следовательно, OA_i — общая касательная к σ_i и α для $i = 1, 2$. Аналогично OB_i — общая касательная к σ_i и β для $i = 1, 2$, причём $OA_1 = OA_2 = OB_1 = OB_2 = \sqrt{Rr}$. Более того, $\angle A_1OB_1 = \angle A_2OB_2$, так как это угол, под которым все решения рода А видны из O. Поэтому и $\angle A_1OA_2 = \angle B_1OB_2$, а так как $OA_1 = OB_1$, то радиусы окружностей α и β равны. Тогда при повороте на $90°$ относительно O окружности α и β совпадут, а тогда A_1 и B_1 совпадут, т. е. $\angle A_1OB_1 = 90°$ (рис. 17).

Далее, пусть O_1 — центр σ_1. Он лежит на биссектрисе прямого угла $\angle A_1OB_1$. Поэтому $\triangle OA_1O_1$ — прямоугольный равнобедренный, откуда

$$OA_1 = A_1O_1 \quad\Leftrightarrow\quad \sqrt{Rr} = \frac{R-r}{2} \quad\Leftrightarrow\quad 4Rr = (R-r)^2 \quad\Leftrightarrow\quad R^2 - 6Rr + r^2 = 0.$$

Следовательно,

$$\frac{R}{r} = 3 + 2\sqrt{2},$$

что соответствует конфигурации 4. Теорема 3 доказана. □

Благодарности

Я признателен Ф. К. Нилову за постановку задачи и постоянное внимание к работе. Неоценимую помощь в подготовке и улучшении данного текста оказал М. Б. Скопенков. Также хотелось бы поблагодарить М. А. Волчкевича за формулировку задачи, из развития которой появилась данная работа.

Список литературы

[1] *Bruen A., Fisher J. C., Wilker J. B.* Apollonius by inversion // Math. Mag. 1983. Vol. 56, № 2. P. 97–103.

[2] *Cecil T. E.* Lie sphere geometry. N. Y.: Universitext, Springer. 1992.

[3] *Zlobec B. J., Mramor Kosta N.* Configurations of cycles and the Apollonius problem // Rocky Mountain J. Math. 2001. Vol. 31, № 2. P. 725–744.

[4] *Гальперин Г., Делман Ч.* Повесть о трёх кругах // Математическое просвещение. Сер. 3. Вып. 20. М.: МЦНМО, 2016. С. 111–134.

[5] *Жижилкин И. Д.* Инверсия. М.: МЦНМО, 2009.

Егор Александрович Морозов, НИУ ВШЭ

gorg.morozov@gmail.com

Выпуклая и комбинаторная геометрия

Пересечения выпуклых тел

Р. И. Просанов

§ 1. Теорема Хелли и её родственники

В настоящей статье мы приведём доказательства некоторых результатов комбинаторной геометрии о взаимном расположении выпуклых тел. Данные результаты хорошо известны специалистам, однако их популяризация представляет определённый интерес. Как это часто бывает с задачами комбинаторной геометрии, весь материал вполне доступен продвинутому школьнику (в частности, эта заметка основана на миникурсе, прочитанном автором на летней школе «Комбинаторика и алгоритмы»). Для достижения наибольшей общности мы сформулируем все результаты для многомерных объектов, однако даже для плоскости они нетривиальны и содержательны, не говоря уже о трёхмерном пространстве. В частности, если читатель не привык к многомерной геометрии, то мы призываем его рассматривать все утверждения на двумерных и трёхмерных примерах. Эту заметку можно воспринимать как небольшой, но достаточно широкий экскурс в современную комбинаторную геометрию: для достижения основного результата нам придётся познакомить читателя с рядом теорем и концепций, имеющих самостоятельную ценность. Мы будем во многом следовать замечательной англоязычной книге Матушека [14]. Отметим, что в последние годы в связи с появлением работ, в которых задачи

Часть данной работы была выполнена в рамках проекта Swiss National Science Foundation 200021_169391.

комбинаторной геометрии решаются топологическими методами, интерес к изложенным результатам лишь усиливается.

Мы предполагаем, что читатель уже знаком с понятием выпуклости. Выпуклость — очень сильное условие, позволяющее избегать нежелательных конструкций, но всё ещё оставляющее большой простор для разнообразия форм. Кроме того, изучение выпуклых тел часто мотивировано различными приложениями математики. Мы надеемся, что читатель также сталкивался с основным фактом комбинаторной геометрии выпуклых тел: известной теоремой Хелли (а если нет, то мы отсылаем его к замечательной статье Протасова [2]). В любом случае напомним её формулировку.

Теорема 1.1 (Хелли, 1913). *Пусть* $F_1, F_2, \ldots, F_n$ *— выпуклые множества в* $\mathbb{R}^d$*, причём любые* $d+1$ *из них имеют общую точку. Тогда все множества имеют общую точку.*

Спустя много лет, прошедших после первого знакомства автора с этим утверждением, оно всё ещё кажется ему удивительным. Несмотря на сколь угодно большое количество тел, огромное разнообразие возможных форм, локальное условие пересечения влечёт глобальное. Благодаря простоте формулировки и изяществу доказательства, теорема Хелли замечательно демонстрирует эстетическую составляющую математики и, видимо, по этой причине является стандартным материалом многих математических кружков.

В данной статье мы обсудим некоторое обобщение теоремы Хелли. Тут стоит заметить, что обобщений у неё существует огромное количество. На пике популярности вокруг теоремы Хелли была развёрнута целая индустрия по производству разнообразных схожих результатов. Подробности можно прочитать в обзоре [1]. Однако теорема, о которой мы будем говорить, выгодно выделяется среди многих. Это так называемая (p,q)-теорема. Её предложили в качестве гипотезы Гуго Хадвигер и Ганс Дебруннер ещё в 1957 году [10], но полностью доказать смогли только в 1992 году выдающиеся математики Нога Алон и Дэниэл Клейтман [5]. Как мы уже отметили, доказательство этой теоремы примечательно прежде всего тем, что на нём можно продемонстрировать целый ряд методов комбинаторной геометрии, нашедших применение во многих других ситуациях.

Теорема 1.2 ((p,q)-теорема, 1992). *Пусть даны натуральные числа* $p \geqslant q \geqslant d+1$ *и* $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ *— семейство выпуклых множеств в* $\mathbb{R}^d$*, причём среди любых* p *из этих множеств найдутся* q *имеющих общую точку. Тогда существует такая константа* $c = c(p,q,d)$ *(не зависящая от* n*) и такое конечное множество точек* X*,* $|X| \leqslant c$*, что любое* F_i *содержит хотя бы одну точку из* X*.*

Другими словами, при наложении некоторых локальных условий на пересечения (куда более слабых, чем в теореме Хелли) все тела можно одновременно «проткнуть» с помощью ограниченного набора точек (вне зависимости от количества тел в системе, которое может стремиться к бесконечности при фиксированных p и q).

В дальнейшем, если для семейства выпуклых множеств выполнено условие этой теоремы, то будем говорить, что на семейство наложено (p, q)-условие.

Точное значение $c(p, q, d)$ известно только в случае $p < 2(q - 1)$ и равно $p - q + 1$, доказательство этого факта предоставляется читателю. Во всех остальных случаях оно неизвестно. В данной статье мы не будем предлагать никаких оценок $c(p, q, d)$. Тогда теорема сводится к тому, что можно выбрать набор точек размера, не зависящего от n. Поэтому можно положить q наименьшим, для которого утверждение имеет смысл, т. е. взять $q = d + 1$. В самом деле, если среди любых p множеств найдутся $q \geqslant d+1$ множеств, имеющих общую точку, то, конечно, можно утверждать, что найдётся $d + 1$ множество, имеющее общую точку. Так что для всех больших q достаточно положить $c(p, q, d) = c(p, d + 1, d)$.

В ближайших разделах мы познакомим читателя с рядом существенных результатов комбинаторной геометрии, а затем продемонстрируем, как это всё приводит к доказательству (p, q)-теоремы.

§ 2. Дробная теорема Хелли

Первый нужный нам факт является другим естественным обобщением теоремы Хелли. Пусть снова дано семейство из n выпуклых множеств в $\mathbb{R}^d$. Всего мы имеем C_n^{d+1} наборов точек этих множеств размера $d+1$. Если для каждого из этих наборов найдётся общая точка, то по теореме Хелли все множества имеют общую точку. Хороший вопрос: а если не все, но «почти все» из таких наборов имеют общую точку, можно ли что-то утверждать про всё семейство? Интуитивно хочется сказать, что тогда и «почти все» множества должны пересекаться в какой-то точке. В этом и состоит утверждение дробной теоремы Хелли, которую мы сейчас докажем.

Теорема 2.1 (Качальский и Лю, 1979, [12]). *Пусть дано число $\alpha > 0$, и пусть $F_1, F_2, \ldots, F_n$ — выпуклые множества в $\mathbb{R}^d$, причём известно, что имеется αC_n^{d+1} наборов $F_{i_1}, F_{i_2}, \ldots, F_{i_{d+1}}$, каждый из которых пересекается в одной точке. Тогда существует такое β, зависящее только от α и d, но не от n, что не менее βn множеств среди F_i имеют общую точку.*

Мы докажем, что достаточно взять $\beta = \alpha/(d + 1)$. Это весьма слабая оценка, в частности, если $\alpha \to 1$, то β стремится к $1/(d + 1)$, хотя, казалось бы, при этом дробная теорема Хелли должна переходить в обычную. И действительно, Гил Калаи [11] доказал, что при $\alpha < 1$ можно взять $\beta = 1 - (1-\alpha)^{1/(d+1)}$, — с такой оценкой мы получим обычную теорему Хелли из дробной. Однако доказательство Калаи существенно выходит за рамки данной статьи.

Доказательство. В дальнейшем под многогранником будем подразумевать выпуклый многогранник (политоп) в пространстве $\mathbb{R}^d$. Если читатель до сих пор не слишком часто сталкивался с многомерными многогранниками, то предлагаем ему рассматривать всё для случая плоскости и трёхмерного пространства, а также ознакомиться с прекрасной книгой [4]. В целом достаточно понимать, что многомерный многогранник — это просто естественное обобщение понятий трёхмерного многогранника и двумерного многоугольника на пространства большей размерности. Будем пользоваться следующим определением многогранника: (выпуклый) многогранник в $\mathbb{R}^d$ есть выпуклая оболочка конечного множества точек из $\mathbb{R}^d$. С понятием выпуклой оболочки можно подробнее ознакомиться в [4].

Мы сейчас проделаем важный трюк, который понадобится нам и в дальнейшем. Покажем, что если доказать дробную теорему Хелли для многогранников, то очень легко получить её и для любых выпуклых множеств.

Пусть дано множество $\{1, \ldots, n\}$ всех индексов, нумерующих наши выпуклые множества $F_1, F_2, \ldots, F_n$. Выберем произвольное подмножество $I \subset \{1, \ldots, n\}$ и положим $F_I = \bigcap_{i \in I} F_i$. Если F_I непусто, то выберем точку $\bar{z}_I$ в этом пересечении (здесь и далее мы пишем горизонтальную черту над символами точек пространства, чтобы отличать их от чисел). В качестве индекса здесь взято не число, а подмножество I исходного множества индексов.

Теперь положим $F'_j = \operatorname{conv}(\{\bar{z}_I : j \in I\})$, где $\operatorname{conv}(X)$ обозначает выпуклую оболочку множества X. Другими словами, для каждого множества F_j мы рассматриваем все подмножества индексов I, содержащие j, и если F_I непусто, то выбираем из него точку $\bar{z}_I$, после чего рассматриваем выпуклую оболочку всех взятых точек, получая некоторый многогранник. Автор оставляет читателю следующие упражнения: для любого j имеем $F'_j \subseteq F_j$, а для любого подмножества индексов I множество $F'_I = \bigcap_{i \in I} F'_i$ непусто тогда и только тогда, когда F_I непусто.

Таким образом, наше множество точек $\bar{z}_I$ и соответствующее ему семейство $F'_1, \ldots, F'_n$ несут в себе всю комбинаторную информацию об устройстве пересечений множеств из нашего исходного семейства. Мы не будем писать штрихи в дальнейшем, а будем просто предполагать, что наше семейство состоит из многогранников.

Для доказательства теоремы определим *лексикографический порядок* на точках из $\mathbb{R}^d$, т. е. научимся сравнивать их друг с другом. Введём обычную систему координат и скажем, что точка $\bar{x} = (x_1, \ldots, x_d)$ лексикографически больше точки $\bar{y} = (y_1, \ldots, y_d)$, если при некотором i имеем $x_j = y_j$ для всех $j < i$, но $x_i > y_i$. Читателю предлагается в качестве упражнения описать, как для данной точки $\bar{x}$ выглядит множество точек в $\mathbb{R}^d$, лексикографически меньших этой точки.

Лексикографическим минимумом множества назовём лексикографически минимальную среди всех точек данного множества, если такая существует. Заметим, что для любого многогранника такая точка заведомо существует и единственна. (Это верно и для произвольного компакта.) Например, рассмотрим многоугольник на плоскости. Возьмём горизонтальную прямую ниже этого многоугольника и будем двигать её вверх, пока прямая не коснётся многоугольника. Если она касается в вершине, это и есть искомая точка минимума, а если она касается по стороне, то возьмём самую левую вершину этой стороны. Читателю предлагается обобщить это рассуждение на многогранники большей размерности.

Выберем некоторое подмножество индексов $I \subset \{1, \ldots, n\}$ размера не меньше d, для которого $F_I \neq \varnothing$. Тогда найдётся такое подмножество $J \subset I$ размера d, для которого лексикографические минимумы множеств F_I и F_J совпадают. Это значит, что точка минимума любого пересечения в нашем семействе всегда задаётся пересечением только каких-то d множеств семейства.

В самом деле, пусть $\bar{u}$ — точка лексикографического минимума в F_I. Рассмотрим множество A точек, меньших $\bar{u}$. Заметим, что это множество выпуклое. Поскольку $\bar{u}$ — минимум в F_I, то $A \cap F_I = \varnothing$. Но тогда по теореме Хелли среди набора $\{F_i\colon i \in I\} \cup A$ должно найтись $d + 1$ множество с пустым пересечением, иначе у всех множеств набора будет общая точка. Среди них обязательно должно быть A, так как F_I непусто. Но тогда оставшиеся d множеств и есть искомые, а их индексы составляют нужное подмножество J.

Для каждого из αC_n^{d+1} наборов индексов I, для которых $F_I \neq \varnothing$, мы выбираем такой набор $J = J(I)$ размера d, что F_J и F_I имеют общую точку лексикографического минимума. Тогда некоторый d-набор J_0 будет сопоставлен не менее чем

$$\frac{\alpha C_n^{d+1}}{C_n^d} = \alpha \frac{n-d}{d+1}$$

различным $(d+1)$-наборам I. Каждый такой $(d+1)$-набор имеет вид $J_0 \cup \{i\}$ для некоторого индекса i. Но тогда пересечение любого такого $(d+1)$-набора будет иметь одну и ту же точку лексикографического минимума, а значит, все они будут иметь общую точку. Получается, мы нашли не менее

$$d+\alpha\frac{n-d}{d+1}>\alpha\frac{n}{d+1}$$

множеств, имеющих общую точку. □

§ 3. О том, как выбрать точку внутри многих симплексов

Другим важным фактом окажется первая лемма о выборе.

Выпуклая оболочка множества из $d+1$ точек, не лежащих в общей $(d-1)$-мерной гиперплоскости, называется *d-симплексом.* Симплекс — многомерное обобщение понятий треугольника и тетраэдра, он является примером простейшего многогранника. Если же мы берём выпуклую оболочку $d+1$ точки, которые лежат в общей d-мерной гиперплоскости, то это будет называться *вырожденным $(d-1)$-симплексом.*

Пусть дано множество из n точек в $\mathbb{R}^d$. Тогда любой набор из $d+1$ точки определяет некоторый симплекс, а всего таких симплексов будет C_n^{d+1}. Ясно, что многие из этих симплексов будут пересекаться, причём достаточно часто в одной точке может пересекаться несколько симплексов. Естественный вопрос: а как много симплексов обязательно должны пересекаться одновременно? В этом и состоит первая лемма о выборе.

Теорема 3.1 (первая лемма о выборе, Барань, 1982, [6]). *Для любого множества S из n точек (n достаточно большое) найдётся точка $\mathbb{R}^d$, лежащая не менее чем в cC_n^{d+1} всех симплексов, определяемых точками S, где $c=c(d)$ — некоторая константа, не зависящая от n.*

Наибольший интерес представляет случай, когда S — множество точек общего положения в $\mathbb{R}^d$. Это значит, что никакие $k+1$ из них не лежат в общей $(k-1)$-мерной гиперплоскости при любом $k<d$. Если это условие не выполнено, то многие симплексы будут вырождены.

Для плоскости существует следующая точная оценка.

Теорема 3.2 (Борош, Фюреди, 1984, [7]). *Для любого множества S из n точек общего положения на плоскости (n достаточно большое) найдётся точка, лежащая не менее чем в $\frac{2}{9}C_n^3$ треугольников, определяемых точками S.*

Для больших размерностей сам Барань получил оценку $c\geqslant\frac{1}{(d+1)^d}$. Интерес к этой задаче поднялся с новой силой после новаторской работы Громова [9], в которой он принципиально новыми методами получил оценку $c\geqslant\frac{2d}{(d+1)(d+1)!}$ (см. также подробности в [15]).

Ниже мы сперва комбинаторно докажем лемму о выборе в общем случае (с плохой оценкой на c), а затем приведём адаптацию чудесного короткого доказательства Бориса Буха [8] теоремы Бороша — Фюреди (в незначительно ослабленной форме).

Предварительно нам понадобится следующий классический результат.

Теорема 3.3 (теорема Каратеодори). *Пусть X — конечное множество точек*в $\mathbb{R}^d$, $\bar{x} \in \operatorname{conv}(X)$. *Тогда существует такое множество $X_0 \subseteq X$, $|X_0| \leqslant d+1$, что $\bar{x} \in \operatorname{conv}(X_0)$.*

Доказательство. Воспользуемся тем, что $\operatorname{conv}(X)$ — выпуклый многогранник. Без ограничения общности, пусть его размерность равна d. Проведём индукцию по d. База очевидна. Для перехода индукции выберем произвольную точку $\bar{x}_0 \in X$. Луч $\overrightarrow{\bar{x}_0\bar{x}}$ начинается в многограннике $\operatorname{conv}(X)$ и пересекает его границу в некоторой точке $\bar{x}'$. Она лежит в некоторой грани нашего многогранника, т. е. в многограннике $\operatorname{conv}(X')$ размерности $d_0 < d$ (где $X' \subset X$). Тогда по предположению индукции $x' \in \operatorname{conv}(X_0')$, где $X_0' \subseteq X' \subset X$, $|X_0'| \leqslant d_0 + 1 \leqslant d$. Положим $X_0 = X_0' \cup \{\bar{x}_0\}$. Читателю предоставляется проверить, что тогда $x \in \operatorname{conv}(X_0)$. □

Теорема Каратеодори верна и без предположения о конечности множества X; убедиться в этом мы снова предоставляем читателю.

Теперь приступим к доказательству первой леммы о выборе.

Доказательство. Пусть $\mathcal{F}$ — множество всех симплексов с вершинами в S (возможно, вырожденных). Мы хотим доказать, что $\mathcal{F}$ удовлетворяет условиям дробной теоремы Хелли для некоторого α — отсюда автоматически будет следовать первая лемма о выборе.

Всего мы имеем $C_{C_n^{d+1}}^{d+1}$ наборов симплексов размера $d+1$. Это количество — многочлен по n степени $(d+1)^2$.

Будем пока рассуждать чисто комбинаторно, не привлекая никакой геометрии. Пусть $t = d(d+1)^2 + 1$. Выберем некоторое подмножество $A \subset S$ размера t. Пусть $s = d^2(d+1) + 1$. Тогда любые $d+1$ подмножеств в A размера s имеют общую точку. В самом деле, в $d+1$ множеств размера s входят с учётом кратностей $s(d+1) = d^2(d+1)^2 + d + 1$ точек. Но если общей точки нет, то каждая точка входит не более чем в d множеств, а всего точек с учётом кратностей не больше, чем $td = d^2(d+1)^2 + d$, — противоречие.

Итак, любые $d+1$ подмножеств размера s содержат общую точку, поэтому их выпуклые оболочки пересекаются хотя бы в одной точке. Тогда по теореме Хелли выпуклые оболочки всех подмножеств в A размера s имеют общую точку $\bar{x}$. В частности, $\bar{x} \in \operatorname{conv}(A)$. По теореме Каратеодори можно выбрать такое подмножество A_1 размера $d+1$, что $\bar{x} \in \operatorname{conv}(A_1)$. (Теорема Каратеодори позволяет выбрать такое множество размера не больше $d+1$, но если оно меньше, то можно добавить туда произвольные точки из A.) Однако множество $A \backslash A_1$ всё ещё содержит не менее s точек. Значит, $\bar{x} \in \operatorname{conv}(A \backslash A_1)$. Снова выберем такое $A_2 \subseteq A \backslash A_1$ размера $d+1$, что $\bar{x} \in \operatorname{conv}(A_2)$. Мы можем продолжать так делать, пока у нас остаётся не менее s точек, а значит, можно повторить эту операцию $d+1$ раз. Получается, что мы можем выбрать из A такой набор из $d+1$ множеств размера $d+1$, не имеющих общих вершин, что выпуклые оболочки этих множеств пересекаются в общей точке $\bar{x}$.

Итак, из любого множества размера t мы можем выбрать пересекающийся $(d+1)$-набор из симплексов, причём все вершины этих симплексов различны. Каждый такой набор, в свою очередь, будет сопоставлен не более чем $C_n^{t-(d+1)^2}$ множествам размера t. В конечном счёте получаем не менее $C_n^t / C_n^{t-(d+1)^2}$ пересекающихся $(d+1)$-наборов. Это выражение — многочлен по n степени $(d+1)^2$. Теперь ясно, что для некоторого $\alpha > 0$ и любого достаточно большого n это выражение больше, чем $\alpha C_{C_n^{d+1}}^{d+1}$, что и требовалось доказать. $\square$

Теперь докажем отдельно версию для плоскости.

Доказательство. Мы докажем чуть более слабое утверждение: в качестве c можно взять любое число, меньшее $2/9$.

Пусть дано множество S из n точек общего положения. Покажем сначала, что найдутся такие три прямые, пересекающиеся в одной точке (необязательно из S), что в каждой из 6 частей, на которые они делят плоскость (без учёта самих прямых), лежит не менее $\lfloor n/6 \rfloor - 3 \geqslant n/6 - 4$ точек (а при достаточно аккуратном рассуждении можно получить $\lfloor n/6 \rfloor - 1$).

Направлением назовём некоторый вектор, задаваемый точкой на единичной окружности. Для некоторого направления $\bar{v}$, которое будем считать горизонтальным, рассмотрим множество всех прямых, параллельных $\bar{v}$ и таких, что по каждую сторону от всякой такой прямой лежит не меньше $\lfloor n/2 \rfloor - 1$ точек. Читателю предлагается доказать самостоятельно, что такие прямые обязательно существуют. Множество всех таких прямых образует некоторую полосу (возможно нулевой толщины, когда искомая прямая только одна). Пусть l — средняя прямая из этой полосы. Обозначим через S_+ и S_- части множества S, лежащие сверху и снизу от прямой l. Покажем, что на ней найдётся такая точка P, что через неё можно провести прямую, делящую S_+ и S_- в отношении $1:2$ одновременно.

Выберем направление $\bar{u} \neq \pm\bar{v}$. Для такого направления аналогично предыдущему существует некоторая полоса прямых, параллельных $\bar{u}$ и таких, что они делят множество S_+, причём сверху лежит не менее

$$\left\lfloor \frac{|S_+|}{3} \right\rfloor - 1 \geqslant \left\lfloor \frac{n}{6} \right\rfloor - 2$$

точек из S_+, а снизу точек не менее

$$\left\lfloor \frac{2|S_+|}{3} \right\rfloor - 1 \geqslant \left\lfloor \frac{n}{3} \right\rfloor - 2.$$

Пусть эта полоса пересекает прямую l по отрезку $I_{\bar{u}+}$. Построим аналогичную полосу для множества S_-, но только чтобы, наоборот, сверху каждой прямой из полосы было около $2/3$ всех точек из S_-, а снизу около $1/3$. Обозначим через $I_{\bar{u}-}$ отрезок пересечения этой полосы с l. Пусть $\bar{u}$ пробегает все направления от $-\bar{v}$ до $\bar{v}$ по часовой стрелке. Проследим за изменениями положения отрезков $I_{\bar{u}-}$ и $I_{\bar{u}+}$. Ясно, что каждый из них пробежит (возможно, изменяя размеры) по всей прямой l, но в противоположных направлениях. А значит, где-то эти отрезки пересекутся, т. е. искомая точка P найдётся. Несложно заметить, что все такие точки образуют некоторый отрезок (возможно, нулевой длины). Выберем в качестве P середину этого отрезка и проведём через неё ту самую прямую, которая делит S_+ и S_- в отношении примерно $1:2$.

Обозначим через S_1 ту часть S_+, где теперь находится примерно две трети его точек, а через S_2 — такую же часть множества S_-. Из P можно провести некоторое множество лучей, делящих S_1 почти пополам (как и раньше — с точностью до одной точки), и некоторое множество лучей, делящих почти пополам множество S_2. Понятно, что в обоих случаях все такие лучи образуют некоторый угол. Если среди них найдутся два луча, составляющих вместе прямую, то мы нашли то, что искали. Иначе мы проведём в каждом угле биссектрису и обозначим через r_1 и r_2 прямые, которые содержат эти биссектрисы. Обозначим через γ *ориентированный угол* между этими прямыми: мы берём меньший из углов, но если он получается переходом по часовой стрелке от r_1 к r_2, то берём его со знаком плюс, а если против часовой, то со знаком минус.

Заметим теперь, что прямая l, точка P и угол γ зависят от направления $\bar{v}$ непрерывно. Повернём теперь $\bar{v}$ по часовой стрелке к $-\bar{v}$. Прямая l и точка P при этом перейдут сами в себя, а вот прямые r_1 и r_2 поменяются местами, а значит, наш ориентированный угол γ поменяет знак. По теореме о промежуточном значении непрерывной функции найдётся направление, для которого угол между прямыми r_1 и r_2 равен 0, т. е. соответствующие лучи, поскольку они всегда направлены в разные стороны, образуют прямую. Таким образом мы найдём искомые три прямые, P — это точка их пересечения.

Покажем, что точка P искомая. В самом деле, выберем в каждой из шести частей по точке из S в порядке обхода: A, B, C, D, E и F. Эти шесть точек образуют всего 20 треугольников. Заметим, что точка P лежит не менее чем в восьми из этих треугольников. Действительно, она лежит в треугольниках ACE и BDF. Рассмотрим теперь все четырёхугольники типа $ABCE$ — три точки подряд и четвёртая через одну. Проведём диагональ BE. Очевидно, что P лежит или в треугольнике ABE, или в треугольнике BCE. Добавится ещё шесть треугольников. Тем не менее каждый такой треугольник может участвовать во многих шестиугольниках (но не более чем в $(n/6+20)^3$, так как каждая из шести частей плоскости содержит не более $n/6+20$ точек из S). Поэтому P лежит не менее чем в $\dfrac{8(n/6-4)^6}{(n/6+20)^3}$ треугольниках.

Это выражение порядка $n^3/27$. Такой же старший член у многочлена $\frac{2}{9}C_n^3$. Тогда для любого $c<2/9$ и достаточно большого n это выражение больше, чем cC_n^3. Следовательно, точка P будет гарантированно лежать в cC_n^3 треугольниках, что и требовалось доказать. □

§ 4. Дробные упаковки и трансверсали

В этом разделе мы познакомим читателя с понятиями, важнейшими не только для комбинаторной геометрии, но и для комбинаторики в целом.

(p,q)-Теорема относится к частному случаю следующего общего вопроса: дано некоторое семейство множеств, мы хотим выбрать такой набор элементов этих множеств, чтобы каждое из них содержало хотя бы один элемент выбранного набора. Этот набор называется *трансверсалью* семейства. Комбинаторика очень часто изучает вопросы, связанные с построением трансверсалей для определённых множеств.

Дадим все необходимые определения в максимально общей ситуации. Пусть X — произвольное множество, а $\mathcal{F}$ — семейство его подмножеств (в комбинаторике такой объект называется ещё гиперграфом). *Трансверсалью* назовём такое подмножество $T\subset X$, что $F\cap T\neq\varnothing$ для всех F из семейства $\mathcal{F}$. *Трансверсальным числом* семейства $\tau(\mathcal{F})$ назовём минимальный размер трансверсали $\mathcal{F}$ (который может быть равен бесконечности). Например, можно взять в качестве X множество всех точек в $\mathbb{R}^d$, а в качестве $\mathcal{F}$ — некоторое конечное семейство выпуклых множеств. Тогда (p,q)-теорема утверждает, что при (p,q)-условии, наложенном на $\mathcal{F}$, трансверсальное число этого семейства не зависит от n, т. е. от размера $\mathcal{F}$.

Нам понадобится другая важная характеристика семейства $\mathcal{F}$, тесно связанная с предыдущей. *Упаковкой* назовём такой набор $\mathcal{P}\subset\mathcal{F}$, что никакие два подмножества из $\mathcal{P}$ не пересекаются. В свою очередь, *упаковоч-*

ным числом $\nu(\mathcal{F})$ назовём размер максимальной упаковки в $\mathcal{F}$. На всякий случай отметим: трансверсаль — это подмножество самого X, а упаковка — подмножество (подсемейство) в $\mathcal{F}$.

Например, если X — множество вершин некоторого графа, а $\mathcal{F}$ — множество его рёбер, то упаковка называется *паросочетанием.* Задача построения оптимального паросочетания имеет огромное практическое значение, так как применяется в различных областях: от анализа компьютерных сетей до моделирования транспортных потоков. (*Вопрос:* а что будет называться трансверсалью графа?)

Некоторые из читателей могут быть знакомы с известной теоремой Кёнига. На нашем языке она говорит, что если $\mathcal{F}$ — множество рёбер некоторого двудольного графа[1], то $\nu(\mathcal{F}) = \tau(\mathcal{F})$. Читателю предлагается показать, что любая трансверсаль по размеру должна быть не меньше, чем размер любой упаковки: $\nu(\mathcal{F}) \leqslant \tau(\mathcal{F})$.

В общем случае о связи этих величин больше ничего нельзя сказать наверняка. Например, $\tau(\mathcal{F})$ даже при $\nu(\mathcal{F}) = 1$ может быть сколь угодно большим. Хорошим упражнением для читателя является построение такого примера.

Задача поиска или хотя бы оценки τ или ν для заданного $\mathcal{F}$ имеет важное прикладное значение. Но вычисление точных значений этих величин есть алгоритмически сложная (NP-полная) задача. Чтобы получить их оценку, придётся ввести следующие понятия. Пусть X — конечное множество. Припишем всем элементам веса́ от 0 до 1, т. е. определим некоторую функцию из множества X в отрезок $[0; 1]$. Существует естественный способ сопоставить каждому подмножеству $Y \subset X$ такую функцию: припишем элементу *вес* 1, если он входит в Y, и 0, если он не входит. Такая функция называется *характеристической функцией* подмножества Y. *Вес* произвольного множества равен сумме весов его элементов. Трансверсаль T для $\mathcal{F}$ с каждым $F \in \mathcal{F}$ пересекается хотя бы по одному элементу. А значит, взяв в качестве весовой функции характеристическую функцию множества T, получим, что суммарный вес каждого F равен хотя бы 1. Если теперь для произвольной весовой функции вес F не меньше 1, то скажем, что она «покрывает» F. Назовём функцию *дробной трансверсалью* семейства $\mathcal{F}$, если она покрывает каждое $F \in \mathcal{F}$. Общим весом $S(\phi)$ весовой функции назовём суммарный вес всех точек в X:

$$S(\phi) = \sum_{x \in X} \phi(x).$$

[1] Граф называется двудольным, если его вершины можно разбить на два множества так, что вершины любого ребра принадлежат разным множествам.

Дробным трансверсальным числом $\tau^*(\mathcal{F})$ назовём минимальный общий вес дробной трансверсали для $\mathcal{F}$ (а точнее — точную нижнюю грань всех таких весов). Поскольку, как было описано выше, любая трансверсаль переходом к характеристической функции порождает дробную трансверсаль, то верно неравенство $\tau^*(\mathcal{F}) \leqslant \tau(\mathcal{F})$.

Рассмотрим пример. Возьмём граф «треугольник»: в качестве X возьмём его вершины, а в качестве $\mathcal{F}$ его рёбра. Тогда ясно, что $\tau(\mathcal{F}) = 2$. Но несложно заметить, что можно выбрать дробную трансверсаль с суммарным весом $3/2$, если приписать каждой вершине вес $1/2$. Читателю предлагается доказать, что лучше выбрать весовую функцию уже не получится, т. е. $\tau^*(\mathcal{F}) = 3/2$.

Вычисление τ^* может дать хорошую оценку на τ. Помимо нижней оценки $\tau^*(\mathcal{F}) \leqslant \tau(\mathcal{F})$ также известна полученная Ловасом [13] верхняя оценка

$$\tau(\mathcal{F}) \leqslant \tau^*(\mathcal{F})(1 + \ln \delta(\mathcal{F})),$$

где $\delta(\mathcal{F})$ есть максимальное количество множеств в $\mathcal{F}$, имеющих общую точку. Заметим, что вычисление τ^* в конкретной ситуации есть алгоритмически быстрая задача (решаемая за полиномиальное время).

Аналогично можно определить дробные упаковки. Вместо того, чтобы приписывать веса элементам из X, будем приписывать их множествам из $\mathcal{F}$, т. е. определим весовую функцию из $\mathcal{F}$ в $[0; 1]$. Подобно случаю трансверсалей, для произвольного подсемейства $\mathcal{P} \subset \mathcal{F}$ определена характеристическая весовая функция, приписывающая множеству 1, если оно входит в $\mathcal{P}$, и 0 в противном случае. Теперь нужно выделить свойство, которое описывает характеристические функции упаковок. Семейство множеств $\mathcal{P}$ является упаковкой, если входящие в него множества не пересекаются, а значит, каждый элемент X покрывается не более чем одним множеством из семейства. Получаем, что суммарный вес множеств над каждым элементом X не превосходит 1. Мы можем определить *дробную упаковку*: это такая весовая функция ψ, что для каждого $x \in X$

$$\sum_{F \in \mathcal{F}:\, x \in F} \psi(F) \leqslant 1.$$

Общий вес дробной упаковки — это $S(\psi) = \sum_{F \in \mathcal{F}} \psi(F)$, а *дробное упаковочное число* $\nu^*(\mathcal{F})$ — это максимальный общий вес дробной упаковки в $\mathcal{F}$. Аналогично из определения получаем неравенство $\nu(\mathcal{F}) \leqslant \nu^*(\mathcal{F})$.

Прежде чем читатель продвинется дальше, автор советует ему убедиться, что для недавно разобранного примера $\nu^*(\mathcal{F})$ также равно $3/2$.

Оказывается, за этим совпадением стоит куда более общий принцип. Имеет место

Теорема 4.1 [14]. *Если X — конечное множество, то для любого семейства $\mathcal{F}$ выполнено равенство $\nu^*(\mathcal{F}) = \tau^*(\mathcal{F})$. Более того, в обоих случаях это значение достигается на некоторых функциях, которые принимают только рациональные значения.*

Доказательство данной теоремы не слишком сложно, но всё же выходит за рамки материала данной заметки. Упомяну, что она вытекает из так называемой *двойственности линейного программирования*. Оказывается, задачи нахождения $\nu^*(\mathcal{F})$ и $\tau^*(\mathcal{F})$ можно сформулировать как задачи нахождения минимума и максимума некоторых линейных векторных функций при определённых условиях. Существует «двойственный переход», который в данной ситуации переводит задачу нахождения $\nu^*(\mathcal{F})$ в задачу нахождения $\tau^*(\mathcal{F})$, и известна теорема, утверждающая, что решения исходной задачи и двойственной к ней совпадают.

Мы дали все определения лишь для конечных множеств, но в геометрических задачах в качестве X обычно берётся некоторое бесконечное множество. Например, чтобы сформулировать (p, q)-теорему в этих терминах, мы определяли выше X как множество всех точек $\mathbb{R}^d$. Конечно, дробные трансверсали и упаковки можно определить и для бесконечных множеств, однако это требует достаточной аккуратности. Тем не менее, нам это не понадобится. Оказывается, уже данных выше определений достаточно для получения интересных геометрических приложений.

§ 5. Как (p, q)-теорема связана с дробными упаковками

Пусть $\mathcal{F} = \{F_1, F_2, \ldots, F_n\}$ — некоторое семейство выпуклых множеств в $\mathbb{R}^d$. Как и в начале доказательства дробной теоремы Хелли, для каждого подмножества индексов $I \subset \{1, \ldots, n\}$, если пересечение множеств с такими номерами непусто, выберем в этом пересечении некоторую точку $\overline{x}_I$. Возьмём в качестве X совокупность всех этих точек (это множество конечно) и сопоставим каждому множеству $F \in \mathcal{F}$ множество $F' = F \cap X$. Обозначим через $\mathcal{F}'$ семейство, состоящее из всех таких F'. Легко заметить, что набор множеств $F_{i_1}, \ldots, F_{i_k}$ пересекается тогда и только тогда, когда пересекаются $F'_{i_1}, \ldots, F'_{i_k}$. В частности, трансверсаль для $\mathcal{F}'$ — это трансверсаль для $\mathcal{F}$. По сути, пара $(X, \mathcal{F}')$ — это конечный гиперграф, который содержит всю информацию о пересечениях множеств из $\mathcal{F}$. В дальнейшем мы будем опускать штрих, но, говоря о пересечении нескольких множеств из $\mathcal{F}$, всегда будем помнить, что можно выбрать точку пересечения, лежащую в X. Также будем говорить о дробных трансверсалях и дробных упаковках системы $\mathcal{F}$, имея в виду, что мы определяем их по отношению ко множеству X.

В этом разделе мы докажем, что если $\mathcal{F}$ удовлетворяет $(p, d+1)$-условию, то $\nu^*(\mathcal{F}) \leqslant c$, где c — некоторая константа, не зависящая от n.

Сперва заметим, что семейство $\mathcal{F}$ удовлетворяет условию дробной теоремы Хелли. В самом деле, каждый набор из p множеств включает как минимум один пересекающийся $(d+1)$-набор. С другой стороны, каждый такой набор входит в C_{n-d-1}^{p-d-1} p-наборов. Тогда ясно, что найдутся не менее

$$\frac{C_n^p}{C_{n-d-1}^{p-d-1}} = \frac{C_n^{d+1}}{C_p^{d+1}}$$

пересекающихся $(d+1)$-наборов. Так что условие дробной теоремы Хелли выполнено с $\alpha = (C_p^{d+1})^{-1}$.

Поэтому существует такое β, не зависящее от n, что не менее чем βn множеств из $\mathcal{F}$ имеют общую точку (ясно, что эту точку можно выбрать в X).

Очевидно, что число β связано с общим весом дробных упаковок. Например, если дробная упаковка принимает на каждом множестве одинаковое значение, то это значение не может быть больше, чем $1/(\beta n)$ (иначе у точки пересечения βn множеств будет вес больше 1), а значит, общий вес такой упаковки не может быть больше, чем $1/\beta$.

Пусть ψ — оптимальная дробная трансверсаль, общий вес которой равен $\nu^*(\mathcal{F})$. В прошлом разделе упоминалось, что её можно выбрать так, чтобы она принимала только рациональные значения. Приведём все эти значения к некоторому общему знаменателю D. Пусть $F \in \mathcal{F}$, тогда $\psi(F)$ имеет вид $m(F)/D$. Рассмотрим новое семейство $\widetilde{\mathcal{F}}$: возьмём туда $m(F)$ копий каждого элемента $F \in \mathcal{F}$. Каждую копию теперь мы будем рассматривать как отдельное множество.

Заметим, что для этого семейства выполнено $(d(p-1)+1, d+1)$-условие. В самом деле, возьмём $d(p-1)+1$ множество из $\widetilde{\mathcal{F}}$. Тогда этот набор или содержит копии p различных множеств из $\mathcal{F}$, или он содержит не меньше $d+1$ копий одного и того же множества. В любом случае найдётся $d+1$ пересекающееся множество.

А раз для $\widetilde{\mathcal{F}}$ выполнено $(d(p-1)+1, d+1)$-условие, то для него верна дробная теорема Хелли: пусть N — это размер $\widetilde{\mathcal{F}}$, т. е. $N = \sum_{F \in \mathcal{F}} m(F)$; тогда существует такое β', не зависящее от n, что не меньше $\beta' N$ множеств из $\widetilde{\mathcal{F}}$ имеют общую точку (мы также можем выбрать эту точку из X). Обозначим эту точку $\bar{a}$. Заметим, что $\nu^*(\mathcal{F}) = N/D$. Имеем неравенства

$$1 \geqslant \sum_{F \in \mathcal{F}:\, \bar{a} \in F} \psi(F) = \sum_{F \in \mathcal{F}:\, \bar{a} \in F} \frac{m(F)}{D} \geqslant \frac{\beta' N}{D} = \beta' \nu^*(\mathcal{F}).$$

Отсюда получаем, что $\nu^*(\mathcal{F}) \leqslant 1/\beta'$.

Мы научились оценивать сверху $\nu^*(\mathcal{F})$ при $(p, d+1)$-условии, но как быть с $\tau(\mathcal{F})$? Для этого нам понадобятся некоторые новые понятия.

§ 6. Слабые ε-сети и завершение доказательства

Пусть, как и прежде, X — некоторое конечное множество, $\mathcal{F}$ — семейство его подмножеств, а ε — некоторое положительное число. Множество $T \subset X$ называется ε-сетью, если $F \cap T \neq \varnothing$ для любого такого $F \in \mathcal{F}$, что $|F| \geqslant \varepsilon|X|$. Таким образом, ε-сеть есть трансверсаль для подсемейства всех множеств из $\mathcal{F}$, имеющих размер не менее чем $\varepsilon|X|$. Оказывается, размер ε-сети не зависит от $|X|$, но зависит от некоторой комбинаторной характеристики семейства $\mathcal{F}$, называемой *размерностью Вапника — Червоненкиса*, однако объяснение этого понятия лежит за пределами данной статьи (см. подробности в [3]). Во многих задачах достаточно вместо трансверсали построить ε-сеть для некоторого ε, а для этого существуют достаточно эффективные алгоритмы.

Нам, однако, понадобится схожее, но более геометрическое понятие. Пусть X — конечное множество точек в $\mathbb{R}^d$. Скажем, что множество точек T (которое теперь уже не обязательно лежит в X) является *слабой ε-сетью* для X, если всякое выпуклое множество C, содержащее не менее $\varepsilon|X|$ точек из X, содержит точку из T.

Теорема 6.1. *Для всякого ε и всякого конечного X на плоскости существует слабая ε-сеть, размер которой равен $\lceil c\varepsilon^{-(d+1)}\rceil$ для некоторой константы c, т. е. зависит только от ε, но не от размера множества X.*

Доказательство. Доказательство есть попросту многократное применение первой леммы о выборе (или теоремы Бороша — Фюреди для случая плоскости). Пусть множество X имеет размер n. Будем пошагово строить нужную слабую ε-сеть. Положим $T_0 = \varnothing$. Пусть на i-м шаге мы выбрали множество T_i, но оно всё ещё не слабая ε-сеть. Тогда найдётся выпуклое множество C_i, которое содержит как минимум εn точек из X, но ни одной точки из T_i. Пусть $X_i = C_i \cap X$, $n_i = |X_i|$. Применим к этому множеству первую лемму о выборе и найдём точку $\bar{a}_i$, лежащую как минимум в

$$c_0 C_{n_i}^{d+1} \geqslant c' n_i^{d+1} \geqslant c' \varepsilon^{d+1} n^{d+1}$$

симплексах, определяемых X_i, где c_0 и c' — некоторые константы (в самом деле, слева в этом неравенстве стоит многочлен степени $d+1$ по n_i, а значит, подбором константы можно добиться, чтобы при достаточно больших n он был больше, чем n^{d+1}). Положим $T_{i+1} = T_i \cup \bar{a}_i$. Всего множество X определяет $C_n^{d+1} \leqslant c'' n^{d+1}$ симплексов, а за один шаг мы покрываем $c' \varepsilon^{d+1} n^{d+1}$ из них. Значит, мы заведомо остановимся за $\lceil c\varepsilon^{-(d+1)}\rceil$ шагов, где $c = c''/c'$. □

Рассмотрим семейство $\mathcal{F}$ выпуклых тел на плоскости и соответствующее множество X, как в § 5. Покажем, что $\tau(\mathcal{F}) \leqslant c\tau^*(\mathcal{F})^{d+1}$, где c — константа из теоремы 6.1. В силу этой теоремы достаточно доказать, что любая слабая $1/\tau^*(\mathcal{F})$-сеть конечного множества точек, которое мы сейчас построим, является трансверсалью для $\mathcal{F}$ — теорема о слабых ε-сетях говорит, что можно выбрать такую сеть размера $c\tau^*(\mathcal{F})^{d+1}$.

Пусть ϕ — дробная трансверсаль с общим весом $\tau^*(\mathcal{F})$, принимающая рациональные значения. Аналогично рассуждениям в предыдущем разделе приведём все эти значения к общему знаменателю D и положим

$$\phi(\bar{x}) = \frac{m(\bar{x})}{D}.$$

Построим новое множество $\widetilde{X}$ (а точнее, «мультимножество»), взяв в него $m(\bar{x})$ копий каждой точки $\bar{x} \in X$. Любую такую копию мы будем считать отдельной точкой. Заметим, что и доказательство теоремы 6.1, и доказательство первой леммы о выборе, на которую оно опирается, работает без изменений для мультимножеств, т. е. для случая, когда мы допускаем повторения точек. Покажем, что всякое множество из $\mathcal{F}$ содержит не менее $|\widetilde{X}|/\tau^*(\mathcal{F})$ точек из $\widetilde{X}$, тогда любая слабая $1/\tau^*(\mathcal{F})$-сеть этого множества будет трансверсалью семейства $\mathcal{F}$. Для $F \in \mathcal{F}$ положим $\widetilde{F} = F \cap \widetilde{X}$. Тогда $|\widetilde{F}| = \sum_{\bar{x} \in F} m(\bar{x})$. Поскольку ϕ — дробная трансверсаль, имеем

$$1 \leqslant \sum_{\bar{x} \in X} \phi(\bar{x}) = \sum_{\bar{x} \in X} \frac{m(\bar{x})}{D} = \frac{|\widetilde{F}|}{D}.$$

Но

$$|\widetilde{X}| = \sum_{\bar{x} \in X} m(\bar{x}) = D\tau^*(\mathcal{F}).$$

Выражая отсюда D и подставляя в предыдущее неравенство, окончательно получаем $|\widetilde{F}| \geqslant |\widetilde{X}|/\tau^*(\mathcal{F})$, что и требовалось доказать.

Мы можем завершить доказательство (p, q)-теоремы. Для семейства выпуклых множеств $\mathcal{F}$ в $\mathbb{R}^d$ его трансверсальное число не превосходит $c\tau^*(\mathcal{F})^{d+1}$, а значит, и $c\nu^*(\mathcal{F})^{d+1}$. Но если на $\mathcal{F}$ наложено (p, q)-условие, то на него наложено $(p, d+1)$-условие, что влечёт ограниченность $\nu^*(\mathcal{F})$ некоторой константой, не зависящей от n. Тем самым (p, q)-теорема полностью доказана.

Благодарности

Я благодарен всем участникам курса, посвящённого (p, q)-теореме, на летней школе «Комбинаторика и алгоритмы 2016», и Андрею Купавскому, рассказавшему мне об этой теме.

Список литературы

[1] *Данцер Л., Грюнбаум Б., Кли В.* Теорема Хелли и её применения. М.: Мир, 1968.

[2] *Протасов В. Ю.* Теорема Хелли и вокруг неё // Квант. 2009. № 3. С. 8–14.

[3] *Райгородский А. М.* Системы общих представителей в комбинаторике и их приложения в геометрии. М.: МЦНМО, 2009.

[4] *Циглер Г.* Теория многогранников. М.: МЦНМО, 2014.

[5] *Alon N., Kleitman D.* Piercing convex sets and the Hadwiger — Debrunner (p, q)-problem // Adv. Math. 1992. Vol. 96, № 1. P. 103–112.

[6] *Bárány I.* A generalization of Carathéodory's theorem // Discrete Math. 1982. Vol. 40. P. 141–152.

[7] *Boros E., Füredi Z.* The number of triangles covering the center of an n-set // Geom. Dedicata. 1984. Vol. 17. P. 69–77.

[8] *Bukh B.* A point in many triangles // Electron. J. Combin. 2006. Vol. 13, Note 10.

[9] *Gromov M.* Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry // Geom. Funct. Anal. 2010. Vol. 20, № 2. P. 416–526.

[10] *Hadwiger H., Debrunner H.* Ausgewählte Einzelprobleme der kombinatorischen Geometrie in der Ebene // Enseignement Math. (2). 1955. Bd. 1, S. 56–89.

[11] *Kalai G.* Intersection patterns of convex sets // Israel J. Math. 1984. Vol. 48, № 2–3. P. 161–174.

[12] *Katchalski M., Liu A.* A problem of geometry in $\mathbb{R}^n$ // Proc. Amer. Math. Soc. 1979. Vol. 75, № 2. P. 284–288.

[13] *Lovász L.* On the ratio of optimal integral and fractional covers // Discrete Math. 1975. Vol. 13, № 4. P. 383–390.

[14] *Matoušek J.* Lectures on discrete geometry. New York: Springer-Verlag, 2002. (Graduate Texts in Math.; Vol. 212).

[15] *Matoušek J., Wagner U.* On Gromov's method of selecting heavily covered points // Discrete Comput. Geom. 2014. Vol. 52, № 1. P. 1–33.

Роман Игоревич Просанов, Фрибурский университет (Швейцария), Московский физико-технический институт

rprosanov@mail.ru

Доказательство теоремы Радона при помощи понижения размерности

Е. С. Колпаков

Известна классическая теорема Радона [1]:

Дано целое $d \geqslant 1$ и $d+2$ точек в d-мерном пространстве $\mathbb{R}^d$. Тогда эти точки можно разбить на два непересекающихся подмножества, чьи выпуклые оболочки имеют непустое пересечение.

Первоначальное алгебраическое доказательство этой теоремы см. в [3]. Оно обычно и приводится. В настоящей статье содержится другое её доказательство — через понижение размерности. Ещё одно доказательство через понижение размерности, несколько более сложное, приводится в статье [2]. В нём используется

Лемма об отделении. *Даны $d+2$ точек в d-мерном пространстве. Тогда существует гиперплоскость, натянутая на d точек из данных, которая разделяет две оставшиеся точки.*

В настоящей статье лемма об отделении не используется. Оба доказательства через понижение размерности дают более сильный результат — количественную теорему Радона.

Разбиение множества точек на два подмножества, выпуклые оболочки которых пересекаются *ровно по одной точке*, называется *радоновским.*

Количественная теорема Радона. *Для любого целого $d \geqslant 1$ и $d+2$ точек общего положения в d-мерном пространстве $\mathbb{R}^d$ радоновское разбиение множества этих точек существует и единственно.*

Теорема Радона связана со следующими теоремами, которые могут быть доказаны через понижение размерности (см. [4, 5]).

Работа выполнена при частичной поддержке Добрушинской студенческой стипендии.

Теорема Конвея — Гордона — Закса. *Для любых шести точек в пространстве, никакие четыре из которых не лежат в одной плоскости, найдутся два зацеплённых треугольника с вершинами в этих точках.*

Теорема ван Кампена — Флореса. *Среди любых семи точек в четырёхмерном пространстве $\mathbb{R}^4$ можно выбрать две непересекающиеся тройки точек так, что треугольники с вершинами в них пересекаются.*

Доказательство количественной теоремы Радона.

Проведём индукцию по d. *База:* $d=1$. Даны три точки на прямой. Тогда есть ровно одно радоновское разбиение: одно из множеств состоит из двух крайних точек, другое из средней точки.

Переход от $d-1$ *к* d. Обозначим через $A_1, \ldots, A_{d+2}$ данные точки в пространстве $\mathbb{R}^d$. Меняя, если нужно, нумерацию точек, можно выбрать гиперплоскость α так, чтобы точка A_{d+2} лежала в одном полупространстве относительно α, а точки $A_1, \ldots, A_{d+1}$ — в другом полупространстве. Тогда точка A_{d+2} не лежит в симплексе $A_1 \ldots A_{d+1}$. Для $1 \leqslant i \leqslant d+1$ обозначим через A_i' пересечение α с $A_{d+2}A_i$.

Выпуклая оболочка множества X будет обозначаться $\langle X \rangle$.

1) *Доказательство существования радоновского разбиения.* По предположению индукции имеется радоновское разбиение множества точек

$$A_1', \ldots, A_{d+1}' = I' \sqcup J'$$

в $(d-1)$-мерном пространстве α. Поскольку точки $A_1, \ldots, A_{d+1}$ находятся в общем положении, пересечение I' с J' состоит из единственной точки, которую обозначим Y' (рис. 1).

Пусть $A_1, \ldots, A_{d+1} = I \sqcup J$ — соответствующее разбиение множества точек $A_1, \ldots, A_{d+1}$. А именно, пусть I — множество таких точек A_i, что A_i' лежит в I'. Аналогично определим множество J. Тогда отрезок $A_{d+2}Y'$ лежит в симплексах $\langle \{A_{d+2}\} \cup J \rangle$ и $\langle \{A_{d+2}\} \cup I \rangle$. Следовательно, луч $A_{d+2}Y'$ пересекает каждый из симплексов $\langle I \rangle$ и $\langle J \rangle$. Обозначим $Y_1 := A_{d+2}Y \cap \langle I \rangle$ и $Y_2 := A_{d+2}Y \cap \langle J \rangle$.

Не ограничивая общность рассуждения, считаем, что Y_1 лежит на отрезке $A_{d+2}Y_2$. Тогда симплекс $\langle I \rangle$ пересекает симплекс $\langle \{A_{d+2}\} \cup J \rangle$ по точке Y_1. Поскольку точки $A_1, \ldots, A_{d+2}$ находятся в общем положении, получаем, что Y_1 — единственная точка пересечения $\langle I \rangle$ и $\langle \{A_{d+2}\} \cup J \rangle$.

Значит, $I \sqcup (\{A_{d+2}\} \cup J)$ или $(\{A_{d+2}\} \cup I) \sqcup J$ является радоновским разбиением множества точек $A_1, \ldots, A_{d+2}$.

2) *Доказательство единственности радоновского разбиения.* Предположим, что радоновских разбиений хотя бы два. Пусть I и J — одно из них, $\widetilde{I}$ и $\widetilde{J}$ — другое. Положим $X = \langle I \rangle \cap \langle J \rangle$.

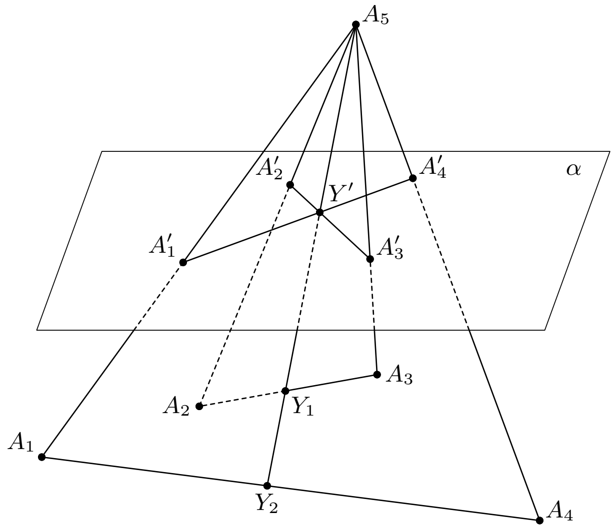


Рис. 1. На рисунке изображён случай размерности $d = 3$. Точка A_5 отделена плоскостью α от точек A_1, A_2, A_3, A_4. Отрезки A_5A_1, A_5A_2, A_5A_3, A_5A_4 пересекают плоскость α в точках A'_1, A'_2, A'_3, A'_4 соответственно. Отрезки $A'_1A'_4$ и $A'_2A'_3$ пересекаются в точке Y'. Луч A_5Y' пересекает отрезки A_2A_3 и A_1A_4 в точках Y_1 и Y_2 соответственно. Точка Y_1 лежит на отрезке A_5Y_2, поэтому отрезок A_2A_3 пересекает треугольник $A_1A_4A_5$ в точке Y_1

Без ограничения общности $A_{d+2} \in J$. Положим $X' = A_{d+2}X \cap \alpha$. Пусть I' — множество таких точек A'_i, что A_i лежит в I, а J' — множество таких точек A'_i, что A_i лежит в J. Тогда X' является единственной точкой пересечения $\langle I' \rangle$ и $\langle J' \rangle$. Значит, $I' \sqcup J'$ есть радоновское разбиение множества точек $A'_1, \ldots, A'_{d+1}$. Аналогично из множеств $\widetilde{I}$ и $\widetilde{J}$ строится другое радоновское разбиение множества точек $A'_1, \ldots, A'_{d+1}$, отличное от $I' \sqcup J'$.

Таким образом, у множества точек $A'_1, \ldots, A'_{d+1}$ из $(d-1)$-мерного пространства α есть два радоновских разбиения, что противоречит предположению индукции. Значит, число радоновских разбиений множества точек $A_1, \ldots, A_{d+2}$ не больше 1. Переход индукции доказан. $\square$

Благодарности

Автор выражает благодарность А. Б. Скопенкову за полезные замечания и А. Акопяну за сообщение о статье [2].

Список литературы

[1] *Данцер Л., Грюнбаум Б., Кли В.* Теорема Хелли и её применения. М.: Мир, 1968. С. 12, 21.

[2] *Peterson B. B.* The Geometry of Radon's Theorem // American Mathematical Monthly. 1972. Vol. 79, № 9. P. 949–963.

[3] *Radon J.* Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten // Math. Ann. 1921. Bd. 83. S. 113–115.

[4] *Skopenkov A.* On van Kampen — Flores, Conway — Gordon — Sachs and Radon theorems // `https://arxiv.org/abs/1704.00300v1`

[5] *Skopenkov A.* Realizability of hypergraphs and Ramsey link theory // `https://arxiv.org/abs/1402.0658`

Егор Сергееевич Колпаков, МГУ и НМУ

`kolpak317bel@gmail.com`

Оклеивание тетраэдра квадратами

А. А. Балакин

§ 1. Введение

В данной работе доказана следующая теорема.

Теорема 1. *Поверхность правильного тетраэдра нельзя оклеить конечным числом квадратов без просветов и наложений.*

Иногда мы будем говорить не об оклеивании, а о разрезании поверхности тетраэдра на квадраты. Здесь и далее эти понятия означают одно и то же. Формальное определение приведено в следующем параграфе.

Идея доказательства теоремы заключается в том, чтобы сопоставить разрезанию поверхности тетраэдра периодическое разрезание плоскости. Затем разрезанию плоскости сопоставляется разрезание некоторого L-образного шестиугольника на части, из которых складывается несколько квадратов. Далее вводится понятие так называемой x-площади, которая обобщает обычную площадь, однако может быть отрицательной для некоторых многоугольников, но всегда неотрицательна для квадратов.

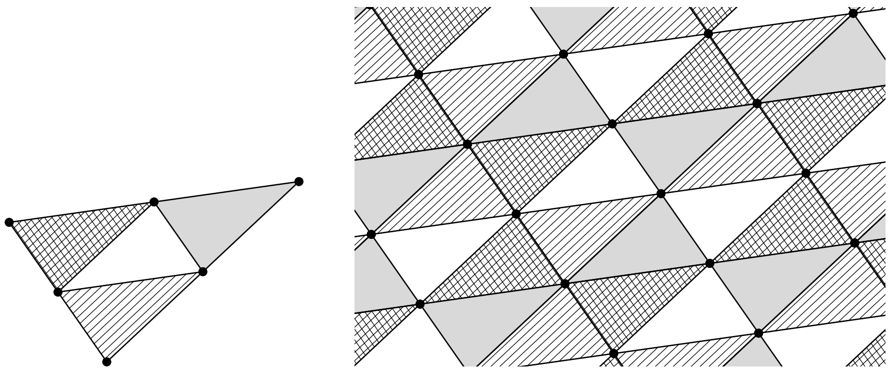
В принципе, теорема 1 следует из приведённой в данной работе леммы 1 и теоремы 10 из статьи Keньёна [1] с довольно сложным доказательством. Здесь мы фактически приведём элементарное доказательство результата Кеньёна для частного случая, которое, скорее всего, может быть обобщено.

Некоторые результаты и методы этой работы относятся не только к правильному тетраэдру, но и к любым равногранным тетраэдрам (то есть тетраэдрам, все четыре грани которых равны). Например, в следующем параграфе мы рассматриваем понятия, определённые для всех равногранных тетраэдров. Другой подход к разрезаниям поверхностей на квадраты, связанный с электрическими цепями, можно найти в работе [4].

Автор частично поддержан грантом Президента РФ МК-6137.2016.1.

§ 2. От тетраэдра к плоскости

Мотивировка. Возьмём равногранный тетраэдр и покрасим его грани в четыре цвета[1]. Разрежем его поверхность вдоль трёх рёбер, исходящих из одной вершины, и развернём на плоскость. Получится «большой» треугольник, состоящий из четырёх «маленьких» треугольников разных цветов, на которые он разбивается своими средними линиями (рис. 1а). Эти «маленькие» треугольники были гранями исходного тетраэдра. Далее замостим «большим» треугольником плоскость, отражая его относительно середин его сторон, затем относительно середин сторон новых треугольников и т. д. Получится разбиение плоскости на «маленькие» треугольники четырёх цветов (рис. 1б). Каждый «маленький» треугольник граничит по сторонам с треугольниками трёх оставшихся цветов.



а) б)

Рис. 1. Разноцветное замощение

Определение. Пусть плоскость разбита на треугольники четырёх цветов, как показано на рис. 1б. Вершины этих треугольников образуют решётку на плоскости. Назовём две точки на плоскости *эквивалентными*, если одну можно перевести в другую композицией нескольких центральных симметрий относительно узлов этой решётки. Заметим, что если точки эквивалентны, то они лежат в треугольниках одного цвета, и каждый треугольник может быть переведён композицией центральных симметрий относительно узлов решётки в любой другой того же цвета. Значит, если мы возьмём на плоскости треугольник, состоящий из треугольников разбиения всех четырёх цветов (рис. 1а), в нём будут присутствовать точки, являющиеся представителями всех классов эквивалентности. Назовём такой треугольник

[1] Цвета на рисунках заменены различной штриховкой.

фундаментальным. С другой стороны, фундаментальный треугольник является развёрткой равногранного тетраэдра на плоскость. Назовём отображение из фундаментального треугольника на поверхность тетраэдра, переводящее точку треугольника в её прообраз при развёртке, *склейкой.* Композицию отображения, переводящего точку плоскости в эквивалентную ей точку фундаментального треугольника, с последующей склейкой, назовём *накрытием.*

Определение. Прообразы всех вершин тетраэдра при накрытии образуют решётку на плоскости. Будем говорить, что эта решётка *порождена* данным тетраэдром.

Если среди внутренних точек квадрата на плоскости нет узлов решётки, порождённой данным равногранным тетраэдром, то образ этого квадрата при накрытии назовём *квадратом на поверхности тетраэдра.*

Оклеить поверхность тетраэдра квадратами без просветов и наложений (*разрезать* на квадраты) значит представить её в виде объединения конечного числа квадратов на поверхности тетраэдра, не имеющих общих внутренних точек.

Лемма 1. *Поверхность равногранного тетраэдра можно разрезать на квадраты тогда и только тогда, когда можно разрезать плоскость на квадраты так, чтобы разрезание переходило в себя при центральных симметриях относительно узлов порождённой этим тетраэдром решётки и внутренние точки квадратов не были узлами этой решётки.*

Доказательство. По определениям накрытия и квадрата на поверхности тетраэдра, прообразом квадрата на поверхности тетраэдра при накрытии будет множество квадратов на плоскости, переходящее в себя при центральной симметрии относительно любого узла порождённой решётки. И если поверхность тетраэдра представлена в виде объединения квадратов, то и плоскость будет представлена в виде объединения их прообразов.

При этом любое множество квадратов на плоскости, внутренние точки которых не являются узлами решётки, при накрытии переходит в множество квадратов на поверхности тетраэдра. Если же плоскость разрезана на квадраты и разрезание переходит в себя при симметрии относительно каждого узла решётки, причём внутренние точки квадратов не являются узлами решётки, то при накрытии все квадраты, содержащие прообраз фиксированной точки, перейдут в один квадрат на поверхности тетраэдра. Увидеть это можно следующим образом: если мы отразим квадрат на плоскости относительно узла решётки, то его образ при накрытии не изменится, а все квадраты, содержащие прообраз фиксированной точки, можно получить композицией таких преобразований из любого из них. Это означает, что поверхность тетраэдра будет оклеена квадратами без просветов и наложений. □

Пример. Сделаем небольшое отступление и приведём пример тетраэдра, поверхность которого можно разрезать на квадраты. А именно, возьмём равногранный тетраэдр, грани которого равны треугольнику с вершинами в точках A, B и C, имеющих на плоскости соответственно декартовы координаты $(0,0)$, $(2,2)$ и $(3,0)$, и покажем, как разрезать плоскость симметрично относительно узлов порождённой этим тетраэдром решётки.

Разобьём плоскость на прямоугольники следующим образом. Возьмём решётку, порождённую этим тетраэдром, и проведём через каждую её вершину прямую, параллельную стороне AB. Назовём прямую, проходящую через A, чётной, а соседние прямые — нечётными. Далее назовём соседей нечётных прямых чётными, и т. д. Опустим из всех вершин нечётных прямых перпендикуляры на соседние прямые (рис. 2). Получилось разбиение плоскости на прямоугольники, симметричное относительно всех узлов решётки. Нетрудно проверить, что стороны этих прямоугольников равны $2\sqrt{2}$ и $\frac{3}{2}\sqrt{2}$, т. е. они относятся как $4:3$, а значит, их можно разрезать на $4 \cdot 3 = 12$ квадратов. Тогда по лемме 1 и заданный равногранный тетраэдр можно оклеить квадратами.

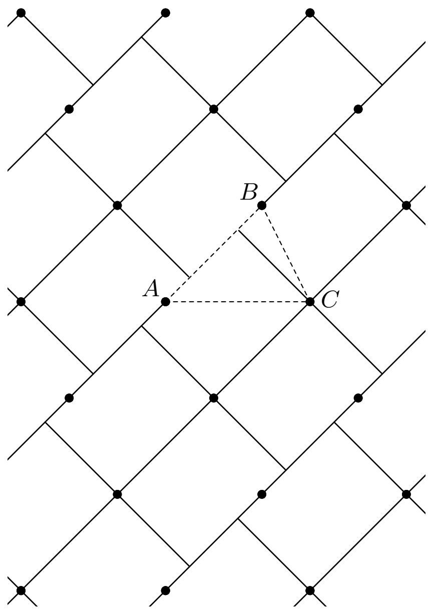


Рис. 2. Разбиение на прямоугольники

Теперь вернёмся к обсуждению понятий, необходимых для доказательства теоремы 1.

Определение. Возьмём решётку, порождённую некоторым равногранным тетраэдром. Прообразом фиксированной вершины при накрытии будет подрешётка вдвое большего размера. Будем говорить, что такая подрешётка *порождена* вершиной.

Замечание. Нетрудно проверить, что композицией двух центральных симметрий относительно узлов решётки, порождённой тетраэдром, является параллельный перенос на удвоенный вектор, соединяющий два центра симметрии, т. е. на удвоенный вектор решётки. Значит, прообраз разрезания тетраэдра переходит в себя при параллельных переносах на векторы подрешётки, порождённой вершиной тетраэдра.

Обозначение. Правильный тетраэдр порождает правильную треугольную решётку на плоскости. Зафиксируем какую-нибудь его вершину. Она порождает правильную треугольную подрешётку, стороны которой в два раза больше сторон исходной. Наконец, у этой подрешётки есть прямоугольная подрешётка, прямоугольник которой имеет отношение сторон $\sqrt{3}$, меньшая из которых равна стороне треугольника правильной подрешётки. Обозначим такую прямоугольную решётку буквой Г.

Далее мы в основном будем рассматривать решётку Г. Заметим также, что прообраз разрезания тетраэдра переходит в себя при параллельном переносе на любой вектор этой решётки.

§ 3. От плоскости к уголкам

Мотивировка. Пусть на плоскости с фиксированной решёткой дано направление разрезания: стороны квадратов, на которые мы хотим разрезать плоскость, либо параллельны, либо перпендикулярны этому направлению. Если мы докажем невозможность разрезания в каждом данном направлении, мы докажем и невозможность разрезания вообще.

Определение. Нарисуем решётку Г так, чтобы вертикальные стороны её прямоугольников относились к горизонтальным как $\sqrt{3}$. Пусть дано негоризонтальное направление u разрезания. Между каждыми соседними по горизонтали узлами нарисуем «ступеньку» (два перпендикулярных

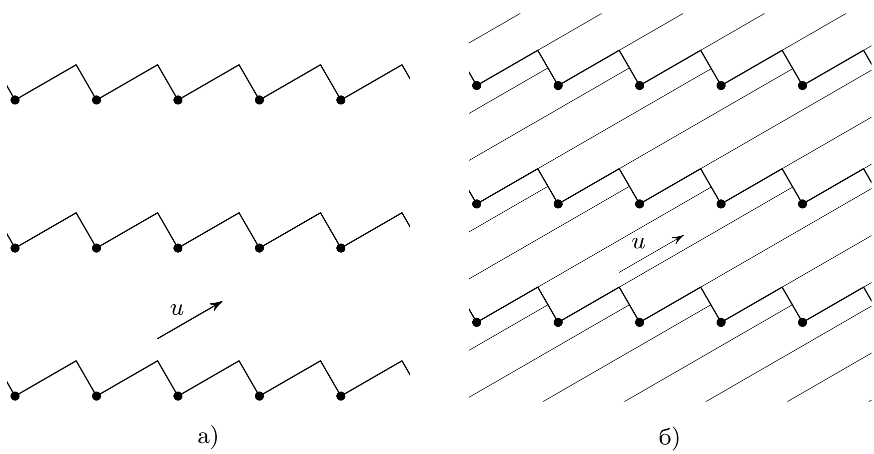


Рис. 3. Разбиение на уголки

отрезка с общим концом), одна из сторон которой параллельна u, а другая перпендикулярна (рис. 3а). Далее проведём через каждую вершину отрезки, параллельные u, до пересечения со «ступенькой» выше. Получим разбиение плоскости на прямоугольники или L-образные шестиугольники, которые будем называть *уголками* (рис. 3б).

Замечание. В случае когда уголок вырождается в прямоугольник, будем тоже называть его «уголком», считая, что одна из его сторон равна нулю. В случае же, когда направление u горизонтально, заменим его на вертикальное.

Определение. Будем называть уголок *разрезаемым*, если можно взять несколько квадратов (не обязательно равных), как-то разрезать каждый из них на конечное число прямоугольников, а затем из всех получившихся прямоугольников составить данный уголок.

Лемма 2. *Если плоскость можно разрезать на квадраты вдоль заданного направления так, чтобы разбиение переходило в себя при всех параллельных переносах на векторы решётки* Γ, *то уголок, построенный по этому направлению, разрезаем.*

Доказательство. Заметим, что уголки, на которые мы разбили плоскость, совмещаются параллельным переносом на векторы решётки Γ. Более того, разрезания, проведённые внутри каждого уголка, также совмещаются. Это означает, что если квадрат на плоскости разбивается линиями уголков на несколько частей, то равные им части лежат в каждом из уголков. Следовательно, мы можем разрезать уголок на несколько частей, из которых можно сложить квадраты. И так как при этом все разрезания будут параллельны сторонам квадратов, то и обратно — мы сможем разрезать квадраты на прямоугольники и сложить из них уголок. А это и есть определение разрезаемости. $\square$

§ 4. Определение x-площади

Нам потребуется несколько обозначений для определения x-площади. О том, как это определение естественно возникает в одном из доказательств известной теоремы Дена, можно прочитать в работе [3].

Определение 1. Зафиксируем некоторый уголок. Пусть он получается вырезанием прямоугольника $c \times d$ из прямоугольника $a \times b$. Пусть он разрезан на прямоугольники и $r_1 < r_2 < r_3 < \ldots < r_k$ — все длины сторон этих прямоугольников. Обозначим

$$P = \{a, a\sqrt{3}, b, b\sqrt{3}, c, c\sqrt{3}, d, d\sqrt{3}, r_1, r_1\sqrt{3}, \ldots, r_k, r_k\sqrt{3}\}.$$

Найдём такие числа $t_1, t_2, \ldots, t_n \in P$, чтобы любое число $p \in P$ единственным образом представлялось в виде

$$p = y_0 a + y_1 a\sqrt{3} + y_2 t_1 + \ldots + y_{n+1} t_n$$

с рациональными $y_0, y_1, \ldots, y_{n+1}$. Это можно сделать следующим образом [2, решение задачи 25.8.18]: выпишем в строку все числа из P, начиная с a и $a\sqrt{3}$, и подчеркнём те числа, которые не представляются в виде линейной комбинации предыдущих с рациональными коэффициентами. Это и будут искомые $t_1, t_2, \ldots, t_n$.

Иными словами, дополним a и $a\sqrt{3}$ до *базиса* в пространстве линейных комбинаций чисел из множества P с рациональными коэффициентами. Назовём числа, представимые в виде такой комбинации, *хорошими*. Заметим, что если число z хорошее, то и число $z\sqrt{3}$ тоже хорошее.

Пусть даны вещественное число x и прямоугольник с хорошими сторонами

$$z_0 a + z_1 a\sqrt{3} + z_2 t_1 + \ldots + z_{k+1} t_n$$

и

$$w_0 a + w_1 a\sqrt{3} + w_2 t_1 + \ldots + w_{k+1} t_n,$$

где z_i и w_i рациональны. Его *x-площадью* (или *площадью Гамеля*) назовём число $(z_0 + z_1 x)(w_0 + w_1 x)$.

Лемма 3. *Если прямоугольник разрезан на прямоугольники с хорошими сторонами, то его x-площадь равна сумме x-площадей этих прямоугольников.*

Доказательство [2, решение задачи 25.8.16]. Нетрудно убедиться, что сумма x-площадей двух прямоугольников с хорошими сторонами равна x-площади их объединения. Действительно, пусть имеется прямоугольник с x-площадью S, который состоит из двух прямоугольников с общей стороной

$$z = z_0 a + z_1 a\sqrt{3} + z_2 t_1 + \ldots + z_{k+1} t_n$$

и другими сторонами

$$w = w_0 a + w_1 a\sqrt{3} + w_2 t_1 + \ldots + w_{k+1} t_n,$$
$$v = v_0 a + v_1 a\sqrt{3} + v_2 t_1 + \ldots + v_{k+1} t_n.$$

Тогда сумма их x-площадей равна

$$(z_0 + z_1 x)(w_0 + w_1 x) + (z_0 + z_1 x)(v_0 + v_1 x) =$$
$$= (z_0 + z_1 x)((w_0 + v_0) + (w_1 + v_1)x) = S.$$

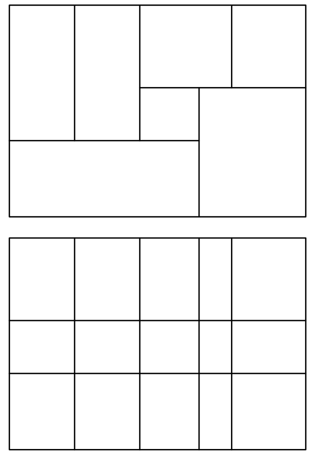


Рис. 4. Продление линий

Пусть теперь количество прямоугольников в разрезании больше двух. Продолжим каждый разрез, как показано на рис. 4. Тогда каждый прямоугольник нового разрезания также будет иметь хорошие стороны. Рассмотрим горизонтальные слои из последовательно приложенных друг к другу по общей стороне прямоугольников. Используя уже доказанное свойство аддитивности x-площади для двух прямоугольников с общей стороной, легко доказать по индукции, что x-площадь любого такого слоя равна сумме x-площадей прямоугольников, составляющих этот слой. Теперь приложим уже эти слои друг к другу и применим только что доказанное утверждение об аддитивности x-площади ряда прямоугольников. Получим, что x-площадь разрезаемого прямоугольника равна сумме x-площадей горизонтальных слоёв. Эта сумма равна сумме x-площадей всех прямоугольников разрезания. □

Определение. Пусть уголок разрезан на конечное число прямоугольников с хорошими сторонами. Назовём его *x-площадью* сумму x-площадей прямоугольников, на которые он разрезан.

Следствие (из леммы 3). *Если уголок, полученный вырезанием прямоугольника $c \times d$ из прямоугольника $a \times b$, разрезан на прямоугольники с хорошими сторонами, то сумма x-площадей этих прямоугольников равна сумме x-площадей прямоугольников $a \times (b - d)$ и $(a - c) \times d$.*

Доказательство. Нарисуем разрезание уголка на прямоугольники и продлим сторону уголка длины c до пересечения с границей. Получилось новое разрезание уголка на прямоугольники с хорошими сторонами. В частности, уголок состоит из двух прямоугольников $a \times (b-d)$ и $(a-c) \times d$ с хорошими сторонами, каждый из которых, в свою очередь, разбит на несколько прямоугольников с хорошими сторонами. Значит, по лемме 3, x-площади этих двух прямоугольников равны сумме x-площадей прямоугольников, из которых они состоят. Тогда сумма x-площадей прямоугольников $a \times (b - d)$ и $(a - c) \times d$ равна сумме x-площадей всех прямоугольников из нового разрезания. Аналогично получаем, что такова же сумма x-площадей прямоугольников из разрезания, данного в условии леммы. □

Лемма 4. *Для любого действительного x верно, что x-площадь квадрата с хорошей стороной неотрицательна.*

Доказательство. Если сторона квадрата равна

$$z_0 a + z_1\sqrt{3} + z_2 t_1 + \ldots + z_{k+1} t_n,$$

то его x-площадь равна $(z_0 + z_1 x)^2 \geqslant 0$. □

§ 5. x-Площадь уголка

Пусть плоскость разбита на квадраты вдоль заданного направления u так, что разрезание переходит в себя при сдвиге на любой вектор решётки Γ. Если u горизонтально, заменим его на вертикальное направление. Введём прямоугольную систему координат так, чтобы началом координат был какой-нибудь узел решётки Γ, ось ординат была параллельна вектору u, а ось абсцисс — перпендикулярному ему вектору v (рис. 5). Обозначим четыре стороны уголка через a, b, c и d, как показано на рис. 5. Нетрудно проверить, что тогда узлы решётки Γ, в которые упирается уголок, имеют координаты (a, d) и (c, b). Обозначим координаты вертикального вектора стороны прямоугольника решётки Γ через (e, f).

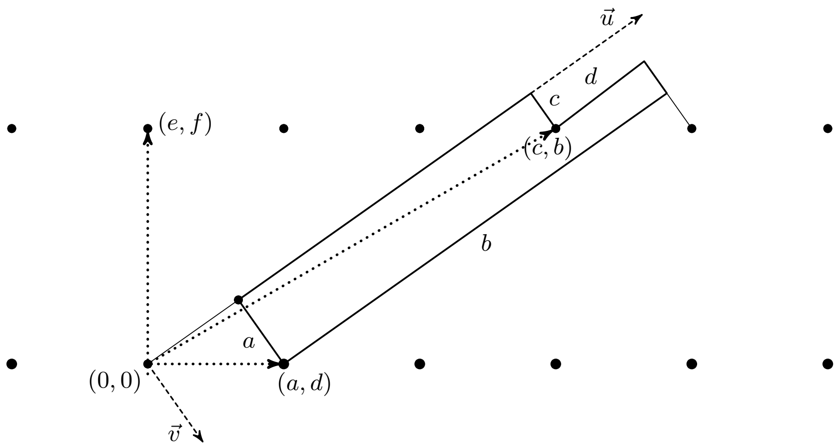


Рис. 5. Обозначения сторон и координат

Заметим, что $a \neq 0$. Иначе направление u горизонтально, а его мы условились заменить на вертикальное.

Лемма 5. *Для любого направления u существует такое число x, что x-площадь уголка (см. определение в § 3) отрицательна.*

Доказательство. Зафиксируем направление разрезания и само разрезание. По этим данным построим базис (см. § 4). Посчитаем x-площадь уголка, изображённого на рис. 4, в зависимости от x. Обозначим через d_0

и d_1 соответственно коэффициенты при a и $a\sqrt{3}$ в разложении числа d по базису. Аналогично введём a_0 и a_1, b_0 и b_1 и c_0 и c_1. В частности, $a_0=1$ и $a_1=0$.

Вектор (e,f) получается из вектора (a,d) поворотом против часовой стрелки на 90° и растяжением в $\sqrt{3}$ раз. Значит, $(e,f)=\sqrt{3}(-d,a)$, т. е. $e=-\sqrt{3}d$, $f=a\sqrt{3}$. Тогда e и f тоже хорошие. Аналогично предыдущим обозначениям введём e_0 и e_1, f_0 и f_1.

По следствию из леммы 3, x-площадь уголка равна

$$(a_0+a_1x)((b_0+b_1x)-(d_0+d_1x))+((a_0+a_1x)-(c_0+c_1x))(d_0+d_1x)=\\ =(a_0+a_1x)(b_0+b_1x)-(c_0+c_1x)(d_0+d_1x).$$

По построению, так как вершины решётки (a,d) и (c,d) лежат на соседних горизонтальных уровнях решётки (рис. 4), выполняется равенство между координатами соответствующих векторов $(c,b)=(e,f)+m(a,d)$ для некоторого целого m. Это значит, что x-площадь уголка равна

$$(a_0+a_1x)((f_0+f_1x)+m(d_0+d_1x))-(d_0+d_1x)((e_0+e_1x)+m(a_0+a_1x))=\\ =(a_0+a_1x)(f_0+f_1x)-(d_0+d_1x)(e_0+e_1x).$$

А так как $(a_0+a_1x)=1$, $(e_0+e_1x)=-3d_1-d_0x$ и $(f_0+f_1x)=x$, имеем

$$(a_0+a_1x)(f_0+f_1x)-(d_0+d_1x)(e_0+e_1x)=\\ =x-(-3d_1-d_0x)(d_0+d_1x)=3d_0d_1+(d_0^2+3d_1^2+1)x+d_0d_1x^2.$$

Возможны следующие три случая.

I. Если $d_0d_1<0$, то 0-площадь отрицательна.

II. Если $d_0d_1=0$, то x-площадь равна $(d_0^2+3d_1^2+1)x$, тогда (-1)-площадь отрицательна.

III. Если $d_0d_1>0$, то $(-d_0/d_1)$-площадь равна $-\dfrac{d_0}{d_1}$, т. е. отрицательна. □

§ 6. Доказательство основной теоремы

Теорема 1. *Поверхность правильного тетраэдра нельзя оклеить конечным числом квадратов без просветов и наложений.*

Доказательство. Если поверхность правильного тетраэдра можно разрезать на квадраты, то по леммам 1 и 2 уголок, построенный по решётке Γ и направлению u, разрезаем. По лемме 3, x-площадь этого уголка равна сумме x-площадей нескольких квадратов (разрезав уголок на прямоугольники, можно собрать несколько квадратов). Но, по лемме 4, x-площадь квадрата неотрицательна для любого x. А по лемме 5 найдётся такое x, что x-площадь уголка меньше нуля. Полученное противоречие доказывает, что поверхность правильного тетраэдра нельзя разрезать на квадраты. □

На этом фоне возникает естественный вопрос: поверхности каких равногранных тетраэдров можно оклеить конечным числом квадратов без просветов и наложений? Этот вопрос пока открыт, однако метод, приведённый в первом разделе данной работы, позволяет сопоставить каждой оклейке равногранного тетраэдра периодическое разрезание плоскости. Последнее, в свою очередь, близко к разрезанию тора на квадраты, что подробно разобрано в [1].

Гипотеза. *Пусть все грани тетраэдра равны треугольнику с вершинами* $(0,0)$, $(1,0)$ *и* (x,y). *Поверхность этого тетраэдра можно оклеить квадратами без просветов и наложений тогда и только тогда, когда точка* (x,y) *лежит либо на окружности с рациональным центром и рациональным радиусом, не пересекающей ось* Ox, *либо на горизонтальной прямой на рациональном расстоянии от оси* Ox.

Задача. Дан конверт в форме прямоугольника $a \times b$. При каких вещественных a и b его можно оклеить квадратными марками без просветов и наложений с обеих сторон? Квадраты разрешается перегибать через край прямоугольника, их размеры могут различаться.

Благодарности

Автор благодарен М. Б. Скопенкову, без которого данная работа не началась бы и не состоялась в том виде, в каком она имеет быть сейчас, а также Ф. А. Шарову за ценные замечания.

Список литературы

[1] *Kenyon R.* Tiling with squares and square-tileable surfaces https://pdfs.semanticscholar.org/eec7/31c9882dcb80a2fb1dbe2073b85e3ed13753.pdf

[2] *Скопенков М. Б., Малиновская О. А., Дориченко С. А., Шаров Ф. А.* Собери квадрат // Элементы математики в задачах: через олимпиады и кружки — к профессии / Ред. А. А. Заславский, А. Б. Скопенков, М. Б. Скопенков. М.: МЦНМО, 2018. С. 542–554.

[3] *Шаров Ф. А.* X-площадь // Представлено в «Квант».

[4] *Chien E., Luo F.* Rectangle tilings of closed surfaces from discrete harmonic 1-chains https://pdfs.semanticscholar.org/47a9/0f45fb72847f8caac69c924fd30c7aa74901.pdf

Андрей Алексеевич Балакин, НИУ Высшая школа экономики
a99bland@yandex.ru

Наш семинар: математические сюжеты

Коды и олимпиады

С. Б. Гашков

§ 1. Введение

Прикладная математика — это когда вы ищете решение задачи, а чистая — это когда вы ищете задачу для решения.

Финн, «Анна и Чёрный рыцарь»

Пусть читатель попробует самостоятельно решить хотя бы некоторые из приведённых далее задач. Некоторые из них в том или ином виде предлагались на различных олимпиадах [1)] (в скобках указано, где именно, и эти задачи можно найти в [2, 8–11]). Что между ними общего?

Задача 1 (ММО 1954, второй тур, 10.5). Рассматриваются всевозможные десятизначные десятичные числа, записываемые при помощи цифр 1 и 2. Разбейте их на два класса так, чтобы сумма любых двух чисел из одного класса содержала в своей записи не менее двух троек.

Задача 2 (ММО 1967, второй тур, 8.3). Для зашифровки телеграфных сообщений требуется разбить всевозможные десятичные «слова» — наборы из десяти точек и тире — на две группы так, чтобы любые два слова одной группы отличались не менее чем в трёх разрядах. Укажите способ такого разбиения или докажите, что его не существует.

[1)] ММО означает Московскую, а ВМО — Всесоюзную математическую олимпиаду.

Задача 3 (олимпиада ФРГ, 1970/71 г., второй тур). В племени Мумбо-Юмбо все имена различны, состоят из букв А и Б, имеют длину n и отличаются друг от друга не менее чем тремя буквами. Докажите, что в племени не более $2^n/(n+1)$ человек. Может ли эта граница достигаться? В соседнем племени АББА имена отличаются не менее чем двумя буквами. Докажите, что в племени не более 2^{n-1} человек. Может ли эта граница достигаться?

Задача 4. Множество из всех k-значных n-разрядных упорядоченных наборов $(x_1, \ldots, x_n)$, $x_i \in \{0, \ldots, k-1\}$, назовём n-мерной k-ичной шахматной доской. Будем говорить, что ладья, находящаяся в поле с координатами $(x_1, \ldots, x_n)$, бьёт любое поле $(x_1, \ldots x_{i-1}, y, x_{i+1}, \ldots, x_n)$, $y = 0, \ldots, k-1$, $i = 1, \ldots, n$. Обозначим через $m(n, k)$ минимальное число ладей, бьющих все поля доски. Докажите, что

а) $m(2, k) = k$;

б) (ВМО 1971, 9.6, 10.6) $m(3, k) = \lfloor k^2/2 \rfloor$;

в) $m(n, k) \leqslant k^n/((k-1)n+1)$;

г) если k равно степени простого числа, то неравенство пункта (в) обращается в равенство при $n = 1 + k + \ldots + k^i$, $i = 1, 2, \ldots$

Задача 5 (волшебный веер Эдуарда Люка[2]). Зрителю предлагается задумать число от 1 до 31. Фокусник предлагает ему сказать, в каких полосках веера он видит задуманное число, а в каких — нет.

На первой полоске написаны все нечётные числа.

На второй — числа 2, 3, 6, 7, 10, 11, 14, 15, 18, 19, 22, 23, 26, 27, 30, 31.

На третьей — числа 4, 5, 6, 7, 12, 13, 14, 15, 20, 21, 22, 23, 28, 29, 30, 31.

На четвёртой — числа 8, 9, 10, 11, 12, 13, 14, 15, 24, 25, 26, 27, 28, 29, 30, 31, и на последней — все числа от 16 до 31.

Как фокусник угадывает задуманное число?

Задача 6 (теорема Заранкевича[3]). Пусть $k = k_{a,b}(n, m)$ — наибольшее число единиц в таблице с n строками и m столбцами, заполненной нулями и единицами и не содержащей a строк и b столбцов, на пересечении которых стояли бы сплошь единицы. Тогда k — наибольшее такое число, что[4]

$$n\binom{k/n}{b} \leqslant (a-1)\binom{m}{b}.$$

[2] Эта задача из книги по занимательной математике, опубликованной в XIX в. французским математиком Э. Люка (1842–1891), автором многих красивых задач и теорем.

[3] К. Заранкевич (1902–1959) — польский математик, опубликовавший задачу на эту тему в 50-е годы. Сама теорема впервые появилась в статьях венгерского математика П. Эрдёша (1913–1996). Несколько олимпиадных задач, фактически являющихся частными случаями этой теоремы, можно найти в [4].

[4] Далее через $\binom{x}{y}$ обозначается количество сочетаний из x по y (другое распространённое обозначение — C_x^y).

Все эти задачи относятся к интересной области современной прикладной математики — теории кодирования. Как они решаются, станет понятно при чтении соответствующих разделов этой статьи.

§ 2. Что такое коды?

> Изначально его убеждение состояло в том, что перед ним — именно шифр, ибо то, что алхимики и некроманты в старину часто пользовались тайнописью, давно стало общеизвестным фактом: видимо, эти искатели истины пытались либо уберечь секреты от соперников, либо укрыть их от ревнивых глаз церковных властей.
>
> *Джон Гласби, «Чёрное зеркало»*

Двоичным кодом C с *блоковой длиной* n можно назвать любое множество двоичных наборов длины n (называемых также *кодовыми словами*)[5]. *Расстоянием* между словами a, b называется число позиций, в которых эти слова различаются. Например, $d(a, b) = 3$ в случае $a = (11001)$, $b = (00011)$. Очевидно, что для любых $a, b, c \in B^n$ выполняется *неравенство треугольника* $d(a, b) + d(b, c) \geqslant d(a, c)$, а также ещё два условия: $d(a, b) = d(b, a)$ и $d(a, b) = 0$ только в случае $a = b$. *Минимальным расстоянием* кода C называется число $d(C) = \min_{a,b\in C} d(a, b)$, равное минимальному расстоянию между его различными словами. В задаче 3 на самом деле речь идёт о коде с минимальным расстоянием, не меньшим 3, и о коде с минимальным расстоянием, не меньшим 2. В задаче 2 также фигурируют два кода с расстоянием, не меньшим 3. Задача 1 легко решается, если увидеть её связь с кодом с минимальным расстоянием 2.

Интерес для кодирования с исправлением ошибок представляют только коды с расстоянием, большим единицы. Например, код C, состоящий из всех наборов чётного веса. *Весом* набора называется число ненулевых позиций в нём. Этот код (он даёт решение второй половины задачи 3) можно использовать для обнаружения одной ошибки. Предположим, что надо передать по ненадёжному каналу связи двоичное слово $x = (x_1, \ldots, x_{n-1})$. Известно, что в полученном слове может быть одна ошибка (замена двоичного символа 0 или 1 на противоположный). Если передавать в точности слово x, то ошибку заметить невозможно. Но если к слову x добавить

---

[5] Код C можно рассматривать как подмножество множества вершин двоичного n-мерного куба $B^n = \{0, 1\}^n$.

один *проверочный символ* $c_n = x_1 \oplus \ldots \oplus x_n$ (где $\oplus$ — операция *сложения по модулю два*[6]) и передать закодированное слово

$$c = (c_1, \ldots, c_n) = (x_1, \ldots, x_{n-1}, c_n),$$

то ошибку легко обнаружить, так как если её не было, то $c_1 \oplus \ldots \oplus c_n = 0$, а если ошибка была, то эта сумма равна единице. Указанный код называется *проверкой на чётность*[7].

Бинарным линейным кодом длины n называется любое такое множество двоичных векторов длины n, что покомпонентная сумма по модулю два любых двух его векторов всегда принадлежит коду. Число единиц в сумме двух векторов по модулю два очевидно равно расстоянию между этими векторами. Бинарный линейный код можно рассматривать как линейное пространство над полем $\{0, 1\}$ из двух элементов. Размерность этого пространства называется *размерностью кода*. Простейшая часть их теории является переформулировкой теорем линейной алгебры. В задачах 9, 10 как раз речь идёт о линейных кодах (и их размерностях).

Можно рассматривать не только двоичные (бинарные), но и q-ичные коды при $q > 2$. *Расстоянием* между q-ичными векторами x, y называется число $\rho(x, y)$ координат, в которых эти векторы не совпадают. Так определённое расстояние совпадает в случае $q = 2$ с введённым выше, и для него выполнено неравенство треугольника $\rho(x, y) \leqslant \rho(x, z) + \rho(y, z)$. Кодовым расстоянием называется, как и в бинарном случае, минимальное расстояние между разными кодовыми векторами.

Если код имеет кодовое расстояние $d = 2t + 1$, то он может исправлять вплоть до t ошибок. Действительно, если при передаче кодового слова x в нём произошло t ошибок, то мы получим искажённое слово x', для которого $\rho(x', x) = t$. По искажённому слову можно однозначно восстановить кодовое слово, так как если из двух разных кодовых слов x, y получено одно и то же искажённое не более чем t ошибками слово $x' = y' = z$, то согласно неравенству треугольника

$$2t + 1 = d \leqslant \rho(x, y) \leqslant \rho(x, z) + \rho(y, z) \leqslant 2t,$$

а это невозможно. Восстановление кодового слова по искажённому слову называется *декодированием*[8].

[6] По определению $x_1 \oplus \ldots \oplus x_n$ равно остатку от деления обычной суммы $x_1 + \ldots + x_n$ на два. Сложение по модулю два отличается от обычного только равенством $1 \oplus 1 = 0$.

[7] Его часто применяют для проверки целостности информации. Он обнаруживает наличие ошибки, но не находит и не исправляет её.

[8] Декодирование — непростой процесс, и он далее не рассматривается.

Если кодовое расстояние равно d, то код обнаруживает $d-1$ ошибку (если принятое слово является кодовым, то ошибок не было, потому что для получения другого кодового слова нужно сделать не менее d ошибок, а по предположению их не больше $d-1$; если же полученное слово не кодовое, то ошибки были).

Интересно для любых n, d найти q-ичный код максимальной мощности[9] с блоковой длиной n и расстоянием d. Обозначим его мощность $m_q(n,d)$. Для произвольного нечётного $d=2t+1$ легко получить следующую верхнюю оценку (из которой следует решение первой половины задачи 3 и п. (в) задачи 4):

Теорема 1 (граница Хэмминга — Рао[10], граница сферической упаковки). *Для любых n, q, d*

$$m_q(n,d) \leqslant \frac{q^n}{\sum\limits_{i=0}^{t} \binom{n}{i}(q-1)^i}.$$

В частности, при $q=2$

$$m_2(n,d) \leqslant \frac{2^n}{\sum\limits_{i=0}^{t} \binom{n}{i}}.$$

Доказательство. Достаточно рассмотреть *шары радиуса t с центрами в кодовых словах* и заметить, что они не пересекаются, а каждый из них состоит из $\sum_{i=0}^{t}\binom{n}{i}(q-1)^i$ наборов (каждый набор в шаре однозначно определяется не более чем t позициями, в которых он отличается от центра, и значение в каждой позиции определяется $q-1$ способом). Поэтому общее число слов во всех шарах равно

$$m_q(n,d)\biggl(\sum_{i=0}^{t} \binom{n}{i}(q-1)^i\biggr),$$

но оно не может быть больше q^n — общего числа q-ичных слов, откуда следует нужное неравенство. $\square$

Коды, для которых достигается эта граница, называются *совершенными* (или *плотно упакованными*, потому что они порождают совершенную упаковку многомерного q-ичного куба шарами).

---

[9] *Мощностью кода* называется число элементов в нём.

[10] Ричард Хэмминг (1915–1998) — американский математик, Кальямпуди Радхакришна Рао (р. 1920) — индийский специалист по математической статистике.

Нижняя граница сферической упаковки была указана Гилбертом[11].

Теорема 2. *Для любых n, q, d*

$$m_q(n,d) \geqslant \frac{q^n}{\sum\limits_{i=0}^{d-1} \binom{n}{i}(q-1)^i}.$$

В частности, при $q=2$

$$m_2(n,d) \geqslant \frac{2^n}{\sum\limits_{i=0}^{d-1} \binom{n}{i}}.$$

Доказательство. Рассмотрим максимальный код с расстоянием d (его мощность равна по определению $m_q(n,d)$) и построим шар радиуса $d-1$ с центром в каждом кодовом слове. Выше было показано, что мощность каждого такого шара равна $\sum_{i=0}^{d-1}\binom{n}{i}(q-1)^i$. Объединение этих шаров должно покрывать весь q-ичный куб (непокрытая вершина была бы удалена от центров шаров, т. е. от кодовых слов, на расстояние не меньше d, и её можно было бы добавить к коду, не уменьшая его расстояния и увеличивая мощность, что противоречит его максимальности), поэтому

$$m_q(n,d)\left(\sum_{i=0}^{d-1}\binom{n}{i}(q-1)^i\right) \geqslant q^n,$$

откуда и следует нужная нам оценка[12]. □

Идеи доказательства границ сферической упаковки давно известны в геометрии[13] и используются также в теории приближений[14]. Эти идеи давно попали и в сборники олимпиадных задач.

Задача 7. Докажите, что на стол размера 12×22 можно положить не менее 74 монет единичного радиуса.

Задача 8 (первый вопрос — ММО 1958, второй тур, 8.5, 9.4). Обозначим через a наибольшее число непересекающихся кругов диаметра 1,

[11] Эдгар Гилберт (E. Gilbert, 1923–2013) — известный американский специалист по дискретной математике. Не нужно путать Э. Гилберта со знаменитым немецким математиком Д. Гильбертом (D. Hilbert, 1862–1943).

[12] Советский математик Р. Р. Варшамов (1927–1999) чуть уточнил эту оценку для любых линейных кодов, поэтому её часто называют *границей Варшамова — Гилберта*.

[13] Неравенства Бликфельдта для плотности упаковки шаров в пространстве, см. [15].

[14] Для установления неравенств между *энтропией* и *ёмкостью метрических пространств*.

центры которых лежат внутри многоугольника M, через b — наименьшее число кругов радиуса 1, которыми можно покрыть весь многоугольник M, а через c — наибольшее число непересекающихся кругов радиуса 1, центры которых лежат внутри многоугольника M. Что больше: a или b? b или c?

Задача 9 (ММО 1980, 9.2). На пульте имеется несколько кнопок, с помощью которых осуществляется управление световым табло. После нажатия любой кнопки некоторые лампочки на табло переключаются (для каждой кнопки есть свой набор лампочек, причём наборы могут пересекаться). Докажите, что число состояний, в которых может находиться табло, равно некоторой степени числа 2. (Два состояния табло различны, если они различаются состоянием хотя бы одной лампочки).

Задача 10. Комиссия составила N списков, и оказалось, что для любых двух списков найдётся третий, который получается из этих двух сначала объединением, а затем последующим вычёркиванием тех, кто входил в оба списка. Докажите, что N на единицу меньше некоторой степени двойки.

§ 3. Границы для кодов с большими расстояниями

Однако все его попытки отыскать ключ к шифру проваливались раз за разом, и тогда Смит понял, что придерживался изначально ложной посылки.

Джон Гласби, «Чёрное зеркало»

Для двоичных кодов в случае больших расстояний оценка теоремы 1 становится очень грубой и её можно существенно уточнить. Справедливы следующие утверждения.

Лемма 1. *Выполнено неравенство* $m(n, d) \leqslant 2m(n-1, d)$.

Доказательство. У кода мощности $m(n, d)$ не менее половины кодовых слов имеют одинаковую n-ю компоненту (нуль или единицу). Если её отбросить, то получим код мощности не меньше $m(n, d)/2$ с расстоянием не меньшим d. □

Лемма 2. *Если d нечётно, то* $m(n, d) = m(n+1, d+1)$.

Доказательство. Добавим к каждому слову максимального кода с длиной n и расстоянием d ещё одну компоненту так, чтобы вес полученного слова был чётным (рассмотрим *расширенный код*). Мощность кода не изменится, а расстояние между любыми словами кода станет чётным.

Так как оно не уменьшилось, то оно будет не меньше $d+1$ (и равно $d+1$ там, где было равно d). Отсюда $m(n,d) \leqslant m(n+1,d+1)$.

Обратно, пусть имеем код мощности $m(n+1,d+1)$. Возьмём в нём два слова на расстоянии $d+1$ и выберем компоненту, которая в этих двух словах различна. Отбросив её во всех словах, получим код мощности $m(n+1,d+1)$ с длиной n и минимальным расстоянием $(d+1)-1=d$. Значит, $m(n,d) \geqslant m(n+1,d+1)$, отсюда следует нужное равенство. □

Теорема 3 (граница Плоткина[15]).

1) *При* $2d > n \geqslant d$

$$m(n,d) \leqslant 2\left\lfloor \frac{d}{2d-n} \right\rfloor;$$

2) *при* $2d=n$ *справедливо неравенство* $m(n,d) \leqslant 2n$*;*

3) *при нечётном* d *и* $2d+1 > n \geqslant d$

$$m(n,d) \leqslant 2\left\lfloor \frac{d+1}{2d+1-n} \right\rfloor;$$

4) *при* $2d+1=n$ *справедливо неравенство* $m(n,d) \leqslant 2n+2$.

Доказательство. Рассмотрим максимальный код мощности $m = m(n,d)$ с расстоянием d и оценим сумму R попарных расстояний между его словами. Очевидно, $R \geqslant dm(m-1)/2$, причём равенство возможно лишь для эквидистантных кодов (у которых попарные расстояния между словами равны). Если через h_i обозначить число кодовых слов, у которых i-я компонента равна 1, то $R = \sum_{i=1}^{n} h_i(m-h_i)$ (количество пар слов, различающихся в i-й компоненте, равно $h_i(m-h_i)$). Согласно неравенству между средним геометрическим и средним арифметическим $h_i(m-h_i) \leqslant m^2/4$, поэтому при целых h_i имеем $h_i(m-h_i) \leqslant \lfloor m^2/4 \rfloor$ и $R \leqslant n\lfloor m^2/4 \rfloor$. Таким образом,

$$\frac{dm(m-1)}{2} \leqslant \frac{nm^2}{4}.$$

Значит, при $2d > n$, если m чётно, имеем

$$m \leqslant \frac{2d}{2d-n} \leqslant 2\left\lfloor \frac{d}{2d-n} \right\rfloor,$$

а если m нечётно, то

$$m+1 \leqslant \frac{2d}{2d-n} \leqslant 2\left\lfloor \frac{d}{2d-n} \right\rfloor.$$

При $2d=n$ имеем согласно лемме 1

$$m = m(2d,d) \leqslant 2m(2d-1,d) \leqslant 4\left\lfloor \frac{d}{2d-(2d-1)} \right\rfloor = 4d = 2n.$$

[15] М. Плоткин — американский специалист по теории кодирования, доказал эту теорему около 1960 г.

Если d нечётно, то при $2d+1>n$ по лемме 2

$$m(n,d)=m(n+1,d+1)\leqslant 2\left\lfloor\frac{d+1}{2d+1-n}\right\rfloor,$$

а при $2d+1=n$

$$m=m(2d+1,d)=m(2d+2,d+1)\leqslant 4(d+1)=2(n+1). \qquad \square$$

В. И. Левенштейн[16] доказал (см., например, [6, 7, 14]), что равенства в указанных оценках возможны тогда и только тогда, когда существуют матрицы Адамара любого порядка n, кратного 4. О матрицах Адамара см. раздел 3.1.

Задача 11 (на основе ММО 1993, 10.5). В ботаническом определителе растения описываются 100 бинарными признаками. Определитель считается хорошим, если любые два растения отличаются более чем по половине признаков. Докажите, что в хорошем определителе описано а) не более 50 растений, б) не более 34 растений.

Задача 12. В ботаническом определителе растения описываются 128 бинарными признаками. Определитель считается точным, если любые два растения отличаются не более чем в половине признаков. Докажите, что в точном определителе описано не более 256 растений, и покажите, что существует такой определитель, содержащий ровно 256 растений.

Задача 13. (i) (ММО 1948, второй тур, 9–10 кл., задача 4). Какое наибольшее число лучей можно провести из одной точки в трёхмерном пространстве так, чтобы все попарные углы между ними были тупыми?

(ii) Будем говорить, что два набора чисел $(a_1,\ldots,a_n)$ и $(b_1,\ldots,b_n)$ длины n образуют «тупой угол», если $a_1b_1+\ldots+a_nb_n<0$. Докажите, что если любые два из m наборов образуют тупой угол, то $m\leqslant n+1$.

Задача 14 (ВМО 1970, 9.4). Из цифр 1 и 2 составили 5 n-разрядных чисел так, что у каждых двух чисел совпали цифры ровно в m разрядах, но ни в одном разряде не совпали все пять чисел. Докажите, что $2/5\leqslant m/n\leqslant 3/5$.

Задача 15. Из цифр 0 и 1 составлены $N=2k+1$ различных n-разрядных чисел, у любых двух из которых совпадают ровно m разрядов. Докажите, что

$$\frac{m}{n}\leqslant\frac{N+1}{2N},$$

а если $N=2k$, то

$$\frac{m}{n}\leqslant\frac{N}{2(N-1)}.$$

[16] Владимир Иосифович Левенштейн (1935–2017) — известный российский специалист по теории кодирования.

Задача 16. Из цифр 0 и 1 составлены N различных n-разрядных чисел, причём нет ни одного разряда, в котором они все совпадают. Докажите, что $m/n \geqslant 2/N$.

3.1. Матрицы Адамара

> Золотая карта теперь была в круглых дырочках, словно швейцарский сыр. Все они располагались в узлах координатной сетки мсье Декарта, однако не все узлы были пробиты. Результат являл собой странное смешение упорядоченного и случайного; таким, наверное, предстаёт чётко отпечатанный, но зашифрованный текст.
>
> *Нил Стивенсон, «Система мира»*

Жак Адамар (выдающийся французский математик, 1865–1963) пришёл к этим матрицам, решая экстремальную задачу: найти среди всех матриц данного размера $n \times n$ с элементами, по модулю не превосходящими единицы, матрицу с максимальным определителем. Оказалось, что если n кратно 4 (или равно 2), то такая матрица состоит из ± 1, причём скалярное произведение любых двух (различных) строк и любых двух столбцов равно нулю. Под *скалярным произведением* двух векторов $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ понимается число $x_1y_1 + \ldots + x_ny_n$. Векторы с нулевым скалярным произведением называются *ортогональными*. Матрицы с указанными свойствами называются *матрицами Адамара*. Легко доказать, что смена знака на противоположный в любой строке превращает матрицу Адамара в другую матрицу Адамара[17]. Аналогичное утверждение верно и для столбцов. Поэтому иногда в определение матриц Адамара добавляют условие, что верхняя строка и левый столбец должны состоять из единиц.

Из условия ортогональности любой строки верхней единичной строке следует, что во всех остальных строках поровну плюс и минус единиц, значит, n чётно. В теории кодирования используют матрицы, которые получаются из матриц Адамара заменой минус единиц на нули. Заменяя нули и единицы на мальчиков и девочек и решив задачу 18, читатель легко докажет, что если размер матрицы Адамара $n > 2$, то он кратен четырём.

Для построения матриц Адамара придуманы хитроумные методы, но гипотеза о том, что для любого n, кратного четырём, существуют такие матрицы порядка n, пока не доказана.

Приведём здесь только одну, самую простую конструкцию матриц Адамара, которая позволяет строить их для любого $n = 2^k$. Для $n = 2$, очевидно,

[17] Надо только не забыть проверить ортогональность столбцов.

матрица Адамара имеет вид
$$A_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$
Если матрица A_{2^k} размера $2^k \times 2^k$ уже построена, то матрицу $A_{2^{k+1}}$ можно составить из четырёх блоков размера $2^k \times 2^k$ следующим образом:
$$A_{2^{k+1}} = \begin{pmatrix} A_{2^k} & A_{2^k} \\ A_{2^k} & -A_{2^k} \end{pmatrix}.$$

Задача 17. Докажите по индукции, что построенная последовательность матриц действительно состоит из матриц Адамара[18].

Теперь читатель легко решит последний пункт задачи 12.

Задача 18 (ВМО 1983, 9.5). Группа детского сада построилась парами друг за другом. При этом оказалось, что в каждой из двух колонн стоит поровну мальчиков и девочек, а число пар, в которых стоят девочка и мальчик, равно числу остальных пар. Докажите, что число детей в группе делится на 8.

Задача 19 (ММО 1959, второй тур, 7.5). Даны числа $x_i = \pm 1$, $i = 1, \ldots, n$. Докажите, что если $x_1x_2 + x_2x_3 + \ldots + x_{n-1}x_n + x_nx_1 = 0$, то n делится на 4.

Задача 20 (ММО 1971, второй тур, 10.1). В вершинах правильного n-угольника стоят числа 1 или -1. Если его повернуть на произвольный угол $2k\pi/n$, $k = 1, \ldots, n-1$, перемножить числа в совместившихся вершинах и все такие произведения сложить, то результат будет равен нулю. Докажите, что n есть квадрат целого числа.

В задаче[19] 20, как теперь нетрудно догадаться, на самом деле тоже идёт речь о матрицах Адамара, а точнее, о таких матрицах с дополнительным свойством цикличности. На олимпиаде также предлагалось найти все такие матрицы. Ответ к этой задаче жюри не знало, и задача получила название «проблемы Зелевинского»[20]. Ещё раньше она была известна как проблема Райзера о циркулянтных матрицах Адамара. Она не решена и по сей день.

3.2. Эквидистантные равновесные коды

Под этим термином скрывается довольно простое понятие, о котором фактически идёт речь в следующей задаче (рассматривавшейся в [4]).

---

[18] Эту конструкцию ещё до Адамара предложил Дж. Сильвестр (1814–1897), поэтому такие матрицы можно называть матрицами Адамара — Сильвестра.

[19] Задача 19 по существу является её упрощённым вариантом и близка к задаче 18.

[20] На олимпиаду её предложил известный математик А. В. Зелевинский (1953–2013), в прошлом призёр международной олимпиады, а в то время студент второго курса.

Задача 21. Даны 10 множеств из 4 элементов каждое, причём объединение любых двух содержит ровно 7 элементов. Сколько элементов может быть в объединении всех этих множеств? Укажите все возможные значения.

Ответ: $1+3\cdot 10=31$ и $1+3+3^2=13$.

Пусть объединение этих множеств состоит из m элементов. Сопоставив каждому из этих множеств набор из нулей и единиц с 4 единицами на позициях, соответствующих элементам данного множества, получим множество вершин двоичного m-мерного куба, лежащих в его четвёртом слое (т. е. имеющих веса, равные 4). Условие задачи означает, что попарные расстояния между этими вершинами равны 7. Коды, лежащие в одном слое, т. е. состоящие из наборов равного веса, называются равновесными, а коды, у которых расстояния между любыми двумя кодовыми словами равны, называются эквидистантными. Задача описания всех равновесных эквидистантных кодов крайне сложна.

Задача 22. Докажите, что максимальный эквидистантный код кодовой длины $n=q^2+q+1$ с весом $q+1$ и расстоянием $2q+1$ имеет мощность не более q^2+q+1. Равенство возможно тогда и только тогда, когда существует проективная плоскость порядка q.

§ 4. Простейший пример кода, исправляющего одну ошибку

Всё следует сделать настолько простым,
насколько это возможно, но не проще.

Альберт Эйнштейн

Клод Шеннон[21] выдвинул идею *помехоустойчивого кодирования* (кодирования с исправлением ошибок). Рассмотрим пример кода, исправляющего одну ошибку (простейший нетривиальный частный случай кода Хэмминга). Пусть нужно передать двоичное слово (x_1,x_2,x_3,x_4) Добавим к нему проверочные символы $x_5=x_1+x_3+x_4$, $x_6=x_1+x_2+x_4$, $x_7=x_1+x_2+x_3$ (знак $+$ здесь обозначает сложение по модулю два; символы x_1,x_2,x_3,x_4 называются *информационными*). Процедура вычисления по информационным символам проверочных и составления из них кодового слова (закодированного сообщения) называется *кодированием* (так же называется и само отображение исходного сообщения в кодовое слово).

[21] Клод Элвуд Шеннон (1916–2001) — американский математик и инженер, основоположник теории информации.

На языке матриц в рассматриваемом примере кодирование сводится к умножению матрицы M на транспонированный вектор $(x_1, x_2, x_3, x_4)^T$ (т. е. вектор, расположенный в столбец):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}.$$

Передаём закодированное сообщение $c = (x_1, \ldots x_7)$ и получаем *зашумлённое сообщение* $r = c + e$, где $e = (e_1, \ldots, e_7)$ — вектор ошибок. В нашем примере он имеет вес 1, так как по предположению ошибка может произойти (если произойдёт) только в одной позиции. Например, возможно $e = e_3 = (0, 0, 1, 0, 0, 0, 0)$. Тогда

$$r = c + e = (c_1, c_2, c_3 + 1, c_4, c_5, c_6, c_7) = (c_1, c_2, \bar{c}_3, c_4, c_5, c_6, c_7),$$

где $\overline{0} = 1$, $\overline{1} = 0$. Число 3 будет в рассматриваемом случае *позицией ошибки*. Для определения позиции ошибки (а значит, и нахождения самой ошибки) можно вычислить *проверочные суммы*

$$S_1 = r_1 + r_3 + r_4 + r_5,$$
$$S_2 = r_1 + r_2 + r_4 + r_6,$$
$$S_3 = r_1 + r_2 + r_3 + r_7.$$

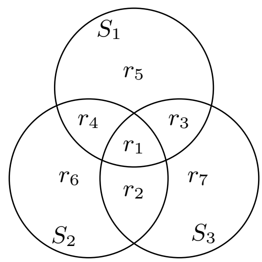


Рис. 1. Код Хэмминга с блоковой длиной 7

Наглядно все эти суммы изображены на рис. 1. Каждая сумма содержится в своём круге. На матричном языке эта процедура равносильна умножению матрицы на вектор:

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \\ r_7 \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}.$$

Указанная матрица H называется *проверочной матрицей* кода. В краткой нотации умножение матрицы H на вектор r^T записывается как $S = Hr^T$ (где S — вектор-столбец $(S_1, S_2, S_3)^T$).

Заметим, что матрица H выбрана так, что $Hc^T = 0$ (где 0 — нулевой вектор-столбец) для любого кодового вектора c. Это можно проверить непосредственно (и даже не пользуясь матричным языком, просто подставив в указанные выше три суммы S_i вместо x_5, x_6, x_7 их выражения через (x_1, x_2, x_3, x_4). На матричном языке это также легко показать. Заметим, что матрица M представима в виде

$$\begin{pmatrix} E \\ A \end{pmatrix},$$

где E — единичная матрица[22] размера 4×4, а A — матрица размера 3×4. Матрица H составлена из той же подматрицы A и единичной подматрицы размера 3×3, поэтому

$$\begin{aligned} Hc^T &= A \cdot (x_1, x_2, x_3, x_4)^T + E \cdot (x_5, x_6, x_7)^T = \\ &= A \cdot (x_1, x_2, x_3, x_4)^T + (x_5, x_6, x_7)^T = 0^T, \end{aligned}$$

так как в силу равенства

$$M \cdot (x_1, x_2, x_3, x_4)^T = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)^T$$

имеем

$$A \cdot (x_1, x_2, x_3, x_4)^T = (x_5, x_6, x_7)^T.$$

Используя легко проверяемые свойства матричного умножения и сложения, получаем

$$S = Hr^T = H(c^T + e^T) = He^T = H_i,$$

где e — вектор с единственной единицей в i-й позиции, H_i — i-й столбец матрицы H (это равенство можно проверить и непосредственно, заменяя r_j в суммах S_i на $c_j + e_j = x_j + e_j$). Заметим теперь, что все столбцы матрицы H различны и отличны от нуля[23]. Поэтому по столбцу H_i можно однозначно определить его номер, а значит, и позицию ошибки. Если же ошибки не было, то очевидно, что $S = Hr^T = Hc^T = 0$, и это равенство можно проверить, сравнив S с нулевым вектором-столбцом (если же S ненулевой, то ошибка, очевидно, была). Для определения позиции ошибки по вычисленному вектору S (называемому *синдромом*) можно заготовить таблицу, содержащую номера позиций ошибок, например в двоичной записи. Эта таблица состоит из восьми строк, которые занумерованы двоичными набо-

[22] Квадратная матрица называется *единичной*, если на главной диагонали стоят единицы, а в остальных местах нули. Главная диагональ — это диагональ, идущая от левого верхнего угла матрицы к правому нижнему углу. Если умножить единичную матрицу на вектор, то в результате получится тот же вектор.

[23] Это стало возможным благодаря тому, что всего имеется 8 различных двоичных столбцов высоты 3.

рами длины 3. Если ошибки нет, то номер можно считать равным нулю. Если проверочная матрица имеет вид

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix},$$

то эта таблица не требуется, так как синдром S совпадает в этом случае с двоичным номером позиции ошибки. Разумеется, матрицу M также придётся тогда изменить.

Построенный код является частным случаем кодов Хэмминга, которые рассматриваются далее. В случае когда длина кодовых слов равна 31, идея построения кода очень близка к задаче 5, предложенной французским математиком Люка ещё в XIX веке (это станет ясно после прочтения § 5). Да и следующие три задачи тоже к ней очень близки.

Задача 23 (из задачной базы московских олимпиад). В n-элементном множестве выбрано $5n$ различных двухэлементных подмножеств. Докажите, что, объединяя их попарно, можно получить не менее $45n$ трёхэлементных подмножеств.

Задача 24. Найдите среди 63 монет фальшивую, выполнив всего 6 взвешиваний, если известно, что она легче остальных. За одно взвешивание можно взвесить сразу несколько монет. План взвешиваний требуется составить заранее.

Задача 25 (ММО 1990, 8.5). Табло, состоящее из 64 лампочек, управляется 64 кнопками: каждая лампочка — своей кнопкой. За одно включение можно одновременно нажать любой набор кнопок и записать, какие лампочки при этом зажглись. За какое наименьшее количество включений можно узнать обо всех лампочках табло: какая лампочка какой кнопкой управляется?

Обратим внимание на некоторые свойства построенного выше кода. Его мощность равна $2^4 = 16$, сумма любых двух кодовых слов по модулю два опять является кодовым словом (т. е. этот код линейный), расстояние кода равно трём. Меньше трёх оно быть не может, иначе бы он не исправлял ошибки, но можно расстояние вычислить и явно, заметив, что $d(a, b) = d(a + b, 0)$. Поскольку код содержит кодовое слово веса нуль (нулевое слово) и кодовое расстояние равно трём, в коде нет слов веса 1 или 2. Для каждого кодового слова рассмотрим шар радиусом 1 с центром в этом слове. Этот шар содержит, кроме центра, ещё 7 двоичных наборов (вершин семимерного двоичного куба), получающихся, если в центральном наборе заменить ровно один из семи его символов на противоположный.

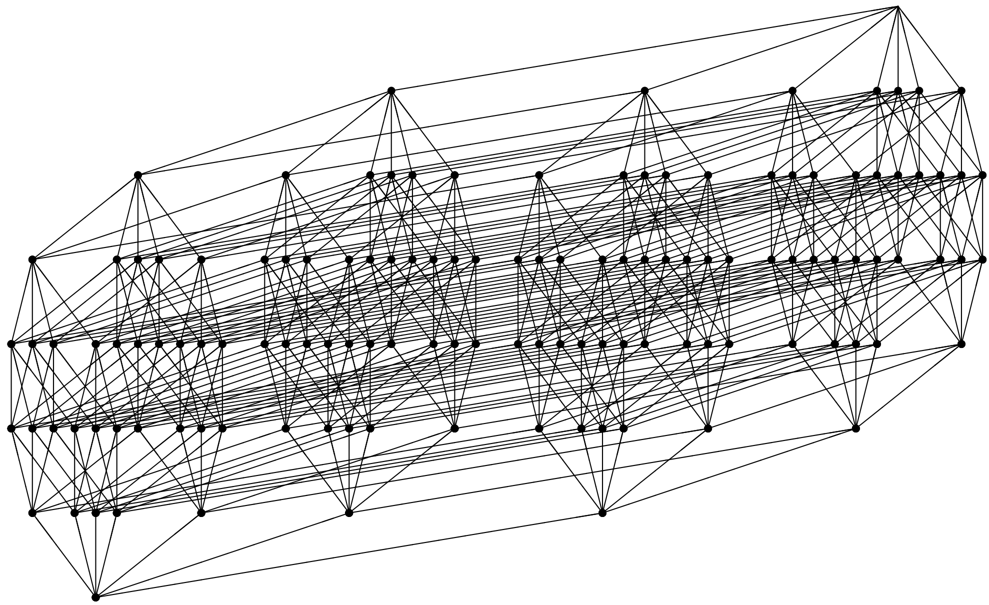


Рис. 2. Семимерный куб

Эти шары с центрами в кодовых словах не пересекаются[24] (не имеют общих вершин) и поэтому в совокупности содержат $2^4 \cdot 8 = 2^7$ различных вершин куба, т. е. все его вершины (в семимерном кубе 2^7 вершин). Такие точные покрытия многомерного куба непересекающимися шарами называются *совершенными*, а соответствующие им коды — *совершенными кодами*.

В [4] исходя только из свойства совершенности указанного кода объяснено, как можно однозначно определить число кодовых вершин на третьем слое семимерного куба[25], см. рис. 2.

Там найден его *весовой спектр*, а именно количество слов каждого заданного веса. Он имеет вид

$$a_0 = 1, \quad a_1 = 0, \quad a_2 = 0, \quad a_3 = 7,$$
$$a_4 = 7, \quad a_5 = 0, \quad a_6 = 0, \quad a_7 = 1.$$

Также там объяснено, что кодовые слова веса три определяют интересную комбинаторную конфигурацию — *систему троек Штейнера*[26], и показано, что эта система троек изоморфна конфигурации семи трёхто-

[24] Если шары с центрами a, b имеют общую вершину c, то $d(a,b) \leqslant d(a,c) + d(c,b) \leqslant 2$, что невозможно.

[25] k-й слой куба состоит из всех вершин веса k и очевидно содержит $\binom{7}{k}$ вершин.

[26] Якоб Штейнер (1796–1863) — выдающийся швейцарский геометр.

чечных прямых в проективной плоскости на семи точках — так называемой плоскости Фано[27] (рис. 3).

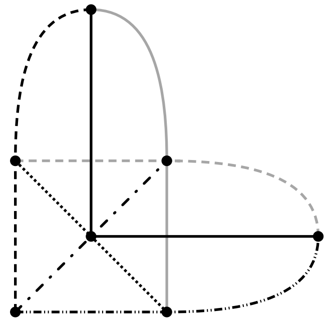


Рис. 3. Плоскость Фано

Коды, все вершины которых лежат на одном слое куба, называются *равновесными кодами*. Таким образом, максимальный код веса три с блоковой длиной семь также имеет мощность 7. Кодовые слова веса четыре также определяют интересную комбинаторную конфигурацию. Каждое из них задаёт четырёхэлементное подмножество в множестве $\{1, 2, 3, 4, 5, 6, 7\}$, состоящее из номеров позиций единиц в этом слове. Система из этих семи четвёрок образует пример *блок-схемы*, в которой каждая пара элементов принадлежит в точности двум четвёркам[28].

Указанные конфигурации (блок-схемы) третьего и четвёртого слоёв обладают ещё одним интересным свойством. Они двойственны друг к другу, а именно: четвёрки из второй блок-схемы являются дополнениями троек из первой блок-схемы до множества $\{1, \ldots, 7\}$.

§ 5. Код Хэмминга

Два выходных подряд я приходил и обнаруживал, что все мои данные выгружены и ровным счётом ничего не сделано. Я был в ярости, потому что мне были нужны ответы, а два выходных оказались потеряны зря. Тогда я сказал себе: «Чёрт, если машина может обнаружить ошибку, то что мешает ей определить, где эта ошибка произошла, и исправить её?»

Ричард Хэмминг, по: Дж. Маккормик, «Девять алгоритмов, которые изменили мир»

Начнём с двоичного кода Хэмминга. Его можно построить, например, так. Пусть $n = 2^m - 1$, $k = n - m$, и пусть $x_1, \ldots, x_k$ — информационный вектор, который нужно закодировать. Добавим к нему проверочные символы $x_{k+1}, \ldots, x_n$, для вычисления которых умножим вектор-столбец $X = (x_1, \ldots, x_k)^T$ на (m, k)-матрицу M_n (с m строками и k столбцами), столбцы которой представляют из себя все возможные наборы длины m из нулей и единиц, содержащие хотя бы две единицы (таких наборов ровно

[27] Джино Фано (1871–1951) — выдающийся итальянский математик.

[28] О блок-схемах см., например, [4, 14, 16].

$2^m - 1 - m = n - m = k$). Добавим к этой матрице ещё m столбцов, содержащих ровно одну единицу каждый (можно считать, что они образуют квадратную матрицу с единицами по диагонали), и получим (m, n)-матрицу H_n, столбцами которой являются все возможные ненулевые наборы длины m из нулей и единиц. Так же как и в § 4, можно проверить, что для любого кодового слова $x = (x_1, \ldots, x_n)$ справедливо матричное равенство

$$H_n x^T = 0.$$

Матрица H (далее индекс n опускаем) называется проверочной матрицей этого кода. Она позволяет не только проверить, является ли данное слово c кодовым, но и позволяет найти ошибку, если она была одна. Действительно, как и в § 4, имеем

$$S = Hr^T = H(c^T + e^T) = He^T = H_i,$$

где e — вектор ошибок, имеющий одну единицу в i-й позиции, а H_i — столбец матрицы H с номером i. Так как все столбцы различны и отличны от нулевого, то по синдромному вектору S можно однозначно найти ошибку.

Как и в § 4, рассмотрим некоторые свойства построенного кода. Его мощность равна 2^k, он линейный, размерность его как линейного пространства над полем из двух элементов равна k (потому что если базис пространства состоит из k векторов, то пространство состоит из всех возможных их сумм, которых ровно 2^k, включая и пустую сумму, равную по определению нулю), расстояние кода равно трём (меньше трёх оно быть не может, иначе бы он не исправлял ошибки). Так как код имеет кодовое слово веса нуль (нулевое слово) и расстояние его равно трём, то кодовых слов веса 1 или 2 в нём нет. Для каждого кодового слова рассмотрим шар радиуса 1 с центром в этом слове. Этот шар содержит, кроме центра, ещё n двоичных наборов (вершин n-мерного двоичного куба). Эти шары с центрами в кодовых словах не пересекаются, поэтому в совокупности содержат $2^k \cdot (n+1) = 2^{k+m} = 2^n$ различных вершин куба, т. е. все эти вершины. Поэтому код Хэмминга H_n является совершенным (и плотно упакованным). Очевидно также, что он лежит на границе сферической упаковки. Обратно, любой максимальный код с расстоянием три имеет такие же параметры, что и код Хэмминга. Действительно, если граница оценки $m(n, 3) \leqslant 2^n/(n+1)$ достигается, то $n + 1 = 2^m$ (иначе данная дробь не будет целым числом), откуда $m(n, 3) = 2^{n-m}$. Тем самым полностью решена задача 3, и теперь легко решить задачу 26.

Задача 26. Всегда ли можно угадать число от 1 до 2048, задав 15 вопросов с ответом «да» или «нет», если на один из них может быть дан неправильный ответ? Вопросы требуется составить заранее.

Ещё одна задача на ту же тему:

Задача 27 [19, p. 120]. У трёх мудрецов шляпы двух цветов. Ведущий надевает мудрецам шляпы так, что в результате каждый видит шляпы всех остальных мудрецов, но не видит своей шляпы и не знает её цвета. По команде ведущего они одновременно называют цвет. Каждый мудрец должен назвать цвет, исходя только из того, какие цвета он видит у остальных, но им разрешается пасовать, что означает отказ от угадывания. Мудрецы выигрывают только при условии, что хотя бы один из них угадал цвет и при этом никто не назвал цвет неправильно. Перед тестом мудрецам сообщили правила и дали возможность договориться о том, как действовать во время теста. Оптимальная стратегия — это стратегия, которая для всевозможных раскладов шляп даёт наибольшее число выигрышей.

а) Предложите стратегию мудрецов, при которой они выигрывают больше чем в половине случаев.

б) Найдите оптимальную стратегию и докажите, что она оптимальна.

5.1. q-ичные коды Хэмминга

Я бедная девушка, у которой плохо с арифметикой! Выше двух для меня сразу начинается высшая математика!

Дмитрий Емец, «Таня Гроттер»

В некоторых случаях существуют также совершенные q-ичные коды. Границы Хэмминга они могут достигать, например, когда $1+(q-1)n=q^m$. Тогда мощность кода будет равна q^{n-m}, где n — его блоковая длина. Такие коды можно построить в случае существования конечного поля из q элементов[29].

Полем называется любое множество, на котором можно определить операции сложения и умножения так, что эти операции удовлетворяют тем же законам, что и операции сложения и умножения рациональных чисел, а именно, переместительному: $a+b=b+a$, $ab=ba$, сочетательному: $(a+b)+c=a+(b+c)$, $(ab)c=a(bc)$, распределительному: $a(b+c)=ab+ac$, и удовлетворяют тождествам $a+0=a$, $a\cdot 1=a$, а также имеют однозначно определённые обратные операции вычитания $a-b$ и деления a/b, удовлетворяющие тождествам $(a-b)+b=a$, $(a/b)b=a$. Из алгебры известно, что порядок (число элементов) q любого конечного поля — степень простого

[29] Есть гипотеза, что при q, не равном степени простого числа, таких совершенных кодов не существует.

числа p, $q=p^n$, причём для любого такого q существует единственное[30] поле порядка q, называемое полем Галуа[31] и обозначаемое $\mathrm{GF}(q)$. Например, $\mathrm{GF}(2)=(\{0,1\},\oplus,\cdot)$.

Построим над полем $\mathrm{GF}(q)$ коды, являющиеся обобщением кодов Хэмминга. Для каждого ненулевого вектора v длины m над полем $\mathrm{GF}(q)$ рассмотрим множество коллинеарных ему векторов λv, где $\lambda\in\mathrm{GF}(q)\setminus\{0\}$. Это множество образует прямую в пространстве $\mathrm{GF}(q)^n$, проходящую через начало координат. Выберем на каждой такой прямой (а их будет $n=(q^m-1)/(q-1)$, так как они имеют только одну общую точку — начало координат) любую точку, отличную от начала координат, например можно выбрать точку, у которой последняя координата равна 1. Это всегда можно сделать, за исключением случая, когда все точки прямой имеют нулевую последнюю координату. В этом случае можно выбрать точку, у которой предпоследняя координата равна единице. Если таковой не найдётся и предпоследняя координата равна нулю, можно найти точку, у которой третья с конца координата равна единице, и т. д. Указанная точка на прямой (и соответствующий радиус-вектор) определяется однозначно. Все остальные получаются теперь умножением на элементы поля (при умножении на нуль получается нулевой вектор). Очевидно, что любой ненулевой вектор из пространства $\mathrm{GF}(q)^n$ получается из одного из указанных выше векторов $v_1,\ldots v_n$ умножением на ненулевой элемент поля, причём такое представление определено однозначно.

Рассмотрим (m,n)-матрицу H_n над полем $\mathrm{GF}(q)$, столбцы которой являются указанными векторами $v_1,\ldots,v_m$. Определим код как множество таких векторов c, что $Hc^T=0$ (т. е. как нулевое пространство этой матрицы). Тогда матрица H_n является проверочной матрицей указанного кода. Код исправляет одну ошибку, потому что, как и в § 4, имеем

$$S=Hr^T=H(c^T+e^T)=He^T=e_iH_i,$$

где e — вектор ошибок, имеющий один ненулевой символ e_i в i-й позиции, а H_i — i-й столбец матрицы H. Так как все столбцы различны и отличны от нулевого, а результаты их умножения на ненулевые элементы поля тоже различны, по вектору S можно однозначно найти и позицию ошибки i, и её величину e_i. Поскольку матрица H_n (после подходящей перестановки столбцов) содержит единичную (m,m)-подматрицу, размерность кода (размерность нулевого подпространства) равна $n-m$, а значит, его мощность равна q^{n-m} (множество всех линейных комбинаций базисных векторов над полем

[30] С точностью до изоморфизма.

[31] Эварист Галуа (1811–1832) — великий французский математик.

$\mathrm{GF}(q)$ имеет мощность q^{n-m}), т. е. код находится точно на границе сферической упаковки и, следовательно, является совершенным кодом. Тем самым решён п. (г) задачи 4. Решение п. (б) читатель может найти в [2, с. 165–166].

§ 6. Коды Рида — Соломона

В полях Галуа, полных цветов, примитивные корни танцуют часами[32].

С. Б. Вейнстейн (IEEE Transactions of Information Theory, 1971)

Пусть $q > 2$ — простое число. Напомним, что q-ичным линейным кодом C длины n и размерности k (сокращённо $[n, k]$-кодом) называется любое линейное k-мерное подпространство C пространства $\mathrm{GF}(q)^n$ всех n-мерных векторов над полем $\mathrm{GF}(q)$. Легко видеть, что кодовое расстояние линейного кода равно минимальному весу, который может иметь ненулевой кодовый вектор Если $[n, k]$-код имеет кодовое расстояние d, то он называется $[n, k, d]$-кодом.

Простейший вариант построения кодов Рида — Соломона[33] над полем $\mathrm{GF}(q)$ (сокращённо RS-кодов) следующий. Пусть $k < n \leqslant q$. Сопоставим каждому вектору $a = (a_0, \ldots, a_{k-1}) \in \mathrm{GF}(q)^k$ многочлен

$$a(x) = a_0 + a_1 x + \ldots + a_{k-1} x^{k-1}$$

степени $k - 1$ над полем $\mathrm{GF}(q)$. Пусть $x_1, \ldots, x_n \in \mathrm{GF}(q)$ — различные элементы этого поля. Рассмотрим (линейное) отображение $l\colon \mathrm{GF}(q)^k \to \mathrm{GF}(q)^n$, определяемое равенством

$$l(a) = (a(x_1), \ldots, a(x_n)) \in \mathrm{GF}(q)^n.$$

Образ $l(\mathrm{GF}(q)^k) \subset \mathrm{GF}(q)^n$ этого отображения — линейный код C, называемый RS-кодом. В силу неравенства $n > k$ многочлен $a(x)$ степени $k-1$ однозначно восстанавливается по своим значениям в n точках, поэтому отображение $l\colon \mathrm{GF}(q)^k \to C$ взаимно однозначно. Значит, мощность кода C равна q^k, поэтому его размерность равна k. Кодовое расстояние $d(C) \geqslant n - k + 1$, так как для любого ненулевого многочлена $a(x)$ вектор его значений $(a(x_1), \ldots, a(x_n)) \in C \subset \mathrm{GF}(q)^n$ имеет вес, не меньший $n - k + 1$, потому что ненулевой многочлен степени $k - 1$ имеет не более $k - 1$ корней.

---

[32] In Galois Fields, full of flowers, primitive elements dance for hours.

[33] Ирвинг Рид (1923–2012) и Густав Соломон (1930–1996) — американские специалисты по теории кодирования.

На самом деле $d(C) = n - k + 1$, потому что, согласно так называемой границе Синглтона, для любого $[n, k]$-кода $d(C) \leqslant n - k + 1$. Коды, лежащие на этой границе, называются *кодами с максимальным расстоянием* (maximim-distance separable (MDS) — кодами). Такими являются RS-коды. Бинарные коды не достигают этой границы.

Соответствующая теорема может быть сформулирована ещё и так (определение $m_q(n, d)$ см. в § 2):

Теорема 4 (граница Синглтона[34]), или проекционная граница).

$$m_q(n, d) \leqslant q^{n-d+1}.$$

Доказательство. Достаточно спроектировать все q^k кодовых слов на подпространство $\mathrm{GF}(q)^{k-1}$, для чего надо заполнить нулями последние $n - k + 1$ координат кодового вектора, не изменяя первые $k - 1$ координат. В силу принципа Дирихле какие-то два кодовых слова при этом будут иметь одинаковую проекцию, т. е. первые $k - 1$ координат у них совпадают, поэтому расстояние между ними будет не больше $n - k + 1$. Значит, $d \leqslant n - k + 1$ (причём это верно и для нелинейных кодов мощности больше q^{k-1}). □

Теперь легко решаются задачи 28 и 29.

Задача 28. На n-мерной k-ичной шахматной доске поставлена $k^m + 1$ ладья. Докажите, что найдутся две ладьи, координаты которых различаются не более чем в $n - m$ позициях. В частности, если $m = n - 1$, то найдутся две ладьи, угрожающие друг другу.

Задача 29. Если k равно степени простого числа и $n > m + 1$, то можно в задаче 28 так расставить k^{m+1} ладью, что координаты любых двух будут отличаться не менее чем в $n - m$ позициях. В частности, если $m = n - 2$, то на n-мерной k-ичной шахматной доске можно расставить k^{n-1} ладей так, чтобы они не били друг друга. При $n = 2$ это очевидно.

А к задаче 30 дадим ответ и указание.

Задача 30. Вы заходите в комнату, где лежит шахматная доска, на некоторых клетках которой стоят фигуры (или шашки). Вам сообщают координаты одной клетки доски и дают задание передать информацию об этой клетке вашему другу, который войдёт в эту комнату позже вас. Взятие фигуры или установка на пустую клетку фигуры считается ходом. У вас не будет возможности после посещения этой комнаты встретиться с другом или сообщить ему что-либо. До его прихода позиция на доске

[34]) Ричард Коллом Синглтон (1928–2007) — американский специалист по теории кодирования.

изменяться не будет. Но перед посещением этой комнаты у вас есть возможность договориться с другом о совместных действиях. За какое наименьшее количество ходов можно решить поставленную задачу?

Ответ: достаточно одного хода.

Указание. Координаты клетки можно «закодировать» двоичным набором длины 6. Расстановка фигур на доске определяется двоичным набором длины 64. Воспользуйтесь проверочной матрицей линейного [64, 58]-кода, т. е. кода Хэмминга, расширенного добавлением нулевой координаты. Это двоичная $(6, 64)$-матрица, все столбцы которой различны (можно удалить из неё нулевой столбец и использовать $(6, 63)$-матрицу — проверочную матрицу обычного кода Хэмминга). Если умножить двоичный набор длины 64 на эту матрицу, изменив его в одном месте, можно получить любой заданный набор. Этот набор можно выбрать так, чтобы он «кодировал» нужную клетку.

Ещё одна задача на эту тему:

Задача 31 [19, p. 121]. Шпион, засланный в чужую страну, может использовать для общения с центром только передачи местной радиостанции, ежедневно передающей в эфир одно сообщение длиной 255 бит. У шпиона есть доступ к тексту передаваемого сообщения до его выхода в эфир, но всё, что он может сделать, — это изменить в сообщении один из битов (или вообще ничего не менять). Сколько битов информации сможет ежедневно передавать шпион в Центр? (Разумеется, договариваться с центром о способе шифровки/дешифровки он может и должен заранее.)

§ 7. Теорема Заранкевича и декодирование

> Гамильтон наугад открыл страницу и убедился, что книга, судя по всему, является компендиумом самых диких идей...
>
> *Ричард Тирни, «Крик во тьме»*

Код C с блоковой длиной n называется (e, l)-*списочно-декодируемым*, если для любого слова y длины n в шаре $B(y, e)$ радиуса e с центром в y (в метрике Хэмминга) находится не более l кодовых слов, т. е. $|B(y, e) \cap C| \leqslant l$. Судан и Гурусвами[35] [18] доказали[36], что справедлива

[35] Мадлу Судан (р. 1966), Венкатесан Гурусвами (р. 1976) — индийско-американские специалисты по информатике.

[36] На самом деле они доказали существенно больше, а именно, предложили также эффективный алгоритм, перечисляющий для любого кодового слова список слов, удалённых от него на расстояние, не большее e.

Теорема 5. *Любой (n,k,d)-код $(e, n(d-e))$-списочно-декодируем при $e<n-\sqrt{(n-d)n}$. В частности, любой $RS(n,k)_q$-код $(e, n(d-e))$-списочно-декодируем при $e < n - \sqrt{(k-1)n}$.*

Доказательство. Очевидно, что

$$n-d = \sqrt{n-d}\sqrt{n-d} < \sqrt{(n-d)n} \quad \text{и} \quad n-\sqrt{(n-d)n} \geqslant \frac{d}{2},$$

так как

$$n-\frac{d}{2} = \frac{n+n-d}{2} \geqslant \sqrt{(n-d)n}$$

согласно неравенству между средним арифметическим и средним геометрическим.

Теорема 5 выводится из теоремы Заранкевича, сформулированной в виде задачи 6. Пусть $c_j \in B(y,e)\cap C$, $j \leqslant m$, — все кодовые слова, лежащие в шаре $B(y,e)$, где $y=(y_1,\ldots,y_n)$ — произвольное слово длины n в кодовом алфавите. Нужно показать, что $m \leqslant n(d-e)$.

Определим (n,m)-матрицу A из нулей и единиц так, что $a_{i,j}=1 \Leftrightarrow y_i=c_{j,i}$, где $c_{j,i}$ — i-я координата вектора c_j. Так как для любых j_1, j_2 в силу определения минимального кодового расстояния $d(c_{j_1}, c_{j_2}) \geqslant d$, то число совпадающих координат у этих векторов не больше $n-d$, поэтому в j_1-м и j_2-м столбцах матрицы A имеется не более $n-d$ общих единиц. Значит, матрица не содержит $(n-d+1, 2)$-матриц второго порядка, состоящих из единиц. Поэтому число k единиц в ней, согласно теореме Заранкевича, таково, что

$$n\frac{\frac{k}{n}\left(\frac{k}{n}-1\right)}{2} \leqslant \frac{m(m-1)(n-d)}{2},$$

откуда

$$k(k-n) \leqslant nm(m-1)(n-d).$$

Слова y, c_j совпадают не менее чем в $t=n-e>\sqrt{(n-d)n}$ позициях, так как по условию $d(c_j,y)\leqslant e$, поэтому в каждом столбце матрицы A не меньше t единиц, откуда следует, что $k \geqslant mt$. Можно считать, что $k > n$, иначе

$$m < \sqrt{\frac{n}{n-d}} \leqslant \sqrt{n} \leqslant n(d-e).$$

Поэтому

$$mt(mt-n) \leqslant k(k-n) \leqslant m(m-1)n(n-d),$$

откуда

$$m^2t^2 - mnt \leqslant m(m-1)n(n-d).$$

Значит,

$$m^2(t^2 - n(n-d)) \leqslant mn(t-n+d),$$
$$m \leqslant \frac{n(t-n+d)}{t^2 - n(n-d)} \leqslant n(t+d-n) = n(d-e) < n^2.$$

Теорема доказана. □

§ 8. Алфавитное кодирование

Алфавиты стали более знакомыми, и теперь, кроме букв, стали попадаться и цифры — в порядках, которые я не сразу узнала.

Скарлетт Томас, «Наваждение Люмаса»

Теория кодирования — обширная область, имеющая тесные связи с алгеброй, теорией чисел, комбинаторикой, теорией графов, теорией вероятности, теорией информации. Она не ограничивается теорией кодов, исправляющих ошибки. Важным её разделом является, например, алфавитное кодирование. Оно применяется не с целью коррекции ошибок, а, например, для сжатия текста[37)], а в старые времена применялось и для шифрования[38)]. Так называемый генетический код[39)] тоже можно рассматривать как схему алфавитного кодирования. Задачи 32, 33, 34 относятся как раз к этой области.

Задача 32 (ММО 1962, второй тур, 9.5). Даны 2^n конечных последовательностей из нулей и единиц, причём ни одна из них не является началом никакой другой. Докажите, что сумма длин этих последовательностей не меньше $n \cdot 2^n$.

Задача 33. В генеалогическом древе князя Рюрика ни у кого из его потомков не было больше k сыновей. На годовщину рождения князя собрались вместе все n княжичей и вычислили суммарную длину своих ветвей генеалогического древа. Докажите, что она оказалась не меньше $n \log_k n$.

Задача 34 (студенческая олимпиада мехмата). В списке из n слов, составленных из букв k-буквенного алфавита, ни одно из которых не является началом другого, количество слов любой данной длины l, $l = 1, 2, \ldots, m$,

---

[37)] Вероятно, впервые для этой цели была применена азбука Морзе, которую можно рассматривать как схему алфавитного кодирования, впрочем, не префиксную и не обладающую свойством однозначности декодирования.

[38)] Например, шифры простой замены, шифр Полибия, шифр Бэкона можно рассматривать как схемы алфавитного кодирования.

[39)] Сопоставление аминокислотам троек символов из алфавита А, Г, Ц, Т.

равно n_l. Докажите, что $n_1/k + \ldots + n_m/k^m \leqslant 1$. Неравенство может обращаться в равенство[40].

В задаче 32 на самом деле речь идёт о средней длине элементарного кодового слова при кодировании n-буквенного алфавита двухбуквенным. Процедура кодирования слов в данном n-буквенном алфавите, например двоичными словами, заключается в замене каждой буквы этого алфавита двоичным кодовым словом. Если ни одно кодовое слово не является началом другого, то такое кодирование называется *префиксным*. Это условие на кодовые слова позволяет легко выполнять декодирование[41]. В задаче 32 на самом деле ищется префиксный код с минимальной средней длиной элементарного кодового слова.

Эту задачу можно решить в более общей постановке — когда буквам данного алфавита сопоставляются такие числа $p_i > 0$, что $p_1 + \ldots + p_n = 1$ (p_i — это вероятности появления этих букв), а для кодирования используется k-ичный алфавит. Тогда средняя длина элементарного кодового слова будет равна $l_c = p_1 l_1 + \ldots + p_n l_n$, где l_i — длина кодового слова. Для неё Шеннон получил следующую оценку:

Теорема 6. *Справедливо неравенство $l_c \geqslant H_k(p_1, \ldots, p_n)$, где*

$$H_k(p_1, \ldots, p_n) = -(p_1 \log_k p_1 + \ldots + p_n \log_k p_n)$$

— энтропия Шеннона.

Доказательство появится чуть позже, а вначале укажем связь между префиксными кодами в k-ичном алфавите и корневыми k-арными деревьями[42]. Так называются деревья с ориентированными рёбрами, в которых из корня (единственной вершины нулевого яруса) и любой внутренней вершины i-го яруса выходит не более k рёбер, направленных в вершины $(i+1)$-го яруса.

Индукцией по номеру яруса легко доказать, что число вершин на i-м ярусе не превосходит k^i. По определению, в каждую вершину, кроме корня, входит ровно одно ребро. Вершины, из которых не выходит ни одного ребра, назовём листьями[43]. Для каждого листа существует единственный

[40] Эта задача предлагалась на олимпиаде в 1980-е годы, когда ещё не было ни курса дискретной математики, ни учебника [17].

[41] Разумеется, вместо префиксных кодов можно использовать постфиксные, т. е. такие, в которых ни одно элементарное кодовое слово не является концом другого. Менее очевидно, что существуют и не префиксные и не постфиксные коды, которые тоже обладают свойством однозначности декодирования.

[42] О которых идёт речь в задаче 33.

[43] Используется также термин «висячие вершины».

путь, ведущий в него из корня. Число рёбер в нём обозначим l_i, где i — номер листа (в произвольной нумерации). Если для каждой внутренней вершины сопоставить выходящим из неё рёбрам различные числа (метки) из множества (алфавита) $A_k = \{0, \ldots, k-1\}$, то пути из корня в лист i можно сопоставить слово w_i в алфавите A_k длины l_i, которое получится, если выписать все метки рёбер этого пути, начиная с ребра, выходящего из корня. Совокупность полученных слов w_i образует набор элементарных кодовых слов схемы алфавитного кодирования, определяемой по данному дереву. Очевидно, что эта схема является префиксной.

Верно и обратное, а именно, по любой префиксной схеме алфавитного кодирования можно построить корневое дерево, которому будет указанным выше образом сопоставляться как раз упомянутая префиксная схема алфавитного кодирования. Аккуратное доказательство этого утверждения проводится по индукции и оставляется читателю в качестве ещё одной задачи. В силу указанной связи между префиксными кодами и деревьями, утверждения про префиксные коды можно переформулировать как утверждения про деревья. Поэтому из теоремы 6 следуют оценки не только задачи 32, но и задачи 33 (которую можно рассматривать как обобщение задачи 32). А задачу 34 можно считать следствием теоремы 7:

Теорема 7 (неравенство Крафта[44]). *Пусть в корневом k-арном дереве длины всех путей от корня к листьям равны l_i, $i = 1, \ldots, n$. Тогда справедливо неравенство $k^{-l_1} + \ldots + k^{-l_n} \leqslant 1$.*

Доказательство. Представим, что из каждого листа i, как из корня, вырастает ветка, имеющая вид полного k-арного дерева высоты $l - l_i$, где $l = \max_i l_i$ (назовём k-арное дерево *полным высоты* h, если из каждой внутренней вершины выходит ровно k рёбер, а листья образуют h-й ярус). Тогда на этой ветке будет k^{l-l_i} листьев, и все листья «расцветшего» дерева образуют в нём l-й ярус. Всего листьев будет $k^{l-l_1} + \ldots + k^{l-l_n} \leqslant k^l$, так как в l-м ярусе не может быть больше k^l листьев. Остаётся поделить обе части неравенства на k^l. □

Теперь можно доказать теорему 6. Воспользуемся тем, что для функции $\log_k x$ при подходящем[45] значении K справедливо неравенство $\log_k x \leqslant K(x-1)$, имеющее простой геометрический смысл: график функции (за исключением точки касания) лежит ниже касательной к нему, проведённой через точку графика $(x=1, y=\log_k 1=0)$. Подставляя в неравенство

[44] То же самое неравенство для произвольного кода с однозначным декодированием доказывается более сложно и называется неравенством Макмиллана.

[45] Точная формула для K далее не существенна.

$x = k^{-l_i}/p_i$, получаем, что

$$-l_i + \log_k\Bigl(\frac{1}{p_i}\Bigr) \leqslant K\Bigl(\frac{k^{-l_i}}{p_i} - 1\Bigr),$$

откуда

$$l_i \geqslant K\Bigl(1 - \frac{k^{-l_i}}{p_i}\Bigr) + \log_k\Bigl(\frac{1}{p_i}\Bigr).$$

Далее,

$$p_i l_i \geqslant K(p_i - k^{-l_i}) + p_i \log_k\Bigl(\frac{1}{p_i}\Bigr).$$

Складывая эти неравенства, получаем

$$\begin{aligned} &p_1 l_1 + \ldots + p_n l_n \geqslant \\ &\geqslant K(p_1 + \ldots + p_n - k^{-l_1} - \ldots - k^{-l_n}) + p_1 \log_k\Bigl(\frac{1}{p_1}\Bigr) + \ldots + p_n \log_k\Bigl(\frac{1}{p_n}\Bigr) = \\ &\qquad = K(1 - k^{-l_1} - \ldots - k^{-l_n}) + H_k(p_1, \ldots, p_n) \geqslant H_k(p_1, \ldots, p_n) \end{aligned}$$

согласно теореме 7. Таким образом, теорема 6 доказана.

Благодарности

Автор благодарен А. В. Устинову и рецензенту редколлегии за полезные советы.

Список литературы

[1] *Берлекэмп Э. Р.* Алгебраическая теория кодирования. М.: Мир, 1971.

[2] *Васильев Н. Б., Егоров А. А.* Задачи Всесоюзных математических олимпиад. М.: Наука, 1988.

[3] *Гальперин Г. А., Толпыго А. К.* Московские математические олимпиады. М.: Просвещение, 1986.

[4] *Гашков С. Б.* Разностные множества, конечные геометрии, матрицы Заранкевича и экстремальные графы // Математическое просвещение. Сер. 3. Вып. 21. М.: МЦНМО, 2017. С. 145–185.

[5] *Гашков С. Б.* Графы-расширители и их применения в теории кодирования // Математическое просвещение. Сер. 3. Вып. 13. М.: МЦНМО, 2009. С. 104–126.

[6] *Левенштейн В. И.* Элементы теории кодирования // Дискретная математика и математическая кибернетика. Т. 1. М.: Наука, 1974.

[7] *Мак-Вильямс Ф. Дж., Слоэн Н. Дж. А.* Теория кодов, исправляющих ошибки. М.: Связь, 1979.

[8] *Прасолов В. В. и др.* Московские математические олимпиады 1935–1957 гг. М.: МЦНМО, 2010.

[9] *Прасолов В. В. и др.* Московские математические олимпиады 1958–1967 гг. М.: МЦНМО, 2013.

[10] *Бегунц А. В. и др.* Московские математические олимпиады 1981–1992 гг. М.: МЦНМО, 2017.

[11] *Фёдоров Р. М. и др.* Московские математические олимпиады 1993–2005 гг. / 3-е изд. М.: МЦНМО, 2017.

[12] *Садовничий В. А., Григорьян А. А., Конягин С. В.* Задачи студенческих математических олимпиад. М.: МГУ, 1987.

[13] *Сидельников В. М.* Теория кодирования. М.: Физматлит, 2008.

[14] *Таранников Ю. В.* Комбинаторные свойства дискретных структур и приложения к криптологии. М.: МЦНМО, 2011.

[15] *Фейеш Тот Л.* Расположения на плоскости, на сфере и в пространстве. М.: Физматгиз, 1958.

[16] *Холл М.* Комбинаторика. М.: Мир, 1970.

[17] *Чашкин А. В.* Дискретная математика. М.: Академия, 2012.

[18] *Guruswami V., Sudan M.* Improved decoding of Reed — Solomon and algebraic-geometric codes // IEEE Trans. Inform. Theory. 1999. Vol. 45. P. 1757–1767.

[19] *Winkler P.* Mathematical puzzles: a connoisseur's collection. Natick, USA: Taylor and Francis Inc., 2004.

Сергей Борисович Гашков, мехмат МГУ

`sbgashkov@gmail.com`

О вычислении конечных тригонометрических сумм

Н. Н. Осипов

Введение

О явном вычислении (или, как ещё говорят, о вычислении в замкнутой форме) конечных *тригонометрических сумм* написано довольно много работ (см., например, статьи [18–23, 26, 27], а также библиографию к ним). Цель настоящей статьи — познакомить читателя с наиболее простыми методами нахождения таких сумм.

Зачем нужно точно вычислять тригонометрические суммы или зачем могут потребоваться тригонометрические тождества с суммами? Ведь для большинства случаев точные формулы не нужны, а достаточно каких-то оценок и неравенств. Это, конечно, верно, но иногда, для получения особо точных оценок, тождества могут сильно помочь.

Например, классические теоремы Чебышёва и Коркина — Золотарёва об *экстремальных многочленах* (алгебраических многочленах, наименее уклоняющихся от нуля на данном отрезке в C-норме и L_1-норме соответственно) легко доказываются, если воспользоваться тригонометрическими тождествами

$$\sum_{j=0}^{2n-1}(-1)^j\cos^k\frac{\pi j}{n}=\begin{cases}0, & \text{если } 0\leqslant k<n,\\ \dfrac{n}{2^{n-2}}, & \text{если } k=n\end{cases}$$

(см., например, [2] или, в более элементарном изложении, [5] в случае теоремы Чебышёва). В статье [6] с помощью разнообразных тригонометрических тождеств выводится другой классический результат — *неравенство Бернштейна* для производной (тригонометрического, а затем и алгебраического) многочлена.

Типичные элементарные задачи, приводящие к конечным тригонометрическим суммам, обычно связаны с геометрией правильного многоугольника. Их можно решать с помощью геометрических рассуждений (и обычной алгебры), но есть и более алгоритмический подход, использующий *алгебру комплексных чисел* (см., например, § 10 в книге [10]). Вот простейший пример такой задачи: найти сумму всех векторов, исходящих из центра правильного N-угольника в его вершины. Эта задача сводится к вычислению тригонометрической суммы

$$\sum_{j=0}^{N-1} \zeta^j.$$

Здесь и далее ζ — фиксированный *первообразный корень* N-й степени из единицы, например

$$\zeta = \cos\frac{2\pi}{N} + i\sin\frac{2\pi}{N} = \exp\frac{2\pi i}{N}.$$

Далее мы будем рассматривать в основном *стандартные тригонометрические суммы* — суммы вида

$$S(N) = \sum_{j=0}^{N-1} R(\zeta^j), \tag{0.1}$$

где $R(z)$ — некоторая *рациональная функция*. Именно к таким суммам приводят наиболее содержательные геометрические задачи про правильный N-угольник.

При вычислении стандартных тригонометрических сумм очень эффективными оказываются методы *теории функций комплексного переменного*: разложение в степенные ряды (Тейлора или Лорана) и теория вычетов. Всё это содержится в любом учебнике по ТФКП (см., например, [12]), и мы будем предполагать, что читатель знаком с соответствующей техникой.

В теореме 1 (см. § 1) предлагается простой регулярный способ вычисления сумм вида (0.1). Все удобства (как и недостатки) этого способа читатель сможет оценить только в процессе решения конкретных задач — с этой целью в § 2 приводится небольшая коллекция задач, а в § 3 даны решения, указания и ответы к ним. Самыми яркими экземплярами в данной коллекции являются, несомненно, задача 2.12 о тождествах М. Рисса и задача 2.13 о тождестве Эйзенштейна. Тождества Рисса играют важную роль при доказательстве упомянутого выше неравенства Бернштейна (подробности см. в [6]). Тождество Эйзенштейна — это один из примеров так называемых *теорем взаимности* для тригонометрических сумм. Подробный рассказ об общих теоремах взаимности выходит за рамки этой статьи (их доказательство обычно требует более тонкой и менее элементарной техники, образцы которой читатель может найти в статьях [18] и [22]).

Элементарный подход к вычислению стандартных тригонометрических сумм основан на *теории многочленов*: здесь главным инструментом является теорема о разложении правильной рациональной дроби в сумму простейших рациональных дробей (см., например, главу 3 в учебнике [4]). В принципе, такой подход может быть доступен даже школьникам (в первую очередь, конечно, ученикам математических классов).

Для вычисления стандартных тригонометрических произведений, аналогичных (0.1), также достаточно теории многочленов, но мы только слегка затронем эту тему (см. теорему 2 в § 1 и задачи 1.2, 2.15–2.17).

В завершающем § 4 мы рассмотрим тригонометрические суммы нескольких специальных типов, но ограничимся лишь отдельными примерами и описанием общего подхода к вычислению.

Значения сумм первых двух типов (см. разделы 4.1 и 4.2) выражаются в терминах различных *арифметических функций*, наиболее известными примерами которых являются функция Эйлера $\varphi(N)$ и функция Мёбиуса $\mu(N)$. Техника вычисления широко использует аппарат теории *мультипликативных функций* (в частности, *свёртку Дирихле* и *формулу обращения Мёбиуса*). В подробном изложении основные факты об арифметических (в частности, о мультипликативных) функциях читатель может найти в соответствующих главах учебников [9] и [15]. Быстро навести справки можно, например, по ссылке [29].

Третий тип специальных тригонометрических сумм — это суммы, аналогичные *квадратичным суммам Гаусса* (см. раздел 4.3). Несмотря на полностью элементарную конструкцию, значения таких сумм иногда могут иметь весьма нетривиальный смысл в рамках *теории дивизоров* мнимых квадратичных полей.

§ 1. Немного теории

Мы начнём с одной геометрической задачи для школьников, решение которой непосредственно сводится к вычислению некоторой тригонометрической суммы, и заодно продемонстрируем один элементарный подход к нахождению таких сумм.

Задача 1.1. На окружности единичного радиуса равномерно расположены N точек, где N нечётно. Одну из этих точек соединили отрезками с остальными и затем подсчитали сумму величин, обратных квадратам расстояний от центра окружности до проведённых отрезков. Докажите, что получилось $N^2 - 1$.

Решение. Пусть $A_0, A_1, \ldots, A_{N-1}$ — данные точки, причём A_0 — та точка, которую соединяли со всеми остальными. Легко видеть, что указанные расстояния суть $\cos(\alpha_j/2)$, где α_j — угол, под которым видна хорда

A_0A_j из центра окружности. Ясно, что

$$\alpha_j = \begin{cases} \dfrac{2\pi j}{N}, & \text{если } j < \dfrac{N}{2}, \\ \dfrac{2\pi(N-j)}{N}, & \text{если } j > \dfrac{N}{2}. \end{cases}$$

Таким образом, задача сводится к вычислению суммы

$$S(N) = \sum_{j=0}^{N-1} \frac{1}{\cos^2 \dfrac{\pi j}{N}} \tag{1.1}$$

(для удобства суммирования мы добавили единицу), а точнее, к доказательству равенства $S(N) = N^2$. Точки A_j можно интерпретировать как комплексные числа ζ^j, где $j = 0, 1, \ldots, N-1$, а сумму $S(N)$ — записать в виде

$$S(N) = \sum_{j=0}^{N-1} \frac{4\zeta^j}{(\zeta^j+1)^2}.$$

Чтобы вычислить $S(N)$, заметим, что

$$S(N) = \lim_{r\to 1} \sum_{j=0}^{N-1} \frac{4r\zeta^j}{(r\zeta^j+1)^2}.$$

Зачем понадобилось переходить к пределу? Дело в том, что при $|r| < 1$ можно (а при $r = 1$ — нельзя) написать следующую цепочку равенств:

$$\sum_{j=0}^{N-1} \frac{4r\zeta^j}{(r\zeta^j+1)^2} = 4\sum_{j=0}^{N-1}\sum_{k=0}^{\infty}(-1)^{k+1}k(r\zeta^j)^k = 4\sum_{k=0}^{\infty}(-1)^{k+1}kr^k\sum_{j=0}^{N-1}\zeta^{kj} ==$$

$$= 4N^2\sum_{l=0}^{\infty}(-1)^{l+1}lr^{Nl} = 4N^2\frac{r^N}{(r^N+1)^2}.$$

Здесь мы воспользовались легко проверяемым равенством

$$\sum_{j=0}^{N-1}\zeta^{kj} = \begin{cases} 0, & \text{если } k \not\equiv 0 \pmod N, \\ N, & \text{если } k \equiv 0 \pmod N, \end{cases} \tag{1.2}$$

а также разложением в ряд

$$\frac{z}{(z+1)^2} = \sum_{k=0}^{\infty}(-1)^{k+1}kz^k,$$

имеющим место при $|z| < 1$. Теперь, поскольку

$$\lim_{r\to 1}\left(4N^2\frac{r^N}{(r^N+1)^2}\right) = N^2,$$

требуемое соотношение доказано. $\square$

Как вариант, можно было бы вычислить сумму

$$\sum_{j=0}^{N-1}\frac{4\zeta^j}{(z-\zeta^j)^2} = -4\sum_{j=0}^{N-1}\frac{1}{z-\zeta^j} + 4z\sum_{j=0}^{N-1}\frac{1}{(z-\zeta^j)^2} \tag{1.3}$$

и затем в найденное выражение подставить $z = -1$. Первая сумма в правой части равенства (1.3) есть *логарифмическая производная* $P'(z)/P(z)$ многочлена

$$P(z) = \prod_{j=0}^{N-1}(z-\zeta^j) = z^N - 1.$$

Таким образом,

$$\sum_{j=0}^{N-1}\frac{1}{z-\zeta^j} = \frac{Nz^{N-1}}{z^N-1}.$$

Вторая сумма находится дифференцированием по z полученного равенства:

$$\sum_{j=0}^{N-1}\frac{1}{(z-\zeta^j)^2} = \frac{Nz^{N-2}(N+z^N-1)}{(z^N-1)^2}.$$

В итоге имеем

$$\sum_{j=0}^{N-1}\frac{4\zeta^j}{(z-\zeta^j)^2} = \frac{4N^2z^{N-1}}{(z^N-1)^2}.$$

Подставив сюда $z = -1$, получим

$$\sum_{j=0}^{N-1}\frac{4\zeta^j}{(1+\zeta^j)^2} = N^2.$$

Попутно можно также обнаружить следующие равенства:

$$\sum_{j=0}^{N-1}\frac{1}{1+\zeta^j} = \frac{N}{2}, \quad \sum_{j=0}^{N-1}\frac{1}{(1+\zeta^j)^2} = \frac{-N^2+2N}{4}.$$

Действуя в том же духе, нетрудно продолжить этот ряд равенств:

$$\sum_{j=0}^{N-1}\frac{1}{(1+\zeta^j)^3} = \frac{-3N^2+4N}{8}, \quad \sum_{j=0}^{N-1}\frac{1}{(1+\zeta^j)^4} = \frac{N^4-22N^2+24N}{48} \quad \text{и т. д.}$$

Этот способ вычисления суммы (1.1) отличается от предыдущего тем, что полностью элементарен (обходится без разложений в степенные ряды), но нам повезло — многочлен $P(z)$ оказался на самом деле «малочленом», и потому с ним удобно вычислять. Вообще, если аналог многочлена $P(z)$ допускает «разумную» форму записи, то такой способ вычисления соответствующих тригонометрических сумм будет работать (см. ниже раздел 4.1). С другой стороны, чтобы получить эту «разумную» форму записи, иногда могут потребоваться разложения в ряды (как, например, в разделе 4.3; см. также решение задачи 2.14).

Естественное желание обобщить (т. е. фактически предложить *алгоритм вычисления* тригонометрических сумм типа (1.1)) рано или поздно приведёт к утверждению наподобие следующей теоремы.

Теорема 1. *Пусть* $\zeta = \exp(2\pi i/N)$, $R(z)$ — *рациональная функция, множество полюсов* P_R *которой не содержит точек* ζ^j $(j=0,1,\ldots,N-1)$. *Тогда*

$$\frac{1}{N}\sum_{j=0}^{N-1} R(\zeta^j) = \sum_{0\neq z_*\in P_R} \operatorname*{res}_{z=z_*} F(z) + \operatorname*{res}_{z=0} F(z) + \operatorname*{res}_{z=\infty} F(z), \tag{1.4}$$

где

$$F(z) = \frac{R(z)}{z(1-z^N)}.$$

Доказательство получается в результате применения *теоремы о полной сумме вычетов* к рациональной функции $F(z)$. В качестве упражнения читателю рекомендуется аккуратно вычислить вычеты $F(z)$ в точках $z=\zeta^j$ (см. ниже (1.5)). □

Теорема 1 является более элементарным вариантом способа вычисления тригонометрических сумм с помощью вычетов, предлагаемого в задаче 30.04 из сборника задач [8]. Напомним общее правило вычисления вычетов, которого обычно достаточно для практических приложений: если z_* — полюс функции $F(z)$ кратности $k \geqslant 1$, то

$$\operatorname*{res}_{z=z_*} F(z) = \lim_{z\to z_*} \frac{G^{(k-1)}(z)}{(k-1)!},$$

где $G(z) = (z-z_*)^k F(z)$. В частности, при $z_* = \zeta^j$ и функции $F(z)$, определённой выше, получим

$$\operatorname*{res}_{z=z_*} F(z) = -\frac{R(\zeta^j)}{N}. \tag{1.5}$$

Кроме того, если $R(z)=f(z)/g(z)$ и $z_*\neq 0$ — *простой* полюс $R(z)$, $z_*^N \neq 1$, то

$$\operatorname*{res}_{z=z_*} F(z) = \frac{f(z_*)}{g'(z_*)z_*(1-z_*^N)}.$$

С помощью теоремы 1 сумма (1.1) вычисляется очень легко: имеем

$$R(z) = \frac{4z}{(z+1)^2}, \quad F(z) = \frac{4}{(z+1)^2(1-z^N)},$$

поэтому

$$\operatorname*{res}_{z=0} F(z) = \operatorname*{res}_{z=\infty} F(z) = 0, \quad \operatorname*{res}_{z=-1} F(z) = N$$

и, таким образом, $S(N) = N^2$.

Теорему 1 легко модифицировать на случай, когда среди полюсов $R(z)$ присутствуют некоторые ζ^j, т. е. корни из единицы N-й степени. А именно, из суммы в левой части равенства (1.4) нужно всего лишь удалить не имеющие смысла слагаемые, соответствующие этим ζ^j. В качестве упражнения по вычислению вычетов читателю предлагается показать, что при чётном N сумма (1.1) без слагаемого с номером $j = N/2$ равна $(N^2-1)/3$.

Попутно отметим, что нахождение аналогичных (0.1) *тригонометрических произведений* оказывается существенно проще. Поскольку рациональная функция есть отношение многочленов, можно ограничиться рассмотрением случая, когда $R(z)$ — многочлен. Как показывает следующая теорема, здесь достаточно обычной элементарной алгебры.

Теорема 2. *Пусть $\zeta = \exp(2\pi i/N)$, $R(z)$ — многочлен степени k со старшим коэффициентом 1 и корнями $\theta_1, \ldots, \theta_k$. Тогда*

$$\prod_{j=1}^{N-1} R(\zeta^j) = (-1)^{k(N-1)} \prod_{l=1}^{k} \frac{\theta_l^N - 1}{\theta_l - 1}. \tag{1.6}$$

Если некоторые из корней θ_l равны 1, *то соответствующие сомножители в правой части равенства* (1.6) *следует считать равными N.*

Доказательство получается с помощью разложений

$$R(z) = \prod_{l=1}^{k}(z-\theta_l), \quad \frac{z^N-1}{z-1} = (-1)^{N-1}\prod_{j=1}^{N-1}(\zeta^j - z)$$

и непосредственного перемножения. □

В качестве иллюстрации рассмотрим один важный пример, который пригодится нам далее (см. раздел 4.3).

Задача 1.2. Пусть N нечётно, $N = 4m + r$, где $r \in \{1, 3\}$. Докажите, что

$$\prod_{l=1}^{(N-1)/2} 2\sin\frac{2\pi(2l-1)}{N} = (-1)^m\sqrt{N}. \tag{1.7}$$

Решение. К произведению P в левой части равенства (1.7) теорема 2 неприменима, но с её помощью можно найти P^2. Заметим, что

$$P = (-1)^{(N-1)/2} \prod_{l=1}^{(N-1)/2} 2\sin\frac{2\pi(N-(2l-1))}{N}.$$

Когда l меняется от 1 до $(N-1)/2$, переменные $2l-1$ и $N-(2l-1)$ вместе пробегают все значения от 1 до $N-1$, поэтому

$$P^2 = (-1)^{(N-1)/2} \prod_{l=1}^{(N-1)/2} 2\sin\frac{2\pi(N-(2l-1))}{N} \prod_{l=1}^{(N-1)/2} 2\sin\frac{2\pi(2l-1)}{N} =$$

$$= (-1)^{(N-1)/2} \prod_{j=1}^{N-1} 2\sin\frac{2\pi j}{N} = (-1)^{(N-1)/2} \prod_{j=1}^{N-1} \frac{\zeta^j - \zeta^{-j}}{i} = \prod_{j=1}^{N-1} (\zeta^{2j} - 1),$$

где $\zeta = \exp(2\pi i/N)$. Последнее произведение вычисляется с помощью теоремы 2 и оказывается равным N. Итак, $P = \pm\sqrt{N}$. Осталось найти знак P. Он определяется числом тех l от 1 до $(N-1)/2$, для которых

$$\sin\frac{2\pi(2l-1)}{N} < 0.$$

Это число может быть выражено как

$$\left[\frac{N-1}{2}\right] - \left[\frac{N+2}{4}\right] = m + \left[\frac{r-1}{2}\right] - \left[\frac{r+2}{4}\right]$$

(квадратные скобки обозначают взятие целой части) и в каждом из случаев $r = 1$ и $r = 3$ оказывается равным m. Отсюда следует (1.7). □

Дальнейшее развитие (в элементарном русле) сюжета о тригонометрических произведениях читатель может найти, например, в статье [20].

§ 2. Подборка задач

Предлагаемые ниже задачи так или иначе связаны с вычислением конечных тригонометрических сумм. Основной инструмент для решения задач — это, конечно, теорема 1, хотя иногда возможны и другие варианты.

2.1. Корни из единицы

Следующие две задачи предназначены исключительно для разминки. Первая задача — ещё одно упражнение в применении фундаментального равенства (1.2). Вторая задача дана скорее в довесок к первой, но может быть интересна и сама по себе.

Задача 2.1. Пусть $z_1, \ldots, z_m$ — некоторый набор комплексных корней степени N из единицы, т. е. $z_l^N = 1$ ($l = 1, \ldots, m$). Предположим, что для каждого $j = 1, \ldots, N-1$ выполняется равенство

$$\sum_{l=1}^{m} z_l^j = 0.$$

Докажите, что m делится на N, а заданный набор состоит из m/N раз повторяющихся чисел $1, \zeta, \ldots, \zeta^{N-1}$, где $\zeta = \exp(2\pi i/N)$.

Задача 2.2[1]. Дан многочлен с целыми коэффициентами и старшим коэффициентом 1. Известно, что все его комплексные корни по модулю равны 1. Докажите, что все корни многочлена суть корни из 1.

2.2. Правильный N-угольник

Пусть правильный N-угольник $A_0A_1 \ldots A_{N-1}$ ($N > 1$) вписан в окружность радиуса 1. В следующей серии задач рассматриваются суммы

$$S_k^{(1)}(P) = \sum_{j=0}^{N-1} |PA_j|^k, \quad S_k^{(2)}(P) = \sum_{j=0}^{N-1} (-1)^j |PA_j|^k, \quad S_k^{(3)}(P) = \sum_{j=0}^{N-1} h_j(P)^k,$$

где точка P принадлежит данной окружности, $h_j(P)$ — расстояние от центра окружности до прямой PA_j.

Задача 2.3[2]. Докажите, что:

а) если k чётно и $0 \leqslant k < 2N$, то $S_k^{(1)}(P)$ не зависит от P;

б) если $k \equiv N \pmod 2$, $0 \leqslant k < N$ и P лежит на дуге $A_{N-1}A_0$, то $S_k^{(2)}(P) = 0$;

в) $S_k^{(3)}(P) = 2^{-k} S_k^{(1)}(-P)$.

Задача 2.4. Вычислите:

а) $S_{2N}^{(1)}(P)$; б) $S_N^{(2)}(P)$, если P лежит на дуге $A_{N-1}A_0$.

Задача 2.5. Докажите, что если P лежит на дуге $A_{N-1}A_0$, то

$$\sum_{j=0}^{N-2} \frac{1}{|PA_j| \cdot |PA_{j+1}|} = \frac{1}{|PA_0| \cdot |PA_{N-1}|}.$$

Задача 2.6. Вычислите:

а) $S_{-2}^{(1)}(P)$; б) $S_{-1}^{(2)}(P)$, если P лежит на дуге $A_{N-1}A_0$ и N нечётно.

---

[1] XXII Российский фестиваль юных математиков, 2011 г.

[2] Случай $k = 1$ п. б) хорошо известен (см., например, задачу 609(а) в книге [14]).

2.3. Тригонометрические суммы: техника вычисления

Здесь предлагаются технически более сложные задачи, поэтому можно не стесняться в выборе средств для их решения. Разумеется, мы рекомендуем в первую очередь попробовать теорему 1.

Задача 2.7. В условиях теоремы 1 дайте элементарный способ вычисления суммы $\sum_{j=0}^{N-1} R(\zeta^j)$, основанный на разложении рациональной функции $R(z)$ в сумму простейших дробей.

Задача 2.8. Пусть

$$S_k = \sum_{j=0}^{N-1} \cos^k \frac{2\pi j}{N}.$$

а) Докажите, что если k нечётно и $1 \leqslant k < N$, то $S_k = 0$.
б) Докажите, что если N нечётно, то $S_{-1} = (-1)^{(N-1)/2} N$.

Задача 2.9. При нечётном N вычислите сумму

$$\sum_{j=0}^{N-1} \operatorname{tg}\left(\theta + \frac{\pi j}{N}\right),$$

где число θ таково, что все слагаемые имеют смысл.

Задача 2.10. Докажите, что

$$\sum_{j=1}^{N-1} \operatorname{ctg}^2 \frac{\pi j}{N} = \frac{(N-1)(N-2)}{3}.$$

Задача 2.11[3]. Пусть $0 \leqslant k \leqslant N$. Докажите, что

$$\sum_{j=0}^{N-1} \frac{\sin \dfrac{\pi k(4j+1)}{2N}}{1 - \cos \dfrac{\pi(4j+1)}{2N}} = kN.$$

Задача 2.12[4]. Пусть $0 \leqslant k \leqslant n$. Докажите *тождества Рисса*:

$$\sum_{j=0}^{2n-1} (-1)^{j+1} \frac{\cos \dfrac{\pi k(2j-1)}{2n}}{\sin^2 \dfrac{\pi(2j-1)}{4n}} = 0, \qquad \sum_{j=0}^{2n-1} (-1)^{j+1} \frac{\sin \dfrac{\pi k(2j-1)}{2n}}{\sin^2 \dfrac{\pi(2j-1)}{4n}} = 4kn.$$

[3] IV Соросовская олимпиада (I тур, задача 10 для 11 класса).
[4] См. [6, задача 35].

Задача 2.13[5]. Докажите *тождество Эйзенштейна*:

$$l\sum_{j=1}^{k-1}\frac{\operatorname{tg}\dfrac{\pi lj}{k}}{\operatorname{tg}\dfrac{2\pi j}{k}}+k\sum_{j=1}^{l-1}\frac{\operatorname{tg}\dfrac{\pi kj}{l}}{\operatorname{tg}\dfrac{2\pi j}{l}}=-\frac{(k-l)^2}{2},$$

где k, l — нечётные взаимно простые натуральные числа.

Задача 2.14. Докажите, что при нечётном N

$$\sum_{j=0}^{N-1} j\operatorname{tg}\frac{\pi j}{N}=\frac{N}{2}\sum_{j=1}^{N-1}(-1)^j\operatorname{ctg}\frac{\pi j}{N}.$$

2.4. Тригонометрические произведения

Для решения следующих задач читателю предлагается применять теорему 2 как наиболее эффективное средство (хотя, конечно, возможны и другие способы решения).

Задача 2.15. Докажите, что

$$\prod_{j=1}^{N-1}\sin\frac{\pi j}{N}=\frac{N}{2^{N-1}}.$$

Задача 2.16. При нечётном N докажите, что

$$\prod_{j=1}^{N-1}\operatorname{tg}\frac{\pi j}{N}=(-1)^{(N-1)/2}N.$$

Задача 2.17. Докажите, что

$$\prod_{j=1}^{N-1}\Bigl(5-4\cos\frac{2\pi j}{N}\Bigr)=(2^N-1)^2.$$

§ 3. Решения, указания и ответы

2.1. Положим $z_l=\zeta^{k_l}$, $k_l\in\{0,1,\ldots,N-1\}$. При фиксированном $k\in\{0,1,\ldots,N-1\}$ рассмотрим новый набор чисел $w_1,\ldots,w_m$, где $w_l=\zeta^{-k}z_l=\zeta^{k_l-k}$. Ясно, что для каждого $j=1,\ldots,N-1$

$$\sum_{l=1}^{m} w_l^j=\zeta^{-kj}\sum_{l=1}^{m} z_l^j=0.$$

[5] Цит. по статье [18].

Суммируя эти равенства, получим

$$\sum_{j=1}^{N-1}\sum_{l=1}^{m}\zeta^{(k_l-k)j}=\sum_{l=1}^{m}\sum_{j=1}^{N-1}\zeta^{(k_l-k)j}=0,\quad \text{или}\quad \sum_{l=1}^{m}\sum_{j=0}^{N-1}\zeta^{(k_l-k)j}=m.$$

Из (1.2) следует, что

$$\sum_{j=0}^{N-1}\zeta^{(k_l-k)j}=\begin{cases} N, & \text{если } k_l=k,\\ 0, & \text{если } k_l\neq k.\end{cases}$$

Значит, количество k_l, равных k, равно m/N и тем самым одинаково для всех k.

2.2. Пусть $z_1, \ldots, z_n$ — эти корни. Для каждого $j\in\mathbb{N}$ рассмотрим точку

$$P^{(j)}=\bigl(s_1(z_1^j,\ldots,z_n^j),\ldots,s_n(z_1^j,\ldots,z_n^j)\bigr),$$

где $s_k(x_1,\ldots,x_n)$ — *элементарные симметрические многочлены*. Как следует из *основной теоремы о симметрических многочленах*, $P^{(j)}\in\mathbb{Z}^n$ для любого j. Кроме того, из условия следует, что k-я координата всех точек $P^{(j)}$ ограничена биномиальным коэффициентом C_n^k $(k=1,\ldots,n)$. Значит, $P^{(j)}=P^{(l)}$ для некоторых $j\neq l$, откуда

$$(z_1^j,\ldots,z_n^j)=(z_{\alpha_1}^l,\ldots,z_{\alpha_n}^l)$$

для некоторой перестановки $\alpha=(\alpha_1,\ldots,\alpha_n)$. Если эта перестановка является тождественной, то утверждение задачи очевидно. Если она циклическая, скажем $\alpha=(2,3,\ldots,n,1)$, то

$$j\arg z_k\equiv l\arg z_{k+1}\pmod{2\pi},\qquad k=1,\ldots,n,$$

где $z_{n+1}=z_1$. Отсюда видно, что все $\arg z_k$ соизмеримы с π, и утверждение задачи верно. Осталось сослаться на тот факт, что всякую перестановку можно разложить в произведение циклических.

2.3. Пусть $A_j=\zeta^j$, где $\zeta=\exp(2\pi i/N)$. Точку P будем считать комплексным числом с условием $|P|=1$. Пусть также $\xi=\exp(\pi i/N)$, так что $\zeta=\xi^2$ и $\xi^N=-1$.

а) Обозначим $m=k/2$. Из условия $|P|=1$ следует, что

$$|PA_j|^2=|P-\zeta^j|^2=2-\frac{P}{\zeta^j}-\frac{\zeta^j}{P}=-\frac{(\zeta^j-P)^2}{P\zeta^j}.$$

Далее можно рассуждать двумя способами.

Первый способ. Имеем

$$S_k^{(1)}(P)=\sum_{j=0}^{N-1}|P-\zeta^j|^k=\sum_{j=0}^{N-1}\Big(2-\frac{P}{\zeta^j}-\frac{\zeta^j}{P}\Big)^m=$$
$$=\sum_{j=0}^{N-1}\sum_{l=0}^{m}(-1)^lC_m^l2^{m-l}\Big(\frac{P}{\zeta^j}+\frac{\zeta^j}{P}\Big)^l=\sum_{l=0}^{m}(-1)^lC_m^l2^{m-l}\sigma_l,$$

где введено обозначение

$$\sigma_l=\sum_{j=0}^{N-1}\Big(\frac{P}{\zeta^j}+\frac{\zeta^j}{P}\Big)^l.$$

Поскольку $0\leqslant l\leqslant m<N$, достаточно убедиться, что σ_l не зависит от P при $0\leqslant l<N$. В самом деле,

$$\sigma_l=\sum_{j=0}^{N-1}\sum_{t=0}^{l}C_l^tP^{2t-l}\zeta^{(l-2t)j}=\sum_{t=0}^{l}C_l^tP^{2t-l}\sum_{j=0}^{N-1}\zeta^{(l-2t)j}=\begin{cases}0, & l \text{ нечётно},\\ NC_l^{l/2}, & l \text{ чётно},\end{cases}$$

так как $|l-2t|\leqslant l<N$. Таким образом,

$$S_k^{(1)}(P)=N\sum_{t=0}^{[m/2]}2^{m-2t}C_m^{2t}C_{2t}^t.$$

Второй способ. Имеем

$$S_k^{(1)}(P)=\sum_{j=0}^{N-1}|P-\zeta^j|^k=\sum_{j=0}^{N-1}\frac{(-1)^m(\zeta^j-P)^{2m}}{P^m\zeta^{mj}}=\sum_{j=0}^{N-1}R(\zeta^j),$$

где

$$R(z)=\frac{(-1)^m(z-P)^{2m}}{P^mz^m}.$$

Далее воспользуемся теоремой 1. В данном случае

$$F(z)=\frac{(-1)^m(z-P)^{2m}}{P^mz^{m+1}(1-z^N)}.$$

Так как $m<N$, то $\operatorname{res}_{z=\infty}F(z)=0$. Кроме того,

$$\operatorname*{res}_{z=0}F(z)=\frac{(-1)^m}{P^m}\,c_m,$$

где c_m — коэффициент при z^m в разложении

$$\frac{(z-P)^{2m}}{1-z^N} = (z^{2m} + \ldots + (-1)^m C_{2m}^m P^m z^m + \ldots + P^{2m})(1 + z^N + z^{2N} + \ldots).$$

Из неравенства $m < N$ следует, что $c_m = (-1)^m C_{2m}^m P^m$. Значит,

$$\operatorname*{res}_{z=0} F(z) = C_{2m}^m.$$

Следовательно, $S_k^{(1)}(P) = NC_{2m}^m$.

Замечание. Второй способ приводит к ответу в замкнутой форме, зато первый способ позволяет получить формулу

$$S_k^{(1)}(P) = N \sum_{t=0}^{[m/2]} C_m^{2t} C_{2t}^t |P|^{2t}(|P|^2+1)^{m-2t},$$

где P произвольна (при $|P| = 1$ правая часть упрощается до NC_{2m}^m).

б) Пусть $m = (N-k)/2$. Для точки P дуги $A_{N-1}A_0$ имеем

$$|PA_j| = |P - \zeta^j| = \lambda(P) \cdot \frac{\zeta^j - P}{\xi^j}, \tag{3.1}$$

где $\lambda(P) = |1-P|/(1-P)$. Следовательно,

$$S_k^{(2)}(P) = \sum_{j=0}^{N-1} (-1)^j |P - \zeta^j|^k = \lambda^k(P) \sum_{j=0}^{N-1} \zeta^{mj}(\zeta^j - P)^k = \lambda^k(P) \sum_{j=0}^{N-1} R(\zeta^j), \tag{3.2}$$

где $R(z) = z^m(z-P)^k$. Далее воспользуемся теоремой 1. В данном случае

$$F(z) = \frac{z^{m-1}(z-P)^k}{1-z^N}.$$

Так как $m \geqslant 1$, то $\operatorname{res}_{z=0} F(z) = 0$, а так как $m + k - 1 = (N+k)/2 - 1 \leqslant N - 2$, то $\operatorname{res}_{z=\infty} F(z) = 0$. Таким образом, $S_k^{(2)}(P) = 0$.

в) Из геометрических соображений ясно, что $2h_j(P) = |P'A_j|$, где P' — точка, противоположная P: $P' = -P$.

2.4. *Указание.* а) См. решение п. а) задачи 2.3 в случае $k = 2N$. б) См. решение п. б) задачи 2.3 в случае $k = N$.

Ответ. а) $NC_{2N}^N + (-1)^N N(P^N + P^{-N})$; б) $N\lambda^N(P)(1 + (-1)^N P^N)$.

2.5. Опираясь на формулу (3.1), получим равенство

$$\sum_{j=0}^{N-2} \frac{1}{|PA_j| \cdot |PA_{j+1}|} = \lambda^{-2}(P) \sum_{j=0}^{N-2} \frac{\xi^{2j+1}}{(\zeta^j - P)(\zeta^{j+1} - P)}, \quad \text{где} \quad \lambda(P) = \frac{|1-P|}{1-P}.$$

Далее можно рассуждать двумя способами.

Первый способ. Сумму можно найти с помощью *телескопического суммирования*:

$$\sum_{j=0}^{N-2}\frac{\xi^{2j+1}}{(\zeta^j-P)(\zeta^{j+1}-P)}=\xi\sum_{j=0}^{N-2}\frac{\zeta^j}{(\zeta^j-P)(\zeta^{j+1}-P)}=$$

$$=\frac{\xi}{1-\zeta}\sum_{j=0}^{N-2}\left(\frac{1}{\zeta^{j+1}-P}-\frac{1}{\zeta^j-P}\right)=\frac{\xi}{1-\zeta}\left(\frac{1}{\zeta^{N-1}-P}-\frac{1}{1-P}\right)=\frac{\xi^{N-1}}{(1-P)(\zeta^{N-1}-P)}.$$

Осталось заметить, что

$$\lambda^{-2}(P)\frac{\xi^{N-1}}{(1-P)(\zeta^{N-1}-P)}=\frac{1}{|PA_0|\cdot|PA_{N-1}|}.$$

Второй способ. Указание. С помощью теоремы 1 вычислите сумму

$$\sum_{j=0}^{N-1}\frac{\zeta^j}{(\zeta^j-P)(\zeta^{j+1}-P)}.$$

2.6. а) Как видно из решения п. а) задачи 2.3,

$$S_{-2}^{(1)}(P)=-\sum_{j=0}^{N-1}\frac{P\zeta^j}{(\zeta^j-P)^2}=-\sum_{j=0}^{N-1}R(\zeta^j),\quad \text{где}\quad R(z)=\frac{Pz}{(z-P)^2}.$$

Далее применим теорему 1. В данном случае $F(z)=\dfrac{P}{(z-P)^2(1-z^N)}$. Легко видеть, что

$$\operatorname*{res}_{z=0}F(z)=\operatorname*{res}_{z=\infty}F(z)=0,\qquad \operatorname*{res}_{z=P}F(z)=N\frac{P^N}{(1-P^N)^2}.$$

Таким образом,

$$S_{-2}^{(1)}(P)=-N^2\frac{P^N}{(1-P^N)^2}.$$

б) Воспользуемся равенством (3.2) при $m=(N+1)/2$ и $k=-1$. Имеем

$$R(z)=\frac{z^{(N+1)/2}}{z-P},\qquad F(z)=\frac{z^{(N-1)/2}}{(z-P)(1-z^N)}.$$

Тогда

$$\operatorname*{res}_{z=0}F(z)=\operatorname*{res}_{z=\infty}F(z)=0,\qquad \operatorname*{res}_{z=P}F(z)=\frac{P^{(N-1)/2}}{1-P^N}.$$

Следовательно,

$$S_{-1}^{(2)}(P)=N\lambda^{-1}(P)\frac{P^{(N-1)/2}}{1-P^N}.$$

2.7. *Указание.* Достаточно уметь вычислять при $k = 1, 2, \ldots$ суммы

$$S_k(z) = \sum_{j=0}^{N-1} \frac{1}{(z - \zeta^j)^k}.$$

2.8. а) Решение 1. Если $k \geqslant 1$ нечётно, то $\cos^k x$ представляется в виде линейной комбинации $\cos kx$, $\cos{(k-2)x}$, $\ldots$, $\cos x$. Поэтому достаточно показать, что

$$\sum_{j=0}^{N-1} \cos\frac{2\pi kj}{N} = 0$$

при $1 \leqslant k < N$. Но это очевидным образом следует из равенства (1.2).

Решение 2. Имеем

$$S_k = \frac{1}{2^k}\sum_{j=0}^{N-1}(\zeta^j + \zeta^{-j})^k = \frac{1}{2^k}\sum_{j=0}^{N-1}\sum_{l=0}^{k} C_k^l \zeta^{(k-2l)j} = \frac{1}{2^k}\sum_{l=0}^{k} C_k^l \sum_{j=0}^{N-1}\zeta^{(k-2l)j} = 0,$$

так как $0 < |k - 2l| < N$, и поэтому внутренняя сумма равна нулю.

Решение 3. Применим теорему 1. В данном случае

$$R(z) = \frac{(z^2+1)^k}{2^k z^k}, \quad F(z) = \frac{(z^2+1)^k}{2^k z^{k+1}(1 - z^N)}.$$

Степень числителя равна $2k$, а степень знаменателя есть $k + 1 + N \geqslant 2k + 2$. Поэтому $\operatorname{res}_{z=\infty} F(z) = 0$. Разложим $F(z)$ в ряд Лорана в точке $z = 0$. Поскольку k нечётно и $k < N$, коэффициент при z^{-1} будет нулевым, т. е. $\operatorname{res}_{z=0} F(z) = 0$.

б) Вычислим S_{-1} с помощью теоремы 1. Имеем

$$R(z) = \frac{2z}{z^2+1}, \quad F(z) = \frac{2}{(z^2+1)(1-z^N)},$$

так что $\operatorname{res}_{z=0} F(z) = \operatorname{res}_{z=\infty} F(z) = 0$. Кроме того,

$$\operatorname*{res}_{z=i} F(z) + \operatorname*{res}_{z=-i} F(z) = \frac{1}{i(1-i^N)} + \frac{1}{(-i)(1-(-i)^N)} = (-1)^{(N-1)/2}.$$

Следовательно, $S_{-1} = (-1)^{(N-1)/2} N$.

2.9. Решение 1. Применим теорему 1. В данном случае

$$R(z) = \frac{1}{i}\cdot\frac{az-1}{az+1}, \quad F(z) = \frac{1}{i}\cdot\frac{az-1}{az+1}\cdot\frac{1}{z(1-z^N)},$$

где $a = \exp(2\theta i)$. Учитывая нечётность N, находим

$$\operatorname*{res}_{z=-1/a} F(z) = \frac{1}{i}\cdot\frac{2a^N}{a^N+1}, \quad \operatorname*{res}_{z=0} F(z) = -\frac{1}{i}, \quad \operatorname*{res}_{z=\infty} F(z) = 0.$$

Следовательно,

$$\frac{1}{N}\sum_{j=0}^{N-1} R(\zeta^j) = \frac{1}{i}\cdot\frac{2a^N}{a^N+1} - \frac{1}{i} = \frac{1}{i}\cdot\frac{a^N-1}{a^N+1} = \operatorname{tg}(N\theta).$$

Таким образом, искомая сумма равна $N\operatorname{tg}(N\theta)$.

Решение 2. *Указание.* При нечётном N справедливо тождество

$$\operatorname{tg}(Nx) = \frac{C_N^1\operatorname{tg}x - C_N^3\operatorname{tg}^3 x + \ldots - (-1)^{(N+1)/2}C_N^N\operatorname{tg}^N x}{C_N^0 - C_N^2\operatorname{tg}^2 x + \ldots - (-1)^{(N+1)/2}C_N^{N-1}\operatorname{tg}^{N-1} x}.$$

Подставьте сюда $x = \theta + \pi j/N$ для $j = 0, 1, \ldots, N-1$.

Замечание. При чётном N сумма равна $-N\operatorname{ctg}(N\theta)$.

2.10. Воспользуемся теоремой 1 в её модифицированной версии. Здесь

$$R(z) = -\frac{(z+1)^2}{(z-1)^2}, \quad F(z) = -\frac{(z+1)^2}{z(z-1)^2(1-z^N)}.$$

Легко видеть, что $\operatorname{res}_{z=0} F(z) = -1$, $\operatorname{res}_{z=\infty} F(z) = 0$. Положим

$$G(z) = (z-1)^3 F(z) = \frac{(z+1)^2(1-z)}{z(1-z^N)}.$$

Тогда, как показывает прямое, но несколько утомительное вычисление,

$$\operatorname*{res}_{z=1} F(z) = \frac{G''(1)}{2} = \frac{N^2+2}{3N}.$$

Таким образом, имеем

$$\frac{1}{N}\sum_{j=1}^{N-1} R(\zeta^j) = \frac{N^2+2}{3N} - 1 = \frac{(N-1)(N-2)}{3N},$$

откуда и следует утверждение задачи.

Замечание. Как вариант, можно было бы доказать равенство

$$\sum_{j=0}^{N-1}\operatorname{ctg}^2\left(\theta + \frac{\pi j}{N}\right) = N\operatorname{ctg}^2(N\theta) + \frac{N(N-1)}{\sin^2(N\theta)}$$

(что даже технически проще) и затем, перенеся $\operatorname{ctg}^2\theta$ в правую часть, перейти к пределу при $\theta \to 0$.

2.11. Применим теорему 1. Положим $a = \exp\left(-\frac{\pi i}{2N}\right)$, так что $\zeta = a^{-4}$ и $a^N = -i$. Нетрудно проверить, что

$$\frac{\sin\dfrac{\pi k(4j+1)}{2N}}{1-\cos\dfrac{\pi(4j+1)}{2N}} = R(\zeta^j), \quad \text{где} \quad R(z) = \frac{a^{2k}-z^{2k}}{ia^{k-1}z^{k-1}(a-z)^2}.$$

Таким образом, в данном случае

$$F(z) = \frac{a^{2k}-z^{2k}}{ia^{k-1}z^k(a-z)^2(1-z^N)}.$$

Вычет в (простом!) полюсе $z = a$ находится легко: $\operatorname{res}_{z=a} F(z) = k(1+i)$. Для вычисления вычета в полюсе $z = 0$ воспользуемся рядами

$$\frac{1}{(a-z)^2} = \frac{1}{a^2} + \frac{2z}{a^3} + \frac{3z^2}{a^4} + \ldots, \qquad \frac{1}{1-z^N} = 1 + z^N + z^{2N} + \ldots,$$

а также тем, что $k \leqslant N$. Перемножив ряды, найдём

$$\operatorname*{res}_{z=0} \frac{1}{z^k(a-z)^2(1-z^N)} = \frac{k}{a^{k+1}}.$$

Отсюда

$$\operatorname*{res}_{z=0} F(z) = \frac{a^{k+1}}{i} \cdot \operatorname*{res}_{z=0} \frac{1}{z^k(a-z)^2(1-z^N)} = \frac{a^{k+1}}{i} \cdot \frac{k}{a^{k+1}} = -ki.$$

Наконец, вычислим вычет в точке $z = \infty$. Поскольку степень числителя $F(z)$ есть $2k-1$, а степень знаменателя есть $k+1+N \geqslant 2k+1$, получим $\operatorname{res}_{z=\infty} F(z) = 0$. Таким образом,

$$\frac{1}{N}\sum_{j=0}^{N-1} R(\zeta^j) = k(1+i) - ki = k,$$

что и требовалось доказать.

2.12. Воспользуемся теоремой 1 при $N = 2n$. Положим $\xi = \exp\frac{\pi i}{2n}$, так что $\zeta = \xi^2$ и $\xi^{2n} = -1$. Удобнее сразу находить комплексную сумму

$$S = \frac{1}{4n}\sum_{j=0}^{2n-1}(-1)^{j+1}\frac{\exp\dfrac{\pi k(2j-1)i}{2n}}{\sin^2\dfrac{\pi(2j-1)}{4n}} =$$

$$= \frac{1}{2n}\sum_{j=0}^{2n-1}(-1)^{j+1}\frac{\exp\dfrac{\pi k(2j-1)i}{2n}}{1-\cos\dfrac{\pi(2j-1)}{2n}} = \frac{1}{2n}\sum_{j=0}^{2n-1} R(\zeta^j),$$

для которой

$$R(z) = \frac{2z^{n+k+1}}{\xi^{k-1}(z-\xi)^2}, \quad F(z) = \frac{2z^{n+k}}{\xi^{k-1}(z-\xi)^2(1-z^{2n})}.$$

Если $0 \leqslant k \leqslant n$, то $\mathrm{res}_{z=0} F(z) = \mathrm{res}_{z=\infty} F(z) = 0$. Кроме того, $\mathrm{res}_{z=\xi} F(z) = G'(\xi) = ki$, что получается непосредственным дифференцированием функции

$$G(z) = \frac{2z^{n+k}}{\xi^{k-1}(1-z^{2n})}.$$

В итоге $S = ki$, что и требовалось.

2.13. Рассмотрим рациональную функцию

$$R(z) = \frac{z^2+1}{(z^2-1)(z^l+1)},$$

к которой применим теорему 1 при $N = k$. Обозначим $\zeta_* = \exp(2\pi i/l)$. Имеет место следующее равенство:

$$\frac{1}{k}\sum_{j=1}^{k-1} R(\zeta^j) = \operatorname*{res}_{z=1} F(z) + \operatorname*{res}_{z=-1} F(z) + \sum_{j=1}^{l-1} \operatorname*{res}_{z=-\zeta_*^j} F(z) + \operatorname*{res}_{z=0} F(z) + \operatorname*{res}_{z=\infty} F(z),$$

где

$$F(z) = \frac{z^2+1}{z(z^2-1)(z^l+1)(1-z^k)}.$$

Очевидно, что $\mathrm{res}_{z=0} F(z) = -1$, $\mathrm{res}_{z=\infty} F(z) = 0$. Чуть сложнее убедиться в том, что для каждого $j = 1, \ldots, l-1$

$$\operatorname*{res}_{z=-\zeta_*^j} F(z) = -\frac{1}{l} R_*(\zeta_*^j), \quad \text{где} \quad R_*(z) = \frac{z^2+1}{(z^2-1)(z^k+1)}.$$

И, наконец, самое сложное вычисление:

$$\operatorname*{res}_{z=1} F(z) = \frac{k+l}{4k}, \quad \operatorname*{res}_{z=-1} F(z) = \frac{k+l}{4l}$$

(из соображений симметрии достаточно получить только первое равенство). Итак,

$$\frac{1}{k}\sum_{j=1}^{k-1} R(\zeta^j) = \frac{k+l}{4k} + \frac{k+l}{4l} - \frac{1}{l}\sum_{j=1}^{l-1} R_*(\zeta_*^j) - 1,$$

что можно переписать в виде

$$\frac{1}{k}\sum_{j=1}^{k-1} R(\zeta^j) + \frac{1}{l}\sum_{j=1}^{l-1} R_*(\zeta_*^j) = \frac{(k-l)^2}{4kl}.$$

Осталось домножить это равенство на $-2kl$ и взять вещественную часть. Читателю рекомендуется аккуратно проверить всё, что касается вычисления вычетов функции $F(z)$.

Замечание 1. Если взять

$$R(z) = \frac{z+1}{(z-1)(z^l-1)},$$

то аналогичным образом возникнет тождество

$$l\sum_{j=1}^{k-1} \operatorname{ctg}\frac{\pi j}{k} \operatorname{ctg}\frac{l\pi j}{k} + k\sum_{j=1}^{l-1} \operatorname{ctg}\frac{\pi j}{l} \operatorname{ctg}\frac{k\pi j}{l} = \frac{k^2+l^2+1}{3} - kl, \tag{3.3}$$

где k, l — любые взаимно простые натуральные числа. С помощью тождества (3.3) можно доказать теорему взаимности для так называемых *сумм Дедекинда* (см., например, статью [23], а также § 3.7 в книге [16])

$$s(l,k) = \sum_{r=0}^{k-1} \rho\Big(\frac{r}{k}\Big)\rho\Big(\frac{lr}{k}\Big),$$

где функция $\rho(x)$ вещественного аргумента x определена следующим образом:

$$\rho(x) = \begin{cases} \{x\} - \dfrac{1}{2}, & \text{если } x \notin \mathbb{Z}, \\ 0, & \text{если } x \in \mathbb{Z} \end{cases}$$

(здесь $\{x\}$ — дробная часть x). Для этого нужно предварительно установить равенство

$$s(l,k) = \frac{1}{4k}\sum_{j=1}^{k-1} \operatorname{ctg}\frac{\pi j}{k} \operatorname{ctg}\frac{l\pi j}{k}, \tag{3.4}$$

где k, l предполагаются взаимно простыми. Равенство (3.4), в свою очередь, доказывается на основе представления

$$\rho\Big(\frac{r}{k}\Big) = \frac{1}{k}\sum_{j=1}^{k-1}\Big(\frac{1}{1-\zeta^j} - \frac{1}{2}\Big)\zeta^{rj}, \quad \zeta = \exp\Big(\frac{2\pi i}{k}\Big),$$

которое легко проверить, применив теорему 1 при $N = k$ к рациональной функции $R(z) = z^r/(1-z)$ (подробный вывод равенства (3.4) см. в [23]).

Замечание 2. Можно предложить такой рецепт получения подобных теорем взаимности для тригонометрических сумм: применять теорему 1 при $N = k$ к рациональной функции $R(z) = h(z)/(z^l \pm 1)$, где $h(z)$ — какая-нибудь фиксированная рациональная функция. Конкретные примеры теорем взаимности читатель может найти в статье [18].

2.14. Пусть $\xi = \exp(\pi i/N)$ и $\zeta = \xi^2 = \exp(2\pi i/N)$. Тогда

$$\operatorname{tg}\frac{\pi j}{N} = -2\operatorname{Im}\frac{1}{\zeta^j+1}, \quad \operatorname{ctg}\frac{\pi j}{N} = -2\operatorname{Im}\frac{1}{\zeta^j-1}.$$

Решение 1. Нам понадобится следующий аналог равенства (1.2):

$$\sum_{j=0}^{N-1} j\zeta^{kj} = \begin{cases} \dfrac{N}{\zeta^k-1}, & \text{если } k \not\equiv 0 \pmod N, \\ \dfrac{N(N-1)}{2}, & \text{если } k \equiv 0 \pmod N). \end{cases} \tag{3.5}$$

Вычислим сумму $\sum_{j=0}^{N-1}\dfrac{j}{r\zeta^j+1}$, где $|r|<1$, используя разложение в ряд

$$\frac{1}{z+1} = \sum_{k=0}^{\infty}(-1)^k z^k.$$

Имеем

$$\sum_{j=0}^{N-1}\frac{j}{r\zeta^j+1} = \sum_{j=0}^{N-1}\sum_{k=0}^{\infty} j(-1)^k r^k \zeta^{kj} = \sum_{k=0}^{\infty}(-1)^k r^k \sum_{j=0}^{N-1} j\zeta^{kj}.$$

Пусть $k = Nl + j$, где $l \geqslant 0$ и $0 \leqslant j \leqslant N-1$. С помощью (3.5) находим:

$$\sum_{k=0}^{\infty}(-1)^k r^k \sum_{j=0}^{N-1} j\zeta^{kj} =$$

$$= \sum_{l=0}^{\infty}(-1)^{Nl} r^{Nl}\frac{N(N-1)}{2} + \sum_{j=1}^{N-1}\sum_{l=0}^{\infty}(-1)^{Nl+j} r^{Nl+j}\frac{N}{\zeta^j-1} =$$

$$= \frac{1}{r^N+1}\left(\frac{N(N-1)}{2} + N\sum_{j=1}^{N-1}\frac{(-1)^j r^j}{\zeta^j-1}\right).$$

Отсюда при $r \to 1$ вытекает равенство

$$\sum_{j=0}^{N-1}\frac{j}{\zeta^j+1} = \frac{N(N-1)}{4} + \frac{N}{2}\sum_{j=1}^{N-1}\frac{(-1)^j}{\zeta^j-1}. \tag{3.6}$$

Осталось перейти к мнимым частям.

Решение 2. Сначала докажем тождество

$$\sum_{j=0}^{N-1}\frac{q^j}{\zeta^j+1} = \frac{q^N-1}{2}\sum_{j=0}^{N-1}\frac{(-1)^j}{q\zeta^j-1}. \tag{3.7}$$

Действительно, имеем

$$\sum_{j=0}^{N-1} \frac{q^j}{r\zeta^j+1} = \sum_{j=0}^{N-1} q^j \sum_{k=0}^{\infty}(-1)^k r^k \zeta^{kj} =$$

$$= \sum_{k=0}^{\infty}(-1)^k r^k \sum_{j=0}^{N-1}(q\zeta^k)^j = \sum_{k=0}^{\infty}(-1)^k r^k \frac{q^N-1}{q\zeta^k-1} =$$

$$= \sum_{j=0}^{N-1}\sum_{l=0}^{\infty}(-1)^{Nl+j} r^{Nl+j} \frac{q^N-1}{q\zeta^j-1} = \frac{q^N-1}{r^N+1}\sum_{j=0}^{N-1}\frac{(-1)^j r^j}{q\zeta^j-1}.$$

Теперь тождество (3.7) получается предельным переходом при $r \to 1$.

Далее можно получить равенство (3.6), для чего нужно продифференцировать тождество (3.7) по q и затем перейти к пределу при $q \to 1$.

Решение 3. Здесь мы опираемся на такой аналог равенства (1.2):

$$\sum_{j=0}^{N-1} \xi^{kj} = \begin{cases} \dfrac{1-(-1)^k}{1-\xi^k}, & \text{если } k \not\equiv 0 \pmod{2N}, \\ N, & \text{если } k \equiv 0 \pmod{2N}. \end{cases}$$

Как следствие, при нечётном N имеет место равенство

$$\sum_{j=1}^{N-1} \xi^{(N+2k)j} = \frac{1+\xi^{N+2k}}{1-\xi^{N+2k}} = \frac{1-\zeta^k}{1+\zeta^k}.$$

Будем вычислять сумму

$$\sum_{j=1}^{N-1} \frac{(-1)^j}{\zeta^j-1} = \sum_{j=1}^{N-1}\frac{\xi^{Nj}}{\xi^{2j}-1}.$$

Имеем

$$\sum_{j=1}^{N-1}\frac{\xi^{Nj}}{(r\xi^j)^2-1} = -\sum_{k=0}^{\infty} r^{2k}\sum_{j=1}^{N-1}\xi^{(N+2k)j} = \sum_{k=0}^{\infty} r^{2k}\frac{\zeta^k-1}{\zeta^k+1} =$$

$$= \sum_{j=0}^{N-1}\sum_{l=0}^{\infty} r^{2Nl+2j}\frac{\zeta^j-1}{\zeta^j+1} = \frac{1}{1-r^{2N}}\sum_{j=0}^{N-1} r^{2j}\frac{\zeta^j-1}{\zeta^j+1}.$$

Теперь с помощью *правила Лопиталя* при $r \to 1$ получим

$$\sum_{j=1}^{N-1}\frac{(-1)^j}{\zeta^j-1} = -\frac{1}{N}\sum_{j=0}^{N-1} j\frac{\zeta^j-1}{\zeta^j+1} = \frac{2}{N}\sum_{j=0}^{N-1}\frac{j}{\zeta^j+1} - \frac{N-1}{2},$$

что эквивалентно равенству (3.6).

Замечание. Ещё один вариант доказанного равенства выглядит так:

$$\sum_{j=0}^{N-1} j^2 \operatorname{tg}\frac{\pi j}{N} = \frac{N^2}{2}\sum_{j=1}^{N-1}(-1)^j \operatorname{ctg}\frac{\pi j}{N}.$$

2.15–2.17. *Указание.* Данные произведения равны

$$\frac{(-1)^{N-1}}{2^{N-1}}\prod_{j=1}^{N-1}(\zeta^j-1), \quad (-1)^{(N-1)/2}\prod_{j=1}^{N-1}\frac{\zeta^j-1}{\zeta^j+1}, \quad \prod_{j=1}^{N-1}(2\zeta^j-1)(\zeta^j-2)$$

соответственно, где $\zeta=\exp(2\pi i/N)$.

§ 4. Вариации и обобщения

В этом параграфе мы рассмотрим некоторые другие типы тригонометрических сумм.

Суммы первого типа отличаются от стандартных сумм (0.1) тем, что суммирование происходит только по тем $j=0,1,\ldots,N-1$, которые взаимно просты с N.

В суммах второго типа суммирование распространяется на все $j=0,1,\ldots,N-1$, но j-е слагаемое содержит дополнительный множитель, зависящий от (j,N) — наибольшего общего делителя чисел j и N.

Суммы третьего типа аналогичны классическим квадратичным суммам Гаусса. Мы дополнительно будем считать, что $N=p>2$ — простое число.

4.1. Суммирование по приведённой системе вычетов

В стандартной тригонометрической сумме (0.1) суммирование происходит по *полной системе вычетов* $Z_N=\{0,1,\ldots,N-1\}$ по модулю N. Можно взять в качестве области суммирования *приведённую систему вычетов* по модулю N:

$$Z_N^*=\{j\in Z_N\colon (j,N)=1\}.$$

Пусть, как и выше, $\zeta=\exp(2\pi i/N)$. Рассмотрим тригонометрическую сумму

$$S^*(N)=\sum_{j\in Z_N^*}R(\zeta^j), \tag{4.1}$$

число слагаемых в которой равно $\varphi(N)$. Легко видеть, что

$$S(N)=\sum_{j=0}^{N-1}R(\zeta^j)=\sum_{d|N}\sum_{j'\in Z_{N/d}^*}R(\zeta_d^{j'})=\sum_{d|N}S^*\Big(\frac{N}{d}\Big)=\sum_{d|N}S^*(d), \tag{4.2}$$

где $\zeta_d = \zeta^d$ есть первообразный корень из единицы степени N/d. Полученное соотношение можно использовать для вычисления суммы $S^*(N)$.

Действительно, непосредственное применение формулы обращения Мёбиуса (см., например, § 3.2 в книге [9]) даёт

$$S^*(N) = \sum_{d|N} \mu(d) S\Big(\frac{N}{d}\Big). \tag{4.3}$$

Таким образом, значения специальной суммы (4.1) могут быть выражены через значения стандартной суммы (0.1). Отметим, что формулу (4.3) можно также получить с помощью хорошо известного из комбинаторики *принципа включений — исключений* (см., например, по ссылке [30]).

Правая часть в равенстве (4.3) допускает дальнейшее упрощение, если функция $S(N)$ мультипликативна, т. е. выполняется равенство

$$S(N_1 N_2) = S(N_1) S(N_2) \tag{4.4}$$

при условии $(N_1, N_2) = 1$. В этом случае функция $S^*(N)$ также будет мультипликативной (как свёртка Дирихле двух мультипликативных функций). Для $N = p^\alpha$, где p — простое число, имеем

$$\sum_{d|N} \mu(d) S\Big(\frac{N}{d}\Big) = \sum_{\beta=0}^{\alpha} \mu(p^\beta) S(p^{\alpha-\beta}) = S(p^\alpha) - S(p^{\alpha-1}).$$

Следовательно, если

$$N = \prod_{p|N} p^\alpha \tag{4.5}$$

— *каноническое разложение* числа N, то

$$S^*(N) = \prod_{p|N} (S(p^\alpha) - S(p^{\alpha-1})).$$

Более того, если функция $S(N)$ *вполне мультипликативна* (равенство (4.4) верно для любых N_1 и N_2), то предыдущая формула упрощается до

$$S^*(N) = S(N) \prod_{p|N} \Big(1 - \frac{1}{S(p)}\Big)$$

(в предположении, что $S(p) \neq 0$ для любого простого числа p).

В качестве примера вычислим сумму

$$S^*(N) = \sum_{j \,\in\, Z_N^*} \frac{1}{\cos^2 \dfrac{\pi j}{N}}, \tag{4.6}$$

где N нечётно. Так как в данном случае $S(N)=N^2$ (см. решение задачи 1.1) и эта функция вполне мультипликативна, то сразу получим

$$S^*(N) = N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right). \tag{4.7}$$

Способ решения задачи 1.1 через разложения в ряды подсказывает альтернативный путь получения формулы (4.7), но здесь необходим аналог равенства (1.2) для суммы

$$c_N(k) = \sum_{j \in Z_N^*} \zeta^{kj},$$

которая известна как *сумма Рамануджана*. В частности, имеем

$$c_N(0) = \varphi(N), \quad c_N(1) = \mu(N).$$

Первая из этих формул очевидна по определению. Для доказательства второй заметим, что при любом k сумма Рамануджана $c_N(k)$ является мультипликативной функцией от N (потому что такова правая часть равенства (1.2) как функция от N). Поэтому $c_N(k)$ можно находить по общему правилу, описанному выше. В случае $k = 1$ мы таким способом легко получим вторую из формул. Общую формулу

$$c_N(k) = \frac{\varphi(N)\mu(N/(k,N))}{\varphi(N/(k,N))}$$

читателю предлагается вывести, опираясь на формулу для $c_N(1)$. Суммы Рамануджана $c_N(k)$ имеют важные приложения в теории чисел и поэтому представляют самостоятельный интерес (см. оригинальную статью [26], а также § А.7 в книге [24]).

Если для некоторой суммы (0.1) имеет место равенство $S(N) = N^a$, то для соответствующей суммы (4.1) будем иметь $S^*(N) = J_a(N)$, где

$$J_a(N) = N^a \prod_{p|N} \left(1 - \frac{1}{p^a}\right)$$

— так называемые *функции Жордана* (см., например, упр. 17 на с. 48 книги [15]). В частности, $J_1(N)$ — это функция Эйлера $\varphi(N)$. Читателю предлагается убедиться, что для $a = 1$ и каждого чётного $a \geqslant 2$ существует рациональная функция $R_a(z)$, для которой

$$\sum_{j=0}^{N-1} R_a(\zeta^j) = N^a. \tag{4.8}$$

Например:

$$R_1(z) = \frac{2}{z+1}, \quad R_2(z) = \frac{4z}{(z+1)^2}, \quad R_4(z) = -\frac{8z(z^2-4z+1)}{(z+1)^4} \quad \text{и т. д.}$$

Ещё один способ получения формулы (4.7) таков. Многочлен

$$\prod_{j\,\in\, Z_N^*} (z-\zeta^j) = \Phi_N(z)$$

называется *круговым многочленом* порядка N (см., например, § 13 в книге [11]). С помощью формулы обращения Мёбиуса круговой многочлен может быть записан в виде явной рациональной дроби с целыми коэффициентами:

$$\Phi_N(z) = \prod_{d|N} (z^{N/d}-1)^{\mu(d)}. \tag{4.9}$$

Следовательно,

$$\sum_{j\,\in\, Z_N^*} \frac{1}{z-\zeta^j} = \frac{\Phi_N'(z)}{\Phi_N(z)}, \quad \sum_{j\,\in\, Z_N^*} \frac{1}{(z-\zeta^j)^2} = -\Big[\frac{\Phi_N'(z)}{\Phi_N(z)}\Big]'.$$

Из представления (4.9) следует, что

$$\frac{\Phi_N'(z)}{\Phi_N(z)} = N \sum_{d|N} \frac{\mu(d)}{d} \cdot \frac{z^{N/d-1}}{z^{N/d}-1}. \tag{4.10}$$

При $z=-1$ и нечётном N правая часть последнего равенства равна

$$-\frac{1}{2}\cdot N \sum_{d|N} \frac{\mu(d)}{d} = -\frac{1}{2}\cdot N \prod_{p|N}\Big(1-\frac{1}{p}\Big) = -\frac{1}{2}\cdot J_1(N).$$

Таким образом, имеем

$$\sum_{j\,\in\, Z_N^*} \frac{1}{1+\zeta^j} = \frac{J_1(N)}{2}.$$

Далее, опираясь на равенство

$$N^a \sum_{d|N} \frac{\mu(d)}{d^a} = J_a(N),$$

аналогично можно показать, что

$$\sum_{j\,\in\, Z_N^*} \frac{1}{(1+\zeta^j)^2} = -\frac{J_2(N)}{4} + \frac{J_1(N)}{2}$$

и, вообще, для любого натурального s сумма $\sum_{j\,\in\, Z_N^*} \dfrac{1}{(1+\zeta^j)^s}$ есть линейная комбинация функций Жордана $J_1(N)$ и $J_a(N)$ с чётными $a \leqslant s$. В итоге

$$S^*(N) = \sum_{j\,\in\, Z_N^*} \frac{4\zeta^j}{(1+\zeta^j)^2} = 4\sum_{j\,\in\, Z_N^*} \frac{1}{1+\zeta^j} - 4\sum_{j\,\in\, Z_N^*} \frac{1}{(1+\zeta^j)^2} = J_2(N),$$

как и ожидалось. Читателю предлагается доказать, что для любого натурального s аналогичная сумма $\sum_{j\,\in\, Z_N^*} \dfrac{1}{(1-\zeta^j)^s}$ также будет линейной комбинацией функций

Жордана $J_1(N)$ и $J_a(N)$ с чётными $a \leqslant s$ (теперь N — любое натуральное число, большее единицы)[6]. Так, например, имеем

$$\sum_{j\,\in\, Z_N^*}\frac{1}{1-\zeta^j}=\frac{J_1(N)}{2},\quad \sum_{j\,\in\, Z_N^*}\frac{1}{(1-\zeta^j)^2}=-\frac{J_2(N)}{12}+\frac{J_1(N)}{2}\quad \text{и т. д.}$$

Полученные формулы для сумм (4.1) (в частности, формула (4.7) для суммы (4.6)) окажутся полезными при вычислении тригонометрических сумм ещё одного класса.

4.2. Суммы с наибольшими общими делителями

Одним из обобщений стандартных сумм (0.1) являются суммы вида

$$S^{**}(N)=\sum_{j=0}^{N-1}D((j,N))R(\zeta^j),$$

где $D(N)$ — какая-нибудь арифметическая функция. Очевидным образом модифицируя цепочку равенств (4.2), получим следующую формулу:

$$S^{**}(N)=\sum_{d|N}S^*(d)D\Big(\frac{N}{d}\Big), \tag{4.11}$$

где $S^*(N)$ — сумма (4.1). При наличии дополнительных свойств (типа мультипликативности) у функций $S^*(N)$ и $D(N)$ формула (4.11) допускает дальнейшее упрощение в том же стиле, что и формула (4.3). Не вдаваясь в подробности, рассмотрим конкретный пример, а именно сумму

$$S^{**}(N)=\sum_{j=0}^{N-1}\frac{(j,N)}{\cos^2\dfrac{\pi j}{N}}, \tag{4.12}$$

где N нечётно (см. [25]). В данном случае $D(N)=N$, $S^*(N)=J_2(N)$ (см. формулу (4.7)). Поскольку обе функции мультипликативны, такой же будет и функция $S^{**}(N)$. Это позволяет стандартным способом получить следующую формулу для суммы (4.12) (в терминах канонического разложения (4.5) числа N):

$$S^{**}(N)=N\sum_{d|N}\frac{J_2(d)}{d}=N\prod_{p|N}\Big(p^\alpha+p^{\alpha-1}-\frac{1}{p}\Big). \tag{4.13}$$

В частности, из формулы (4.13) следует, что сумма (4.12) всегда есть целое число.

---

[6] Для начала нужно ответить на вопрос: как подставить $z=1$ в правую часть равенства (4.10)?

Как второй пример рассмотрим сумму

$$S^{**}(N) = \sum_{j=0}^{N-1} (j, N)\zeta^{kj} = N \sum_{d|N} \frac{c_d(k)}{d}.$$

Здесь также возможны дальнейшие упрощения. Например, при $k = 1$ имеем

$$N \sum_{d|N} \frac{c_d(1)}{d} = N \sum_{d|N} \frac{\mu(d)}{d} = \varphi(N).$$

Фактически речь идёт о вычислении *дискретного преобразования Фурье* для последовательности (j, N) при $j = 0, 1, \ldots, N-1$ (подробности см. в [27]).

И ещё один пример: формула

$$\sum_{j=0}^{N-1} (j, N) R_a(\zeta^j) = N \sum_{d|N} \frac{J_a(d)}{d}, \tag{4.14}$$

где N нечётно, а рациональные функции $R_a(z)$ удовлетворяют условию (4.8). Читателю предлагается упростить правую часть равенства (4.14), а также обобщить результаты на случай степенной функции $D(N) = N^b$ или какой-нибудь другой мультипликативной функции.

4.3. Суммы Гаусса и их аналоги

В этом разделе пусть $N = p > 2$ — простое число. Сумма вида

$$g_a(p) = \sum_{j=0}^{p-1} \left(\frac{j}{p}\right) \zeta^{aj},$$

где $a \in \mathbb{Z}$, называется *суммой Гаусса*[7]. Здесь $\zeta = \exp(2\pi i/p)$, а $\left(\frac{j}{p}\right)$ обозначает *символ Лежандра*: если $j \equiv 0 \pmod p$, то $\left(\frac{j}{p}\right) = 0$, иначе

$$\left(\frac{j}{p}\right) = \begin{cases} 1, & \text{если сравнение } x^2 \equiv j \pmod p \text{ разрешимо,} \\ -1 & \text{в противном случае} \end{cases}$$

(см. любой учебник по элементарной теории чисел, например [9]).

Пусть для краткости $g_1(p) = g(p)$. Легко показать, что

$$g_a(p) = \left(\frac{a}{p}\right) g(p),$$

[7] Точнее, квадратичной суммой Гаусса. Детальное изложение основных фактов о квадратичных суммах Гаусса имеется, например, в главе 6 книги [1]. Существует естественный аналог сумм Гаусса для *конечных полей*, о чём можно прочитать в книге [17].

а также дать другое представление суммы Гаусса при $a \not\equiv 0 \pmod p$:

$$g_a(p) = \sum_{j=0}^{p-1} \zeta^{aj^2}.$$

Гораздо труднее вычислить сумму $g(p)$. Точнее, нетрудно установить равенство

$$g^2(p) = (-1)^{(p-1)/2} p,$$

откуда $g(p) = \pm\sqrt{p}$ или $g(p) = \pm i\sqrt{p}$, но настоящую проблему составляет отыскание *знака* суммы $g(p)$. На самом деле, как доказал сам Гаусс,

$$g(p) = \begin{cases} \sqrt{p}, & \text{если } p \equiv 1 \pmod 4, \\ i\sqrt{p}, & \text{если } p \equiv 3 \pmod 4. \end{cases} \tag{4.15}$$

Различные доказательства этого факта и его обобщений можно найти в книгах [1, 3, 7, 15, 17] (см. также серию задач на с. 22–24 книги [13]).

Приведём самое, на наш взгляд, элементарное доказательство, которое принадлежит Кронекеру (см. [1, с. 95–98]). Воспользуемся равенством

$$\prod_{l=1}^{(p-1)/2} (\zeta^{2l-1} - \zeta^{-2l+1}) = \begin{cases} \sqrt{p}, & \text{если } p \equiv 1 \pmod 4, \\ i\sqrt{p}, & \text{если } p \equiv 3 \pmod 4, \end{cases}$$

которое является другой записью равенства (1.7) при $N = p$ (см. задачу 1.2). Из сказанного выше про $g(p)$ следует, что

$$g(p) = \varepsilon \prod_{l=1}^{(p-1)/2} (\zeta^{2l-1} - \zeta^{-2l+1}), \quad \text{где } \varepsilon = \pm 1.$$

Нужно доказать, что всегда $\varepsilon = 1$. Для этого рассмотрим многочлен

$$f(x) = \sum_{j=0}^{p-1} \left(\frac{j}{p}\right) x^j - \varepsilon \prod_{l=1}^{(p-1)/2} (x^{2l-1} - x^{p-2l+1}).$$

Понятно, что многочлен $f(x)$ имеет целые коэффициенты, при этом $f(1) = f(\zeta) = 0$. Как известно, круговой многочлен

$$\Phi_p(x) = x^{p-1} + \ldots + x + 1$$

неприводим[8] над полем рациональных чисел $\mathbb{Q}$. Значит, $f(x)$ должен делиться на произведение $(x-1)\Phi_p(x) = x^p - 1$, т. е. $f(x) = (x^p - 1)g(x)$ для некоторого

[8] Это верно для любого кругового многочлена $\Phi_N(x)$, но в случае когда $N = p$ — простое число, утверждение о неприводимости нетрудно доказывается с помощью хорошо известного *критерия Эйзенштейна* (см., например, § 6.2 в книге [11]).

многочлена $g(x)$ с целыми коэффициентами. Сделаем здесь замену переменной $x=1+t$ и посмотрим на коэффициент при $t^{(p-1)/2}$ слева и справа. Слева он равен

$$\sum_{j=0}^{p-1}\Big(\frac{j}{p}\Big)C_j^{(p-1)/2}-\varepsilon\prod_{l=1}^{(p-1)/2}(4l-2-p),$$

а справа он делится на p (из-за множителя $(1+t)^p-1$, все коэффициенты которого, за исключением старшего, кратны p). Как следствие,

$$\sum_{j=0}^{p-1}\Big(\frac{j}{p}\Big)C_j^q\equiv\varepsilon\prod_{l=1}^{q}(4l-2)\ (\mathrm{mod}\ p), \tag{4.16}$$

где для удобства записи введено обозначение $q=(p-1)/2$. Имеем

$$q!C_j^q=j(j-1)\dots(j-q+1)=j^q+a_{q-1}j^{q-1}+\dots+a_1j,$$
$$q!\prod_{l=1}^{q}(4l-2)=(p-1)!.$$

По *теореме Эйлера* $\Big(\frac{j}{p}\Big)\equiv j^q\ (\mathrm{mod}\ p)$, а по *теореме Вильсона* $(p-1)!\equiv-1\ (\mathrm{mod}\ p)$. Следовательно, сравнение (4.16) после домножения на $q!$ приводится к виду

$$\sum_{j=0}^{p-1}(j^{2q}+a_{q-1}j^{2q-1}+\dots+a_1j^{q+1})\equiv-\varepsilon\quad(\mathrm{mod}\ p). \tag{4.17}$$

Наконец, заметим, что

$$\sum_{j=0}^{p-1}j^{2q-1}\equiv\dots\equiv\sum_{j=0}^{p-1}j^{q+1}\equiv0\ (\mathrm{mod}\ p),$$
$$\sum_{j=0}^{p-1}j^{2q}=\sum_{j=0}^{p-1}j^{p-1}\equiv-1\ (\mathrm{mod}\ p)$$

(в справедливости этих сравнений легко убедиться, если воспользоваться цикличностью группы ненулевых классов вычетов по модулю p). В итоге сравнение (4.17) сводится к сравнению $-1\equiv-\varepsilon\ \ (\mathrm{mod}\ p)$, откуда и следует, что $\varepsilon=1$.

Используя классический результат (4.15), докажем равенство

$$\sum_{j=0}^{p-1}\mathrm{tg}\,\frac{\pi j^2}{p}=\begin{cases}0, & \text{если } p\equiv1\ (\mathrm{mod}\ 4),\\ -(-1)^{(p^2-1)/8}2\sigma_0(p)\sqrt{p}, & \text{если } p\equiv3\ (\mathrm{mod}\ 4),\end{cases} \tag{4.18}$$

где

$$\sigma_0(p)=\sum_{b=1}^{(p-1)/2}\Big(\frac{b}{p}\Big).$$

Для этого будем вычислять сумму

$$S^\dagger(p)=\sum_{j=0}^{p-1}\frac{1}{1+\zeta^{j^2}}$$

методом разложения в ряд. Имеем

$$\sum_{j=0}^{p-1}\frac{1}{1+r\zeta^{j^2}}=\sum_{j=0}^{p-1}\sum_{k=0}^{\infty}(-1)^k r^k\zeta^{kj^2}=\sum_{k=0}^{\infty}(-1)^k r^k\sum_{j=0}^{p-1}\zeta^{kj^2}=$$

$$=\sum_{l=0}^{\infty}\sum_{a=0}^{p-1}(-1)^{pl+a}r^{pl+a}\sum_{j=0}^{p-1}\zeta^{(pl+a)j^2}=$$

$$=\frac{p}{1+r^p}+\sum_{l=0}^{\infty}(-1)^l r^{pl}\sum_{a=1}^{p-1}(-1)^a r^a\sum_{j=0}^{p-1}\zeta^{aj^2}=$$

$$=\frac{p}{1+r^p}+\frac{g(p)}{1+r^p}\sum_{a=1}^{p-1}(-1)^a r^a\Big(\frac{a}{p}\Big).$$

Переходя к пределу при $r\to 1$, получим

$$S^\dagger(p)=\frac{p}{2}+\frac{g(p)}{2}\sum_{a=1}^{p-1}(-1)^a\Big(\frac{a}{p}\Big).$$

Преобразуем сумму в правой части полученного равенства:

$$\sum_{a=1}^{p-1}(-1)^a\Big(\frac{a}{p}\Big)=\sum_{b=1}^{(p-1)/2}\Big(\frac{2b}{p}\Big)-\sum_{b=1}^{(p-1)/2}\Big(\frac{p-2b}{p}\Big)=\Big(\frac{2}{p}\Big)\Big(1-\Big(\frac{-1}{p}\Big)\Big)\sigma_0(p).$$

Напомним, что

$$\Big(\frac{-1}{p}\Big)=(-1)^{(p-1)/2},\qquad \Big(\frac{2}{p}\Big)=(-1)^{(p^2-1)/8}.$$

В частности, если $p\equiv 1\pmod 4$, то

$$\sum_{a=1}^{p-1}(-1)^a\Big(\frac{a}{p}\Big)=0.$$

Более того, в этом случае имеем

$$0=\sum_{a=1}^{p-1}\Big(\frac{a}{p}\Big)=\sum_{b=1}^{(p-1)/2}\Big(\frac{b}{p}\Big)+\sum_{b=1}^{(p-1)/2}\Big(\frac{p-b}{p}\Big)=\Big(1+\Big(\frac{-1}{p}\Big)\Big)\sigma_0(p)=2\sigma_0(p),$$

откуда $\sigma_0(p)=0$. Как следствие, $S^\dagger(p)=p/2$ при $p\equiv 1\pmod 4$.

Если же $p \equiv 3 \pmod 4$, то

$$S^{\dagger}(p) = \frac{p}{2} + i(-1)^{(p^2-1)/8}\sigma_0(p)\sqrt{p}.$$

Теперь для получения (4.18) достаточно заметить, что

$$\operatorname{tg}\frac{\pi j^2}{p} = -2\operatorname{Im}\frac{1}{1+\zeta^{j^2}}$$

для каждого $j = 0, 1, \ldots, p-1$, и просуммировать эти равенства.

Ещё один похожий, но технически более сложный пример — это равенство

$$\sum_{j=0}^{p-1}\frac{1}{\cos^2\dfrac{\pi j^2}{p}} = \begin{cases} p^2 - (-1)^{(p^2-1)/8}8\sigma_1(p)\sqrt{p}, & \text{если } p \equiv 1 \pmod 4,\\ p^2, & \text{если } p \equiv 3 \pmod 4,\end{cases} \tag{4.19}$$

где

$$\sigma_1(p) = \sum_{b=1}^{(p-1)/2} b\left(\frac{b}{p}\right).$$

Для доказательства сумму в (4.19) нужно записать в виде

$$S^{\ddagger}(p) = \sum_{j=0}^{p-1}\frac{4\zeta^{j^2}}{(1+\zeta^{j^2})^2},$$

после чего аналогичным образом получить формулу

$$S^{\ddagger}(p) = p^2 - g(p)\left(1 + \left(\frac{-1}{p}\right)\right)\sum_{a=1}^{p-1}(-1)^a a\left(\frac{a}{p}\right).$$

Результат дальнейших преобразований таков:

$$\sum_{a=1}^{p-1}(-1)^a a\left(\frac{a}{p}\right) = 2\left(\frac{2}{p}\right)\left(1 + \left(\frac{-1}{p}\right)\right)\sigma_1(p) - p\left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\sigma_0(p).$$

Кроме того, если вычислять сумму

$$\sum_{a=1}^{p-1} a\left(\frac{a}{p}\right)$$

двумя способами, то можно получить равенство

$$\left(1 - \left(\frac{-1}{p}\right)\right)\sigma_1(p) + p\left(\frac{-1}{p}\right)\sigma_0(p) = 2\left(\left(\frac{2}{p}\right) - \left(\frac{-2}{p}\right)\right)\sigma_1(p) + p\left(\frac{-2}{p}\right)\sigma_0(p).$$

При $p \equiv 1 \pmod 4$ оно оказывается тривиальным, а при $p \equiv 3 \pmod 4$ позволяет выразить $\sigma_1(p)$ через $\sigma_0(p)$:

$$\sigma_1(p) = \begin{cases} \dfrac{p\sigma_0(p)}{3}, & \text{если } p \equiv 3 \pmod 8, \\ 0, & \text{если } p \equiv 7 \pmod 8. \end{cases}$$

Читателю предлагается восстановить детали этих рассуждений.

Впрочем, лучше рассуждать так. Полученное нами для $|r| < 1$ равенство

$$\sum_{j=0}^{p-1} \frac{1}{1 + r\zeta^{j^2}} = \frac{p}{1 + r^p} + \frac{g(p)}{1 + r^p} \sum_{a=1}^{p-1} (-1)^a r^a \left(\frac{a}{p}\right)$$

можно истолковать как равенство рациональных дробей от r (т. е. как тождество). Подставляя сюда $r = -1/z$ и затем сокращая на z, получим

$$\sum_{j=0}^{p-1} \frac{1}{z - \zeta^{j^2}} = \frac{D_p(z)}{z^p - 1}, \tag{4.20}$$

где введено обозначение

$$D_p(z) = pz^{p-1} + g(p) \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) z^{p-a-1}.$$

Далее тождеством (4.20) можно распорядиться стандартным образом — начать дифференцировать по z и подставлять $z = \pm 1$ с целью вычислить суммы вида

$$\sum_{j=0}^{p-1} \frac{1}{(1 + \zeta^{j^2})^s}, \quad \sum_{j=1}^{p-1} \frac{1}{(1 - \zeta^{j^2})^s},$$

где s — любое натуральное число. Вот третий пример (в дополнение к примерам (4.18) и (4.19)), который можно обнаружить на этом пути:

$$\sum_{j=1}^{p-1} \operatorname{ctg} \frac{\pi j^2}{p} = \begin{cases} 0, & \text{если } p \equiv 1 \pmod 4, \\ \dfrac{2\sigma_0(p)}{3\sqrt{p}}, & \text{если } p \equiv 3 \pmod 8, \\ 2\sigma_0(p)\sqrt{p}, & \text{если } p \equiv 7 \pmod 8. \end{cases}$$

Доказательство предоставляется читателю в качестве очередного упражнения.

Целые числа $\sigma_0(p)$ при $p > 3$ и $p \equiv 3 \pmod 4$ имеют весьма нетривиальный смысл (см. теорему 4 на с. 383 книги [3]): если h — число классов дивизоров мнимого квадратичного поля $\mathbb{Q}(\sqrt{-p})$, то

$$h = \begin{cases} \dfrac{\sigma_0(p)}{3}, & \text{если } p \equiv 3 \pmod 8, \\ \sigma_0(p), & \text{если } p \equiv 7 \pmod 8. \end{cases} \tag{4.21}$$

Как следствие, имеем $\sigma_0(p)>0$ для всех $p\equiv 3 \pmod 4$. Интересно отметить, что элементарное доказательство этого неравенства (не апеллирующее к равенству (4.21)), по-видимому, ещё не найдено. Относительно недавно[9] последовательность $\sigma_0(p)$ (где p пробегает ряд нечётных простых чисел) была зарегистрирована на хорошо известном сайте OEIS [28]. Что касается последовательности целых чисел $\sigma_1(p)$, то на данный момент запись об этой последовательности на сайте OEIS отсутствует.

Впечатляющее количество примеров тригонометрических тождеств типа (4.18) и (4.19) имеется в статье [19].

Список литературы

[1] *Айерлэнд К., Роузен М.* Классическое введение в современную теорию чисел. М.: Мир, 1987.

[2] *Андреев Н. Н., Юдин В. А.* Наименее уклоняющиеся от нуля многочлены и кубатурные формулы чебышевского типа // Тр. МИАН. 2001. Т. 232. С. 45–57.

[3] *Боревич З. И., Шафаревич И. Р.* Теория чисел. М.: Наука, 1985.

[4] *Винберг Э. Б.* Курс алгебры. М.: МЦНМО, 2017.

[5] *Гашков С.* Задача Чебышёва и тригонометрические многочлены // Квант. 1990. №6. С. 25–27.

[6] *Гашков С. Б.* Неравенство Бернштейна, тождество Рисса и формула Эйлера для ряда обратных квадратов // Математическое просвещение. Сер. 3. Вып. 18. М.: МЦНМО, 2014. С. 143–171.

[7] *Дэвенпорт Г.* Мультипликативная теория чисел. М.: Наука, 1971.

[8] Сборник задач по теории аналитических функций / Под ред. М. А. Евграфова. М.: Наука, 1972.

[9] *Нестеренко Ю. В.* Теория чисел. М.: Academia, 2008.

[10] *Понарин Я. П.* Алгебра комплексных чисел в геометрических задачах. М.: МЦНМО, 2014.

[11] *Прасолов В. В.* Многочлены. М.: МЦНМО, 2014.

[12] *Привалов И. И.* Введение в теорию функций комплексного переменного. М.: Наука, 1984.

[13] *Степанов С. А.* Арифметика алгебраических кривых. М.: Наука, 1991.

[14] *Шарыгин И. Ф.* Геометрия. Планиметрия. 9–11 кл.: От учебной задачи к творческой: Пособие для учащихся. М.: Дрофа, 2001.

[15] *Apostol T. M.* Introduction to analytic number theory. Springer, 1976.

[16] *Apostol T. M.* Modular functions and Dirichlet series in number theory. Springer, 1990.

[9] Для сравнения отметим, что тождество (4.18) было открыто В. Лебегом в 1850 году.

[17] *Berndt B. C., Evans R. J., Williams K. S.* Gauss and Jacobi sums. New York: John Wiley & Sons, Inc., 1998. (Canadian Mathematical Society Series of Monographs and Advanced Texts).

[18] *Berndt B. C., Yeap B. P.* Explicit evaluations and reciprocity theorems for finite trigonometric sums // Adv. in Appl. Math. 2002. Vol. 29, № 3. P. 358–385.

[19] *Berndt B. C., Zaharescu A.* Finite trigonometric sums and class numbers // Math. Ann. 2004. Vol. 330, № 4. P. 551–575.

[20] *Chamberland M.* Finite trigonometric product and sum identities // Fibonacci Quart. 2012. Vol. 50, № 3. P. 217–221.

[21] *Cohen E.* Trigonometric sums in elementary number theory // Amer. Math. Monthly. 1959. Vol. 66. P. 105–117.

[22] *da Fonseca C. M., Kowalenko V.* On a finite sum with powers of cosines // Appl. Anal. Discrete Math. 2013. Vol. 7. P. 354–377.

[23] *Grosswald E.* Dedekind — Rademacher sums // Amer. Math. Monthly. 1971. Vol. 78. P. 639–644.

[24] *Nathanson M. B.* Additive number theory: the classical bases. New York: Springer-Verlag, 1996. (Graduate Texts in Math.; Vol. 165).

[25] *Osipov N. N.* Problem 12003 // Amer. Math. Monthly. 2017. Vol. 124, № 8. P. 754.

[26] *Ramanujan S.* On certain trigonometrical sums and their applications in the theory of numbers [Trans. Cambridge Philos. Soc. 1918. Vol. 22, № 13. P. 259–276] // Collected papers of Srinivasa Ramanujan. Providence, RI: AMS Chelsea Publ., 2000. P. 179–199.

[27] *Schramm W.* The Fourier transform of functions of the greatest common divisor // Integers. 2008. Vol. 8. # A50.

[28] `http://oeis.org/A178153`

[29] `https://en.wikipedia.org/wiki/Arithmetic_function`

[30] `https://en.wikipedia.org/wiki/Inclusion-exclusion_principle`

Николай Николаевич Осипов,
Сибирский федеральный университет, Красноярск
nnosipov@rambler.ru

Нам пишут

Простое доказательство изопериметрической теоремы

В этом письме приводятся чёткая формулировка и короткое доказательство основного результата работы [1] Е. И. Алексеевой (см. ниже), а также проясняется его связь с изопериметрической теоремой для плоскости Лобачевского. Доказательство по сути не отличается от приведённого в [1]. Однако ввиду красоты и важности результата короткое доказательство, освобождённое от ненужных деталей, может быть интересно читателю.

Теорема 1 [1]. *На плоскости Лобачевского среди треугольников ABC с заданными длинами двух сторон AB и AC максимальную площадь имеет тот, у которого угол A равен сумме углов B и C.*

Изопериметрическая теорема для плоскости Лобачевского утверждает, что среди фигур данной площади, ограниченных спрямляемыми кривыми, наибольший периметр имеет круг. Эта теорема и её многомерные аналоги давно и хорошо известны [4]. Классическому рассуждению Я. Штейнера об изопериметрах (см., например, [2, с. 19–22] или [3, с. 30–31]) недостаёт именно теоремы 1, чтобы оно стало доказательством изопериметрической теоремы для плоскости Лобачевского. Действительно, в рассуждении

Это письмо составлено из отзывов В. О. Бугаенко и О. В. Шварцмана на работу [1], которые были переданы Е. И. Алексеевой в 2009 г. для улучшения этой работы, представленной на Московскую математическую конференцию школьников, см. www.mccme.ru/mmks. (Хотя рецензирование на ММКШ анонимное, рецензенты любезно согласились на публикацию настоящего письма с их именами.)

Штейнера в силу этой теоремы нужно рассматривать треугольники ABC с равенством $\angle A = \angle B + \angle C$. Это равенство означает, что на отрезке BC имеется точка D, для которой $\angle BAD = \angle B$, $\angle CAD = \angle C$, что эквивалентно равенствам $DB = DA = DC$. Таким образом, точка A лежит на окружности с диаметром BC, а это и требуется в рассуждении Штейнера. Поэтому мы не исключаем, что теорема 1 была известна специалистам или любителям элементарной математики, хотя никаких ссылок найти не удалось.

Доказательство теоремы 1. Обозначим через α, β, γ углы треугольника ABC. Воспользуемся моделью Пуанкаре в круге. Вершину A поместим в центр модели. Рассмотрим евклидову окружность ω и евклидову прямую, содержащие гиперболические прямые BC и AB соответственно. Они пересекаются в двух точках B и B' (рис. 1).

Докажем, что *площадь гиперболического треугольника ABC равна удвоенной величине евклидова угла $AB'C$*, который мы обозначим через τ. Действительно, угол между хордой BC и окружностью ω также равен τ, поскольку угол между хордой и касательной равен соответствующему вписанному углу. Так как сумма углов евклидова треугольника ABC равна $\alpha + \beta + \gamma + 2\tau = \pi$, имеем $S(ABC) = \pi - (\alpha + \beta + \gamma) = 2\tau$.

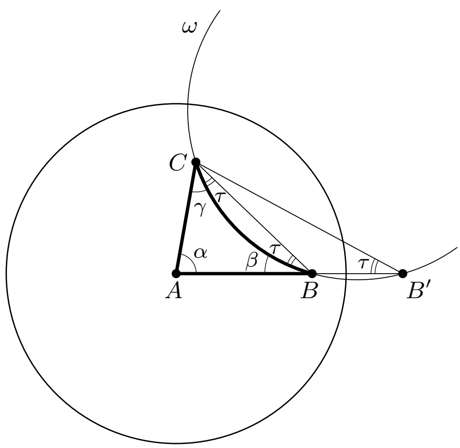


Рис. 1

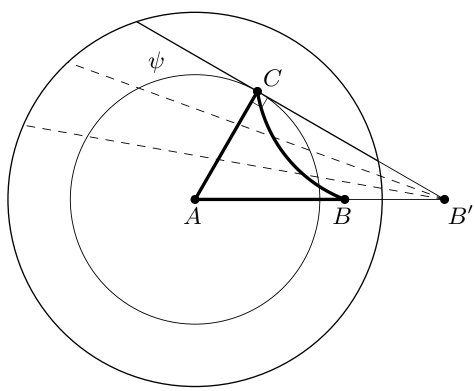


Рис. 2

Таким образом, треугольник ABC имеет максимальную площадь тогда и только тогда, когда угол $AB'C$ максимален. Поскольку длины сторон AB и AC фиксированы, а меняется лишь угол между ними, можно считать фиксированными точки A и B; тогда точка C может перемещаться по окружности ψ с центром A. Очевидно, что угол $AB'C$ максимален, если евклидова прямая $B'C$ касается окружности ψ (рис. 2).

Это, в свою очередь, означает, что евклидов угол ACB' прямой. Последнее условие равносильно тому, что $\pi/2 = \angle CAB' + \angle CB'A = \alpha + \tau$. Сопоставив это с выведенной ранее формулой $S(ABC) = 2\tau$ и формулой $S(ABC) = \pi - \alpha - \beta - \gamma$ для площади треугольника, получаем требуемое $\alpha = \beta + \gamma$. □

Список литературы

[1] *Алексеева Е.* Гиперболические треугольники максимальной площади с двумя заданными сторонами // Математическое просвещение. Сер. 3. Вып. 14 М.: МЦНМО, 2010. С. 175–183. См. также: *Alekseeva J. I.* Hyperbolic triangles of the maximum area with two fixed sides, `http://arxiv.org/abs/0911.5319`

[2] *Крыжановский Д. А.* Изопериметры. М.: Физматгиз, 1959.

[3] *Протасов В. Ю.* Максимумы и минимумы в геометрии. М.: МЦНМО, 2012.

[4] *Schmidt E.* Beweis der isoperimetrischen Eigenschaft der Kugel im hyperbolischen und sphärischen Raum jeder Dimensionenzahl // Math. Zeitschrift. 1943. Bd. 49. S. 1–109.

Аркадий Борисович Скопенков, МФТИ, НМУ

`http://www.mccme.ru/~skopenko`

Решение кубических уравнений с помощью трансцендентных функций

А. А. Гольдман

Мы приведём метод точного (или символьного) решения кубического уравнения с вещественными коэффициентами с помощью тригонометрических и гиперболических функций (в *неприводимом* и *приводимом* случаях соответственно), см. теоремы 1 и 2.

При решении кубических уравнений используются либо алгебраические функции (Алг.), либо трансцендентные (Трансц.). Для каждого случая приводимости и каждого типа функций есть свой способ решения:

	Неприводимый случай	Приводимый случай
Трансц.	Метод Виета	Гиперболические функции
Алг.	Метод дель Ферро (комплексный)	Метод дель Ферро

В обоих случаях приводимости корни выражаются через радикалы с помощью метода дель Ферро [3, 10]. Для решения уравнений в неприводимом случае с помощью трансцендентных функций используют метод Виета. В этой работе мы рассматриваем метод решения кубических уравнений в приводимом случае через трансцендентные функции, а именно через гиперболические функции.

Более общий результат Ритта см. в [2, 6]. Для него уже нужны существенно более сложные идеи и методы.

Метод Виета хорошо известен в учебной литературе [1, 3–5, 7, 10, 11]. В этих источниках его формулировка отличается от предложенной далее. Метод гиперболических функций, описываемый в данной работе, также известен. Формулировка метода Виета, предложенная в работе, также может быть найдена в интернете [8].

Перейдём к точным формулировкам.

Уравнение $ay^3 + by^2 + cy + d = 0$ сводится к уравнению вида

$$x^3 + px + q = 0$$

путём деления на a и замены $y = x - b/(3a)$. При $p = 0$ решение очевидно. *Далее считаем* $p \neq 0$. Обозначим через

$$D = -108\Big(\Big(\frac{p}{3}\Big)^3 + \Big(\frac{q}{2}\Big)^2\Big)$$

дискриминант кубического уравнения $x^3 + px + q = 0$. В зависимости от его знака уравнение решается по-разному:

- при $D > 0$ (*неприводимый случай*) методом Виета, использующим тригонометрические функции (см. теорему 1);
- при $D < 0$ (*приводимый случай*) аналогично, используя гиперболические функции (см. теорему 2);
- при $D = 0$ применимы оба способа.

При $D < 0$ уравнение имеет один вещественный корень, при $D > 0$ оно имеет три различных вещественных корня, при $D = 0$ все три корня вещественные и среди них есть равные [9, с. 142].

Положим $Q = \dfrac{3\sqrt{3}q}{2p\sqrt{|p|}}$.

Теорема 1 (Виет). *Пусть p и q — вещественные числа, $p \neq 0$, $D \geqslant 0$. Тогда уравнение $x^3 + px + q = 0$ имеет вещественные корни*

$$x_k = \sqrt{\frac{-4p}{3}}\cos\Big(\frac{\arccos Q + 2\pi k}{3}\Big), \quad \text{где} \quad k = -1, 0, 1.$$

Напомним определения функций, используемых в следующей теореме:

$$\operatorname{ch} x = \frac{e^x + e^{-x}}{2}, \quad \operatorname{sh} x = \frac{e^x - e^{-x}}{2};$$
$$\operatorname{arch} b = a, \text{ где } \operatorname{ch} a = b \text{ и } a \geqslant 0; \quad \operatorname{arsh} b = a, \text{ где } \operatorname{sh} a = b.$$

Теорема 2. *Пусть p и q — вещественные числа и $D < 0$. Тогда уравнение $x^3 + px + q = 0$ имеет ровно один вещественный корень*

$$x = \begin{cases} -\operatorname{sgn} q\sqrt{\dfrac{-4p}{3}}\operatorname{ch}\dfrac{\operatorname{arsh} Q}{3}, & p < 0, \\ -\sqrt{\dfrac{4p}{3}}\operatorname{sh}\dfrac{\operatorname{arsh} Q}{3}, & p > 0. \end{cases}$$

Приведём доказательство только для теоремы 2. Теорема 1 доказывается аналогично.

Доказательство теоремы 2. 1) $p<0$. Сделав замену $x=\sqrt{\frac{-4p}{3}}y$, перейдём к уравнению $4y^3-3y=Q$. Функция $\operatorname{ch} x$ принимает значения на $[1;+\infty)$. Замена на $\operatorname{ch} x$ возможна, если $Q\geqslant 1$. При $D<0$ и $p<0$ это условие выполняется, если $q<0$. Поэтому делаем замену $y=-\operatorname{sgn} q \operatorname{ch} t$, где $t\geqslant 0$. Как известно, $\operatorname{ch} 3x=4\operatorname{ch}^3 x-3\operatorname{ch} x$, поэтому $\operatorname{ch} 3t=Q$, отсюда $t=\frac{\operatorname{arch} Q}{3}$. Значит,

$$y=-\operatorname{sgn} q \operatorname{ch}\frac{\operatorname{arch} Q}{3} \quad \text{и} \quad x=-\operatorname{sgn} q\sqrt{\frac{-4p}{3}}\operatorname{ch}\frac{\operatorname{arch} Q}{3}.$$

2) $p>0$. Сделав замену $x=\sqrt{\frac{4p}{3}}y$, перейдём к уравнению $4y^3+3y=-Q$. Теперь, заменив y на $\operatorname{sh} t$ (функция $\operatorname{sh} x$ принимает значения на $\mathbb{R}$), имеем $\operatorname{sh} 3x=4\operatorname{sh}^3 x+3\operatorname{sh} x$, поэтому $\operatorname{sh} 3t=-Q$, отсюда $t=-\frac{\operatorname{arsh} Q}{3}$. Значит,

$$y=-\operatorname{sh}\frac{\operatorname{arsh} Q}{3} \quad \text{и} \quad x=-\sqrt{\frac{4p}{3}}\operatorname{sh}\frac{\operatorname{arsh} Q}{3}. \qquad \square$$

Список литературы

[1] *Гиндикин С.* «Великое искусство» // Квант. 1976. № 9. С. 2–10.

[2] *Хованский А. Г.* Полиномы Чебышёва и их обращения // Математическое просвещение. Сер. 3. Вып. 17. М.: МЦНМО, 2013. С. 93–106.

[3] *Клумова И., Фукс Д.* Формула существует, но... // Квант. 1976. № 9. С. 11–16.

[4] *Краснодемская А.* Графическое решение кубических уравнений // Квант. 1976. № 9. С. 18–19.

[5] *Резников А.* Формула Кардано и геометрия // Квант. 1976. № 9. С. 17.

[6] *Ritt J. F.* On algebraic functions which can be expressed in terms of radicals // Trans. AMS. Vol. 24, № 1. 1922. P. 21–30. https://arxiv.org/abs/1610.05968

[7] *Соловьёв Ю.* Вызов Ван Роумена // Квант. 1986. № 6. С. 18–20.

[8] https://ru.wikipedia.org/wiki/ТригонометрическаяформулаВиета

[9] *Винберг Э. Б.* Курс алгебры. М.: МЦНМО, 2017.

[10] *Янкелевич В.* «Неприводимый» случай // Квант. 1971. № 11. С. 20–21.

[11] Элементы математики в задачах: через олимпиады и кружки — к профессии / Под ред. А. А. Заславского, А. Б. Скопенкова и М. Б. Скопенкова. М.: МЦНМО, 2017.

Артур Андреевич Гольдман, факультет компьютерных наук НИУ ВШЭ
artur-goldman@rambler.ru

Задачник

(составитель А. Я. Канель-Белов)

Условия задач

В этом разделе вниманию читателей предлагается подборка задач разной степени сложности, в основном трудных. Надеемся, что эти задачи окажутся интересными для читателей «Математического просвещения», в том числе для сильных школьников, интересующихся математикой.

Мы обращаемся с просьбой ко всем читателям, имеющим собственные подборки таких задач, присылать их в редакцию. Мы с удовольствием будем публиковать свежие авторские задачи.

В скобках после условия задачи указывается автор (уточнения со стороны читателей приветствуются).

1. Лягушка совершает прыжки, каждый — на метр. Направление каждого прыжка выбирается случайно (считаем, что случайная величина, равная углу поворота, распределена равномерно на отрезке $[-\pi; \pi]$). С какой вероятностью после трёх прыжков лягушка окажется на расстоянии не больше 1 м от начальной точки?

   (*American Mathematical Contest, 2010*)

2. а) Монетка подбрасывается до тех пор, пока не выпадет две решки подряд. Уже сделано 9 бросков, и игра ещё не закончена. Какова вероятность, что десятый бросок окажется последним? (*В. К. Ковальджи*)

   б) Рассмотрим решётку $\Omega = \{(x, y, z) \in \mathbb{Z}^3 \colon y+1 \geqslant x \geqslant y \geqslant z \geqslant 0\}$. Лягушка прыгает по ней параллельно координатным осям, длина прыжка равна единице. Каково число возможных путей длины $3n$ из точки $(0, 0, 0)$ в точку (n, n, n)? (*А. Вершик, Ф. Петров*)

3. Может ли быть, что три человека, находящиеся на расстоянии 0, 1 и 2 от начала дороги, пройдут, не обгоняя друг друга, до точек, находящихся на расстоянии 1000, 1001 и 1002 от начала дороги так, чтобы

последний всё время видел первого, но ни в какой момент не видел второго (дорога идёт в одном направлении по горизонтали, но может подниматься и спускаться)? (*Н. Н. Константинов*)

4. Докажите, что середины сторон произвольного описанного четырёхугольника, отличного от квадрата, лежат на эллипсе, касающемся вписанной окружности в двух точках (возможно, мнимых). (*А. А. Заславский*)

5. а) Назовём два различных натуральных числа m и n *родственными*, если они имеют одни и те же простые делители, причём числа $m-1$ и $n-1$ обладают тем же свойством. Докажите, что существует бесконечно много пар родственных чисел. (*Фольклор*)

   б) Найдите все такие пары многочленов $P(x)$, $Q(x)$ с комплексными коэффициентами, что P делит Q^2+1, а Q делит P^2+1. (*IMC-2018, предложил R. Angelo*)

6. Все вершины выпуклого 10^9-угольника имеют целые координаты. Докажите, что его диаметр не меньше 10^{12}. (*А. Я. Канель-Белов*)

7. Даны многочлены $P(x), Q(x)$ с неотрицательными коэффициентами и старшим коэффициентом 1. Докажите, что все их коэффициенты равны 0 или 1, если $P(x)Q(x) = 1 + x + x^2 + \ldots + x^n$. (*Б. И. Каневский, В. А. Сендеров*)

8. При каких n любой треугольник можно разрезать на n равных треугольников? (*А. Ю. Сойфер*)

9. Пусть α, β — положительные числа. Рассмотрим такую симметрическую матрицу (a_{ij}), что
$$a_{ij} = \frac{1}{i+j+\alpha} \cdot \frac{1}{i+j+\beta}.$$
   Докажите, что эта матрица положительно определена. (*А. А. Логунов*)

10. а) Полицейский ловит Гангстера в городе, представляющем собой квадрат 10×10, разбитый улицами на квадратные клетки — кварталы. Полицейский видит Гангстера, только если на него натыкается, и оба они ездят только по улицам. Скорость Полицейского в 10 000 раз больше скорости Гангстера. Может ли Полицейский поймать Гангстера за ограниченное время? (*А. Я. Канель-Белов*)

    б) Тот же вопрос, если потребовать, чтобы путь Полицейского был *конечнозвенной ломаной*. (*А. Я. Канель-Белов*)

11. Найдите предел
$$\lim_{R\to\infty} \sum_{(a,b)\,\in\, D_R} \frac{(-1)^{a+b}}{a^2+b^2},$$
если D_R есть множество целочисленных точек (x, y), находящихся от нуля на положительном расстоянии, не превосходящем R.
(*IMC-2018, предложил R. Angelo*)

12. С переменной x и действительными числами разрешается провести не более 100 операций сложения, умножения и возведения в любую натуральную степень. Можно ли получить многочлен с любым данным числом действительных корней? (*Фольклор*)

Дополнение к задачнику

Хорошая задача обычно существует не сама по себе, она ценна своими связями. Наиболее содержательные из них открывают новые сюжеты и темы, в рамках которых возникают новые задачи, открываются новые грани. Именно поэтому их решение обогащает и оказывается столь полезным, помимо чисто интеллектуальной тренировки.

Публикуем дополнения к очередным задачам.

В выпуске 5 была опубликована

Задача 5.5. Докажите, что пересечение 10 правильных тетраэдров, вписанных в додекаэдр (вершины тетраэдров являются вершинами додекаэдра), есть икосаэдр. (*А. Я. Белов*)

Она связана с задачей, которую В. А. Сендеров дал на матбое между школами 2 и 179 в далёком 1979 году:

Задача 5.5′. Дан додекаэдр. Какое наименьшее число его движений нужно взять, чтобы любое движение додекаэдра можно было представить в виде их композиции? (*Фольклор*)

В выпуске 6 была опубликована

Задача 6.1. На пир собрались 100 людоедов, и они стали есть друг друга. Оказалось, что из любых 10 найдётся один, оказавшийся в желудке у другого. Докажите, что найдётся цепочка из 12 вложенных людоедов. (*В. А. Сендеров*)

В этой связи интересен следующий апгрейд задачи VI Турнира городов (осень 1984 г., основной вариант, 9–10 класс, № 4):

Задача 6.1′. Каждый вечер некоторые дамы дают приём (всего n дам), а остальные ходят на приём к тем дамам, которые в этот вечер принимают. Каково минимальное число вечеров, чтобы каждая дама попала к каждой? (*А. Я. Канель-Белов*)

В номере 8 была опубликована

Задача 8.11. Ряд $\sum a_n$ *сходится в среднем*, если существует предел средних арифметических его частичных сумм

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n} s_k}{n}, \quad s_k=\sum_{m=1}^{k} a_k.$$

Пусть ряд $\sum a_n$ сходится в среднем и при этом а) $a_n{=}o(1/n)$ (т. е. $\lim na_n{=}0$); б) $a_n = O(1/n)$ (т. е. $\exists\, C > 0\colon |na_n| < C$).

Докажите, что тогда ряд $\sum a_n$ сходится. (*А. Я. Белов*)

Естественно рассмотреть родственную ситуацию, а именно, когда вместо *сходимости в среднем* рассматривается существование предела

$$\lim_{r\to 1-}\sum_{n=1}^{\infty} a_n r^n. \tag{0.1}$$

Тогда возникает

Задача 8.11$'$. Предположим, что предел (0.1) существует. Докажите, что тогда ряд $\sum a_n$ сходится, если а) $a_n = o(1/n)$; б) $a_n = O(1/n)$.

В выпуске 11 была опубликована

Задача 11.4. d-Мерная ладья бьёт по прямым вдоль осей координат. Расстановку ладей в k-мерном кубе назовём *полной*, если ладьи не бьют друг друга и количество ладей максимально возможное.

а) Какое максимальное число ладей можно расставить в d-мерном кубе $n \times \ldots \times n$ так, чтобы они не били друг друга?

б) Слоем трёхмерного куба $n \times n \times n$ назовём квадрат $n \times n$, состоящий из клеток с одинаковой третьей координатой. Пусть первые k слоёв заполнены полно (т. е. в них стоят nk ладей). Докажите, что эту расстановку можно продолжить до полной расстановки всего куба. Верно ли аналогичное утверждение для четырёхмерного куба?

в) В трёхмерном кубе $n \times n \times n$ расставили ладьи и зафиксировали угловую клетку. Каково максимальное число подкубов с полной расстановкой и той же угловой клеткой? Аналогичный вопрос для d-мерного куба.

(*А. Я. Канель*)

В связи с этой задачей возникла

Задача 11.4$'$. Сколько полных расстановок ладей в трёхмерном кубе $4 \times 4 \times 4$? (*А. Ю. Эвнин*)

В выпуске 11 была также опубликована

Задача 11.9. Можно ли множество рациональных точек единичной сферы раскрасить в два цвета (чёрный и белый) так, что если три точки отвечают концам трёх ортогональных векторов, то одна из них будет чёрной, а две другие — белыми? (*Д. Муштари*)

Отметим, что множество **всех** точек единичной сферы так раскрасить нельзя (см. решение задачи 10.8 в предыдущем выпуске 22). Задача 10.8 содержится в книге Ю. И. Манина «Доказуемое и недоказуемое» (с. 89, 94–97) в связи с доказательством отсутствия скрытых параметров в квантовой механике.

В выпуске 22 опубликована также заметка Д. Х. Муштари «О правильной раскраске 16-мерной сферы» (с. 218–219), где показано, что не существует правильной раскраски множества рациональных точек 16-мерной единичной сферы (когда любые две точки, отвечающие перпендикулярным векторам, раскрашены в разные цвета). Здесь возникает семейство интересных задач.

Решение задачи 11.9 использует идеи, содержащиеся в заметке Д. Х. Муштари. С ней связана

Задача 11.9′. Существует ли вписанная в решётку нечётнозвенная замкнутая ломаная, все звенья которой имеют одинаковую длину? (*А. К. Ковальджи*)

В выпуске 14 была опубликована

Задача 14.6. $f(x, y)$ — бесконечно дифференцируемая функция от двух переменных с локальным минимумом в нуле. Других критических точек у неё нет. Верно ли, что этот минимум глобальный? (Точка называется *критической* [1]), если в ней обе частные производные $\partial f/\partial x$ и $\partial f/\partial y$ обращаются в нуль.) (*Фольклор*)

Данная задача в своё время была распространена среди учеников замечательного математика Н. А. Бобылева, увы, ныне покойного. Другая близкая задача, также распространённая среди его учеников, такова.

Задача 14.6′. Многочлен $P(x, y)$ от двух переменных принимает только положительные значения. Может ли он принимать все положительные значения? (*Фольклор*)

В номере 14 была также опубликована

Задача 14.12. По некоторым рёбрам клеток плоской решётки проведены перегородки. Пьяница с равной вероятностью идёт из квадрата, где он находится, в любой соседний квадрат, куда он может пройти. Докажите, что он с вероятностью 1 вернётся в исходную точку. (*А. Я. Канель, М. Б. Скопенков*)

---

[1]) Для бесконечно дифференцируемой функции.

Эта задача оказывается связана со следующей задачей.

Задача 14.12$'$. Из резисторов спаяли цепь. Может ли сопротивление цепи увеличиться, если, ничего не размыкая, к каким-то двум клеммам припаять ещё один резистор? (*Фольклор*)

В семнадцатом выпуске была опубликована

Задача 17.4. $\mathcal{A}$ — отображение плоскости в себя, сохраняющее расстояние (т. е. $|XY| = |\mathcal{A}(X)\mathcal{A}(Y)|$ для любых точек X, Y плоскости). Доказать, что $\mathcal{A}$ — отображение плоскости на себя (т. е. каждая точка имеет прообраз при этом отображении). (*Фольклор*)

С ней тесно связана

Задача 17.4$'$. $\mathcal{A}$ — отображение плоскости в себя, сохраняющее *единичные* расстояния (если $|XY| = 1$, то $|\mathcal{A}(X)\mathcal{A}(Y)| = |XY| = 1$). Доказать, что $\mathcal{A}$ сохраняет *все* расстояния. (*Фольклор*)

В выпуске 18 была опубликована

Задача 18.1. На плоскости начерчен угол величиной в n градусов, где $n < 180$ — натуральное число.

(а) Для каких n этот угол можно разделить с помощью циркуля и линейки на n равных углов?

(б) Пусть m — произвольное натуральное число. Для каких n угол A можно разделить с помощью циркуля и линейки на m равных углов?

(в) Пусть m и n — натуральные взаимно простые числа и $m/n < 180$, а k — ещё одно натуральное число. На плоскости начерчен угол A величиной в m/n градусов. Для каких k угол A можно разделить с помощью циркуля и линейки на k равных углов? (*Г. А. Гальперин*)

С ней связана весьма поучительная классическая задача:

Задача 18.1$'$. Какие правильные многоугольники можно вписать в (не обязательно квадратную) решётку? (*Фольклор*)

В прошлом выпуске 22 была опубликована

Задача 22.12. а) Любую ли фигуру из зеркального стекла можно осветить изнутри одной лампочкой?

б) Любой ли многоугольник из зеркального стекла можно осветить изнутри одной лампочкой? (*Фольклор*)

С этим сюжетом связана следующая

Задача 22.12$'$. Докажите, что луч света, бегающий в зеркальном эллипсе, касается либо некоторого эллипса, либо некоторой гиперболы. (*Фольклор*)

Решения задач из прошлых выпусков

11.4. Условие. d-Мерная ладья бьёт по прямым вдоль осей координат. Расстановку ладей в k-мерном кубе назовём *полной*, если ладьи не бьют друг друга и количество ладей максимально возможное.

а) Какое максимальное число ладей можно расставить в d-мерном кубе $n \times \ldots \times n$ так, чтобы они не били друг друга?

б) Слоем трёхмерного куба $n \times n \times n$ назовём квадрат $n \times n$, состоящий из клеток с одинаковой третьей координатой. Пусть первые k слоёв заполнены полно (т. е. в них стоят nk ладей). Докажите, что эту расстановку можно продолжить до полной расстановки всего куба. Верно ли аналогичное утверждение для четырёхмерного куба?

в) В трёхмерном кубе $n \times n \times n$ расставили ладьи и зафиксировали угловую клетку. Каково максимальное число подкубов с полной расстановкой и той же угловой клеткой? Аналогичный вопрос для d-мерного куба.

(*А. Я. Канель*)

а) Ответ. k^{n-1}.

Решение. Оценка. Легко видеть, что больше расставить нельзя: рассмотрим проекцию на $(n-1)$-мерную грань. Каждая её клетка будет покрыта проекциями ладей не более одного раза — иначе две ладьи окажутся на одной вертикали и будут бить друг друга.

Пример. Начнём со стандартного олимпиадного решения для трёхмерного куба с ребром 8. Цифра обозначает номер слоя, в котором стоит ладья над данной клеткой:

$$\begin{vmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 7 & 0 & 1 & 2 & 3 & 4 & 5 \\ 5 & 6 & 7 & 0 & 1 & 2 & 3 & 4 \\ 4 & 5 & 6 & 7 & 0 & 1 & 2 & 3 \\ 3 & 4 & 5 & 6 & 7 & 0 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 & 7 & 0 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 0 \end{vmatrix}.$$

Здесь при переходе к следующему слою происходит циклический сдвиг. Аналогичным образом, в четырёхмерном пространстве при переходе к следующему трёхмерному слою со всеми его двумерными слоями также происходит циклический сдвиг и т. д. Мы предоставляем это воображению читателя и приведём изящное описание этой конструкции.

Разместим ладьи в точках с координатами $x_1, \ldots, x_n$, где

$$\sum x_i \equiv 0 \mod k.$$

Легко видеть, что указанное расположение ладей удовлетворяет условию задачи (и соответствует чисто комбинаторному описанию, данному выше).

Следующее упражнение относится к кодам, регистрирующим ошибки. Последовательность битов означает вершину куба. Если в одном бите случается ошибка, то происходит сдвиг на ребро. Можно отметить половину вершин куба, попарно не соединённых рёбрами (кстати, ровно двумя способами), а больше — нельзя.

Упражнения.

1. В n-мерном кубе покрашено более половины вершин. Ребро называется *покрашенным*, если покрашены обе ограничивающие его вершины. Докажите, что покрашено не менее n рёбер.
2. Докажите, что минимальное количество ладей в n-мерном кубе, которое может бить все его клетки, составляет $\lceil n^2/2 \rceil$, т. е. наименьшее целое число, не меньшее чем $n^2/2$.

б) Решение. Пусть k слоёв трёхмерного куба $n \times n \times n$ полно заполнены. Нам достаточно показать, что можно полно заполнить ещё один слой. В каждой горизонтали нового слоя будет k запрещённых (соответственно, $n - k$ разрешённых) позиций, куда можно ставить ладью; то же верно и для столбцов. Задача сводится к известной *теореме Холла о паросочетаниях*.

Аналогичный результат для куба размерности больше 3 нам неизвестен. Мы будем признательны читателям за соображения на эту тему.

Упражнения.

1. Сто сумасшедших красят доску 100×100 в 100 цветов, соблюдая единственное правило: в одной строке и в одном столбце не может быть двух клеток, раскрашенных одинаково. Могут ли сумасшедшие правильно раскрасить доску, если уже раскрашено 100 клеток?
2. Единичный квадрат разбит двумя способами на n равновеликих многоугольников. Докажите, что можно выбрать n точек так, что в каждой из полученных $2n$ частей окажется по одной точке.
3. (*Н. С. Келлин*) Сто замов получили взыскания от ста завов. При этом каждый зав наложил по одному взысканию на 15 замов, а каждый зам

получил по одному взысканию от 15 завов. Докажите, что директор может снять 1400 взысканий так, что у каждого зама останется по одному взысканию и все взыскания будут наложены разными завами.

в) Ответ. Размеры двух подкубов с полными расстановками ладей и общим углом должны различаться не менее чем в два раза. Тогда кубов с полной расстановкой ладей будет не более чем $\log_2 n$.

Решение. Отметим, что для любого s любой s-мерный клетчатый слой k-мерного куба с полной расстановкой ладей сам имеет полную расстановку ладей (их в нём будет n^{s-1}). Взяв $s = 2$, сводим задачу к следующей очевидной лемме.

Лемма. *Пусть в углу квадрата $n \times n$ отмечен клеточный подквадрат $m \times m$, где $n/2 < m < n$. Тогда невозможно в большом квадрате $n \times n$ вне подквадрата $m \times m$ поместить n ладей так, чтобы они не били друг друга (максимальное число ладей, которые можно поместить, равно $2(n - m)$).*

Можно привести пример k-мерного куба с ребром 2^n, в котором имеется n подкубов с общим углом и полной расстановкой ладей. Для этого напомним определение *побитового сложения*. Если даны два числа M и N в двоичной записи, то результатом побитового сложения (XOR) $L = M \oplus N$ будет такое L, i-я ($i = 1, 2, \ldots$) двоичная цифра которого равна 0, если i-е цифры в M и N одинаковы, и 1, если они разные.

Искомая расстановка задаётся уравнением, немного отличающимся от решения п. а): $\bigoplus_i x_i = 0$. Наглядно её можно представить по индуктивной конструкции: куб с ребром 2^n разбивается на 2^k кубиков с ребром 2^{n-1}, которые красятся в шахматном порядке. В чёрных осуществляется уже построенная полная расстановка ладей, а белые остаются пустыми.

(*А. Я. Канель*)

11.5. Условие. На плоскости отметили n непересекающихся отрезков и $n + 2$ точки, которые не лежат на этих отрезках. Докажите, что найдутся две точки, которые «видят» друг друга (т. е. соединяющий их отрезок не пересекает отмеченные отрезки). (*М. Л. Концевич*)

Решение. Прежде всего, можно считать отрезки находящимися в *общем положении*, т. е. никакие три прямые, на которых лежат эти отрезки, не пересекаются в одной точке, никакие две не параллельны и никакая отмеченная точка не лежит на этих прямых. Это достигается малым шевелением. Далее, будем продлевать отрезки до пересечения друг с другом. Если при этом конец отрезка упирается в другой отрезок, то движение этого конца останавливается. Если при этом расширении точки «видят» друг друга, то они «видели» друг друга и раньше.

В результате получается граф на плоскости с $2n$ вершинами (от каждого отрезка по 2 конца), все вершины которого имеют степень 3 или 1. Следовательно, у него не более $3n$ рёбер и, с учётом формулы Эйлера, не более $n+1$ клетки. В одну из них попадут две точки, и они будут «видеть» друг друга.

Комментарий. Данная задача связана с вычислением средней площади части, на которую разбивают плоскость случайно расположенные трещины в процессе разрастания. При разрастании конец трещины, когда она упирается в другую, перестаёт двигаться. Средняя площадь части обратна среднему числу частей в пересчёте на единицу площади, а это среднее число и оказывается примерно равным среднему числу трещин в единице площади, как видно из решения задачи 11.5.

Обобщение результата на размерности выше 2 неизвестно.

Упражнение. а) Докажите, что все грани выпуклого многогранника не могут одновременно иметь больше 5 сторон. Кроме того, все вершины не могут одновременно иметь степень выше 5.

б) Треугольник разбит на выпуклые четырёхугольники. Докажите, что найдётся вершина разбиения степени 3.

в) Выпуклый 2018-угольник разбит на выпуклые семиугольники. Докажите, что четыре стороны одного из них выходят на границу.

г) Плоскость разбита на выпуклые семиугольники единичного диаметра. Докажите, что внутри круга радиуса 100 их не меньше миллиарда.

Список литературы

[1] *Матерон Ж.* Случайные множества и интегральная геометрия. М.: Мир, 1978.

[2] *Сантало Л. А.* Интегральная геометрия и геометрические вероятности. М.: Наука, 1983.

(*А. Я. Канель-Белов*)

13.4. Условие. Для каких $\lambda \in [0,1]$ для любой непрерывной функции $f\colon [0,1] \to \mathbb{R}$, такой, что $f(0)=f(1)$, обязательно найдётся такое $x \in [0, 1-\lambda]$, что $f(x) = f(x+\lambda)$? (*Фольклор*)

Ответ. Для всех λ вида $\dfrac{1}{n}$, где $n \in \mathbb{N}$, и только для них.

Решение. Переформулируем вопрос: Для каких $\lambda \in [0,1]$ для любой непрерывной функции $f\colon [0,1] \to \mathbb{R}$, такой, что $f(0) = f(1)$, обязательно найдётся горизонтальная хорда длины λ, концы которой лежат на графике этой функции?

Определим функцию $g_\alpha(x)$ следующим образом: $g_\alpha(x)=f(x)-f(x+\alpha)$. Покажем, что

$$\forall\, n\in\mathbb{N}\ \exists\, x\in\Big[0, 1-\frac{1}{n}\Big]\colon g_{1/n}(x)=0.$$

Рассмотрим

$$g_{1/n}(0),\ g_{1/n}\Big(\frac{1}{n}\Big),\ g_{1/n}\Big(\frac{2}{n}\Big),\ \ldots,\ g_{1/n}\Big(\frac{n-1}{n}\Big).$$

Без ограничения общности считаем, что $g_{1/n}(0)\geqslant 0$.

Если $g_{1/n}(0)=0$, то утверждение доказано.

Пусть $g_{1/n}(0)>0$. Заметим, что

$$\sum_{k=0}^{n-1} g_{1/n}\Big(\frac{k}{n}\Big)=f(0)-f\Big(\frac{1}{n}\Big)+f\Big(\frac{1}{n}\Big)-f\Big(\frac{2}{n}\Big)+f\Big(\frac{n-1}{n}\Big)-f(1)=f(0)-f(1)=0.$$

Так как $g_{1/n}(0)>0$, то $\exists\, k\in\mathbb{N}\colon g_{1/n}\Big(\dfrac{k}{n}\Big)<0$. Но тогда по теореме о промежуточном значении

$$\exists\, x\in\Big[0, \frac{k}{n}\Big]\colon g_{1/n}(x)=0, \quad \text{т.е.} \quad \exists\, x\in\Big[0, 1-\frac{1}{n}\Big]\colon f(x)=f\Big(x+\frac{1}{n}\Big),$$

что и требовалось.

Докажем, что при $\lambda\in[0,1]\setminus\Big\{\dfrac{1}{n}\mid n\in\mathbb{N}\Big\}$ ответ отрицательный. Положим $k=[1/\lambda]$, $\delta=1-kc$. Построим на плоскости ломаную, состоящую из двух семейств параллельных отрезков. Выберем произвольное $a\in(-1/k,0)$. Соединим отрезком точки $[0,0]$ и $[\delta,-ka]$, а другим отрезком — точки $[\delta,-ka]$ и $[\lambda,a]$. Построенный участок перенесём параллельно себе так, чтобы начальная точка перенесённого участка совпала с концом исходного. Будем действовать аналогично, пока не достроим ломаную до точки с абсциссой $k\lambda$. На участке $[k\lambda,1]$ построим отрезок, соединяющий точки $(k\lambda,ka)$ и $(1,0)$. Он будет параллелен и равен (с точностью до параллельного переноса) отрезкам первого семейства.

Построенная ломаная является графиком некоторой функции

$$h(x)\colon[0,1]\to[-1,1].$$

Из построения следует, что если $x, x+\lambda\in[0,1]$, то $h(x+\lambda)=h(x)+a$. Это означает, что не существует горизонтальной хорды длины λ с концами на графике. Тогда такой хорды нет и у функции

$$f(x)=\frac{h(x)+1}{2}\colon[0,1]\to[0,1],$$

что и требовалось. (*М. Третьяков, Г. Юшков, ученики 10 класса школы № 57 г. Москвы*)

13.7. Условие. Существует ли множество из $2(2n-1)$ точек на плоскости, не лежащих на одной прямой, которое можно разбить на $2n-1$ пару точек так, чтобы любая прямая, проходящая через точки из разных пар, проходила бы ещё через одну точку этого множества? (*Фольклор*)

Ответ. Да.

Решение. Рассмотрим правильный $(2n-1)$-угольник на плоскости и ещё $2n-1$ точек на бесконечно удалённой прямой, отвечающих направлениям его сторон. Отметим, что, поскольку $2n-1$ число нечётное, каждая диагональ параллельна ровно одной стороне. Кроме того, диагоналям, проведённым из вершины A, отвечают все возможные направления сторон, кроме направления стороны, противоположной вершине A. Сопоставим каждой вершине точку бесконечно удалённой прямой, отвечающую направлению противоположной стороны. Получится искомое расположение на проективной плоскости.

Переведя проективным преобразованием бесконечно удалённую прямую в обычную, сохранив все прямые, соединяющие наши точки, и сами наши точки в качестве обычных (не бесконечно удалённых), мы получим требуемую конфигурацию точек и на обычной (аффинной) плоскости.

Комментарий. При конструировании часто бывает полезно уносить объект на бесконечность. (*А. Я. Канель-Белов*)

15.1. Условие. «Задачи на устный счёт».

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 погрешностью 20 %. (*В. И. Арнольд*)

а) Ответ. 2.

Решение. Как известно, $2^{10}=1024$ или $2^{10}/1000=1{,}024$, и нам достаточно найти первую цифру числа $1{,}024^{40}$. Заметим, что

$$1{,}025^{40}=\left(1+\frac{1}{40}\right)^{40}\approx e=2{,}718281828\ldots$$

При этом

$$\left(\frac{1{,}024}{1{,}025}\right)^{40}=\left(1-\frac{1}{1025}\right)^{40}\approx 0{,}96.$$

Итак, первая цифра равна 2, как и у числа e.

Упражнение (*В. О. Бугаенко*). При переходе улицы в неположенном месте вероятность быть сбитым машиной равна 0,1 %. Хулиган Вася в течение 3 лет каждый день переходит улицу в неположенном месте. Оцените его шансы остаться в живых.

б) Ответ. $[2^{\sqrt{15}}] = 14$.

Решение. Покажем, что на самом деле $14 < 2^{\sqrt{15}} < 15$.

Воспользуемся формулой бинома Ньютона для показателя $1/2$:

$$\sqrt{15} = 4 \cdot \sqrt{1 - \frac{1}{16}} = 4 \cdot \left(1 - \frac{1}{32} - \frac{1}{2048} - \ldots\right),$$

откуда

$$2^{\sqrt{15}} = 2^{4 - \frac{1}{8} - \frac{1}{512} - \cdots}, \quad 2^{\sqrt{15}}\, 2^{\frac{1}{8} + \frac{1}{512} + \cdots} = 16. \qquad (*)$$

Второй сомножитель меньше $2^{\frac{1}{8} + \frac{1}{256}} = 2^{\frac{33}{256}}$, а это выражение меньше 1,1, так как по формуле бинома Ньютона

$$1{,}1^{\frac{256}{33}} = 1 + 0{,}1 \cdot \frac{256}{33} + 0{,}01 \cdot \frac{256}{33} \cdot \frac{223}{33} \cdot \frac{1}{2} + \ldots = 1 + \frac{256}{330}\left(1 + \frac{223}{660}\right) > 1 + \frac{3}{4} \cdot \frac{4}{3} = 2.$$

С другой стороны, $14 \cdot 1{,}1 < 16$. Значит, $2^{\sqrt{15}} > 14$.

Заметим теперь, что

$$2^{1/8} > 1 + \frac{1}{8} - \frac{7}{128} > 1{,}07,$$

и ввиду $(*)$ имеем $2^{\sqrt{15}} \cdot 1{,}07 < 16$. В то же время $15 \cdot 1{,}07 > 16$. Значит, $2^{\sqrt{15}} < 15$.

в) Ответ. $\sqrt[3]{60} < 2 + \sqrt[3]{7}$.

Решение. Заметим, что $\sqrt[3]{60} = 4\sqrt[3]{1 - \frac{1}{16}}$ и $2 + \sqrt[3]{7} = 2 + 2\sqrt[3]{1 - \frac{1}{8}}$.

Разлагая оба выражения в ряд, имеем, что первые два главных члена совпадают, а третий член уже оказывается существенно больше у второго выражения (перед этим членом стоит знак «плюс»), и это различие не может быть скомпенсировано остальными членами.

г) Ответ. $\frac{\sqrt{2}}{2}$.

Решение. Заметим, что

$$\int\limits_0^{2\pi} \sin^{100} x\, dx = \int\limits_0^{2\pi} \cos^{100} x\, dx.$$

Этот интеграл надо рассмотреть в окрестности двух точек максимума, где $\cos x = \pm 1$, и в этих точках его поведение одинаково. В окрестности нуля $\cos x$ и $e^{-x^2/2}$ совпадают с точностью до членов четвёртого порядка, поэтому наш интеграл примерно равен выражению

$$2\int\limits_{-\infty}^{+\infty} e^{-50x^2} dx = 2\sqrt{\frac{\pi}{50}} = \sqrt{\frac{4\pi}{50}} \approx \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}.$$

Пренебречь членами в выражениях для $\cos x$ и $e^{-x^2/2}$ порядка 4 и выше можно в силу «правила трёх сигм»: 98 % площади под графиком функции e^{-x^2/σ^2} заключено в промежутке $[-3\sigma, +3\sigma]$, а в этих пределах остальные члены несущественны (у нас $\sigma^2 = 1/50$). (*А. Я. Белов*)

Упражнения. Латинскими буквами обозначаются вещественные числа. Если условие задачи является утверждением, то задача состоит в том, чтобы это утверждение доказать.

1. Если $x, y \geqslant \sqrt{2}$, то $\dfrac{1}{\sqrt{1+x^2}} + \dfrac{1}{\sqrt{1+y^2}} \geqslant \dfrac{2}{\sqrt{1+xy}}$.

Указание. Замена $x = e^{u/2}$, $y = e^{v/2}$ переводит неравенство в

$$\frac{f(u)+f(v)}{2} \geqslant f\Big(\frac{u+v}{2}\Big),$$

где $f(t) = (1+e^t)^{-1/2}$, $e^u \geqslant 2$, $e^v \geqslant 2$. Но $f''(t) > 0$ при $e^t > 2$.

2. $x^4 + x^3 + x^2 + x + \dfrac{1}{3} > 0$.

Указание. $x^2\Big(x^2 + x + \dfrac{1}{4}\Big) + \Big(\dfrac{3}{4}x^2 + x + \dfrac{1}{3}\Big) = \ldots + \dfrac{3}{4}\Big(x^2 + 2\cdot\dfrac{2}{3}x + \dfrac{4}{9}\Big)$.

3. Если $|x-a| + |y-b| < c$, то $|xy - ab| < \Big(|a| + |b| + \dfrac{c}{4}\Big)c$.

Указание. Положим $z := x - a$ и $t := y - b$. Тогда $4|zt| \leqslant (|z| + |t|)^2 < c^2$.

4. Что больше: $\log_2 5$ или $\log_3 13$?

Указание. $\log_2 5 < 7/3 < \log_3 13$.

5. а) $\operatorname{ctg} x + \dfrac{2}{\pi} > \dfrac{1}{x}$ при $0 < x < \dfrac{\pi}{2}$; б) $1 > \dfrac{2x}{\pi \sin x} > e^{\frac{2x}{\pi}-1}$ при $0 < x < \dfrac{\pi}{2}$.

Указание. а) Рассмотрите $f(x) = \operatorname{ctg} x + \dfrac{2}{\pi} - \dfrac{1}{x}$. Имеем $f(\pi/2) = 0$ и $f'(x) < 0$ при $0 < x < \pi/2$.

б) Подумайте о геометрическом смысле левого неравенства. Для доказательства правого заметьте, что если $f > 0$, $g > 0$, то $f > g$ равносильно $\ln f > \ln g$. Далее воспользуйтесь методом решения и результатом п. а).

6. а) Избавьтесь от иррациональности в знаменателе дроби $\dfrac{1}{\sqrt[3]{2}+\sqrt[3]{3}+\sqrt[3]{4}}$.
б) Что больше: $\sqrt[3]{413}$ или $6+\sqrt[3]{3}$?

а) Указание. Воспользуйтесь равенством

$$a^3 + b^3 + c^3 - 3abc = (a+b+c)(a^2 + b^2 + c^2 - ab - bc - ca).$$

б) Ответ. $6+\sqrt[3]{3}$.

Первое решение. См. указание к п. а). Если хотя бы два из трёх чисел a, b, c различны, то знаки чисел $a+b+c$ и $a^3+b^3+c^3-3abc$ совпадают.

Второе решение. Поскольку $6=\sqrt[3]{216}$, задача эквивалентна сравнению чисел $\sqrt[3]{413}$ и $f\left(\frac{213}{2}\right)$, где

$$f(x)=\sqrt[3]{\frac{219}{2}+x}+\sqrt[3]{\frac{219}{2}-x}.$$

На отрезке $\left[0, \frac{219}{2}\right]$ функция $f(x)$ монотонно убывает, и

$$f(0)>\sqrt[3]{413}>f\left(\frac{219}{2}\right).$$

Поэтому достаточно найти такое число x, что $f(x)=\sqrt[3]{413}$. Решая уравнение

$$\sqrt[3]{a+x}+\sqrt[3]{a-x}=\sqrt[3]{b},$$

получаем $x^2=a^2-\dfrac{(b-2a)^3}{27b}$, откуда следует ответ.

Третье решение. Сравним 413 и $219+18t^2+108t$, где $t=\sqrt[3]{3}$. Это значит, что сравниваются числа 0 и $P_2(t):=18t^2+108t-194$, т. е. α и t, где α — положительный корень многочлена P_2.

Замечание. Третье решение опирается на рациональность одного из слагаемых. Однако сравнение чисел a и $(\sqrt[3]{b}+\sqrt[3]{c})^3$, где a, b, $c\in\mathbb{Q}$, легко приводится к сравнению чисел $\sqrt[3]{x}$ и $y+\sqrt[3]{z}$, где x, y, $z\in\mathbb{Q}$. Аналогично для рациональных a, b, c при $a>0$ можно выяснить, что больше: $\sqrt[n]{a}$ или $b+\sqrt[3]{c}$.

7. Что больше: а) 300! или 200^{300}? б) 300! или 100^{300}?

а) Ответ. 200^{300}. Указание. $2k/3>(k!)^{1/k}$ при $k\geqslant 3$.

б) Ответ. 300!. Указание. С помощью неравенства $(1+1/k)^k<3$ докажите по индукции, что $(k!)^{1/k}>k/3$. (*В. А. Сендеров*)

15.2. Условие. Может ли сумма двух периодических функций с минимальными периодами 1 и $\sqrt{2}$ снова быть периодической функцией? (*Фольклор*)

Ответ. Да, может.

Решение. Определим следующие функции $f(x)$, $g(x)$.

$$f(a+b\sqrt{2})=b,\ \text{если } a,b\in\mathbb{Z},\quad f(x)=0 \text{ иначе};$$
$$g(a+b\sqrt{2})=a,\ \text{если } a,b\in\mathbb{Z},\quad g(x)=0 \text{ иначе}.$$

Тогда минимальным периодом функции f является 1, а минимальным периодом функции g является $\sqrt{2}$. Сумма функций $f+g$ имеет вид:

$$(f+g)(a+b\sqrt{2})=a+b,\ \text{где } a,b\in\mathbb{Z},\quad (f+g)(x)=0 \text{ иначе}.$$

Тогда число $\sqrt{2}-1$ является периодом функции $f+g$ (проверяется непосредственно). (*Г. Юшков, ученик 10 класса школы № 57 г. Москвы*)

15.4. Условие. Во все точки целочисленной решётки на плоскости вбиты гвозди. На плоскость положили отрезок длины 2011, не задевающий ни одного из 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*)

а) Ответ. Можно.

Решение. Будем поворачивать отрезок относительно начала по часовой стрелке, пока не коснёмся (т. е. не подойдём к ней «бесконечно близко») точки решётки. Если при этом коснулись только одной, то будем двигать его дальше, пока не коснёмся второй.

Легко видеть, что, повернув чуть-чуть и перенеся отрезок параллельно самому себе, отрезок можно вывести целиком в нужную сторону относительно системы коснувшихся гвоздей и продолжить этот процесс, повернув вектор дальше по часовой стрелке. Процесс закончится, поскольку всякий раз происходит переход от близости к одному направлению, заданному парой точек решётки, находящихся на расстоянии не больше 2011, к другому направлению. А таких направлений конечное число.

б) Лемма. *Существует такое R, что в кольце $R \leqslant r \leqslant \sqrt{R^2+2011}$ с центром в начале координат нет точек целочисленной решётки.*

Доказательство. Пусть $p_1, p_2, \ldots, p_{2012}$ — первые 2012 простых чисел вида $4k+3$. По китайской теореме об остатках существует такое натуральное число N, что $N+j$ $(j=1,\ldots,2012)$ делится на p_j и не делится на p_j^2. Как известно, натуральное число является суммой двух квадратов целых чисел тогда и только тогда, когда любой его простой делитель вида $4k+3$ входит в его разложение в чётной степени. Поэтому $R=\sqrt{N+1}$ удовлетворяет условиям леммы. □

Поместим теперь начало отрезка в точку $(R,0)$, а конец поместим в точку $(R,\sqrt{2011})$. Тогда при вращении его относительно начала координат $(0,0)$ он весь будет находиться в кольце, указанном в лемме, и не заденет целочисленной точки.

Упражнения.

1. Решите уравнение в целых числах: $x^2+4=y^3$. (*В. А. Сендеров*)

2. Докажите, что для любого k существует окружность с центром в начале координат, на которой лежит ровно $4k$ точек с целыми координатами. (Задача допускает решение без использования гауссовых чисел.)

(*А. Я. Белов*)

17.2. Условие. а) Найти 300-ю цифру после запятой числа $\sqrt[3]{0{,}\underbrace{99\ldots9}_{100}}$.

б) С помощью калькулятора найти первую цифру числа 2^{10^6}.

(*А. Я. Белов*)

Решение. а) $\sqrt[3]{0{,}\underbrace{99\ldots9}_{100}} = \sqrt[3]{1-10^{-100}}$. Разлагая последнее выражение в ряд по формуле бинома Ньютона[1)], имеем:

$$\sqrt[3]{1-10^{-100}} = 1 - \frac{10^{-100}}{3} - \frac{10^{-200}}{9} - \frac{5\cdot 10^{-300}}{81} - \frac{10\cdot 10^{-400}}{243} - \ldots$$

Из разложения видно, что эта цифра равна 5.

б) Первая цифра числа $M = 2^{10^6}$ равна k, если при некотором натуральном l выполняется неравенство $k\cdot 10^l \leqslant M < (k+1)\cdot 10^l$, т. е. $\{10^6 \lg 2\} \in (\lg k, \lg(k+1))$. Остаётся вычислить $\lg 2$ на калькуляторе, сдвинуть десятичную точку на 6 позиций вправо, отбросить целую часть и сравнить результат с числами $0, \lg 2, \ldots, \lg 9, 1$.

В итоге имеем:

$$\lg 2 = 0{,}30102999566, \quad \{10^6\ \lg 2\} = 0{,}99566, \quad \lg 9 = 0{,}95424250943,$$

так что первая цифра числа 2^{10^6} равна 9. (*А. Я. Белов*)

Упражнение. Что больше: $A = \sqrt{2+\sqrt[3]{3}} + \sqrt{2-\sqrt[3]{3}}$ или а) $2\sqrt{2}$; б) $2\sqrt[3]{2}$?

а) Ответ. $\sqrt{2+\sqrt[3]{3}} + \sqrt{2-\sqrt[3]{3}} < 2\sqrt{2}$.

Решение. Надо воспользоваться выпуклостью подграфика функции $\sqrt{x}$.

б) Ответ. $\sqrt{2+\sqrt[3]{3}} + \sqrt{2-\sqrt[3]{3}} > 2\sqrt[3]{2}$.

Решение. $A^2 = \left(\sqrt{2+\sqrt[3]{3}} + \sqrt{2-\sqrt[3]{3}}\right)^2 = 4 + 2\sqrt{4-\sqrt[3]{3^2}}$. Результат надо сравнить с $\sqrt[3]{2^8}$, получаем сравнение $2 + \sqrt{4-\sqrt[3]{3^2}} \vee \sqrt[3]{2^5}$.

Нужно показать, что $4-\sqrt[3]{3^2} > 4(\sqrt[3]{2^2}-1)^2$, т. е. $-\sqrt[3]{3^2} > 4(\sqrt[3]{2^4}-\sqrt[3]{2^5})$, или

$$\frac{\sqrt[3]{3^2}}{8} < \sqrt[3]{2}\,(\sqrt[3]{2}-1).$$

[1)] Эта формула для натуральных показателей была известна и до Ньютона. Его заслуга — в обобщении формулы для любого вещественного и комплексного показателя.

Таким образом, надо проверить неравенство

$$\frac{9}{1024} < (\sqrt[3]{2}-1)^3.$$

Заметим, что

$$\frac{9}{1024} < \frac{1}{100} < \frac{1}{729/8} = \frac{1}{(4{,}5)^3}.$$

Поэтому достаточно установить, что $\sqrt[3]{2}-1 > 1/4{,}5 = 2/9$ или $\sqrt[3]{2} > 11/9$. Возводя обе части неравенства в куб, получаем очевидное неравенство $2 > 1331/729$.

Комментарий. Для успешного решения задач необходима культура выкладок. Взяв старые учебники и тренируясь на наиболее громоздких выражениях, можно узнать много мелких идей и простых приёмов, что приводит к заметному улучшению результатов и экономии времени.

(*А. Я. Белов*)

20.5. Условие. Имеется девять точек на плоскости: три красные, три синие, три зелёные. Никакие четыре не лежат на одном цикле (окружности или прямой). Дугами циклов будем называть дуги окружностей и отрезки прямых. Докажите, что существуют три непересекающиеся дуги цикла, на каждой из которых лежит по точке каждого цвета. (*A. Oppenheim*)

Решение. Положим на плоскость сферу достаточно большого радиуса, касающуюся плоскости по южному полюсу в одной из отмеченных точек. Спроектируем точки на эту сферу стереографически, центр проекции — в северном полюсе. Циклы перейдут в циклы: пересекающиеся — в пересекающиеся, непересекающиеся — в непересекающиеся. Задача перенесена на сферу.

Окружим точки шарами малого радиуса так, чтобы никакие четыре точки разных шаров не попали на одну плоскость. Каждый шар равномерно заполним веществом, отвечающим цвету его центра.

Имеет место знаменитая

Теорема о бутерброде с ветчиной. *В n-мерном пространстве даны n ограниченных множеств. Тогда существует плоскость, разделяющая каждое из них на два подмножества равной меры (в трёхмерном пространстве это означает: «поровну хлеба, поровну масла, поровну ветчины»).*

Применим эту теорему. Получим плоскость P, проходящую через три разноцветные точки. Она пересекается со сферой по циклу и делит сферу на две части — «верхнюю» и «нижнюю», в каждой из которых находится по точке каждого цвета. Итак, имеется цикл на P, «верхняя» тройка точек и «нижняя» тройка точек. Дуги соответствующих циклов не пересекаются, что и требовалось.

Упражнение. Дан выпуклый многоугольник. Среди окружностей, проходящих через три его последовательные вершины, выбирается окружность наибольшего радиуса. Докажите, что она содержит его целиком.

Комментарий к решению задачи 20.5. Идея проекции на сферу была предложена М. Л. Концевичем, когда он был школьником. Так решается и следующая задача.

Дано n точек общего положения (никакие три не лежат на одной прямой и никакие четыре не лежат на одной окружности). Тогда имеется окружность, проходящая через три из них и не содержащая ни одной отмеченной точки внутри, а также окружность, проходящая через три и содержащая внутри все остальные. Имеют место и все промежуточные случаи. (*А. Я. Канель-Белов*)

21.4. Условие. На плоскости отмечено n прямых общего положения, т. е. никакие три не пересекаются в одной точке, никакие две не параллельны и любые три точки пересечения либо не лежат на одной прямой, либо лежат на одной из отмеченных. Сколькими различными способами можно добавить *хорошую* прямую (т. е. не проходящую через точку пересечения двух отмеченных)? Два способа считаются *одинаковыми*, если один можно получить из другого, непрерывно двигая новую прямую так, чтобы она всё время оставалась *хорошей*. (*Фольклор*)

Ответ. $\dfrac{n-1}{8}(n^3-5n^2+22n-8)$.

Решение. Перейдём к проективной плоскости. Пространство прямых на проективной плоскости отличается от пространства прямых на обычной (аффинной) плоскости добавлением одного элемента (бесконечно удалённой прямой). В силу проективной двойственности прямой-«экватору» отвечает «полюс»-точка, соответствующая паре векторов, ортогональных к плоскости экватора. При этом трём прямым, проходящим через точку, отвечают три точки, лежащие на одной прямой, и наоборот. Пространство прямых на проективной плоскости изоморфно ей самой. По двойственности задача для проективной плоскости сводится к следующей.

На проективной плоскости отмечено n точек общего положения (никакие три не лежат на одной прямой и любые три прямые, проходящие через пары точек, либо не имеют общей точки, либо проходят через одну из отмеченных). Сколькими различными способами можно добавить *хорошую* точку (т. е. не лежащую на прямой, соединяющей две отмеченных точки)? Два способа одинаковы, когда один можно получить из другого, непрерывно двигая новую точку так, чтобы она всё время оставалась *хорошей*.

Для решения этой задачи достаточно попарно соединить все отмеченные точки и подсчитать число областей, на которые делят проективную плоскость эти прямые.

У нас $L = \binom{n}{2} = \frac{n(n-1)}{2}$ прямых, попарно соединяющих наши точки. При этом L прямых общего положения делят проективную плоскость на столько же частей, на сколько $L-1$ прямых общего положения делят обычную плоскость (одну из прямых можно объявить «бесконечно удалённой»). Получается $L(L-1)/2+1$ частей, из которых в случае обычной плоскости $2L$ частей неограниченных и, соответственно, $L(L-5)/2+1$ ограниченных.

Наши прямые, однако, не являются прямыми общего положения — имеется n «фокусов» (т. е. отмеченных точек), в которых сходятся по $n-1$ прямых. В каждом из них «схлопнулось» по $(n-1)(n-6)/2+1$ частей, общее количество схлопнувшихся частей равно $n((n-1)(n-6)/2+1)$. В итоге имеем, что искомое число частей составляет

$$\frac{\binom{n}{2}\cdot\left(\binom{n}{2}-1\right)}{2} + 1 - n\cdot\left(\frac{(n-1)(n-6)}{2}+1\right) = \frac{n-1}{8}(n^3-5n^2+22n-8).$$

Упражнения.

1. В Москве 9 высотных зданий. Приезжий математик хочет найти точку, из которой все здания видны в заданном порядке (считая от МГУ по часовой стрелке). Для любого ли заданного порядка это возможно? А если зданий 5, 6, 7 или 8?

   Указание. Рассмотрите *вырожденные положения* наблюдателя — когда он попадает на прямую, соединяющую два здания. Рассмотрите разбиение плоскости такими прямыми.

2. На сфере отмечено n точек *общего положения*, т. е. никакие три не лежат на одной прямой (дуге большого круга) и, как следствие, никакие две не противоположны. Сколькими способами можно добавить *хорошую* (т. е. не проходящую через отмеченные точки) прямую, если два способа одинаковы, когда один можно получить из другого, непрерывно двигая прямую так, чтобы она всё время оставалась *хорошей*.

3. На единичной сфере отмечено несколько кривых, суммарной длиной меньше π. Докажите, что существует прямая, ни одну из них не пересекающая.

Комментарий. См. статью: *Канель А., Ковальджи А.* Треугольники и катастрофы // Квант. 1992. № 11. С. 42–50, а также `https://www.google.com/search?q="Теорема+Робертса+о+треугольниках"` (*А. Я. Канель-Белов*)

22.1. Условие. Пусть для функции $g(x)$ при всех $x \geqslant 1$ выполняется равенство $g(x)^{g(x)} = x$. Найдите такую элементарную функцию $h(x)$, что $\lim_{x\to\infty}(g(x) - h(x)) = 0$. (*А. Я. Канель-Белов*)

Ответ. $h(x) = \dfrac{\ln x}{\ln\ln x}$.

Решение. Нам потребуется

Лемма. *Пусть $f'(x) \geqslant 1$ в некоторой области значений x. Тогда $|f(A) - f(B)| \geqslant |A - B|$ для любых A, B из этой области.*

Для доказательства леммы достаточно представить ось абсцисс как проволочку, которую функция f растягивает с коэффициентом, равным производной в данной точке. Если он везде больше или равен единице, то длина образа $|f(A) - f(B)|$ не меньше длины исходного отрезка $|A - B|$.

Поскольку при $x>1$ производная функции $x\ln x$ больше единицы, имеем

Следствие. *При $x, y > 1$ выполняется неравенство*

$$|x\ln x - y\ln y| \geqslant |x - y|.$$

Логарифмируя равенство из условия задачи, получаем $g(x)\ln g(x) = \ln x$. Для $h(x) = \dfrac{\ln x}{\ln\ln x}$ в силу следствия имеем

$$|g(x) - h(x)| < |g(x)\ln g(x) - h(x)\ln h(x)| = |\ln x - h(x)\ln h(x)| = \\ = \left|\ln x - \left(\ln x - \frac{\ln x \ln\ln\ln x}{\ln\ln x}\right)\right| \to 0 \quad \text{при } x \to \infty.$$

Комментарий. В ряде случаев, хотя функция не элементарна, можно найти её асимптотику в виде элементарной функции.

(*А. Я. Канель-Белов*)

22.2. Условие. а) Людям житья не стало от вампиров. Чтобы спасти поголовье людей, добрый волшебник наложил заклятие: 1) каждую ночь может выходить на охоту только один вампир; 2) если вампир ест человека, то следующим днём он превращается в человека и следующей ночью может быть съеден другим вампиром.

Вампир готов стать человеком, но риск быть съеденным другим вампиром ему неприемлем, при этом всем известно, что вампиры бесконечно умные и доверяют интеллектуальным качествам товарищей. Изначально имеется n вампиров. При каких n вампир выйдет на охоту?

б) Пусть вампиры умеют организовываться в банды по $n_1, \ldots, n_k$ вампиров. Если банда решает ночью выйти на охоту, то каждый ест по своему человечку и утром они становятся людьми. Число n называется *экологически равновесным* или просто *экологичным*, если никто из вампиров не собирается выйти на охоту.

Рассмотрим игру: имеется n фантиков, разрешается брать по $n_1, \ldots, n_k$ фантиков. Тот, кто не может сделать ход, — проигрывает. Докажите, что множество экологичных чисел в задаче про вампиров совпадает с множеством чисел в задаче о фантиках, проигрышных для начинающих.

(*А. Я. Канель-Белов*)

Решение. а) Ответ. При нечётных n. Один вампир гарантированно выходит на охоту. Если их два, то каждый опасается выйти: останется его товарищ, который гарантированно выйдет на охоту и может его съесть. Если вампиров три, то каждый понимает, что ему можно охотиться — оставшиеся два товарища будут бояться друг друга. Поэтому кто-нибудь из вампиров выйдет на охоту. Если их четыре, то первый боится охотиться — ибо понимает, что остаются три и в силу предыдущих рассуждений следующей ночью случится охота и т. д. Легко убедиться по индукции, что при нечётном n какой-нибудь вампир выходит на охоту, а при чётном они боятся друг друга.

б) Пусть множества *экологичных чисел* и *проигрышных для начинающего* определены для всех n, меньших некоторого M, и для таких чисел совпадают. Нам надо проверить, что число M либо принадлежит обоим множествам, либо ни одному не принадлежит.

Если $M - n_i$ при некотором i является числом проигрышным для начинающего, то первый делает ход, забирает n_i фантиков и выигрывает. Если же $M - n_i$ ни для какого i не является числом проигрышным для начинающего, то число M проигрышно для начинающего, ибо любой ход приводит к проигрышу.

Аналогичным образом, если $M - n_i$ при некотором i является экологичным числом, то банда из n_i вампиров сможет смело выйти на охоту. Если же $M - n_i$ ни для какого i не является экологичным числом, то, если банда из n_i вампиров выйдет на охоту, после этого останется неэкологичное число $M - n_i$ вампиров, и какая-то банда вампиров выйдет на охоту следом, что первой банде неприемлемо.

Мы видим, что принадлежность числа M множеству экологичных чисел равносильна принадлежности множеству чисел, проигрышных для начинающего, так что задачи про вампиров и фантики изоморфны.

22.4. Условие. Дана кососимметрическая матрица A. Докажите, что некоторая неотрицательная нетривиальная линейная комбинация её столбцов образует вектор, все координаты которого неотрицательны.

(*И. В. Митрофанов*)

Решение 1. Предположим противное. Пусть $v_1, v_2, \ldots, v_n$ — векторы столбцов матрицы.

Определим подмножество $S \subset \mathbb{R}^n$ как множество всевозможных сумм вида $a_1v_1 + a_2v_2 + \ldots + a_nv_n$ при неотрицательных вещественных $a_1, \ldots, a_n$, не все из которых нули. Множество точек в $\mathbb{R}^n$ со всеми положительными координатами обозначим $\mathbb{R}^n_+$. Если S и $\mathbb{R}^n_+$ не пересекаются, то нам потребуется

Теорема о разделяющей гиперплоскости. *Пусть X и Y — два непересекающихся выпуклых непустых подмножества $\mathbb{R}^n$. Тогда существуют такой ненулевой вектор v и такое число c, что $\langle x, v\rangle \leqslant c$ и $\langle y, v\rangle \geqslant c$ для любых $x \in X$ и $y \in Y$.*

Так как множества S и $\mathbb{R}^n_+$ содержат вместе с любым вектором все сонаправленные ему, то вектор v, полученный из теоремы, будет давать неположительные скалярные произведения с векторами из S и неотрицательные с векторами из $\mathbb{R}^n_+$.

Пусть v имеет координаты $\lambda_1, \ldots, \lambda_n$. Понятно, что $\lambda_i \geqslant 0$ при всех i. Координаты вектора $(\lambda_1, \ldots, \lambda_n)A$ — неположительные числа, так как это координаты скалярных произведений v и v_i. Так как A кососимметрична, то

$$A(\lambda_1, \ldots, \lambda_n)^T = -((\lambda_1, \ldots, \lambda_n)A)^T$$

— нетривиальная комбинация столбцов матрицы без отрицательных координат.

Решение 2. Пусть K_+ — множество ненулевых векторов с неотрицательными координатами, K_- — множество ненулевых векторов с неположительными координатами, $e_1, \ldots, e_n$ — базисные векторы.

Предположим, что для любого $v \in K_+$ найдётся e_i такой, что $e_iAv < 0$. Рассмотрим отображение $f\colon \mathbb{R}^n \to \mathbb{R}^n$, определённое как

$$f(v) = \min(e_1Av, 0)e_1 + \ldots + \min(e_nAv, 0)e_n$$

(иными словами, вектор v переходит в Av, а затем все положительные координаты заменяются нулями).

Отображение f непрерывно и переводит K_+ в K_-. Обозначив $s(v)$ сумму координат вектора v, рассмотрим отображение

$$g\colon v \to \frac{f(v)}{s(f(v))}.$$

Оно переводит в себя множество неотрицательных векторов с суммой координат 1. Но такое множество векторов гомеоморфно $(n-1)$-мерному шару, поэтому здесь применима

Теорема Брауэра о неподвижной точке. *Любое непрерывное отображение замкнутого шара в себя в конечномерном евклидовом пространстве имеет неподвижную точку.*

Таким образом, существует вектор $v \in K_+$, для которого $g(v) = v$. Это значит, что $f(v) = \lambda v$ для некоторого $\lambda > 0$. Так как A кососимметрична, то $\langle v, Av\rangle = 0$. Следовательно,

$$\langle f(v), f(v) - Av\rangle = \langle \lambda v, \lambda v\rangle - \lambda\langle v, Av\rangle \neq 0.$$

Но из формулы для $f(v)$ следует, что каждая координата равна нулю хотя бы у одного из векторов $f(v)$, $f(v) - Av = \Delta$. Противоречие.

Комментарий. Предположим, что два игрока играют в следующую игру: первый загадывает номер столбца матрицы A, второй загадывает номер её строки, после чего они открываются, и первый получает от второго столько рублей, какое число написано на пересечении загаданных строки и столбца (или отдаёт, если число меньше нуля). Тогда утверждение задачи можно понимать следующим образом: у первого игрока есть такая *вероятностная стратегия* (т. е. распределение вероятностей по столбцам), что при любой вероятностной стратегии второго игрока матожидание выигрыша первого неотрицательно. Если мы возьмём стратегию для первого и аналогично построенную стратегию для второго, то получим *равновесие Нэша* этой игры, матожидание выигрышей в нём будет нулевым.

(*И. Митрофанов*)

Опечатки, замеченные в выпуске 21

Страница,	строка	Напечатано	Следует читать
24	15 сверху	1937	1936
25	4 сверху	1968–69	1958–59
33	12 снизу	Зимина	Зимана
98	2 снизу (не считая сноски)	ций	ций. На самом деле, каждой точке этой области соответствует некоторый
166	4 сверху	независимо с $A_4 \cap A_5$	независимо с $A_3 \cap A_4 \cap A_5$
169	2 сверху	$(\overline{A}_2 \cap \ldots \cap \overline{A}_d)$	$(\overline{A}_2 \cup \ldots \cup \overline{A}_d)$
181	15 снизу	10	5

Информация для авторов

1. Сборник «Математическое просвещение» предназначен для широкого круга научных работников, преподавателей, учащихся и всех, кто интересуется математикой. Издание публикует материалы по различным областям математики, а также по проблемам её истории и преподавания, интересные и доступные указанной аудитории.

2. Сборник «Математическое просвещение» не публикует существенно новые научные результаты, оценка которых доступна лишь специалистам в соответствующей области. Не публикуются также материалы по текущим вопросам преподавания математики в учебных заведениях.

3. Материалы принимаются по электронной почте на адрес `matpros@yandex.ru` в виде двух файлов (`pdf` и `tex`) с дополнительными файлами рисунков и т. п., если требуется. Допускается присылка статей, набранных в `Word`.

4. Просим обратить внимание, что материалы принимаются в чёрно-белом исполнении.

5. Просим авторов кратко пояснять в начале статьи, в чём её цель и почему тема статьи представляет интерес.

6. Редакция благодарна авторам за оформление ссылок на литературу как в предыдущих выпусках, см. `http://www.mccme.ru/free-books/matpros.html`

7. В конце статьи необходимо указать для каждого из авторов:

— фамилию, имя, а также отчество (если есть) полностью,

— место работы/обучения,

— электронный адрес для дальнейшей переписки.

8. Авторы задач вместе с условием представляют письменное решение (хотя бы набросок).

9. Авторы опубликованных статей имеют право на 2 экземпляра сборника каждый, просим обращаться по адресу `matpros@yandex.ru`

Научно-популярное издание

Издательство Московского центра
непрерывного математического образования

119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241–08–04.

Отпечатано в ООО «Типография „Миттель-Пресс“».
г. Москва, ул. Руставели, д. 14, стр. 6.
Тел./факс +7 (495) 619–08–30, 647–01–89.
E-mail: mittelpress@mail.ru

Подписано в печать 22.11.2018 г. Формат 70×100 1/16. Бумага офсетная.
Печать офсетная. Печ. л. 15. Тираж 800 экз. Заказ №