

# МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ

Третья серия

ВЫПУСК 13

Москва  
Издательство МЦНМО  
2009

УДК 51.009  
ББК 22.1  
М34

*Первый тираж издания осуществлен  
при финансовой поддержке  
Российского фонда фундаментальных исследований  
по проекту № 09-01-07050*

## Редакционная коллегия

Бугаенко В. О.	Винберг Э. Б.	Вялый М. Н.
Гальперин Г. А.	Глейзер Г. Д.	Гусейн-Заде С. М.
Дориченко С. А.	Егоров А. А.	Ильяшенко Ю. С.
Канель-Белов А. Я.	Константинов Н. Н.	Прасолов В. В.
Розов Н. Х.	Сосинский А. Б.	Тихомиров В. М.
Френкин Б. Р.	Яценко И. В.	

ГЛАВНЫЙ РЕДАКТОР: Э. Б. Винберг      ОТВ. СЕКРЕТАРЬ: М. Н. Вялый

АДРЕС РЕДАКЦИИ:

119002, Москва, Б. Власьевский пер., д. 11, к. 301

(с пометкой «Математическое просвещение»)

EMAIL: [matpros@mccme.ru](mailto:matpros@mccme.ru)      WEB-PAGE: [www.mccme.ru/free-books](http://www.mccme.ru/free-books)

М34 **Математическое просвещение.** Третья серия, вып. 13. —

М.: МЦНМО, 2009. — 192 с.

ISBN 978-5-94057-486-6

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

УДК 51.009

ББК 22.1

ISBN 978-5-94057-486-6

© МЦНМО, 2009.

## СОДЕРЖАНИЕ

### Математический мир

<i>Я. С. Дубнов (1887–1957)</i> . . . . .	5
Н. Р. Брумберг <i>Я. С. Дубнов — страницы биографии</i> . . . . .	10
А. Г. Порошкин <i>О работе профессора Я. С. Дубнова в Коми государственном педагогическом институте (1952–1954 уч. годы)</i> . . . . .	13

### Группы и топология

Ант. А. Клячко <i>Комбинаторная теория групп и геометрия</i> . . . . .	18
М. Н. Вялый, О. В. Шварцман <i>Фуксовы группы: от топологии к геометрии</i> . . . . .	33

### Наш семинар: математические сюжеты

Ю. И. Ионин <i>Конечные проективные плоскости</i> . . . . .	50
Г. Г. Магарил-Ильяев, В. М. Тихомиров <i>Метод Ньютона и его приложения к решению уравнений и теории экстремума</i> . . . . .	80
С. Б. Гашков <i>Графы-расширители и их применения в теории кодирования</i> . . . . .	104
Ю. Л. Притыкин <i>Колмогоровская сложность</i> . . . . .	127
В. А. Гурвич <i>Метрические и ультраметрические пространства сопротивлений</i> . . . . .	134
П. В. Бибигов <i>Описанные циклические линии треугольника в геометрии Лобачевского</i> . . . . .	142
Ф. В. Петров <i>Вписанные четырёхугольники и трапеции в абсолютной геометрии</i> . . . . .	149
А. В. Акопян <i>О некоторых классических конструкциях в геометрии Лобачевского</i> . . . . .	155

### Нам пишут

А. В. Стояновский <i>Некоторые аддитивные соотношения в треугольнике Паскаля</i> . . . . .	171
А. А. Азамов <i>О жордановых многоугольниках</i> . . . . .	174

### Задачный раздел

<i>Условия задач</i> . . . . .	179
<i>Решения задач из предыдущих выпусков</i> . . . . .	182



---

---

## Математический мир

---

---

Я. С. ДУБНОВ (1887 – 1957)



Пятьдесят лет тому назад, в конце 1957 года начали выходить выпуски второй серии «Математического просвещения»<sup>1)</sup>. Всего тогда вышло шесть выпусков (два в 1957 и по одному в последующие четыре года; потом издание прекратило свое существование). Одним из энтузиастов, которому издание «Математического просвещения» было обязано своим возрождением, был Яков Семенович Дубнов, светлая личность и активный деятель математического просвещения. При его непосредственном участии были

---

<sup>1)</sup>Первая серия была начата в 1934 году и закончилась с началом войны.

подготовлены три выпуска «Математического просвещения». В декабре 1957 года Якову Семеновичу исполнялось 70 лет. Ему захотелось провести свои постюбилейные дни подальше от Москвы, и он поехал в Саратов для чтения лекций. После одной из лекций Яков Семенович внезапно скончался. Жизни и творчеству Я. С. Дубнова посвящен очерк его коллеги и друга А. М. Лопшица: «Яков Семенович Дубнов — ученый, педагог, человек», опубликованный в пятом выпуске второй серии «Математического просвещения».

Наше издание «Математического просвещения» имеет подзаголовок «Третья серия». При его рождении в редакционной статье, открывающей первый том нового издания, было сказано, что редакционная коллегия надеется продолжить славные традиции первых двух серий. Представляется уместным рассказать здесь об одном из организаторов предыдущей серии — Якове Семеновиче Дубнове.<sup>2)</sup>

Ниже печатаются некоторые материалы о его жизни и творчестве: «Жизнеописание Дубнова Якова Семеновича», написанное им самим при поступлении в Коми государственный педагогический институт (мне представляется, что читателям нашего издания может быть интересно, как иногда складывалась жизнь представителей поколения наших дедов и прадедов), диплом об окончании Я. С. Дубновым Новороссийского университета со сведениями о том, чему учили в университетах в предреволюционные годы (эти материалы были любезно предоставлены старшим преподавателем кафедры матанализа, зам. декана Коми гос. пединститута Б. Д. Лившицем при посредничестве Я. Э. Юдовича, за что мы выражаем им искреннюю благодарность), статья Н. Р. Брумберг, дополняющая статью Лопшица некоторыми материалами, которые в те годы не могли быть напечатаны, и воспоминания о Я. С. Дубнове профессора Сыктывкарского университета А. Г. Порошкина.

*В. М. Тихомиров*

---

<sup>2)</sup> Я. С. Дубнов был автором следующих статей во второй серии «Математического просвещения»: «Тригонометрия в школьном курсе геометрии» (вып. 1), «К проблеме создания новых учебников по математике для средней школы» (вып. 3), «Два новых учебника по алгебре» и «Содержание и методы преподавания математического анализа и аналитической геометрии в средней школе» (вып. 5), «Модель евклидовой геометрии на гиперболической плоскости» (вып. 6).

## ЖИЗНЕОПИСАНИЕ ДУБНОВА ЯКОВА СЕМЕНОВИЧА

Родился 30. XI. 1887 в г. Мстиславле Могилевской губернии в семье литератора, впоследствии историка, С. М. Дубнова. В 1891 г. семья переехала в Одессу, где я в 1902 г. поступил в классическую гимназию. В 1906 году окончил ее и поступил на физ.-мат. факультет Новороссийского (Одесского) университета, где специализировался по математике, главным образом под руководством профессоров В. Ф. Кагана и С. О. Шатуновского. В 1910 г. получил серебряную медаль за сочинение на тему, предложенную факультетом. В том же году, за участие в студенческих антиправительственных выступлениях был исключен из университета, отбыл  $1\frac{1}{2}$  месяца тюремного заключения и был выслан в Вильно под надзор полиции. В 1913 г. получил разрешение временно вернуться в Одессу для сдачи экстерном гос. экзаменов, получил диплом первой степени и снова был выслан в административном порядке из Одессы. В результате недосмотра полиции, мне удалось в 1914 г. поселиться в Москве, где я с небольшими перерывами живу поныне. В июне 1914 г. начал преподавание математики на частных общеобразовательных курсах. С начала 1918 г., продолжая преподавание, был консультантом Отдела реформы школы Наркомпроса и участвовал в составлении первых вариантов программ по математике для единой трудовой школы. В 1919–20 гг. провел год в Спас-Деменске Калужской губернии, где преподавал математику в средней школе. В начале 1920 г. вернулся в Москву и преподавал математику на рабфаке Московского государственного университета вплоть до 1923 г. С 1921 г. начал преподавание высшей математики в Московском электротехническом институте связи, позже преобразованном в энергетический факультет Московского высшего технического училища. В этом вузе я преподавал до 1929 г., последние два года в должности доцента. С 1923 по 1928 г. проходил аспирантуру в МГУ, завершившуюся диссертацией «Теория прямолинейных конгруенций»; после установления ученых степеней получил за эту диссертацию степень кандидата физ.-мат. наук. С 1923 по 1937 г. преподавал математику в педагогическом факультете 2-го МГУ, после преобразованном в Московский государственный педагогический институт им. В. И. Ленина. В этом вузе в 1930 г. я был утвержден постановлением ГУС'а (Государственного Ученого Совета) и. о. профессора. С 1928 по 1952 г. работал в Московском государственном университете, где в 1931 г. постановлением ГУС'а был назначен действительным членом научно-исследовательского института математики и механики (НИИММ МГУ) с окладом профессора. По представлению этого института, ВАК Наркомпроса присвоил мне 1. II. 1936 г. ученую степень доктора физ.-мат. наук без защиты диссертации и ученое звание действительного члена НИИММ МГУ.

С утверждением кафедр, состоял профессором кафедры дифференциальной геометрии по Х.1948 г., когда по сокращению штата кафедры был переведен на должность старшего научного сотрудника НИИММ МГУ, однако, оставаясь в кафедре дифференциальной геометрии, продолжая преподавание на почасовой оплате. Ввиду упразднения звания действительного члена НИИ, ученый совет механико-математического факультета МГУ в закрытом голосовании представил меня к званию профессора, а Ученый Совет МГУ утвердил это избрание и передал в ВАК МВО. Однако, эта комиссия не приняла дела к рассмотрению на том формальном основании, что в кафедре дифференциальной геометрии отсутствует профессорская вакансия. В начале 1952 г. в НИИММ МГУ была упразднена должность старшего научного сотрудника и на этом прервалась моя работа в МГУ.

По совместительству с этой работой, после прекращения преподавания в Московском государственном пединституте, я работал в ряде учреждений педагогического уклона: 1) в Загорском учительском институте (зав. каф. математики в 1940–41 гг.), 2) в Московском городском институте усовершенствования учителей (лектор в 1938–39 гг. и в 1947–49 гг.), 3) в научно-исследовательском институте методики преподавания Академии педагогических наук (старший научный сотрудник в 1946–48 гг.) и др.

Семейные обстоятельства: жена, врач Б. А. Дубнова, кандидат наук, была репрессирована в 1936 г., ныне находится на поселении в Коми АССР, проживает в Княжпогосте. Две замужние дочери, преподавательницы математики и физики, проживают в Москве, из них младшая со своей семьей — при мне. В ноябре 1941 г. я был вместе с этой дочерью эвакуирован Московским государственным университетом (где во время войны работал без перерывов) сначала в Ашхабад, затем в Свердловск, в июне 1943 г. рэвакуирован.

Общественная работа: работал в профсоюзных органах. В частности, в 1919–20 гг. был председателем Районного бюро союза работников просвещения в г. Спас-Деменске; в 1934 г. — председателем бюро секции научных работников в НИИММ МГУ. Ныне состою выборным членом бюро школьной секции Московского математического общества.

Под моим руководством проходили аспирантуру и защитили кандидатские диссертации 7 человек (один из них ныне — доктор). Кроме того, ко мне в разное время были прикомандированы 5 человек из провинции; из них 4 человека защитили кандидатские диссертации, один подготовил ее. Мною опубликовано 27 научных работ [...], 3 учебника для высшей школы (из них один в 10 изданиях, второй в 4-х), 4 статьи по вопросам методики преподавания, несколько десятков статей в старом и новом издании БСЭ.



## ДИПЛОМ

Предъявитель сего, бывший студент физико-математического факультета ИМПЕРАТОРСКОГО Новороссийского Университета, Дубнов Хаим-Яков Симонович, сын мещанина, вероисповедания иудейского, в бытность студентом этого факультета подвергался полукурсовому испытанию, на котором оказал следующие успехи: по аналитической геометрии (I и II ч.ч.) — *весьма удовлетворительные*; по высшей алгебре — *весьма удовлетворительные*; по введению в анализ — *весьма удовлетворительные*; по началам теории определителей — *весьма удовлетворительные*; по началам теории чисел — *весьма удовлетворительные*. С разрешения Министерства Народного Просвещения от 15 сентября 1912 года, за №36156, подвергался в сентябре и октябре того же года испытанию в качестве экстерна в Физико-Математической испытательной комиссии при названном Университете, по отделению математических наук, при чем оказал следующие успехи: на письменном испытании: по математике (I и II ч.ч.) — *весьма удовлетворительные*; на устных испытаниях: по дифференциальному исчислению (I и II ч.ч.) — *весьма удовлетворительные*; по интегральному исчислению (I и II ч.ч.) — *весьма удовлетворительные*; по физике (I и II ч.ч.) — *весьма удовлетворительные*; по химии — *удовлетворительные*; по механике — *весьма удовлетворительные*; по математике (интегрирование уравнений, определенные интегралы, вариационное исчисление, исчисление конечных разностей) — *весьма удовлетворительные*, по астрономии — *удовлетворительные*, по физической географии с метеорологией — *весьма удовлетворительные*; по дифференциальной геометрии — *весьма удовлетворительные*; по теории чисел — *весьма удовлетворительные*.

## Я. С. Дубнов — страницы биографии

Н. Р. Брумберг

Профессор Дубнов скоропостижно скончался 13 декабря 1957 года. Он читал лекции по методике преподавания геометрии в Саратовском университете и умер в гостинице сразу после последней лекции.

Когда я ещё училась в школе, в мои руки попала его тоненькая брошюра «Ошибки в геометрических доказательствах». Она сразу подкупила меня ясностью изложения, четкостью мышления, обилием геометрических парадоксов, обрушивающихся на читателя. Между прочим, сначала она так и называлась: «Парадоксы в геометрических доказательствах». Но, поскольку первое издание появилось в 1953 году, во время борьбы с засильем иностранных слов и терминов в русском языке, пришлось заменить слово «парадоксы» на «ошибки». Среди школьников она разошлась моментально. Вслед за первым появилось второе издание. Затем — переводы на иностранные языки. Многие из моих однокурсников вспоминали, чем была для них эта тоненькая книжка. Впечатления самые яркие.

Мне не довелось встретиться с Я. С. Дубновым в мои университетские годы. Значительно позже, и притом достаточно случайно, жизнь столкнула меня с людьми, которые хорошо знали этого человека: его дочь Викторией Яковлевной, его бывшими аспирантами и коллегами. Меня поразила его судьба, захватило обаяние личности, захотелось больше узнать о нём и рассказать другим. Помимо устных источников, я черпала сведения из статьи В. В. Вагнера и А. М. Лопшица к тридцатилетию со дня смерти Якова Семеновича в «Трудах семинара по векторному и тензорному анализу» (вып. XI, 1961, изд. МГУ), из статьи И. М. Яглома в журнале «Математика в школе» (№6 за 1987 год), посвященной столетию со дня рождения Якова Семеновича, и наконец, из приложения «Одесса математическая» к израильской газете «Вести» от 13.01.2007.

Яков Семенович родился 1 декабря 1887 года в маленьком живописном городке Мстиславле в Белоруссии, в Могилёвской губернии. Его отец был известным историком, писателем, публицистом, деятелем русской культуры. Он погиб в рижском гетто в возрасте 80 с лишним лет, разделив судьбу своего народа. Яков Семенович горько сожалел, что не взял отца в Москву: разрешение было получено, но началась война, и он не успел.

Когда Якову исполнилось пять лет, Дубновы переехали в Одессу. Там прошли и его школьные годы. Он учился в частной гимназии, поскольку как раз в это время в казённых гимназиях была введена процентная норма для еврейских детей. Директором этой гимназии был известный геометр Вениамин Фёдорович Каган, который впоследствии возглавил кафедру дифференциальной геометрии в Московском университете. Там же преподавал математику замечательный педагог и блистательный лектор С. О. Шатуновский. Яша Дубнов окончил гимназию в 19 лет и сразу поступил в Новороссийский университет (там же, в Одессе) к своему учителю Кагану. Как и многие студенты, он участвовал в студенческих волнениях и был столь активен, что его арестовали и полтора месяца он провёл в царской тюрьме. Свойственная ему необычайная скромность не позволила Якову Семеновичу позже назвать себя политкаторжанином и приобрести определённый статус при советской власти. Когда Дубнов вышел из тюрьмы, его сразу исключили из университета с волчьим билетом, а ведь у него уже была готова научная работа для подачи на конкурс!

Только в 1914 году он решился уехать в Москву по подложному паспорту. Все это время он не переставал заниматься исследовательской работой. На жизнь зарабатывал частными уроками. Тогда многие, как и сейчас, выходили из положения подобным образом. Лишь в 1920 году, уже при советской власти, он начал преподавать в различных московских вузах. С 1928 года он — преподаватель на кафедре дифференциальной геометрии в первом МГУ (нынешнем Московском государственном университете). С 1931 года — профессор на этой кафедре. Руководил кафедрой его любимый учитель профессор В. Ф. Каган.

Многие годы непрерывного преподавания, научных занятий, создания курсов методики преподавания — и вот неожиданный поворот судьбы! В 1936 г. арестована и репрессирована его жена Бася Ароновна. И вот профессор Дубнов ездит на двухчасовые свидания в Красноярский край. Можно вообразить, сколько времени занимала такая поездка! Помимо этого, раз в году можно было подавать прошение о пересмотре дела. Иногда за ответом (по вызову из НКВД) отправлялась дочь Виктория. Однажды 18-летняя девочка, студентка физфака, пришла на Кузнецкий мост в приёмную НКВД за ответом. «К сожалению, всё остаётся по-прежнему», — сказал ей майор НКВД. Студентка потеряла самообладание, затопала ногами и закричала на майора: «Как?! Прошло 12 лет, и вы не смогли исправить человека! Куда же годится ваша организация?» Майор опешил. Сказал только: «Уходи, девчонка!» Девушка благополучно удалилась, а мать после этого отпустили на поселение в Княжпогост в Коми АССР, «списали по состоянию здоровья».

Пока Яков Семенович имел ту или иную возможность работать в Москве, он из Москвы не уезжал. Но в 1952 г. его ставку сократили. Он уехал

работать в Сыктывкар, жену перевели туда и дали им комнату в общежитии. Заметим, ректор этого института обладал известной смелостью — не боялся брать на работу ссыльных и их близких. А кого же ещё найдёшь в такой глуши? Благодаря этому обстоятельству, в частности, в Сыктывкаре сложилась неплохая математическая школа. Она существует и до сих пор.

После смерти Сталина Басе Ароновне разрешили вернуться в Москву. Дубновы возвращаются. Яков Семёнович снова работает в МГУ до самой своей внезапной кончины в Саратове.

Еще один эпизод из жизни Дубнова, рассказанный мне его дочерью. На общем собрании студентов и преподавателей второго МГУ он осмелился заступиться за Льва Седова, сына Троцкого — «изменника родины». Вышинский, ректор университета, предложил его исключить. Но Дубнов за него заступился, сказал, что он отличник и в прекрасных отношениях с остальными студентами. Это происходило при общем молчании всех присутствующих. При всех он обвинил ректора в не порядочности. Разумеется, Вышинскому удалось добиться своего (исключения студента), но не на этот раз. Дома Дубнов рассказал об этом жене, и она с грустью сказала: «Ну теперь тебя, Яша, посадят». Тем не менее, для него это осталось без последствий. А Л. Седов не избежал судьбы и погиб.

Все друзья и коллеги Дубнова — Н. В. Ефимов, М. А. Крейнес, Л. А. Люстерник, П. К. Рашевский и другие — отзывались о нем с большим пиететом. Остается лишь сожалеть, что мне пока слишком мало удалось узнать об этом человеке.

## О работе профессора Я. С. Дубнова в Коми государственном педагогическом институте (1952–1954 уч. годы)

А. Г. Порошкин

С профессором Я. С. Дубновым мне довелось поработать два учебных года на кафедре математики Коми государственного педагогического института. После окончания в 1951 г. этого института и года работы в сельской школе я был принят на кафедру математики в должности ассистента и приступил к работе в сентябре 1952 г.

Яков Семёнович по приказу директора института (ректоры в те времена были только в университетах) был принят на работу с 1 сентября 1952 г., но приступил к работе в конце сентября или в начале октября. Помню, что он читал курсы дифференциальной геометрии (оба учебных года), аналитической геометрии (в 1953–54 уч. году), проективной геометрии и оснований геометрии, вёл практические занятия по первым двум курсам. Курсы двух геометрий (дифференциальной и аналитической) читались физикам и математикам отдельно. В те годы на первый курс принималось по 50 студентов-математиков и столько же физиков, но ко второму курсу число студентов сильно сокращалось.

Яков Семёнович проводил кружковые занятия со студентами III–IV курсов по решению задач Московских математических олимпиад школьников (в стране тогда подобные олимпиады проводились только в Москве, Ленинграде и, возможно, в Киеве). На первом занятии кружка предупредил участников: «Пусть вас не обижает то, что будем решать школьные задачи. Имейте в виду, что в лице московских школьников, участвующих в олимпиадах, вы имеете серьёзных конкурентов». Эти занятия посещались также учащимися старших классов городских школ.

Яков Семёнович нередко читал студентам популярные лекции. Темы их были всегда интересными. Помню одну из них: «Для чего нужно знать высшую математику учителю математики средней школы?». Все подобные лекции студенты посещали охотно и слушали с большим интересом, хотя и проводились они вне расписания.

Семинара специально для преподавателей кафедры Яков Семёнович не организовывал, хотя по вопросам преподавания такой семинар мог бы

существовать. Однако заседания кафедры нередко практически переходили в семинары при самом активном участии Якова Семёновича. На них он в непродолжительных своих выступлениях высказывал интересные мысли, предлагал обсуждать их, активно обсуждал планы открытых лекций (или сами лекции) молодых преподавателей (их на кафедре было большинство), давал полезные советы, вносил предложения по изменениям учебных программ и перестройке некоторых курсов. Помню, на одном из заседаний он сделал интересный доклад о преподавании начал дифференциального исчисления. Яков Семёнович считал, что начинать этот раздел следует с понятия дифференциала, а уж затем знакомить с производной.<sup>1)</sup> Он предлагал ввести понятие «подъём прямой» (по-видимому, вместо углового коэффициента). Выказано было много интересного, но время и работа стёрли почти всё из памяти.

Кафедрой математики в Коми пединституте Я. С. Дубнов не заведовал. Конечно, по принятым правилам при наличии на кафедре доктора наук, профессора, он должен был бы руководить кафедрой. Но руководство института, видимо, предполагая, что Яков Семёнович долго не задержится в Сыктывкаре, да ещё и при отсутствии нормальных бытовых условий, сочло возможным оставить заведующим кафедрой своего выпускника, фронтовика Г. П. Лапина, успешно защитившего в 1950 г. кандидатскую диссертацию и имевшего уже ученое звание доцента. Это могло быть и ориентиром для будущих выпускников факультета, настроить их на продолжение учёбы в аспирантуре: Смотрите! И нашим выпускникам это по силам! Надо сказать, что до Г. П. Лапина на кафедре не было ни одного кандидата наук с конца 30-х годов, тем более из выпускников своего института.

В течение 1952 г. на кафедре математики работал еще один профессор из Ленинградского горного института, Андрей Митрофанович Журавский, осужденный в 1942 г. по «Делу №555» на 10 лет<sup>2)</sup> и в феврале 1952 г. высланный в Сыктывкар. В январе 1953 г. он возвратился в Ленинград, в 1954 г. был реабилитирован, а затем восстановлен в прежней должности профессора и заведующего кафедрой. Беседы Я. С. Дубнова и А. И. Журавского часто проходили на кафедре в основном во внеурочное время и были для нас всегда интересны; их помощь, советы преподавателям были неоценимы. Их занятия, посещаемые нами, были для нас прекрасными уроками, показывающими, как надо работать со вчерашними школьниками.

---

<sup>1)</sup>По этому поводу см. статью Я. С. Дубнова «Содержание и методы преподавания математического анализа и аналитической геометрии в средней школе» в сборнике «Математическое просвещение», сер. 2, вып. 5, 1960, с. 17–56, 72.

<sup>2)</sup>Подробно об этом можно прочитать в статье Я. Голованова «Палачи и жертвы дела №555», опубликованной в журнале «Огонёк» №5 за 1992 год.

Яков Семёнович был прекрасным лектором, какие нечасто встречаются даже среди профессоров. Работавший в Коми пединституте с 1937 г. заведующим кафедрой физики доцент Г. П. Балин с особой радостью принял сообщение о том, что к нам в институт едет работать московский профессор Я. С. Дубнов. В конце 20-х – начале 30-х годов Г. П. Балин обучался во втором МГУ (впоследствии Московском государственном пединституте), где слушал лекции Якова Семёновича. Он говорил, что Яков Семёнович был единственным преподавателем, на лекциях которого была 100-процентная посещаемость (в то время в вузах было свободное посещение занятий и лекции других преподавателей посещало не более трети студентов).

Мы, студенты послевоенных лет, пользовались прекрасными учебниками и задачками Якова Семеновича по курсам дифференциальной геометрии и математического анализа (эти книги были в нашей библиотеке, думаю, что они сохранились и на кафедрах), были как бы знакомы с ним заочно, и нам естественно было интересно встретиться с ним, поработать вместе с известным ученым и профессором.

Приезд Якова Семёновича в Сыктывкар объясняется следующим. Его жена, известный врач, кандидат медицинских наук, в 1936 г. была осуждена и отбыла срок в лагере. В конце 40-х годов она была переведена на поселение в посёлок Железнодорожный (Княжпогост), в 100 км от Сыктывкара, а в 1952 г. ей было разрешено проживание в г. Сыктывкаре, где с 1932 года существовал педагогический институт. Сюда в этот институт и приехал Яков Семенович. Проживали они в четырёхэтажном студенческом общежитии института в центре города. (В то время немало наших преподавателей, и не только молодых, но и семейных, проживали в общежитиях. Квартиры даже в деревянных домах ждать приходилось долго).

По-видимому, в 1954 г. жене Якова Семёновича разрешили выехать в Москву, и они уехали из Сыктывкара в августе 1954 г. Яков Семёнович, видя тяжелое положение с кадрами на кафедре математики (на кафедре оставалось 8 преподавателей, из них один кандидат физико-математических наук доцент, заведующий кафедрой), возможно, давал согласие продолжить работу в КГПИ по совместительству<sup>3)</sup>: приезжать на небольшой срок, читать лекции, принимать экзамены (подобное практиковалось в конце 40-х годов, когда на таких условиях приглашали из Горьковского пединститута доцентов Н. А. Фролова (читал курс математического анализа) и В. И. Костина (читал курс «Основания геометрии»). Согласись на это руководство института — была бы огромная помощь для кафедры. Однако оно отказало в этой просьбе, предложило остаться на работе на

<sup>3)</sup> В личном деле Якова Семеновича имеется заявление о переводе его в 1954–55 учебном году на 0,5 ставки.

постоянной основе. Разумеется, Яков Семёнович на это не согласился. Кто бы согласился променять Московский университет на пединститут в Сыктывкаре, куда в то время можно было долететь только самолётом! Вернувшись в Москву, Яков Семёнович организовал выпуск научно-методических сборников «Математическое просвещение», в основном предназначенных для активных учителей, любопытных студентов и даже школьников. Через несколько лет после его кончины выпуск сборников прекратился. Кажется, в 1962 г., я был свидетелем беседы профессора В. А. Залгаллера с участником семинара в ЛГУ. На вопрос: почему прекратился выпуск «Математического просвещения» В. А. Залгаллер ответил примерно следующее: «Потому что не стало инициатора создания этого сборника, душой болевшего за него, — Якова Семёновича Дубнова».

Никогда ни с кем у Якова Семёновича не было конфликтов, хотя свою точку зрения он всегда твёрдо отстаивал.

Студенты вспоминали: на одной из первых лекций в аудитории начался шумок, он прервал рассказ и заявил: «Впрочем, если кому-то не интересно слушать меня, я не возражаю, если они покинут аудиторию». Голос его был довольно громкий, язык — дай Бог так говорить многим нашим дикторам радио и телевидения. Оформление доски идеальное, как говорят, снимай с доски все формулы и чертежи и переноси на страницу учебника — ничего печатать не нужно. Лекции читал замечательно и все студенты от них были в восторге. До сих пор оставшиеся еще в живых бывшие ученики вспоминают его добрым словом. Был всегда доброжелателен к студентам, но требований не снижал: то, что ты сдаёшь на экзамене, ты должен понимать и знать. Помню, был такой случай. После экзамена в одной группе он шутя и с сожалением говорил на кафедре: «Получили 20 двоек из 25 возможных». Но через несколько дней, после переэкзаменовки, та же группа оставила Якову Семеновичу хорошее впечатление; он весьма довольный рассказывал на кафедре об этом: вот, дескать, подготовились как следует и получился другой результат. Между прочим, принимая экзамен, он весь был занят беседой с экзаменуемым. Создавалось впечатление, что аудитория оставлена им без внимания. Об этом как-то тактично намекнул Якову Семёновичу наш декан М. А. Выборов, зайдя к нему на экзамен: дескать, студенты могут воспользоваться источниками. Яков Семёнович спокойно ответил: «Ничего, это не так важно. Во время ответа сразу станет ясно, сознательно усвоен материал или нет». Он с большим удовлетворением ставил высокие оценки знающим студентам и комментировал на кафедре их ответы. Память его была изумительна.

Яков Семёнович нередко выступал и на заседаниях Совета института, и его выступления всегда вызывали большой интерес членов Совета и присутствующих.



На стендах Коми пединститута висят портреты профессоров Я. С. Дубнова и А. М. Журавского. Оба они оставили яркий след в истории факультета, института, несмотря на непродолжительность их работы в нашем коллективе.

---

---

# Группы и топология

---

---

## Комбинаторная теория групп и геометрия

Ант. А. Клячко

В этой статье, предназначенной для детей от пятнадцати до девяности девяти лет, в популярной форме рассказывается о некоторых связях между комбинаторной теорией групп и геометрией: о геометрической интерпретации вывода следствий из соотношений и о теории малых сокращений.

### 1. ЧТО ТАКОЕ СООТНОШЕНИЕ МЕЖДУ ЭЛЕМЕНТАМИ ГРУППЫ?

Между элементами группы могут существовать разные зависимости. Например, элементы  $a$  и  $b$  могут коммутировать ( $ab = ba$ ) или удовлетворять более сложному соотношению  $(a^{2009}b^{-2009})^{2009} = 1$ . Изучением всевозможных зависимостей, или соотношений, занимается *комбинаторная теория групп*. Эта наука тесно связана с геометрией. О некоторых таких связях рассказывается в этой статье.

Давайте попытаемся аккуратно определить понятие соотношения между элементами группы. Пусть  $X$  — некоторый алфавит (то есть просто некоторое множество) и  $X^{-1}$  — копия алфавита  $X$ . При этом мы считаем, что  $X \cap X^{-1} = \emptyset$  и для каждой буквы  $x \in X$  однозначно определена некоторая буква  $x^{-1} \in X^{-1}$ , причём отображение  $x \mapsto x^{-1}$  является взаимно однозначным соответствием между множествами  $X$  и  $X^{-1}$ . Мы будем предполагать, что  $X = \{x_1, x_2, \dots\}$  и  $X^{-1} = \{x_1^{-1}, x_2^{-1}, \dots\}$ . Рассмотрим объединённый алфавит  $X^{\pm 1} = X \cup X^{-1}$  и некоторое слово  $w(x_1, x_2, \dots)$  в этом алфавите (это значит, что  $w$  представляет собой конечную последовательность, каждый элемент которой есть либо одна из букв  $x_i$ , либо одна из букв  $x_i^{-1}$ ). Если теперь  $a_1, a_2, \dots$  — элементы некоторой группы  $G$ , то естественным образом определяется значение

$w(a_1, a_2, \dots)$  слова  $w(x_1, x_2, \dots)$  на этом наборе элементов  $a_1, a_2, \dots$  — надо вместо каждой буквы  $x_i$  подставить элемент  $a_i$ , вместо  $x_i^{-1}$  подставить  $a_i^{-1}$  и всё перемножить. Например, если в качестве  $G$  взять симметрическую группу  $S_3$ , а в качестве слова  $w$  взять коммутатор  $w(x_1, x_2) = [x_1, x_2] = x_1^{-1}x_2^{-1}x_1x_2$ , то значением этого слова на паре транспозиций будет тройной цикл:  $w((12), (23)) = (12)^{-1}(23)^{-1}(12)(23) = (321)$ .

Если значение слова  $w(x_1, x_2, \dots)$  на наборе элементов  $a_1, a_2, \dots$  равно единице в группе  $G$ , то говорят, что элементы  $a_1, a_2, \dots$  удовлетворяют соотношению  $w(x_1, x_2, \dots) = 1$ . Иногда соотношением мы будем для краткости называть само слово  $w(x_1, x_2, \dots)$ , написанное в левой части этого равенства.

При работе с соотношениями используют обычные сокращения. Например, соотношение  $x_1^{-1}x_2x_1^{-1}x_2x_2x_2 = 1$  можно записать в виде  $(x_1^{-1}x_2)^2x_2^2 = 1$  или  $(x_1^{-1}x_2)^2 = x_2^{-2}$ .

## 2. ЧТО ТАКОЕ СЛЕДСТВИЕ СООТНОШЕНИЙ?

Бывает так, что какое-то соотношение следует из других соотношений. Например, соотношение коммутативности  $xy = yx$  следует из соотношений  $x^2 = 1, y^2 = 1$  и  $(xy)^2 = 1$ . Действительно, эти равенства означают, что элементы  $x, y$  и  $xy$  совпадают со своими обратными. С другой стороны, в любой группе имеет место тождество  $(xy)^{-1} = y^{-1}x^{-1}$ . Таким образом,  $xy = yx$ , что и требовалось.

Следующее определение выглядит вполне естественным.

СЕМАНТИЧЕСКОЕ ОПРЕДЕЛЕНИЕ СЛЕДСТВИЯ. Соотношение

$$w(x_1, x_2, \dots) = 1$$

называют *следствием* семейства соотношений  $\mathcal{V}$ , если в любой группе  $G$  любой набор элементов, удовлетворяющий всем соотношениям семейства  $\mathcal{V}$ , удовлетворяет также соотношению  $w(x_1, x_2, \dots) = 1$ .

Мы назвали это определение семантическим, поскольку в нём речь идёт о смысле определяемого понятия. Определения, в которых не упоминается смысл, а говорится только о внешнем виде, о формальных признаках, называют синтаксическими. Например, если мы скажем, что *хороший студент* — это студент, который интересуется наукой и способен решать нетривиальные задачи, то это будет семантическое определение. Если же мы скажем, что *хороший студент* — это студент, у которого в зачётке нет троек, то это будет синтаксическое определение. Разумеется, эти два определения неэквивалентны. Вряд ли вам удастся придумать удовлетворительные синтаксические определения таких понятий, как *интересная книга, красивая женщина, трудная задача* и т. п. В математике

форма и содержание более тесно связаны, чем в реальной жизни. В частности, понятие следствия можно определить чисто синтаксически. Для этого нам понадобится кое-какая нехитрая техника работы со словами. Слово в алфавите  $X^{\pm 1}$  называется *несократимым*, если оно не содержит подслов вида  $xx^{-1}$  и  $x^{-1}x$ , где  $x \in X$ . Несократимое слово, полученное из слова  $w$  последовательным вычёркиванием двухбуквенных подслов указанного вида, мы будем называть *несократимой формой* слова  $w$  и обозначать  $\text{НФ}(w)$ . Нетрудно проверить, что несократимая форма любого слова определена однозначно (то есть окончательный результат сокращений не зависит от того, в каком порядке эти сокращения производятся). Слово, полученное из слова  $u$  приписыванием к нему справа слова  $v$ , мы будем обозначать  $uv$ . *Формально обратным* к слову  $u = x_1^{\varepsilon_1} x_2^{\varepsilon_2} \dots x_n^{\varepsilon_n}$ , где  $x_i \in X$  и  $\varepsilon_i \in \{\pm 1\}$ , называют слово  $u^{-1} = x_n^{-\varepsilon_n} x_{n-1}^{-\varepsilon_{n-1}} \dots x_1^{-\varepsilon_1}$ .

СИНТАКСИЧЕСКОЕ ОПРЕДЕЛЕНИЕ СЛЕДСТВИЯ. Соотношение

$$w(x_1, x_2, \dots) = 1$$

называют *следствием* семейства соотношений  $\mathcal{V}$ , если найдутся неотрицательное целое число  $k$ , слова  $v_i \in \mathcal{V}$ , слова  $u_i$  в алфавите  $\{x_1, x_2, \dots\}$  и числа  $\varepsilon_i \in \{\pm 1\}$  такие, что

$$\text{НФ}(w) = \text{НФ}(u_1^{-1} v_1^{\varepsilon_1} u_1 u_2^{-1} v_2^{\varepsilon_2} u_2 \dots u_k^{-1} v_k^{\varepsilon_k} u_k).$$

При  $k = 0$  правая часть этого равенства по определению считается пустым словом.

Два определения следствия на самом деле эквивалентны. В частности, продолжая пример, с которого мы начали этот параграф, несложно показать, что соотношение  $xy = yx$  является синтаксическим следствием соотношений  $x^2 = 1$ ,  $y^2 = 1$  и  $(xy)^2 = 1$ . Действительно, это вытекает из следующего представления коммутатора в виде произведения трёх квадратов:

$$[x, y] = x^{-1} y^{-1} x y = \text{НФ}(x^{-2} x y^{-2} x^{-1} (x y)^2). \quad (*)$$

УТВЕРЖДЕНИЕ 1. Для любого семейства соотношений  $\mathcal{V}$  множество синтаксических следствий этих соотношений совпадает с множеством семантических следствий.

ДОКАЗАТЕЛЬСТВО. То, что каждое синтаксическое следствие является семантическим, совершенно очевидно. Если для элементов  $a_1, a_2, \dots$  некоторой группы  $G$  справедливы равенства  $v_i(a_1, a_2, \dots) = 1$ , то справедливы также равенства  $u_i(a_1, a_2, \dots)^{-1} (v_i(a_1, a_2, \dots))^{\pm 1} u_i(a_1, a_2, \dots) = 1$  и, следовательно, верно равенство

$$\prod_i u_i(a_1, a_2, \dots)^{-1} (v_i(a_1, a_2, \dots))^{\pm 1} u_i(a_1, a_2, \dots) = 1.$$

А переход к несократимой форме вообще не меняет значений слов в группе  $G$ .

Для доказательства обратной импликации рассмотрим множество всех слов в нашем групповом алфавите как полугруппу относительно операции приписывания. Назовём слова  $u$  и  $v$  *эквивалентными* и будем писать  $u \sim v$ , если соотношение  $uv^{-1} = 1$  является синтаксическим следствием соотношений семейства  $\mathcal{V}$ . Это отношение эквивалентности согласовано с операцией приписывания, то есть

$$u \sim u_1, v \sim v_1 \implies uv \sim u_1v_1.$$

Значит, на множество классов эквивалентности корректно определена операция приписывания, относительно которой это множество  $G$  будет полугруппой и даже группой, поскольку  $uu^{-1} \sim 1$  (символом 1 мы здесь обозначаем пустое слово). При этом однобуквенные слова (точнее, их классы эквивалентности) удовлетворяют всем соотношениям семейства  $\mathcal{V}$ . Значит, каждое семантическое следствие  $w = 1$  этих соотношений также должно выполняться на однобуквенных словах. Но по построению это возможно лишь в том случае, когда  $w = 1$  является синтаксическим следствием рассматриваемых соотношений. Детальное доказательство мы оставляем читателям в качестве несложного упражнения.

Синтаксическое определение, в отличие от семантического, делает совершенно очевидным следующий факт.

**ТЕОРЕМА КОМПАКТНОСТИ.** Каждое следствие семейства соотношений является следствием некоторого конечного подсемейства этого семейства.

### 3. ПРИ ЧЁМ ТУТ ГЕОМЕТРИЯ?

Всё, что мы говорили до сих пор о группах, легко переносится на другие алгебраические системы. Заинтересованный читатель может самостоятельно сформулировать и доказать аналогичные факты о соотношениях между элементами, например, кольца или алгебры.

Отличительная черта теории групп, о которой пойдёт речь в этом параграфе, состоит в наличии геометрической интерпретации вывода следствий из соотношений. Эта особенность теории групп позволяет задействовать геометрическую интуицию при решении чисто алгебраических задач.

Начнём с примера. На рисунке 1 слева нарисован четырёхугольник, рёбра которого ориентированы (то есть на них имеются стрелки) и помечены буквами  $x$  и  $y$ . На границе этого четырёхугольника написан коммутатор  $[x, y]$ , то есть мы прочитаем слово  $x^{-1}y^{-1}xy$ , если будем обходить

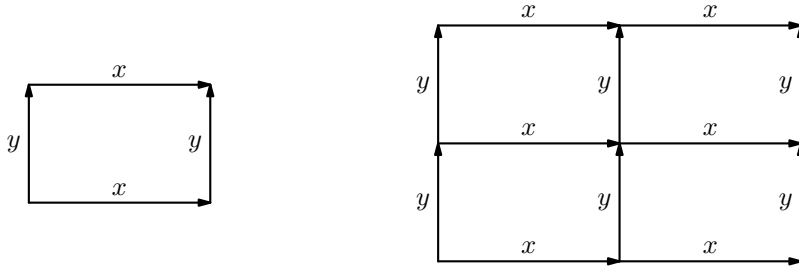


Рис. 1.

границу этого четырёхугольника против часовой стрелки (начав с правой верхней вершины) и читать метки рёбер, причём вместо букв  $x$  и  $y$  читать  $x^{-1}$  и  $y^{-1}$ , соответственно, в тех случаях, когда мы проходим соответствующее ребро против стрелки. Если мы склеим четыре таких четырёхугольника, как показано на рисунке 1 справа, то получим диаграмму, на границе которой написан коммутатор  $[x^2, y^2]$ , то есть  $x^{-2}y^{-2}x^2y^2$ . Этот рисунок, как мы увидим, может служить доказательством того, что соотношение  $[x^2, y^2] = 1$  является следствием соотношения  $[x, y] = 1$ . Более того, у любого следствия любого набора соотношений существует такое геометрическое доказательство.

Необходимые определения выглядят следующим образом. *Картой* мы будем называть конечный связный плоский (то есть нарисованный на плоскости) граф. Такой граф делит плоскость на конечное число областей, называемых *клетками*, ровно одна из которых неограничена. Неограниченную клетку называют *внешней*, остальные клетки называют *внутренними*. Границу внешней клетки называют также *границей карты*. Под *длиной* пути (составленного из рёбер) между двумя вершинами в графе понимается число рёбер на этом пути.

Пусть  $X$  — некоторый алфавит и  $\mathcal{V}$  — некоторый набор слов (соотношений) в алфавите  $X^{\pm 1}$ . Допустим, что у нас имеется карта, рёбра которой ориентированы (то есть на них нарисованы стрелки) и помечены буквами алфавита  $X$ . Под *меткой клетки* мы понимаем слово в алфавите  $X^{\pm 1}$ , которое читается при обходе границы этой клетки против часовой стрелки (при этом буквы, отвечающие рёбрам, которые проходятся против направления ребра, следует заменять на обратные). Метка клетки определена с точностью до циклического сдвига.

Такая размеченная карта называется *диаграммой ван Кампена* над набором соотношений  $\mathcal{V}$ , если метка каждой внутренней клетки является либо одним из соотношений набора  $\mathcal{V}$ , либо формально обратным к соотношению из  $\mathcal{V}$ .

ГЕОМЕТРИЧЕСКОЕ ОПРЕДЕЛЕНИЕ СЛЕДСТВИЯ. Соотношение  $w = 1$  называют *следствием* семейства соотношений  $\mathcal{V}$ , если  $w$  является меткой внешней клетки некоторой диаграммы ван Кампена над семейством  $\mathcal{V}$ .

На рисунке 2 изображено несколько примеров «геометрических импликаций». Что нарисовано справа внизу, догадайтесь сами.

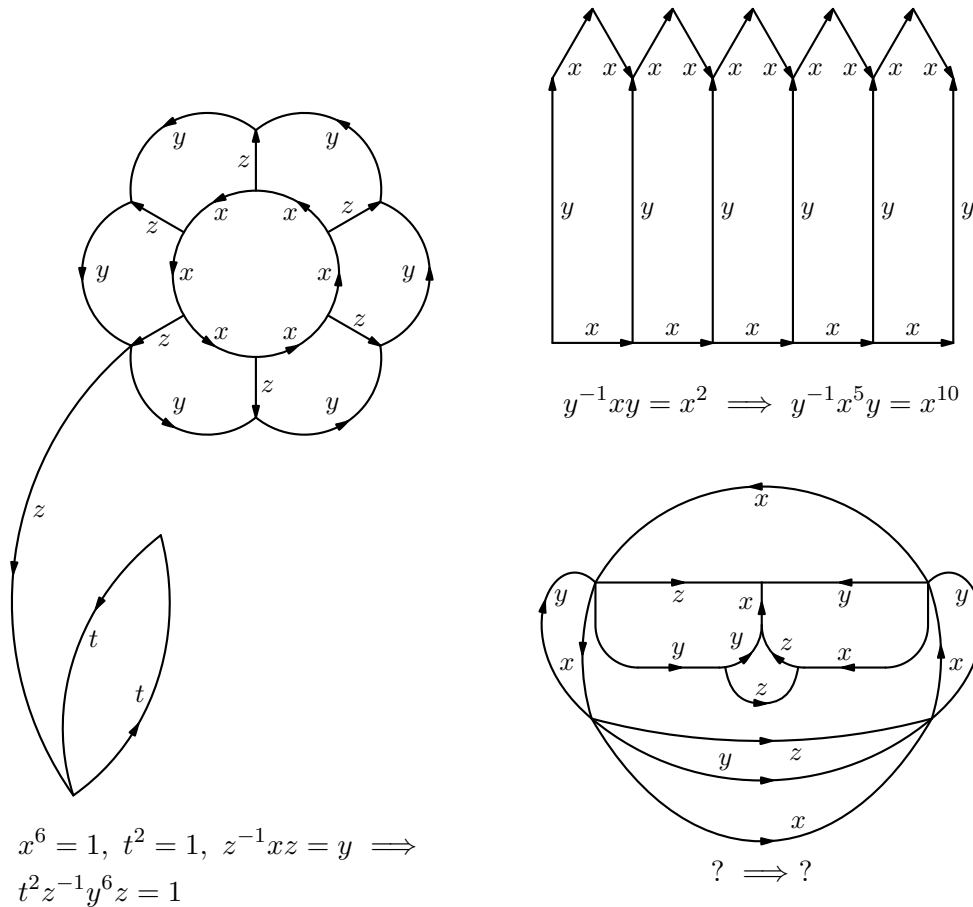


Рис. 2.

ЛЕММА ВАН КАМПЕНА. Для любого семейства соотношений  $\mathcal{V}$  множество геометрических следствий этих соотношений совпадает с множеством синтаксических (а значит, и семантических) следствий.

Мы не будем доказывать эту лемму, но рассмотрим один пример. Формула (\*) показывает, что соотношение  $[x, y] = 1$  является синтаксическим следствием соотношений  $x^2 = 1, y^2 = 1$  и  $(xy)^2 = 1$ . На рисунке 3

слева показана диаграмма, состоящая из отдельных клеток с хвостиками, на границе которой написано слово  $x^{-2}xy^{-2}x^{-1}(xy)^2$ , фигурирующее в правой части формулы (\*).

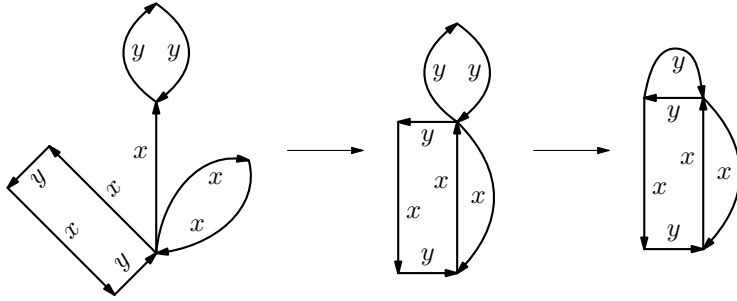


Рис. 3.

Произведя «схлопывания» (рис. 4), мы получим диаграмму, изображённую на рисунке 3 справа, на контуре которой написан коммутатор  $[x, y]$ . Тем самым мы показали, что данное синтаксическое следствие является геометрическим.



Рис. 4.

Если мы хотим, наоборот, показать, что геометрическое следствие является синтаксическим, мы должны «расчленить» данную диаграмму на отдельные клетки с хвостиками. На рисунке 5 показано, как расчленяется диаграмма, изображённая на рисунке 1.

Расчленинная диаграмма даёт следующее выражение коммутатора  $[x^2, y^2]$  через сопряжённые к простым коммутаторам:

$$[x^2, y^2] = (x^{-1}[x, y]x) \cdot ((yx)^{-1}[x, y](yx)) \cdot [x, y] \cdot (y^{-1}[x, y]y).$$

В качестве упражнения мы предлагаем читателю расчленить диаграмму, изображённую на рисунке 6 (заимствованном из книжки [2]), и получить представление коммутатора  $[[x, y], y]$  в виде произведения четырёх кубов. Из возможности такого представления нетрудно вывести, что каждая конечно порождённая группа, в которой все неединичные элементы имеют порядок три, является конечной (и разрешимой).



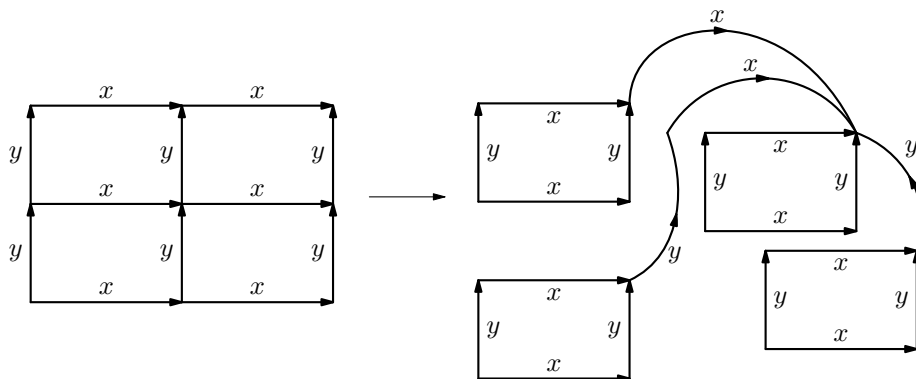


Рис. 5.

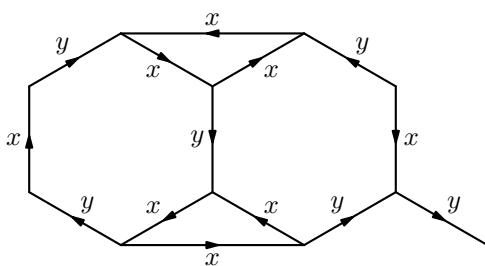


Рис. 6.

#### 4. УСЛОВИЯ МАЛОГО СОКРАЩЕНИЯ

На самом деле, проверить, что что-то является следствием чего-то, — это не самая трудная задача. Гораздо труднее бывает доказать, что что-то НЕ следует из чего-то. Для доказательства отсутствия импликации нужно построить группу, в которой некоторые соотношения выполняются, а некоторые нет. Например, посмотрев на симметрическую группу  $S_3$  и две транспозиции в ней, можно сделать вывод, что соотношение  $[x, y]^{1000} = 1$  не является следствием соотношений  $x^2 = 1$  и  $y^2 = 1$ .

Беда, однако, в том, что построить группу с нужными свойствами бывает не так просто. Более того, может случиться, что никакие традиционные примеры групп заведомо не годятся для достижения нужной цели. Например, импликация

$$y^{-1}x^2y = x^3 \implies [y^{-1}xy, x] = 1$$

выполняется в каждой конечной группе (докажите!) и в каждой линейной группе (то есть в группе невырожденных матриц над полем), но, тем не

менее, эта импликация неверна в классе всех групп. Другими словами, существует очень хитрая группа, в которой для некоторых двух элементов первое соотношение выполняется, а второе нет.

Геометрическая интерпретация вывода следствий из соотношений может помочь доказать отсутствие некоторых импликаций. Самым известным инструментом здесь является *теория малых сокращений*, простейший вариант которой мы рассмотрим в этом параграфе.

Возьмём некоторый набор соотношений  $\mathcal{V}$ . В дальнейшем мы будем предполагать, что все соотношения несократимы и набор  $\mathcal{V}$  является *симметризованным*, то есть вместе с каждым соотношением  $v \in \mathcal{V}$  набор  $\mathcal{V}$  содержит формально обратное соотношение  $v^{-1}$  и все циклические перестановки соотношения  $v$ . Это предположение не ограничивает общности, поскольку и формально обратное соотношение, и циклические перестановки соотношения являются очевидными следствиями исходного соотношения.

Скажем, что симметризованный набор слов (соотношений)  $\mathcal{V}$  удовлетворяет *условию малого сокращения*  $S'(\lambda)$ , где  $\lambda$  — некоторое положительное вещественное число, если для любых двух различных слов  $v_1$  и  $v_2$  из набора  $\mathcal{V}$  длина их общего начала меньше, чем  $\lambda \cdot |v_1|$ .

Рассмотрим, например, набор соотношений

$$\begin{aligned} x^{-1}y^{-1}x y x^{-1}y^{-1}x y &= 1, & y^{-1}x y x^{-1}y^{-1}x y x^{-1} &= 1, \\ x y x^{-1}y^{-1}x y x^{-1}y^{-1} &= 1, & y x^{-1}y^{-1}x y x^{-1}y^{-1}x &= 1, \\ y^{-1}x^{-1}y x y^{-1}x^{-1}y x &= 1, & x^{-1}y x y^{-1}x^{-1}y x y^{-1} &= 1, \\ y x y^{-1}x^{-1}y x y^{-1}x^{-1} &= 1, & x y^{-1}x^{-1}y x y^{-1}x^{-1}y &= 1, \end{aligned} \quad (**)$$

получающийся симметризацией из соотношения  $[x, y]^2 = 1$ . Каждое из слов этого набора имеет длину 8, а общее начало двух слов этого набора состоит не более чем из одной буквы. Таким образом, этот набор удовлетворяет условию  $S'(\lambda)$  тогда и только тогда, когда  $\lambda > \frac{1}{8}$ .

Следующий результат представляет собой простейший (и важнейший) частный случай известной *леммы Гриндлингера*.<sup>1)</sup>

**ОСНОВНАЯ ТЕОРЕМА.** *Если симметризованный набор соотношений  $\mathcal{V}$  удовлетворяет условию  $S'(\frac{1}{6})$  и соотношение  $w = 1$ , где  $w$  — непустое несократимое слово, является следствием набора соотношений  $\mathcal{V}$ , то некоторая циклическая перестановка слова  $w$  имеет с одним из слов*

<sup>1)</sup>Мартин Давидович Гриндлингер (Martin Greendlinger) родился в Нью-Йорке в 1932 году. После защиты диссертации переехал в СССР, долгое время работал в Тульском и Ивановском педагогических институтах; потом вернулся в США. Много информации о его жизни можно найти в рассекреченных архивах ФБР.

$v \in \mathcal{V}$  общее начало  $u$ , длина которого больше, чем половина длины слова  $v$ :

$$|u| > \frac{|v|}{2}.$$

Например, из этой теоремы следует, что соотношение  $[x^{1000}, y^{1000}]^{1000} = 1$  не является следствием соотношения  $[x, y]^2 = 1$ . Действительно, набор соотношений (\*\*), получающийся из соотношения  $[x, y]^2 = 1$  симметризацией удовлетворяет, как мы видели, условию  $C' \left(\frac{1}{6}\right)$ ; следовательно, согласно основной теореме у каждого следствия есть циклическая перестановка, несократимая форма которого обладает с одним из соотношений набора (\*\*) общим подсловом длины большей, чем  $\frac{8}{2} = 4$ . Но максимальные общие подслова циклических перестановок слова  $[x^{1000}, y^{1000}]^{1000}$  и слов набора (\*\*) имеют длину два (это слова  $x^{-1}y^{-1}$ ,  $y^{-1}x$ ,  $xy$  и  $yx^{-1}$ ). Таким образом, мы получаем следующий факт.

**СЛЕДСТВИЕ.** *Существует группа  $G$ , в которой для некоторых двух элементов  $a, b \in G$*

$$[a, b]^2 = 1, \quad \text{но} \quad [a^{1000}, b^{1000}]^{1000} \neq 1.$$

Подобных следствий, разумеется, мы можем получить сколько угодно. На самом деле, можно сказать даже больше:

*если конечный симметризованный набор соотношений  $\mathcal{V}$  удовлетворяет условию  $C' \left(\frac{1}{6}\right)$ , то для любого соотношения  $w = 1$  мы можем быстро определить, является ли оно следствием набора соотношений  $\mathcal{V}$ .*

Действительно, посмотрим на все циклические перестановки слова  $w$ . Если ни одна из них не имеет ни с одним из слов  $v$  набора  $\mathcal{V}$  общего начала длины большей, чем  $|v|/2$ , то, в силу основной теоремы, мы заключаем, что соотношение  $w = 1$  не является следствием соотношений данного набора.

Рассмотрим теперь случай, когда слово  $w$  (или какая-то его циклическая перестановка) имеет длинное общее начало  $u$  с одним из соотношений  $v \in \mathcal{V}$ :

$$w \equiv uf, \quad \mathcal{V} \ni v \equiv ug, \quad |u| > |g|.$$

В этом случае соотношение  $v = 1$  может быть переписано в виде  $u = g^{-1}$ , и соотношение  $w = 1$  (то есть  $uf = 1$ ) является следствием соотношений из  $\mathcal{V}$  тогда и только тогда, когда таким следствием является соотношение  $g^{-1}f = 1$ . Но это соотношение короче чем исходное, поэтому, заменив слово  $w$  на слово  $g^{-1}f$  и повторив эту процедуру несколько раз, в конце концов мы получим ответ. Мы либо придём к пустому слову и таким образом установим, что  $w = 1$  является следствием набора соотношений

$\mathcal{V}$ , либо получим слово, не имеющее длинных общих подслов со словами набора  $\mathcal{V}$  и заключим, что  $w = 1$  не является следствием соотношений набора  $\mathcal{V}$ . Этот простой метод выявления следствий называется *алгоритмом Дэна*.

Для доказательства основной теоремы вам понадобится одно общее наблюдение, касающееся диаграмм ван Кампена. Если в диаграмме есть две клетки, имеющие общее ребро и «зеркально симметричные» относительно этого ребра (рис. 7, слева), то такую пару клеток называют *сократимой парой*. Дело в том, что сократимую пару всегда можно сократить, как показано на рисунке 7. Поэтому диаграммы, о которых идёт речь в определении геометрического следствия, можно считать *приведёнными*, то есть не содержащими сократимых пар.

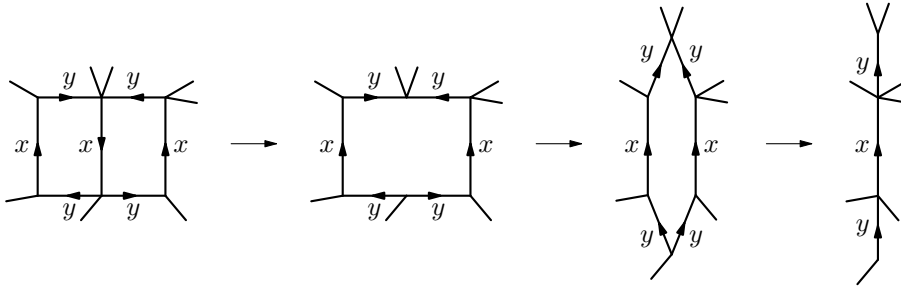


Рис. 7.

В приведённой диаграмме ван Кампена над набором соотношений, удовлетворяющим условию  $C' \left( \frac{1}{6} \right)$ , никакие две внутренние клетки не могут иметь общего участка границы, длина которого больше или равна  $\frac{1}{6}$  периметра одной из этих клеток. Это замечание подсказывает следующее определение.

Если некоторая карта не имеет вершин степени один (то есть каждая вершина лежит по крайней мере на двух рёбрах) и каждый связный общий участок границы любых двух внутренних клеток карты имеет длину (в смысле количества рёбер) меньше, чем  $\frac{1}{6}$  от периметра каждой из этих клеток, то про такую карту говорят, что она *удовлетворяет условию*  $C' \left( \frac{1}{6} \right)$ .

С учётом сделанного выше замечания основная теорема сводится к следующему простому, чисто геометрическому факту.

**УТВЕРЖДЕНИЕ 2.** Пусть имеется карта, удовлетворяющая условию  $C' \left( \frac{1}{6} \right)$ . Тогда найдётся клетка, граница которой имеет общий связный



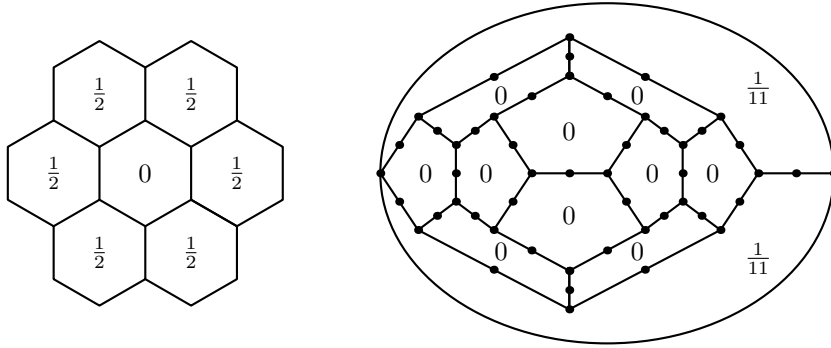


Рис. 9.

## 5. УГЛЫ, КРИВИЗНЫ И ДОКАЗАТЕЛЬСТВО УТВЕРЖДЕНИЯ 2

Рассмотрим некоторую клетку  $D$  карты и вершину  $v$ , лежащую на границе этой клетки. Такую пару  $\alpha$  мы будем называть *углом клетки  $D$  при вершине  $v$* . Число углов при каждой вершине совпадает со степенью этой вершины (то есть с числом рёбер, исходящих из неё), а число углов клетки совпадает с периметром этой клетки.

Следующее простое, но полезное утверждение называют иногда *весовым тестом*.

**ЛЕММА.** Если каждому углу  $\alpha$  каждой клетки некоторой карты поставлено в соответствие число  $\nu(\alpha)$  (которое мы будем называть величиной угла  $\alpha$ ), то

$$\sum_v K(v) + \sum_D K(D) = 4. \quad (***)$$

Здесь суммирование распространяется на все вершины  $v$  и все клетки  $D$  рассматриваемой карты, а величины  $K(v)$  и  $K(D)$ , называемые кривизнами соответствующей вершины и клетки, определяются так:

$$K(v) \stackrel{\text{def}}{=} 2 - \sum_{\alpha} \nu(\alpha), \quad K(D) \stackrel{\text{def}}{=} 2 - \sum_{\alpha} (1 - \nu(\alpha)),$$

где первая сумма распространяется на все углы при вершине  $v$ , а вторая — на все углы клетки  $D$ .

**ДОКАЗАТЕЛЬСТВО.** Заметим, что результат суммирования (\*\*\*) не зависит от величин углов. Действительно, если  $\alpha$  — угол клетки  $D$  при вершине  $v$ , то величина  $\nu(\alpha)$  входит в сумму (\*\*\*) два раза: один раз со знаком минус, когда мы считаем кривизну вершины  $v$ , а другой раз со знаком плюс, когда мы считаем кривизну клетки  $D$ .

Таким образом, равенство (\*\*\*) достаточно доказать для случая, когда все величины углов равны единице. В этом случае наше равенство принимает вид

$$2 \cdot (\text{число вершин}) - (\text{сумма степеней всех вершин}) + 2 \cdot (\text{число клеток}) = 4.$$

Сумма степеней всех вершин совпадает с удвоенным числом рёбер карты. Поэтому доказываемое равенство сводится к формуле Эйлера

$$(\text{число вершин}) - (\text{число рёбер}) + (\text{число клеток}) = 2,$$

которая легко доказывается по индукции.

Величины  $\nu(\alpha)$  можно представлять себе как обычные величины углов, измеренные в радианах, умноженных на  $\pi$ . Нулевая кривизна будет тогда соответствовать обычной евклидовой геометрии (например, сумма всех углов треугольника будет равна единице); отрицательная кривизна делает ситуацию похожей на геометрию Лобачевского, а положительная кривизна — на сферическую геометрию.

Приступим теперь к доказательству утверждения 2. Пусть у нас имеется карта, удовлетворяющая условию  $C' \left( \frac{1}{6} \right)$ . Припишем углам величины по следующим правилам:

- а) всем углам внешней клетки припишем величину 1;
- б) углу внутренней клетки, находящемуся при вершине степени  $d = k + l$ , при которой имеется  $k$  углов внешней клетки и  $l$  углов внутренних клеток, припишем величину  $2 - \frac{k}{l}$ .

Из этих правил вытекает, что

- каждая вершина имеет нулевую кривизну;
- внешняя клетка имеет кривизну 2.

Согласно лемме это означает, что должна найтись внутренняя клетка с положительной кривизной.

Внутренняя клетка, не граничащая с внешней, будет иметь отрицательную кривизну, поскольку из условия  $C' \left( \frac{1}{6} \right)$  следует, что у такой клетки будет по крайней мере 7 соседних клеток и, следовательно, по крайней мере 7 её углов имеют величину  $\leq \frac{2}{3}$ , а величины остальных углов не превосходят единицу (рисунок 10, слева). Вообще, несложно заметить, что кривизна внутренней клетки  $D$  может быть положительной лишь в случае, когда эта клетка имеет ровно один связный участок общей границы с внешней клеткой и граничит ещё не более чем с тремя внутренними



Рис. 10.

клетками. Из условия  $C' \left( \frac{1}{6} \right)$  следует, что в таком случае клетка  $D$  выходит на границу всей карты более чем половиной своего периметра, что и доказывает утверждение 2, а вместе с ним и основную теорему.

На рисунке 10 представлены некоторые из возможных вариантов расположения внутренней клетки относительно внешней. В центре клетки написана её кривизна.

Познакомиться поближе с теорией малых сокращений, её обобщениями и применением других геометрических соображений в комбинаторной теории групп можно по книгам [1] и [2].

Автор очень признателен А. Ю. Ольшанскому, который нашёл несколько существенных неточностей в предварительной версии этой статьи. Автор благодарит также Д. В. Баранова, А. О. Захарова, О. В. Куликову, Е. В. Френкель и других участников семинара по теории групп МГУ за то, что они любезно согласились прочитать этот текст и высказать свои замечания.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Линдон Р., Шуш П. Комбинаторная теория групп. – М.: Мир, 1980.
- [2] Ольшанский А. Ю. Геометрия определяющих соотношений в группах. – М.: Наука, 1989.

---

Ант. А. Клячко, механико-математический факультет Московского государственного университета  
 Москва 119991, Ленинские горы, МГУ, мех-мат  
 email: klyachko@mech.math.msu.su



## Фуксовы группы: от топологии к геометрии

М. Н. Вялый      О. В. Шварцман

Группа  $\Gamma$  движений плоскости Лобачевского  $\Lambda^2$  называется *дискретной*, если орбита  $\{\Gamma x\}$  любой точки  $x \in \Lambda^2$  является дискретным множеством плоскости. Дискретные группы движений плоскости Лобачевского называются *фуксовыми группами*. С помощью фуксовых групп бесформенным топологическим поверхностям удалось придать строгую геометрическую форму. К радости как самой поверхности, так и почти всей современной математики, новый геометрический костюм оказался нежестким.

Хотя наша цель — фуксовы группы, начать нужно с поверхностей. Начнем немедленно, а в качестве отличного путеводителя по теме рекомендуем книгу В. Г. Болтянского и В. А. Ефремовича [1].

### 1. ПОВЕРХНОСТИ КАК ПРЕДМЕТ ИСКУССТВА КРОЙКИ И СКЛЕЙКИ

1.1. Топология не видит особой разницы между поверхностью гири и бублика. Обе они напоминают поверхность под именем тор, изображенную на рис. 1.

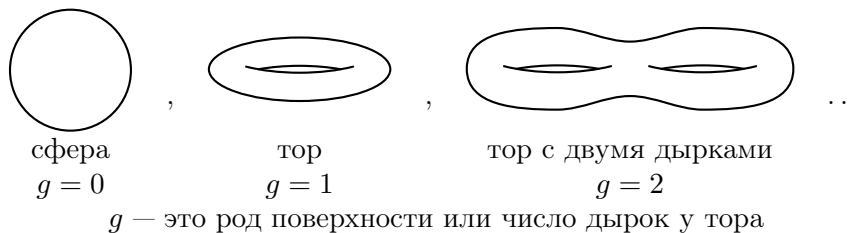


Рис. 1.

Рассматривая поверхности гири и бублика, тополог скажет, что они *гомеоморфны* (и считаются в топологии одинаковыми). Более точно, фигуры  $A$  и  $B$  называются гомеоморфными, если между их точками можно установить взаимно однозначное и взаимно непрерывное соответствие  $f: A \rightarrow B$ . Последнее означает, что как само отображение  $f$ , так и его теоретико-множественное обращение, являются непрерывными отображениями.

Представим клоунскую гирию, поверхность которой сделана из эластичной пленки. С топологической точки зрения эту пленку можно как угодно деформировать, не допуская разрывов. Имея в виду такую свободу, не так уж сложно вообразить деформацию, после которой поверхность гирии переходит в поверхность бублика.

1.2. Раньше квартиры обогревались в зимнюю стужу многосекционными батареями. Поверхность такой батареи (если пожертвовать техническими мелочами) представляет из себя с топологической точки зрения одну из поверхностей ряда, представленного на рис. 1 (число дыр равно удвоенному числу секций минус один).

Гомеоморфна ли, скажем, шестисекционная батарея трехсекционной? Чувствуется, что нет. Но как это доказать? Для этого можно использовать *топологический инвариант*, т. е. величину, одинаковую для всех гомеоморфных поверхностей. Таким инвариантом, различающим поверхности, окажется *характеристика Эйлера* (см. ниже раздел 2).

1.3. Пора договориться о том, что в этом очерке будет называться поверхностью. Это будет связная, компактная фигура в  $E^3$ , причем такая, что у каждой ее точки есть окрестность, гомеоморфная открытому диску на плоскости.

Впрочем, поверхности с краем тоже важны: если на торе вырезать «круглую» дыру, то получится компактная поверхность с краем, которая называется ручкой. Склеивая две ручки их краями, получим тор с двумя дырами или крендель (третья слева поверхность на рис. 1). Следующую поверхность в ряду, намеченном на рис. 1, можно получить, склеивая продырявленный крендель и ручку. Итак, шаг вправо на рис. 1 осуществляется операцией приклеивания ручки. Запомним это.

Аналогичным образом можно получить двухсекционную батарею склейкой двух односекционных батарей.

1.4. Поверхность можно получить топологической склейкой сторон многоугольника. Нарисуем на плоскости  $2n$ -угольник  $L$ , ограничивающий топологический диск  $D$ . Разобьем произвольным образом множество его сторон на пары, отметив одной и той же буквой стороны, попавшие в одну пару. Выберем направление обхода многоугольника, после чего можно говорить о начале и конце стороны. Стороны, попавшие в одну пару, предназначены для топологической склейки. Что это означает? Каждая сторона гомеоморфна отрезку. Склеить стороны в паре  $(a, a)$  означает рассмотреть такой гомеоморфизм  $\varphi_a$  указанных отрезков, который начало одного переводит в конец другого. Способ склейки — важный момент в нашей истории: если хотя бы в одной паре приклеить начало к началу, то склейку нельзя будет вложить в  $E^3$  без самопересечений (доказать это непросто).

Вершина 2 склеивается с вершиной 5, а вершина 1 — с вершиной 6.

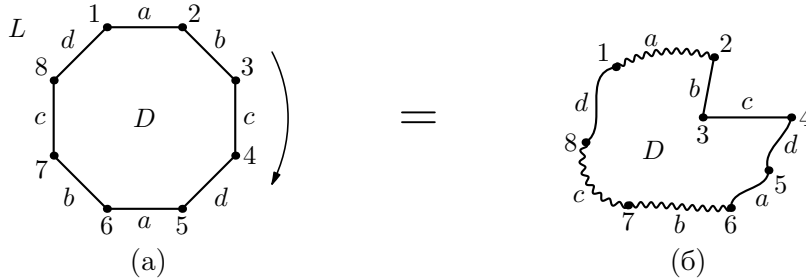


Рис. 2.

Многоугольник  $L$  с разбитыми на пары сторонами мы будем называть *выкройкой*. Через  $X(L)$  мы будем обозначать склейку по выкройке  $L$ .

На рис. 2(а) изображена восьмиугольная выкройка  $L$ . Строгая форма этой выкройки не должна вводить в заблуждение: с равным успехом мы могли бы работать и с чем-то таким, как на рис. 2(б).

Обход многоугольника  $L$  дает циклическое слово  $w = abcdabcd$  (потому циклическое, что его можно читать с любого места), составленное из букв на его сторонах. Ясно, что циклическое слово  $w$  полностью определяет выкройку.

Можно доказать, что любая поверхность  $S$  допускает такую выкройку  $L$  (и далеко не одну!), что  $X(L) = S$ . Наметим план доказательства, оставляя детали читателю.

Прежде всего нужно триангулировать поверхность, т.е. разбить ее на конечное число топологических треугольников  $T_i$  (фигур, гомеоморфных плоскому треугольнику), правильно примыкающих друг к другу. Это означает, что никакие два треугольника разбиения не имеют общей внутренней точки, и что любая сторона треугольника принадлежит в точности двум треугольникам разбиения. Возможность триангуляции вытекает из компактности поверхности.

Искомую выкройку  $L$  будем строить по шагам. Сначала выберем какой-нибудь треугольник  $T_1$  и сопоставим ему плоский треугольник  $L_1$ . На следующем шаге в качестве  $T_2$  возьмем любой треугольник разбиения, имеющий общую сторону с треугольником  $T_1$ . Склеим  $T_1$  и  $T_2$  по общей стороне и сопоставим полученному объединению плоский четырехугольник  $L_2$ . Далее в качестве  $T_3$  возьмем любой из треугольников, имеющих общую сторону с  $T_1$  или  $T_2$ , и сопоставим ему многоугольник  $L_3$ , полученный приклеиванием треугольника к соответственной стороне  $L_2$ . Добавляя таким образом на каждом шаге новый треугольник, мы будем получать всякий раз многоугольник, ограничивающий диск.

Если новый треугольник имеет более одной стороны, общей с предыдущими треугольниками, то вторую (а, если нужно, и третью) пару общих сторон мы отмечаем одинаковыми буквами, запоминая, что их нужно склеить.

Так как поверхность связна, то этот процесс закончится не раньше, чем мы переберем все треугольники разбиения (почему?). Поэтому у последнего многоугольника  $L_t = L$  не останется неотмеченных сторон. Значит, склейка сторон многоугольника  $L$  восстанавливает поверхность  $S$ .

## 2. ЭЙЛЕРОВА ХАРАКТЕРИСТИКА ВЫКРОЙКИ. ЭЛЕМЕНТАРНЫЕ ПРЕОБРАЗОВАНИЯ. КАНОНИЧЕСКИЕ ФОРМЫ

2.1. Назовем две вершины выкройки эквивалентными, если они склеиваются в одну точку. Если  $v$  — число классов эквивалентных вершин, то назовем число  $e(L) = 1 + v - n$  *эйлеровой характеристикой* выкройки  $L$  (напомним, что число сторон  $L$  равно  $2n$ ).

Введем два элементарных преобразования выкроек, каждое из которых не меняет ни эйлеровой характеристики выкройки, ни топологического типа ее склейки. Последнее означает, что если выкройки  $L, L'$  получены одна из другой с помощью элементарного преобразования, то склейки  $X(L)$  и  $X(L')$  оказываются гомеоморфными.

Элементарное преобразование первого типа уничтожает пару смежных одинаково отмеченных сторон в склейке  $L$  с числом сторон не меньше, чем четыре. На рисунке 3 показано, как действует это преобразование. Заметим, что при уничтожении сторон четырехугольника мы получаем двуугольник. В элементарной геометрии такой фигуры, конечно, нет. Однако топологический двуугольник может иметь «кривые» стороны. При склейке двуугольной выкройки получается сфера.

Мы оставляем читателю проверку того, что  $e(L) = e(L')$  и склейка  $X(L)$  гомеоморфна  $X(L')$ . Отметим также, что преобразование первого типа всегда уменьшает оба числа  $n$  и  $v$  на 1.

Элементарное преобразование второго типа разрезает многоугольник  $L$  по диагонали на две части, в каждой из которых содержится сторона,

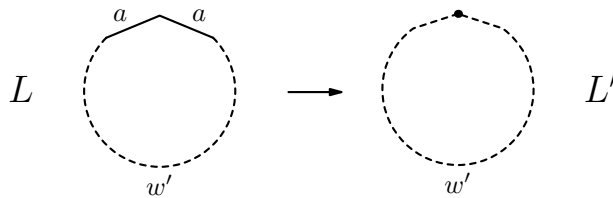


Рис. 3.

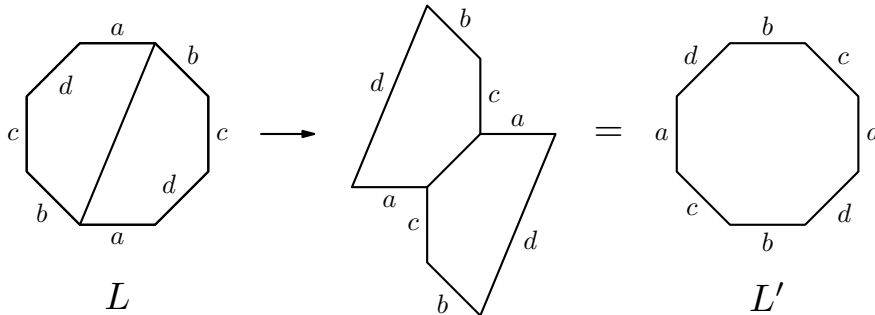


Рис. 4.

отмеченная, скажем, буквой  $d$  (см. рис. 4). Затем левый кусок приклеивается к правому по стороне  $d$ . Новая выкройка  $L'$  «поглощает» сторону  $d$ , а из диагонали рождаются две новые стороны, которые наследуют имя исчезнувшей стороны (см. рис. 4). Легко проверить, что такое преобразование не меняет чисел  $n$  и  $v$ , так что и эйлерова характеристика не изменяется. Кроме того, выкройка  $L'$  склеивается в ту же поверхность, что и  $L$  (мы ведь просто сначала распороли, а затем склеили то же самое в другом порядке).

С помощью элементарных преобразований первого и второго типа мы можем добиться того, чтобы

- а) никакие две одинаковые буквы в циклическом слове  $w$  не стояли рядом (если речь идет о поверхности, отличной от сферы);
- б) все вершины выкройки были эквивалентны ( $v = 1$ ).

Мы пропускаем доказательство этого утверждения. Читатель может самостоятельно решить эту задачу или обратиться к начальным курсам по топологии, например, [2, гл. 1].

Выкройку, удовлетворяющую условиям а) и б), назовем *удобной*.

2.2. Удобную выкройку можно привести к стандартному виду (разложение на ручки, см. [7]).

**ТЕОРЕМА 1.** *Любую удобную выкройку с помощью элементарных преобразований второго типа можно привести к виду*

$$w = \underbrace{ababcdcdefef\dots}_{4g \text{ букв}}$$

(см. рис. 5). Эйлерова характеристика такой выкройки равна  $2 - 2g$ .

На стандартной выкройке физически ощутимо прослеживается  $g$ -кратное приклеивание ручки в  $E^3$  ( $abab$  — это ведь тор). Но с помощью этой

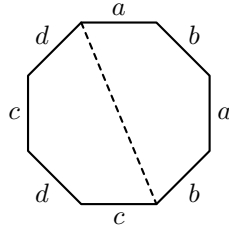


Рис. 5.

операции строится весь ряд поверхностей на рис. 1. Следовательно, любая поверхность  $S$  гомеоморфна ровно одной поверхности этого ряда, и эйлерова характеристика определяет топологический тип поверхности.

2.3. Кроме стандартной, есть и другие полезные канонические выкройки для поверхности рода  $g$  (эйлерова характеристика  $2 - 2g$ ).

**ТЕОРЕМА 2.** Любую удобную выкройку  $L$  с эйлеровой характеристикой  $e(L) = 2 - 2g$  с помощью элементарных преобразований второго типа можно привести к виду

$$w = ui, \quad \text{где } u \text{ — слово из } 2g \text{ букв.}$$

**ЗАМЕЧАНИЕ.** При  $g = 1$  такая выкройка совпадает со стандартной.

Обе теоремы доказываются одновременно индукцией по  $g$ , которая базируется на сделанном выше замечании. Индуктивный переход опирается на следующую ключевую лемму.

**ЛЕММА.** Пусть в удобной выкройке встречаются две разделяющие друг друга пары букв:  $w = \dots a \dots b \dots a \dots b \dots$ . Тогда можно преобразовать выкройку так, что эти буквы будут непосредственно следовать друг за другом:  $w = \dots abab \dots$ .

**ДОКАЗАТЕЛЬСТВО.** См. рис. 6.

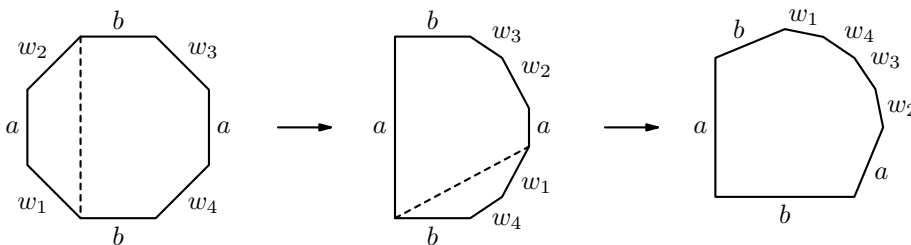
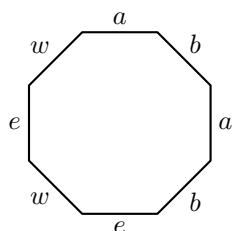


Рис. 6.

Читателю мы оставляем проверку того факта, что в удобной выкройке разделяющие пары действительно встречаются. (Подсказка: докажите, что если в выкройке нет разделяющих пар, то это выкройка сферы.) Отсюда с учетом леммы сразу получается доказательство теоремы 1.

Продолжим доказательство теоремы 2. Согласно индуктивному предположению можно считать, что перед нами выкройка вида, изображенного на рис. 7.



Под словом  $w$  спрятана целая ломаная, а  $e, a, b$  — это просто буквы.

Рис. 7.

С помощью двух элементарных преобразований второго типа получаем каноническую выкройку (рис. 8).

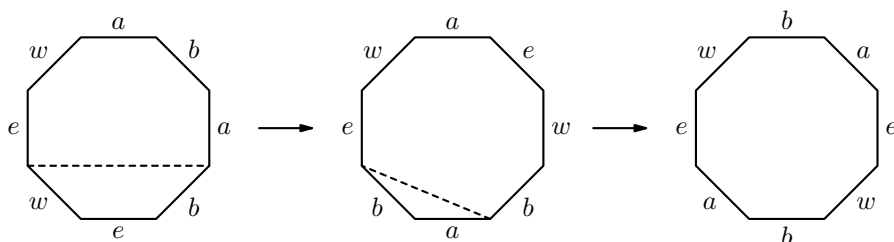


Рис. 8.

### 3. ГЕОМЕТРИЧЕСКИЕ СКЛЕЙКИ И ФУКСОВЫ ГРУППЫ

3.1. Рассмотрим на плоскости Лобачевского правильный восьмиугольник  $\mathcal{P}_0$  с углом  $\pi/4$  при вершине.

В этом параграфе мы построим фуксову группу  $\Gamma$ , обладающую следующими свойствами:

- 1) всевозможные сдвиги  $\{\gamma\mathcal{P}_0\}$ ,  $\gamma \in \Gamma$ , восьмиугольника  $\mathcal{P}_0$  с помощью элементов из группы  $\Gamma$  образуют замощение плоскости Лобачевского плитками, конгруэнтными  $\mathcal{P}_0$ ;
- 2) никакие две различные внутренние точки  $\mathcal{P}_0$  не переводятся друг в друга движениями из группы  $\Gamma$ ;

- 3) напротив, мы укажем элементы группы  $\Gamma$ , отождествляющие противоположные стороны  $\mathcal{P}_0$  (склеивающие их геометрически) по выкройке  $w = abcdabcd$ .

Из свойств 1) и 2) следует, что орбита  $\{\Gamma x\}$  любой точки  $x \in \Lambda^2$  пересекает многоугольник  $\mathcal{P}_0$  не менее, чем в одной точке, а внутренность  $\mathcal{P}_0$  — не более, чем в одной точке<sup>1)</sup>. Таким образом, пространство орбит группы  $\Gamma$  на плоскости  $\Lambda^2$  или, что то же самое, факторпространство  $X = \Lambda^2/\Gamma$ , есть не что иное, как склейка сторон  $\mathcal{P}_0$  по выкройке, т. е. поверхность рода 2. Можно сказать, что фундаментальный многоугольник  $\mathcal{P}_0$  — это геометрическая выкройка поверхности рода 2, аналогичная топологической выкройке, рассмотренной в п. 1.4 (рисунок 2).

Как будет ясно из дальнейшего, все наши построения проходят для любого правильного  $4g$ -угольника с углом  $\pi/2g$  при вершине. В этом случае геометрическая склейка  $X = \Lambda^2/\Gamma$  является поверхностью рода  $g$ . При  $g = 1$  события развиваются на евклидовой плоскости, где противоположные стороны единичного квадрата склеиваются с помощью группы  $\Gamma = \mathbb{Z}^2$  целочисленных параллельных переносов плоскости. Начиная с рода 2, уже нет возможности нарисовать правильный  $4g$ -угольник с углом  $\pi/2g$  на евклидовой плоскости. Зато такую возможность предоставляет плоскость Лобачевского  $\Lambda^2$ , а роль группы склейки переходит к фуксовой группе  $\Gamma$ .

Чтобы следить за выполнением этой программы, полезно в качестве нового путеводаителя выбрать отличную статью Э. Б. Винберга [3].

3.2. Рассмотрим на плоскости Лобачевского треугольник  $S$  с углами  $\pi/2$ ,  $\pi/8$ ,  $\pi/8$ , и через  $T$  обозначим группу, порожденную отражениями  $r_1$ ,  $r_2$ ,  $r_3$  относительно сторон этого треугольника. Про группу  $T$  (треугольный гиперболический калейдоскоп по Э. Б. Винбергу) известно, в частности, следующее.

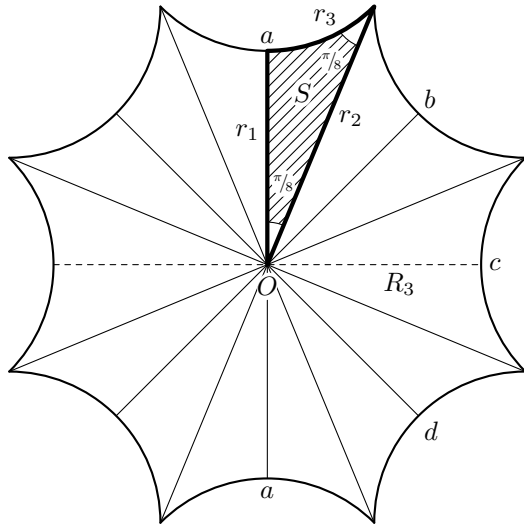
а) Группа  $T$  задается образующими  $r_1$ ,  $r_2$ ,  $r_3$  с определяющими соотношениями  $r_1^2 = r_2^2 = r_3^2 = (r_1 r_2)^8 = (r_2 r_3)^8 = (r_1 r_3)^2 = 1$  (определения образующих и соотношений приводятся ниже в разделе 4 на с. 43).

б) Многократные отражения треугольника  $S$  относительно его сторон приводят к замощению плоскости  $\Lambda^2$  конгруэнтными копиями  $S$ . Это означает, что  $\Lambda^2 = \bigcup_{t \in T} t(S)$ ,  $t$  пробегает всю группу  $T$ , причем треугольники  $t_1(S)$  и  $t_2(S)$  при  $t_1 \neq t_2$  либо не имеют общих точек, либо имеют единственную общую вершину, либо прилегают друг к другу по целой стороне.

На рисунке 9 изображен фрагмент замощения, состоящий из 16 треугольников, выложенных в правильный восьмиугольник  $\mathcal{P}_0$  с углом  $\pi/4$  и центром в точке  $O$ .

<sup>1)</sup>В теории дискретных групп преобразований всякий многоугольник с такими свойствами называется *фундаментальным многоугольником* для соответствующей группы.





Отражения  $r_1$  и  $r_2$  порождают подгруппу диэдра  $D_8$ , оставляющую точку  $O$  на месте. Треугольники, полученные из треугольника  $S$  с помощью движений из группы диэдра, заполняют правильный восьмиугольник  $\mathcal{P}_0$ .

Рис. 9.

в) Подгруппа  $T_0$  группы  $T$ , порожденная отражениями  $r_1$  и  $r_2$ , состоит ровно из тех движений группы  $T$ , которые оставляют точку  $O$  на месте.

Ясно, что  $T_0$  — это группа симметрии правильного восьмиугольника  $\mathcal{P}_0$ , т. е. группа диэдра  $D_8$ . Кроме того,  $\mathcal{P}_0 = \bigcup_{t \in T_0} t(S)$ .

г) Никакие две разные точки треугольника  $S$  не эквивалентны под действием группы  $T$ , т. е. не переводятся одна в другую нетождественным движением из этой группы.

Заметим, что, окружив каждую точку  $t(O)$ ,  $t \in T$ , правильным восьмиугольником  $\mathcal{P}_{t(O)}$ , мы получаем замощение плоскости  $\Lambda^2$  правильными восьмиугольниками, конгруэнтными восьмиугольнику  $\mathcal{P}_0$  (для краткости будем говорить о замощении плиткой  $\mathcal{P}_0$ ). Это замощение инвариантно под действием группы  $T$ .

3.3. 1) Построим группу  $\Gamma$ . Для этого рассмотрим отображение  $\varphi$  множества образующих  $\{r_1, r_2, r_3\}$  группы  $T$  в группу  $T_0$ , которое отражения  $r_1$  и  $r_2$  переводит в себя, а отражение  $r_3$  — в то отражение  $R_3$  из группы диэдра  $T_0$ , зеркало которого ортогонально зеркалу отражения  $r_1$  (см. рис. 9).

Образы  $\varphi(r_i)$ ,  $i = 1, 2, 3$ , удовлетворяют всем определяющим соотношениям группы  $T$  (проверьте это!). Следовательно, отображение  $\varphi$  продолжается до гомоморфизма  $\varphi: T \rightarrow T_0$  группы  $T$  на группу  $T_0$  (следует учесть, что отражения  $r_1$  и  $r_2$  порождают  $T_0$ ).

Ядро Кер  $\varphi = \Gamma$  и есть искомая группа. Отметим на будущее, что  $\Gamma \cap T_0 = \{1\}$ , поскольку на  $T_0$  гомоморфизм  $\varphi$  действует тождественно.

2) Покажем, что никакие разные внутренние точки  $\mathcal{P}_0$  не эквивалентны под действием группы  $\Gamma$ . Предположим, что это не так, и  $\gamma(x) = y$  для пары внутренних точек  $x \neq y$  и некоторого движения  $\gamma$  из  $\Gamma$ . Если точки  $x$  и  $y$  принадлежат одному треугольнику, то мы вступаем в противоречие с п. 3.2г). Если же они лежат в разных треугольниках, то с помощью движения  $t_0$  из группы  $T_0$  добьемся того, чтобы точки  $t_0x$  и  $y$  лежали внутри  $\mathcal{P}_0$  в одном треугольнике. Но  $(\gamma t_0^{-1})(t_0x) = y$ . Снова принимая во внимание п. 3.2г), заключаем, что такое возможно, только если  $t_0(x) = y$ , а  $\gamma t_0^{-1}$  элемент группы  $T$ , переводящий плитку  $\mathcal{P}_0$  в себя (почему?) и оставляющий точку  $y$  на месте. Но такой элемент лежит в группе  $T_0$  согласно п. 3.2в). Следовательно, группе  $T_0$  принадлежит и элемент  $\gamma \in \Gamma$ . Так как  $\Gamma \cap T_0 = \{1\}$ , получаем  $\gamma = 1$  и  $x = y$ . Противоречие.

3) Покажем, наконец, что в группе  $\Gamma$  есть элементы, отождествляющие противоположные стороны  $\mathcal{P}_0$ , и что никаких «лишних» склеек точек границы  $\mathcal{P}_0$  группа  $\Gamma$  не производит.

На этом этапе удобно изменить разметку сторон восьмиугольника  $\mathcal{P}_0$ . А именно, стороны, попавшие в одну пару, обозначать одной и той же буквой алфавита, но один раз — строчной, а другой — прописной буквой. Такая разметка изображена на рис. 10.

Обозначим через  $a$  движение из группы  $\Gamma$ , равное композиции  $r_3 R_3$  двух отражений из группы  $T$ . (Контрольный вопрос: почему  $a$  принадлежит группе  $\Gamma$ ?) Движение  $a$  переводит сторону  $A$  в сторону  $a$  в соответствии с правилом топологической склейки из п. 1.4, а плитку  $\mathcal{P}_0$  — в плитку, смежную с ней по стороне  $a$ . Аналогично,  $A = a^{-1} = R_3 r_3$  переводит  $\mathcal{P}_0$  в плитку, смежную с ней по стороне  $A$ .

Аналогично определяются движения  $b, B, c, C, d, D$  из группы  $\Gamma$ .

Предположим, что группа  $\Gamma$  склеила две точки границы (см. рис. 10), т. е.  $\gamma(x) = y$ . При этом элемент  $\gamma$  не переводит плитку  $\mathcal{P}_0$  в себя, так как

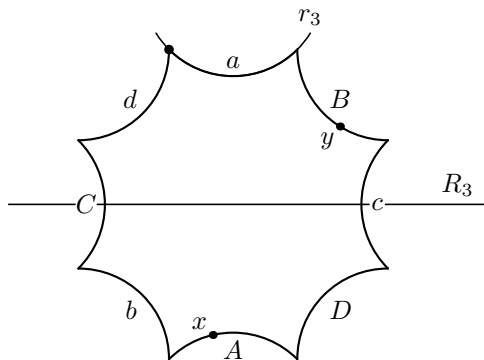


Рис. 10.

группа симметрий  $\mathcal{P}_0$  есть  $T_0$ . Следовательно,  $\gamma$  переводит плитку  $\mathcal{P}_0$  в плитку, смежную с ней по стороне  $B$ . Тогда движение  $B^{-1}\gamma$  лежит в  $\Gamma$  и переводит  $\mathcal{P}_0$  в себя. Следовательно,  $B^{-1}\gamma = 1$ , т. е.  $\gamma = B$ , а точка  $x$  обязана лежать на стороне  $b$ .

Намеченная нами программа полностью выполнена.

УПРАЖНЕНИЕ. Докажите, что группа  $\Gamma$  действует на плоскости Лобачевского без неподвижных точек.

#### 4. ОБРАЗУЮЩИЕ И ОПРЕДЕЛЯЮЩИЕ СООТНОШЕНИЯ ДЛЯ ФУКСОВОЙ ГРУППЫ $\Gamma$

Напомним, что элементы, порождающие группу, называются ее образующими. Между образующими имеются, вообще говоря, соотношения. Множество соотношений называется определяющим, если любое другое соотношение есть их следствие (смотри статью А. А. Клячко [5] в этом номере).

Рассмотрим построенную в п. 3 фуксову группу  $\Gamma$ , которая геометрически склеивает поверхность рода 2 по выкройке  $w = aBcDAbCd$ .

ТЕОРЕМА. *Группа  $\Gamma$  порождается элементами  $a, b, c, d, A, B, C, D$  с единственным нетривиальным определяющим соотношением*

$$abcdABCD = 1.$$

(Тривиальные определяющие соотношения суть  $aA = bB = cC = dD = 1$ .)

4.1. Докажем, что группа  $\Gamma$  порождается движениями  $a, b, c, d, A, B, C, D$ , каждое из которых переводит плитку  $\mathcal{P}_0$  в одну из смежных с ней плиток замощения.

Рассмотрим элемент  $\gamma \in \Gamma$  и плитку  $\gamma(\mathcal{P}_0) = \mathcal{P}_\gamma$ . Заметим, что есть ровно один элемент группы  $\Gamma$  с таким свойством, так как только единичный элемент группы  $\Gamma$  сохраняет плитку  $\mathcal{P}_0$ . Соединим плитку  $\mathcal{P}_0$  с плиткой  $\mathcal{P}_\gamma$  цепью плиток  $(\mathcal{P}_0, \mathcal{P}_1, \dots, \mathcal{P}_k = \mathcal{P}_\gamma)$ , в которой плитки  $\mathcal{P}_{i-1}$  и  $\mathcal{P}_i$  смежны по стороне.

Покажем, что элемент  $\gamma$  можно представить в виде произведения преобразований  $a, b, c, d, A, B, C, D$ . Для этого воспользуемся тем, что плитка  $\mathcal{P}_0$  отмечена циклическим словом  $w$ . Предположим, что в соседнюю плитку  $\mathcal{P}_1$  мы переходим через сторону плитки  $\mathcal{P}_0$ , отмеченную буквой  $B$ . Тогда  $\mathcal{P}_1 = B(\mathcal{P}_0)$ . Оказавшись в плитке  $\mathcal{P}_1$ , напишем на только что пройденной стороне букву  $b$ , а на остальных сторонах — такие буквы, чтобы получилось (циклическое) слово  $w$ . Это можно сделать единственным образом. Общая сторона плиток  $\mathcal{P}_1$  и  $\mathcal{P}_2$  окажется отмеченной какой-то из наших букв. Пусть это будет буква  $c$ . Тогда  $Bc(\mathcal{P}_0) = \mathcal{P}_2$ . Именно  $Bc$ , а не  $cB$ ,

как вам, быть может, показалось! Дело в том, что движение, переводящее плитку  $\mathcal{P}_1$  в плитку  $\mathcal{P}_2$ , лишь сопряжено  $c$  и равно  $BcB^{-1}$ .

Далее продолжаем по тому же правилу. Войдя в плитку  $\mathcal{P}_2$ , напомним только что пройденной стороне букву  $C$  и восстановим циклическое слово  $w$  на плитке  $\mathcal{P}_2$ . Если сторона, через которую предстоит пройти в  $\mathcal{P}_3$ , оказалась отмеченной буквой  $A$ , то  $BcA(\mathcal{P}_0) = \mathcal{P}_3$ .

Добравшись до плитки  $\mathcal{P}_\gamma$ , мы получим искомое представление  $\gamma$  в виде произведения элементов  $a, b, c, d, A, B, C, D$ .

ЗАМЕЧАНИЕ. Мы изобразили аналогичный процесс на рисунке 11. Поскольку рисовать замощение плоскости Лобачевского восьмиугольниками тяжело, без ущерба для сути дела мы нарисовали замощение евклидовой

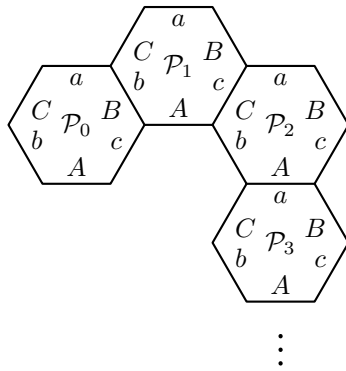


Рис. 11.

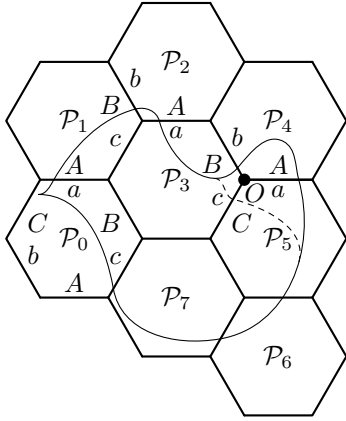
плоскости шестиугольными сотами, которые склеивают тор по выкройке  $aBcAbC$ . Группа склейки  $\Gamma$  в этом случае есть дискретная группа параллельных переносов плоскости на векторы решетки (подумайте, какие именно).

В процессе доказательства мы по каждой цепочке плиток построили слово, составленное из букв  $a, b, c, d, A, B, C, D$ . Ясно, что можно обратить эту конструкцию и построить по любому такому слову соответствующую цепочку плиток: если, например, наше слово равно  $AbCca$ , то по нему строится цепочка  $\mathcal{P}_0, \mathcal{P}_1 = A(\mathcal{P}_0), \mathcal{P}_2 = Ab(\mathcal{P}_0), \mathcal{P}_3 = AbC(\mathcal{P}_0), \mathcal{P}_4 = AbCc(\mathcal{P}_0) = Ab(\mathcal{P}_0), \mathcal{P}_5 = AbCca(\mathcal{P}_0)$ .

4.2. Переходим к определяющим соотношениям. Любое соотношение в группе  $\Gamma$  имеет вид

$$\omega = 1,$$

где в левой части записано слово в алфавите  $\{a, b, c, d, A, B, C, D\}$ . Согласно п. 4.1 такое слово  $\omega$  определяет цепочку плиток, которая начинается



Случай замощения евклидовой плоскости шестиугольниками. Соотношение (а, точнее, цепочку плиток  $\mathcal{P}_0, \mathcal{P}_1, \dots, \mathcal{P}_7, \mathcal{P}_0$ ) мы «снимаем с гвоздя»  $O$ , заменяя фрагмент  $\dots BA \dots$  на  $\dots c \dots$ , используя соотношения обхода  $BAC = 1$  или  $BA = c$ .

Рис. 12.

и заканчивается плиткой  $\mathcal{P}_0$ . Такую цепь назовем циклом (см. рис. 12). Можно считать, что в слове  $\omega$  не встречаются фрагменты  $Aa$  или  $aA$ . Это равносильно тому, что в соответствующей цепочке плиток отсутствуют фрагменты вида  $\mathcal{P}\mathcal{P}'\mathcal{P}$ , которые естественно назвать возвращениями.

Каждое соотношение приводит к циклу плиток, и, наоборот, с каждым циклом связано соотношение. В замощении имеются естественные циклы, образованные обходами вокруг вершины замощения. Все вершины замощения эквивалентны при действии группы  $\Gamma$  (почему?). Следовательно, все соотношения обхода равносильны.

Легко проверить, используя правило разметки плиток из п. 4.1, что соотношение обхода вершины плитки  $\mathcal{P}_0$ , отмеченной черной точкой на рис. 10, выглядит так:

$$abcdABCD = 1.$$

Покажем, что соотношение обхода является определяющим. Идея доказательства проста и восходит, по-видимому, к Г. М. Кокстеру: на рисунке 12 изображена петля, проходящая через цикл плиток без возвращений, соответствующий некоторому соотношению в группе  $\Gamma$ . Вершины замощения, попавшие внутрь петли, Кокстер называл «гвоздями». Гвозди мешают «стянуть» наш цикл в цикл  $\mathcal{P}_0\mathcal{P}_0$ , который отвечает пустому слову.

Рисунок 12 учит нас, что соотношение обхода вокруг «гвоздя»  $O$  позволяет снять цикл  $(\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_4, \mathcal{P}_5, \mathcal{P}_6, \mathcal{P}_7)$  с гвоздя  $O$ , заменив его циклом  $(\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3, \mathcal{P}_5, \mathcal{P}_6, \mathcal{P}_7)$ .

Последовательное применение операции «снятия с гвоздя» с помощью соотношений обхода приводит нас к цели: слово, отвечающее начальному циклу, приводится к пустому слову с помощью соотношений обхода.

На этом доказательство теоремы об образующих и определяющих соотношениях группы  $\Gamma$  заканчивается.

## 5. ОТМЕЧЕННЫЕ ФУКСОВЫ ГРУППЫ РОДА 2

5.1. В предыдущем разделе мы построили такую фуксову группу  $\Gamma$ , что

- а) группа  $\Gamma$  действует на плоскости Лобачевского  $\Lambda^2$  без неподвижных точек;
- б) факторпространство  $\Lambda^2/\Gamma$  гомеоморфно поверхности рода 2 (аналогичное построение имеет место и для любого рода  $g > 1$ ).

Назовем такую группу фуксовой группой рода 2. В нашем построении фуксова группа рода 2 появилась не одна, а вместе с фундаментальным правильным гиперболическим восьмиугольником  $\mathcal{P}$  с суммой внутренних углов, равной  $2\pi$ , стороны которого отмечены по образцу  $w = aBcDAbCd$  рисунка 10.

Такого сорта отмеченные многоугольники (обязательно несамопересекающиеся, но уже не обязательно правильные и не обязательно выпуклые!) играют заметную роль в жизни фуксовых групп, а потому мы дадим им имя.

**ОПРЕДЕЛЕНИЕ.** Гиперболический восьмиугольник без самопересечений с суммой внутренних углов, равной  $2\pi$ , стороны которого отмечены по образцу  $w = aBcDAbCd$ , назовем *каноническим фундаментальным восьмиугольником* (или канонической геометрической выкройкой).

Напомним также, что одинаковой буквой (прописной и строчной) отмечаются равные стороны. Более того, следуя одной идее Алисы<sup>2)</sup>  $B$  одновременно обозначает и то единственное сохраняющее ориентацию движение плоскости, которое сторону  $b$  переводит в сторону  $B$ .

5.2. А. Пуанкаре доказал, что любой канонический восьмиугольник ( $4g$ -угольник) служит фундаментальной областью фуксовой группы  $\Gamma$  рода 2 (рода  $g > 1$ ). Группа  $\Gamma$  порождается элементами  $a, b, c, d$  с единственным определяющим соотношением  $abcdABCD = 1$  (или  $abcd = dcba$ ).

Верно и обратное утверждение: для любой фуксовой группы рода 2 существует фундаментальный канонический восьмиугольник. Попробуем это объяснить.

5.3. Пусть  $\Gamma$  — фуксова группа рода 2. Зафиксируем точку  $y_0 \in \Lambda^2$  и рассмотрим множество  $\mathcal{F}$  таких точек  $y$ , что  $\rho(y, y_0) \leq \rho(\gamma y, y_0)$  для любого  $\gamma \in \Gamma$  (здесь  $\rho$  — расстояние на плоскости Лобачевского).

<sup>2)</sup> «собака по имени Собака» Просто и удобно.

Несложно показать (но мы этого делать не будем), что множество  $\mathcal{F}$  является выпуклым (а потому простым) гиперболическим фундаментальным многоугольником для группы  $\Gamma$  (так называемый фундаментальный многоугольник Вороного – Дирихле). Таким образом мы обзаводимся геометрической выкройкой поверхности рода 2 в виде несамопересекающегося многоугольника. Так как группа  $\Gamma$  действует на плоскости  $\Lambda^2$  без неподвижных точек, то выкройка Вороного – Дирихле  $\mathcal{F}$  не содержит двух смежных сторон, отмеченных одной и той же буквой (почему?). Как объяснялось в разделе 2, выкройку  $\mathcal{F}$  с помощью элементарных топологических преобразований можно сделать сначала удобной, а затем — канонической. Заменяя на каждом шаге этой цепочки гомеоморфизм склейки сторон на движение из группы  $\Gamma$ , отождествляющее эти стороны, мы получим в конце цепочки канонический восьмиугольник. В самом деле, каждое геометрическое элементарное преобразование приводит к новому фундаментальному многоугольнику для группы  $\Gamma$ . Отсюда следует, что после каждого шага у нас должен получаться многоугольник  $\mathcal{P}$  без самопересечений.

5.4. Назовем пару  $(\Gamma, \mathcal{P})$ , состоящую из фуксовой группы рода 2 и ее фундаментального канонического восьмиугольника  $\mathcal{P}$ , *геометрически отмеченной фуксовой группой рода 2*. В множестве геометрически отмеченных фуксовых групп рода 2 введем отношение эквивалентности, объявив пары  $(\Gamma, \mathcal{P})$  и  $(\Gamma_1, \mathcal{P}_1)$  эквивалентными, если существует такое сохраняющее ориентацию движение плоскости  $t$ , что

а)  $\Gamma_1 = t\Gamma t^{-1}$ ;

б)  $t$  переводит отмеченный восьмиугольник  $\mathcal{P}$  в отмеченный восьмиугольник  $\mathcal{P}_1$ .

Фактормножество или множество классов эквивалентных пар обозначим через  $Q(2)$ . С другой стороны, обозначим через  $\mathcal{P}(2)$  множество канонических восьмиугольников, рассматриваемых с точностью до их равенства как отмеченных.

Тонкая теорема Пуанкаре и рассмотрения п. 5.3 говорят нам о том, что множества  $Q(2)$  и  $\mathcal{P}(2)$  находятся во взаимно однозначном соответствии. Множество  $\mathcal{P}(2)$  можно превратить в топологическое пространство с естественной топологией геометрической сходимости восьмиугольников.

Как устроено пространство  $Q(2)$  в окрестности любой своей точки  $(\Gamma_0, \mathcal{P}_0)$ ? Очевидно, что оно состоит из классов фуксовых групп, канонические многоугольники которых можно получить «малым шевелением» вершин многоугольника  $\mathcal{P}_0$ . Важно заметить, что

а) многоугольники, полученные из  $\mathcal{P}_0$  «малым шевелением» его вершин, остаются несамопересекающимися;

- б) канонический многоугольник  $\mathcal{P}$  однозначно восстанавливается по любой своей вершине  $x$  с помощью движений из группы  $\Gamma$  так, как это изображено на рис. 13.

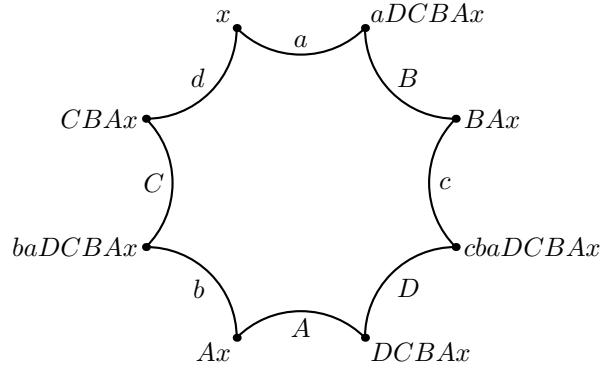


Рис. 13.

Следующий эвристический метод подсчета параметров может быть сделан строгим и дает правильное представление о запасе отмеченных фуксовых групп рода 2.

Каждая вершина многоугольника  $\mathcal{P}$  зависит от двух параметров (положение точки на гиперболической плоскости). Получается 16 параметров для «шевеления». Но между ними имеется пять зависимостей: ведь нужно позаботиться о том, чтобы полученный многоугольник был каноническим, т. е. обеспечить равенство противоположных сторон и добиться того, чтобы сумма внутренних углов была равна  $2\pi$ . Итого  $4 + 1 = 5!$  Если еще учесть группу движений плоскости Лобачевского, которая зависит от 3 параметров, то у нас остается восемь ( $16 - 5 - 3 = 8$ ) независимых параметров.

А как теперь узнать, от скольких параметров зависит пространство «чистых» (не отмеченных каноническим многоугольником  $\mathcal{P}$ ) фуксовых групп рода 2? Для этого нужно учесть следующее обстоятельство: давайте пошевелим вершину  $x$  многоугольника  $\mathcal{P}_0$ , а затем построим с помощью новой точки  $x'$  и старых движений ( $a_0, b_0, c_0, d_0, A_0, B_0, C_0, D_0$ ) из группы  $\Gamma_0$  новый канонический многоугольник  $\mathcal{P}'$  ровно так, как это сделано на рис. 13. Ясно, что при этом  $\mathcal{P}'$  окажется каноническим многоугольником той же фуксовой группы  $\Gamma_0$ . Поэтому при таких «шевелениях» мы не получим новой точки в пространстве фуксовых групп. Отсюда следует, что пространство фуксовых групп рода 2 зависит от  $8 - 2 = 6$  вещественных параметров (для рода  $g$  — от  $(6g - 6)$  параметров).

Строгое и изящное геометрическое изложение этого круга вопросов заинтересованный читатель сможет найти в статье С. М. Натансона [6].



## СПИСОК ЛИТЕРАТУРЫ

- [1] В. Г. Болтянский, В. А. Ефремович. *Наглядная топология* (под редакцией С. П. Новикова). М.: Наука, 1982.
- [2] Ю. Г. Борисович, Н. М. Близняков, Я. А. Израилевич, Т. Н. Фоменко. *Введение в топологию*. Изд. 2е. М.: Наука. Физматлит, 1995.
- [3] Э. Б. Винберг. *Калейдоскопы и группы отражений* // Математическое просвещение. Сер. 3, вып. 7, 2003. С. 45–63.
- [4] Б. А. Дубровин, С. П. Новиков, А. Т. Фоменко. *Современная геометрия. Методы и приложения* (гл. 4, §20). М.: Наука, 1979.
- [5] А. А. Клячко *Комбинаторная теория групп и геометрия* // Математическое просвещение. Сер. 3, вып. 13, 2009. С. 18–32.
- [6] С. М. Натанзон. *Инвариантные прямые фуксовых групп* // УМН, 1972. Т. 27, №4. С. 145–160.
- [7] Г. Зейферт, В. Трельфалль. *Топология*. Изд 2е. М.; Ижевск: РХД, 2001.
- [8] Ч. Коснёвски. *Начальный курс алгебраической топологии* (глава 11). М.: Мир, 1983.
- [9] С. Г. Смирнов. *Прогулки по замкнутым поверхностям*. М.: МЦНМО, 2003.
- [10] H.S.M. Coxeter. *Discrete groups generated by reflections* // Ann. Math., 1934. Vol. 35, no 3. P. 588–621.

---

---

# Наш семинар: математические сюжеты

---

---

## Конечные проективные плоскости

Ю. И. Ионин

### §1. ОПРЕДЕЛЕНИЕ ПРОЕКТИВНОЙ ПЛОСКОСТИ

Истоки проективной геометрии можно найти в попытках художников Возрождения создать правдоподобные двумерные образы трехмерных объектов. Зритель видит точку на картине посредством луча, направленного из этой точки в глаз зрителя. Если глаз зрителя находится в начале координат, то каждая точка в плоскости картины определяется прямой, проходящей через эту точку и начало координат, а каждая прямая в плоскости картины определяется плоскостью, проходящей через эту прямую и начало координат.

Прямые и плоскости, проходящие через начало координат, естественным образом отождествляются с одномерными и двумерными подпространствами трехмерного векторного пространства.

**ОПРЕДЕЛЕНИЕ 1.1.** Пусть  $V$  — трехмерное векторное пространство над полем  $F$ . Точками *проективной плоскости над полем  $F$*  являются все одномерные подпространства пространства  $V$ , а прямыми — все двумерные подпространства. Проективная плоскость над  $F$  обозначается  $PG(2, F)$ . Если  $F$  — конечное поле порядка  $q$ , то проективная плоскость над  $F$  обозначается  $PG(2, q)$ .

Для того чтобы сделать это определение более наглядным, рассмотрим плоскость  $z = 1$  в трехмерном координатном пространстве над полем  $F$ . Каждая прямая, проходящая через начало координат и не лежащая в плоскости  $xy$ , однозначно определяется точкой пересечения  $(a, b, 1)$  с плоскостью  $z = 1$ , и мы обозначим  $(a, b)$  соответствующую этой прямой точку

проективной плоскости  $PG(2, F)$ . Каждая прямая, проходящая через начало координат и лежащая в плоскости  $xy$ , однозначно определяется ее угловым коэффициентом  $a$ , и мы обозначим  $(a)$  соответствующую точку в  $PG(2, F)$ . Для углового коэффициента прямой  $x = 0$  в плоскости  $xy$  мы используем символ  $\infty$ . Таким образом, множество  $P$  всех точек проективной плоскости  $PG(2, F)$  может быть представлено как

$$P = \{(a, b): a, b \in F\} \cup \{(a): a \in F\} \cup \{(\infty)\}. \quad (1)$$

Каждая плоскость, проходящая через начало координат, кроме плоскости  $xy$ , однозначно определяется прямой, по которой она пересекает плоскость  $z = 1$ . Если эта прямая задается в плоскости  $z = 1$  уравнением  $y = xa + b$ , мы обозначаем  $[a, b]$  соответствующую прямую в  $PG(2, F)$ . Если же эта прямая задается уравнением  $x = a$ , то соответствующая прямая в  $PG(2, F)$  обозначается  $[a]$ . Наконец, прямая в  $PG(2, F)$ , соответствующая плоскости  $xy$ , обозначается  $[\infty]$ .

Таким образом, множество  $\mathcal{L}$  всех прямых проективной плоскости  $PG(2, F)$  может быть представлено как

$$\mathcal{L} = \{[a, b]: a, b \in F\} \cup \{[a]: a \in F\} \cup \{[\infty]\}. \quad (2)$$

Прямые в  $PG(2, F)$  можно представить как множества точек следующим образом:

$$[a, b] = \{(x, xa + b): x \in F\} \cup \{a\}, \quad (3)$$

$$[a] = \{(a, y): y \in F\} \cup \{(\infty)\}, \quad (4)$$

$$[\infty] = \{(a): a \in F\} \cup \{(\infty)\}. \quad (5)$$

Подобное представление проективной плоскости называется ее *координатизацией*.

Точки  $(a)$ ,  $a \in F$ , и  $(\infty)$  называют *идеальными точками* проективной плоскости, так что на каждой прямой, кроме  $[\infty]$ , лежит одна идеальная точка. Прямую  $[\infty]$  называют *идеальной прямой*. «Неидеальные» точки и прямые образуют  $AG(2, F)$  — *аффинную плоскость над полем  $F$* .

Через любые две точки проективной плоскости проходит одна и только одна прямая, и любые две прямые пересекаются в единственной точке. Мы примем эти свойства за основу аксиоматического определения проективной плоскости.

**ОПРЕДЕЛЕНИЕ 1.2.** *Проективная плоскость  $\Pi = (P, \mathcal{L})$  состоит из множества  $P$ , элементы которого называются точками, из множества  $\mathcal{L}$ , элементы которого называются прямыми, и из бинарного отношения  $\in$  между точками и прямыми. Если  $x \in L$ , где  $x$  — точка,  $L$  — прямая, мы говорим « $x$  лежит на  $L$ », « $L$  проходит через  $x$ », « $x$  и  $L$  инцидентны» и т. п.*

Точки и прямые проективной плоскости должны удовлетворять следующим аксиомам.

P1. Через любые две точки проходит одна и только одна прямая.

P2. Любые две прямые имеют одну и только одну общую точку.

P3. Существуют четыре точки, никакие три из которых не лежат на одной прямой.

Если  $x, y \in P$ ,  $x \neq y$ , то  $xy$  — это прямая, проходящая через  $x$  и  $y$ ; если  $L, M \in \mathcal{L}$ ,  $L \neq M$ , то  $LM$  — это точка пересечения  $L$  и  $M$ .

Плоскость Фано. На рис. 1 изображена проективная плоскость, в которой точки — это вершины, середины сторон и центр правильного треугольника, а прямые — стороны, медианы и вписанная окружность. Эта проективная плоскость называется *плоскостью Фано*. Если плоскость Фано координатизировать, как показано на рисунке, она превращается в  $PG(2, 2)$ .

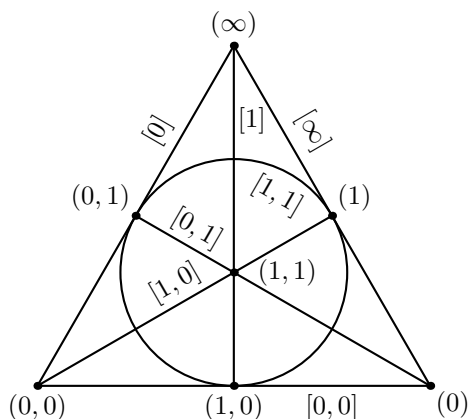


Рис. 1.

Как показывает следующее упражнение, аксиома P3 нужна для того чтобы избежать нескольких вырожденных примеров.

УПРАЖНЕНИЕ. Предположим, что система  $\Pi = (P, \mathcal{L})$  точек и прямых удовлетворяет аксиомам P1 и P2, но не удовлетворяет аксиоме P3. Докажите, что имеет место один (или более) из следующих случаев: (а)  $P = \emptyset$ ; (б) есть только одна прямая (и все точки лежат на этой прямой); (в) существуют  $x \in P$  и  $L \in \mathcal{L}$  такие, что все точки, кроме  $x$ , лежат на  $L$  и все прямые, кроме  $L$ , проходят через  $x$ .

Аксиома P2 может быть получена из P1 заменой точек прямыми, а прямых — точками. Если  $a, b, c, d$  — четыре точки, никакие три из которых не лежат на одной прямой (аксиома P3), то  $ab, bc, cd, da$  — четыре

прямые, никакие три из которых не пересекаются в одной точке. Поэтому для проективной плоскости справедлив следующий *принцип двойственности*: если в верном утверждении о проективных плоскостях заменить точки прямыми, а прямые — точками, то снова получится верное утверждение о проективных плоскостях.

Проективная плоскость называется *конечной*, если множество ее точек (и, следовательно, прямых) конечно. В проективной плоскости над конечным полем порядка  $q$  каждая прямая проходит через  $q + 1$  точек и каждая точка лежит на  $q + 1$  прямых. Аналогичное утверждение справедливо для любой конечной проективной плоскости.

**ТЕОРЕМА 1.3.** Для любой конечной проективной плоскости  $\Pi = (P, \mathcal{L})$  существует натуральное число  $n \geq 2$  — *порядок*  $\Pi$ , такое что

- (а) каждая прямая проходит через  $n + 1$  точек,
- (б) каждая точка лежит на  $n + 1$  прямых,
- (в)  $|P| = |\mathcal{L}| = n^2 + n + 1$ .

**ДОКАЗАТЕЛЬСТВО.** Пусть  $L, M \in \mathcal{L}$ ,  $L \neq M$ , и пусть  $a$  — точка, не лежащая ни на  $L$ , ни на  $M$  (такая точка существует в силу аксиомы Р3). Зададим отображение  $f$  множества точек на прямой  $L$  в множество точек на прямой  $M$  следующим образом:

если  $x \in L$ , то  $f(x)$  — точка пересечения прямых  $xa$  и  $M$ .

Проверьте, что  $f$  — биекция, откуда следует, что число точек на любой прямой одно и то же. Представим это число как  $n + 1$ . В силу принципа двойственности через любую точку проективной плоскости проходят  $n + 1$  прямых. Зафиксируем точку  $a$ . Каждая прямая, проходящая через  $a$ , проходит через  $n$  других точек, и все точки плоскости лежат на этих прямых, так что  $|P| = (n + 1)n + 1 = n^2 + n + 1$ . В силу принципа двойственности  $|\mathcal{L}| = n^2 + n + 1$ . Если  $n = 0$  или  $1$ , то  $|P| < 4$ , что противоречит аксиоме Р3, так что  $n \geq 2$ .

С понятием проективной плоскости тесно связано понятие аффинной плоскости.

**ОПРЕДЕЛЕНИЕ 1.4.** *Аффинная плоскость* состоит из множества точек и из множества прямых, удовлетворяющих следующим аксиомам.

- A1. Через любые две точки проходит одна и только одна прямая.
- A2. Если точка  $x$  не лежит на прямой  $L$ , то существует единственная прямая  $M$ , проходящая через  $x$  и не имеющая с  $L$  общих точек.
- A3. Существуют три точки, не лежащие на одной прямой.

В аффинной плоскости две прямые *параллельны*, если они не имеют общих точек или совпадают. Присоединив к каждой прямой в аффинной

плоскости идеальную точку таким образом, что две прямые получают одну и ту же идеальную точку в том и только в том случае, если они параллельны, и присоединив к плоскости идеальную прямую, образованную всеми идеальными точками, мы получим проективную плоскость. Обратное, удалив из проективной плоскости одну прямую и все лежащие на ней точки, мы получим аффинную плоскость.

## §2. ДРУГИЕ ПРИМЕРЫ ПРОЕКТИВНЫХ ПЛОСКОСТЕЙ

Соотношения (1)–(5) предыдущего параграфа описывают проективную плоскость над полем  $F$ . Оказывается, те же соотношения описывают проективную плоскость, если поле  $F$  заменено почтиполем.

**ОПРЕДЕЛЕНИЕ 2.1.** *Почтиполе* — это множество  $F$  с двумя бинарными операциями — сложением и умножением, удовлетворяющими следующим условиям:

- (а)  $F$  — абелева группа относительно сложения, нейтральный элемент которой обозначается  $0$ ;
- (б)  $F^* = F \setminus \{0\}$  — группа относительно умножения;
- (в)  $(a + b)c = ac + bc$  для любых  $a, b, c \in F$ .

Таким образом, почтиполе, в отличие от поля, не требует коммутативности умножения и одного из дистрибутивных законов.

**ПРИМЕР.** Пусть  $q$  — нечетное простое число или степень нечетного простого числа и пусть  $F$  — поле порядка  $q^2$ . Чтобы получить почтиполе, мы сохраним все элементы поля  $F$ , оставим неизменной операцию сложения и введем новую операцию умножения, обозначаемую  $\circ$ :

$$a \circ b = \begin{cases} ab, & \text{если } b \text{ является квадратом элемента поля } F, \\ a^q b, & \text{если } b \text{ не является квадратом элемента поля } F. \end{cases}$$

**УПРАЖНЕНИЕ.** Проверьте, что  $(F, +, \circ)$  — почтиполе. При проверке ассоциативности умножения воспользуйтесь тем, что в конечном поле произведение двух неквадратов — квадрат. Дистрибутивный закон следует из тождества  $(a + b)^q = a^q + b^q$ .

Наименьшее почтиполе (не являющееся полем) состоит из девяти элементов. Его мультипликативную группу можно представить как  $\{\pm 1, \pm i, \pm j, \pm k\}$  с кватернионным умножением:  $i^2 = j^2 = k^2 = -1$ ,  $ij = -ji = k$ ,  $jk = -kj = i$ ,  $ki = -ik = j$ .

Пусть  $F$  — произвольное конечное почтиполе. Определим множество точек  $P$  и множество прямых  $\mathcal{L}$  соотношениями (1)–(5). Мы утверждаем, что  $\Pi = (P, \mathcal{L})$  — проективная плоскость.

Аксиома P3 очевидна, а проверка аксиомы P1 проводится точно так же, как и в случае поля. Однако, если мы попытаемся проверить аксиому P2 «в лоб», мы обнаружим, что требуется отсутствующий дистрибутивный закон. Мы «пойдем в обход» и сначала заметим, что из соотношений (3)–(5) следует, что каждая прямая состоит из  $n + 1$  точек, где  $n$  — число элементов почтиполя  $F$ , что через каждую точку проходят  $n + 1$  прямых и что общее число прямых равно  $n^2 + n + 1$ . Теперь мы зафиксируем прямую  $L$  и подсчитаем число пар  $(x, M)$ , где  $M$  — прямая, отличная от  $L$ , а  $x$  — общая точка  $L$  и  $M$ . Если мы сначала выберем точку  $x \in L$ , а затем прямую  $M \neq L$ , проходящую через  $x$ , мы получим, что число таких пар равно  $(n + 1)n$ . Поскольку имеется  $n^2 + n$  прямых  $M \neq L$  и ни одна из них (в силу аксиомы P1) не пересекает  $L$  более чем в одной точке, мы получаем, что каждая такая прямая пересекает  $L$  ровно в одной точке.

Приведенное рассуждение существенно использует конечность почтиполя  $F$ . В бесконечном случае для получения проективной плоскости на почтиполе налагается дополнительное требование *планарности*: для любых  $a, b, c \in F$ , уравнение  $xa + xb = c$  имеет единственное решение. Всякое конечное почтиполе планарно.

Если  $F$  — почтиполе, то множество  $K = \{x \in F : x(a + b) = xa + xb \text{ для всех } a, b \in F\}$  называется *ядром*  $F$  и обозначается  $\ker(F)$ .

УПРАЖНЕНИЕ. Проверьте, что  $K = \ker(F)$  — поле,  $F$  — векторное пространство над  $K$ .

Если  $F$  — конечное поле, то  $F$  является конечномерным векторным пространством над конечным полем  $K$ . Отсюда следует, что число элементов любого конечного почтиполя — степень простого числа.

Мы приведем еще один пример построения проективной плоскости.

Пусть  $F$  — поле,  $V$  — векторное пространство четной размерности  $2d$  над  $F$ . Будем считать элементы множества  $V$  точками проективной плоскости. Прямой назовем любое множество точек вида  $U + x$ , где  $x \in V$ , а  $U$  —  $d$ -мерное подпространство пространства  $V$ . Прямые вида  $U + x$  и  $U + y$  с одним и тем же подпространством  $U$  назовем параллельными. Множество всех прямых таким образом разбивается на классы параллельных прямых. Присоединим к каждой прямой идеальную точку так, что две прямых получают одну и ту же идеальную точку в том и только в том случае, если они параллельны. Кроме того, образуем идеальную прямую, состоящую из всех идеальных точек.

УПРАЖНЕНИЕ. Проверьте, что множество точек и множество прямых, описанных в предыдущем абзаце, образуют проективную плоскость.

## §3. КООРДИНАТИЗАЦИЯ

Начав с поля или почтиполя  $F$ , мы построили проективную плоскость, точки и прямые которой удовлетворяют соотношениям (1)–(5). Можно ли получить проективную плоскость, исходя из менее жесткой алгебраической структуры? В этом параграфе мы начнем с проективной плоскости, введем на ней координаты, базирующиеся на бесструктурном множестве  $F$  и затем используем свойства проективной плоскости для определения операций сложения и умножения на  $F$ . Хотя нижеследующие построения могут быть проведены для любой проективной плоскости, мы ограничимся конечным случаем.

Пусть  $\Pi = (P, \mathcal{L})$  — проективная плоскость порядка  $n \geq 2$  и пусть  $F$  — множество из  $n$  элементов. Выберем два элемента в множестве  $F$  и обозначим их  $0$  и  $1$ . Пусть  $\infty$  — символ, не являющийся элементом  $F$ . Выберем в плоскости  $\Pi$  *начальный четырехугольник*, т. е. упорядоченную четверку точек, никакие три из которых не лежат на одной прямой. Обозначим эти точки последовательно  $(0, 0)$ ,  $(0)$ ,  $(\infty)$  и  $(1, 1)$ . Обозначим шесть прямых, проходящих через пары обозначенных точек следующим образом:  $(0, 0)(0) = [0, 0]$ ,  $(0, 0)(\infty) = [0]$ ,  $(0)(\infty) = [\infty]$ ,  $(1, 1)(\infty) = [1]$ ,  $(1, 1)(0) = [0, 1]$ ,  $(0, 0)(1, 1) = [1, 0]$ . (Напомним, что если  $x$  и  $y$  — различные точки, то  $xy$  — проходящая через них прямая, а если  $L$  и  $M$  — различные прямые, то  $LM$  — их общая точка.) Если бы  $F$  было полем или почтиполем, то указанные прямые представляли бы соответственно ось абсцисс, ось ординат, идеальную прямую и прямые  $x = 1$ ,  $y = 1$  и  $y = x$ . Мы теперь можем координатизировать еще три точки:  $(1, 0) = [1][0, 0]$ ,  $(0, 1) = [0][0, 1]$ ,  $(1) = [\infty][1, 0]$  (рис. 2).

На идеальной прямой  $[\infty]$  остается  $n - 2$  точек, кроме  $(0)$ ,  $(\infty)$  и  $(1)$ , а в множестве  $F$  есть  $n - 2$  элементов, кроме  $0$  и  $1$ . Обозначим оставшиеся точки идеальной прямой символами  $(a)$ , где  $a \in F$ ,  $a \neq 0$ ,  $a \neq 1$ , так чтобы разные точки обозначались разными символами.

Мы готовы завершить координатизацию плоскости  $\Pi$ . Для каждого  $a \in F$  положим  $(0, 0)(a) = [a, 0]$ ,  $[1][a, 0] = (1, a)$ ,  $(0)(1, a) = [0, a]$ ,  $[0, a][1, 0] = (a, a)$ ,  $(a, a)(\infty) = [a]$ ,  $[a][0, 0] = (a, 0)$ ,  $[0, a][0] = (0, a)$  (рис. 2). Наконец, для любых  $a, b \in F$  положим

$$(a, b) = [a][0, b], \quad [a, b] = (a)(0, b).$$

Если  $F$  — поле, то  $(a, a+b)$  — это точка пересечения прямых  $[a]$  и  $[1, b]$ , а  $(a, ab)$  — это точка пересечения прямых  $[a]$  и  $[b, 0]$ . Поэтому мы определяем сложение и умножение на произвольном координатирующем множестве  $F$  так, чтобы следующие соотношения выполнялись для любых  $a, b \in F$ :

$$(a, a + b) = [a][1, b], \quad (a, ab) = [a][b, 0]. \quad (6)$$



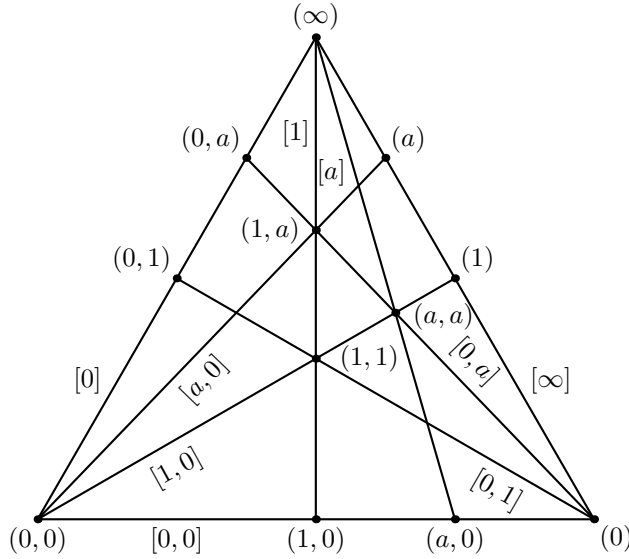


Рис. 2.

Итак, у нас есть множество  $F$  с операциями сложения и умножения и при этом точки и прямые плоскости  $\Pi$  удовлетворяют условиям (1), (2), (4) и (5). Мы будем называть множество  $F$  *координатным кольцом* проективной плоскости  $\Pi$ . Предостережение: координатное кольцо не обязано быть кольцом в обычном алгебраическом смысле этого термина. При этом примеры «плохих» координатных колец совсем не очевидны (мы их получим в последующих параграфах). Отметим также, что алгебраические свойства координатного кольца могут зависеть от выбора начального четырехугольника.

**ТЕОРЕМА 3.1.** Введенные операции сложения и умножения в координатном кольце  $F$  обладают следующими свойствами:

- (а)  $a + 0 = 0 + a = a$ ,  $a \cdot 1 = 1 \cdot a = a$ ,  $a \cdot 0 = 0 \cdot a = 0$  для любого  $a \in F$ ;
- (б) для любых  $a, b \in F$  каждое из уравнений  $a + x = b$ ,  $x + a = b$  имеет единственное решение;
- (с) для любого  $a \in F^* = F \setminus \{0\}$  и для любого  $b \in F$  каждое из уравнений  $ax = b$ ,  $xa = b$  имеет единственное решение.

Мы докажем по одному свойству из каждой группы и предоставим читателю доказательство остальных свойств.

- (а)  $(1, 1 \cdot a) = [1][a, 0] = (1, a)$  (рис. 2). Поэтому  $1 \cdot a = a$ .
- (б)  $a + x = b \Leftrightarrow (a, a + x) = (a, b) \Leftrightarrow [a][1, x] = (a, b) \Leftrightarrow (a, b) \in [1, x]$ .

Последнее условие означает, что  $[1, x]$  — единственная прямая, проходящая через точки  $(1)$  и  $(a, b)$ , что однозначно определяет  $x$ .

(с)  $xa = b \Leftrightarrow (x, xa) = (x, b) \Leftrightarrow [x][a, 0] = (x, b) \Leftrightarrow (x, b) \in [a, 0]$ .  
Последнее условие означает, что  $(x, b)$  — точка пересечения различных ( $a \neq 0$ ) прямых  $[a, 0]$  и  $[0, b]$ , так что  $x$  определен однозначно.

Нам предстоит долгий и не всегда успешный путь от этих простых свойств до аксиом поля или почтиполя, и центральную роль на этом пути будет играть знаменитая теорема проективной геометрии.

#### §4. ТЕОРЕМА ДЕЗАРГА

Жерар Дезарг (1591–1661) — французский инженер, архитектор и математик — считается одним из создателей проективной геометрии. Для формулировки теоремы Дезарга нам понадобится следующее определение.

**ОПРЕДЕЛЕНИЕ 4.1.** Треугольник в проективной плоскости — это упорядоченная тройка точек, не лежащих на одной прямой.

Треугольники  $x_1x_2x_3$ ,  $y_1y_2y_3$ , такие что  $x_1 \neq y_1$ ,  $x_2 \neq y_2$ ,  $x_3 \neq y_3$ , называются *центрально перспективными*, если прямые  $x_1y_1$ ,  $x_2y_2$ ,  $x_3y_3$  попарно различны и проходят через общую точку.

Треугольники  $x_1x_2x_3$ ,  $y_1y_2y_3$ , такие что  $x_1x_2 \neq y_1y_2$ ,  $x_2x_3 \neq y_2y_3$ ,  $x_3x_1 \neq y_3y_1$ , называются *аксиально перспективными*, если точки пересечения прямых  $x_1x_2$  и  $y_1y_2$ ,  $x_2x_3$  и  $y_2y_3$ ,  $x_3x_1$  и  $y_3y_1$  попарно различны и лежат на одной прямой.

Треугольники  $x_1x_2x_3$ ,  $y_1y_2y_3$  на рис. 3 центрально и аксиально перспективны.

**ТЕОРЕМА ДЕЗАРГА.** Два треугольника в проективной плоскости (над полем вещественных чисел) центрально перспективны в том и только в том случае, если они аксиально перспективны.

Заметим, что определения центрально перспективных и аксиально перспективных треугольников взаимно двойственны. (Строго говоря, в случае аксиальной перспективы следует заменить треугольники трехсторонниками, образованными прямыми, проходящими через пары вершин треугольника.) Поэтому для получения теоремы Дезарга во всей полноте достаточно доказать, что центрально перспективные треугольники аксиально перспективны.

Предположим, что треугольники  $x_1x_2x_3$ ,  $y_1y_2y_3$  на рис. 3 центрально перспективны и обозначим через  $z_1$ ,  $z_2$ ,  $z_3$  точки пересечения прямых  $x_2x_3$  и  $y_2y_3$ ,  $x_3x_1$  и  $y_3y_1$ ,  $x_1x_2$  и  $y_1y_2$  соответственно. Представим себе, что рис. 3 является двумерной проекцией трехмерного объекта, т. е. что прямые  $x_1y_1$ ,  $x_2y_2$ ,  $x_3y_3$  на самом деле не лежат в одной плоскости. Тогда

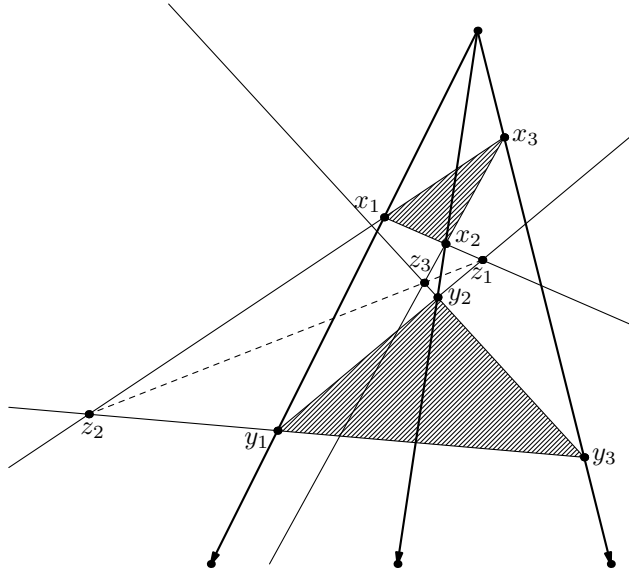


Рис. 3.

точки  $z_1, z_2, z_3$  лежат как в плоскости  $x_1x_2x_3$ , так и в плоскости  $y_1y_2y_3$ , и потому лежат на прямой, являющейся пересечением этих плоскостей. Поскольку проектирование на плоскость под разумно выбранным углом не нарушает ни инцидентности, ни прямолинейности, из трехмерной теоремы Дезарга следует двумерная.

Приведенное рассуждение является наброском доказательства, которое может быть формализовано для проективной плоскости над произвольным полем. Мы, однако, приведем другое доказательство, которое послужит нам и в дальнейшем.

**ТЕОРЕМА 4.2.** Теорема Дезарга справедлива для проективной плоскости  $PG(2, F)$ , где  $F$  — поле.

**ДОКАЗАТЕЛЬСТВО.** Мы будем опираться на определение 1.1, так что отмеченные точки на рис. 3 представляют одномерные подпространства трехмерного векторного пространства  $V$  над полем  $F$ , а прямые — двумерные подпространства. Пусть прямые  $x_1y_1, x_2y_2, x_3y_3$  пересекаются в точке  $c$  (т. е. пересечение соответствующих двумерных подпространств одномерно). Мы должны показать, что точки  $z_1, z_2, z_3$  лежат на одной прямой. Если точка  $c$  лежит на каждой из прямых  $z_1z_2, z_2z_3, z_3z_1$ , то доказывать нечего, так что мы предположим, что  $c \notin z_2z_3$ .

В каждом одномерном подпространстве, отмеченном на рис. 3, выберем по ненулевому вектору и обозначим этот вектор той же буквой,

но жирным шрифтом. Так как векторы  $\mathbf{x}_1$ ,  $\mathbf{y}_1$  и  $\mathbf{c}$  принадлежат одному двумерному подпространству, они линейно зависимы, и мы можем выбрать эти векторы в их одномерных подпространствах так, что  $\mathbf{y}_1 = \mathbf{x}_1 + \mathbf{c}$ . Векторы  $\mathbf{x}_1, \mathbf{z}_2, \mathbf{z}_3$  линейно независимы (в противном случае точки  $x_1, z_2, z_3$  лежали бы на одной прямой, что легко приводится к противоречию). Поэтому мы можем положить  $\mathbf{y}_1 = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{z}_2 + \alpha_3 \mathbf{z}_3$ , где  $\alpha_1, \alpha_2, \alpha_3 \in F$ . Рассмотрим невырожденный линейный оператор  $\varphi$  на  $V$ , задаваемый условиями  $\varphi(\mathbf{x}_1) = \mathbf{y}_1$ ,  $\varphi(\mathbf{z}_2) = \mathbf{z}_2$ ,  $\varphi(\mathbf{z}_3) = \mathbf{z}_3$ . Тогда  $\varphi(\mathbf{c}) = \varphi(\mathbf{y}_1 - \mathbf{x}_1) = (\alpha_1 - 1)\mathbf{y}_1 + \alpha_2 \mathbf{z}_2 + \alpha_3 \mathbf{z}_3 = (\alpha_1 - 1)\mathbf{y}_1 + \mathbf{y}_1 - \alpha_1 \mathbf{x}_1 = \alpha_1 \mathbf{c}$ .

Так как оператор  $\varphi$  линеен и невырожден, он отображает одномерные подпространства в одномерные и двумерные в двумерные. В частности,  $\varphi(x_1) = y_1$ ,  $\varphi(z_2) = z_2$ ,  $\varphi(z_3) = z_3$  и  $\varphi(c) = c$ . Векторы  $\mathbf{z}_2, \mathbf{z}_3$  образуют базис прямой  $z_2 z_3$  (как двумерного подпространства), так что оператор  $\varphi$  тождественен на этой прямой. Если  $L$  — любая прямая, проходящая через  $c$ , то, так как  $\varphi$  оставляет на месте  $c$  и точку пересечения  $L$  с прямой  $z_2 z_3$ , мы получаем, что  $\varphi(L) = L$  (что вовсе не означает, что  $\varphi$  тождественен на  $L$ ).

Так как точка  $x_2$  лежит на прямых  $cx_2$  и  $x_1 z_3$ ,  $\varphi(x_2)$  — это точка пересечения прямых  $cx_2$  и  $y_1 z_3$ , т.е.  $\varphi(x_2) = y_2$ . Аналогично,  $\varphi(x_3) = y_3$ . Следовательно,  $\varphi(x_2 x_3) = y_2 y_3$ . Пусть  $z$  — точка пересечения прямых  $x_2 x_3$  и  $z_2 z_3$ . Тогда  $\varphi(z)$  лежит на прямой  $y_2 y_3$  и в то же время  $\varphi(z) = z$ . Следовательно,  $z = z_1$  и точка  $z_1$  лежит на прямой  $z_2 z_3$ , что и требовалось доказать.

## §5. ДЕЗАРГОВЫ ПЛОСКОСТИ

Приведенное доказательство теоремы Дезарга для проективной плоскости  $PG(2, F)$  существенным образом использует тот факт, что координатное кольцо  $F$  этой плоскости является полем. Как мы увидим далее, для других проективных плоскостей теорема Дезарга не выполняется.

**ОПРЕДЕЛЕНИЕ 5.1.** Проективная плоскость, в которой выполняется Теорема Дезарга, называется *дезарговой*.

Мы употребляем термин «теорема» по отношению к теореме Дезарга в этом определении, только отдавая дань традиции. По существу, дезаргова проективная плоскость — это проективная плоскость, удовлетворяющая дополнительной аксиоме

**D.** Два треугольника центрально перспективны в том и только в том случае, если они аксиально перспективны.

Перефразируя Льва Николаевича Толстого, *все дезарговы плоскости похожи друг на друга, каждая недезаргова плоскость недезаргова по-своему*. Следующая теорема подтверждает первую часть этого тезиса.

ТЕОРЕМА 5.2. Любое координатное кольцо конечной дезарговой проективной плоскости является полем.

Пусть  $F$  — координатное кольцо конечной дезарговой проективной плоскости  $\Pi = (P, \mathcal{L})$ . Мы должны проверить, что операции сложения и умножения, заданные формулами (6), удовлетворяют аксиомам поля. Справедливость нескольких аксиом установлена теоремой 3.1, так что остаются следующие аксиомы: (F1)  $a + b = b + a$ , (F2)  $(a + b) + c = a + (b + c)$ , (F3)  $(ab)c = a(bc)$ , (F4)  $(a + b)c = ac + bc$ , (F5)  $a(b + c) = ab + ac$ , (F6)  $ab = ba$ . При доказательстве свойств (F1)–(F5) мы многократно используем теорему Дезарга. Каждый раз мы выбираем три прямые  $L_1, L_2, L_3$ , проходящие через одну и ту же точку, называемую *полюсом*. Кроме того, мы выбираем точки  $x_1, y_1 \in L_1$ ,  $x_1 \neq y_1$ , точки  $x_2, y_2 \in L_2$ ,  $x_2 \neq y_2$ , и точки  $x_3, y_3 \in L_3$ ,  $x_3 \neq y_3$ . Прямые  $x_1x_2$  и  $y_1y_2$  пересекаются в точке  $z_3$ , прямые  $x_2x_3$  и  $y_2y_3$  пересекаются в точке  $z_1$ , прямые  $x_3x_1$  и  $y_3y_1$  пересекаются в точке  $z_2$ . Применяя теорему Дезарга, мы получаем, что точки  $z_1, z_2, z_3$  лежат на одной прямой, называемой *осью*.

В ходе доказательства мы будем называть прямые  $L_1$  и  $L_2$  *параллельными*,  $L_1 \parallel L_2$ , если эти прямые пересекают идеальную прямую  $[\infty]$  в одной и той же точке. Отметим также, что основные идеи доказательства позаимствованы из книги [3] (в которой, между прочим, речь идет не о конечных плоскостях).

*Коммутативность сложения.*

ЛЕММА 5.3. Для любых  $a, b \in F^*$ , прямые  $(a, 0)(0, a)$  и  $(b, 0)(0, b)$  параллельны.

ДОКАЗАТЕЛЬСТВО. Пусть  $L_1 = [0, 0]$ ,  $L_2 = [1, 0]$ ,  $L_3 = [0]$  с полюсом  $(0, 0)$ . Пусть  $x_1 = (a, 0)$ ,  $y_1 = (b, 0)$ ,  $x_2 = (a, a)$ ,  $y_2 = (b, b)$ ,  $x_3 = (0, a)$ ,  $y_3 = (0, b)$  (рис. 4). Тогда  $x_1x_2 = [a]$ ,  $y_1y_2 = [b]$ , так что  $z_3 = (\infty)$ ;  $x_2x_3 = [0, a]$ ,  $y_2y_3 = [0, b]$ , так что  $z_1 = (0)$ . Следовательно, ось — это прямая  $(0)(\infty) = [\infty]$  и точка пересечения прямых  $x_3x_1$  и  $y_3y_1$  лежит на этой прямой, что и требовалось доказать.

Мы обозначим  $(-1)$  общую идеальную точку всех прямых  $(a, 0)(0, a)$ ,  $a \in F^*$ .

ЛЕММА 5.4. Точка  $(x, y)$  лежит на прямой  $[-1, c]$  в том и только в том случае, если  $x + y = c$ .

ДОКАЗАТЕЛЬСТВО. Прямая  $[-1, c]$  проходит через точки  $(-1)$  и  $(0, c)$ . В силу леммы 5.3 она проходит также через точку  $(c, 0)$ .

Пусть  $(x, y) \in [-1, c]$ ,  $x \neq 0$ ,  $y \neq 0$ ,  $x \neq c$ ,  $y \neq c$ . Пусть  $L_1 = [0]$ ,  $L_2 = [-1, c]$ ,  $L_3 = [0, c]$  с полюсом  $(0, c)$ . Пусть  $x_1 = (0, 0)$ ,  $y_1 = (0, y)$ ,  $x_2 = (c, 0)$ ,  $y_2 = (x, y)$ ,  $x_3 = (c, c)$ ,  $y_3 = (x, c)$  (рис. 5). Тогда  $x_1x_2 = [0, 0]$ ,

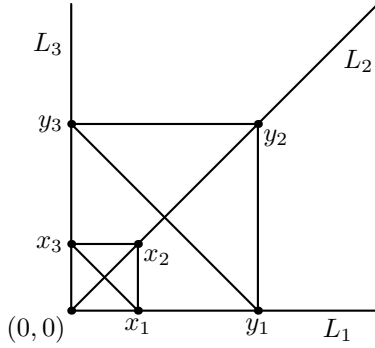


Рис. 4.

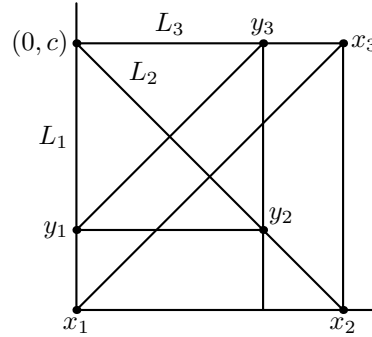


Рис. 5.

$y_1y_2 = [0, y]$ , так что  $z_3 = (0)$ ;  $x_2x_3 = [c]$ ,  $y_2y_3 = [x]$ , так что  $z_1 = (\infty)$ . Следовательно, ось — это прямая  $[\infty]$ , и точка пересечения прямых  $x_3x_1 = [1, 0]$  и  $y_3y_1$  лежит на оси. Так как (1) — идеальная точка на прямой  $x_3x_1$ , то (1) лежит и на прямой  $y_3y_1$ , а так как  $y_1 = (0, y)$ , мы получаем, что  $y_3y_1 = [1, y]$ .

Следовательно,  $y_3 \in [1, y]$ , т.е.  $y_3 = (x, c)$  — это точка пересечения прямых  $[x]$  и  $[1, y]$ . По определению сложения  $x + y = c$ .

Так как для каждого  $x \in F$  уравнение  $x + y = c$  имеет решение (теорема 3.1), то точки  $(x, z)$ , где  $x + z \neq c$ , не лежат на прямой  $[1, c]$ .

Мы готовы доказать коммутативность сложения, применив еще раз теорему Дезарга. Пусть  $a \neq 0$ ,  $b \neq 0$ . Пусть  $L_1 = [a]$ ,  $L_2 = [1, 0]$ ,  $L_3 = [0, a]$  с полюсом  $(a, a)$ . Пусть  $x_1 = (a, 0)$ ,  $y_1 = (a, b)$ ;  $x_2 = (0, 0)$ ,  $y_2 = (b, b)$ ;  $x_3 = (0, a)$ ,  $y_3 = (b, a)$  (рис. 6). Тогда  $x_1x_2 = [0, 0]$ ,  $y_1y_2 = [0, b]$ , так что  $z_3 = (0)$ ;  $x_2x_3 = [0]$ ,  $y_2y_3 = [b]$ , так что  $z_1 = (\infty)$ . Опять идеальная прямая  $[\infty]$  оказывается осью. В силу леммы 5.3 прямая  $x_1x_3$  пересекает ось в

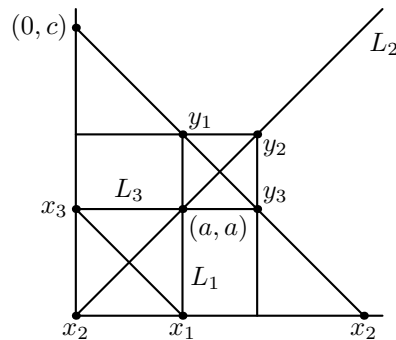


Рис. 6.

точке  $(-1)$ , так что прямая  $y_1y_3$  проходит через  $(-1)$ . Пусть эта прямая пересекает ось ординат в точке  $(0, c)$ . Из леммы 5.4 следует, что  $a + b = c$  и  $b + a = c$ , т. е.  $a + b = b + a$ .

ЛЕММА 5.5. Пусть  $a, b \in F^*$ .

- (а) Все прямые  $(a + t, 0)(t, b)$ , где  $t \in F$ , параллельны друг другу.
- (б) Все прямые  $(0, a + t)(b, t)$ , где  $t \in F$ , параллельны друг другу.

ДОКАЗАТЕЛЬСТВО. (а) Мы докажем, что все такие прямые параллельны прямой  $(a, 0)(0, b)$ . Предположим сначала, что  $a = b$ . В силу леммы 5.4 прямая  $(a + t, 0)(t, a)$  совпадает с прямой  $[-1, a + t]$  и потому (лемма 5.3) параллельна прямой  $(a, 0)(0, a)$ .

Пусть  $a \neq b$  и  $t \neq 0$ . Положим  $L_1 = [0, 0]$ ,  $L_2 = [0, a]$ ,  $L_3 = [0, b]$  (с полюсом  $(0)$ ),  $x_1 = (a, 0)$ ,  $y_1 = (a + t, 0)$ ,  $x_2 = (0, a)$ ,  $y_2 = (t, a)$ ,  $x_3 = (0, b)$ ,  $y_3 = (t, b)$  (рис. 7). Прямые  $x_1x_2$  и  $y_1y_2$  пересекаются в точке  $(-1)$  (леммы 5.3 и 5.4), прямые  $x_2x_3$  и  $y_2y_3$  пересекаются в точке  $(\infty)$ , так что прямые  $x_1x_3$  и  $y_1y_3$  должны пересечься в точке, лежащей на идеальной прямой  $[\infty]$ , что и требовалось доказать.

Утверждение (б) доказывается аналогично.

*Ассоциативность сложения.* Пусть  $a, b, c \in F^*$ . Пусть  $d \in F^*$ . Рассмотрим точки  $x_1 = (b, 0)$ ,  $y_1 = (b + c, 0)$ ,  $x_2 = (0, d)$ ,  $y_2 = (c, d)$ ,  $x_4 = (a + b, 0)$ ,  $y_4 = ((a + b) + c, 0)$ ,  $x_5 = (b, d)$ ,  $y_5 = (b + c, d)$ . Пусть  $x_3$  — точка пересечения прямых  $x_1x_5$  и  $x_2x_4$ , а  $y_3$  — точка пересечения прямых  $y_1y_5$  и  $y_2y_4$  (рис. 8). Из леммы 5.5 следует, что  $x_1x_2 \parallel y_1y_2$  и  $x_2x_4 \parallel y_2y_4$ . Кроме того,  $x_1x_5 \parallel y_1y_5$ . Следовательно, треугольники  $x_1x_2x_3$ ,  $y_1y_2y_3$  аксиально перспективны.

Так как прямые  $x_1y_1$ ,  $x_2y_2$  проходят через  $(0)$ , через эту же точку (теорема Дезарга) проходит прямая  $x_3y_3$ . Следовательно, треугольники  $x_3x_4x_5$ ,

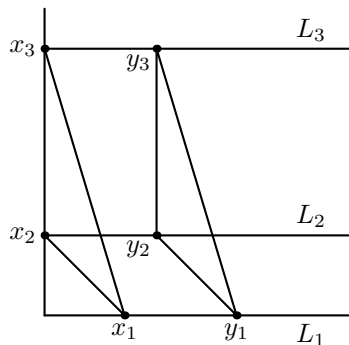


Рис. 7.

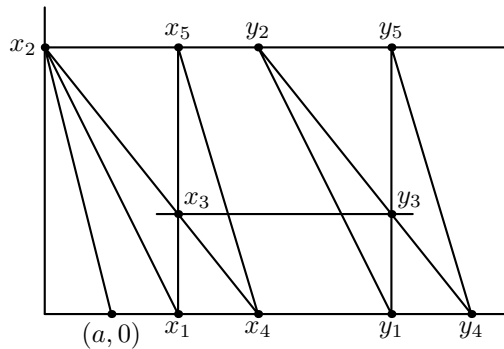


Рис. 8.

$y_3y_4y_5$  центрально перспективны. Так  $x_3x_4 \parallel y_3y_4$  и  $x_3x_5 \parallel y_3y_5$ , то по теореме Дезарга  $x_4x_5 \parallel y_4y_5$ .

Из леммы 5.5 следует, что  $x_4x_5 \parallel (a, 0)(0, d)$ . Следовательно,  $y_4y_5 \parallel (a, 0)(0, d)$ . С другой стороны,  $((b + c) + a, 0)y_5 \parallel (a, 0)(0, d)$  (лемма 5.5). Следовательно,  $y_4 = ((b + c) + a, 0)$ , т. е.  $((a + b) + c, 0) = ((b + c) + a, 0) = (a + (b + c), 0)$ , так что  $(a + b) + c = a + (b + c)$ .

Прежде чем перейти к свойствам умножения, мы получим более наглядное представление произведения.

**ЛЕММА 5.6.** Для любых  $a, b \in F^*$  прямые  $(1, 0)(0, b)$  и  $(a, 0)(0, ab)$  параллельны.

**ДОКАЗАТЕЛЬСТВО.** Если  $a = 1$ , утверждение леммы очевидно; если  $b = 1$ , см. лемму 5.3. Пусть  $a \neq 1, b \neq 1$ . Рассмотрим прямые  $L_1 = [0]$ ,  $L_2 = [b, 0]$ ,  $L_3 = [0, 0]$  с полюсом  $(0, 0)$ . Пусть  $x_1 = (0, b)$ ,  $y_1 = (0, ab)$ ,  $x_2 = (1, b)$ ,  $y_2 = (a, ab)$ ,  $x_3 = (1, 0)$ ,  $y_3 = (a, 0)$  (рис. 9). Так как  $x_1x_2 \parallel y_1y_2$  и  $x_2x_3 \parallel y_2y_3$ , то по теореме Дезарга  $x_1x_3 \parallel y_1y_3$ .

*Ассоциативность умножения.*

При доказательстве формулы  $(ab)c = a(bc)$  мы можем предположить, что  $a, b, c$  не равны ни 0, ни 1. Пусть  $x_1 = (b, 0)$ ,  $y_1 = (ab, 0)$ ,  $x_2 = (0, b)$ ,  $y_2 = (0, ab)$ ,  $x_4 = (1, 0)$ ,  $y_4 = (a, 0)$ ,  $x_5 = (0, bc)$ ,  $y_5 = (0, (ab)c)$ . Пусть  $x_3$  — точка пересечения прямых  $x_1x_5$  и  $x_2x_4$ ,  $y_3$  — точка пересечения прямых  $y_1y_5$  и  $y_2y_4$  (рис. 10).

Из леммы 5.3 следует, что  $x_1x_2 \parallel y_1y_2$ , а из леммы 5.6 следует, что  $x_4x_2 \parallel y_4y_2$ ,  $x_1x_5 \parallel y_1y_5$ . Следовательно, треугольники  $x_1x_2x_3$  и  $y_1y_2y_3$  аксиально перспективны. По теореме Дезарга прямые  $x_1y_1$ ,  $x_2y_2$  и  $x_3y_3$  пересекаются в одной точке. Поэтому треугольники  $x_3x_4x_5$  и  $y_3y_4y_5$  центрально перспективны. Применяя еще раз теорему Дезарга, мы получим,

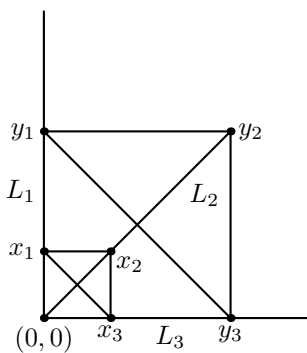


Рис. 9.

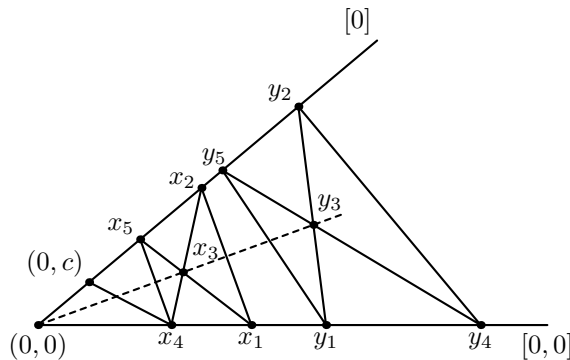


Рис. 10.



что  $x_4x_5 \parallel y_4y_5$ . Теперь из леммы 5.6 следует, что  $y_5 = (0, a(bc))$ , так что  $(ab)c = a(bc)$ .

*Правая дистрибутивность:*  $(a + b)c = ac + bc$ .

Мы можем предположить, что  $a, b, c \in F^*$ . Рассмотрим прямые  $(1, 0)(0, c)$ ,  $(a, 0)(0, ac)$ ,  $(b, 0)(0, bc)$  (рис. 11). В силу леммы 5.6 эти прямые параллельны. Проведем параллельную им прямую через точку  $(b, ac)$ . Пусть эта прямая пересекает оси координат в точках  $(x, 0)$  и  $(0, y)$ . Применим лемму 5.5 дважды:

$$(a + b, 0)(b, ac) \parallel (a, 0)(0, ac), \quad (0, bc + ac)(b, ac) \parallel (0, bc)(b, 0).$$

Первая из этих параллельностей означает, что  $x = a + b$ , вторая означает, что  $y = bc + ac = ac + bc$ . Следовательно,  $(a + b, 0)(0, ac + bc) \parallel (1, 0)(0, c)$  и поэтому  $ac + bc = (a + b)c$  (лемма 5.6).

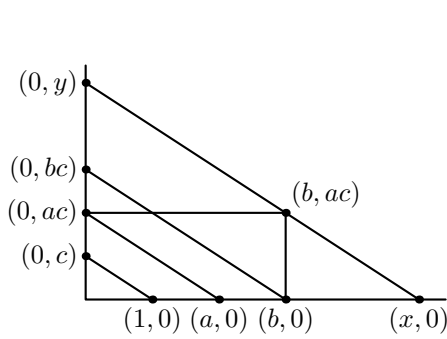


Рис. 11.

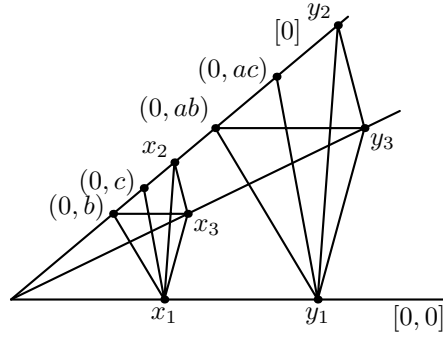


Рис. 12.

*Левая дистрибутивность:*  $a(b + c) = ab + ac$ .

Мы предположим, что  $a, b, c \in F^*$  и  $a \neq 1$ . Рассмотрим точки  $x_1Q$   $Q = (1, 0)$ ,  $y_1 = (a, 0)$ ,  $x_2 = (0, b+c)$ ,  $y_2 = (0, a(b+c))$ ,  $x_3 = (1, b)$ ,  $y_3 = (a, ab)$  (рис. 12). Прямые  $x_1y_1 = [0, 0]$  и  $x_2y_2 = [0]$  проходят через точку  $(0, 0)$ . По определению умножения  $(1, b) = (1, 1 \cdot b) \in [b, 0]$  и  $(a, ab) \in [b, 0]$ , так что прямая  $x_3y_3$  тоже проходит через  $(0, 0)$ . Так как  $x_1x_2 \parallel y_1y_2$  (лемма 5.6) и  $x_3x_1 \parallel y_3y_1$ , то из теоремы Дезарга следует, что  $x_3x_2 \parallel y_3y_2$ , т. е.  $(1, b)(0, b + c) \parallel (a, ab)(0, a(b + c))$ . Применим лемму 5.5:

$$(1, b)(0, b + c) \parallel (1, 0)(0, c), \quad (a, ab)(0, ab + ac) \parallel (a, 0)(0, ac).$$

В силу леммы 5.6 прямые  $(a, 0)(0, ac)$  и  $(1, 0)(0, c)$  параллельны, так что мы получаем, что

$$(a, ab)(0, ab + ac) \parallel (1, 0)(0, c) \parallel (1, b)(0, b + c) \parallel (a, ab)(0, a(b + c)).$$

Следовательно,  $a(b + c) = ab + ac$ .

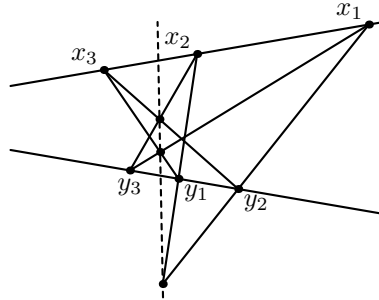


Рис. 13.

*Коммутативность умножения.*

Заметим, что при выводе аксиом поля мы до сих пор не пользовались конечностью данной проективной плоскости. Коммутативность умножения мы теперь получаем «даром» благодаря знаменитой теореме Веддерберна, которая утверждает, что в конечном случае коммутативность умножения является следствием остальных аксиом поля.

В бесконечном случае теорема Дезарга недостаточна для вывода коммутативности умножения — вместо нее требуется более сильная

**ТЕОРЕМА ПАППА.** Если точки  $x_1, x_2, x_3$  проективной плоскости лежат на одной прямой, точки  $y_1, y_2, y_3$  лежат на другой прямой и все шесть точек отличны от точки пересечения этих прямых, то точки пересечения прямых  $x_1y_2$  и  $y_1x_2$ , прямых  $x_2y_3$  и  $y_2x_3$ , прямых  $x_3y_1$  и  $y_3x_1$  лежат на одной прямой (рис. 13).

Однако, рассмотрение бесконечных проективных плоскостей не входит в нашу задачу, и мы перейдем к конечным недезарговым плоскостям.

## §6. НЕДЕЗАРГОВЫ ПЛОСКОСТИ

Чтобы определить, насколько недезарговой может быть недезаргова проективная плоскость, мы вернемся к доказательству теоремы Дезарга для плоскостей  $PG(2, F)$ . Большая часть приведенного рассуждения (за исключением последнего абзаца) состояла в доказательстве следующего утверждения:

если даны точка  $c$  и прямая  $A$ , то для любых точек  $x, y$ , лежащих на одной прямой с точкой  $c$ , отличных от  $c$  и не лежащих на  $A$ , существует биективное отображение множества всех точек плоскости в себя, которое (1) переводит каждую прямую в прямую, (2) переводит в себя все точки прямой  $A$  и все прямые, проходящие через  $c$ , и (3) переводит точку  $x$  в точку  $y$ .

ОПРЕДЕЛЕНИЕ 6.1. Пусть  $\Pi = (P, \mathcal{L})$  — проективная плоскость. Биективное отображение  $f: P \rightarrow P$  называется *коллинеацией* плоскости  $\Pi$ , если для любой прямой  $L$  множество  $f(L) = \{f(x): x \in L\}$  является прямой (точнее, множеством всех точек, лежащих на некоторой прямой).

Пусть  $c$  — точка,  $A$  — прямая в плоскости  $\Pi$ . Коллинеация  $f$  называется  $(c, A)$ -коллинеацией, если  $f(L) = L$  для любой прямой  $L$ , проходящей через  $c$ , и  $f(x) = x$  для любой точки  $x$ , лежащей на  $A$ . Точка  $c$  называется *центром* (или *полосом*), а прямая  $A$  — *осью* коллинеации  $f$ . Если  $c \notin A$ , то  $(c, A)$ -коллинеация называется  $(c, A)$ -гомологией. Если  $c \in A$ , то  $(c, A)$ -коллинеация называется  $(c, A)$ -элацией.

Пусть  $f$  —  $(c, A)$ -коллинеация проективной плоскости  $\Pi$ . Предположим, нам известен образ  $y = f(x)$  некоторой точки  $x$ , отличной от  $c$  и не лежащей на  $A$ . Пусть  $u$  — любая точка, не лежащая на  $A$  и не лежащая на прямой  $cx$  (рис. 14). Чтобы найти  $v = f(u)$ , мы построим точку пересечения прямых  $ux$  и  $A$ . Прямая, проходящая через эту точку и точку  $y$ , должна пересечь прямую  $cu$  в точке  $v$ . Поскольку теперь мы знаем образ точки  $u$ , мы можем применить аналогичное построение для нахождения образа любой точки прямой  $cx$ .

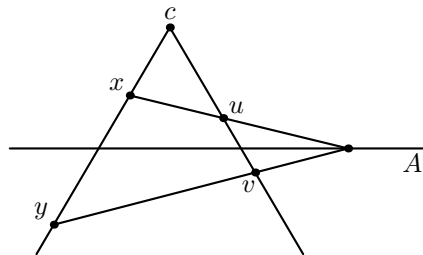


Рис. 14.

Таким образом,  $(c, A)$ -коллинеация однозначно определяется образом одной точки (отличной от  $c$  и не лежащей на  $A$ ).

ОПРЕДЕЛЕНИЕ 6.2. Пусть  $\Pi = (P, \mathcal{L})$  — проективная плоскость,  $c \in P$ ,  $A \in \mathcal{L}$ . Плоскость  $\Pi$  называется  $(c, A)$ -транзитивной, если для любых точек  $x, y$ , лежащих на одной прямой с точкой  $c$ , отличных от  $c$  и не лежащих на  $A$ , существует  $(c, A)$ -коллинеация  $f$  такая, что  $f(x) = y$ .

Пусть  $\Pi = (P, \mathcal{L})$  — конечная проективная плоскость порядка  $n$ ,  $c \in P$ ,  $A \in \mathcal{L}$ . Нетрудно видеть, что все  $(c, A)$ -коллинеации плоскости  $\Pi$  образуют группу (относительно композиции отображений). Поскольку  $(c, A)$ -коллинеация однозначно определяется образом одной точки, отличной от  $c$  и не лежащей на  $A$ , и число таких точек равно  $n$  или  $n - 1$ , в зависимости от того, лежит ли  $c$  на  $A$ , мы получаем следующее утверждение.

**ТЕОРЕМА 6.3.** Пусть  $\Pi = (P, \mathcal{L})$  — конечная проективная плоскость порядка  $n$ ,  $c \in P$ ,  $A \in \mathcal{L}$ . Плоскость  $\Pi$  является  $(c, A)$ -транзитивной в том и только в том случае, если порядок группы всех  $(c, A)$ -коллинеаций равен  $n - 1$  (если  $c \notin A$ ) или  $n$  (если  $c \in A$ ).

Таким образом, доказательство теоремы Дезарга, приведенное в §4, показывает, что проективная плоскость над полем  $(c, A)$ -транзитивна для любой точки  $c$  и любой прямой  $A$ . Это и другие соображения привели Ленца и Барлотти в 50-х годах прошлого века к идее классифицировать проективные плоскости согласно структуре множества  $T(\Pi)$  всех пар  $(c, A)$ , для которых плоскость  $\Pi$  является  $(c, A)$ -транзитивной. Множество  $T(\Pi)$  подчиняется ряду ограничений. Для примера приведем без доказательства следующий результат.

**ТЕОРЕМА 6.4.** Если  $(c, A), (d, A) \in T(\Pi)$ , где  $c \neq d$ , то  $(a, A) \in T(\Pi)$  для любой точки  $a$  на прямой  $cd$ .

В настоящее время *классификация Ленца – Барлотти* для конечных проективных плоскостей насчитывает 15 классов. Два крайних класса — дезарговы плоскости, для которых  $T(\Pi)$  состоит из всех пар  $(c, A)$ , и *антидезарговы плоскости*, для которых  $T(\Pi) = \emptyset$ .

## §7. АНТИДЕЗАРГОВА ПЛОСКОСТЬ

Можно доказать, что все проективные плоскости порядка 2, 3, 4, 5, 7, и 8 — дезарговы. Не существует проективной плоскости порядка 6 (см. §8). Поскольку существует почтиполе  $F$  порядка 9, не являющееся полем (см. §2), равенства (1)–(5) определяют недезаргову плоскость порядка 9 над  $F$ . (Если бы эта плоскость была дезарговой, то  $F$  было бы полем.) Можно показать, однако, что эта плоскость не антидезаргова. В этом параграфе мы используем то же самое почтиполе  $F$  для построения антидезарговой проективной плоскости порядка 9.

Почтиполе  $F$  состоит из 9 элементов:  $0, \pm 1, \pm i, \pm j, \pm k$ , которые перемножаются как кватернионы. По сложению  $F$  — абелева группа. Всю таблицу сложения для  $F$  можно вывести из следующих правил: (1)  $a + a = -a$  для всех  $a \in F$ ; (2)  $i - j = j - k = k - i = 1$ . Например,  $i + j = i - j + j + j = 1 - j = -k$ . Ядро  $K$  почтиполя  $F$  состоит из элементов  $0, 1, -1$ .

Пусть  $V$  — множество всех ненулевых  $3 \times 1$  матриц  $A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ , где  $a_1, a_2, a_3 \in F$ , так что  $|V| = 9^3 - 1 = 728$ . Матрицы  $A$  и  $At$ , где  $t \in F$ ,  $t \neq 0$ , назовем эквивалентными. Тогда множество  $V$  разбивается на  $728/8 = 91$  класс эквивалентности. Каждый класс мы назовем точкой будущей проективной плоскости  $\Pi = (P, \mathcal{L})$ , так что  $|P| = 91 = 9^2 + 9 + 1$ . Мы будем обозначать через  $\langle A \rangle$  точку, содержащую  $A \in V$ . Если  $M$  —

невырожденная  $3 \times 3$  матрица над  $F$ , то  $MA \in V$  для любой матрицы  $A \in V$  и, более того, из  $\langle A \rangle = \langle B \rangle$  следует  $\langle MA \rangle = \langle MB \rangle$ . (При проверке последнего факта потребуется как раз тот дистрибутивный закон, который выполняется в  $F$ .) Поэтому мы можем считать, что матрица  $M$  действует на множестве  $P$ .

Выберем невырожденную  $3 \times 3$  матрицу  $M$  над полем  $K$ , обладающую следующим свойством:  $M^{13}$  — диагональная матрица и ни одна из матриц  $M^s$ , где  $1 \leq s \leq 12$ , не является диагональной. Примером такой матрицы может служить  $M = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$ .

Для каждого  $x \in F$ , обозначим через  $L(x)$  множество всех точек  $\langle A \rangle \in P$  таких, что произведение  $1 \times 3$  матрицы  $[1 \ x \ 1]$  и  $3 \times 1$  матрицы  $A$  равно 0. Нетрудно подсчитать, что для  $x \neq 0$  имеется 80 ненулевых матриц  $A$ , удовлетворяющих этому условию, и они образуют 10 классов эквивалентности. Следовательно,  $|L(x)| = 10$ . Теперь для  $x = 1$  и для каждого  $x \in F \setminus K$  определим 13 прямых  $L_s(x)$ , где  $0 \leq s \leq 12$ , следующим образом:  $L_s(x) = \{M^s a : a \in L(x)\}$ . Обозначим через  $\mathcal{L}$  множество всех таких прямых. Тогда  $|\mathcal{L}| = 7 \cdot 13 = 91$ . Мы опустим проверку того, что  $\Pi = (P, \mathcal{L})$  — проективная плоскость.

Назовем точку  $a \in P$  *ближней*, если  $a = \langle A \rangle$ , где все элементы матрицы  $A$  принадлежат полю  $K$ . Все точки, не являющиеся ближними, мы назовем *дальними*. Можно проверить, что каждая прямая содержит либо четыре, либо одну ближнюю точку. Соответственно мы назовем прямую либо *ближней*, либо *дальней*. Дальнейший анализ плоскости  $\Pi$  показывает, что каждая ближняя точка лежит на четырех ближних прямых, а каждая дальняя точка лежит на одной ближней прямой. Отсюда следует, что ближние точки и ближние прямые образуют проективную плоскость порядка 3. Наконец, преодолев определенные технические трудности, можно показать, что никакая коллинеация плоскости  $\Pi$  не переводит ближнюю точку в дальнюю.

**ТЕОРЕМА 7.1.** Проективная плоскость  $\Pi$  антидезаргова.

**ДОКАЗАТЕЛЬСТВО.** Пусть  $s \in P$  и  $A \in \mathcal{L}$ . В силу сделанного выше замечания о коллинеациях плоскости  $\Pi$  достаточно показать, что существует прямая  $L$ , проходящая через  $s$ , и две точки на  $L$ , отличные от  $s$  и не лежащие на  $A$ , одна из которых — ближняя, а другая — дальняя. Если  $s$  — дальняя точка, то в качестве  $L$  возьмем любую дальнюю прямую, проходящую через  $s$  и отличную от  $A$ . Если  $s$  — ближняя точка, то  $L$  — любая ближняя прямая, проходящая через  $s$  и отличная от  $A$ .

Аналогичное построение можно провести, если  $F$  — любое почтиполе порядка  $q^2$ , описанное в §2. Показатель степени 13 в описании матрицы  $M$  нужно заменить на  $q^2 + q + 1$ . Построенные таким образом антидезарговы

проективные плоскости называются *плоскостями Хьюза*. Построение и свойства плоскостей Хьюза можно найти в книгах [2] и [4]. Антидезаргова и три других проективных плоскости порядка 9 детально проанализированы в элементарно написанной книге [7].

### §8. ТРИ ГИПОТЕЗЫ

Порядок любой известной конечной проективной плоскости является простым числом или степенью простого числа. Однако, единственное известное универсальное ограничение на порядок представлено следующей теоремой, доказанной в 1949 году.

**ТЕОРЕМА БРАКА – РАЙЗЕРА.** Пусть  $n \equiv 1$  или  $2 \pmod{4}$ . Если существует проективная плоскость порядка  $n$ , то  $n$  может быть представлено в виде суммы квадратов двух целых чисел.

Таким образом, не существует проективных плоскостей порядка 6, 14, 21, 22, 30 и многих других порядков. Первый случай, не покрытый этой теоремой, —  $n = 10$ . Огромный компьютерный поиск, завершённый в 1991 году, убедил специалистов, что проективной плоскости порядка 10 не существует (см. [6]).

Следующий порядок, не покрытый теоремой Брака – Райзера, —  $n = 12$ . К настоящему времени известно, что если проективная плоскость порядка 12 существует, то она должна быть антидезарговой.

Вопрос о возможном порядке проективной плоскости тесно связан с *взаимно ортогональными латинскими квадратами*. Латинский квадрат порядка  $n$  — это квадратная матрица порядка  $n$ , в каждом столбце и каждой строке которой встречаются все натуральные числа от 1 до  $n$ . Два латинских квадрата  $A = [a_{ij}]$  и  $B = [b_{ij}]$  одного и того же порядка  $n$  называются ортогональными, если для любой упорядоченной пары  $(k, l)$  натуральных чисел, не превосходящих  $n$ , можно найти индексы  $i$  и  $j$  такие, что  $a_{ij} = k$ ,  $b_{ij} = l$ . Например, следующие три латинских квадрата порядка 4 взаимно ортогональны:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \end{bmatrix}$$

**ТЕОРЕМА 8.1.** Проективная плоскость порядка  $n$  существует в том и только в том случае, если существуют  $n - 1$  взаимно ортогональных латинских квадратов порядка  $n$ .

Интересно, что хотя невозможность проективной плоскости порядка 10 эквивалентна отсутствию девяти попарно ортогональных латинских

квадратов порядка 10, никто пока не сумел найти даже трех таких квадратов (два ортогональных латинских квадрата порядка 10 существуют).

Несмотря на значительные усилия, следующие две гипотезы остаются открытыми.

**ГИПОТЕЗА 1.** Порядок любой конечной проективной плоскости — простое число или степень простого числа.

**ГИПОТЕЗА 2.** Все проективные плоскости простого порядка де-зарговы.

Один из возможных подходов к доказательству (или опровержению) этих гипотез — проанализировать их в применении к каждому из 15 классов Ленца – Барлотти (см. §7). Чем больше множество  $T(\Pi)$  (множество всех пар  $(c, A)$ , для которых плоскость  $\Pi$   $(c, A)$ -транзитивна), тем богаче алгебраическая структура координатного кольца плоскости  $\Pi$  и тем больше шансов, что анализ этой структуры даст необходимую информацию о порядке плоскости. Пять из 15 классов Ленца – Барлотти, возможно, являются пустыми (хотя в каждом из них есть бесконечные плоскости).

**КЛАСС I.2.**  $T(\Pi)$  состоит из единственной пары  $(c, A)$ , где  $c \notin A$ .

**КЛАСС I.3.**  $T(\Pi)$  состоит из двух пар,  $(a, A)$  и  $(b, B)$ , где  $a \in B$ ,  $a \notin A$ ,  $b \in A$ ,  $b \notin B$ .

**КЛАСС I.4.**  $T(\Pi)$  состоит из трех пар, образованных вершинами и противоположными сторонами треугольника.

**КЛАСС II.2.**  $T(\Pi)$  состоит из двух пар,  $(c, A)$  и  $(b, B)$ , где  $b \in A$ ,  $b \notin B$ ,  $c$  — точка пересечения прямых  $A$  и  $B$ .

**КЛАСС III.1.**  $T(\Pi)$  состоит из всех пар  $(x, cx)$ , где  $x$  — произвольная точка фиксированной прямой  $L$ , а  $c$  — фиксированная точка, не лежащая на  $L$ .

**ГИПОТЕЗА 3.** Перечисленные выше классы Ленца – Барлотти не содержат конечных проективных плоскостей.

## §9. $H$ -МАТРИЦЫ

В четырех из пяти перечисленных выше классов Ленца – Барлотти плоскость  $\Pi$   $(c, A)$ -транзитивна, где  $c \notin A$ , т.е.  $\Pi$  допускает группу  $(c, A)$ -гомологий порядка  $n - 1$ , где  $n$  — порядок  $\Pi$ . В этом случае свойства плоскости  $\Pi$  тесно связаны со свойствами определенной матрицы, к описанию которой мы переходим.

Итак, пусть  $\Pi = (P, \mathcal{L})$  — проективная плоскость порядка  $n$ , содержащая такие прямую и не лежащую на ней точку, что соответствующая группа гомологий  $H$  имеет порядок  $n - 1$ . Вспомним, что при координатизации проективной плоскости (см. §3) мы можем выбрать в качестве осей

координат и идеальной прямой любые три прямые, не пересекающиеся в одной точке. Мы предположим, что ось и центр гомологий из группы  $H$  — прямая  $[0]$  и точка  $(0)$ . В качестве координатного кольца мы выберем множество  $F = H \cup \{0\}$  (мы предполагаем, что  $H$  — мультипликативная группа с нейтральным элементом  $1$ ). Точку  $(1, 1)$  мы выберем произвольно и затем найдем точку  $(1)$  на идеальной прямой. Так как все коллинеации из группы  $H$  являются  $((0), [0])$ -гомологиями, они преобразуют идеальную прямую в себя (оставляя на месте точки  $(0)$  и  $(\infty)$ ). Для каждого  $a \in H$  обозначим через  $(a)$  точку  $a(1)$ . После этого координатизация плоскости  $\Pi$  завершается как в §3. Достоинство предложенной координатизации в том, что мы можем описать действие коллинеаций из группы  $H$  на всей плоскости  $\Pi$ .

**ПРЕДЛОЖЕНИЕ 9.1.** Пусть  $a \in H$  и  $x, y \in F$ . Тогда: (1)  $a(\infty) = (\infty)$  и  $a[\infty] = [\infty]$ ; (2)  $a(x) = (ax)$  и  $a[x, 0] = [ax, 0]$ ; (3)  $a(x, 0) = (xa^{-1}, 0)$  и  $a[x] = [xa^{-1}]$ ; (4)  $a(x, y) = (xa^{-1}, y)$  и  $a[x, y] = [ax, y]$ .

Мы докажем (2) и оставим (1), (3) и (4) читателю. Так как  $(0)$  — центр гомологии  $a$ ,  $a(0) = (0) = (a \cdot 0)$ ; если  $x \in H$ , то  $a(x) = a(x(1)) = (ax)(1) = (ax)$ ; так как прямая  $[x, 0]$  проходит через точку  $(x)$  и точку  $(0, 0)$ , лежащую на оси гомологии  $a$ ,  $a[x, 0] = [ax, 0]$ .

Операция умножения в координатном кольце  $F$  совпадает с операцией умножения в группе  $H$ , т. е. для любых  $a, b \in H$ , точка  $(a, ab)$  является точкой пересечения прямых  $[a]$  и  $[b, 0]$ . Для доказательства применим гомологию  $b^{-1}$  к точке  $(a, b)$  и к обоим прямым:  $b^{-1}(a, ab) = (ab, ab)$ ,  $b^{-1}[a] = [ab]$ ,  $b^{-1}[b, 0] = [1, 0]$ . Остается заметить, что  $(ab, ab)$  — точка пересечения прямых  $[ab]$  и  $[1, 0]$ .

Для индексирования строк и столбцов матрицы порядка  $n$  обычно используются числа  $1, 2, \dots, n$ . Однако, любое множество из  $n$  элементов может быть использовано для этой цели и мы рассмотрим матрицу  $C$  порядка  $n$  над координатным кольцом  $F$ , строки и столбцы которой индексированы (в одном и том же порядке) элементами множества  $F$ . При этом удобно считать, что первая строка и первый столбец индексируются нулем. Для любых  $x, y, z \in F$  положим

$$C(x, y) = z \text{ в том и только том случае, если } (1, x) \in [z, y].$$

Мы обозначаем через  $C(x, y)$  элемент матрицы  $C$ , расположенный в строке с индексом  $x$  и столбце с индексом  $y$ . Поскольку существует единственная прямая через точки  $(1, x)$  и  $(0, y)$  и эта прямая содержит единственную идеальную точку  $(z) \neq (\infty)$ , матрица  $C$  определена корректно.

**ПРЕДЛОЖЕНИЕ 9.2.** Пусть  $a, b \in F$ ,  $a \neq b$ . Тогда

$$(1) C(a, a) = 0, C(a, 0) = a \text{ и } C(a, b) \in H;$$



(2) каждая строка и каждый столбец матрицы  $C$  содержат все элементы множества  $F$  по одному разу;

(3) множество  $Q(a, b) = \{C(a, x)C(b, x)^{-1} : x \in F, x \neq a, x \neq b\}$  совпадает с множеством всех неединичных элементов группы  $H$ .

ДОКАЗАТЕЛЬСТВО. (1)  $(1, a) \in [0, a]$ ,  $(1, a) = (1 \cdot a, a) \in [a, 0]$  и  $(1, a) \notin [0, b]$ .

(2) Если  $C(x, b) = C(y, b) = z$ , то  $(1, x), (1, y) \in [z, b]$ . Так как, кроме того,  $(1, x), (1, y) \in [1]$ , мы заключаем, что  $x = y$ . Если  $C(a, x) = C(a, y) = z$ , то точка  $(1, a)$  лежит на прямых  $[z, x]$  и  $[z, y]$ . Так как эти прямые проходят также через точку  $(z)$ , то  $x = y$ . Следовательно, ни один элемент множества  $F$  не повторяется ни в строке, ни в столбце матрицы  $C$ .

(3) Пусть  $x, y \in F$ ,  $x \neq a$ ,  $x \neq b$ ,  $y \neq a$ ,  $y \neq b$ , и пусть  $x \neq y$ . Из предыдущего свойства следует, что  $1 \notin Q(a, b)$ . Пусть  $C(a, x) = s$ ,  $C(b, x) = t$ ,  $C(a, y) = u$ ,  $C(b, y) = v$ . Тогда  $(1, a)$  — точка пересечения прямых  $[s, x]$  и  $[u, y]$ ,  $(1, b)$  — точка пересечения прямых  $[t, x]$  и  $[v, y]$ . Предположим, что  $st^{-1} = uv^{-1} = z$ . Применив гомологию  $z$  к точке  $(1, b)$ , мы получаем, что  $z(1, b) = (z^{-1}, b)$  — точка пересечения прямых  $z[t, x] = [zt, x] = [s, x]$  и  $z[v, y] = [zv, y] = [u, y]$ , т. е.  $(z^{-1}, b) = (1, a)$ ,  $b = a$ . Полученное противоречие доказывает, что ни один элемент множества  $F$  не повторяется в множестве  $Q(a, b)$  и потому  $Q(a, b) = H \setminus \{1\}$ .

ОПРЕДЕЛЕНИЕ 9.3. Пусть  $H$  — мультипликативная группа порядка  $n - 1$  и пусть  $F = H \cup \{0\}$ . Матрица  $C$  порядка  $n$ , строки и столбцы которой индексированы множеством  $F$  называется  $H$ -матрицей, если она обладает свойствами (1)–(3) предложения 9.2.

ПРИМЕР. Пусть  $G = \{1, a, a^2, a^3\}$  — циклическая группа порядка 4, а  $H = \{1, b, b^2\}$  — циклическая группа порядка 3. Тогда  $M$  —  $G$ -матрица,  $N$  —  $H$ -матрица.

$$M = \begin{bmatrix} 0 & a^2 & a^3 & a & 1 \\ 1 & 0 & a^2 & a^3 & a \\ a & 1 & 0 & a^2 & a^3 \\ a^3 & a & 1 & 0 & a^2 \\ a^2 & a^3 & a & 1 & 0 \end{bmatrix} \quad N = \begin{bmatrix} 0 & 1 & b & b^2 \\ 1 & 0 & b^2 & b \\ b & b^2 & 0 & 1 \\ b^2 & b & 1 & 0 \end{bmatrix}$$

Таким образом, начав с проективной плоскости порядка  $n$  и группы гомологий  $H$  порядка  $n - 1$ , мы получили  $H$ -матрицу  $C$ . Справедливо и обратное утверждение, доказательство которого мы оставляем читателю.

ПРЕДЛОЖЕНИЕ 9.4. Пусть  $H$  — мультипликативная группа порядка  $n - 1$ . Для любой  $H$ -матрицы  $C$  существует проективная плоскость порядка  $n$ .

$n$  с координатным кольцом  $F = H \cup \{0\}$ , для которой  $H$  является группой  $((0), [0])$ -гомологий. При этом выполняются следующие свойства:

- (а) умножение ненулевых элементов в  $F$  совпадает с умножением в  $H$ ;
- (б) для любых  $a, b \in F, [a, b] = \{(x, y) : C(y, b) = xa\} \cup \{(a)\}$ ;
- (в) сложение в  $F$  определяется равенством  $C(a + b, b) = a$ .

Следующая теорема дает бесконечное множество  $H$ -матриц.

**ТЕОРЕМА 9.5.** Пусть  $F$  — поле или почтиполе порядка  $n$ . Определим матрицу  $C$  порядка  $n$  над  $F$  со строками и столбцами, индексированными множеством  $F$ :  $C(x, y) = x - y$ .

Тогда  $C$  —  $H$ -матрица, где  $H = F^*$ .

**ДОКАЗАТЕЛЬСТВО.** Первые два свойства  $H$ -матриц очевидны. Проверим свойство (3) из предложения 9.2. Пусть  $a, b \in F, a \neq b$ . Пусть  $x \in F, x \neq a, x \neq b$ . Тогда

$$\begin{aligned} C(x, a)C(x, b)^{-1} &= (x - a)(x - b)^{-1} = ((x - b) - (a - b))(x - b)^{-1} = \\ &= 1 - (a - b)(x - b)^{-1}. \end{aligned}$$

Когда  $x$  пробегает все элементы  $F$ , кроме  $a$  и  $b$ ,  $(x - b)^{-1}$  пробегает все элементы  $F$ , кроме  $(a - b)^{-1}$  и  $0$ , так что  $(a - b)(x - b)^{-1}$  пробегает все элементы  $F$ , кроме  $1$  и  $0$  и потому  $C(x, a)C(x, b)^{-1}$  пробегает все элементы  $F$ , кроме  $0$  и  $1$ , что и требовалось доказать.

Заметим, что матрица  $C$ , описанная в этой теореме, является кососимметрической, если характеристика  $F$  не равна  $2$ . В случае характеристики  $2$  матрица  $C$  симметрична. Оказывается, любая  $H$ -матрица над абелевой группой  $H$  симметрична или кососимметрична. Прежде чем доказать это утверждение, нужно понять, что означает кососимметричность  $H$ -матрицы для произвольной группы  $H$ . Мы называем матрицу  $M$  над полем  $F$  кососимметричной, если  $M^T = -M$ . Заметим, что  $-1$  — единственный элемент второго порядка в мультипликативной группе поля. Поэтому мы даем следующее определение.

**ОПРЕДЕЛЕНИЕ 9.6.** Пусть  $H$  — конечная группа с единственным элементом  $\pi$  второго порядка.  $H$ -матрица  $C$  называется *кососимметричной*, если  $C^T = \pi C$ .

**ТЕОРЕМА 9.7.** Пусть  $H$  — конечная абелева группа. Если группа  $H$  не имеет элементов второго порядка, то любая  $H$ -матрица симметрична. Если группа  $H$  имеет ровно один элемент второго порядка, то любая  $H$ -матрица кососимметрична. Если группа  $H$  имеет два или более элементов второго порядка, то  $H$ -матриц не существует.

Для доказательства теоремы нам понадобится следующая лемма.

ЛЕММА 9.8. Пусть  $\pi$  — произведение всех элементов конечной абелевой группы  $H$ . Если  $H$  имеет ровно один элемент второго порядка, то этот элемент равен  $\pi$ . В противном случае  $\pi = 1$ .

ДОКАЗАТЕЛЬСТВО. Все элементы группы  $H$ , порядок которых равен 1 или 2, образуют подгруппу  $K$ . Все элементы множества  $H \setminus K$  распадаются на пары  $\{x, x^{-1}\}$ . Поэтому произведение всех элементов группы  $H$  равно произведению всех элементов подгруппы  $K$ . Отсюда следует, что  $\pi \in K$ . Если  $H$  не имеет элементов порядка 2, то  $K = \{1\}$ , и потому  $\pi = 1$ . Если  $\tau$  — единственный элемент порядка 2 в  $H$ , то  $K = \{1, \tau\}$ , и потому  $\pi = \tau$ . Предположим, что  $H$  имеет два или более элементов второго порядка и что  $\pi \neq 1$ . Тогда  $|K| \geq 3$ . Пусть  $\tau \in K$ ,  $\tau \neq 1$ ,  $\tau \neq \pi$ . Тогда  $\{1, \pi, \tau, \pi\tau\}$  — подгруппа группы  $K$ . Следовательно,  $|K| \equiv 0 \pmod{4}$ . Все элементы группы  $K$  распадаются на пары  $\{x, \pi x\}$ , причем  $x \cdot \pi x = \pi$ . Следовательно, произведение всех элементов группы  $K$ , т. е.  $\pi$ , равно  $\pi^{|K|/2} = 1$ , так как  $|K|/2$  четно.

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 9.7. Пусть  $F = H \cup \{0\}$ , пусть  $\pi$  — произведение всех элементов группы  $H$  и пусть  $C$  —  $H$ -матрица. Пусть  $a, b \in F$ ,  $a \neq b$ . Из первых двух свойств  $H$ -матриц следует, что

$$\prod_{x \in F, x \neq a, x \neq b} C(a, x) = \pi(C(a, b))^{-1}, \quad \prod_{x \in F, x \neq a, x \neq b} C(b, x) = \pi(C(b, a))^{-1}.$$

Из третьего свойства  $H$ -матриц теперь следует, что

$$\pi(C(a, b))^{-1}(\pi(C(b, a))^{-1})^{-1} = \pi$$

и потому  $C(b, a) = \pi C(a, b)$ .

Если группа  $H$  не имеет элементов второго порядка, то  $\pi = 1$ , так что матрица  $C$  симметрична. Если  $H$  имеет ровно один элемент второго порядка, то матрица  $C$  кососимметрична.

Предположим, что  $H$  имеет два или более элементов второго порядка. Тогда  $\pi = 1$ , так что матрица  $C$  должна быть симметричной. Следовательно, матрица  $C$  имеет выше диагонали и ниже диагонали одно и то же число элементов, равных 1. Так как на диагонали единиц нет, матрица  $C$  содержит четное число единиц. С другой стороны, так как группа  $H$  содержит элементы второго порядка, порядок группы  $H$  четен, и потому порядок матрицы  $C$  нечетен. По определению матрица  $C$  содержит по одной единице в каждой строке и, следовательно, общее число единиц в  $C$  должно быть нечетным. Полученное противоречие доказывает, что если  $H$  содержит два или более элементов второго порядка, то  $H$ -матриц не существует.

Теорема 9.7 устанавливает ограничение на возможные абелевы группы  $(c, A)$ -гомологий в  $(c, A)$ -транзитивной плоскости.

Свойства  $H$ -матриц могут быть использованы для изучения классов Ленца – Барлотти I.3, I.4. и II.2. Доказательство следующей теоремы мы опускаем.

**ТЕОРЕМА 9.9.** Пусть  $H$  — группа всех  $((0), [0])$ -гомологий  $((0), [0])$ -транзитивной конечной проективной плоскости  $\Pi$ ,  $F = H \cup \{0\}$ ,  $C$  — соответствующая  $H$ -матрица.

(а) Плоскость  $\Pi$   $((0), [0], [\infty])$ -транзитивна в том и только в том случае, если  $C(ax, ay) = aC(x, y)$  для любых  $a, x, y \in F$ .

(б) Плоскость  $\Pi$   $([\infty], [0], [0])$ -транзитивна в том и только в том случае, если  $C(xa, ya) = C(x, y)a$  для любых  $a, x, y \in F$ .

(в) Плоскость  $\Pi$   $((0), [0], [\infty])$ - и  $([\infty], [0], [0])$ -транзитивна в том и только в том случае, если группа  $H$  абелева и  $(x + y)z = xz + yz$  для любых  $x, y, z \in F$ .

(г) Плоскость  $\Pi$   $([\infty], [\infty])$ -транзитивна в том и только в том случае, если  $F$  — группа относительно сложения. Если это условие выполняется, то  $C(x + z, y + z) = C(x, y)$  для любых  $x, y, z \in F$ .

Заметим, что в случае (в) для дезарговости плоскости  $\Pi$  недостает «всего лишь» ассоциативности сложения (при наличии которой коммутативность сложения может быть выведена).

## §10. ПОРЯДОК ПРОЕКТИВНОЙ ПЛОСКОСТИ

В этом параграфе мы приведем примеры того, как информация о гомологиях проективной плоскости может привести к доказательству гипотезы 1 (см. §8). Существенную роль при этом будут играть  $H$ -матрицы.

**ТЕОРЕМА 10.1.** Пусть  $\Pi$  — проективная плоскость порядка  $n$ , которая при надлежащей координатизации  $((0), [0])$ - и  $([\infty], [\infty])$ -транзитивна. Пусть  $H$  — группа всех  $((0), [0])$ -гомологий. Если группа  $H$  абелева и сложение в координатном кольце  $F = H \cup \{0\}$  коммутативно, то  $n$  — простое число или степень простого числа.

**ДОКАЗАТЕЛЬСТВО.** Пусть  $C$  —  $H$ -матрица, соответствующая данной координатизации плоскости  $\Pi$ . Пусть  $K$  — мультипликативная группа, изоморфная аддитивной группе координатного кольца  $F$ . Удобно считать, что строки и столбцы матрицы  $C$  индексированы элементами группы  $K$ , а не множества  $F$ . В частности, первый столбец и первая строка имеют индекс 1. Тогда  $C(\alpha, 1) = \alpha$  и, в силу теоремы 9.9(д),  $C(\alpha\gamma, \beta\gamma) = C(\alpha, \beta)$  для любых  $\alpha, \beta, \gamma \in K$ .

*Случай 1:*  $n$  — четное число.

В этом случае группа  $H$  порядка  $n - 1$  не имеет элементов порядка 2, так что матрица  $C$  симметрична. Следовательно, для любого  $\alpha \in K$

$$C(\alpha, 1) = C(1, \alpha) = C(\alpha^{-1}, \alpha\alpha^{-1}) = C(\alpha^{-1}, 1).$$

Так как никакой столбец матрицы  $C$  не содержит двух равных элементов, то  $\alpha = \alpha^{-1}$ , т. е.  $\alpha^2 = 1$ . Так как это верно для любого элемента группы  $K$ , мы получаем, что  $n = |K|$  является степенью числа 2.

Заметим, что коммутативность сложения в  $F$  не играла в этом случае никакой роли.

*Случай 2:  $n$  — нечетное число.*

Этот случай значительно сложнее, и нам понадобится понятие *группового кольца*.

Пусть  $S$  — коммутативное ассоциативное кольцо с единицей и пусть  $G$  — конечная мультипликативная группа. Групповое кольцо  $SG$  группы  $G$  над кольцом  $S$  состоит из формальных линейных комбинаций  $\sum_{x \in G} \alpha_x x$ , где все коэффициенты  $\alpha_x$  принадлежат кольцу  $S$ . Сложение и умножение в  $SG$  определяются следующим образом:

$$\sum_{x \in G} \alpha_x x + \sum_{x \in G} \beta_x x = \sum_{x \in G} (\alpha_x + \beta_x) x, \quad \sum_{x \in G} \alpha_x x \cdot \sum_{y \in G} \beta_y y = \sum_{z \in G} \left( \sum_{xy=z} \alpha_x \beta_y \right) z.$$

Относительно этих операций  $SG$  является ассоциативным кольцом. Мы отождествляем элемент  $\beta$  кольца  $S$  с элементом  $\sum_{x \in G} \alpha_x x$  кольца  $SG$ , где  $\alpha_x = \beta$ , если  $x = 1$ , и  $\alpha_x = 0$ , если  $x \neq 1$ . Кроме того, мы отождествляем элемент  $y$  группы  $G$  с элементом  $\sum_{x \in G} \alpha_x x$  кольца  $SG$ , где  $\alpha_x = 1$ , если  $x = y$ , и  $\alpha_x = 0$ , если  $x \neq y$ . При таком отождествлении единица кольца  $S$ , единица группы  $G$  и единица кольца  $SG$  — один и тот же элемент, обозначаемый 1.

Вернемся к нашему доказательству. Пусть  $p$  — простой делитель числа  $n$ . Нам нужно доказать, что  $n$  — степень  $p$ .

Пусть  $S$  — групповое кольцо группы  $H$  над полем (и, следовательно, кольцом) классов вычетов по модулю  $p$ . Матрица  $C$  может рассматриваться как матрица над кольцом  $S$ . Так как группа  $H$  абелева, кольцо  $S$  коммутативно, так что мы можем рассмотреть групповое кольцо  $SK$  группы  $K$  над кольцом  $S$ . Напомним, что строки и столбцы матрицы  $C$  индексированы элементами группы  $K$  и при этом  $C(\alpha\gamma, \beta\gamma) = C(\alpha, \beta)$  для всех  $\alpha, \beta, \gamma \in K$ . Мы будем называть все матрицы порядка  $n$  над кольцом  $S$ , обладающие таким свойством  *$K$ -инвариантными*.  $K$ -инвариантная матрица  $A$  полностью определяется своим первым столбцом и потому такая матрица однозначно определяется элементом  $\text{supp}(A) = \sum_{\alpha \in K} A(\alpha, 1)\alpha$  кольца  $SK$ . Очевидно,  $\text{supp}(A + B) = \text{supp}(A) + \text{supp}(B)$ . Несколько менее

очевидна формула  $\text{supp}(AB) = \text{supp}(A) \text{supp}(B)$ :

$$\begin{aligned} \text{supp}(A) \text{supp}(B) &= \sum_{\alpha \in K} A(\alpha, 1)\alpha \cdot \sum_{\beta \in K} B(\beta, 1)\beta = \\ &= \sum_{\gamma \in K} \sum_{\alpha\beta=\gamma} A(\alpha, 1)B(\beta, 1)\gamma = \sum_{\gamma \in K} \sum_{\beta \in K} A(\gamma\beta^{-1}, 1)B(\beta, 1)\gamma = \\ &= \sum_{\gamma \in K} \sum_{\beta \in K} A(\gamma, \beta)B(\beta, 1)\gamma = \sum_{\gamma \in K} (AB)(\gamma, 1)\gamma = \text{supp}(AB). \end{aligned}$$

Обозначим чрез  $C^*$   $K$ -инвариантную матрицу, определяемую равенствами  $C^*(\alpha, \beta) = 0$ , если  $\alpha = \beta$ ;  $C^*(\alpha, \beta) = (C(\beta, \alpha))^{-1}$ , если  $\alpha \neq \beta$ . Кроме того, обозначим через  $h$  сумму всех элементов группы  $H$  в кольце  $S$ , через  $I$  — единичную матрицу порядка  $n$  и через  $J$  — матрицу порядка  $n$ , все элементы которой равны 1. Каждый диагональный элемент матрицы  $CC^*$  равен сумме  $n - 1$  единиц, а так как  $S$  — кольцо характеристики  $p$ , то все диагональные элементы матрицы  $CC^*$  равны  $-1$ . Каждый внедиагональный элемент матрицы  $CC^*$  равен  $h - 1$ . Поэтому

$$CC^* = -I + (h - 1)(J - I) = -hI + (h - 1)J.$$

Кроме того,  $J^2 = nJ = O$ ,  $CJ = JC = C^*J = JC^* = hJ$  и  $hC = Ch = hC^* = C^*h = h(J - I)$ . Используя эти равенства, нетрудно показать индукцией по  $m$ , что для любого натурального числа  $m$

$$C^m(CC^*) = (-1)^{m-1}h(I - (m + 2)J).$$

Положив  $m = p - 1$ , мы получаем

$$C^pC^* = h(J - I). \quad (7)$$

В коммутативном кольце  $SK$  характеристики  $p$  выполняется равенство  $(x + y)^p = x^p + y^p$ . Поэтому

$$\text{supp}(C^p) = (\text{supp}(C))^p = \sum_{\alpha \in K} (C(\alpha, 1))^p \alpha^p = \sum_{\alpha \in K, \alpha \neq 1} (C(\alpha, 1))^p \alpha^p.$$

Пусть  $D = C^pC^*$ . Тогда

$$\text{supp}(D) = \sum_{\alpha \in K, \alpha \neq 1} (C(\alpha, 1))^p \alpha^p \cdot \sum_{\beta \in K, \beta \neq 1} C^*(\beta, 1)\beta = \sum_{\gamma \in K} x_\gamma \gamma,$$

где  $x_\gamma = D(\gamma, 1)$ . Уравнение (7) показывает, что  $D(1, 1) = 0$ . Следовательно,

$$0 = \gamma_1 = \sum_{\beta \neq 1, \alpha^p = \beta^{-1}} (C(\alpha, 1))^p C^*(\beta, 1).$$

Равенство этой суммы нулю означает, что никакой элемент группы  $K$ , кроме 1, не является  $p$ -й степенью. Иными словами,  $\alpha^p = 1$  для любого

$\alpha \in K$ . Но тогда  $n = |K|$  — степень простого числа  $p$ , что и требовалось доказать.

## СПИСОК ЛИТЕРАТУРЫ

Наиболее полную информацию о конечных проективных плоскостях и других конечных геометриях можно найти в книгах [2] и [4]. О теоремах Дезарга и Паппа и связанных с ними алгебраических структурах см. [1] и [3]. Книги [1] и [7] особенно хороши для первого знакомства с предметом.  $H$ -матрицы детально изучаются в книге [5] (там  $H$ -матрица фигурирует как *the core of a normalized generalized conference matrix*).

- [1] Артин Э. *Геометрическая алгебра*. М.: Наука, 1969.
- [2] Dembowski P. *Finite Geometries*. Springer, 1968.
- [3] Гильберт Д. *Основания геометрии*. М.: Гостехиздат, 1948.
- [4] Hughes D. R., Piper F. C. *Projective Planes*. Springer, 1982.
- [5] Ionin Y. J., Shrikhande M. S. *Combinatorics of Symmetric Designs*. Cambridge University Press, 2006.
- [6] Lam C. W. H. *The search for a finite projective plane of order 10* // The American Mathematical Monthly, 1991, pp. 305–318.
- [7] Room T. G., Kirkpatrick P. B. *Mini quaternion Geometry*. Cambridge University Press, 1971.

## Метод Ньютона и его приложения к решению уравнений и теории экстремума

Г. Г. Магарил-Ильяев      В. М. Тихомиров

При изучении законов природы, в инженерных расчетах, при решении разнообразных задач управления и экономики, необходимо исследовать и решать *уравнения*, которые описывают изучаемые процессы или явления. Здесь будет рассказано об одном универсальном методе решения уравнений, восходящем к Ньютону (1643 – 1727). Мы пройдем путь, идущий от древних (в частных случаях они пользовались методом, позже описанным Ньютоном) до середины двадцатого века.

### 1. Истоки: МЕТОД РЕШЕНИЯ ОДНОГО НЕЛИНЕЙНОГО УРАВНЕНИЯ С ОДНИМ НЕИЗВЕСТНЫМ

Задача 1.1. *Решить уравнение  $x^2 = 2$ .*

«А чего его решать, — может удивиться читатель, — его решение это ведь  $\pm\sqrt{2}$ , не так ли?» Но  $\sqrt{2}$  — это просто символ, смысл которого надо еще расшифровать. Если спросить чему равен корень из двух, кто-то может сказать: «Корень из двух равен 1.41...». Это приближение  $\sqrt{2}$  десятичной дробью с точностью до сотых. Еще в шестом веке до н.э. пифагорейцы установили, что  $\sqrt{2}$  не дробь, и это стало одним из крупнейших завоеваний древней математики. Решить уравнение  $x^2 = 2$  — значит указать способ (говорят еще *алгоритм*) приближения дробями к числу, квадрат которого равен двум, с любой точностью.

Считается, что первый алгоритм для решения уравнения  $x^2 = 2$ , принадлежит древнегреческому математику Герону, жившему в первом веке нашей эры. Алгоритм Герона описывается так: надо выбрать любую дробь  $x_0$  и затем использовать такую итерационную процедуру  $x_k = \frac{1}{2}(x_{k-1} + \frac{2}{x_{k-1}})$ ,  $k = 1, 2, \dots$ . Если начать с  $x_0 = 1$ , то будем последовательно получать  $x_1 = \frac{3}{2} = 1.5$ ,  $x_2 = \frac{17}{12} = 1.41(6)$ , и если продолжать дальше, то числа  $x_k$  будут со все большей и большей точностью стремиться к числу, квадрат которого равен двум.



Когда говорится о решении уравнения с одним неизвестным, то обычно требуется указать целое число или дробь, являющуюся точным решением уравнения, либо, если это не так, то требуется указать алгоритм приближения к искомому числу дробями с любой степенью точности.

В этом пункте мы будем решать уравнения вида  $F(x) = y$ , где  $x$  и  $y$  — вещественные числа (совокупность которых обозначают  $\mathbb{R}$ ), а  $F$  — функция, определенная, скажем, на некотором интервале  $(a, b)$ , т. е.  $F$  каждому числу  $x \in (a, b)$  ставит в соответствие число  $F(x)$ , что коротко записывают так:  $x \mapsto F(x)$ ,  $x \in (a, b)$  или  $F : (a, b) \rightarrow \mathbb{R}$ . Решить уравнение  $F(x) = y$  — это значит для данного  $y$  найти такое  $x \in (a, b)$ , при котором уравнение становится верным равенством.

Простейшая функция одного переменного — линейная:  $F(x) = Ax$ ,  $A \neq 0$ , в нее переменное входит в первой степени. В этом случае решение уравнения  $Ax = y$  для любого  $y$ , очевидно, дается выражением  $x = A^{-1}y$ . В уравнение  $x^2 = 2$  переменное входит во второй степени. Это уравнение нелинейное. Наша цель — научиться приближенно решать нелинейные уравнения.

Метод решения уравнений  $F(x) = y$ , о котором сейчас будет рассказано, основан на линеаризации, т. е. на решении по ходу дела линейных уравнений.

Пусть дано число  $y$ . Возьмем какое-нибудь число  $x_0$  (так, чтобы  $F(x_0)$  было, по возможности, близко к  $y$ ) и попробуем «добраться» до  $y$  линейно. Точнее говоря, выберем отличное от нуля число  $A$  и в качестве приближенного решения уравнения  $F(x) = y$  возьмем решение линейного уравнения  $A(x - x_0) + F(x_0) = y$ , которое обозначим  $x_1$  и которое, очевидно, имеет вид  $x_1 = x_0 + A^{-1}(y - F(x_0))$ . Геометрически (см. рис. 1)  $x_1$  — абсцисса точки пересечения горизонтальной прямой на уровне  $y$  с прямой, проходящей через точку  $(x_0, F(x_0))$  с угловым коэффициентом  $A$ . Теперь вместо  $x_0$  возьмем точку  $x_1$  и поступая аналогично, найдем точку  $x_2$ . Продолжая

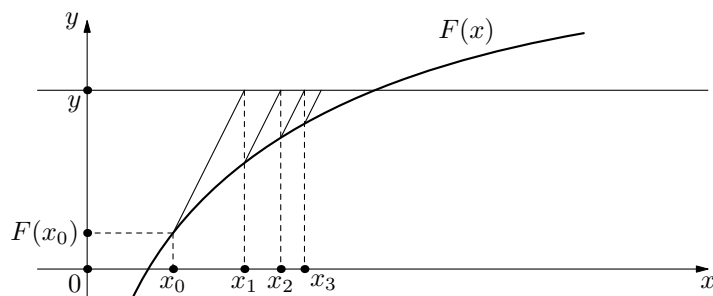


Рис. 1.

этот процесс, получим последовательность

$$x_k = x_{k-1} + A^{-1}(y - F(x_{k-1})), \quad k = 1, 2, \dots, \quad (1.1)$$

которую назовем *модифицированным методом Ньютона* (решения уравнения  $F(x) = y$ ).

**УПРАЖНЕНИЕ 1.** *Сделайте три шага модифицированного метода Ньютона (1.1) для решения уравнения  $x^2 = 2$ , стартуя от точки  $x_0 = 1$  и выбрав  $A = 2$ . Эти вычисления естественно делать с помощью калькулятора. Сравните полученные результаты с числом  $1.414213562\dots$ , представляющим корень из двух с точностью до одной миллиардной (это число появится на калькуляторе, если вы нажмете на  $\sqrt{2}$ ).*

**УПРАЖНЕНИЕ 2.** *Попробуйте определить границы для числа  $A$ , при которых последовательность (1.1) модифицированного метода Ньютона, стартуя от точки  $x_0 = 1$ , будет стремиться к  $\sqrt{2}$ .*

**ИСТОРИЧЕСКИЙ КОММЕНТАРИЙ.** Секретарь Королевского общества Ольденбург в 1676 году обратился к Ньютону за некоторыми разъяснениями, в частности, касательно решения уравнений. Ньютон в ответном письме от 24.10.1676 г., на примере решения уравнения  $F(x) = x^3 - 2x - 5 = 0$ , изложил метод нахождения корня, который в современных обозначениях представляет собой такую итерационную процедуру

$$x_k = x_{k-1} - (F'(x_{k-1}))^{-1}F(x_{k-1}), \quad k = 1, 2, \dots, \quad (1.2)$$

где  $F'$  — это производная  $F$ , а начальную точку  $x_0$  предоставляется выбрать вычислителю, исходя из каких-то соображений целесообразности. Этот метод (его называют *методом Ньютона*) был изложен Ньютоном в сочинении «Анализ с помощью уравнений . . .» (И. Ньютон. Математические работы. — ТТЛ, М. — Л., 1937, с. 9).

Помимо изложенного метода решения уравнений, Ньютон в те годы пользовался также методом приближения функций рядами. Он был не только создателем математического естествознания и одним из родоначальников математического анализа, но и замечательным вычислителем. Он составил для себя таблицы основных, как мы сейчас говорим «элементарных», функций. Он писал Ольденбургу: «Мне стыдно признаться, с какой точностью я проводил свои вычисления».

Производная  $F'(x)$  функции  $F(x) = x^2$  равняется  $2x$ , так что для корня из двух метод Ньютона приводит к такой последовательности:  $x_k = x_{k-1} + \frac{1}{2x_{k-1}}(2 - (x_{k-1})^2) = \frac{1}{2}(x_{k-1} + \frac{2}{x_{k-1}})$ . Именно эта последовательность встречается в сочинениях Герона.

**УПРАЖНЕНИЕ 3.** *Сделайте три шага метода Ньютона (1.2) для решения уравнения  $x^2 = 2$ , стартуя от точки  $x_0 = 1$ , и сравните*

полученный результат с числом 1.414213562, дающим, как уже говорилось, приближение  $\sqrt{2}$  с точностью до одной миллиардной.

Сформулируем основной результат этого пункта. Отметим, что ниже мы пишем  $A^{-1}$  и  $|A^{-1}|$  вместо, скажем,  $1/A$  и  $1/|A|$ . Это связано с тем, что при обобщении данного результата  $A$  уже не будет числом, но форма записи останется прежней. Кроме того, чтобы сохранить единообразие обозначений при дальнейших обобщениях, интервал с центром в точке  $\hat{x}$  радиуса  $\delta > 0$  обозначаем  $U_{\mathbb{R}}(\hat{x}, \delta)$ , т. е.  $U_{\mathbb{R}}(\hat{x}, \delta) = \{x \mid |x - \hat{x}| < \delta\}$ .

**ТЕОРЕМА 1 (ОБ ОБРАТНОЙ ФУНКЦИИ ДЛЯ ФУНКЦИЙ ОДНОГО ПЕРЕМЕННОГО).** Пусть заданы числа  $x_0, \delta > 0, A \neq 0$  и функция  $F: U_{\mathbb{R}}(x_0, \delta) \rightarrow \mathbb{R}$ . Если существует такое число  $\theta \in (0, 1)$ , что для любых  $\xi, x \in U_{\mathbb{R}}(x_0, \delta)$  выполняется неравенство

$$|F(\xi) - F(x) - A(\xi - x)| \leq \frac{\theta}{|A^{-1}|} |\xi - x|, \quad (1.3)$$

то для каждого  $y \in U_{\mathbb{R}}(F(x_0), \delta_0)$ , где  $\delta_0 = \delta(1 - \theta)/|A^{-1}|$ , последовательность (1.1) сходится к такому числу  $\varphi(y) \in U_{\mathbb{R}}(x_0, \delta)$ , что  $F(\varphi(y)) = y$  и при этом  $|\varphi(y) - x_0| \leq K|y - F(x_0)|$ , где  $K = |A^{-1}|/(1 - \theta)$ .

Функция  $\varphi: U_{\mathbb{R}}(F(x_0), \delta_0) \rightarrow U_{\mathbb{R}}(x_0, \delta)$ , построенная в этой теореме, называется *обратной функцией* к  $F$ .

Доказательство опирается на понятия непрерывности функции одного переменного, сходимости последовательности чисел и фундаментальной последовательности чисел. Приведем соответствующие определения.

Пусть функция  $F$  определена на интервале  $U$ . Говорят, что  $F$  непрерывна в точке  $x \in U$ , если для любого числа  $\varepsilon > 0$  найдется такое число  $\delta > 0$ , что  $|F(x + h) - F(x)| < \varepsilon$ , если только  $|h| < \delta$ . Функцию  $F$  называют непрерывной на  $U$ , если она непрерывна в каждой точке  $x \in U$ . Говорят, что числовая последовательность  $\{x_1, x_2, \dots, x_k, \dots\}$  сходится к числу  $c$  или имеет пределом число  $c$  (и пишут  $x_k \rightarrow c$  при  $k \rightarrow \infty$  или  $\lim_{k \rightarrow \infty} x_k = c$ ), если для любого числа  $\varepsilon > 0$  найдется такое число  $N$ , что  $|x_k - c| < \varepsilon$ , как только  $k > N$ . Последовательность  $\{x_1, x_2, \dots, x_k, \dots\}$  называется фундаментальной, если для любого числа  $\varepsilon > 0$  найдется такое число  $N$ , что  $|x_k - x_m| < \varepsilon$ , как только  $k, m > N$ .

В доказательстве будут использованы следующие факты: а) если функция  $F$  определена на интервале  $(c - a, c + a)$  и непрерывна в точке  $c$ , а последовательность  $\{x_1, x_2, \dots, x_n, \dots\}$  сходится к  $c$ , то последовательность  $\{F(x_1), F(x_2), \dots, F(x_n), \dots\}$  сходится к числу  $F(c)$  (этот факт легко следует из самого определения непрерывности  $F$  в точке  $c$ ); б) всякая фундаментальная последовательность сходится (это одно из выражений свойства полноты вещественных чисел); в) предел суммы последовательностей равен сумме пределов и предел произведения последовательности на число равен произведению этого числа на предел последовательности (если  $x_k \rightarrow c, x'_k \rightarrow c'$  при  $k \rightarrow \infty$  и  $a$  — число, то  $x_k + x'_k \rightarrow c + c'$

при  $k \rightarrow \infty$ , а  $ax_k \rightarrow ac$ ; проверка этих свойств совсем проста); d) если  $0 < \theta < 1$ , то  $1 + \theta + \dots + \theta^k = \frac{1 - \theta^{k+1}}{1 - \theta}$  (формула для суммы геометрической прогрессии).

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1. Пусть  $\{x_1, \dots, x_k\}$  — числа, построенные модифицированным методом Ньютона, и при этом все они принадлежат интервалу  $U_{\mathbb{R}}(x_0, \delta)$ . По определению  $x_k$  удовлетворяет уравнению

$$A(x_k - x_{k-1}) + F(x_{k-1}) = y \quad (1.1')$$

и тогда

$$\begin{aligned} |x_{k+1} - x_k| &= \\ &= |A^{-1}(y - F(x_k))| = |A^{-1}||y - F(x_k) - y + F(x_{k-1}) + A(x_k - x_{k-1})| \leq \\ &\leq \theta|x_k - x_{k-1}|. \end{aligned} \quad (1.4)$$

Первое равенство следует из (1.1), второе — из (1.1'), неравенство — следствие (1.3). Тогда  $|x_{k+1} - x_0| \leq |x_{k+1} - x_k| + |x_k - x_{k-1}| + \dots + |x_1 - x_0| \leq (\theta^k + \theta^{k-1} + \dots + 1)|x_1 - x_0| < \frac{|A^{-1}|}{1 - \theta}|y - F(x_0)| < \delta$  (первое неравенство — неравенство для модулей, второе является следствием неравенств типа (1.4) для  $s \leq k$ , третье следует из (1.1) при  $k = 1$  и формулы d) для суммы геометрической прогрессии, последнее — следствие определения  $\delta_0$ ). Таким образом, в силу того, что  $x_0$  принадлежит  $U_{\mathbb{R}}(x_0, \delta)$  и из того, что  $x_s$  принадлежат  $U_{\mathbb{R}}(x_0, \delta)$  при  $0 \leq s \leq k$  следует, что  $x_{k+1}$  принадлежит  $U_{\mathbb{R}}(x_0, \delta)$ . Метод математической индукции позволяет сделать вывод, что все  $x_k$ ,  $k = 0, 1, 2, \dots$ , принадлежат  $U_{\mathbb{R}}(x_0, \delta)$ . Далее, используя те же соображения, что и при доказательстве предыдущего неравенства, получаем для любых  $k$  и  $s$

$$\begin{aligned} |x_{k+s} - x_k| &\leq \\ &\leq |x_{k+s} - x_{k+s-1}| + \dots + |x_{k+1} - x_k| \leq (\theta^{k+s-1} + \dots + \theta^k) < \\ &< \frac{|A^{-1}|\theta^k}{1 - \theta}|y - F(x_0)|. \end{aligned} \quad (1.5)$$

Отсюда следует, что последовательность  $\{x_0, x_1, \dots\}$  фундаментальна, и значит, в силу b) сходится к некоторому числу, которое обозначим  $\varphi(y)$ . Переходя к пределу по  $s$  в соотношении (1.5) при  $k = 0$  и учитывая соотношение  $|y - F(x_0)| < \delta_0$ , получаем, что  $\varphi(y) \in U_{\mathbb{R}}(x_0, \delta)$ . Переходя (используя свойства предела из с) к пределу в равенстве (1.1) (или (1.1')) и учитывая, что из (1.3) следует непрерывность  $F$  в  $U_{\mathbb{R}}(x_0, \delta)$  (действительно, в силу (1.3) для любых  $\xi, x \in U_{\mathbb{R}}(x_0, \delta)$ ), будем иметь  $|F(\xi) - F(x)| = |F(\xi) - F(x) - A(\xi - x) + A(\xi - x)| \leq ((\theta/|A^{-1}|)|\xi - x| + |A||\xi - x|) = (\theta/|A^{-1}| + |A|)|\xi - x|$ , получаем, что  $F(\varphi(y)) = y$ . Снова переходя

к пределу по  $s$  в неравенстве (1.5) при  $k = 0$ , приходим к неравенству  $|\varphi(y) - x_0| \leq K|y - F(x_0)|$ , где  $K = |A^{-1}|/(1 - \theta)$ .

## 2. МЕТОД РЕШЕНИЯ ДВУХ УРАВНЕНИЙ С ДВУМЯ НЕИЗВЕСТНЫМИ

В этом пункте мы сделаем один шаг от единицы к бесконечности, а именно, будем искать решение системы двух уравнений с двумя неизвестными:

$$\begin{cases} F_1(x_1, x_2) = y_1, \\ F_2(x_1, x_2) = y_2, \end{cases}$$

т. е. для данных чисел  $y_1$  и  $y_2$  требуется найти числа  $x_1$  и  $x_2$ , для которых эти уравнения становятся верными равенствами.

Пару чисел  $(z_1, z_2)$  будем трактовать как вектор на плоскости (идущий из начала координат в точку с координатами  $(z_1, z_2)$ ) и записывать его как столбец  $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ . Векторы можно складывать:  $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} z'_1 \\ z'_2 \end{pmatrix} = \begin{pmatrix} z_1 + z'_1 \\ z_2 + z'_2 \end{pmatrix}$  и умножать на числа:  $\alpha \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} \alpha z_1 \\ \alpha z_2 \end{pmatrix}$ . Множество всех таких векторов обозначим  $\mathbb{R}^2$ .

Модулем или длиной вектора  $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$  называется число  $|z| = \sqrt{z_1^2 + z_2^2}$  (согласно теореме Пифагора, это, действительно, длина вектора). Модуль вектора обладает следующими свойствами: (а)  $|x| \geq 0$  и  $|x| = 0$  тогда и только тогда  $x = 0$ , (б)  $|\alpha x| = |\alpha| |x|$  для любого  $\alpha \in \mathbb{R}$ , (с)  $|x+x'| \leq |x| + |x'|$  для любых  $x, x' \in \mathbb{R}^2$ . Первые два свойства очевидны. Последнее следует из известного неравенства Коши – Буняковского для векторов  $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  и  $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ :  $(a_1 b_1 + a_2 b_2) \leq \sqrt{a_1^2 + a_2^2} \sqrt{b_1^2 + b_2^2}$  (которое сразу получается после возведения в квадрат и раскрытия скобок). Действительно, если обе части этого неравенства умножить на два и добавить слева и справа число  $a_1^2 + a_2^2 + b_1^2 + b_2^2$ , то придем к соотношению  $|a + b|^2 \leq (|a| + |b|)^2$ , равносильному неравенству треугольника.

Пусть  $\hat{z} \in \mathbb{R}^2$  и  $\delta > 0$ . Обозначим  $U_{\mathbb{R}^2}(\hat{z}, \delta) = \{z \in \mathbb{R}^2 \mid |z - \hat{z}| < \delta\}$ . Ясно, что это круг без границы с центром в точке  $\hat{z}$  радиуса  $\delta$ .

Предположим, что функции  $F_1$  и  $F_2$  определены в круге  $U_{\mathbb{R}^2}(x_0, \delta)$ . Каждому вектору  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in U_{\mathbb{R}^2}(x_0, \delta)$  сопоставим вектор  $F(x) = \begin{pmatrix} F_1(x) \\ F_2(x) \end{pmatrix}$  (где  $F_i(x) = F_i(x_1, x_2)$ ,  $i = 1, 2$ ). Тогда, как говорят,  $F$  есть отображение из  $U_{\mathbb{R}^2}(x_0, \delta)$  в  $\mathbb{R}^2$  и кратко пишут  $x \mapsto F(x)$ ,  $x \in U_{\mathbb{R}^2}(x_0, \delta)$  или  $F: U_{\mathbb{R}^2}(x_0, \delta) \rightarrow \mathbb{R}^2$ . Написанную выше систему двух уравнений

теперь можно записать так:  $F(x) = y$ , где  $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ . Эта запись, как мы видим, внешне не отличается от одномерной ситуации.

Важную роль среди систем уравнений занимают так называемые *линейные системы*, т. е. системы вида

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = y_1, \\ a_{21}x_1 + a_{22}x_2 = y_2, \end{cases}$$

где  $a_{ij}$ ,  $1 \leq i, j \leq 2$  — заданные числа.

Свяжем с этой системой следующую таблицу (или говорят, матрицу)

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

(при этом числа  $a_{ij}$ ,  $1 \leq i, j \leq 2$  называются элементами матрицы  $A$ ). Для каждого  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$  положим

$$Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}.$$

Этим определено отображение  $A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  (которое мы обозначаем той же буквой  $A$ ), сопоставляющее вектору  $x$  вектор  $Ax$ . Легко проверить, что для любых  $x, x' \in \mathbb{R}^2$  и  $\alpha, \beta \in \mathbb{R}$  справедливо равенство  $A(\alpha x + \beta x') = \alpha Ax + \beta Ax'$ , где  $\alpha A$  и  $\beta A$  — матрицы, получаемые из матрицы  $A$  умножением всех ее элементов соответственно на  $\alpha$  и  $\beta$ . Отображения, обладающие таким свойством, называются *линейными отображениями* или *линейными операторами*.

В терминах матрицы (отображения)  $A$  исходная линейная система записывается в виде  $Ax = y$ , где  $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ . Такую запись называют матричной формой записи линейной системы и внешне она не отличается от одномерной ситуации.

Число  $\det A = a_{11}a_{22} - a_{12}a_{21}$  называется *определителем* или *детерминантом* матрицы  $A$ . Если  $\det A \neq 0$ , то для любого  $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$  методом подстановки легко находится решение линейной системы, оно имеет вид

$$\begin{cases} x_1 = \frac{a_{11}}{\det A}y_1 - \frac{a_{21}}{\det A}y_2, \\ x_2 = -\frac{a_{12}}{\det A}y_1 + \frac{a_{22}}{\det A}y_2. \end{cases}$$

В матричной форме это можно записать как  $x = By$ , где  $B$  — соответствующая матрица. Ее называют *обратной матрицей* к  $A$  и обозначают  $A^{-1}$  (и говорят, что матрица  $A$  обратима). Итак, если матрица  $A$  такова, что

$\det A \neq 0$ , то она обратима и решение линейной системы  $Ax = y$  для любого  $y$  находится по формуле  $x = A^{-1}y$  (что, снова, внешне не отличается от одномерного случая).

Наша цель — научиться находить решение системы двух нелинейных уравнений с двумя неизвестными. Для этого, как мы увидим, нужно, фактически, дословно повторить то, что было сказано про одно уравнение с одним неизвестным, заменяя лишь числа  $x_k$  векторами, модуль числа — длиной вектора, число  $A$  — матрицей  $A$ , а число  $A^{-1}$  — обратной матрицей к  $A$ . Но, помимо этого, в одномерном случае еще участвовало число  $|A^{-1}|$ . В двумерной ситуации — это «длина» (или обычно говорят *норма*) матрицы  $A^{-1}$ . Для каждой матрицы  $C$  с элементами  $c_{ij}$ ,  $1 \leq i, j \leq 2$ , ее норма (обозначаемая  $\|C\|$ ) определяется следующим образом:  $\|C\| = \sqrt{c_{11}^2 + c_{12}^2 + c_{21}^2 + c_{22}^2}$ . Снова, используя неравенство Коши — Буняковского, легко проверить, что  $|Cx| \leq \|C\|\|x\|$  для любого  $x \in \mathbb{R}^2$ .

Опишем теперь метод решения двух уравнений с двумя неизвестными, которые мы записали в виде  $F(x) = y$ . Пусть дан вектор  $y$ . Возьмем вектор  $x_0$  (снова, так, чтобы  $F(x_0)$  было, по возможности, близко к  $y$ ) и обратимую матрицу  $A$ . В качестве приближенного решения уравнения  $F(x) = y$  возьмем решение линейного уравнения  $A(x - x_0) + F(x_0) = y$ , которое обозначим  $x_1$  и которое, очевидно, имеет вид  $x_1 = x_0 + A^{-1}(y - F(x_0))$ . Теперь вместо  $x_0$  возьмем точку  $x_1$  и, поступая аналогично, найдем точку  $x_2$ . Продолжая этот процесс, получим последовательность

$$x_k = x_{k-1} + A^{-1}(y - F(x_{k-1})), \quad k = 1, 2, \dots \quad (2.1)$$

Эту процедуру мы также называем *модифицированным методом Ньютона*.

Сформулируем основной результат этого пункта.

**ТЕОРЕМА 2 (ОБ ОБРАТНОМ ОТОБРАЖЕНИИ ДЛЯ СЛУЧАЯ ДВУХ ПЕРЕМЕННЫХ).** Пусть заданы вектор  $x_0$ , число  $\delta > 0$ , обратимая матрица  $A$  и отображение  $F: U_{\mathbb{R}^2}(x_0, \delta) \rightarrow \mathbb{R}^2$ . Если существует такое число  $\theta \in (0, 1)$ , что для любых  $\xi, x \in U_{\mathbb{R}^2}(x_0, \delta)$  выполняется неравенство

$$|F(\xi) - F(x) - A(\xi - x)| \leq \frac{\theta}{\|A^{-1}\|} |\xi - x|, \quad (2.2)$$

то для каждого  $y \in U_{\mathbb{R}^2}(F(x_0), \delta_0)$ , где  $\delta_0 = \delta(1 - \theta)/\|A^{-1}\|$ , последовательность (2.1) сходится<sup>1)</sup> к такому вектору  $\varphi(y) \in U_{\mathbb{R}^2}(x_0, \delta)$ , что  $F(\varphi(y)) = y$  и при этом  $|\varphi(y) - x_0| \leq K|y - F(x_0)|$ , где  $K = \|A^{-1}\|/(1 - \theta)$ .

Отображение  $\varphi: U_{\mathbb{R}^2}(F(x_0), \delta_0) \rightarrow U_{\mathbb{R}^2}(x_0, \delta)$ , построенное в этой теореме, называется *обратным отображением к  $F$* .

<sup>1)</sup>Определение сходимости векторов будет дано чуть ниже.

Доказательство теоремы 2 опирается на те же понятия, что и теорема 1, но они касаются уже непрерывности функции не одного, а двух переменных, сходимости последовательности не чисел, а векторов и понятия фундаментальной последовательности не чисел, а векторов. Определения всех этих понятий требуют только замены слов «число» на «вектор из  $\mathbb{R}^2$ » и «функция» на «отображение». Но, тем не менее, повторим эти определения.

Говорят, что отображение  $F: U_{\mathbb{R}^2}(x_0, r) \rightarrow \mathbb{R}^2$  непрерывно в точке  $x \in U_{\mathbb{R}^2}(x_0, r)$ , если для любого  $\varepsilon > 0$  найдется такое  $\delta > 0$ , что  $|F(x+h) - F(x)| < \varepsilon$ , как только  $|h| < \delta$ . Отображение  $F$  называют непрерывным на  $U_{\mathbb{R}^2}(x_0, r)$ , если оно непрерывно в каждой точке  $x \in U_{\mathbb{R}^2}(x_0, a)$ . Говорят, что последовательность векторов  $\{x_1, x_2, \dots, x_k, \dots\}$  из  $\mathbb{R}^2$  сходится к вектору  $c$  или имеет пределом вектор  $c$  (и пишут  $x_k \rightarrow c$  при  $k \rightarrow \infty$  или  $\lim_{k \rightarrow \infty} x_k = c$ ), если для любого  $\varepsilon > 0$  найдется такое натуральное  $N = N(\varepsilon)$ , что  $|x_k - c| < \varepsilon$ , как только  $k > N$ . Последовательность  $\{x_1, x_2, \dots, x_k, \dots\}$  векторов из  $\mathbb{R}^2$  называется фундаментальной, если для любого  $\varepsilon > 0$  найдется такое натуральное  $N = N(\varepsilon)$ , что  $|x_k - x_m| < \varepsilon$  как только  $k, m > N$ .

В доказательстве будут использованы аналоги фактов, описанных перед доказательством теоремы 1 (мы нумеруем их в том же порядке): а) если отображение  $F$ , определенное на  $U_{\mathbb{R}^2}(c, r)$ , непрерывно в точке  $c$ , то отсюда легко следует, что если последовательность  $\{x_1, x_2, \dots, x_k, \dots\}$  векторов из  $\mathbb{R}^2$  сходится к  $c$ , то последовательность  $\{F(x_1), F(x_2), \dots, F(x_k), \dots\}$  сходится к вектору  $F(c)$ ; б) всякая фундаментальная последовательность векторов из  $\mathbb{R}^2$  сходится (это, снова, полнота вещественных чисел — сходимости векторов следует из сходимости координат этих векторов); в) предел суммы последовательностей равен сумме пределов и предел произведения последовательности на число равен произведению этого числа на предел последовательности (т. е., если  $x_k \rightarrow c$  и  $x'_k \rightarrow c'$  при  $k \rightarrow \infty$ , то  $x_k + x'_k \rightarrow c + c'$  при  $k \rightarrow \infty$  и если  $x_k \rightarrow c$  при  $k \rightarrow \infty$  и  $a$  — число, то  $ax_k \rightarrow ac$  при  $k \rightarrow \infty$ ; все эти свойства легко следуют из определения предела); г) как и в одномерном случае, используется формула для суммы геометрической прогрессии.

Доказательство теоремы 2 отличается от доказательства теоремы 1 только в одном месте. Второе равенство в соотношении (1.4) надо заменить на неравенство  $|A^{-1}(y - F(x_k))| \leq \|A^{-1}\| |y - F(x_k)|$ , а далее уже ничего не меняется.

**ИСТОРИЧЕСКИЙ КОММЕНТАРИЙ.** Мы без особых затруднений преодолели дистанцию примерно в двести лет. Теоремы об обратных отображениях для многих переменных стали появляться во второй половине девятнадцатого века. Собственно говоря, все началось с теоремы о неявной функции — решении уравнения  $F(x, y) = 0$ , т. е. нахождении функции  $y = \varphi(x)$ , для которой  $F(x, \varphi(x)) = 0$ . Первая теорема о неявной функции (такого рода теоремы будут нами изучаться в следующем пункте) была доказана Улиссом Дини (1845 – 1918) и прочитана им в курсе анализа в пизанском университете в 1877/78 учебном году. Чуть позже появились теоремы об обратных отображениях.



### 3. МЕТОД РЕШЕНИЯ $m$ УРАВНЕНИЙ С $n$ НЕИЗВЕСТНЫМИ

Сделаем еще один шаг и научимся решать систему  $m$  уравнений с  $n$  неизвестными, т. е. систему уравнений вида

$$\begin{cases} F_1(x_1, \dots, x_n) = y_1, \\ F_2(x_1, \dots, x_n) = y_2, \\ \dots \dots \dots \\ F_m(x_1, \dots, x_n) = y_m. \end{cases}$$

Наш метод будет очень схож с тем, что привело нас к цели в одномерном и двумерном случаях.

Здесь мы имеем дело с  $m$  функциями  $n$  переменных, т. е. с функциями, определенными на подмножествах пространства  $\mathbb{R}^n$ . Напомним некоторые понятия и договоримся об обозначениях. Пространство  $\mathbb{R}^n$  — это совокупность всех наборов из  $n$  действительных чисел, расположенных в столбец (называемых векторами или вектор-столбцами), которые, ради экономии места, мы записываем в строку  $x = (x_1, \dots, x_n)^T$ , где  $T$  означает *транспонирование* — замену столбца на строчку. Векторы можно (покоординатно) складывать и умножать на числа. Пусть  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ . Величина  $|x| = \sqrt{x_1^2 + \dots + x_n^2}$  называется *модулем* или *длиной* вектора  $x$ . Эта величина обладает теми же свойствами, которые были отмечены в двумерном случае. Пусть  $\hat{x} \in \mathbb{R}^n$  и  $\delta > 0$ . Множество  $U_{\mathbb{R}^n}(\hat{x}, \delta) = \{x \mid |x - \hat{x}| < \delta\}$  называется *открытым шаром* с центром в точке  $\hat{x}$  радиуса  $\delta$ . Множество в  $\mathbb{R}^n$  называется *открытым*, если с каждой своей точкой оно содержит и некоторый шар с центром в этой точке. Любое открытое множество, содержащее данную точку, называется *окрестностью* этой точки.

Предположим, что функции  $F_i$ ,  $1 \leq i \leq m$ , определены в шаре  $U_{\mathbb{R}^n}(x_0, \delta)$ . Тогда определено отображение  $F: U_{\mathbb{R}^n}(x_0, \delta) \rightarrow \mathbb{R}^m$ , сопоставляющее  $x \in U_{\mathbb{R}^n}(x_0, \delta)$  вектор  $(F_1(x), \dots, F_m(x))^T$ . Написанную систему уравнений теперь кратко можно записать так:  $F(x) = y$ , где  $y = (y_1, \dots, y_m)^T$ .

Как и в предыдущих случаях, важную роль играют системы линейных уравнений:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = y_1, \\ a_{21}x_1 + \dots + a_{2n}x_n = y_2, \\ \dots \dots \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = y_m. \end{cases}$$

Обозначим через  $A$  матрицу этой системы, т. е.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Тогда данная система кратко запишется так:  $Ax = y$ .<sup>2)</sup> Матрица  $A$  задает отображение (которое мы обозначаем той же буквой)  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , сопоставляющее вектору  $x \in \mathbb{R}^n$  вектор  $Ax \in \mathbb{R}^m$ . Это отображение линейно (см. двумерный случай), и его называют также *линейным оператором* (из  $\mathbb{R}^n$  в  $\mathbb{R}^m$ ).

Если бы матрица  $A$  была обратима, то решение уравнения  $Ax = y$  для любого  $y$  давалось бы формулой:  $x = A^{-1}y$ , где  $A^{-1}$  — обратная матрица к  $A$ , и дальнейшие построения, по сути, ничем бы не отличались от рассмотренных случаев. Но это возможно лишь тогда, когда  $m = n$ , а мы хотим решить систему в более общей ситуации, что важно для различных приложений.

Оказывается, что все рассуждения, связанные с применением модифицированного метода Ньютона, остаются без изменения, если  $A^{-1}$  заменить на отображение, которое является правым обратным к  $A$ . Дадим точные определения. Говорят, что линейное отображение  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$  *сюръективно*, если для любого  $y \in \mathbb{R}^m$  существует такое  $x \in \mathbb{R}^n$ , что  $Ax = y$ . На языке линейных систем это означает, что для любого  $y$  линейная система  $Ax = y$  имеет решение, а это (как хорошо известно из линейной алгебры) равносильно тому, что ранг матрицы  $A$  равен  $m$ .

**ЛЕММА 1 (О ПРАВОМ ОБРАТНОМ ОТОБРАЖЕНИИ В КОНЕЧНОМЕРНОМ СЛУЧАЕ).** Пусть  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$  — линейный сюръективный оператор. Тогда существуют отображение  $R: \mathbb{R}^m \rightarrow \mathbb{R}^n$  (правое обратное к  $A$ ) и константа  $\gamma > 0$  такие, что  $AR(y) = y$  и  $|R(y)| \leq \gamma|y|$  для любого  $y \in \mathbb{R}^m$ .

**ДОКАЗАТЕЛЬСТВО** Рассмотрим систему векторов  $e_1 = (1, 0, \dots, 0)^T, \dots, e_m = (0, \dots, 0, 1)^T$  в  $\mathbb{R}^m$ . По условию существуют векторы  $f_k \in \mathbb{R}^n$ ,  $1 \leq k \leq m$ , такие, что  $Af_k = e_k$ . Для каждого  $y = \sum_{k=1}^m y_k e_k \in \mathbb{R}^m$  положим  $R(y) = \sum_{k=1}^m y_k f_k$ . Тогда равенство  $AR(y) = y$  следует из линейности оператора  $A$ , а второе свойство следует из очевидных оценок (и того, что длина вектора не меньше модуля любой его координаты):  $|R(y)| \leq \sum_{k=1}^m |y_k| |f_k| \leq \max_{1 \leq k \leq m} |y_k| \sum_{k=1}^m |f_k| \leq \gamma|y|$ , где  $\gamma = \sum_{k=1}^m |f_k|$ .

Из леммы видно, что правых обратных к  $A$  «много» и константы  $\gamma$  тоже можно варьировать.

Опишем теперь метод решения системы  $m$  уравнений с  $n$  неизвестными, записанной в виде  $F(x) = y$ . Пусть дан вектор  $y$ . Возьмем вектор  $x_0$  и пусть  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$  — сюръективный оператор с правым обратным  $R$ . В качестве приближенного решения уравнения  $F(x) = y$  возьмем какое-нибудь решение линейного уравнения  $A(x - x_0) + F(x_0) = y$ . Вектор  $x_1 = x_0 + R(y - F(x_0))$ , очевидно, будет таким решением (для проверки

<sup>2)</sup>Мы предполагаем, что читатель этого пункта знаком с понятием умножения матриц, в частности, с умножением матрицы на вектор-столбец.

надо применить к обеим частям оператор  $A$ ). Теперь вместо  $x_0$  берем точку  $x_1$  и, поступая аналогично, найдем точку  $x_2$ . Продолжая этот процесс, получим последовательность

$$x_k = x_{k-1} + R(y - F(x_{k-1})), \quad k = 1, 2, \dots \quad (3.1)$$

Эту процедуру мы по-прежнему называем *модифицированным методом Ньютона*.

**ТЕОРЕМА 3** (ОБ ОБРАТНОМ ОТОБРАЖЕНИИ В КОНЕЧНОМЕРНОМ СЛУЧАЕ). Пусть заданы вектор  $x_0 \in \mathbb{R}^n$ , число  $\delta > 0$ , сюръективный линейный оператор  $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$  с правым обратным  $R$  и отображение  $F: U_{\mathbb{R}^n}(x_0, \delta) \rightarrow \mathbb{R}^m$ . Если существует такое число  $\theta \in (0, 1)$ , что для любых  $\xi, x \in U_{\mathbb{R}^n}(x_0, \delta)$  выполняется неравенство

$$|F(\xi) - F(x) - A(\xi - x)| \leq \frac{\theta}{\gamma} |\xi - x|, \quad (3.2)$$

где  $\gamma$  из оценки для  $R$ , то для каждого  $y \in U_{\mathbb{R}^m}(F(x_0), \delta_0)$ , где  $\delta_0 = \delta(1-\theta)/\gamma$ , последовательность (3.1) сходится к такому вектору  $\varphi(y) \in U_{\mathbb{R}^n}(x_0, \delta)$ , что  $F(\varphi(y)) = y$  и при этом  $|\varphi(y) - x_0| \leq K|y - F(x_0)|$ , где  $K = \gamma/(1-\theta)$ .

Отображение  $\varphi: U_{\mathbb{R}^m}(F(x_0), \delta_0) \rightarrow U_{\mathbb{R}^n}(x_0, \delta)$ , построенное в этой теореме, называется *обратным отображением* к  $F$ .

Доказательство фактически ничем не отличается от доказательства теоремы 1, но тем не менее, мы его проведем.

**ДОКАЗАТЕЛЬСТВО** Докажем, что а) векторы  $x_k$  для всех  $k \geq 0$  лежат в  $U_{\mathbb{R}^n}(x_0, \delta)$  и б) что последовательность  $\{x_k\}_{k \geq 0}$  фундаментальна. Утверждение а) докажем по индукции. Элемент  $x_0$ , очевидно, принадлежит  $U_{\mathbb{R}^n}(x_0, \delta)$ . Пусть  $x_s \in U_{\mathbb{R}^n}(x_0, \delta)$ ,  $1 \leq s \leq k$ . Используя последовательно (3.1), оценку для правого обратного, равенство

$$A(x_k - x_{k-1}) = y - F(x_k) \quad (3.3)$$

(которому удовлетворяет  $x_k$  и которое можно получить применяя  $A$  к обеим частям (3.1)), (3.2) и затем итерируя процедуру, получим:

$$\begin{aligned} |x_{k+1} - x_k| &= |R(y - F(x_k))| \leq \gamma|y - F(x_k) - y + F(x_{k-1}) + A(x_k - x_{k-1})| \leq \\ &\leq \theta|x_k - x_{k-1}| \leq \theta^2|x_{k-1} - x_{k-2}| \leq \dots \leq \theta^k|x_1 - x_0|. \end{aligned}$$

Применяя теперь неравенство треугольника, предыдущее неравенство, формулу для суммы геометрической прогрессии, (3.1) при  $k = 1$  (с оценкой для правого обратного) и то, что  $|y - F(x_0)| < \delta_0$ , будем иметь

$$\begin{aligned} |x_{k+1} - x_0| &\leq |x_{k+1} - x_k| + \dots + |x_1 - x_0| \leq \\ &\leq (\theta^k + \theta^{k-1} + \dots + 1)|x_1 - x_0| < \frac{\gamma}{1-\theta}|y - F(x_0)| < \delta \quad (3.4) \end{aligned}$$

т. е.  $x_{k+1} \in U_{\mathbb{R}^n}(x_0, \delta)$  и значит, все  $x_k$ ,  $k \geq 0$ , принадлежат  $U_{\mathbb{R}^n}(x_0, \delta)$ .

Докажем б). Для любых  $k, s \in \mathbb{N}$  имеем:  $|x_{k+s} - x_k| \leq |x_{k+s} - x_{k+s-1}| + \dots + |x_{k+1} - x_k| \leq (\theta^{k+s-1} + \dots + \theta^k)|x_1 - x_0| \leq \frac{\gamma\theta^k}{1-\theta}|y - F(x_0)| < \delta\theta^k$ , откуда вытекает, что последовательность  $\{x_k\}_{n \in \mathbb{N}}$  фундаментальна и тем самым сходится. Обозначим ее предел  $\varphi(y)$ . Переход к пределу в (3.3) (который существует из-за непрерывности  $F$  в  $U_{\mathbb{R}^n}(x_0, \delta)$ ), что доказывается также как в одномерном случае) приводит к равенству  $F(\varphi(y)) = y$ , а переход к пределу в (3.4) обеспечивает неравенство  $|\varphi(y) - x_0| \leq K|y - F(x_0)|$  с  $K = \gamma/(1 - \theta)$ .

#### 4. БЕСКОНЕЧНОМЕРНЫЙ СЛУЧАЙ

В двадцатые годы прошлого века было введено понятие *нормированного пространства*, содержащее в себе многие важнейшие свойства как прямой  $\mathbb{R}$ , так и пространств  $\mathbb{R}^n$ .

Множество  $X$  называется (вещественным) *векторным пространством*, если для его элементов определена операция сложения и умножения элемента на вещественное число. Эти операции удовлетворяют естественным свойствам, к которым мы привыкли, оперируя, например, с векторами на плоскости: коммутативность и ассоциативность сложения, существование нулевого элемента, противоположного элемента и т. п. Элементы векторного пространства (независимо от их природы) часто называют векторами.

Векторное пространство  $X$  называется *нормированным пространством*, если на нем определена функция  $\|\cdot\|_X$ , сопоставляющая вектору  $x$  число  $\|x\|_X$ , называемое *нормой вектора*  $x$ , удовлетворяющее свойствам: 1)  $\|x\|_X \geq 0$  для любого  $x \in X$  и  $\|x\|_X = 0$  тогда и только тогда, когда  $x = 0$ , 2)  $\|\alpha x\|_X = |\alpha|\|x\|_X$  для любых  $x \in X$ ,  $\alpha \in \mathbb{R}$ , 3)  $\|x + x'\|_X \leq \|x\|_X + \|x'\|_X$  для любых  $x, x' \in X$ .

Отметим, что модуль числа, длина вектора в  $\mathbb{R}^2$  и в  $\mathbb{R}^n$  удовлетворяют перечисленным свойствам и тем самым эти пространства являются нормированными.

Пусть  $\hat{x} \in X$  и  $r > 0$ . Множества  $U_X(\hat{x}, r) = \{x \in X \mid \|x - \hat{x}\|_X < r\}$  и  $B_X(\hat{x}, r) = \{x \in X \mid \|x - \hat{x}\|_X \leq r\}$  называются соответственно открытым и замкнутым шаром в  $X$  с центром в  $\hat{x}$  радиуса  $r$ .

Подмножество нормированного пространства называется *открытым*, если с любой своей точкой оно содержит и некоторый шар с центром в этой точке.

Любое открытое множество, содержащее данную точку, называется *окрестностью* этой точки.

Подмножество нормированного пространства называется *замкнутым*, если его дополнение открыто.

УПРАЖНЕНИЕ. Докажите, что открытый (замкнутый) шар является открытым (замкнутым) множеством.

В нормированном пространстве естественным образом определяется понятия сходимости последовательности и непрерывности отображения.

Говорят, что последовательность векторов  $\{x_k\}_{k \in \mathbb{N}}$  в  $X$  сходится к вектору  $\hat{x}$  и пишут  $\lim_{n \rightarrow \infty} x_k = \hat{x}$ , если для любого  $\varepsilon > 0$  найдется натуральное число  $N = N(\varepsilon)$  такое, что неравенство  $\|x_k - \hat{x}\|_X < \varepsilon$  выполняется для всех  $k > N$ .

Пусть  $U$  — открытое подмножество  $X$ ,  $Y$  — другое нормированное пространство и  $F: U \rightarrow Y$ . Говорят, что отображение  $F$  непрерывно в точке  $\hat{x} \in U$ , если для любого  $\varepsilon > 0$  найдется такое число  $\delta = \delta(\varepsilon) > 0$ , что  $\|F(\hat{x} + x) - F(\hat{x})\|_Y < \varepsilon$  как только  $\|x\|_X < \delta$ .

Легко проверить, что если отображение  $F$  непрерывно в точке  $\hat{x}$ , то для любой последовательности  $\{x_k\}_{k \in \mathbb{N}}$ , сходящейся к  $\hat{x}$ , последовательность  $\{F(x_k)\}_{k \in \mathbb{N}}$  сходится к  $F(\hat{x})$ .

Если  $F$  непрерывно в каждой точке  $U$ , то говорят, что  $F$  непрерывно на  $U$ .

Последовательность  $\{x_k\}_{k \in \mathbb{N}}$  векторов из  $X$  называется *фундаментальной* или *последовательностью Коши*, если для любого  $\varepsilon > 0$  найдется такое натуральное число  $N$ , что  $\|x_{k+s} - x_k\| < \varepsilon$  для всех  $k > N$  и  $s \in \mathbb{N}$ .

Нормированное пространство называется *полным нормированным пространством* или *банаховым пространством*, если в нем каждая фундаментальная последовательность является сходящейся.

Пространство  $X = \mathbb{R}^n$  с введенной выше нормой  $\|x\|_X = |x| = (\sum_{k=1}^n x_k^2)^{1/2}$ , является банаховым пространством. Это есть непосредственное следствие полноты вещественных чисел  $\mathbb{R}$ . Важное во многих вопросах пространство  $C([a, b])$  непрерывных функций на  $[a, b]$  с нормой  $\|x(\cdot)\|_{C([a, b])} = \max_{t \in [a, b]} |x(t)|$  также является банаховым пространством. Для доказательства надо взять фундаментальную последовательность  $\{x_k(\cdot)\}$  в  $C([a, b])$ . Тогда для каждого  $t \in [a, b]$  числовая последовательность  $\{x_k(t)\}$  будет фундаментальной и значит, сходящейся. Если обозначить через  $x(\cdot)$  функцию, для которой  $x(t) = \lim_{k \rightarrow \infty} x_k(t)$ , то уже нетрудно проверить, что  $x(\cdot)$  — непрерывная функция и последовательность  $\{x_k(\cdot)\}$  сходится к ней в  $C([a, b])$  (проделайте это).

Вернемся к вопросу о решении уравнений, но уже в нормированных пространствах, т. е. к вопросу о нахождении решения уравнения  $F(x) = y$ , где  $F$  — отображение из нормированного пространства  $X$  в нормированное пространство  $Y$ . Как и раньше, важную роль играют линейные уравнения  $Ax = y$ , где  $A$  — линейный оператор из  $X$  в  $Y$ . Мы сейчас покажем, что

вопрос о нахождении решения уравнения  $F(x) = y$  совершенно аналогичен конечномерному случаю.

**ЛЕММА 2 (О ПРАВОМ ОБРАТНОМ ОТОБРАЖЕНИИ В БЕСКОНЕЧНОМЕРНОМ СЛУЧАЕ).** Пусть  $X$  и  $Y$  — банаховы пространства  $A: X \rightarrow Y$  — линейный непрерывный сюръективный оператор<sup>3)</sup>. Тогда существуют отображение  $R: Y \rightarrow X$  (правое обратное к  $A$ ) и константа  $\gamma > 0$  такие, что  $AR(y) = y$  и  $\|R(y)\|_X \leq \gamma \|y\|_Y$  для всех  $y \in Y$ .

**ДОКАЗАТЕЛЬСТВО** Эта лемма моментально следует из одного из основных принципов линейного анализа — принципа Банаха об открытости. Действительно, согласно этому принципу (см. учебник Колмогорова и Фомина, гл. IV, п. 5), образ единичного открытого шара в  $X$  с центром в нуле при отображении  $A$  содержит некоторый открытый шар в  $Y$  с центром в нуле. Пусть радиус этого шара равен  $r$ . Тогда для любого элемента  $y \in Y$ , по норме меньшего  $r$ , найдется элемент  $x(y) \in X$ , по норме меньший единицы, такой, что  $Ax(y) = y$ . Положим  $R(0) = 0$  и  $R(y) = (2\|y\|_Y/r)x((r/2\|y\|_Y)y)$ , если  $y \in Y$  и  $y \neq 0$ . Ясно, что  $(r/2\|y\|_Y)y$  принадлежит шару радиуса  $r$  с центром в нуле и мы имеем  $AR(y) = y$  и  $\|R(y)\|_X \leq \gamma \|y\|_Y$ , где  $\gamma = 2/r$ .

**ТЕОРЕМА 4 (ОБ ОБРАТНОМ ОТОБРАЖЕНИИ В БЕСКОНЕЧНОМЕРНОМ СЛУЧАЕ).** Пусть  $X$  и  $Y$  — банаховы пространства и заданы вектор  $x_0 \in X$ , число  $\delta > 0$ , сюръективный линейный непрерывный оператор  $A: X \rightarrow Y$  с правым обратным  $R$  и отображение  $F: U_X(x_0, \delta) \rightarrow Y$ . Если существует такое число  $\theta \in (0, 1)$ , что для любых  $\xi, x \in U_X(x_0, \delta)$  выполняется неравенство

$$\|F(\xi) - F(x) - A(\xi - x)\|_Y \leq \frac{\theta}{\gamma} \|\xi - x\|_X, \quad (4.1)$$

где  $\gamma$  из оценки для  $R$ , то для каждого  $y \in U_Y(F(x_0), \delta_0)$ , где  $\delta_0 = \delta(1 - \theta)/\gamma$ , последовательность

$$x_k = x_{k-1} + R(y - F(x_{k-1})), \quad k \in \mathbb{N},$$

сходится к такому вектору  $\varphi(y) \in U_X(x_0, \delta)$ , что  $F(\varphi(y)) = y$  и при этом  $\|\varphi(y) - x_0\|_X \leq K \|y - F(x_0)\|_Y$ , где  $K = \gamma/(1 - \theta)$ .

Отображение  $\varphi: U_Y(F(x_0), \delta_0) \rightarrow U_X(x_0, \delta)$ , построенное в этой теореме, называется *обратным отображением к  $F$* .

Отметим, что если оператор  $A$  обратим, то обратное отображение  $\varphi$  единственно в следующем смысле: если  $x$  такое, что  $F(x) = y$ , то  $x = \varphi(y)$ . Действительно, поскольку оператор  $A$  обратим, то  $R = A^{-1}$  и считаем, что

<sup>3)</sup>Это означает, как и в конечномерном случае, что уравнение  $Ax = y$  имеет решение для любого  $y \in Y$ .

$\gamma = \|A^{-1}\|$ . Тогда согласно (4.1) имеем

$$\begin{aligned}\|x - \varphi(y)\|_X &= \|\Lambda^{-1}(\Lambda(x - \varphi(y)))\|_X \leq \| \Lambda^{-1} \| \|\Lambda(x - \varphi(y))\|_Y = \\ &= \| \Lambda^{-1} \| \|F(x) - F(\varphi(y)) - \Lambda(x - \varphi(y))\|_Y \leq \theta \|x - \varphi(y)\|_X,\end{aligned}$$

т. е.  $x = \varphi(y)$ .

ДОКАЗАТЕЛЬСТВО теоремы 4 является дословным повторением доказательства теоремы 3 с заменой  $|x|$  при  $x \in X$  на  $\|x\|_X$ ,  $|y|$  при  $y \in Y$  на  $\|y\|_Y$  и учетом того, что построенная фундаментальная последовательность  $\{x_k\}$  сходится, поскольку  $X$  — банахово пространство.

ИСТОРИЧЕСКИЙ КОММЕНТАРИЙ. Доказанная нами заключительная теорема принадлежит американскому математику Л. Грейвсу (1896 – 1973). Она была опубликована в 1951 г. Этим мы завершили наш экскурс в вопросы, связанные с решением нелинейных уравнений от Герона, жившего в первом веке нашей эры, до середины прошлого столетия.

## ПРИЛОЖЕНИЯ

Доказанные нами теоремы о существовании решений нелинейных уравнений или, иначе говоря, о существовании обратных функций и отображений, играют в математике огромную роль. В следующих двух разделах будет рассказано о двух фундаментальных приложениях этих результатов к доказательству правила множителей Лагранжа и к доказательству теоремы существования и единственности решения задачи Коши дифференциального уравнения. Первая теорема дает ключ к решению задач на максимум и минимум от истоков до нашего времени, вторая вскрывает фундаментальное свойство детерминизма процессов, описываемых дифференциальными уравнениями.

### 5. ПРАВИЛО МНОЖИТЕЛЕЙ ЛАГРАНЖА.

В первой части этого раздела мы как бы низвергнемся из бесконечности к единице и из университета снова поступаем в школу.<sup>4)</sup> Будем учиться решать задачи на максимум и минимум. Термины «максимум» и «минимум» объединяются общим термином «экстремум», и задачи на максимум и минимум называют *экстремальными задачами*.

Экстремальные задачи стали исследовать на заре развития математики. В «Началах» Евклида (третий век до нашей эры) содержится такая

<sup>4)</sup> Правда во второй части этого раздела мы поднимемся от единицы к двум и далее к любому  $n$ , а в следующем разделе снова вознесемся к бесконечности.

**ЗАДАЧА 5.1.** На основании  $AC$  треугольника  $ABC$  требуется найти такую точку  $F$ , что параллелограмм  $ADEF$ , у которого вершины  $D$  и  $E$  лежат соответственно на сторонах  $AB$  и  $BC$ , имел бы наибольшую площадь.

Переведем эту задачу на формальный язык. Пусть  $b = |AC|$  — длина основания,  $H$  — высота треугольника  $ABC$ ,  $x = |EF|$ ,  $h(x)$  — высота треугольника  $BDE$ . Из подобия треугольников  $DBE$  и  $ABC$  получаем:  $\frac{h(x)}{H} = \frac{x}{b}$ . Площадь параллелограмма равна  $\frac{H(b-x)x}{b}$ . Откуда приходим к такой формулировке: найти максимум функции  $x \mapsto f(x) = \frac{H(b-x)x}{b}$  на отрезке  $[0, b]$ .

Задачи на максимум и минимум встречаются в трудах и других математиков древности, например, у Зенона, Архимеда, Аполлония, Герона. И в новое время (в 16 и 17 веках) было решено множество экстремальных задач. В качестве примера приведем задачу Тартальи (1500 – 1557), который прославился тем, что научился решать уравнения третьей степени.

**ЗАДАЧА 5.2.** Разделить число восемь на две такие части, чтобы произведение их произведения на разность было максимальным.

Вплоть до 17 века каждая экстремальная задача решалась «индивидуально», приемом специально придуманным именно для нее. Первый общий метод нахождения экстремумов функций одного переменного был предложен Ферма (1600 – 1665). Для формулировки теоремы Ферма надо напомнить понятие производной функции одного переменного.

Пусть функция  $f$  определена на некотором интервале  $U$ . Говорят, что  $f$  дифференцируема в точке  $\hat{x} \in U$ , если существует такое число  $a$ , что для всех  $x$ , для которых  $\hat{x} + x \in U$  справедливо представление  $f(\hat{x} + x) = f(\hat{x}) + ax + r(x)$ , где  $r(x)/x \rightarrow 0$  при  $x \rightarrow 0$ . Число  $a$  (определяемое этим представлением однозначно) называется производной функции  $f$  в точке  $\hat{x}$  и обозначается  $f'(\hat{x})$ .

Точка  $\hat{x} \in U$  называется локальным максимумом (минимумом) функции  $f$ , если найдется такой интервал  $(\hat{x} - \delta, \hat{x} + \delta) \subset U$ , что для любой точки  $x$  из этого интервала выполнено неравенство  $f(x) \leq f(\hat{x})$  ( $f(x) \geq f(\hat{x})$ ). Точки локального максимума и локального минимума функции  $f$  называются локальными экстремумами этой функции.

**ТЕОРЕМА ФЕРМА ДЛЯ ФУНКЦИЙ ОДНОГО ПЕРЕМЕННОГО.** Если в точке локального экстремума функция одного переменного дифференцируема, то производная в этой точке равна нулю.

Действительно, пусть  $\hat{x}$  — локальный экстремум  $f$ . Функция  $f$  дифференцируема в  $\hat{x}$  и поэтому  $f(\hat{x} + x) = f(\hat{x}) + f'(\hat{x})x + r(x) = f(\hat{x}) + x(f'(\hat{x}) + r(x)/x)$ . Если, скажем,  $f'(\hat{x}) > 0$  и  $\hat{x}$  — локальный максимум, то для всех



достаточно малых  $x > 0$  будем иметь  $f(\hat{x} + x) > f(\hat{x})$  в противоречии с тем, что  $\hat{x}$  — локальный максимум. Остальные случаи рассматриваются аналогично.

Задачи 5.1 и 5.2 легко решаются с помощью теоремы Ферма. В задаче 5.1 функция  $f(x) = \frac{H(b-x)x}{b}$  определена на отрезке  $[0, b]$  и непрерывна на нем. По теореме Вейерштрасса эта функция достигает на этом отрезке своего абсолютного максимума, причем достигает его внутри отрезка (поскольку она неотрицательна и на концах отрезка равна нулю), а значит, абсолютный максимум должен совпадать с одним из локальных максимумов. Из теоремы Ферма следует, что локальный максимум  $\hat{x}$  один, причем он равен  $\frac{b}{2}$ . Значит, искомая точка  $F$  в задаче Евклида — середина основания  $AC$ .

Решить задачу Тарталья предоставим читателю. Для квадрата большего из чисел ответ Тарталья сформулировал в такой форме: *число восемь надо разделить пополам и квадрат этой половины увеличить на одну треть этого квадрата*.

Теорема Ферма дает возможность найти экстремум в задаче без ограничений. Большинство же интересных задач на экстремум — с ограничениями. Приведем два примера.

**ЗАДАЧА 5.3.** *Найти максимум произведения  $n$  положительных чисел, если их сумма равна единице.*

**ЗАДАЧА 5.4.** *Найти максимум линейной функции на единичном шаре.*

Доказательство неравенства между средним геометрическим и средним арифметическим сводится к решению задачи 5.3, а неравенства Коши — Буняковского — к решению задачи 5.4.

Поставим проблему в общей форме. Пусть на открытом подмножестве  $U \subset \mathbb{R}^n$  определены функции  $f_i$ ,  $i = 0, 1, \dots, m$ . Требуется найти такие точки  $x \in U$ , которые удовлетворяют условиям  $f_i(x) = 0$ ,  $1 \leq i \leq m$ , и в которых  $f_0(x)$  принимает свое максимальное (минимальное) значение. Эту проблему будем записывать как

$$f_0(x) \rightarrow \max(\min), \quad f_1(x) = 0, \dots, f_m(x) = 0. \quad (P)$$

Нас, в основном, будут интересовать локальные максимумы и минимумы. Точка  $x \in U$  называется *допустимой* в задаче  $(P)$ , если  $f_i(x) = 0$ ,  $1 \leq i \leq m$ . Точка  $\hat{x} \in U$  называется *локальным максимумом* (минимумом) в задаче  $(P)$ , если найдется такое  $\delta > 0$ , что  $U_{\mathbb{R}^n}(\hat{x}, \delta) \subset U$  и для любой допустимой точки  $x \in U_{\mathbb{R}^n}(\hat{x}, \delta)$  выполнено неравенство  $f_0(x) \leq f_0(\hat{x})$  ( $f_0(x) \geq f_0(\hat{x})$ ).

Задачи 5.3 и 5.4 записываются в форме задачи (P) следующим образом:

$$x_1 x_2 \cdots x_n \rightarrow \max, \quad x_1 + x_2 + \cdots + x_n = 1, \quad x_i > 0, \quad 1 \leq i \leq n,$$

и

$$y_1 x_1 + y_2 x_2 + \cdots + y_n x_n \rightarrow \max, \quad x_1^2 + x_2^2 + \cdots + x_n^2 = 1.$$

Для того, чтобы сформулировать прием решения задачи (P), описанный Лагранжем, необходимо дать определение дифференцируемости функции многих переменных (а также некоторое усиление этого понятия).

**ДИФФЕРЕНЦИРУЕМОСТЬ И СТРОГАЯ ДИФФЕРЕНЦИРУЕМОСТЬ.** Удобно привести соответствующие определения сразу для отображений конечномерных пространств. Пусть  $U$  — окрестность точки  $\hat{x} \in \mathbb{R}^n$ . Говорят, что отображение  $F: U \rightarrow \mathbb{R}^m$  дифференцируемо в точке  $\hat{x}$ , если существует такая матрица  $A$  размера  $m$  на  $n$ , что для всех  $x \in \mathbb{R}^n$ , для которых  $\hat{x} + x \in U$  справедливо представление  $F(\hat{x} + x) = F(\hat{x}) + Ax + r(x)$ , где  $|r(x)|/|x| \rightarrow 0$  при  $x \rightarrow 0$ . Матрица  $A$  (определяемая этим представлением однозначно) называется *производной отображения  $F$  в точке  $\hat{x}$*  и обозначается  $F'(\hat{x})$ .

Если  $m = 1$ , т. е.  $F$  — функция, то производная есть вектор-строка, элементы которой суть частные производные данной функции по  $x_1, \dots, x_n$  в точке  $\hat{x}$ .

Говорят, что отображение  $F: U \rightarrow \mathbb{R}^n$  строго дифференцируемо в точке  $\hat{x}$ , если оно дифференцируемо в этой точке и для любого  $\varepsilon > 0$  найдется такое  $\delta > 0$ , что для любых  $\xi, x \in U_{\mathbb{R}^n}(\hat{x}, \delta)$  справедливо неравенство

$$|F(\xi) - F(x) - F'(\hat{x})(\xi - x)| \leq \varepsilon |\xi - x|. \quad (5.1)$$

Снова, если  $m = 1$ , то мы получаем определение строгой дифференцируемости для функции.

Если функция определена на некотором открытом подмножестве пространства  $\mathbb{R}^n$ , то аналогично одномерному случаю доказывается

**ТЕОРЕМА ФЕРМА ДЛЯ ФУНКЦИЙ МНОГИХ ПЕРЕМЕННЫХ.** *Если в точке локального экстремума функция дифференцируема, то ее производная в этой точке равна нулю.*

А теперь мы готовы предоставить слово самому Лагранжу. В своей книге «Теория аналитических функций» (1797) он писал:

«Можно высказать следующий общий принцип. Если ищется максимум или минимум некоторой функции многих переменных при условии, что между этими переменными имеется связь, задаваемая одним или несколькими уравнениями, нужно прибавить к функции, о которой говорилось, функции, задающие уравнения связи, умноженные на неопределенные множители, и

искать затем максимум или минимум построенной суммы, как если бы переменные были независимы. Полученные уравнения, присоединенные к уравнениям связи, послужат для определения всех неизвестных.»

Мы позволим себе лишь незначительно изменить рецепт Лагранжа. Во-первых, оказалось удобным умножить на неопределенный множитель и функцию, экстремум которой ищется, т. е. (имея в виду задачу  $(P)$ ) составлять выражение  $\mathcal{L}(x, \lambda_0, \dots, \lambda_m) = \sum_{i=0}^m \lambda_i f_i(x)$ ; которое стали называть *функцией Лагранжа* (а числа  $\lambda_i$  — *множителями Лагранжа*). Во-вторых, мы не будем «искать минимум или максимум» функции Лагранжа, а просто будем применять к задаче на экстремум функции Лагранжа необходимое условие экстремума «как если бы переменные были независимы», т. е. теорему Ферма. Сформулируем точный результат.

**ПРАВИЛО МНОЖИТЕЛЕЙ ЛАГРАНЖА.** Пусть в задаче  $(P)$  все функции  $f_i$ ,  $i = 0, 1, \dots, m$ , строго дифференцируемы в точке  $\hat{x} \in U$ . Тогда, если  $\hat{x}$  — локальный экстремум в этой задаче, то существует такие множители Лагранжа  $\lambda_0, \lambda_1, \dots, \lambda_m$ , не равные одновременно нулю, что производная функции Лагранжа по  $x$  в этой точке равна нулю, т. е.

$$\lambda_0 f'_0(\hat{x}) + \lambda_1 f'_1(\hat{x}) + \dots + \lambda_m f'_m(\hat{x}) = 0.$$

Доказательство Теорема утверждает, что если  $\hat{x}$  — локальный экстремум, то векторы  $f'_i(\hat{x})$ ,  $0 \leq i \leq m$ , линейно зависимы. Докажем теорему от противного: предположим, что эти векторы линейно независимы и придем к противоречию с тем, что  $\hat{x}$  — локальный экстремум. Рассмотрим отображение  $F: U \rightarrow \mathbb{R}^{m+1}$ , определенное по правилу  $F(x) = (f_0(x), f_1(x), \dots, f_m(x))^T$ . По условию функции  $f_i$ ,  $0 \leq i \leq m$ , строго дифференцируемы в  $\hat{x}$ . Отсюда легко следует, что и отображение  $F$  строго дифференцируемо в  $\hat{x}$ . Векторы  $f'_i(\hat{x})$ ,  $i = 0, 1, \dots, m$ , образуют строки матрицы  $F'(\hat{x})$ . Следовательно, ранг этой матрицы равен  $m + 1$  и значит,  $F'(\hat{x})$  — сюръективный оператор. Согласно лемме 1 существует правый обратный  $R$  к  $F'(\hat{x})$ . В силу строгой дифференцируемости  $F$  в точке  $\hat{x}$ , по  $\varepsilon = 1/2\gamma$  (где  $\gamma$  из оценки для правого обратного) можно выбрать такое  $\delta > 0$ , что будет выполняться неравенство (5.1), которое есть в точности неравенство (3.2) в теореме 3 с  $\theta = 1/2$  и  $A = F'(\hat{x})$ . Согласно этой теореме (применительно к нашему случаю) существует отображение  $\varphi: U_{\mathbb{R}^{m+1}}(F(\hat{x}), \delta_0) \rightarrow U_{\mathbb{R}^n}(\hat{x}, \delta)$ , где  $\delta_0 = \delta/2\gamma$ , и константа  $K = 2\gamma$  такие, что  $F(\varphi(y)) = y$  и  $|\varphi(y) - \hat{x}| \leq K|y - F(\hat{x})|$  для всех  $y \in U_{\mathbb{R}^{m+1}}(F(\hat{x}), \delta_0)$ . Так как  $F(\hat{x}) = (f_0(\hat{x}), 0, \dots, 0)^T$ , то для всех достаточно малых по модулю  $\nu$  вектор  $y_\nu = (f_0(\hat{x}) + \nu, 0, \dots, 0)^T$  принадлежит  $U_{\mathbb{R}^{m+1}}(F(\hat{x}), \delta_0)$ . Положим  $x_\nu = \varphi(y_\nu)$ . Тогда  $F(x_\nu) = F(\varphi(y_\nu)) = y_\nu$ , или

$(f_0(x_\nu), f_1(x_\nu), \dots, f_m(x_\nu))^T = (f_0(\hat{x}) + \nu, 0, \dots, 0)^T$ , т.е.  $f_0(x_\nu) = f_0(\hat{x}) + \nu$ ,  $f_i(x_\nu) = 0$ ,  $1 \leq i \leq m$ , и  $|x_\nu - \hat{x}| \leq K|y_\nu - F(\hat{x})| = K|\nu|$ . Это означает, что в любой окрестности точки  $\hat{x}$  существует допустимый в задаче (P) вектор  $x_\nu$  такой, что  $f_0(x_\nu) < f_0(\hat{x})$  ( $f_0(x_\nu) > f_0(\hat{x})$ ), если  $\nu < 0$  ( $\nu > 0$ ), в противоречии с тем, что  $\hat{x}$  — локальный экстремум. Таким образом, векторы  $f'_i(\hat{x})$ ,  $0 \leq i \leq m$ , линейно зависимы и теорема доказана.

Решим теперь, используя данную теорему, задачи 5.3 и 5.4.

Рассмотрим задачу 5.3, но с заменой  $x_i > 0$  на  $x_i \geq 0$ ,  $1 \leq i \leq n$ . В этой задаче множество ограничений компактно (т.е. ограничено и замкнуто), максимизируемая функция непрерывна и, следовательно, по теореме Вейрштрасса решение  $(\hat{x}_1, \dots, \hat{x}_n)$  такой задачи существует. Очевидно, что  $\hat{x}_i \neq 0$ ,  $1 \leq i \leq n$ . Но тогда эта точка является решением и исходной задачи и, в частности, локальным максимумом в задаче 5.3 с единственным ограничением:  $x_1 + \dots + x_n = 1$ . Можно применить правило множителей Лагранжа: существуют такие числа  $\lambda_0$  и  $\lambda_1$ , не равные одновременно нулю, что производная функции Лагранжа  $(x_1, \dots, x_n) \mapsto \lambda_0 x_1 \dots x_n + \lambda_1 (x_1 + \dots + x_n - 1)$  в этой точке равна нулю:  $\lambda_0 \hat{x}_1 \dots \hat{x}_n + \lambda_1 = \dots = \lambda_0 \hat{x}_1 \dots \hat{x}_{n-1} + \lambda_1 = 0$ . Отметим, что  $\lambda_0 \neq 0$ , ибо иначе  $\lambda_1 = 0$ , что невозможно. Отсюда и из выписанных равенств легко следует, что  $\hat{x}_i / \hat{x}_j = 1$  для любых  $i \neq j$ , т.е.  $\hat{x}_1 = \dots = \hat{x}_n$  и так как  $\hat{x}_1 + \dots + \hat{x}_n = 1$ , то  $\hat{x}_i = 1/n$ ,  $1 \leq i \leq n$ . Итак, необходимым условиям удовлетворяет единственная точка  $(1/n, \dots, 1/n)$  и значит, она решение исходной задачи 5.3 и значение этой задачи равно  $1/n^n$ .

Выведем отсюда неравенство для средних. Пусть  $x_1, \dots, x_n$  — произвольные положительные числа. Тогда числа  $x_1 / \sum_{i=1}^n x_i, \dots, x_n / \sum_{i=1}^n x_i$  удовлетворяют ограничениям задачи 5.3 и значит,  $x_1 \dots x_n / (\sum_{i=1}^n x_i)^n \leq 1/n^n$ . Извлекая из обеих частей корень  $n$ -й степени, получаем неравенство

$$\sqrt[n]{x_1 \dots x_n} \leq \frac{x_1 + \dots + x_n}{n}.$$

Перейдем к задаче 5.4. Естественно, мы предполагаем, что не все  $y_i$ ,  $1 \leq i \leq n$ , равны нулю. Решение этой задачи  $(\hat{x}_1, \dots, \hat{x}_n)$  существует по теореме Вейрштрасса (множество ограничений компактно, а максимизируемая функция непрерывна). Согласно правилу множителей Лагранжа существуют такие числа  $\lambda_0$  и  $\lambda_1$ , не равные одновременно нулю, что производная функции Лагранжа  $(x_1, \dots, x_n) \mapsto \lambda_0 (y_1 x_1 + \dots + y_n x_n) + \lambda_1 (x_1^2 + \dots + x_n^2 - 1)$  в точке  $(\hat{x}_1, \dots, \hat{x}_n)$  равна нулю:  $\lambda_0 y_1 + 2\lambda_1 \hat{x}_1 = \dots = \lambda_0 y_n + 2\lambda_1 \hat{x}_n = 0$ . Множитель  $\lambda_0 \neq 0$ , так как в противном случае  $\lambda_1 = 0$  (если бы  $\lambda_1 \neq 0$ , то  $\hat{x}_1 = \dots = \hat{x}_n = 0$  в противоречие с тем, что  $\hat{x}_1 + \dots + \hat{x}_n = 1$ ), что невозможно. Отметим, что и  $\lambda_1 \neq 0$ , ибо в противном случае мы получили бы, что  $y_1 = \dots = y_n = 0$ . Из выписанных равенств следует, что  $\hat{x}_i = (-\lambda_0 / 2\lambda_1) y_i$ ,  $i = 1, \dots, n$ . Так как  $\sum_{i=1}^n y_i \hat{x}_i = (-\lambda_0 / 2\lambda_1) \sum_{i=1}^n y_i^2$

и у нас задача на максимум, то необходимо  $\lambda_0/2\lambda_1 < 0$ . Подставляя  $\hat{x}_i$ ,  $1 \leq i \leq n$ , в ограничения, получаем, что  $\lambda_0/2\lambda_1 = -1/\sqrt{\sum_{i=1}^n y_i^2}$ . Таким образом,  $\hat{x}_i = y_i/\sqrt{\sum_{j=1}^n y_j^2}$ ,  $1 \leq i \leq n$ , и значение задачи равно  $\sqrt{\sum_{j=1}^n y_j^2}$ .

Выведем отсюда неравенство Коши – Буняковского. Пусть  $(x_1, \dots, x_n)$  и  $(y_1, \dots, y_n)$  – произвольные ненулевые наборы чисел. Набор  $(x_1/\sqrt{\sum_{i=1}^n x_i^2}, \dots, x_n/\sqrt{\sum_{i=1}^n x_i^2})$  удовлетворяет ограничениям задачи 5.4 (с данными набором  $(y_1, \dots, y_n)$ ). Тогда по доказанному

$$\sum_{i=1}^n y_i \left( \frac{x_i}{\sqrt{\sum_{j=1}^n x_j^2}} \right) \leq \sqrt{\sum_{j=1}^n y_j^2}.$$

Откуда сразу следует, что

$$\sum_{i=1}^n y_i x_i \leq \left( \sum_{i=1}^n y_i^2 \right)^{1/2} \left( \sum_{i=1}^n x_i^2 \right)^{1/2}.$$

Отметим, что подобным образом можно доказывалось огромное число точных неравенств, возникающих в анализе и геометрии.

УПРАЖНЕНИЕ. Докажите неравенство Гёльдера

$$\sum_{i=1}^n y_i x_i \leq \left( \sum_{i=1}^n |y_i|^p \right)^{1/p} \left( \sum_{i=1}^n |x_i|^q \right)^{1/q},$$

где  $1 < p, q < \infty$  и  $1/p + 1/q = 1$ .

Задача (P) и правило множителей Лагранжа естественным образом обобщаются на бесконечномерный случай. Пусть  $X$  и  $Y$  – нормированные пространства,  $U$  – открытое подмножество  $X$ ,  $f_0: U \rightarrow \mathbb{R}$  и  $F: U \rightarrow Y$ . Рассмотрим задачу

$$f_0(x) \rightarrow \max(\min), \quad F(x) = 0.$$

Аналогично предыдущему определяется понятие локального экстремума. Теорема 4 (вместе с теоремой отделимости для выпуклых множеств) позволяет доказать аналог правила множителей Лагранжа для данного случая (сам этот факт был установлен Л. А. Люстерником в 1934 году). Мы не будем на этом останавливаться, но отметим, что из этой теоремы, в частности, вытекают необходимые условия экстремума в различных задачах классического вариационного исчисления.

## 6. СУЩЕСТВОВАНИЕ РЕШЕНИЙ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Пусть  $(t, x) \rightarrow f(t, x)$  — функция  $(n+1)$ -го переменного ( $t \in \mathbb{R}$ ,  $x \in \mathbb{R}^n$ ) со значениями в  $\mathbb{R}^n$ , непрерывная в окрестности точки  $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$ . Если непрерывно дифференцируемая функция  $\hat{x}(\cdot)$  со значениями в  $\mathbb{R}^n$ , определенная на некотором отрезке  $\Delta = [t_0 - \alpha, t_0 + \alpha]$  такова, что  $\hat{x}(t) = f(t, \hat{x}(t))$  для всех  $t \in \Delta$  и  $\hat{x}(t_0) = x_0$ , то говорят, что  $\hat{x}(\cdot)$  — решение задачи Коши

$$\dot{x} = f(t, x), \quad x(t_0) = x_0. \quad (6.1)$$

на отрезке  $\Delta$ .

Следующий результат есть непосредственное следствие теоремы 4 об обратном отображении.

Выше было определено банахово пространство непрерывных функций  $C([a, b])$ . Нам понадобится аналогичное пространство, но для вектор-функций, т. е. функций со значениями в  $\mathbb{R}^n$ . Совокупность всех непрерывных на отрезке  $\Delta$  функций  $x(\cdot)$  со значениями в  $\mathbb{R}^n$  обозначим через  $C(\Delta, \mathbb{R}^n)$ . Это банахово пространство с нормой  $\|x(\cdot)\|_{C(\Delta, \mathbb{R}^n)} = \max_{t \in \Delta} |x(t)|$ , где  $x(t) = (x_1(t), \dots, x_n(t))^T$ .

**ЛОКАЛЬНАЯ ТЕОРЕМА СУЩЕСТВОВАНИЯ И ЕДИНСТВЕННОСТИ.** Пусть  $t_0 \in \mathbb{R}$ ,  $x_0 \in \mathbb{R}^n$ ,  $a > 0$ ,  $b > 0$ ,  $D = [t_0 - a, t_0 + a] \times B_{\mathbb{R}^n}(x_0, b)$ , функция  $f: D \rightarrow \mathbb{R}^n$  непрерывна на  $D$  и удовлетворяет на этом множестве условию Липшица по  $x$ , т. е. существует такая константа  $L > 0$ , что для всех  $(t, x_i) \in D$ ,  $i = 1, 2$ , справедливо неравенство

$$|f(t, x_1) - f(t, x_2)| \leq L|x_1 - x_2|. \quad (6.2)$$

Тогда найдется такое число  $0 < \alpha \leq a$ , что на отрезке  $\Delta = [t_0 - \alpha, t_0 + \alpha]$  определено единственное решение  $\hat{x}(\cdot)$  задачи Коши (6.1) и при этом

$$\max_{t \in \Delta} |\hat{x}(t) - x_0| \leq 2 \max_{t \in \Delta} \left| \int_{t_0}^t f(\tau, x_0) d\tau \right|.$$

**ДОКАЗАТЕЛЬСТВО** Пусть  $M = M(x_0) = \max_{t \in [t_0 - a, t_0 + a]} |f(t, x_0)|$  и  $\alpha = \min(a, 1/2L, b/2M)$ . Применим теорему 4, обозначив  $X = Y = C(\Delta, \mathbb{R}^n)$ ,  $x_0 = x_0(\cdot)$  (функция, равная тождественно  $x_0$  на  $\Delta$ ),  $\delta = b$ ,  $F: U_X(x_0(\cdot), b) \rightarrow Y$  задается формулой ( $t \in \Delta$ )

$$F(x(\cdot))(t) = x(t) - x_0 - \int_{t_0}^t f(\tau, x(\tau)) d\tau, \quad (i)$$

$A$  — тождественный оператор ( $\gamma = \|A^{-1}\| = 1$ ) и  $\theta = 1/2$  (тем самым  $\delta_0 = b/2$ ).

Для всех  $\xi(\cdot), x(\cdot) \in U_X(x_0(\cdot), b)$  и  $t \in \Delta$  имеем, используя последовательно (i), (6.2) и определение  $\Delta$ :

$$|F(\xi(\cdot))(t) - F(x(\cdot))(t) - (\xi(t) - x(t))| = \left| \int_{t_0}^t (f(\tau, \xi(\tau)) - f(\tau, x(\tau))) d\tau \right| \leq \\ \leq |t - t_0|L \|\xi(\cdot) - x(\cdot)\|_{C(\Delta, \mathbb{R}^n)} \leq \frac{1}{2} \|\xi(\cdot) - x(\cdot)\|_{C(\Delta, \mathbb{R}^n)},$$

т. е. справедливо неравенство (4.1). Далее, снова, в силу (i) и определения  $\Delta$

$$|F(x_0(\cdot))(t)| = \left| \int_{t_0}^t f(\tau, x_0) d\tau \right| \leq |t - t_0|M \leq \frac{b}{2}.$$

Это значит, что  $0 \in U_X(F(x_0(\cdot))(\cdot), b/2)$  и тем самым существует функция  $\hat{x}(\cdot) = \varphi(0)$ , удовлетворяющая условию:  $F(\hat{x}(\cdot))(\cdot) = 0$ , т. е.

$$\hat{x}(t) - x_0 - \int_{t_0}^t f(\tau, \hat{x}(\tau)) d\tau = 0, \quad \text{для любого } t \in \Delta.$$

Таким образом,  $\hat{x}(\cdot)$  — решение задачи Коши (6.1) на отрезке  $\Delta$  и оно единственно в силу замечания после формулировки теоремы 4.

Последнее утверждение сразу следует из соответствующей оценки в теореме.

Отметим, что глобальная теорема существования и единственности для линейной системы уравнений также есть непосредственное следствие теоремы 4.

В непосредственной близости от тех вопросов, которым посвящена данная работа, лежит множество других фундаментальных результатов алгебры, геометрии, анализа, теории дифференциальных уравнений, которые являются фактически следствиями доказанных утверждений. Таковы, например, критерий Сильвестра, приведение симметрических матриц к каноническому виду (и тем самым существенная часть аналитической геометрии), вопросы непрерывной и дифференцируемой зависимости решений дифференциальных уравнений от начальных данных и параметров, представление функций рядами Фурье, решение задач математической физики методом Фурье и многое другое.

---

Г. Г. Магарил-Ильяев, Московский государственный институт радиотехники, электроники и автоматики (технический университет)

В. М. Тихомиров, механико-математический факультет МГУ

# Графы-расширители и их применения в теории кодирования

С. Б. Гашков

## 1. ВВЕДЕНИЕ

Графы-расширители (по английски экспандеры) — это очень интересный класс графов, изучение которых первыми начали московские математики М. С. Пинскер, Л. А. Бассалыго и Г. А. Маргулис в семидесятые годы прошлого века. За прошедшее время эти графы нашли много неожиданных применений, например в теории сложности вычислений и в теории кодирования. Загадочным образом они оказались также связаны с далекими от классической теории графов разделами современной математики, например, с теорией групп и теорией чисел, и являются в настоящее время предметом активных исследований, ведущихся (увы!) в основном зарубежными математиками. Библиография по этой теме насчитывает сотни публикаций. Читателю предлагается обзор некоторых из них, в основном связанных с теорией кодирования. Надеемся, что он будет способствовать активизации подобных исследований в нашей стране.

Часть доказательств упрощена по сравнению с оригинальными статьями.

### 1.1. ЧТО ТАКОЕ ГРАФЫ-РАСШИРИТЕЛИ

Графы-расширители были впервые введены М. С. Пинскером [1] (см. также [2, 3]).

Определить их можно следующим образом. Рассмотрим неориентированный граф  $G$  (возможно с петлями и параллельными ребрами), обычный или двудольный (все необходимые определения из теории графов можно найти, например, в [4]). Для произвольного множества  $S$  его вершин (взятого из одной доли, если он двудольный) обозначим  $\partial S$  множество ребер, выходящих из  $S$  в его дополнение, и назовем его *реберной границей*. Множество концов этих ребер обозначим  $\Gamma(S)$  и назовем *вершинной границей*.

Назовем *реберным (вершинным) отношением расширения* графа  $G$



число

$$h_e(G) = \min_{\{S: |S| \leq \frac{n}{2}\}} \frac{|\partial S|}{|S|}$$

и соответственно число

$$h(G) = \min_{\{S: |S| \leq \frac{n}{2}\}} \frac{|\Gamma(S)|}{|S|}.$$

Последовательность  $G_i$   $d$ -регулярных графов<sup>1)</sup> с растущим числом вершин называется *семейством графов-расширителей*, если существует  $\epsilon > 0$  такое, что для любого  $i$  справедливо неравенство  $h(G_i) \geq \epsilon$  ( $h_e(G_i) \geq \epsilon$ , если рассматривается реберное расширение).

Это самое общее определение, но в разных статьях с разными целями даются разные его варианты, иногда для произвольных, но чаще для двудольных графов, регулярных с одной или с двух сторон (не обязательно равной мощности) и как правило для случая вершинного расширения; часто также специализируются участвующие в определении параметры.

## 1.2. Что такое линейные коды

*Бинарным линейным кодом* длины  $n$  называется любое множество двоичных векторов длины  $n$  такое, что результат покомпонентного сложения по модулю два любых двух его векторов всегда принадлежит коду. Число единиц в сумме двух векторов по модулю два называется *расстоянием* между этими векторами. Минимальное расстояние между разными векторами кода (*кодowymi словами*) называется *расстоянием кода*. Линейный код можно рассматривать как линейное пространство над полем из двух элементов. Размерность этого пространства есть *размерность кода*.

Линейные коды можно определить и над любым конечным полем. Пусть  $GF(q)$  — конечное поле из  $q$  элементов<sup>2)</sup>. Назовем  $q$ -ичным линейным кодом  $C$  длины  $n$  и размерности  $k$  (сокращенно  $[n, k]$ -кодом) любое линейное  $k$ -мерное подпространство  $C$  пространства  $GF(q)^n$  всех  $n$ -мерных векторов над полем  $GF(q)$ . Число  $R = k/n$  называется *пропускной способностью* кода (для того, чтобы воспользоваться кодом, надо произвольное  $q$ -ичное слово длины  $k$  превратить в кодовое слово длины  $n$  с помощью подходящего кодирующего отображения  $GF(q)^k \rightarrow C$ .) *Расстоянием Хемминга* между векторами  $x, y \in GF(q)^n$  называется число

<sup>1)</sup>Т.е. графов, у которых число ребер, выходящих из каждой вершины равно  $d$ .

<sup>2)</sup>Никаких существенных сведений из теории конечных полей далее не понадобится. Полезно, однако, знать, что порядок поля (число элементов в нем) есть степень простого числа, и каждое поле определяется по своему порядку однозначно (с точностью до изоморфизма), благодаря чему поле порядка  $q$  обычно обозначается  $GF(q)$  (первая буква напоминает о первооткрывателе конечных полей Эваристе Галуа).

$\rho(x, y)$  координат, в которых эти вектора не совпадают. Легко видеть, что так определенное расстояние совпадает в случае  $q = 2$  с введенным выше, и для него выполнено неравенство треугольника  $\rho(x, y) \leq \rho(x, z) + \rho(y, z)$ . Кодовым расстоянием называется, как и бинарном случае, минимальное расстояние между разными кодовыми векторами. Легко видеть, что кодовое расстояние линейного кода равно минимальному весу, который может иметь ненулевой кодовый вектор (*весом вектора* называется число его ненулевых координат). Отношение кодового расстояния к длине кода называется *относительным кодовым расстоянием*.

Если  $[n, k]$ -код имеет кодовое расстояние  $d = 2t + 1$  (такие коды часто называются  $[n, k, d]$ -кодами), то он может исправлять вплоть до  $t$  ошибок. Действительно, если при передаче кодового слова  $x$  в нем произошло  $t$  ошибок, то мы получим искаженное слово  $x'$ , для которого  $\rho(x', x) = t$ . По искаженному слову можно однозначно восстановить кодовое слово, так как, если из двух разных кодовых слов  $x, y$  получено одно и то же искаженное не более чем  $t$  ошибками слово  $x' = y' = z$ , то, согласно неравенству треугольника,

$$2t + 1 = d \leq \rho(x, y) \leq \rho(x, z) + \rho(y, z) \leq 2t,$$

а это невозможно. Восстановление кодового слова по искаженному слову называется *декодированием*. Очевидный алгоритм декодирования заключается в переборе всех слов, удаленных от искаженного слова на расстояние не более  $t$ , и проверке их на принадлежность данному коду. Вместо этого можно перебирать все кодовые слова и искать среди них ближайшее к данному искаженному слову. Для произвольного линейного кода эти алгоритмы могут работать слишком медленно. В теории кодирования разработаны множество методов эффективного построения линейных (и не только линейных) кодов с достаточно быстрыми алгоритмами кодирования и декодирования. Из-за ограниченности места для ознакомления с ними мы вынуждены отослать читателя к специальной литературе (например, [5]).

Однако пример одного важного в теоретическом и прикладном отношении кода можно описать в нескольких строчках. Это код Рида-Соломона над полем  $GF(q)$  (сокращенно RS-код). Простейший вариант его построения следующий. Пусть  $k < n \leq q$ . Сопоставим каждому вектору  $a \in GF(q)^k$  многочлен

$$a(x) = a_0 + a_1x + \dots + a_{k-1}x^{k-1} \in GF(q)[X]$$

степени  $k - 1$  над полем  $GF(q)$ . Пусть  $x_1, \dots, x_n \in GF(q)$  различные элементы этого поля. Рассмотрим линейное отображение  $l : GF(q)^k \rightarrow GF(q)^n$ , определяемое равенством  $l(a) = (a(x_1), \dots, a(x_n)) \in GF(q)^n$ . Образ  $l(GF(q)^k) \subset GF(q)^n$  этого отображения — линейный код  $C$ , называе-

мый RS-кодом. В силу неравенства  $n > k$  многочлен  $a(x)$  степени  $k - 1$  однозначно восстанавливается по своим значениям в  $n$  точках, поэтому отображение  $l : GF(q)^k \rightarrow C$  взаимно однозначно, значит, мощность кода  $C$  равна  $q^k$ , поэтому его размерность равна  $k$ . Кодовое расстояние  $d(C) \geq n - k + 1$ , так как для любого ненулевого многочлена  $a(x)$  вектор его значений  $(a(x_1), \dots, a(x_n)) \in C \subset GF(q)^n$  имеет вес, не меньший  $n - k + 1$ , потому что ненулевой многочлен степени  $k - 1$  имеет не более  $k - 1$  корней. На самом деле  $d(C) = n - k + 1$ , потому что согласно так называемой границе Синглтона<sup>3)</sup> для любого  $[n, k]$ -кода  $d(C) \leq n - k + 1$ . Коды, лежащие на этой границе, называются *кодами с максимальным расстоянием*. Как видим, такими являются RS-коды<sup>4)</sup>. Доказательство границы Синглтона довольно просто и поэтому оставляется читателям (есть возможность его прочитать, например, в [5]). Но построение быстрых алгоритмов кодирования и декодирования совсем непросто и мы вынуждены опять отослать читателя, например, к [5].

Недостатком кодов Рида – Соломона является ограниченность их блоковой длины  $n \leq q$ . Далее будет описано построение некоторых классов линейных кодов с произвольной блоковой длиной, у которых и кодовое расстояние, и сложность алгоритма декодирования, и число исправляемых им ошибок растут линейно с ростом блоковой длины. Некоторые из этих кодов близки к границе Синглтона.

Методы построения этих кодов основаны на применении графов-расширителей и некоторых простых результатов линейной алгебры и не требуют знания теории кодирования.

## 2. СПЕКТР ГРАФА И ГРАФЫ-РАСШИРИТЕЛИ

Изучение реберного расширения обычных  $d$ -регулярных графов можно свести вершинному расширению двудольных право- $d$ -регулярных и лево- $2$ -регулярных графов, если заменить обычный граф  $G = (V, E)$  на двудольный граф с долями  $V, E$ , в котором  $v \in V$  и  $e \in E$  соединяются ребром, если и только если в обычном графе  $G$   $v$  и  $e$  инцидентны. Далее для обычных графов рассматривается реберное расширение.

### 2.1. СПЕКТР ГРАФА

Сопоставим произвольному  $d$ -регулярному графу  $G$  его матрицу смежности вершин, т. е. матрицу, элементы которой  $A(G)_{i,j}$  равны числу ребер, соединяющих  $i$  и  $j$ . Эта матрица действительная и симметрическая,

<sup>3)</sup> Это фамилия, а не термин!

<sup>4)</sup> Бинарные коды не достигают этой границы.

поэтому все ее собственные значения действительны. Обозначим их в убывающем порядке  $\lambda_1 \geq \dots \geq \lambda_n$ . Заметим, что  $Ae = de$ , где  $e = (1, \dots, 1)$ , потому что в каждой строке матрицы сумма чисел равна  $d$ , и очевидно для любого  $v$

$$|(Av)_i| = \left| \sum_{j=1}^n A_{ij}v_j \right| \leq d \max |v_i| = d|v|,$$

поэтому для любого  $\lambda_j$  и  $AV_j = \lambda_j V_j$  имеем  $|AV_j| = |\lambda_j| |V_j| \leq d|V_j|$ , откуда  $|\lambda_j| \leq d = \lambda_1$ . Кратность числа  $\lambda_1$  равна числу компонент связности графа. Поэтому граф связан тогда и только тогда, когда  $\lambda_1 > \lambda_2$ .

Граф является двудольным тогда и только тогда, когда  $\lambda_n = -\lambda_1$ . В этом случае все его собственные значения разбиваются на пары противоположных по знаку.

Действительно, рассматривая любую его компоненту связности и предполагая, что  $Av = -dv$ , имеем

$$d|v_i| = |(Av)_i| = \left| \sum_{j=1}^n A_{ij}v_j \right| \leq \sum_{j=1}^n A_{ij}|v_j|,$$

выбирая  $i$  так, чтобы  $|v_i| = |v| = \max_j |v_j|$ , получаем, что

$$d|v| = d|v_i| \leq \sum_{j=1}^n A_{ij}|v_j| \leq \sum_{j=1}^n A_{ij}|v| = d|v|,$$

откуда

$$d|v_i| = \sum_{j=1}^n A_{ij}|v_j| = \sum_{j \in \Gamma(i)} A_{ij}|v_j| \sum_{j=1}^n A_{ij}|v|,$$

значит, во всех вершинах графа, соседних с вершиной в которой  $v$  максимально по модулю, модуль  $v$  имеет то же значение. Двигаясь по вершинам компоненты связности, получаем, что на них везде модуль  $v$  одинаков. Рассматривая множества вершин, в которых значения  $v$  равны, получаем два непересекающихся подмножества рассматриваемой компоненты. Они являются долями двудольного графа, так как внутри них не бывает ребер, иначе для одной из вершин получилось бы неравенство

$$\pm d|v| = -dv_i = (Av)_i = \sum_{j=1}^n A_{ij}v_j = \sum_{j \in \Gamma(i)} A_{ij} \pm |v| > -dv_i.$$

Так как компоненты двудольны, то и весь граф двудолен. Если же он двудолен, то его матрица имеет вид

$$A = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}.$$

Если

$$\lambda v_i = (Av)_i = \sum_{j=1}^n A_{ij}v_j,$$

то, меняя знак у  $v$  на всех вершинах одной доли, получаем, что

$$-\lambda v_i = (Av)_i = \sum_{j=1}^n A_{ij}v_j,$$

т. е.  $-\lambda$  — тоже собственное значение.

Разность  $d - \lambda_2 = \lambda_1 - \lambda_2$  называется *спектральным зазором*. Все графы, имеющие положительный спектральный зазор, являются расширителями. Обозначим  $\mu$  модуль следующего по абсолютной величине собственного значения после максимального значения  $d$  (по определению  $\mu = \max\{|\lambda_2|, |\lambda_n|\}$ , если граф не двудольный, и  $\mu = \lambda_2$ , если он двудольный).

В [6, 7] доказано<sup>5)</sup>, что для любого  $d$ -регулярного  $n$ -вершинного графа справедливо неравенство  $\mu \geq \sqrt{2d-1} - \epsilon_n$ , где  $\epsilon_n \rightarrow 1$ .

Доказательство этого достаточно сложное. В [8] приведена более слабая оценка  $\mu \geq \sqrt{d}(1 - \epsilon_n)$  со следующим простым доказательством.

Пусть  $A$  — матрица смежности вершин графа. Ясно, что диагональные элементы матрицы  $A^k$  есть число путей длины  $k$ , начинающихся и заканчивающихся в соответствующей вершине графа. В частности, диагональные элементы матрицы  $A^2$  не меньше  $d$ , так как всегда есть  $d$  тривиальных циклов длины 2, проходящих через любую вершину. Поэтому след  $\text{tr}(A^2) \geq nd$ . С другой стороны, матрица  $A^2$  имеет собственные числа  $\lambda_i^2$ , поэтому

$$\text{tr}(A^2) = \sum_{i=1}^n \lambda_i^2 = d^2 + \sum_{i=2}^n \lambda_i^2 \leq d^2 + \mu^2(n-1),$$

откуда  $\mu^2 \geq d(n-d)/(n-1) = d(1 - (d-1)/(n-1)) = d(1 - \epsilon_n)$ .

В [6, 7, 9]  $d$ -регулярные графы, для которых  $\mu = \sqrt{2d-1}$ , и двудольные двусторонне  $d$ -регулярные графы, для которых  $\lambda_2 = \sqrt{2d-1}$ , названы *графами Рамануджана*. Другими словами, графами Рамануджана называются графы с наибольшим спектральным зазором. Эти графы достигают наименьшей границы  $\mu = \sqrt{2d-1}$  для  $d$ -регулярных графов. Построены они были независимо в [6, 13] для случая  $d = p + 1$ , где  $p$  — простое. В [14] эти результаты были перенесены на случай  $d = p^k + 1$ . Для других  $d$  графы Рамануджана неизвестны.

Примеры этих графов строятся не очень сложно, но сложно доказываются их правильность. В [7] это занимает почти всю книгу<sup>6)</sup>.

Согласно [6, 13] графом Рамануджана является граф Кэли для группы  $PGL(2, \mathbb{Z}_q)$ , т. е. проективной линейной группы матриц второго порядка над полем  $\mathbb{Z}_q$  со специально выбранным  $(p+1)$ -элементным множеством  $S$ , где  $p$  — простое, не равное  $q$ .

Пусть  $G$  — группа и  $S$  — ее подмножество. *Граф Кэли*  $C(G, S)$  — это ориентированный граф, имеющий  $G$  множеством вершин, в котором упорядоченная пара  $(g, h)$  является ребром, если  $g = hs$ ,  $s \in S$ . Если  $S$  замкнуто относительно инвертирования, т. е.  $s \in S \Rightarrow s^{-1} \in S$ , то граф  $C(G, S)$  можно рассматривать как обычный неориентированный граф. Очевидно, он будет  $|S|$ -регулярным. Если  $S$  — порождающее  $G$  множество, то он будет связным, так как для любых  $g, h$  найдутся  $s_i \in S$ , такие что  $g = hs_1 \dots s_m$ , и эти вершины связаны путем  $h, hs_1, hs_1s_2, \dots, hs_1 \dots s_m = g$ . Очевидно, верно и обратное. Очевидно также, что графы Кэли вершинно транзитивны, т. е.

<sup>5)</sup>Эта теорема принадлежит Алону и Бошпане.

<sup>6)</sup>Про графы там только третья глава, но доказательство опирается на результаты первых двух глав о модулярных формах, в которых решается проблема Рамануджана, откуда видимо и произошло название графов.

любую вершину можно перевести в любую другую подходящим автоморфизмом графа (т. е. перестановкой вершин, переводящей рёбра в рёбра).

Группа  $PGL(2, \mathbb{Z}_q)$  определяется так: берем группу  $GL(2, \mathbb{Z}_q)$  невырожденных матриц второго порядка и вычисляем ее фактор-группу по подгруппе скалярных матриц, т. е. матриц, кратных единичной (эти матрицы образуют центр группы, так как коммутируют со всеми матрицами, а центр очевидно является нормальным делителем и фактор-группа действительно существует). Другими словами, матрицы разбиваются на классы эквивалентности по отношению пропорциональности матриц друг другу в обычном естественном смысле. Читатель может проверить, что  $|PGL(2, \mathbb{Z}_q)| = q(q^2 - 1)$ . Подробности см. [6, 7, 9].

Через второе по величине собственное число  $\mu$  графа можно оценить среднюю степень его подграфа, порожденного данным подмножеством вершин  $S$ .

В [8] обобщенный вариант соответствующей леммы из [10] дан в таком виде: для  $d$ -регулярного графа  $G$  и любых подмножеств  $S, T$  его вершин

$$\left| |E(S, T)| - \frac{d|S||T|}{n} \right| \leq \mu \sqrt{|S||T|},$$

где  $|E(S, T)|$  — число ребер, идущих из  $S$  в  $T$ . При этом ребра, соединяющие вершины  $S \cup T$ , учитываются дважды, так как в [8]  $E(S, T)$  определяется как множество ориентированных ребер, идущих из  $S$  в  $T$  (неориентированные ребра понимаются как пары ребер, ориентированных в разных направлениях).

Эта лемма в [8] названа смешанной леммой о расширителях. В [11] доказано следующее ее обращение.

Если  $G$  —  $d$ -регулярный граф, для которого всегда выполнено неравенство

$$\left| |E(S, T)| - \frac{d|S||T|}{n} \right| \leq \rho \sqrt{|S||T|}$$

при некоторой положительной константе  $\rho$ , то  $\mu(G) = O(\rho(1 + \log(\frac{d}{\rho})))$ , причем оценка точная.

Из леммы о расширителях следует, что для  $n$ -вершинного  $d$ -регулярного расширителя с  $\mu < d/2$  и любого множества вершин  $S$ ,  $|S| \leq n/2$ , степень вершинного расширения

$$h(S) = \frac{|\Gamma(S)|}{|S|} \geq 1 - \frac{\mu}{d} - \frac{|S|}{n},$$

т. е. число соседей  $|\Gamma(S)|$  у множества  $S$  не меньше  $|S|(1 - \mu/d - |S|/n)$ .

Действительно, число ребер  $|E(S, S)| = 2|E(S)|$  удовлетворяет неравенству

$$|E(S, S)| \leq \frac{d|S|^2}{n} + \mu|S|,$$

откуда

$$|E(S, S)|/|S| \leq \frac{d|S|}{n} + \mu,$$

значит, число ребер, идущих из  $S$  вовне его, равно

$$|E(S, \bar{S})|/|S| \geq d \left(1 - \frac{|S|}{n}\right) - \mu,$$

так как  $E(S, S) + E(S, \bar{S}) = E(S, V) = d|S|$  в силу регулярности графа. Но очевидно  $|\Gamma(S)| \geq E(S, \bar{S})/d$ , потому что из каждого соседа выходит не более  $d$  ребер, идущих в  $S$ . Поэтому

$$|\Gamma(S)| \geq E(S, \bar{S})/d = |S| |E(S, \bar{S})| / (d|S|) \geq |S| \left(1 - \frac{|S|}{n} - \frac{\mu}{d}\right).$$

В [12] доказан следующий вариант леммы о расширителях.

Пусть  $G$  — двусторонне  $d$ -регулярный двудольный граф с  $n$  вершинами в каждой доле. Пусть  $S, T$  — подмножества вершин разных долей. Средняя степень  $d_{S,T}$  подграфа с долями  $S, T$  равна  $|2E(S, T)| / (|S| + |T|)$ . Тогда

$$d_{S,T} \leq \frac{d}{n} \frac{2|S||T|}{|S| + |T|} + \mu - \frac{\mu}{n} \frac{|S|^2 + |T|^2}{|S| + |T|}.$$

Эквивалентная формулировка такова:

$$E(S, T) \leq \frac{d|S||T|}{n} + \frac{\mu(|S| + |T|)}{2} - \frac{\mu(|S|^2 + |T|^2)}{2n}.$$

Ниже будет доказано ее обобщение.

### 3. ГРАФЫ-РАСШИРИТЕЛИ И КОДЫ

Далее пойдет речь о применении графов-расширителей в теории кодирования, точнее в построении линейных кодов с большим относительным расстоянием, большой пропускной способностью и линейной сложностью декодирования. Эта тематика тесно связана с так называемыми кодами с низкой плотностью проверок на четность и каскадными кодами, получившими важные промышленные применения в самое последнее время. Существенный вклад в эту теорию еще в семидесятых годах внесли московские кодировщики (см., например, [15, 16]).

#### 3.1. РЕГУЛЯРНЫЕ $(c, d)$ ГРАФЫ

Так называются двудольные графы с долями  $M, N$ ,  $|M| = m$ ,  $|N| = n$  у которых из левой доли  $M$  выходит по  $c$  ребер из каждой вершины, а из правой доли  $N$  — по  $d$  ребер. Очевидно  $cm = dn$ . В [19] при изучении кодов Таннера, о которых будет речь дальше, была получена следующая лемма.

Пусть  $G$  указанный выше граф и  $S \subset M, N \subset N$ . Обозначим  $E(S, T)$  множество ребер из  $S$  в  $T$ . Тогда

$$\left| |E(S, T)| - \frac{d|S||T|}{m} \right| \leq \frac{\mu}{2} \left( |S| + |T| - \frac{|S|^2}{m} - \frac{|T|^2}{n} \right).$$

При  $c = d$  очевидно  $m = n$  и из этого неравенства получается неравенство, приведенное в конце раздела 2.1.

На самом деле можно доказать более сильное неравенство.

$$\left| |E(S, T)| - \frac{d|S||T|}{m} \right| \leq \mu \left( \left( |S| - \frac{|S|^2}{m} \right) \left( |T| - \frac{|T|^2}{n} \right) \right)^{1/2}.$$

Из него с помощью неравенства  $\sqrt{ab} \leq (a+b)/2$  вытекает сформулированное выше неравенство [19].

Докажем его. Матрица графа имеет вид

$$A(G) = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix},$$

где  $A$  есть  $(m, n)$ -матрица, такая, что  $a_{i,j} = 1$  если и только если вершина  $i \in M$  соединяется ребром с вершиной  $j \in N$ . Сначала проверим, что у этой матрицы максимальное собственное значение равно  $\sqrt{cd}$  и собственный вектор (ненормированный) равен

$$(\sqrt{c}, \dots, \sqrt{c}, \sqrt{d}, \dots, \sqrt{d}),$$

где вначале идут  $m$  равных чисел, а потом  $n$  равных чисел. Для любого вектора  $v$  длины  $n + m$  и для  $i \leq m$  очевидно

$$|(A(G)v)_i| = \left| \sum_{j=1}^n A_{i,j} v_{m+j} \right| \leq c \max_j |v_{m+j}|,$$

так как

$$\sum_{j=1}^n A_{i,j} = c.$$

Аналогично для  $i > m$

$$|(A(G)v)_i| = \left| \sum_{j=1}^m A_{j,i-m} v_j \right| \leq d \max_{j \leq m} |v_j|,$$

так как

$$\sum_{j=1}^m A_{j,i} = d.$$

Поэтому для любого  $\lambda_k$  и  $AV_k = \lambda_k V_k$  имеем

$$|\lambda_k| \max_{i \leq m} |V_{k,i}| = \max_{i \leq m} |(AV_k)_i| \leq c \max_{i \leq n} |V_{k,m+i}|$$

и аналогично

$$|\lambda_k| \max_{i > m} |V_{k,i}| = \max_{i > m} |(AV_k)_i| \leq d \max_{i > m} |V_{k,i-m}|$$



откуда

$$|\lambda_k| \leq \frac{c \max_{i \leq n} |V_{k,m+i}|}{\max_{i \leq n} |V_{k,i}|},$$

и аналогично

$$|\lambda_k| \leq \frac{d \max_{i > m} |V_{k,i-m}|}{\max_{i > m} |V_{k,i}|}.$$

Перемножая эти неравенства и выполняя очевидное сокращение дробей, получаем что  $|\lambda_k|^2 \leq cd$ . Вычисляя для вектора

$$v = (\underbrace{\sqrt{c}, \dots, \sqrt{c}}_m, \underbrace{\sqrt{d}, \dots, \sqrt{d}}_n)$$

вектор  $Av$ , рассматривая по отдельности случаи  $i \leq m$  и  $i > m$ , получаем, что в первом случае

$$|(A(G)v)_i| = \left| \sum_{j=1}^n A_{i,j} v_{m+j} \right| = c\sqrt{d} = \sqrt{cd}\sqrt{c} = \sqrt{cd}v_i,$$

и аналогично во втором случае. Поэтому  $v$  — собственный вектор для значения  $\sqrt{cd}$ . Аналогично проверяется, что противоположному собственному значению  $-\sqrt{cd}$  отвечает вектор

$$v = (\underbrace{\sqrt{c}, \dots, \sqrt{c}}_m, \underbrace{-\sqrt{d}, \dots, -\sqrt{d}}_n).$$

Пусть  $1_S$  — характеристический вектор множества  $S$ , т. е. он равен 1 на вершинах из  $S$  и нулю вне их. Аналогично вводится вектор  $1_T$ .

Пусть  $v_i$  — вектора ортонормированного базиса, состоящего из собственных векторов матрицы графа  $A$ . Вектор, соответствующий числу  $\lambda_1 = \sqrt{cd}$ , после нормировки будет равен

$$v_1 = \frac{1}{\sqrt{2cm}} (\underbrace{\sqrt{c}, \dots, \sqrt{c}}_m, \underbrace{\sqrt{d}, \dots, \sqrt{d}}_n),$$

так как норма исходного вектора равна  $cm + dn = 2cm$ . Аналогично, вектор, соответствующий числу  $\lambda_{n+m} = -\sqrt{cd}$ , равен

$$v_{n+m} = \frac{1}{\sqrt{2cm}} (\underbrace{\sqrt{c}, \dots, \sqrt{c}}_m, \underbrace{-\sqrt{d}, \dots, -\sqrt{d}}_n).$$

Рассмотрим вектора

$$1_S - \alpha_1 v_1 - \alpha_{n+m} v_{n+m}, 1_T - \beta_1 v_1 - \beta_{n+m} v_{n+m}.$$

У первого из них координаты при  $i \in S$  равны

$$1 - \frac{|S|}{m},$$

при  $i \notin S, i \leq m$  равны

$$-\frac{|S|}{m},$$

при  $i > m$  равны 0. Поэтому его евклидова норма равна

$$\|1_S - \alpha_1 v_1 - \alpha_{n+m} v_{n+m}\|_2 = \left( \left(1 - \frac{|S|}{m}\right)^2 |S| + (m - |S|) \frac{|S|^2}{m^2} \right)^{1/2} = \left( |S| - \frac{|S|^2}{m} \right)^{1/2}.$$

Аналогично, евклидова норма вектора  $1_T - \beta_1 v_1 - \beta_{n+m} v_{n+m}$  равна

$$\left( |T| - \frac{|T|^2}{n} \right)^{1/2}.$$

Очевидно их координаты в базисе  $\{v_1, \dots, v_{n+m}\}$  равны

$$(0, \alpha_2, \dots, \alpha_{n+m-1}, 0), (0, \beta_2, \dots, \beta_{n+m-1}, 0),$$

поэтому, используя неравенство Коши-Буняковского, имеем

$$\begin{aligned} \sum_{i=2}^{n+m-1} |\alpha_i \beta_i| &\leq \left( \sum_{i=2}^{n+m-1} |\alpha_i|^2 \right)^{1/2} \left( \sum_{i=2}^{n+m-1} |\beta_i|^2 \right)^{1/2} = \\ &= \|1_S - \alpha_1 v_1 - \alpha_{n+m} v_{n+m}\|_2 \|1_T - \beta_1 v_1 - \beta_{n+m} v_{n+m}\|_2 = \\ &= \left( \left( |S| - \frac{|S|^2}{m} \right) \left( |T| - \frac{|T|^2}{n} \right) \right)^{1/2}. \end{aligned}$$

Отсюда

$$\begin{aligned} \left| |E(S, T)| - \frac{d|S||T|}{m} \right| &\leq \mu \left( \sum_{i=2}^{n+m-1} |\alpha_i|^2 \right)^{1/2} \left( \sum_{i=2}^{n+m-1} |\beta_i|^2 \right)^{1/2} = \\ &= \mu \|1_S - \alpha_1 v_1 - \alpha_{n+m} v_{n+m}\|_2 \|1_T - \beta_1 v_1 - \beta_{n+m} v_{n+m}\|_2 = \\ &= \mu \left( \left( |S| - \frac{|S|^2}{m} \right) \left( |T| - \frac{|T|^2}{n} \right) \right)^{1/2}. \end{aligned}$$

### 3.2. Коды ТАННЕРА

Они были введены Таннером в [17], но получили популярность после статьи [18]. Результаты [18] улучшены в [12], и далее в [20, 21]. Результаты [12] перенесены на случай  $(c, d)$ -регулярных двудольных графов в [19].

Дадим определение кода Таннера. Это будет линейный код над произвольным полем  $GF(q)$ . Пусть  $G$  есть  $(c, d)$ -регулярный двудольный граф с долями  $M, N$  мощностей  $m, n$  и множеством ребер  $E$  мощности  $e = cm = dn$ . С каждой вершиной  $v \in M$  ассоциируем линейный  $[c, r_1 c, d_1 = \delta_1 c]$ -код  $C_1$ , где  $c$  — его блоковая длина (длина кодовых слов),  $r_1 c$  — его размерность как линейного пространства ( $r_1$  — это пропускная способность), а  $d_1$  — его расстояние ( $\delta_1$  — относительное минимальное расстояние). Аналогично с каждой вершиной  $v \in N$  ассоциируем линейный  $[d, r_2 d, d_2 = \delta_1 d]$ -код  $C_2$ . Занумеруем ребра и каждому из них сопоставим символ  $x_i, 1 \leq i \leq e$ . Символы, соответствующие ребрам, выходящим из вершины  $v \in M$  обозначим  $x_{v(1)}, \dots, x_{v(c)}$ . Из них составим вектор  $E(v) = (x_{v(1)}, \dots, x_{v(c)})$ . Аналогично для любой  $u \in N$  определим вектор  $E(u) = (x_{u(1)}, \dots, x_{u(d)})$ .

Код Таннера  $\mathcal{C}(G, C_1, C_2)$ , ассоциированный с графом  $G$  и кодами  $C_i$ , определяется как линейный код с блоковой длиной  $e = tc = nd$ , определяемый условиями  $\forall v \in M E(v) \in C_1, \forall u \in N E(u) \in C_2$ . Другими словами, он состоит из всех векторов, у которых проекции на множества координат  $E(v), v \in M, E(u), u \in N$  (т.е. соответствующие подвектора) принадлежат кодам  $C_i$  соответствующей длины. Так как проекция линейной комбинации векторов равна линейной комбинации проекций, то код Таннера линеен.

Оценим его размерность. Если в его определении оставить только ограничения, соответствующие вершинам доли  $M$ , т.е. рассмотреть код определяемый условиями  $\forall v \in M E(v) \in C_1$ , то его размерность будет равна  $r_1tc = r_1e$ , так как подвектора  $E(v)$  попарно не имеют общих координат и указанный набор условий определяет код, который можно рассматривать как прямое произведение  $t$  экземпляров кода  $C_1$ .

Аналогично, код, определяемый условиями  $\forall u \in N E(u) \in C_2$ , имеет размерность  $r_2e$ .

Код  $\mathcal{C}(G, C_1, C_2)$  равен пересечению этих кодов. Известна лемма о том, что пересечение подпространств  $L_i$  размерности  $D_i$  пространства  $L$  размерности  $D$  имеет размерность не меньше  $D_1 + D_2 - D$ .

Применяя эту лемму, получаем, что размерность кода  $\mathcal{C}(G, C_1, C_2)$  не меньше  $r_1e + r_2e - e = (r_1 + r_2 - 1)e$ .

Что касается расстояния кода Таннера, то оценки [18] улучшены в [12] и обобщены в [19]. Последний результат таков: если  $d_i > \mu(G)/2$ , где  $\mu(G)$  — второе по величине собственное число графа  $G$ , то

$$d(\mathcal{C}(G, C_1, C_2)) \geq \frac{m}{d} \left( d_1 d_2 - \frac{\mu(G)}{2} (d_1 + d_2) \right),$$

а если в терминах относительных расстояний, то

$$\begin{aligned} \delta(\mathcal{C}(G, C_1, C_2)) &= \frac{d(\mathcal{C}(G, C_1, C_2))}{e} \geq \frac{1}{cd} \left( d_1 d_2 - \frac{\mu(G)}{2} (d_1 + d_2) \right) = \\ &= \delta_1 \delta_2 - \frac{\mu(G)}{2} \left( \frac{\delta_1}{d} + \frac{\delta_2}{c} \right). \end{aligned}$$

### 3.3. Нижняя граница для минимального расстояния кода ТАННЕРА

Сначала получим следствие границы для плотности ребер в двудольном графе

$$\left| |E(S, T)| - \frac{d|S||T|}{m} \right| \leq \mu \left( \left( |S| - \frac{|S|^2}{m} \right) \left( |T| - \frac{|T|^2}{n} \right) \right)^{1/2}$$

в более удобном и компактном виде. Введем обозначения  $s = |S|/m$ ,  $t = |T|/n$  для относительного числа вершин в  $S, T$ ,  $e = mc = nd = \sqrt{mncd}$  для числа ребер в графе  $G$ ,  $\gamma(G) = \mu(G)/\sqrt{cd}$  для нормализованного второго собственного значения (очевидно  $0 < \gamma(G) < 1$ ). Тогда для относительного числа ребер справедливо неравенство

$$\begin{aligned} \left| \frac{|E(S, T)|}{e} - st \right| &= \left| \frac{|E(S, T)|}{nd} - st \right| = \left| \frac{|E(S, T)|}{e} - \frac{|S||T|}{mn} \right| \leq \\ &\leq \frac{\mu}{\sqrt{cd}} \left( \left( \frac{|S|}{m} - \frac{|S|^2}{m^2} \right) \left( \frac{|T|}{n} - \frac{|T|^2}{n^2} \right) \right)^{1/2} = \\ &= \frac{\mu}{\sqrt{dc}} \sqrt{(s - s^2)(t - t^2)} = \gamma \sqrt{(s - s^2)(t - t^2)}. \end{aligned}$$

Так как очевидно при  $0 < s, t \leq 1$

$$\sqrt{(s - s^2)(t - t^2)} = \sqrt{st} \sqrt{(1 - s)(1 - t)} \leq \sqrt{st}(1 - \sqrt{st}) = \sqrt{st} - st,$$

потому что

$$\sqrt{(1 - s)(1 - t)} + \sqrt{st} \leq \frac{1 - s + 1 - t}{2} + \frac{s + t}{2} = 1,$$

то из предыдущего неравенства следует, что

$$\left| \frac{|E(S, T)|}{e} - st \right| \leq \gamma \sqrt{(s - s^2)(t - t^2)} \leq \gamma(\sqrt{st} - st),$$

в частности

$$\frac{|E(S, T)|}{e} \leq st + \gamma(\sqrt{st} - st) = (1 - \gamma)st + \gamma\sqrt{st}.$$

В конце раздела 3.2 было обещано доказать, что если  $d_i > \mu/2$ , то

$$d(\mathcal{C}(G, C_1, C_2)) \geq \frac{m}{d} \left( d_1 d_2 - \frac{\mu}{2}(d_1 + d_2) \right).$$

Докажем более сильное неравенство

$$\frac{d(\mathcal{C}(G, C_1, C_2))}{e} \geq D = \frac{1}{cd} \frac{d_1 d_2 - \mu \sqrt{d_1 d_2}}{1 - \gamma},$$

где  $e = mc = nd$  — число ребер в  $G$ .

Обозначая относительное расстояние  $d(\mathcal{C}(G, C_1, C_2))/e$  через  $\delta$  и используя обозначения для относительных расстояний  $\delta_1 = d_1/c$ ,  $\delta_2 = d_2/d$ , можно переписать это неравенство как

$$\delta \geq \frac{\delta_1 \delta_2 - \gamma \sqrt{\delta_1 \delta_2}}{1 - \gamma},$$

где  $0 < \gamma = \mu/\sqrt{cd} < 1$ . В частном случае  $c = d$  это неравенство получено в [21]. Заметим, что при  $\gamma \rightarrow 0$

$$\delta \geq \frac{\delta_1 \delta_2 - \gamma \sqrt{\delta_1 \delta_2}}{1 - \gamma} = \delta_1 \delta_2 \left( 1 - \gamma((\delta_1 \delta_2)^{-1/2} - 1) - O(\gamma^2) \right).$$

Так как код Таннера линейен, то достаточно взять ненулевое кодовое слово  $\mathbf{x} \in \mathcal{C}(G, C_1, C_2)$ , рассмотреть соответствующее ему множество ребер  $X = \{e_i \in E : x_i \neq 0\}$

и доказать, что  $|X|/e \geq D$ . Рассмотрим минимальный подграф, содержащий множество  $X$ , т. е. возьмем множества вершин

$$S = \{u \in M : \exists v \in N (u, v) \in X\} \subset M, T = \{v \in N : \exists u \in M (u, v) \in X\} \subset N,$$

и рассмотрим порожденный множеством вершин  $S \cup T$  подграф с множеством ребер  $E(S, T)$ ,  $X \subset E(S, T)$ . Очевидно  $|X| \leq |E(S, T)|$ . Так как  $\mathbf{x}$  — кодовый вектор, то по определению кода Таннера его подвектора, образованные координатами из множеств  $E(v)$ ,  $v \in S$ , принадлежат коду  $C_1$ , а подвектора, образованные координатами из множеств  $E(u)$ ,  $u \in T$ , принадлежат коду  $C_2$ , причем они не равны нулю в силу определения  $S, T$ . Поэтому подвектора, образованные координатами из множеств  $E(v)$ ,  $v \in S$ , содержат не менее  $d_1 = \delta_1 c$  ненулевых координат, значит, общее число ненулевых координат в  $\mathbf{x}$ , т. е.  $|X|$ , не меньше  $d_1|S| = \delta_1 cms = \delta_1 es$ . Аналогично  $|X| \geq \delta_2 et$ . Отсюда очевидно имеем

$$\frac{|X|}{e} \geq \sqrt{\delta_1 \delta_2 st}.$$

В начале этого раздела было доказано, что

$$\frac{|X|}{e} \leq \frac{|E(S, T)|}{e} \leq (1 - \gamma)st + \gamma\sqrt{st}.$$

Поэтому

$$\sqrt{\delta_1 \delta_2 st} \leq \frac{|X|}{e} \leq (1 - \gamma)st + \gamma\sqrt{st},$$

откуда  $\sqrt{st}(\sqrt{\delta_1 \delta_2} - \gamma) \leq (1 - \gamma)st$ , значит

$$\frac{\sqrt{\delta_1 \delta_2} - \gamma}{1 - \gamma} \leq \sqrt{st},$$

следовательно

$$D = \frac{\delta_1 \delta_2 - \gamma \sqrt{\delta_1 \delta_2}}{1 - \gamma} \leq \sqrt{\delta_1 \delta_2 st} \leq \frac{|X|}{e}.$$

### 3.4. ДЕКОДИРОВАНИЕ КОДА ТАННЕРА

Согласно определению кода Таннера  $\mathcal{C}(G, C_1, C_2)$  в  $(c, d)$ -регулярном графе  $G$  множество ребер  $E$  разбито двумя способами на подмножества

$$E = \bigcup_{i=1}^m E_{v_i} = \bigcup_{j=1}^n E_{u_j}.$$

Так как код линейный, то достаточно рассмотреть случай, когда в нулевом кодовом слове произошли ошибки и получилось ненулевое слово  $y \in GF(q)^e$ ,  $e = mc = nd$ . Итеративный алгоритм устранения ошибок выглядит следующим образом.

Для каждой вершины  $v \in M$  берем подвектор  $E(v) = (x_{v(1)}, \dots, x_{v(c)})$  и, так как он принадлежит  $[c, r_1c, \delta_1c]$ -коду  $C_1$ , то применяем алгоритм декодирования для этого кода. Эти процедуры декодирования для всех  $v \in M$  можно делать параллельно. В случае, если вес подвектора  $E(v)$

меньше  $d_1/2$ , то в результате декодирования получается правильное слово, т. е. нулевое, в противном случае может получиться неправильное слово, т. е. ненулевое. Из полученных подслов составляем слово  $z \in GF(q)^e$  — результат первого шага алгоритма декодирования кода Таннера. К полученному слову применяем аналогичную параллельную процедуру, декодирующую подслово  $E(u)$ ,  $u \in N$ , соответствующие вершинам правой доли  $N$  графа  $G$ . В результате получим слово  $w \in GF(q)^e$ . Если  $w = z$ , то алгоритм заканчивает работу и выдает ответ  $w = z$ . Если нет, то к слову  $w$  применяем опять процедуру параллельного декодирования вначале всех подслов  $E(v)$ ,  $v \in M$ , а потом — всех подслов  $E(u)$ ,  $u \in N$  и т. д.

Можно показать, что при любом  $\alpha$ ,  $0 < \alpha < 1$ , таком, что  $\sqrt{\delta_1 \delta_2} > 2\gamma/\alpha$ , этот алгоритм заканчивает работу за

$$\frac{\log_2(mn)}{2 \log_2 \frac{1}{\alpha}} + O_{\mathcal{C}, \alpha}(1), \quad \mathcal{C} = \mathcal{C}(G, C_1, C_2)$$

шагов и правильно декодирует любое слово  $y \in GF(q)^e$  с числом ненулевых символов не более

$$e \frac{\alpha \delta_1 \delta_2 - 2\gamma \sqrt{\delta_1 \delta_2}}{4(1-\gamma)}.$$

Мультипликативные константы в этих оценках зависят от алгоритмов декодирования кодов  $C_1, C_2$  и определяются величинами  $d_i, c, d$ . Числа  $m, n$  можно выбирать при этом сколь угодно большими.

Заметим, что при  $\gamma \rightarrow 0$

$$\frac{\alpha \delta_1 \delta_2 - 2\gamma \sqrt{\delta_1 \delta_2}}{4(1-\gamma)} = \frac{\alpha}{4} \delta_1 \delta_2 \left( 1 - \gamma \left( \frac{2}{\alpha} (\delta_1 \delta_2)^{-1/2} - 1 \right) - O(\gamma^2) \right).$$

Ограничение на вес вектора  $Y$  (число исправляемых ошибок) имеет вид

$$\frac{\alpha \delta_1 \delta_2 - 2\gamma \sqrt{\delta_1 \delta_2}}{4(1-\gamma)} \geq \frac{|Y|}{e}.$$

Для расстояния кода Таннера выше была получена оценка

$$\frac{d(\mathcal{C}(G, C_1, C_2))}{e} \geq \frac{\delta_1 \delta_2 - \gamma \sqrt{\delta_1 \delta_2}}{1-\gamma}.$$

Указанное выше ограничение чуть более, чем в четыре раза хуже. Если взять в качестве  $G$  граф с  $\gamma(G) \rightarrow 0$  (например, при  $c = d$  — граф Рамануджана), то асимптотически отношение расстояния кода к числу исправляемых этим алгоритмом ошибок будет равно четырем. Оценка для числа итераций имеет вид

$$k = \frac{\log_2 \left( \frac{2|Y|}{\sqrt{d_1 d_2}} \right)}{\log_2 \frac{1}{\alpha}} \leq \frac{\log_2 \left( \sqrt{mn} \frac{\alpha \sqrt{\delta_1 \delta_2} - 2\gamma}{2(1-\gamma)} \right)}{\log_2 \frac{1}{\alpha}} = \frac{\log_2(mn)}{2 \log_2 \frac{1}{\alpha}} + O_{\mathcal{C}, \alpha}(1).$$

## 3.5. КОНКРЕТНЫЙ ПРИМЕР КОДА ТАННЕРА

В [12] предложен следующий пример кода Таннера. В качестве  $G$  берем граф Рамануджана (построенный методом [6, 7]). У этого графа  $m = n = q$  — простое число,  $c = d = p + 1$ ,  $p = 223$ , согласно [7]  $q$  следует выбрать так, чтобы  $p$  не было квадратом по модулю  $q$ . Соответствующий граф имеет  $q(q^2 - 1)$  вершин и является двудольным и

$$\gamma(G) = \mu(G)/(p + 1) = 2\sqrt{p}/(p + 1) = \sqrt{223}/112 < 15/112.$$

В качестве кодов  $C_i$  берем [224, 115, 30]-ВСН-код<sup>7)</sup>. Тогда относительное расстояние

$$\frac{d(\mathcal{C}(G, C_1, C_2))}{e} = \delta \geq \frac{\delta_1 \delta_2 - \gamma \sqrt{\delta_1 \delta_2}}{1 - \gamma} = (30/224) \frac{30/224 - \gamma}{1 - \gamma} > 9 \cdot 10^{-5},$$

а пропускная способность (т.е. размерность, деленная на длину блока) не меньше  $2 \cdot 115/224 - 1 = 6/224$ . Блоковая длина  $(p + 1)q(q^2 - 1) > 2.5 \cdot 10^9$ . Условие  $\sqrt{\delta_1 \delta_2} > 2\gamma$  для него выполняется.

## 4. ЭКСПАНДЕРНЫЕ КОДЫ С ПОЧТИ МИНИМАЛЬНЫМ РАССТОЯНИЕМ

В [20] показано, что для любой пропускной способности  $R \in (0, 1]$  и  $\epsilon < \epsilon_0$  можно построить бесконечное семейство кодов с пропускной способностью не менее  $R$  над алфавитом размера  $2^{O(\log(1/\epsilon)/(R\epsilon^4))}$ , и относительным минимальным расстоянием не меньше  $1 - R - \epsilon$ . Кодирование осуществляется в линейное время со сложностью  $(1/\epsilon)^{O(1)}$  на символ. Декодировать можно долю ошибок, не меньшую  $(1 - R - \epsilon)/2$  от длины кода, в линейное время со сложностью  $(1/\epsilon)^{O(1)}$  на символ.

Рот и Скачек в [21] показали, что в этом результате можно улучшить границу для мощности алфавита до  $2^{O(\log(1/\epsilon)/\epsilon^3)}$ .

Опишем их конструкцию чуть в более общем виде. Вспомним конструкцию кода Таннера  $\mathcal{C}(G, C_1, C_2)$  над алфавитом  $F = GF(q)$ . Здесь  $C_1$  линейный  $[c, r_1c, d_1 = \delta_1c]$ -код, где  $c$  — его блоковая длина, ассоциированный с каждой вершиной  $v \in M$ , и с каждой вершиной  $u \in N$  ассоциируем линейный  $[d, r_2d, d_2 = \delta_1d]$ -код  $C_2$ . Занумеруем ребра и каждому из них сопоставим элемент  $x_i \in F$ ,  $1 \leq i \leq e$ . Элементы, соответствующие ребрам, выходящим из вершины  $v \in M$  обозначим  $x_{v(1)}, \dots, x_{v(c)}$ . Из них составим вектор  $E(v) = (x_{v(1)}, \dots, x_{v(c)})$ . Аналогично для любой  $u \in N$  определим вектор  $E(u) = (x_{u(1)}, \dots, x_{u(d)})$ . Код Таннера  $\mathcal{C}(G, C_1, C_2)$  над полем  $GF(q)$ , ассоциированный с графом  $G$  и кодами  $C_i$ , определяется как линейный код с блоковой длиной  $e = mc = nd$ , определяемый условиями  $\forall v \in M E(v) \in C_1, \forall u \in N E(u) \in C_2$ .

Пусть  $\Phi = F^{r_1}$  — новый алфавит. Фиксируем биективное линейное отображение кодирования  $\mathcal{E}$  из  $r_1$ -мерного пространства  $\Phi = F^{r_1}$  над полем  $F$  в кодовое пространство  $C_1$  той же размерности над тем же полем и

<sup>7)</sup>О кодах Боуза-Чоудхури-Хоквингема, см., например, [5].

рассмотрим биективное линейное отображение  $\varphi_{\mathcal{E}} : \mathcal{C} \rightarrow \Phi^m$ , определяемое равенством

$$\varphi_{\mathcal{E}}(c) = (\mathcal{E}^{-1}(E(v)) : v \in M) \in \Phi^m.$$

Образ кода  $\mathcal{C}$  при этом отображение является линейным кодом  $\mathcal{C}_{\Phi}$  над алфавитом  $\Phi$  с блоковой длиной  $m$ . Размерность кода  $\mathcal{C}(G, C_1, C_2)$ , как было показано в разделе 3.2, не меньше  $(r_1 + r_2 - 1)e = (r_1 + r_2 - 1)cm$ . Размерность кода  $\mathcal{C}_{\Phi}$  над алфавитом  $F$  такая же. Блоковая длина над алфавитом  $F$  равна  $r_1cm$ , поэтому пропускная способность над алфавитом  $F$  не меньше

$$(r_1 + r_2 - 1)e/(r_1cm) = (r_1 + r_2 - 1)/r_1 = 1 - 1/r_1 + r_2/r_1.$$

Получим нижнюю оценку для относительного расстояния кода  $\mathcal{C}_{\Phi}$  над алфавитом  $\Phi$ . Очевидно относительное расстояние равно  $d(\mathcal{C}_{\Phi})/m$ . Докажем, что

$$\frac{d(\mathcal{C}_{\Phi})}{m} \geq \frac{\delta_2 - \gamma\sqrt{s/t}}{1 - \gamma}.$$

Так как код  $\mathcal{C}_{\Phi}$  линеен, то достаточно взять произвольное ненулевое кодовое слово

$$\mathbf{c} = (\mathcal{E}^{-1}(E(v)) : v \in M) \in \mathcal{C}_{\Phi} \subset \Phi^m,$$

и оценить снизу его вес, т. е. число ненулевых символов из  $\Phi$  в нем. Рассмотрим соответствующее ему кодовое слово

$$\mathbf{x} = (E(v) : v \in M) \in \mathcal{C}(G, C_1, C_2) \subset F^e,$$

и возьмем соответствующее ему множество ребер  $X = \{e_i \in E : x_i \neq 0\}$ . Вес слова  $\mathbf{c} \in \Phi^m$  очевидно совпадает с числом ненулевых векторов  $E(v)$ ,  $v \in M$ , составляющих вектор  $\mathbf{x} \in \mathcal{C}(G, C_1, C_2)$ . Рассмотрим минимальный подграф, содержащий множество  $X$ , т. е. возьмем множества вершин

$$\begin{aligned} S &= \{u \in M : \exists v \in N (u, v) \in X\} \subset M, \\ T &= \{v \in N : \exists u \in N (v, u) \in X\} \subset N, \end{aligned}$$

и рассмотрим порожденный множеством вершин  $S \cup T$  подграф с множеством ребер  $E(S, T)$ ,  $X \subset E(S, T)$ . Очевидно  $|X| \leq |E(S, T)|$ . Так как  $\mathbf{x}$  — кодовый вектор, то по определению кода  $\mathcal{C}_{\Phi}$  его подвектора, образованные координатами из множеств  $E(v)$ ,  $v \in S$ , принадлежат коду  $C_1$ , а подвектора, образованные координатами из множеств  $E(u)$ ,  $u \in T$ , принадлежат коду  $C_2$ , причем они не равны нулю в силу определения  $S, T$ , к тому же число ненулевых векторов  $E(v)$ ,  $v \in M$ , в точности равно  $|S|$ , и число ненулевых векторов  $E(u)$ ,  $u \in N$ , в точности равно  $|T|$ . Поэтому вес слова  $\mathbf{c} \in \Phi^m$  равен  $|S|$ , откуда для оценки снизу  $d(\mathcal{C}_{\Phi})/m$  достаточно оценить снизу  $|S|/m = s$ .



Так как подвектора, образованные координатами из множеств  $E(v)$ ,  $v \in S$ , содержат не менее  $d_1 = \delta_1 c$  ненулевых координат (это кодовые слова кода  $C_1$ ), значит общее число ненулевых координат в  $\mathbf{x}$ , т.е.  $|X|$ , не меньше  $d_1|S| = \delta_1 cms = \delta_1 es$ . Аналогично  $|X| \geq \delta_2 et$ , где  $t = |T|/n$ . Отсюда имеем

$$\frac{|X|}{e} \geq \max\{\delta_1 s, \delta_2 t\}.$$

В разделе 3.3 было доказано, что

$$\frac{|X|}{e} \leq \frac{|E(S, T)|}{e} \leq (1 - \gamma)st + \gamma\sqrt{st}, \frac{\sqrt{\delta_1 \delta_2} - \gamma}{1 - \gamma} \leq \sqrt{st},$$

следовательно при  $s/t \geq \delta_2/\delta_1$

$$\frac{\delta_2 - \gamma\sqrt{\frac{\delta_2}{\delta_1}}}{1 - \gamma} = \frac{\sqrt{\delta_1 \delta_2} - \gamma}{1 - \gamma} \sqrt{\frac{\delta_2}{\delta_1}} \leq \sqrt{st} \sqrt{s/t} = s,$$

а при  $s/t < \delta_2/\delta_1$  очевидно

$$\delta_2 t \leq \frac{|X|}{e} \leq (1 - \gamma)st + \gamma\sqrt{st},$$

откуда  $\delta_2 t - \gamma\sqrt{st} \leq (1 - \gamma)st$ , значит, и в этом случае

$$\frac{\delta_2 - \gamma\sqrt{\frac{\delta_2}{\delta_1}}}{1 - \gamma} \leq \frac{\delta_2 - \gamma\sqrt{s/t}}{1 - \gamma} \leq s.$$

Теперь можно показать, как для любого достаточно малого  $\epsilon$  выбрать код  $\mathcal{C}_\Phi$  с данной пропускной способностью  $R$  и минимальным относительным расстоянием  $1 - R - \epsilon$ . Для данного  $\epsilon > 0$ , выберем  $8/\epsilon^3 > c = d \geq 4/\epsilon^3$ ,  $2c \geq q > c$ , так чтобы  $c = p + 1$ ,  $p$  — простое число, простое  $p_1 = O(p)$  выбираем так, чтобы  $p$  не было квадратом по модулю  $p_1$ . Тогда согласно [7] соответствующий граф Рамануджана имеет  $n = p_1(p_1^2 - 1) = O(p_1^3) = O(1/\epsilon^9)$  вершин и является двудольным и для него

$$\gamma = \gamma(G) = \mu(G)/(p + 1) = 2\sqrt{p}/(p + 1) < \epsilon^{3/2}.$$

Пусть  $\delta_1 = \epsilon$ ,  $\delta_2 = \delta$ , возьмем любое  $R = r_2 > 1 - \delta$ ,  $r = r_1 > 1 - \delta_1 = 1 - \epsilon$ . Коды  $C_i$  с этими параметрами можно выбрать среди RS-кодов. Тогда код  $\mathcal{C}_\Phi$  имеет пропускную способность не меньше

$$1 - \frac{1}{r_1} + \frac{r_2}{r_1} > 1 - \frac{1}{1 - \epsilon} + \frac{R}{1 - \epsilon} = \frac{R - \epsilon}{1 - \epsilon} > R - \epsilon,$$

и относительное расстояние не меньше

$$\frac{\delta - \gamma\sqrt{\frac{\delta}{\epsilon}}}{1 - \gamma} > \delta - \gamma\sqrt{\frac{\delta}{\epsilon}} > \delta - \epsilon^{3/2}\sqrt{\frac{1}{\epsilon}} = \delta - \epsilon > 1 - R - \epsilon.$$

Таким образом, код  $\mathcal{C}_\Phi$  близок к границе Синглтона при  $\epsilon \rightarrow 0$ . Кроме того,  $|\Phi| = q^{rc} < q^c = 2^{O(\epsilon^{-3} \log 1/\epsilon)}$ .

Однако уже при  $\epsilon = 1/10$  все оценки выходят за пределы разумных. Какой-то смысл имеет, например, выбор  $c = d = 114$ ,  $q = 2^7$ , тогда соответствующий граф Рамануджана при выборе  $p = 113$ ,  $p_1 = 137$  имеет  $n = 137(137^2 - 1) = 2571218$  вершин и является двудольным и для него

$$\gamma = \gamma(G) = \mu(G)/(p+1) = 2\sqrt{p}/(p+1) = \sqrt{113}/57 = 0.186\dots$$

В качестве кода  $C_1$  возьмем  $[114, 57, 58]$ -RS-код, для него  $r_1 = 57/114 = 1/2$ ,  $\delta_1 = 58/114 = 0.508\dots$ , а в качестве  $C_2$  возьмем  $[114, 75, 40]$ -RS-код, для него  $r_2 = 75/114$ ,  $\delta_2 = 40/114 = 20/57$ , тогда условие  $\sqrt{\delta_1\delta_2} > 2\gamma$ , достаточное для хорошей оценки скорости декодирования, выполнено. Тогда код  $\mathcal{C}_\Phi$  имеет длину  $n = 2571218$ , пропускную способность не меньше

$$1 - \frac{1}{r_1} + \frac{r_2}{r_1} = -1 + 75/57 = 18/57 = 6/19,$$

и относительное расстояние не меньше

$$\frac{\delta_2 - \gamma\sqrt{\frac{\delta_2}{\delta_1}}}{1 - \gamma} = \frac{\frac{20}{57} - \frac{\sqrt{113 \cdot 20}}{57\sqrt{29}}}{1 - \frac{\sqrt{113}}{57}} = 0.24\dots$$

Мощность алфавита для этого кода  $|\Phi| = q^{r_1c} = 2^{7 \cdot 57} = 2^{399}$ , т. е. буквы имеют длину 399 бит.

Приведенный пример показывает, что описанные выше конструкции, весьма эффективные теоретически, к сожалению, пока не применимы на практике.

#### 4.1. ДЕКОДИРОВАНИЕ КОДА РОТА-СКАЧЕКА

Оно похоже на декодирование кода Таннера. Согласно определению кода Таннера  $\mathcal{C}(G, C_1, C_2)$  в  $(c, d)$ -регулярном графе  $G$  множество ребер  $E$  разбито двумя способами на подмножества

$$E = \bigcup_{i=1}^m E_{v_i} = \bigcup_{j=1}^n E_{u_j}.$$

Так как код линейный, то достаточно рассмотреть случай, когда в нулевом кодовом слове произошли ошибки и получилось ненулевое слово  $y \in \Phi^m$ . Итеративный алгоритм устранения ошибок выглядит следующим образом.

Для каждой вершины  $v \in M$  берем  $\mathcal{E}^{-1}(E(v)) \in \Phi$  — букву кодового слова, по ней вычисляем подвектор  $E(v) = (x_{v(1)}, \dots, x_{v(c)})$ , и так как он принадлежит  $[c, r_1c, \delta_1c]$  коду  $C_1$ , то применяем алгоритм декодирования для этого кода. Эти процедуры декодирования для всех  $v \in M$  можно делать параллельно. В случае, если вес подвектора  $E(v)$  меньше  $d_1/2 = \delta_1c$ , то в результате декодирования получается правильное слово, т. е. нулевое, в противном случае может получиться ненулевое. Из полученных слов

составляем слово  $z \in GF(q)^e$  — результат первого шага алгоритма. К полученному слову применяем аналогичную параллельную процедуру, декодирующую подслово  $E(u)$ ,  $u \in N$ , соответствующие вершинам правой доли  $N$  графа  $G$ . В результате получим слово  $w \in GF(q)^e$ . Если  $w = z$ , то алгоритм заканчивает работу и выдает ответ  $w = z$ . Если нет, то к слову  $w$  применяем опять процедуру параллельного декодирования вначале всех подслов  $E(v)$ ,  $v \in M$ , а потом — всех подслов  $E(u)$ ,  $u \in N$  и т. д. Покажем, что при  $\sqrt{\delta_1 \delta_2} > 2\gamma$  этот алгоритм заканчивает работу за

$$(\log \sqrt{mn}) / \log(\delta_1 \delta_2 / 4\gamma^2) + O_{\delta_1, \delta_2, \gamma}(1)$$

итераций, выполняет  $O(\sqrt{mn})$  раз процедуры декодирования кодов  $C_i$  и правильно декодирует любое слово  $y \in \Phi^m$  с числом ненулевых символов не более

$$m \frac{\delta_2/2 - \gamma \sqrt{\delta_2/\delta_1}}{1 - \gamma}.$$

Числа  $m, n, c, d$  можно выбирать при этом сколь угодно большими. Тогда можно выбрать граф так, что  $\gamma \rightarrow 0$ , при этом

$$\frac{\delta_2/2 - \gamma \sqrt{\delta_2/\delta_1}}{1 - \gamma} \rightarrow \delta_2/2,$$

если  $\delta_2/\delta_1$  ограничено.

Рассмотрим множество  $Y$  ребер графа  $G$ , соответствующих ненулевым координатам декодируемого вектора  $y \in \Phi^m$ . Выберем минимальные множества вершин  $S \subset M$ ,  $T \subset N$ , такие, что  $Y \subset E(S, T)$ . По условию

$$s = |S|/m < \beta = \frac{\delta_2/2 - \gamma \sqrt{\delta_2/\delta_1}}{1 - \gamma}.$$

Далее, как и в случае кода Таннера, рассмотрим множество  $S_1 \subset S$ , состоящее из вершин степени  $\geq d_1/2$  в подграфе  $(S, T, Y)$  и множество  $W$  ребер, соответствующих ненулевым координатам вектора  $w$ . Тогда  $W \subset E(S_1, T)$ , так как  $W \subset Y \subset E(S, T)$ , а все ребра, выходящие из вершин со степенями, меньшими  $d_1/2$ , не входят во множество  $W$ , так как после декодирования становятся нулевыми (при декодировании слова с весом меньшим  $d_1/2$  из кода  $C_1$  с расстоянием  $d_1$  получается нулевое слово, так как оно есть ближайшее кодовое слово и удалено на расстояние меньше  $d_1/2$ ). Рассмотрим в графе  $E(S_1, T)$  множество  $T_1 \subset T$ , состоящее из вершин степени  $< d_1/2$ . Тогда по аналогичным соображениям имеем, что множество  $Z$  ребер, соответствующих ненулевым координатам вектора  $z$ , содержится в  $E(S_1, T_1)$ .

Поэтому  $\delta_1 s_1 \leq 2((1 - \gamma)s_1 t + \gamma \sqrt{s_1 t})$ , где  $s_1 = |S_1|/m$ ,  $t = |T|/n$ , и аналогично  $\delta_2 t_1 \leq 2((1 - \gamma)s_1 t_1 + \gamma \sqrt{s_1 t_1})$ , где  $t_1 = |T_1|/n$ , а также

$$\sqrt{\delta_1 \delta_2} \sqrt{s_1 t_1} \leq 2((1 - \gamma)st + \gamma \sqrt{st}).$$

Из  $\delta_2 t_1 \leq 2((1 - \gamma)s_1 t_1 + \gamma \sqrt{s_1 t_1})$  следует, что  $\gamma \sqrt{s_1 t_1} \geq t_1(\delta_2/2 - (1 - \gamma)s_1)$ , откуда

$$\sqrt{s_1/t_1} \geq \frac{(\delta_2/2) - (1 - \gamma)s_1}{\gamma} > \frac{(\delta_2/2) - (1 - \gamma)\beta}{\gamma} = \sqrt{\delta_2/\delta_1}.$$

Значит,  $\sqrt{t_1} \leq \sqrt{s_1} \sqrt{\delta_1/\delta_2} \leq \sqrt{\beta} \sqrt{\delta_1/\delta_2}$ , поэтому

$$\sqrt{s_1 t_1} \leq \sqrt{s_1} \sqrt{\beta} \sqrt{\delta_1/\delta_2} \leq \beta \sqrt{\delta_1/\delta_2} = \frac{\sqrt{\delta_1 \delta_2/2} - \gamma}{1 - \gamma}.$$

Выполняя второй шаг итерации, по множествам  $S_1, T_1$  построим множества  $S_2, T_2$  и т. д. Из доказанного выше следует, что

$$\sqrt{s_{i+1} t_{i+1}} \leq \frac{2((1-\gamma)s_i t_i + \gamma \sqrt{s_i t_i})}{\sqrt{\delta_1 \delta_2}}.$$

Полагая для краткости  $x_i = \sqrt{s_i t_i}$ , имеем

$$x_{i+1} \leq f(x_i), \quad f(x) = ax^2 + bx, \quad a = \frac{2(1-\gamma)}{\sqrt{\delta_1 \delta_2}}, \quad b = \frac{2\gamma}{\sqrt{\delta_1 \delta_2}}.$$

Так как положительный корень уравнения  $f(x) = x$  равен

$$\frac{1-b}{a} = \frac{\sqrt{\delta_1 \delta_2/2} - \gamma}{1-\gamma},$$

и  $0 < x_1 = \sqrt{s_1 t_1} < (1-b)/a$ , то

$$x_{i+1} \leq f(x_i) < x_i < x_1 < \frac{1-b}{a}, \quad \frac{x_{i+1}}{x_i} \leq ax_i + b < q = ax_1 + b < 1,$$

откуда  $x_{i+1} \leq x_1 q^i$ , т. е. после каждой итерации алгоритма величина

$$\sqrt{st} = \frac{\sqrt{|S||T|}}{\sqrt{mn}}$$

убывает не медленней, чем в геометрической прогрессии.

Если  $st > 0$ , то очевидно  $st = |S||T|/(mn) \geq 1/(mn)$ , поэтому при  $st < 1/mn$  на самом деле  $st = 0$ , значит,  $|S| = 0$  или  $|T| = 0$ , поэтому  $|Y| = 0$  и алгоритм заканчивает работу, построив из ненулевого вектора (содержащего ошибки) нулевой вектор. Отсюда вытекает для числа итераций оценка  $\log_{1/q} \sqrt{mn} + 1$ , где  $1/q = 1/(ax_1 + b) > 1$ . Чтобы ее улучшить, прологарифмируем неравенство  $\frac{x_{i+1}}{x_i} \leq ax_i + b < ax_1 + b < 1$ , тогда

$$\log \frac{x_{i+1}}{x_i} \leq \log(ax_i + b) = \log b + \log \left(1 + \frac{ax_i}{b}\right) \leq \log b + \frac{ax_i}{b} \leq \log b + \frac{ax_1}{b} q^{i-1},$$

откуда, суммируя, имеем  $\log \frac{x_{i+1}}{x_1} \leq i \log b + \frac{ax_1}{b(1-q)}$ , значит,

$$x_{i+1} \leq \frac{1-b}{a} e^{\frac{1-b}{b(1-q)}} b^i,$$

поэтому для числа итераций справедлива оценка

$$\nu = \log_{1/b} \frac{a\sqrt{mn}}{1-b} + \frac{1-b}{b(1-q) \log(1/b)} + 1.$$

Для числа применений алгоритмов декодирования  $C_i$  очевидно справедлива оценка

$$m \left(1 + \sum_{i=1}^{\nu} s_i\right) + n \left(1 + \sum_{i=1}^{\nu} t_i\right).$$

Применяя неравенства  $\delta_1 s_{i+1} \leq 2((1-\gamma)s_{it_i} + \gamma\sqrt{s_{it_i}})$ ,  $\delta_2 t_{i+1} \leq 2((1-\gamma)s_{it_i} + \gamma\sqrt{s_{it_i}})$ , имеем оценку

$$\begin{aligned} m \left( 1 + \beta + \frac{2}{\delta_1} \sum_{i=1}^{\nu-1} ((1-\gamma)s_{it_i} + \gamma\sqrt{s_{it_i}}) \right) + n \left( 1 + \frac{2}{\delta_2} \sum_{i=1}^{\nu-1} ((1-\gamma)s_{it_i} + \gamma\sqrt{s_{it_i}}) \right) = \\ = O_{\delta_1, \delta_2, \beta}(n+m) \left( (1-\gamma) \sum_{i=1}^{\nu-1} x_i^2 + \gamma \sum_{i=1}^{\nu-1} x_i \right) = \\ = O_{\delta_1, \delta_2, \beta}(n+m) \left( \frac{\gamma}{a} e^{\frac{1-b}{b(1-q)}} + \frac{1-\gamma}{a(1+b)} e^{\frac{2-2b}{b(1-q)}} \right). \end{aligned}$$

Так как суммарная сложность применения каждого из алгоритмов декодирования кодов  $C_i$  равна  $O_{\delta_1, \delta_2, r_1, r_2, c, d}(1)$ , то сложность декодирования кода Рота-Скачека равна

$$O_{\delta_1, \delta_2, r_1, r_2, c, d, \beta}(m) \left( \frac{\gamma}{a} e^{\frac{1-b}{b(1-q)}} + \frac{1-\gamma}{a(1+b)} e^{\frac{2-2b}{b(1-q)}} \right).$$

Автор благодарен М. Н. Вялomu за конструктивную критику, в значительной мере способствовавшую улучшению текста статьи.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Pinsker M. S. On the complexity of a concentrator. In 7th International Teletraffic Conference, pages 318/1-318/4, 1973.
- [2] Бассалыго Л.А., Пинскер М.С. Сложность оптимальной неблокирующей коммутационной схемы. Проблемы передачи информации, 9(1):84-87, 1973.
- [3] Маргулис Г.А. Точные конструкции расширителей. Проблемы передачи информации. 9(4):71-80, 1973.
- [4] Ловас Л., Пламмер М. Прикладные задачи теории графов. М. Мир, 1998.
- [5] Сидельников В.М. Теория кодирования. Физматлит, 2008.
- [6] Lubotzky A., Philips R., Sarnak P. Ramanujan graphs. Combinatorica 8(3), 261-277 (1988).
- [7] Сарнак П. Модулярные формы и их приложения. М. Фазис, 1998.
- [8] Hoory S., Linial N., Wigderson A. Expander graphs and their applications Bull. Amer. Math. Assoc. S-0273-0979(06), 01126-8 (2006).
- [9] Davidoff G., Sarnak P., Valette A. Elementary number theory, group theory and Ramanujan graphs. Cambridge University Press, 2003.
- [10] Alon N., Chung F.R.K. Explicit constructions of linear sized tolerant networks, Discr.Math, 72, 15-19(1988).

- [11] Bilu Y., Linial N. Lifts, discrepancy and nearly optimal spectral gaps. *Combinatorica*, to appear.
- [12] Zemor G. On expander codes. *IEEE Trans. Inform. Theory*, IT-47(2), 835-837 (2001).
- [13] Маргулис Г.А. Точные теоретико-групповые конструкции комбинаторных схем и их приложения в построении расширителей и концентраторов. *Проблемы передачи информации*. 24(1), 39-46 (1988).
- [14] Morgenstern M. Existence and explicit constructions of  $(q + 1)$ -regular Ramanujan graphs for every prime power  $q$ . *J. Comb. Theory, Ser. B*, 62, 44-62 (1994).
- [15] Зяблов В.В. Оценка сложности построения бинарных линейных каскадных кодов *Проблемы передачи информации*. 7(1), 3-10 (1971).
- [16] Dumer I. Concatenated codes and their multilevel generalizations. In *Handbook of coding theory*, v.2, North-Holland, 1988, pp.1911-1988.
- [17] Tanner R.M. A recursive approach to low complexity codes. *IEEE Trans. Inform. Theory*, IT-27(5), 533-547 (1981).
- [18] Sipser M. , Spielman D. A. Expander codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1710-1722, 1996.
- [19] Janwa H., Lal A.K. On Tanner codes: minimum distance and decoding. *AAECC13* (2002).
- [20] Guruswami V., Indyk P. Linear time encodable/decodable codes with near-optimal rate, *IEEE Trans. Inform. Theory*, IT-51(10), 3393-3400 (2005).
- [21] Roth R., Skachek V. Improved nearly-MDS expander codes, *IEEE Trans. Inform. Theory*, (2006).

## Колмогоровская сложность

Ю. Л. Притыкин

Данная заметка представляет собой переработанный конспект лекции (с задачами, решения многих из которых были разобраны на лекции), прочитанной в школе №57 г. Москвы школьникам 11д класса в 2006 году и повторенной школьникам 11в класса в 2008 году. Она основана на находящейся в стадии написания книге по колмогоровской сложности В. А. Успенского, Н. К. Верещагина, А. Шеня [2], которую тем не менее можно рекомендовать для дальнейшего чтения по этой теме, см. файлы по ссылке.

Речь пойдёт о фундаментальном понятии в теории алгоритмов — колмогоровской сложности (другие названия — описательная сложность, дескриптивная сложность, алгоритмическая сложность, колмогоровская энтропия). Его окончательно сформулировал Колмогоров (а также независимо Соломонов и Чейтин) в своих работах 1965 года [1]. Оно формализует интуитивное представление о количестве информации в конечных объектах — в нашем случае это слова из 0 и 1.

Интуитивно ясно, что строку

«0101010101010101010101010101»

можно описать короче — «15 слов 01», в то время как строку

«110110000111010011010011110010»

никак существенно короче не запишешь. Значит, в первой из них информации меньше, чем во второй, хотя они имеют одинаковую длину. Или, например, чтобы задать слово из миллиона нулей, достаточно сказать «1000000 нулей», и вовсе необязательно все эти нули выписывать. Более жизненный пример: часто при работе с файлами большого размера приходится использовать различные архиваторы — `zip`, `gzip`, `rar` и т. д. — для компрессии, то есть получения более короткого описания. Важно, что если потом разархивировать файл, то получится в точности исходный, и никакая информация не потеряется. Только такие способы описания мы и будем рассматривать. Неформально, количеством информации в слове назовём минимальную длину всевозможных его описаний.

Для точного определения способа описания мы воспользуемся понятием алгоритма. Интуитивное представление о том, что такое алгоритм, есть

у каждого (например, программа на почти любом известном вам языке программирования — Pascal, C, C++, Java и т. д.). Мы ограничимся только рассмотрением декомпрессоров, то есть алгоритмов получения по описанию исходного слова (естественно потребовать, чтобы по любому слову декодирование происходило однозначно). *Способом описания* (или *декомпрессором*) будем называть произвольную вычислимую с помощью алгоритма функцию  $D$  на множестве  $A$  всех двоичных слов (включая пустое слово  $\Lambda$ ):  $A = \{\Lambda, 0, 1, 00, 01, 10, 11, 000, \dots\}$ . Значения функции  $D$  тоже лежат в  $A$ . Однако это ограничение можно обойти и распространить  $D$  на многие другие конечные объекты. Для этого можно рассматривать какое-нибудь их кодирование двоичными словами. Мы разрешаем вычислимым функциям быть неопределёнными на некоторых словах (возможно, даже на всех) — содержательно это означает, что соответствующая программа не завершает работу (зацикливается) на некоторых входах.

Через  $|x|$  обозначим длину слова  $x \in A$ .

ОПРЕДЕЛЕНИЕ 1. Назовём число

$$K_D(x) = \min\{|y| : D(y) = x\}$$

сложностью слова  $x$  относительно способа описания  $D$ . (Мы считаем  $\min \emptyset = +\infty$ .)

ЗАДАЧА 1. Какую функцию сложности  $K_D$  определяют следующие способы описания? а)  $D(y) = \Lambda$  для любого слова  $y$  из  $A$ ; б)  $D(\text{bin}(n)) = \underbrace{00\dots 0}_n$  (здесь  $\text{bin}(n)$  — двоичная запись целого положительного числа  $n$ ); в)  $D(y) = y$  для любого слова  $y$  из  $A$ .

Определение сложности очень сильно зависит от способа описания. Подбирая такой способ, мы можем сделать очень маленькой сложность любого конкретного слова или даже семейства слов. Наша задача — выбрать оптимальный способ, который хорош для всех слов одновременно. Естественно считать, что способ описания тем лучше, чем более короткие описания он даёт.

ОПРЕДЕЛЕНИЕ 2. Способ описания  $D_1$  не хуже способа  $D_2$ , если для некоторой константы  $C$  и для всех слов  $x$  имеем

$$K_{D_1}(x) \leq K_{D_2}(x) + C.$$

Мы будем пренебрегать различием на константу. Без этого дальнейшую теорию построить не удаётся (в частности, для аналогичного определения без константы неверна теорема Соломонова — Колмогорова о существовании оптимального способа описания, см. далее).



Прежде чем строить наилучший способ описания, начнём с малого. Пусть есть два способа описания  $D_1$  и  $D_2$ . Как построить способ  $D$ , который не хуже каждого из них? Положим

$$D(0y) = D_1(y), \quad D(1y) = D_2(y).$$

Таким образом, учитываются сразу оба способа описания. Чтобы потом можно было декодировать полученное описание, приписываем к нему информацию о том, каким способом оно было получено — первым или вторым.

**ЗАДАЧА 2.** Проверьте, что  $K_D(x) \leq K_{D_1}(x) + 1$  и  $K_D(x) \leq K_{D_2}(x) + 1$  для любого  $x$ .

Обобщив эту идею, мы можем построить такой оптимальный способ описания  $D$ , что для любого другого способа  $D'$  существует такая константа  $C_{D'}$ , что для любого слова  $x$  имеем

$$K_D(x) \leq K_{D'}(x) + C_{D'}$$

(это утверждение называется теоремой Соломонова – Колмогорова). Действительно, мы можем учитывать все имеющиеся способы записи одновременно, просто указывая в начале описания, какой способ использовался. Положим

$$D(py) = p(y),$$

где  $p$  — произвольный алгоритм (записанный двоичным кодом). Ясно тогда, что  $D$  хуже способа описания, задаваемого программой  $p$ , не более чем на константу — длину текста программы  $p$ .

Таким образом, когда  $D$  подадут на вход текст, он выделяет из его начала текст программы, а потом применяет эту программу к оставшемуся тексту. Однако возникает проблема. Пусть, например, нам дают слово 01. Мы не знаем, применять ли программу 0 к слову 1 или программу 01 к пустому слову (или, может быть, программу с пустым текстом  $\Lambda$  к слову 01). Поэтому мы должны придумать такое описание пары  $p, y$ , чтобы однозначно выделять из него  $p$ . Это несложно: например, будем удваивать каждый символ слова  $p$ , а по его окончании вставим 01. Например, если  $p = 011001$  и  $y = 01001$ , то получим код 0011110000110101001.

**ЗАДАЧА 3.** А) Описанный выше способ кодирования пары слов  $x$  и  $y$  одним словом даёт оценку  $2|x| + |y| + 2$  на длину кода. Как можно было бы эту оценку улучшить (например, получить оценку  $|x| + |y| + 2 \log |x| + 2$  или  $|x| + |y| + \log |x| + 2 \log \log |x| + 2$ , и т. д.)<sup>1)</sup> Б) Можно ли придумать способ кодировать пару слов  $x, y$  словом длины  $|x| + |y| + c$ , где  $c$  — константа?

Оптимальный способ описания построен.

<sup>1)</sup>Всюду под  $\log$  подразумевается  $\log_2$ .

ОПРЕДЕЛЕНИЕ 3. Фиксируем некоторый оптимальный способ описания  $D$ . Назовём

$$K(x) = \min\{|y| : D(y) = x\}$$

колмогоровской сложностью слова  $x$ .

ЗАДАЧА 4. Докажите, что замена оптимального способа описания в этом определении на другой (оптимальных может быть много) приводит к изменению функции сложности не более чем на константу.

ЗАДАЧА 5. Докажите, что колмогоровская сложность любого слова конечна.

В силу произвола в выборе оптимального способа описания в определении все утверждения про колмогоровскую сложность носят асимптотический характер. Рассмотрим пример: как может меняться сложность слова при приписывании к нему символа 1? Ясно, что от этого количество информации в слове меняется несущественно. Можно доказать, что  $K(1x) = K(x) + O(1)$ . (Здесь  $O(1)$  обозначает функцию от  $x$ , ограниченную константой. Далее мы будем использовать общее определение  $O(f(x))$  как такой функции  $g(x)$ , что для некоторой константы  $C$  и всех  $x$ , больших некоторого  $D$ , выполняется неравенство  $g(x) \leq Cf(x)$  (будем для удобства считать, что все рассматриваемые функции неотрицательные).) Неформально, для этого достаточно показать, что можно по каждому из слов  $x$  и  $1x$  легко получать другое, не используя дополнительной информации.

ЗАДАЧА 6. а) Проведите последнее рассуждение формально. б) Докажите, что  $K(xx) = K(x) + O(1)$ . в) Докажите, что  $K(\underbrace{00\dots 0}_n) \leq \log n + O(1)$ .

ЗАДАЧА 7. Докажите, что  $K(x) \leq |x| + O(1)$ .

ЗАДАЧА 8. Пусть  $P$  — произвольный алгоритм. Докажите, что  $K(P(x)) \leq K(x) + O(1)$ .

ЗАДАЧА 9. а) Покажите, что если описания рассматривать не в бинарном алфавите, а в алфавите из 4 символов, то сложность уменьшится вдвое. б) Сформулируйте и докажите аналогичное утверждение для перехода к произвольному конечному алфавиту.

Оказывается, что неравенство в задаче 7 для большинства слов близко к равенству. Действительно, давайте посчитаем, сколько слов может иметь сложность меньше некоторого фиксированного числа  $n$ . Слов сложности 0 не более одного — оно должно иметь описание  $\Lambda$ , слов сложности 1 не более двух — их описания могут быть только 0 и 1, и так далее. Поэтому слов сложности меньше  $n$  не больше, чем  $1 + 2 + 4 + \dots + 2^{n-1} = 2^n - 1$ . Таким образом, для любого натурального  $n$  существует менее  $2^n$  слов  $x$ , для которых  $K(x) < n$ . Отсюда легко получаем, что доля слов сложности

менее  $n - c$  среди всех слов длины  $n$  меньше  $\frac{2^{n-c}}{2^n} = 2^{-c}$ . Например, доля слов сложности меньше 90 среди слов длины 100 меньше  $\frac{1}{1024}$ . Таким образом, большинство слов несжимаемы или почти несжимаемы.

Удивительно, что в полученном утверждении — существует менее  $2^n$  слов сложности меньше  $n$  — нет никаких констант.

ЗАДАЧА 10. Где в этом утверждении зависимость от фиксированного в определении колмогоровской сложности оптимального способа описания?

ЗАДАЧА 11. Покажите, что для некоторого фиксированного  $c$  и всех  $n$  количество слов сложности меньше  $n$  заключено между  $2^{n-c}$  и  $2^n$ .

ЗАДАЧА 12. Покажите, что среднее арифметическое сложностей всех слов длины  $n$  равно  $n + O(1)$ .

ЗАДАЧА 13. А) Докажите, что если  $y$  получено из слова  $x$  длины  $n$  заменой одного символа, то  $K(y) = K(x) + O(\log n)$ . Б) Может ли тем не менее  $K(y)$  существенно отличаться от  $K(x)$  (например, в 1000 раз)?

ЗАДАЧА 14. Докажите, что  $K(xy) \leq K(x) + K(y) + 2 \log K(x) + O(1)$ .

ЗАДАЧА 15. А) Докажите, что  $K(x, y) \leq K(x) + K(y) + O(\log n)$  для любых  $x$  и  $y$  длины не более  $n$  (здесь под парой слов  $x, y$  понимается её кодирование двоичным словом в некоторой раз навсегда выбранной стандартной кодировке). Б) Можно ли доказать неравенство  $K(x, y) \leq K(x) + K(y) + O(1)$ ? (Указание: можно ли доказать  $K(x, y) \leq |x| + |y| + O(1)$ ?) В) Как соотносятся между собой  $K(x, y)$  и  $K(xy)$ ?

ЗАДАЧА 16. А) Докажите  $2K(xyz) \leq K(xy) + K(xz) + K(yz) + O(\log n)$  для всех слов  $x, y, z$  длины не более  $n$ . Б) Пусть тело в трёхмерном пространстве имеет объём  $V$ , а площади его ортогональных проекций на плоскости  $xOy$ ,  $xOz$  и  $yOz$  (в прямоугольной системе координат) равны соответственно  $S_{xy}$ ,  $S_{xz}$  и  $S_{yz}$ . Докажите неравенство  $V^2 \leq S_{xy} \cdot S_{xz} \cdot S_{yz}$ .

Было бы очень полезно уметь вычислять  $K(x)$ . Это можно пытаться делать следующим образом. Возьмём алгоритм оптимальной декомпрессии (тот самый, который был зафиксирован в определении колмогоровской сложности). Запустим этот алгоритм на пустом слове. Если получилось  $x$ , значит,  $K(x) = 0$ . Если нет, запустим алгоритм на словах 0 и 1. Если на каком-то из них выдан ответ  $x$ , значит,  $K(x) = 1$ . И так далее: если мы уже знаем, что  $K(x) > n - 1$ , и нашлось слово длины  $n$ , на котором алгоритм выдаёт  $x$ , значит,  $K(x) = n$ . Однако есть проблема: уже на первом шаге, пытаясь применить программу к пустому слову, мы можем столкнуться с трудностью — алгоритм может не завершить работу. И до всего следующего дело уже не дойдёт.

ЗАДАЧА 17. Покажите, что оптимальный способ описания — не всюду определённая функция.

Утверждение этой задачи может показаться странным, ведь если способ описания определён не всюду, мы можем доопределить его в некоторых точках — ясно, что от этого он может только улучшиться. Однако формального противоречия здесь нет (только философское) — чтобы функция осталась вычислимой, её можно доопределить лишь в конечном количестве точек.

В заключение приведём любопытное рассуждение, показывающее, что вычислять функцию  $K(x)$  не удаётся (ни вышеприведённым способом, ни каким-либо другим).

Пусть алгоритмически вычислимая функция  $f$  оценивает снизу колмогоровскую сложность, то есть  $f(x) \leq K(x)$  для любого  $x$ . Покажем, что такая оценка может быть только тривиальной: для некоторого  $C$  и для всех  $x$  выполнено  $f(x) \leq C$ . Предположим обратное — пусть  $f$  не ограничена. Пользуясь этим, построим следующий алгоритм  $P$ . Получив на вход натуральное число  $n$ , он запускает одновременно алгоритм для вычисления  $f$  на всех словах (отдельный вопрос — как запускать одновременно счётное число программ; обдумайте этот вопрос самостоятельно). Время от времени на каких-то из слов алгоритм выдаёт ответ, и мы можем проверить, верно ли для этого слова  $x$ , что  $f(x) \geq n$ . Поскольку по предположению  $f$  не ограничена, обязательно когда-нибудь найдётся слово, для которого это неравенство действительно выполнено. Первое найденное такое слово алгоритм  $P$  и выдаёт в качестве ответа.

Имеем  $f(P(n)) \geq n$ , а значит, и  $K(P(n)) \geq n$ . С другой стороны, по задаче 8 имеем  $K(P(n)) \leq K(n) \leq \log n + O(1)$ . Неравенство  $\log n + O(1) \geq n$  нарушается при достаточно больших  $n$  — противоречие.

Чтобы доказать невычислимость функции  $K$ , осталось заметить, что она является собственной оценкой снизу.

**Задача 18.** Докажите, что функция, равная на слове  $x$  его кратчайшему описанию (при оптимальном способе описания), не вычислима алгоритмически. (Во многом именно поэтому мы ограничиваемся рассмотрением декомпрессоров.)

**Задача 19.** Существует ли алгоритм, которому можно на вход подать тексты двух программ, задающих способы описания  $D_1$  и  $D_2$ , про которые известно, что  $D_1$  не хуже  $D_2$ , и этот алгоритм выдаст константу в неравенстве  $K_{D_1}(x) \leq K_{D_2}(x) + C$  из определения?

**Задача 20.** Пусть  $B(n) = \min\{N \in \mathbb{Z} : K(m) > n \text{ для любого } m > N\}$  (регулятор сходимости  $K(m)$  к  $\infty$ ). Тогда для любой алгоритмически вычислимой функции  $f: \mathbb{N} \rightarrow \mathbb{N}$  почти для всех  $n$  верно  $B(n) \geq f(n)$  (то есть  $B$  растёт быстрее любой вычислимой функции!).

**Задача 21.** Пусть  $0 < \alpha < 1$ . А) Докажите, что существуют сколь угодно длинные слова, у которых каждое начало  $x$  имеет сложность

$K(x) > \alpha|x| + O(1)$ . б) Докажите, что существуют сколь угодно длинные слова, у которых каждое подслово  $x$  сложно:  $K(x) > \alpha|x| + O(1)$ .

## СПИСОК ЛИТЕРАТУРЫ

- [1] А. Н. Колмогоров. *Три подхода к определению понятия «количество информации»* // Проблемы передачи информации, т. 1, №1, 1965. С. 3–11.
- [2] В. А. Успенский, Н. К. Верещагин, А. Шень. *Колмогоровская сложность*. Неопубликованная книга, см.:  
<http://kolmsem.math.ru/rus/materials/materials.html>

## Метрические и ультраметрические пространства сопротивлений

В. А. Гурвич

Данная заметка посвящена формулировке и доказательству следующего факта и его обобщений: сопротивления между полюсами сети удовлетворяют неравенству треугольника

$$\mu(a, b) \leq \mu(a, c) + \mu(c, b). \quad (1)$$

Упоминание физических терминов не должно вводить читателя в заблуждение: речь идет о математическом утверждении. Ниже мы его сформулируем строго и дадим план доказательства.

Изложение построено как цепочка связанных друг с другом упражнений, задач и указаний к ним и следует заметке [2].

1. ЗАКОН ПРОВОДИМОСТИ. Имеется ребро  $e$  с концами  $a$  и  $b$  (проводник). Закон проводимости — это функция  $f_e$ , которая напряжению  $y_e$  ставит в соответствие ток  $y_e^* = f_e(y_e)$ .

Здесь мы будем рассматривать только степенной закон проводимости:

$$y_e^* = f_e(y_e) = \lambda_e^s |y_e|^r \operatorname{sign}(y_e) = \frac{|y_e|^r}{\mu_e^s} \operatorname{sign}(y_e). \quad (2)$$

Параметры  $\lambda_e$  и  $\mu_e = \lambda_e^{-1}$  называются соответственно проводимостью и сопротивлением ребра  $e$ . Параметры  $r$  и  $s$  мы будем называть показателем и нормировочной степенью закона проводимости. Будем предполагать, что все четыре параметра строго положительны.

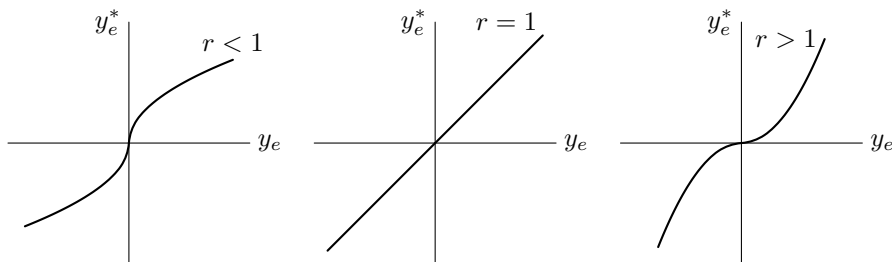


Рис. 1.

Заметим, что при  $r = s = 1$  получается обычный закон Ома (ток пропорционален напряжению). Заметим также, что в формуле (2) можно было бы обойтись и без параметра  $s$ , но он нам понадобится в дальнейшем.

УПРАЖНЕНИЕ 1. Проверьте, что функция  $f_e$

- (i) непрерывна и принимает все вещественные значения;
- (ii) строго возрастает:  $f_e(y_e) > f_e(y'_e)$  если  $y_e > y'_e$ ,
- (iii) нечетна:  $f_e(-y_e) = -f_e(y_e)$ ; кроме того,
- (iv) обратная функция  $f_e^{-1}$  также является степенной;
- (v) найдите параметры  $r'$  и  $s'$  обратной функции.

2. ОСНОВНЫЕ ПЕРЕМЕННЫЕ. *Схема проводников* — неориентированный связный взвешенный граф  $G = (V, E, \mu)$  с множеством вершин  $V$ , множеством ребер  $E$  и весами (сопротивлениями) ребер  $\mu_e$ .

Задача расчета схемы — это задача решения некоторой системы уравнений. Переменные в этой системе делятся на четыре группы, по две переменные для каждой вершины и каждого ребра:

- ▷ потенциалы вершин  $x_v$ ;
- ▷ токи на ребрах  $y_e^*$ ;
- ▷ напряжения на ребрах  $y_e$ ;
- ▷ суммарные токи в вершинах  $x_v^*$ .

Все системы уравнений, которые мы будем рассматривать, включают в себя уравнения, которые линейно выражают напряжения через потенциалы и суммарные токи через токи на ребрах. Для записи этих уравнений нам придется произвольным образом ориентировать рёбра и ввести функцию инцидентности  $\text{inc}(v, e)$ :

$$\text{inc}(v, e) = \begin{cases} +1, & \text{если вершина } v \text{ является началом ребра } e; \\ -1, & \text{если вершина } v \text{ является концом ребра } e; \\ 0, & \text{во всех других случаях.} \end{cases} \quad (3)$$

Упомянутые уравнения имеют вид

$$y_e = \sum_{v \in V} \text{inc}(v, e)x_v, \quad (4)$$

$$x_v^* = \sum_{e \in E} \text{inc}(v, e)y_e^*. \quad (5)$$

Заметим, что уравнение (4) для ребра  $e = (a, b)$  можно записать проще  $y_e = \text{inc}(e, a)x_a + \text{inc}(e, b)x_b$  или еще проще:  $y_e = x_a - x_b$  (в последнем случае предполагается, что ребро  $e$  ориентировано от  $a$  к  $b$ ).

УПРАЖНЕНИЕ 2. Докажите, что из уравнений (4,5) следует равенство

$$\sum_{v \in V} x_v x_v^* = \sum_{e \in E} y_e y_e^*. \quad (6)$$

3. ДВУХПОЛЮСНЫЕ СХЕМЫ. Зафиксируем две различные вершины  $a, b \in V$ , которые будем называть *полюсами*. Добавим к уравнениям (4,5) следующие уравнения

$$x_v^* = 0, \quad \text{для } v \in V \setminus \{a, b\}, \quad (7)$$

выражающие первый закон Кирхгофа. Также зафиксируем потенциалы в полюсах

$$x_a = x_a^0, \quad x_b = x_b^0. \quad (8)$$

Зафиксируем параметры  $r$  и  $s$ . Кроме перечисленных выше линейных уравнений для каждого ребра  $e \in E$  запишем закон проводимости (2).

Заметим, что полученная система уравнений обладает тем свойством, что значения потенциалов вершин  $x_v, v \in V \setminus \{a, b\}$ , однозначно определяют значения всех остальных переменных.

ЗАДАЧА 1. Докажите, что для любых  $x_a^0, x_b^0$  существует единственное решение системы уравнений (4,5,7,8) и (2) для каждого ребра  $e$ .

ПОДСКАЗКА: примените метод последовательных приближений. Подберите такое начальное распределение потенциалов вершин, чтобы потом все потенциалы (нестрого) монотонно возрастали. Докажите, что в любой вершине последовательность потенциалов, во-первых, сходится и, во-вторых, куда нужно (т. е. к решению системы).

Если  $x_a \neq x_b$ , то в вершинах  $a$  и  $b$  первый закон Кирхгофа не выполняется.

ЗАДАЧА 2. Докажите, что  $x_a^* + x_b^* = 0$ . Более того, если  $x_a > x_b$ , то все токи из  $a$  вытекают (т. е.  $y_e^* \text{inc}(a, e) > 0$  для каждого ребра  $e$ , инцидентного  $a$ ), а в  $b$  втекают (т. е.  $y_e^* \text{inc}(b, e) < 0$  для каждого ребра  $e$ , инцидентного  $b$ ); если  $x_a < x_b$ , то выполняются противоположные неравенства; если же  $x_a = x_b$ , то  $x_v = x_a = x_b$  и  $x_v^* = 0$  для всех  $v \in V$ , а  $y_e = y_e^* = 0$  для всех  $e \in E$ .

Величина  $y(a, b) = x_a - x_b$  называется напряжением, а  $y^*(a, b) = x_a^* = -x_b^*$  — током в двухполюсной схеме  $(G, a, b)$ .

ЗАДАЧА 3. Докажите, что ток  $y^*(a, b)$  и напряжение  $y(a, b)$  также связаны степенным законом проводимости с теми же самыми параметрами  $r$  и  $s$ , т. е.

$$y^*(a, b) = f_{a,b}(y(a, b)) = \lambda(a, b)^s |y(a, b)|^r \text{sign}(y(a, b)) = \frac{|y(a, b)|^r}{\mu(a, b)^s} \text{sign}(y(a, b)). \quad (9)$$



Величины  $\lambda(a, b)$  и  $\mu(a, b) = \lambda(a, b)^{-1}$  называются соответственно проводимостью и сопротивлением двухполюсной схемы  $(G, a, b)$ .

ЗАМЕЧАНИЕ. Задача 3 объясняет, почему мы ограничились степенным законом проводимости. Можно показать (см. [1]), что для любого другого семейства непрерывных функций аналогичное утверждение не выполняется.

УПРАЖНЕНИЕ 3. Проверьте, что

- (i)  $\mu(a, b) = \mu(b, a)$  и
- (ii)  $\mu(a, b) > 0$  при  $a \neq b$ ; иными словами,  $y^*(a, b) > 0$  при  $x_a > x_b$ .

По определению положим  $\mu(a, b) = 0$  при  $a = b$ .

УПРАЖНЕНИЕ 4. Проверьте, что если зафиксировать потенциал  $x_a$  в одной вершине  $a \in V$  и потребовать выполнения первого закона Кирхгофа в остальных  $x_v^* = 0, v \in V \setminus \{a\}$ , то  $x_v = x_a$  для всех  $v \in V, y_e = y_e^* = 0$  для всех  $e \in E$  и  $x_a^* = 0$ .

4. ПОСЛЕДОВАТЕЛЬНОЕ И ПАРАЛЛЕЛЬНОЕ СОЕДИНЕНИЕ РЕБЕР. Рассмотрим две простейших двухполюсных схемы, изображенных на рис. 2.

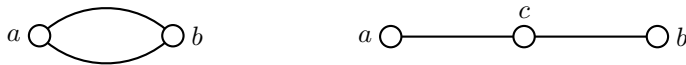


Рис. 2.

ЗАДАЧА 4. Выразите сопротивления этих схем через сопротивления ребер.

5. НЕРАВЕНСТВО ТРЕУГОЛЬНИКА. Зафиксируем теперь не две, а три вершины  $a, b, c \in V$ .

ЗАДАЧА 5. Докажите неравенство

$$\mu(a, c)^{s/r} + \mu(c, b)^{s/r} \geq \mu(a, b)^{s/r}. \tag{10}$$

Опишем план решения этой задачи. Зафиксируем в двухполюсной сети  $(G, a, b)$  потенциалы  $x_a = x_a^0$  и  $x_b = x_b^0$ . Пусть для определенности  $x_a^0 \geq x_b^0$ . Согласно задаче 1 все остальные потенциалы определены однозначно, в частности, потенциал вершины  $c$  принимает некоторое значение  $x_c^0$ .

Рассмотрим еще две двухполюсные схемы:  $(G, a, c)$  с потенциалами  $x_a^0, x_c^0$  и  $(G, c, b)$  с потенциалами  $x_c^0, x_b^0$ . Будем искать распределения потенциалов в  $(G, a, c)$  методом последовательных приближений, а в качестве начального приближения возьмем распределение потенциалов в схеме  $(G, a, b)$ .

Заметим, что при этом потенциалы вершин  $a$  и  $c$  не меняются.

ЗАДАЧА 6. Докажите, что потенциалы остальных вершин монотонно возрастают; в каждой вершине последовательность потенциалов сходится; в результате получается искомое распределение.

ЗАДАЧА 7. Выведите неравенства для токов

$$y^*(a, b) \geq y^*(a, c) \quad \text{и} \quad y^*(a, b) \geq y^*(c, b). \quad (11)$$

ЗАДАЧА 8. Выведите неравенство (10) из (11).

ЗАДАЧА 9. Покажите, что при  $s \geq r$  из (10) вытекает неравенство треугольника  $\mu(a, c) + \mu(c, b) \geq \mu(a, b)$ , а при  $s/r \rightarrow \infty$  в пределе получается ультраметрическое неравенство  $\max(\mu(a, c), \mu(c, b)) \geq \mu(a, b)$ .

6. ЧЕТЫРЕ ПРИМЕРА МЕТРИКИ СОПРОТИВЛЕНИЙ.

6.1. При  $r = s = 1$  получаем обычные «электрические» сопротивления. Таким образом, как и было обещано, сопротивления между вершинами графа  $G$  образуют метрику.

6.2. Граф  $G = (V, E)$  моделирует сеть дорог. Сопротивление  $\mu_e$  интерпретируется как длина дороги  $e \in E$  (или время проезда по ней); а  $\mu(a, b)$  — это длина кратчайшего пути из  $a$  в  $b$  (или минимальное время проезда). В этом случае неравенство треугольника очевидно.

ЗАДАЧА 10. Покажите, что этот случай описывается предельным переходом  $r = s \rightarrow 0$  или более общо  $s \rightarrow 0, r/s \rightarrow 1$ .

6.3. Пусть теперь граф  $G = (V, E)$  моделирует трубопровод. Каждое ребро  $e \in E$  — это труба, по которой транспортируется жидкость или газ;  $\lambda_e$  интерпретируется как пропускная способность трубы  $e$ , т. е. количество жидкости или газа, которое можно пропустить по  $e$  в единицу времени, а  $\lambda(a, b)$  — это пропускная способность всей двухполосной схемы, т. е. максимальное количество жидкости или газа, которое можно прокачать из  $a$  в  $b$  за единицу времени.

ЗАДАЧА 11. Покажите, что этот случай описывается предельным переходом  $s = 1, r \rightarrow 0$ . При этом величины  $\mu$  удовлетворяют ультраметрическому неравенству. Что означает это неравенство в терминах проводимостей, т. е. максимальных возможных потоков между  $a$  и  $b$ ,  $b$  и  $c$ ,  $c$  и  $a$ ? (См. [4, теорема 9.1].)

6.4. Пусть теперь граф  $G = (V, E)$  моделирует систему рек и каналов, а  $\lambda_e$  обозначает *ширину*  $e$ , т. е. максимальный размер (ширину, вес, водоизмещение) корабля, который еще может пройти по  $e$ ; при этом  $\lambda(a, b)$  — это ширина всей двухполосной схемы, т. е. максимальный размер корабля, который еще может пройти из  $a$  в  $b$ .

ЗАДАЧА 12. Покажите, что этот случай описывается предельным переходом  $s \rightarrow 0$ ,  $r/s \rightarrow 0$ . При этом величины  $\mu$  удовлетворяют ультраметрическому неравенству. Что оно означает в терминах проводимостей?

ЗАМЕЧАНИЕ. В моделях 6.3 и 6.4 предполагается, что вершины никак не ограничивают возможности транспортировки, а в 6.2 они не увеличивают длины пути.

ПОДСКАЗКА К ЗАДАЧАМ 10–12. Начните со схем, изображенных на рисунке 2, и подберите параметры  $r$  и  $s$  для этих схем.

7. ЛИНЕЙНЫЙ СЛУЧАЙ. Обычные электрические сопротивления отвечают случаю  $r = s = 1$ .

В этом случае можно усилить рассуждение про двухполюсные схемы из п. 3, где было показано, что ток и напряжение между полюсами связаны законом проводимости с теми же параметрами  $r$  и  $s$ , что и на каждом ребре. Другими словами, двухполюсную схему можно заменить одним ребром.

ЗАМЕЧАНИЕ. Минти [5] показал, что двухполюсную схему можно заменить одним ребром не только для степенного, но и для произвольного монотонного закона проводимости. Точнее говоря, если закон проводимости на каждом ребре  $e$  задается монотонно возрастающей функцией  $f_e$ , то двухполюсную схему с полюсами  $a, b$  можно заменить ребром  $(a, b)$  с монотонной функцией проводимости  $f_{(a,b)}$ .

Теперь рассмотрим  $k$ -полюсную схему в линейном случае. Зафиксируем  $k \geq 2$  различных полюсов  $a_1, \dots, a_k \in V$ . Добавим к уравнениям (4,5) первый закон Кирхгофа для вершин, которые не являются полюсами:

$$x_v^* = 0, \quad \text{для } v \in V \setminus \{a_1, \dots, a_k\}, \quad (12)$$

а в полюсах зафиксируем потенциалы

$$x_{a_i} = x_i^0, \quad 1 \leq i \leq k. \quad (13)$$

УПРАЖНЕНИЕ 5. Докажите, что существует единственное решение линейной системы уравнений (2, 4, 5, 12, 13).

Две  $k$ -полюсные схемы  $(G, a_1, \dots, a_k)$  и  $(G', a'_1, \dots, a'_k)$  будем называть *эквивалентными*, если при одинаковых потенциалах в полюсах суммарные токи в полюсах тоже одинаковы, т. е. при условиях  $x_{a_i} = x_{a'_i} = x_i^0$  в уравнениях (13) для решений систем уравнений (2, 4, 5, 12, 13) выполнено  $x_{a_i}^* = x_{a'_i}^*$  для  $1 \leq i \leq k$ .

ЗАДАЧА 13. Для каждой  $k$ -полюсной схемы с  $n$  вершинами существует эквивалентная ей  $k$ -полюсная схема с  $k$  вершинами.

Эту задачу можно решить, явно указав редукцию  $(n + 1)$ -вершинной схемы к  $n$ -вершинной, где  $n \geq k$ .

Занумеруем вершины  $(n + 1)$ -вершинной схемы  $G$  от 0 до  $n$  и обозначим через  $\lambda_{ij}$  проводимость на ребре  $(i, j)$ . (Если такого ребра нет, то проводимость равна 0.) Построим  $n$ -вершинную схему  $G'$ , вершины которой занумерованы от 1 до  $n$ , а проводимости заданы формулой

$$\lambda'_{ij} = \lambda_{ij} + \frac{\lambda_{0i}\lambda_{0j}}{\sum_{\ell=1}^n \lambda_{0\ell}}. \quad (14)$$

ЗАДАЧА 14. Докажите, что схемы  $G$  и  $G'$  эквивалентны.

Опишем план решения этой задачи.

Во-первых, заметим, что в силу линейности при параллельном соединении ребер проводимости складываются. Поэтому достаточно преобразовать  $(n + 1)$ -вершинную звезду в эквивалентную  $n$ -вершинную схему. Звезда на вершинах  $0, 1, \dots, n$  содержит  $n$  ребер  $(0, i)$ .

Во-вторых, поскольку система уравнений линейна, то достаточно рассмотреть базисные наборы потенциалов в вершинах  $1, \dots, n$  вида  $x_i = 1$ , а  $x_\ell = 0$  при  $\ell \neq i$ . Проводимости  $\lambda'_{i\ell}$  можно найти, вычислив потенциал в вершине 0. При этом оказывается, что  $\lambda'_{ij}$  получает одно и то же значение для двух базисных наборов: с  $x_i = 1$  и с  $x_j = 1$ .

ЗАМЕЧАНИЕ. При  $n = 3$  замена 4-вершинной звезды на треугольник называется Y- $\Delta$  преобразованием.

УПРАЖНЕНИЕ 6. Выведите неравенство треугольника в линейном случае из эквивалентности произвольной 3-полосной схемы некоторой 3-вершинной схеме.

В случае двухполосной схемы из формулы (14) можно получить формулу Кирхгофа для проводимости  $\lambda(a, b)$ . Для этого введём матрицу Кирхгофа  $M$  для  $n$ -вершинной схемы:  $M_{ij} = \lambda_{ij}$  при  $i \neq j$  и  $M_{ii} = -\sum_j \lambda_{ij}$ . Рассмотрим два минора матрицы Кирхгофа:  $\Delta_1(a, b)$  — это определитель матрицы, полученной вычеркиванием строки  $a$  и столбца  $b$ , а  $\Delta_2(a, b)$  — это определитель матрицы, полученной вычеркиванием строк  $a, b$  и столбцов  $a, b$ .

ЗАДАЧА 15. Выведите из формулы (14) формулу Кирхгофа для проводимостей

$$\lambda(a, b) = \left| \frac{\Delta_1(a, b)}{\Delta_2(a, b)} \right|. \quad (15)$$

БЛАГОДАРНОСТИ. Автор признателен М. Н. Вялому за полезные идеи и улучшение изложения.

## СПИСОК ЛИТЕРАТУРЫ

- [1] Гвишиани А. Д., Гурвич В. А. *Динамические задачи классификации и выпуклое программирование в приложениях*. М.: Наука, 1992.
- [2] Гвишиани А. Д., Гурвич В. А. *Метрические и ультраметрические пространства сопротивлений* // УМН, 1987. Т. 42, вып. 6. С. 187–188.
- [3] Ляшко О. *Почему не уменьшается сопротивление?* // Квант, №1, 1985. С. 10–15.
- [4] Ху Т. *Целочисленное программирование и потоки в сетях*. М.: Мир, 1974.
- [5] Minty G. J. *Monotone networks* // Proceedings of the Royal Society of London, 1960. Ser. A, vol. 257. P. 194–212.

# Описанные циклические линии треугольника в геометрии Лобачевского

П. В. Бибииков

Хорошо известно, что на плоскости Лобачевского вокруг треугольника можно описать либо *окружность*, либо *эквидистанту*, либо *орцикл* (эти кривые называются *циклическими линиями*). В этой статье мы докажем критерий, позволяющий определить, какую именно циклическую линию можно описать вокруг данного треугольника.

## 1. КРАТКИЕ НАПОМИНАНИЯ

Все рассуждения мы будем проводить в модели Пуанкаре в верхней полуплоскости (см. [1]). Напомним, что *плоскостью Лобачевского* называется фиксированная полуплоскость (которую обычно называют *верхней*) относительно некоторой прямой. Эта прямая вместе с бесконечно удаленной точкой называется *абсолютом*. Точки абсолюта называются *бесконечно удаленными*. *Прямыми* в модели Пуанкаре являются полуокружности с центром на абсолюте и вертикальные лучи, перпендикулярные абсолюту (рис. 1). *Угол* между двумя неевклидовыми прямыми по определению полагается равным евклидовому углу между соответствующими кривыми.

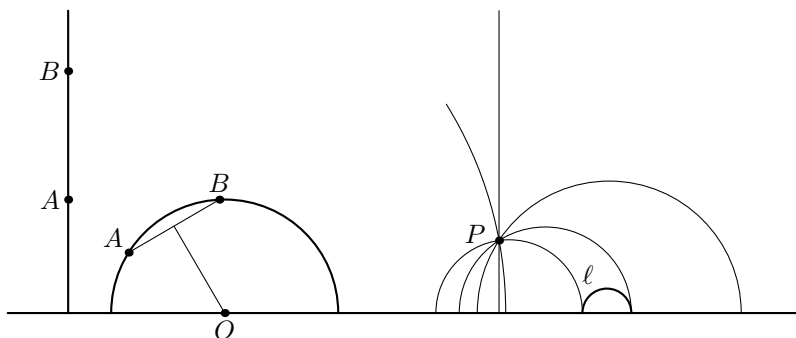


Рис. 1.

Расстояние между точками  $A$  и  $B$  может быть вычислено по формуле

$$\rho(A, B) = \ln \frac{r' + r}{r' - r},$$

где  $r = AB$ ,  $r' = A'B$  и  $A'$  — точка, симметричная точке  $A$  относительно абсолюта.

В геометрии Лобачевского также можно определить простейшие кривые. Однако, в отличие от евклидовой геометрии, в геометрии Лобачевского элементарные кривые более разнообразны и не ограничиваются одной лишь окружностью.

**УПРАЖНЕНИЕ 1.** Докажите, что всякая евклидова окружность, целиком лежащая в верхней полуплоскости, является также неевклидовой окружностью, и наоборот, каждая неевклидова окружность является одновременно и евклидовой окружностью.

Итак, если евклидова окружность целиком лежит в верхней полуплоскости, то она является неевклидовой окружностью. Возникает естественный вопрос: *а чем является евклидова окружность, пересекающая абсолюта?* Для того, чтобы ответить на этот вопрос, напомним еще одно определение.

**ОПРЕДЕЛЕНИЕ 1.** *Эквидистантой* называется множество точек, расположенных на заданном расстоянии  $h$  от данной прямой  $p$  и лежащих в заданной полуплоскости относительно этой прямой. Прямая  $p$  называется *базой эквидистанты*, величина  $h$  — *высотой*, а фиксированная полуплоскость — *полуплоскостью эквидистанты*.

2. *Орициклом* называется кривая, пересекающая все прямые, имеющие общую бесконечно удаленную точку, под прямым углом.

**УПРАЖНЕНИЕ 2.** Докажите, что евклидова окружность, касающаяся абсолюта, является орициклом, а окружность, пересекающая абсолюта в двух точках, — эквидистантой (рис. 2).

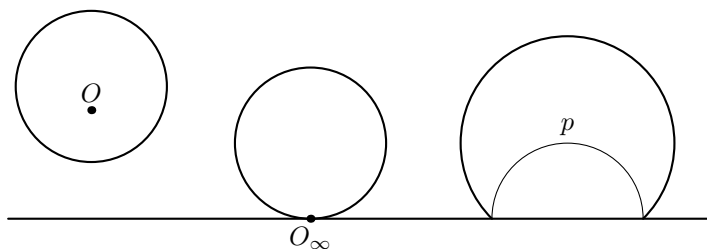


Рис. 2.

Отметим, что наклонные лучи с началом на абсолюте также являются эквидистантами, а прямые, параллельные абсолюту, — орициклами.

УПРАЖНЕНИЕ 3. Докажите это.

Итак, циклические кривые — это в точности те кривые, которые изображаются в модели Пуанкаре дугами окружности. Поскольку в геометрии Евклида вокруг любого треугольника можно описать единственную окружность, получаем следующую теорему для геометрии Лобачевского.

ТЕОРЕМА 1. *Вокруг любого треугольника можно описать либо единственную окружность, либо единственный орицикл, либо единственную эквидистанту (рис. 3).*

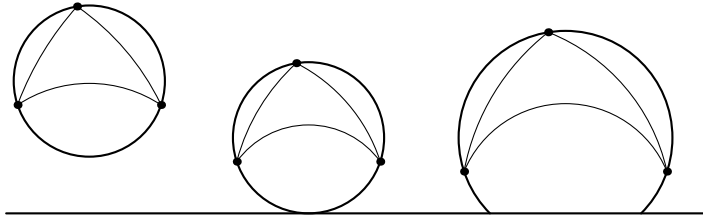


Рис. 3.

В частности, в геометрии Лобачевского существуют треугольники, вокруг которых нельзя описать окружность! (Например, центральный и правый треугольники на рис. 3.)

УПРАЖНЕНИЕ 4. Подумайте, почему в геометрии Лобачевского не проходит евклидово доказательство о существовании описанной окружности треугольника. Как нужно изменить теорему об описанной окружности, чтобы она стала верной в геометрии Лобачевского?

Возникает естественный вопрос: как по элементам треугольника понять, *какую именно* циклическую линию можно описать? В следующем разделе мы дадим ответ на этот вопрос.

## 2. КРИТЕРИЙ ФОРМЫ ОПИСАННОЙ ЦИКЛИЧЕСКОЙ ЛИНИИ ТРЕУГОЛЬНИКА

Мы начнем наши рассуждения со случая орицикла.

Пусть вокруг треугольника  $ABC$  со сторонами<sup>1)</sup>  $a$ ,  $b$  и  $c$  описан орицикл  $\xi$  (рис. 4а). Поскольку сторона  $AB$  имеет наибольшую длину, точка  $C$  лежит на дуге  $\overset{\frown}{AB}$  орицикла  $\xi$ . Это означает, что  $\overset{\frown}{BC} + \overset{\frown}{AC} = \overset{\frown}{AB}$ .

<sup>1)</sup>В этом разделе мы считаем, что  $c$  — наибольшая сторона в треугольнике.



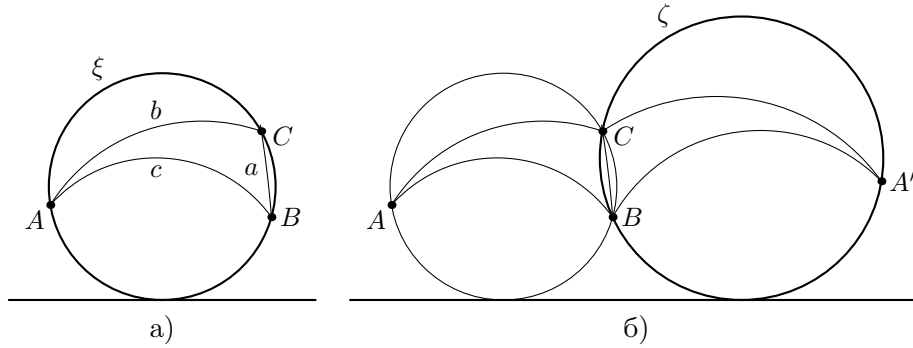


Рис. 4.

Теперь воспользуемся формулой длины дуги орицикла  $\overset{\smile}{AB}$

$$\overset{\smile}{AB} = 2 \operatorname{sh} \frac{\rho(A, B)}{2}$$

(см. [3]) и получим необходимое условие для существования описанного орицикла

$$\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} = \operatorname{sh} \frac{c}{2}. \quad (1)$$

Оно же является и достаточным. В самом деле, пусть в треугольнике  $ABC$  стороны  $a, b$  и  $c$  удовлетворяют соотношению (1). Рассмотрим орицикл  $\zeta$ , проходящий через точки  $B$  и  $C$  (рис. 4б). Длина дуги  $\overset{\smile}{BC}$ , очевидно, равна  $2 \operatorname{sh} \frac{a}{2}$ . Отметим на орицикле  $\zeta$  точку  $A'$ , такую, что  $A'C = b$  и точка  $C$  лежит на дуге  $\overset{\smile}{BA'}$ . Тогда  $\overset{\smile}{A'C} = 2 \operatorname{sh} \frac{b}{2}$ , а значит, длина дуги  $\overset{\smile}{A'B}$  равна  $2 \operatorname{sh} \frac{a}{2} + 2 \operatorname{sh} \frac{b}{2} = 2 \operatorname{sh} \frac{c}{2}$ . Отсюда следует, что  $A'B = c$ . Итак, мы доказали, что  $\triangle ABC = \triangle A'BC$ , а значит, описанной циклической линией треугольника  $ABC$  является орицикл.

Теперь подумаем, каким будет критерий в случае окружности или эквидистанты. При непрерывном изменении положений вершин треугольника величина  $\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} - \operatorname{sh} \frac{c}{2}$  изменяется непрерывно. Отсюда нетрудно вывести, используя (1), что знак этой величины одинаков для всех треугольников, вписанных в окружность, и для всех треугольников, вписанных в эквидистанту. Поэтому достаточно найти знак выражения  $\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} - \operatorname{sh} \frac{c}{2}$  для какого-то одного треугольника, вписанного в данную циклическую линию.

УПРАЖНЕНИЕ 5. Проверьте неравенство  $\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} - \operatorname{sh} \frac{c}{2} > 0$  для какого-нибудь треугольника, вписанного в окружность.

УПРАЖНЕНИЕ 6. Проверьте неравенство  $\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} - \operatorname{sh} \frac{c}{2} < 0$  для какого-нибудь треугольника, вписанного в эквидистанту.

Таким образом, окончательный критерий формы описанной циклической линии выглядит следующим образом.

ТЕОРЕМА 2. 1) *Вокруг треугольника можно описать окружность тогда и только тогда, когда*

$$\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} > \operatorname{sh} \frac{c}{2}.$$

2) *Вокруг треугольника можно описать орицикл тогда и только тогда, когда*

$$\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} = \operatorname{sh} \frac{c}{2}.$$

3) *Вокруг треугольника можно описать эквидистанту тогда и только тогда, когда*

$$\operatorname{sh} \frac{a}{2} + \operatorname{sh} \frac{b}{2} < \operatorname{sh} \frac{c}{2}.$$

### 3. ВНЕВПИСАННЫЕ ЦИКЛИЧЕСКИЕ ЛИНИИ

Помимо описанной окружности, в евклидовой геометрии есть еще несколько замечательных окружностей, связанных с треугольником: вписанная окружность и три внеписанные окружности. Естественный вопрос: *существуют ли эти окружности в геометрии Лобачевского?*

Оказывается, что со вписанной окружностью все в порядке: а именно, в любой неевклидов треугольник можно вписать единственную окружность, центр которой совпадает с точкой пересечения биссектрис треугольника.

УПРАЖНЕНИЕ 7. Докажите это утверждение.

А вот с внеписанными окружностями ситуация похожа на ситуацию с описанной окружностью. Например, существуют треугольники, для которых нет *ни одной* внеписанной окружности (попробуйте нарисовать соответствующий пример). Однако верна следующая теорема.

ТЕОРЕМА 3. *У любого треугольника существуют три внеписанные циклические линии, т. е. три циклические линии, касающиеся одной из сторон треугольника и продолжений двух других сторон (рис. 5).*

УПРАЖНЕНИЕ 8. Докажите эту теорему.

Естественным представляется вопрос о форме внеписанной циклической линии, касающейся фиксированной стороны треугольника. Из теоремы 2 можно очень просто получить критерий, позволяющий определить форму внеписанной циклической линии.

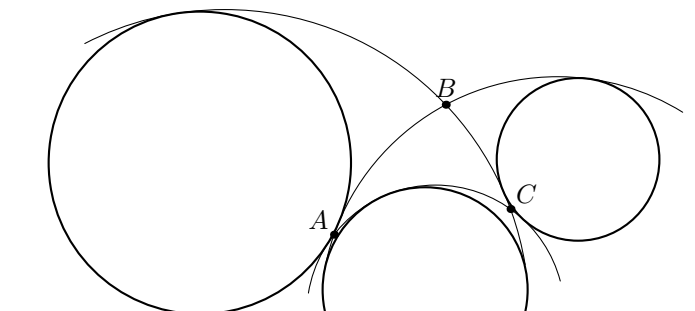


Рис. 5.

Пусть  $ABC$  — произвольный треугольник и  $\omega$  — его внеписанная циклическая линия, касающаяся стороны  $AB$  и продолжений сторон  $CA$  и  $CB$  в точках  $C_1$ ,  $B_1$  и  $A_1$  соответственно (рис. 6). Заметим, что  $\omega$  является описанной циклической линией треугольника  $A_1B_1C_1$ , поэтому к нему можно применить теорему 2 и получить искомый критерий. Для этого нужно лишь вычислить длины сторон треугольника  $A_1B_1C_1$ .

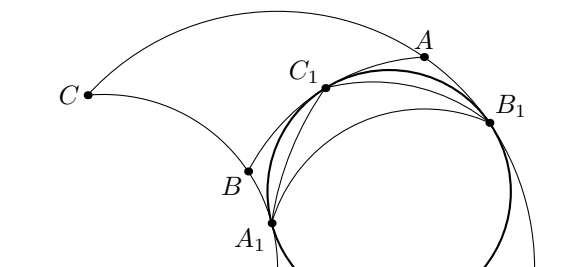


Рис. 6.

Итак, пусть стороны треугольника  $A_1B_1C_1$  равны  $a_1$ ,  $b_1$  и  $c_1$ , стороны треугольника  $ABC$  равны  $a$ ,  $b$  и  $c$ , а углы равны  $\alpha$ ,  $\beta$  и  $\gamma$ .

УПРАЖНЕНИЕ 9. Применив гиперболическую теорему синусов, докажете следующие равенства:

$$\operatorname{sh} \frac{a_1}{2} = \operatorname{sh}(p - b) \cos \frac{\alpha}{2}, \quad \operatorname{sh} \frac{b_1}{2} = \operatorname{sh}(p - a) \cos \frac{\beta}{2}, \quad \operatorname{sh} \frac{c_1}{2} = \operatorname{sh} p \sin \frac{\gamma}{2}.$$

Значит, форма внеписанной линии определяется знаком выражения

$$\operatorname{sh}(p - b) \cos \frac{\alpha}{2} + \operatorname{sh}(p - a) \cos \frac{\beta}{2} - \operatorname{sh} p \sin \frac{\gamma}{2}.$$

Однако этот критерий можно упростить (см. [2, с. 75]).

УПРАЖНЕНИЕ 10. Докажите, что для треугольника с вневписанным орициклом выполняются равенства

$$\frac{\operatorname{sh}(p-b) \cos \alpha/2}{\sin \alpha/2} = \frac{\operatorname{sh}(p-a) \cos \beta/2}{\sin \beta/2} = \frac{\operatorname{sh} p \sin \gamma/2}{\cos \gamma/2} = 1.$$

Поэтому критерий приобретает вид, «двойственный» теореме 2.

ТЕОРЕМА 4. Пусть  $ABC$  — произвольный треугольник, и  $\omega$  — его вневписанная циклическая линия, касающаяся стороны  $AB$  и продолжений сторон  $CA$  и  $CB$ . Тогда

1)  $\omega$  является окружностью тогда и только тогда, когда

$$\sin \frac{\alpha}{2} + \sin \frac{\beta}{2} > \cos \frac{\gamma}{2};$$

2)  $\omega$  является орициклом тогда и только тогда, когда

$$\sin \frac{\alpha}{2} + \sin \frac{\beta}{2} = \cos \frac{\gamma}{2};$$

3)  $\omega$  является эквидистантой тогда и только тогда, когда

$$\sin \frac{\alpha}{2} + \sin \frac{\beta}{2} < \cos \frac{\gamma}{2}.$$

#### СПИСОК ЛИТЕРАТУРЫ

- [1] С. Г. Гиндикин. *Рассказы о физиках и математиках*. М.: МЦНМО, НМУ, 2001.
- [2] В. В. Прасолов. *Геометрия Лобачевского*. Изд. 3е. М.: МЦНМО, 2004.
- [3] А. С. Смогоржевский. *О геометрии Лобачевского*. М.: ГИТТЛ, 1957.

## Вписанные четырёхугольники и трапеции в абсолютной геометрии

Ф. В. Петров

Обсуждаются аналоги вписанных четырёхугольников и трапеций  
в абсолютной геометрии.

Следующие известные еще древним грекам теоремы классической (евклидовой) геометрии играют в ней роль, которую нелегко преувеличить:

ТЕОРЕМА О ВПИСАННОМ ЧЕТЫРЕХУГОЛЬНИКЕ. Следующие свойства выпуклого четырёхугольника  $ABCD$  на плоскости равносильны:

- (i) точки  $A, B, C, D$  лежат на одной окружности;
- (ii)  $\angle ABD = \angle ACD$ ;
- (iii)  $\angle ABC + \angle ADC = \angle BAD + \angle BCD = \pi$ .

ТЕОРЕМА О ТРАПЕЦИИ. Следующие свойства выпуклого четырёхугольника  $ABCD$  на плоскости равносильны:

- (i)  $BC \parallel AD$ ;
- (ii)  $\angle BCA = \angle CAD$ ;
- (iii)  $\angle ABC + \angle BAD = \angle BCD + \angle CDA = \pi$ .

Доказательства обеих теорем существенно используют пятый постулат Евклида о единственности прямой, параллельной данной, проходящей через данную точку вне прямой. Теорема о трапеции по существу ему равносильна.

Однако оказывается, что эти теоремы можно видоизменить так, чтобы они выполнялись и в абсолютной геометрии (то есть доказать без использования пятого постулата).

Это приятное обстоятельство позволяет перенести многие евклидовы теоремы, использующие вписанные четырёхугольники и трапеции, на случай плоскости Лобачевского. Впечатляющие примеры такого рода читатель может найти в статье А. В. Акопяна [1] в этом сборнике.

Автор вполне допускает, что эти теоремы были получены давно, еще до открытия неевклидовой геометрии, в попытках доказать пятый постулат. По крайней мере, для этого были все возможности. Часть теоремы о вписанном четырёхугольнике фигурирует в задачнике З. А. Скопеца и В. А. Жарова [3] и в книге В. В. Прасолова и В. М. Тихомирова [2].

Всюду в нашей заметке, если не оговорено противное, речь идет о плоскости в абсолютной геометрии. Напомним, что в абсолютной геометрии верны такие факты, как признаки равенства треугольников, неравенство треугольника, свойство «против большей стороны лежит больший угол», а сумма углов любого треугольника не превосходит  $\pi$  (развернутого угла).

Начнем со следующего определения-теоремы, обобщающего понятие вписанного четырехугольника.

**ТЕОРЕМА 1 («Вписанный четырехугольник»).** *Следующие свойства выпуклого четырехугольника  $ABCD$  равносильны:*

- (i)  $\angle DBC + \angle BDA = \angle CAD + \angle ACB$ ;
- (ii)  $\angle ABD + \angle BDC = \angle BAC + \angle ACD$ ;
- (iii)  $\angle BAD + \angle BCD = \angle ABC + \angle ADC$ .

Четырехугольники с такими свойствами будем называть «вписанными».

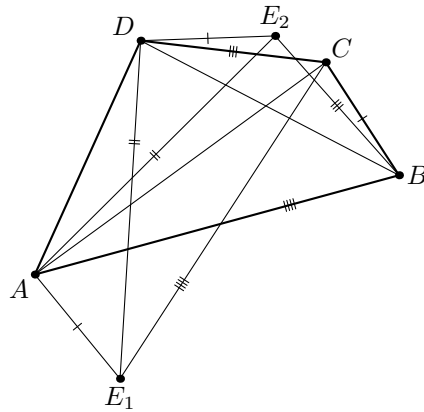


Рис. 1.

**ДОКАЗАТЕЛЬСТВО.** Покажем, как, например, из (i) следует (ii) (заметим, что равенство (iii) есть сумма (i) и (ii), так что любые два из трех равенств влекут третье автоматически). Пусть  $E_1$  — точка, симметричная  $B$  относительно серединного перпендикуляра к  $AC$ , а  $E_2$  — точка, симметричная  $C$  относительно серединного перпендикуляра к  $BD$ . Тогда равны пары треугольников  $E_1AC$  и  $BCA$ ;  $E_2DB$  и  $CBD$ . Имеем

$$\begin{aligned} \angle E_1AD &= \angle E_1AC + \angle CAD = \angle ACB + \angle CAD = \angle DBC + \angle BDA = \\ &= \angle BDE_2 + \angle BDA = \angle ADE_2, \end{aligned}$$

так что треугольники  $E_1AD$  и  $E_2DA$  равны по двум сторонам и углу

между ними, откуда  $E_1D = E_2A$ . Далее, треугольники  $E_1CD$  и  $ABE_2$  равны по трем сторонам, так что

$$\begin{aligned} \angle ABD + \angle BDC &= \angle ABD + \angle DBE_2 = \angle ABE_2 = \angle E_1CD = \\ &= \angle E_1CA + \angle ACD = \angle BAC + \angle ACD, \end{aligned}$$

что и требовалось доказать.

Вывод (i) и (ii) из (iii) аналогичен: надо рассмотреть точки, симметричные  $C$  и  $B$  относительно серединных перпендикуляров к  $AB$  и  $CD$  соответственно.

Заметим, что если четырехугольник  $ABCD$  действительно вписан в окружность с центром  $O$ , то он удовлетворяет этим равенствам. В самом деле, если, для определенности,  $O$  лежит внутри  $ABCD$ , то каждая из сумм  $\angle ABC + \angle ADC$  и  $\angle BAD + \angle BCD$  равна сумме углов при основаниях равнобедренных треугольников  $AOB$ ,  $BOC$ ,  $COD$ ,  $DOA$ . К сожалению, обратное неверно: существуют вписанные четырехугольники, не вписанные, однако, ни в какую окружность. Это связано с тем, что не любой треугольник имеет описанную окружность (серединные перпендикуляры к сторонам не обязаны пересекаться). Отметим, что на плоскости Лобачевского есть и другие кривые, помимо окружностей, для которых вписанные в них четырехугольники удовлетворяют требованиям теоремы. Читатель, хорошо знакомый с геометрией Лобачевского, догадался, что речь об *эквилидистантах*. Напомним, что эквидистанта задается прямой  $l$ , полуплоскостью с границей  $l$  и расстоянием  $d$  как геометрическое место точек в полуплоскости, расстояние от которых до  $l$  равно  $d$ . Несложно видеть, что вписанный в эквидистанту четырехугольник  $ABCD$  удовлетворяет например (iii). Доказательство такое же, как для вписанного в окружность четырехугольника, только роль равнобедренных треугольников с вершиной в центр окружности играют равнобедренные четырехугольники Саккери  $AVH_bH_a$  и т. д. ( $H_a$ ,  $H_b$  — проекции  $A$ ,  $B$  на  $l$ ).

Прояснить сходство между окружностями и эквидистантами проще всего, рассмотрев модель Пуанкаре плоскости Лобачевского, например, в полуплоскости. Тогда окружности будут выглядеть как евклидовы окружности, не пересекающие абсолют, а эквидистанты — как евклидовы окружности, пересекающие абсолют в двух точках. Любой «вписанный четырехугольник» вписан либо в окружность, либо в эквидистанту, либо в орицикл — окружность, касающуюся абсолюта (здесь мы называем евклидовыми окружностями все евклидовы окружности и прямые, не перпендикулярные абсолюту).

Обозначим через  $\sphericalangle XYZ$  величину  $\angle XYZ - \angle YXZ - \angle YZX$ .

Заметим, что теорему 1 можно переформулировать как равносильность равенств  $\sphericalangle ABD = \sphericalangle ACD$ ,  $\sphericalangle ADB = \sphericalangle ACB$ ,  $\sphericalangle ABC = -\sphericalangle ADC$

и т. д. Если для фиксированной пары точек  $X, Z$  считать, что  $\sphericalangle XYZ = \sphericalangle XYZ - \sphericalangle YXZ - \sphericalangle YZX$  для  $Y$  по одну сторону от  $XZ$  и  $\sphericalangle XYZ = -\sphericalangle XYZ + \sphericalangle YXZ + \sphericalangle YZX$  по другую сторону, то «окружностью» (то есть окружностью или эквидистантой), проходящей через точки  $A$  и  $B$ , будет геометрическое место точек  $Z$  таких, что  $\sphericalangle AZB = \text{const}$ .

УПРАЖНЕНИЕ.. Докажите, что треугольник  $ABC$  определяется значениями  $AB, AC, \sphericalangle BAC$  однозначно с точностью до равенства.

Перейдем к обобщению теоремы о трапеции.

ТЕОРЕМА 2 («ТРАПЕЦИЯ»). Следующие свойства выпуклого четырехугольника  $ABCD$  равносильны:

- (i)  $\sphericalangle BAD + \sphericalangle ABC = \sphericalangle BCD + \sphericalangle ADC$ ;
- (ii)  $\sphericalangle CBD + \sphericalangle CAD = \sphericalangle BCA + \sphericalangle BDA$ .

Такие четырехугольники будем называть «трапециями» с основаниями  $BC$  и  $AD$ .

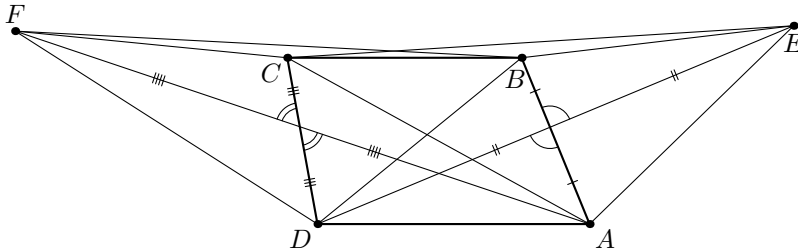


Рис. 2.

ДОКАЗАТЕЛЬСТВО. Выведем (ii), предполагая (i). Пусть точка  $E$  симметрична точке  $D$  относительно середины  $AB$ , точка  $F$  симметрична  $A$  относительно середины  $CD$ . Равны пары треугольников  $ABE$  и  $BAD$ ;  $CDA$  и  $DCF$ . Далее, по двум сторонам и углу равны треугольники  $EBC$  и  $FCB$  ( $\sphericalangle EBC = \sphericalangle EBA + \sphericalangle ABC = \sphericalangle BAD + \sphericalangle ABC = \sphericalangle BCD + \sphericalangle ADC = \sphericalangle BCD + \sphericalangle DCF = \sphericalangle BCF$ ). Вообще говоря, первое равенство верно только если  $\sphericalangle BAD + \sphericalangle ABC \leq \pi$ . Это так:  $\sphericalangle BAD + \sphericalangle ABC$  есть полусумма углов четырехугольника  $ABCD$ , которая в абсолютной геометрии не превосходит  $2\pi$ ). Значит,  $EC = BF$  и по трем сторонам равны треугольники  $EAC$  и  $FDB$ , так что  $\sphericalangle BAC + \sphericalangle ADB = \sphericalangle EAC = \sphericalangle FDB = \sphericalangle CDB + \sphericalangle ACD$ . Вычитая это из (i) получаем (ii).

Для вывода (i) из (ii) надо рассмотреть точки, симметричные  $D$  относительно середины  $AC$  и  $A$  относительно середины  $BD$  и провести аналогичное вычисление. Предоставим это читателю.



Отметим, что в условиях теоремы треугольники  $ABD$  и  $ACD$  равновелики. Это можно понимать в том смысле, что равны суммы их углов (на плоскости Лобачевского площадь треугольника пропорциональна величине  $\pi -$  (сумма его углов)), а евклидов случай рассматривать отдельно. Но можно понимать это более единообразным и при том элементарным способом: мы доказали, что эти треугольники *равнодополняемы*, то есть дополняются равными треугольниками до многоугольников, которые в свою очередь разбиваются на равные треугольники (именно, невыпуклый шестиугольник  $AEBCOD$  разбивается на два треугольника, равных  $ABD$ , и треугольник  $BCO$ , а с другой стороны — на треугольники  $EBC$ ,  $AEC$ ,  $AOD$ . Аналогичные два разбиения есть и для шестиугольника  $BCFDAO$ .) На плоскости Лобачевского выполнена теорема Больяи – Гервина о том, что равновеликие многоугольники всегда равнодополняемы и даже равноставлены (могут быть разбиты на соответственно равные треугольники) [4], так что вообще говоря и другие теоремы о равновеликости должно быть можно доказывать, не прибегая к понятиям анализа.

Сформулируем также отдельно следующее следствие теоремы 2:

**ЛЕММА 3.** *Выпуклый четырехугольник  $ABCD$  является трапецией с основаниями  $BC$  и  $AD$  (то есть  $\angle A + \angle B = \angle C + \angle D$ ) если и только если треугольники  $ABD$  и  $ACD$  равновелики.*

**ДОКАЗАТЕЛЬСТВО.** Уже доказано, что в трапеции такие треугольники равновелики. Докажем, что если они равновелики, то  $\angle A + \angle B = \angle C + \angle D$ . Предположим, что, например,  $\angle A + \angle B > \angle C + \angle D$ . Пусть точка  $X$  движется по стороне  $CD$  от  $C$  к  $D$ . Заметим, что разность  $\angle ABX + \angle BAD - \angle BXD - \angle XDA$  меняется непрерывно, причем при  $X = C$  она отрицательна, а при  $X$  рядом с  $D$  — положительна (если  $X$  близко к  $D$ , то  $\angle BXD$  стремится к  $\pi - \angle BDC > \pi - \angle CDA$ , так что  $\angle BXD + \angle XDA > \pi \geq (\angle ABX + \angle BAD + \angle BXD + \angle XDA)/2$ , так как сумма углов четырехугольника не превосходит  $2\pi$ ). Значит, найдется такое положение точки  $X$  на отрезке  $CD$ , для которого упомянутая разность равна нулю и  $ABXD$  — трапеция. Но тогда получится, что равновелики треугольники  $ACD$ ,  $ABD$  и  $AXD$ , что невозможно.

### БЛАГОДАРНОСТЬ

Автор благодарен Арсению Акопяну за внимание к работе и полезные обсуждения.

## СПИСОК ЛИТЕРАТУРЫ

- [1] А. В. Акопян. *О некоторых классических конструкциях в геометрии Лобачевского* // Математическое просвещение. Сер. 3, вып. 13, 2009. С. 155–170.
- [2] В. В. Прасолов, В. М. Тихомиров. *Геометрия*. М.: МЦНМО, 1998.
- [3] З. А. Скопец, В. А. Жаров. *Задачи и теоремы по геометрии. Планиметрия*. Учпедгиз, 1962.
- [4] R. Hartshorne. *Geometry: Euclid and Beyond*. Springer, 2000.

## О некоторых классических конструкциях в геометрии Лобачевского

А. В. Акопян\*

В статье пойдет речь об аналогах классических конструкций евклидовой геометрии в геометрии Лобачевского. Мы покажем, что в абсолютной геометрии выполнены такие теоремы как теорема о трёх колпаках, пересечения общих хорд трёх окружностей. Будет приведен аналог прямой и окружности Эйлера треугольника, а также доказана знаменитая теорема Фейербаха.

### 1. ВВЕДЕНИЕ

В последнее десятилетие существенно возрос интерес к так называемой классической геометрии. Было получено много новых и интересных теорем, а также написано множество книжек и статей. Однако, практически никак не исследованы даже самые элементарные (по нынешним меркам) конструкции в геометрии Лобачевского и в сферической геометрии.

В данной статье мы сформулируем несколько аналогов классических теорем евклидовой геометрии. Основным результатом является аналог теоремы Фейербаха о касании вписанной окружности и окружности девяти точек Эйлера (напомним, что эта окружность проходит через середины сторон, точки падения высот, а также через середины отрезков, соединяющих вершины с ортоцентром треугольника).

Формулировку аналога этой теоремы мы дадим в следующем разделе. В разделе 3 мы объясняем смысл этой аналогии. В разделе 4 мы докажем существование окружности Эйлера. Далее в разделе 5 будут построены аналоги понятий степени точки, инверсии и гомотетии, с помощью которых будет доказано существование прямой Эйлера в разделе 6. В разделе 7 мы докажем теорему Фейербаха, а в разделе 8 — теорему о «трёх колпаках» и с помощью неё докажем одно интересное свойство точек Фейербаха.

---

\*Работа была проделана при поддержке РФФИ, грантов №06-01-00648 и №08-01-00565-а, а также Фонда поддержки молодых ученых «Конкурс Мёбиуса».

Сразу отметим, что в дальнейших рассуждениях мы будем пользоваться тем, что любые две прямые пересекаются и всегда существует окружность, проходящая через любые три точки. В геометрии Лобачевского это, вообще говоря, неверно. Прямые могут расходиться, а окружности могут вырождаться в орициклы или эквидистанты (которые можно считать окружностями с центрами на абсолюте и за <sup>1</sup>). Рассмотрение «расходящихся» случаев только усложнит доказательство, хотя не будет содержать никаких принципиально новых идей. Для придания полноты рассуждениям мы всегда можем сказать, что данная теорема следует из теоремы об аналитическом продолжении, поскольку рассмотренные нами случаи достаточно обширны (содержат внутреннюю точку в конфигурационном пространстве).

Действительно, геометрию Лобачевского можно пополнить псевдоточками и поставить им в соответствие пучки прямых (прямых, перпендикулярных некоторой одной прямой). В модели Клейна это прямые, проходящие через одну точку, а в модели Пуанкаре окружности, содержащие эти прямые, должны принадлежать одному пучку. Поэтому утверждения, сформулированные на языке «точки лежат на одной прямой» и «прямые пересекаются в одной точке», в модели Пуанкаре можно перевести на язык «пучки содержат общий элемент» (уравнения задающие их линейно зависимы) или «прямые принадлежат одному пучку». Окружностями можно считать множество пучков (то есть «точек»), образующихся при симметриях относительно прямых из другого пучка («центра» окружности). Так различные утверждения можно записывать на аналитическом языке и смотреть на область, в которой выполняются требуемые равенства, указывающие, например, на совпадение некоторых точек. Поскольку все встречающиеся функции будут аналитическими, то обращение в ноль на области, размерность которой совпадает с размерностью конфигурационного пространства, означает тождественное обращение в ноль.

Тем не менее, по ходу рассуждений мы будем стараться обойтись «малой кровью» и показывать, как можно полностью разобрать утверждения, не прибегая к столь мощным инструментам.

## 2. ФОРМУЛИРОВКА

В 1822 году Карл Вильгельм Фейербах обнаружил, что окружность девяти точек Эйлера касается вписанной и трёх невписанных окружностей треугольника. Эта теорема теперь называется теоремой Фейербаха. У неё есть множество интересных и достаточно разных доказательств.

---

<sup>1</sup>Подробнее об этом вы можете прочесть в статьях Ф. В. Петрова [7] и П. В. Бибикова [3] в этом сборнике.

Например, она сразу следует из обращения теоремы Кэзи (обобщения теоремы Птолемея). Другое доказательство можно получить, проделав инверсию с центром в середине одной из сторон треугольника.

Не такое простое, но элементарное и очень содержательное, доказательство этой теоремы было получено Владимиром Протасовым в [10], где он показал, что теорема Фейербаха является частным случаем так называемой «теоремы о сегменте».

У теоремы Фейербаха есть много интересных обобщений, см., например, [5] или [6].

В 2008 г. П. В. Бибииков сформулировал гипотезу, которая является обобщением теоремы Фейербаха на случай геометрии Лобачевского.

*ГИПОТЕЗА (ПАВЕЛ БИБИИКОВ). Обозначим через  $M_a$ ,  $M_b$  и  $M_c$  основания биссекторов (чевиан, делящих площадь треугольника пополам) треугольника  $ABC$  в геометрии Лобачевского. Тогда окружность, проходящая через точки  $M_a$ ,  $M_b$  и  $M_c$ , касается вписанной и трёх внеписанных окружностей треугольника  $ABC$ .*

Назовем эту окружность *окружностью Эйлера* треугольника  $ABC$ .

Мы докажем эту гипотезу. Кроме того, будет показано существование прямой, на которой лежат центр описанной окружности треугольника, точка пересечения биссекторов, центр окружности Эйлера, а также точка пересечения чевиан, соединяющих вершины треугольника с повторными точками пересечения окружности Эйлера со сторонами. По мнению автора, этой прямой вполне подходит название прямой Эйлера.

### 3. ПСЕВДОМЕДИАНЫ И ПСЕВДОВЫСОТЫ

Прежде чем говорить о теореме Фейербаха, следует разобраться, через какие шесть точек на сторонах треугольника проходит окружность Эйлера.

Биссекционные отрезки, как и медианы в евклидовой плоскости, делят площадь треугольника пополам. В геометрии Лобачевского это означает, что равны суммы углов двух треугольников, на которые биссектор делит исходный треугольник. После этого естественно сформулировать какое-нибудь угловое определение высот евклидового треугольника и перенести это определение на случай геометрии Лобачевского.

Пусть  $H_a$  — основание высоты, опущенной из вершины  $A$  на сторону  $BC$ . Тогда в евклидовом случае будет выполнено следующее равенство:

$$\begin{aligned} \angle AH_a B - (\angle H_a A B + \angle A B H_a) &= \angle AH_a C - (\angle H_a A C + \angle A C H_a) \\ &= \frac{1}{2}(\pi - (\angle A + \angle B + \angle C)) = 0. \end{aligned} \quad (1)$$

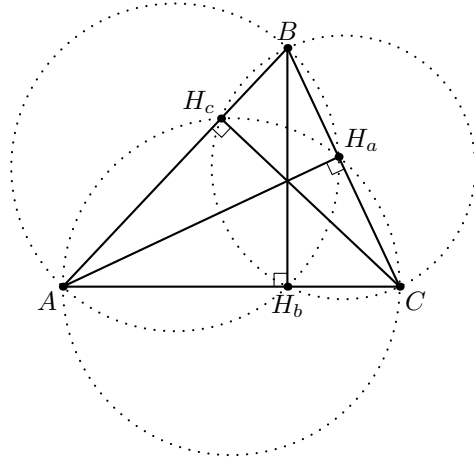


Рис. 1.

Назовем *псевдовысотами* чевианы с аналогичным свойством в геометрии Лобачевского. Убедимся в существовании псевдовысот.

Обозначим через  $\sphericalangle XYZ$  величину  $\sphericalangle XYZ - \sphericalangle ZXY - \sphericalangle YZX$ , где углы берутся со знаком:  $\sphericalangle XYZ > 0$  тогда и только тогда, когда поворот по кратчайшей дуге от  $X$  к  $Z$  вокруг  $Y$  происходит против часовой стрелки.

Рассмотрим треугольник  $ABC$ , в котором вершины  $A, B, C$  идут по часовой стрелке. Используя указанное выше правило знаков для углов, можно проверить, что при движении точки  $X$  на прямой  $BC$  в направлении от  $B$  к  $C$  величина  $\sphericalangle BXA$  непрерывно уменьшается, а величина  $\sphericalangle AXC$  непрерывно увеличивается, причем  $\sphericalangle BXA - \sphericalangle AXC > 0$  для достаточно далеких точек  $X$  на луче  $CB$  и  $\sphericalangle BXA - \sphericalangle AXC < 0$  для достаточно далеких точек  $X$  на луче  $B$ . Поэтому найдется такая точка  $H_a$ , для которой  $\sphericalangle BXA = \sphericalangle AXC$ . Это и есть основание псевдовысоты на прямой  $BC$ .

Обозначим через  $H_b$  и  $H_c$  основания псевдовысот на прямых  $CA$  и  $AB$  соответственно. Для них будут выполняться равенства  $\sphericalangle CH_bB = \sphericalangle BH_bA$  и  $\sphericalangle AH_cC = \sphericalangle CH_cB$ .

В следующем разделе мы покажем, что через эти три точки проходит окружность Эйлера.

Заметим, что

$$\sphericalangle BH_aA + \sphericalangle AH_aC = \sphericalangle CH_bB + \sphericalangle BH_bA = \sphericalangle AH_cC + \sphericalangle CH_cB = S_{\triangle ABC},$$

откуда получаем

$$\sphericalangle BH_aA = \sphericalangle AH_aC = \sphericalangle CH_bB = \sphericalangle BH_bA = \sphericalangle AH_cC = \sphericalangle CH_cB = \frac{1}{2}S_{\triangle ABC}.$$

Несложно доказать в абсолютной геометрии, что четырехугольники  $ABH_aH_b$ ,  $ACH_aH_c$  и  $BCH_bH_c$  являются вписанными (правильное определение вписанного четырехугольника см. в статье Ф. В. Петрова [7] в этом выпуске). Для этого воспользуемся аналогом теоремы о вписанном угле. Эта теорема есть во многих книжках (см. например [12], или [11], или упомянутую статью [7]), тем не менее, мы здесь напомним её доказательство.

ЛЕММА 4. Для любой точки  $X$  на дуге  $AB$  некоторой фиксированной окружности, величина  $\sphericalangle AXB$  постоянна.

ДОКАЗАТЕЛЬСТВО. Пусть  $O$  — центр этой окружности, а  $X$  — произвольная точка на дуге  $AB$ . Легко видеть, что  $O$  лежит на серединном перпендикуляре к  $AB$ , а кроме того выполнено равенство

$$\sphericalangle OAB = \frac{1}{2}(\sphericalangle AXB - \sphericalangle BAX - \sphericalangle XBA) = \frac{1}{2}\sphericalangle AXB. \quad (2)$$

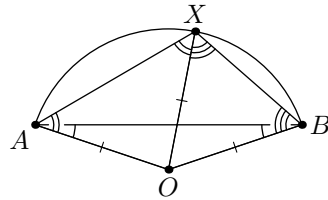


Рис. 2.

То есть величина  $\sphericalangle AXB$  не зависит от того, какую точку  $X$  на дуге мы берем, и всегда равна  $2\sphericalangle OAB$ .

Можно показать и обратное, что если  $\sphericalangle AXB = \sphericalangle AYB$ , то четыре точки  $A, B, X$  и  $Y$  лежат на одной окружности в модели Пуанкаре, т. е. на *циклической линии* — окружности, эквидистанте или орицикле.

В дальнейшем нам пригодится еще такое следствие теоремы о вписанном угле:

ЛЕММА 5. Существует единственная тройка точек  $X_a, X_b$  и  $X_c$  на сторонах треугольника  $ABC$ , что четырехугольники  $ABX_aX_b$ ,  $ACX_cX_a$  и  $BCX_cX_b$  вписанные.

ДОКАЗАТЕЛЬСТВО. Действительно, из теоремы о вписанном угле получаем

$$\begin{aligned} \sphericalangle BX_aA = \sphericalangle BX_bA = S_{\triangle ABC} - \sphericalangle CX_bB = S_{\triangle ABC} - \sphericalangle CX_cB = \\ = \sphericalangle AX_cC = \sphericalangle AX_aC = S_{\triangle ABC} - \sphericalangle BX_aA. \end{aligned} \quad (3)$$

То есть  $\sphericalangle BX_aA = \frac{1}{2}S_{\triangle ABC} = \sphericalangle BH_aA$ . Значит, точки  $H_a$  и  $X_a$  совпадают. Аналогично рассуждаем для двух других точек.

## 4. ПАРАЛЛЕЛЬНОСТЬ И АНТИПАРАЛЛЕЛЬНОСТЬ

Напомним, что для евклидова угла  $XOY$  можно определить понятие антипараллельности. Пусть на прямой  $OX$  заданы точки  $A$  и  $B$ , а на прямой  $OY$  точки  $C$  и  $D$ , и пусть четырехугольник  $ABCD$  будет вписанным. В этом случае мы назовем отрезки  $AD$  и  $BC$  антипараллельными. В евклидовой геометрии имеется простой критерий вписанности четырехугольника, а именно сумма противоположных углов должна равняться  $\pi$ . Поэтому легко понять, что если два отрезка  $A'D'$  и  $B'C'$  параллельны двум антипараллельным отрезкам  $AD$  и  $BC$  соответственно, то они также будут антипараллельны между собой.

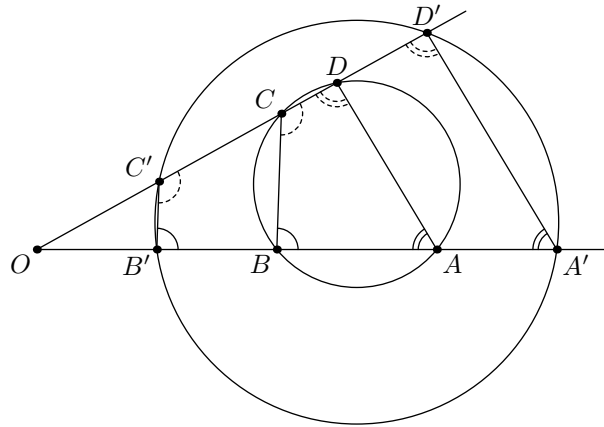


Рис. 3.

Перенесем это понятие на случай плоскости Лобачевского. Так же, как и в евклидовом случае, назовем отрезки  $AD$  и  $BC$  антипараллельными, если четырехугольник  $ABCD$  — вписанный. Аналогично евклидовому случаю, можно воспользоваться критерием вписанности и получить, что отрезки антипараллельны тогда и только тогда, когда  $\angle DAO - \angle ODA = \angle OCB - \angle CBO$ , что равносильно  $\sphericalangle DAO = \sphericalangle OCB$ .

Теперь уже можно ввести понятие параллельности отрезков относительно угла. Отрезки  $BC$  и  $AD$  параллельны, если

$$\angle DAO - \angle ODA = \angle CBO - \angle OCB, \text{ что равносильно } \sphericalangle DAO = -\sphericalangle OCB. \quad (4)$$

Понятно, что любые два отрезка, антипараллельные некоторому отрезку, будут параллельны между собой.

Заметим, что в статье Ф. В. Петрова [7] четырехугольники  $ABCD$  с параллельными отрезками  $AD$  и  $BC$  называются



ЛЕММА 6 ([7, ЛЕММА 3]). *Площади треугольников  $ABC$  и  $ABD$  равны тогда и только тогда, когда выполнено следующее условие на сумму углов четырехугольника  $ABCD$ :*

$$\angle A + \angle D = \angle B + \angle C. \quad (5)$$

Из этой леммы мы получаем, что если  $M_a$  и  $M_b$  — основания биссекционных отрезков, то отрезки  $M_aM_b$  и  $AB$  параллельны относительно угла  $ACB$ . Поскольку же  $H_aH_b$  антипараллелен  $AB$  относительно того же угла, то  $H_aH_b$  антипараллелен  $M_aM_b$ .

Теперь легко доказать, что  $H_a$ ,  $H_b$  и  $H_c$  являются повторными точками пересечения окружности  $M_aM_bM_c$  со сторонами треугольника. Действительно, обозначим повторные точки пересечения со сторонами за  $X_a$ ,  $X_b$  и  $X_c$ . Тогда  $X_aX_b$  будет антипараллелен  $M_aM_b$ , а значит антипараллелен и  $AB$ , т. е. четырехугольник  $X_aX_bAB$  будет вписанным. Аналогичное вписанными будут и четырехугольники  $ACX_aX_c$  и  $BCX_bX_c$ . Но по лемме 5 такая тройка точек  $X_a$ ,  $X_b$ ,  $X_c$  единственна и это тройка  $H_a$ ,  $H_b$  и  $H_c$ .

В модели Пуанкаре можно указать на естественную связь между параллельностью и антипараллельностью. Последняя, как мы помним, соответствует случаю вписанного четырехугольника. Существует теорема Лексея, доказательство которой можно найти например в [13].

ТЕОРЕМА 7 (ТЕОРЕМА ЛЕКСЕЯ В ГЕОМЕТРИИ ЛОБАЧЕВСКОГО). *Пусть точки  $A^*$  и  $B^*$  инверсны точкам  $A$  и  $B$  относительно окружности абсолюта модели Пуанкаре. Пусть  $\omega$  — некоторая евклидова окружность, проходящая через точки  $A^*$  и  $B^*$ . Тогда для любой точки  $X$ , лежащей на окружности  $\omega$ , площадь треугольника  $XAB$  (в смысле геометрии Лобачевского) будет постоянной, т. е. не будет зависеть от точки  $X$ .*

Из этой теоремы и леммы 6 следует, что если отрезки  $AB$  и  $DC$  параллельны, то точки  $A^*$ ,  $B^*$ ,  $C$ ,  $D$  лежат на одной окружности в плоскости модели Пуанкаре.

Заметим также, что лемму 6 можно вывести из теоремы 7.

## 5. СТЕПЕНЬ ТОЧКИ, ГОМОТЕТИЯ И ИНВЕРСИЯ

Сопоставим каждому отрезку  $AB$  его *псевдодлину*  $d_E(A, B)$  следующим образом. Поместим точку  $A$  в центр диска Пуанкаре (радиус диска подразумевается равным 1), тогда  $d_E(A, B)$  определим как евклидову длину отрезка  $AB$ .

Нетрудно проверить, что псевдодлина не зависит от выбора движения, переводящего точку  $A$  в центр диска Пуанкаре. Кроме того, псевдодлина симметрична:  $d_E(A, B) = d_E(B, A)$ .

С помощью псевдодлины можно определить понятия степени точки относительно окружности, гомотетии и инверсии. Так, *степенью точки*  $A$  относительно окружности  $\omega$  назовем произведение псевдодлин отрезков, соединяющих  $AB$  и  $AC$ , где  $BC$  — хорда окружности  $\omega$ , проходящая через  $A$ . Корректность этого определения (то есть независимость величины  $d_E(A, B) \cdot d_E(A, C)$  от выбора прямой, проходящей через  $A$ ) становится очевидной после переноса точки  $A$  в центр модели Пуанкаре, в этом случае теорема становится просто евклидовой.

Теперь можно определить радикальную ось двух окружностей как множество точек, у которых степени относительно двух данных окружностей равны.

ЛЕММА 8. *Радикальная ось является прямой.*

Это очевидно для двух пересекающихся окружностей (у них есть общая хорда). Для непересекающихся окружностей можно, например, выйти в пространство, надстроив над окружностями пересекающиеся сферы (радикальной осью которых будет плоскость, проходящая через их пересечение). Подробнее об этой конструкции см. [1].

Отметим, что теперь можно переносить из евклидовой геометрии теоремы, доказательство которых использует радикальные оси. Например, так можно доказать теорему Брианшона для окружности в геометрии Лобачевского (см. доказательство теоремы Брианшона с помощью радикальных осей в [1] или [9]).

Из леммы 8 сразу следует, что псевдовысоты пересекаются в одной точке — радикальном центре окружностей  $ABH_aH_b$ ,  $ACH_aH_c$  и  $BCH_bH_c$ .

Аналогично, точка пересечения биссекторов будет радикальным центром окружностей  $A^*B^*M_aM_b$ ,  $A^*C^*M_aM_c$  и  $B^*C^*M_bM_c$  (поскольку все прямые, проходящие через  $A^*$ , проходят и через  $A$ , то прямая  $A^*M_a$  будет ничем иным как биссектором  $AM_a$ ).

Определим теперь *гомотетию с центром в точке*  $A$  и *коэффициентом*  $k$  как преобразование, переводящее точку  $P$  в точку  $P'$ , лежащую на луче  $AP$  (в случае положительного  $k$ ) и такую, что  $d_E(A, P') = k \cdot d_E(A, P)$ .

Аналогично определим *инверсию с центром в точке*  $A$  и *радиусом*  $r$  как преобразование, переводящее точку  $P$  в точку  $P'$ , лежащую на луче  $AP$  и такую, что  $d_E(A, P') \cdot d_E(A, P) = r^2$ .

Легко видеть, что эти преобразования переводят окружности модели Пуанкаре в окружности, а прямые, проходящие через центр, — в прямые.

## 6. ПРЯМАЯ ЭЙЛЕРА

Обозначим через  $O$  центр описанной окружности треугольника  $ABC$ , через  $E$  — центр окружности Эйлера, через  $M$  и  $H$  — точки пересечения

биссекторов и псевдовысот соответственно. Покажем, что эти точки лежат на одной прямой, которую мы и будем называть *прямой Эйлера*.

Поместим точку  $M$  в центр модели Пуанкаре. Тогда из критерия вписанности для гиперболических и евклидовых четырехугольников следует, что (евклидовы) прямые  $AB$  и  $M_aM_b$  будут параллельны в евклидовом смысле. Следовательно, верны следующие равенства:

$$\frac{d_E(M, M_a)}{d_E(M, A)} = \frac{d_E(M, M_b)}{d_E(M, B)} = \frac{d_E(M, M_c)}{d_E(M, C)}. \quad (6)$$

То есть описанная окружность треугольника  $ABC$  переходит в окружность Эйлера при гиперболической гомотетии с центром в  $M$  и коэффициентом  $-\frac{d_E(M, M_a)}{d_E(M, A)}$ . Конечно, центр описанной окружности при этом не должен переходить в точку  $E$  (т. е. центр образа окружности), но, тем не менее, он обязан лежать на прямой  $ME$ . Кроме того, он должен лежать и на прямой  $MO$ . Получаем, что эти две прямые совпадают и точки  $O$ ,  $E$  и  $M$  лежат на одной прямой.

Аналогично рассуждаем для точки  $H$ . Поскольку  $H$  является радикальным центром окружностей  $ABH_aH_b$ ,  $ACH_aH_c$  и  $BCH_bH_c$  выполнено соотношение

$$d_E(H, H_a) \cdot d_E(H, A) = d_E(H, H_b) \cdot d_E(H, B) = d_E(H, H_c) \cdot d_E(H, C). \quad (7)$$

То есть окружность Эйлера является образом описанной окружности треугольника  $ABC$  после композиции поворота вокруг точки  $H$  на угол  $\pi$  и инверсии с центром в  $H$  и квадратом радиуса  $d_E(H, H_a) \cdot d_E(H, A)$ . Отсюда видно, что  $H$  также лежит на прямой  $OE$ .

Приведенное рассуждение аналогично доказательству существования прямой Эйлера в евклидовой геометрии. Поэтому, по мнению автора, эта прямая может считаться аналогом прямой Эйлера в геометрии Лобачевского.

## 7. ТЕОРЕМА ФЕЙЕРБАХА

Теперь сформулируем теорему о сегменте (см. [10]), которая будет главным инструментом при доказательстве теоремы Фейербаха.

**ТЕОРЕМА 9 (ТЕОРЕМА О СЕГМЕНТЕ, ПРОТАСОВ [10]).** Пусть даны две прямые  $a$  и  $b$  и окружность  $\omega$ , касающаяся их обеих, а также некоторое фиксированное число  $\varphi$ . Каждой касательной  $l$  к  $\omega$  поставим в соответствие окружность  $\omega_l$  следующим образом. Пусть прямая  $l$  пересекает прямые  $a$  и  $b$  в точках  $A$  и  $B$  соответственно, тогда  $\omega_l$  — это такая окружность, проходящая через точки  $A$  и  $B$ , что величина дуги  $AB$  в ней равна  $\varphi$ . Тогда окружность  $\omega_l$  касается двух фиксированных окружностей, каждая из которых в свою очередь касается прямых  $a$  и  $b$ .

Посмотрим на теорему Фейербаха в модели Пуанкаре с евклидовой точки зрения. Забудем про основания биссекторов, они нам не понадобятся, и будем рассматривать только псевдовысоты  $AA'$ ,  $BB'$ ,  $CC'$  треугольника  $ABC$ . Увидим мы следующую теорему:

**ТЕОРЕМА 10.** Пусть даны два такие треугольника  $ABC$  и  $A'B'C'$ , что четырехугольники  $ABA'B'$ ,  $ACA'C'$  и  $BCB'C'$  — вписанные. Тогда существует четыре окружности, касающиеся окружностей, описанных около треугольников  $ABC'$ ,  $AB'C$ ,  $A'BC$  и  $A'B'C'$ .

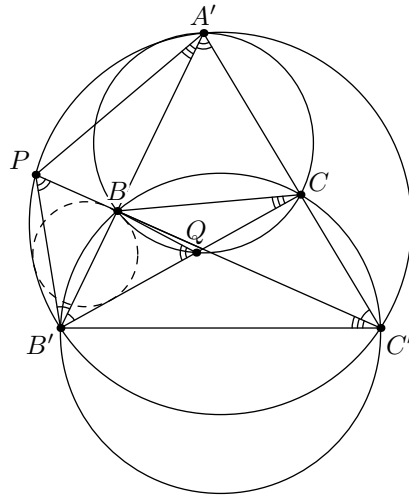


Рис. 4.

**ДОКАЗАТЕЛЬСТВО.** Сделаем инверсию с центром в точке  $A$ . Тогда окружности  $ABC'$  и  $AB'C$  перейдут в прямые, остальные же окружности останутся окружностями. Таким образом, две из этих четырех окружностей можно считать прямыми.

Итак, пусть имеется следующая конструкция. Точки  $B$ ,  $B'$ ,  $C'$  и  $C$  лежат на одной окружности. Прямые  $BB'$  и  $CC'$  пересекаются в точке  $A'$ . Требуется доказать, что существует четыре окружности, касающиеся прямых  $B'C$ ,  $C'B$  и описанных окружностей треугольников  $A'BC$  и  $A'B'C'$ .

Пусть  $P$  — точка повторного пересечения прямой  $C'B$  и окружности  $A'B'C'$ . Аналогично,  $Q$  — точка повторного пересечения прямой  $CB'$  и окружности  $BCA'$ . Тогда с помощью теоремы о вписанном угле получаем, что

$$\angle B'PB = \angle B'A'C' = \angle B'QB \quad \text{и} \quad (8)$$

$$\angle PB'A = \angle PC'A' = \angle BB'Q. \quad (9)$$

Мы видим, что четырехугольник  $B'PVQ$  представляет из себя дельтоид, поэтому существует две окружности, касающиеся его сторон. Заметим также, что

$$\angle PC'B' = \angle BCB'. \quad (10)$$

То есть величины дуг  $PB$  и  $QB'$  в окружностях  $A'C'B'$  и  $A'BC$  равны. А значит, по теореме о сегменте для прямых  $C'B$  и  $CB'$  существуют четыре окружности (по две на каждую вписанную в дельтоид окружность), касающиеся прямых  $B'C$ ,  $C'B$  и окружностей  $A'BC$  и  $A'B'C'$ .

## 8. ТЕОРЕМА О ТРЁХ КОЛПАКАХ И ОДНО СВОЙСТВО ТОЧЕК ФЕЙЕРБАХА

Покажем, что в геометрии Лобачевского выполнена теорема о трёх колпаках, в некотором смысле двойственная к теореме о пересечении радикальных осей трёх окружностей. Она оказывается очень полезной во многих ситуациях. В частности, с ее помощью мы получим одно замечательное свойство точек Фейербаха.

**ТЕОРЕМА 11 (ТЕОРЕМА О ТРЁХ КОЛПАКАХ).** Пусть  $\omega_1$ ,  $\omega_2$  и  $\omega_3$  — три окружности в геометрии Лобачевского. Пусть  $P_1$  — точка пересечения общих внешних касательных к окружностям  $\omega_2$  и  $\omega_3$ . Аналогично определим точки  $P_2$  и  $P_3$ . Тогда точки  $P_1$ ,  $P_2$  и  $P_3$  лежат на одной прямой. Теорема остается верной, если ровно две из трёх точек будут определяться как пересечение внутренних касательных.

Данную теорему можно доказать, непосредственно воспользовавшись гиперболическим аналогом теоремы Менелая. Однако гораздо приятнее перенести классическое евклидово рассуждение, связанное с выходом в трёхмерное пространство.

**ДОКАЗАТЕЛЬСТВО.** Надстроим над каждой окружностью сферу того же радиуса и рассмотрим общую ко всем трем сферам касательную плоскость. Легко понять, что эта плоскость проходит через точки  $P_1$ ,  $P_2$  и  $P_3$ , поскольку содержит образующие соответствующих конусов. Но, поскольку плоскости пересекаются по прямой, касательная плоскость пересекает исходную по прямой, содержащей  $P_1$ ,  $P_2$  и  $P_3$ .

Поскольку точки  $P_1$ ,  $P_2$  и  $P_3$  являются центрами гомотетий соответствующих окружностей, получаем следующую теорему.

**ТЕОРЕМА 12 (ТЕОРЕМА МОНЖА).** Пусть  $\omega_1$ ,  $\omega_2$  и  $\omega_3$  — три окружности. Пусть  $P_1$  — центр положительной гомотетии окружностей  $\omega_2$  и  $\omega_3$ . Аналогично определим точки  $P_2$  и  $P_3$ . Тогда точки  $P_1$ ,  $P_2$  и  $P_3$  лежат на одной прямой. Теорема остается верной, если ровно две из трёх гомотетий будут отрицательными.

В приведенном доказательстве мы существенно опирались на то, что центр такой гомотетии существует (а если не существует, то использовали теорему об аналитическом продолжении). Однако для пересекающихся окружностей или в случае, когда одна окружность находится внутри другой, это, вообще говоря, не так очевидно. Давайте покажем, как можно избежать этих неприятностей.

*Центром гомотетии* окружностей  $\omega_1$  и  $\omega_2$  назовем такую точку, что прямые, проходящие через неё, пересекают  $\omega_1$  и  $\omega_2$  под равными углами. Таких точки две и, в зависимости от того как расположены эти углы относительно прямой, центр будет либо положительный, либо отрицательный.

Легко показать, что если через точку проходит две прямые, пересекающие  $\omega_1$  и  $\omega_2$  под равными углами, то все прямые, проходящие через эту точку, обладают этим свойством (достаточно сделать окружности симметричными относительно диаметра диска Пуанкаре). Одной из таких прямых является прямая, проходящая через центры окружностей. Она пересекает  $\omega_1$  и  $\omega_2$  под прямыми углами. В качестве второй такой прямой можно взять диаметр диска Пуанкаре, проходящий через евклидов центр гомотетии (положительный или отрицательный) двух окружностей (без ограничения общности можно считать, что оба центра окружности не лежат на одном диаметре). Он пересекает окружности под равными углами, а значит, представляет в модели Пуанкаре прямую плоскости Лобачевского, обладающую тем же свойством. Получили две прямые, пересекающие  $\omega_1$  и  $\omega_2$  под равными углами. Их пересечение и будет центром гомотетии.

Отсюда сразу следует теорема Монжа: прямая, пересекающая пары окружностей  $\omega_1$  и  $\omega_2$ , и  $\omega_2$  и  $\omega_3$  под равными углами, также пересекает под равными углами окружности  $\omega_1$  и  $\omega_3$ , т. е. проходит через их центр гомотетии.

Теорема Монжа оказывается чрезвычайно полезной при доказательстве различных утверждений о пересечении нескольких прямых в одной точке или принадлежности нескольких точек одной прямой. Вот одно из таких утверждений.

**ТЕОРЕМА 13.** *Пусть дан треугольник  $ABC$  и окружность  $\omega$ . Пусть окружность  $\omega_a$  касается лучей  $AB$  и  $AC$ , а также окружности  $\omega$  в точке  $P_a$  (внутренним образом). Аналогично определяются окружности  $\omega_b$ ,  $\omega_c$  и точки  $P_b$  и  $P_c$ . Тогда прямые  $AP_a$ ,  $BP_b$  и  $CP_c$  пересекаются в одной точке.*

**ДОКАЗАТЕЛЬСТВО.** Пусть точка  $P$  — это центр положительной гомотетии окружности  $\omega$  и вписанной окружности треугольника  $ABC$ . Тогда из теоремы Монжа следует, что точки  $A$ ,  $P$ ,  $P_a$  лежат на одной прямой. Аналогично рассуждаем для прямых  $BP_b$  и  $CP_c$ . Таким образом, все эти три прямые проходят через точку  $P$ .

Из доказательства становится ясно, что эта теорема допускает ряд вариаций. Например, можно потребовать, что все касания с окружностью  $\omega$  — внешние. Тогда точкой пересечения будет центр отрицательной гомотетии вписанной окружности и  $\omega$ .

Поскольку центр гомотетии вписанной окружности и окружности Эйлера лежит на прямой  $EI$ , из теоремы 13 следует такая теорема:

**ТЕОРЕМА 14.** Пусть на плоскости Лобачевского дан треугольник  $ABC$ . Пусть  $F$ ,  $F_a$ ,  $F_b$  и  $F_c$  — точки касания его окружности Эйлера со вписанной и соответствующими вневыписанными окружностями. Пусть  $I$  — это центр вписанной окружности треугольника. Тогда прямые  $FI$ ,  $AF_a$ ,  $BF_b$  и  $CF_c$  пересекаются в одной точке.

## 9. А ЧТО ЖЕ ПРОИСХОДИТ НА СФЕРЕ?

Читатель, хорошо знакомый с геометрией Лобачевского, знает, что в «алгебраическом смысле» геометрия Лобачевского и сферическая геометрия устроены одинаково. Поэтому все описанные в этой статье рассуждения можно с легкостью применить и к случаю сферической геометрии. Это было бы даже в некотором смысле удобнее по двум причинам. Во-первых, сферическая геометрия «реальна». Сферу легко представить или пощупать (взять, например, глобус), она не требует сложных определений, теоремы выглядят более наглядными и мотивированными. Но при этом сфера обладает одним существенным недостатком: её очень неудобно рисовать на плоскости. Можно сделать, например, стереографическую проекцию и получить модель «сферы на плоскости», которая во многом аналогична модели Пуанкаре. Но получившаяся модель всё равно будет достаточно неудобна для работы хотя бы потому, что прямые очень сложно будет «на глаз» отличить от обычных окружностей (прямыми в этой модели будут окружности, относительно которых степень точки  $(0; 0)$  равна  $-1$ ).

Для иллюстрации «похожести геометрий» покажем, как для сферы доказывается теорема Лекселя о геометрическом множестве таких точек  $X$ , что площадь треугольников  $ABX$  постоянна.

Прежде всего заметим, что эквидистантами на сфере являются обычные окружности (вот уж где действительно видно, что эквидистанты и окружности — это «одно и то же»). Если посмотреть на глобус и считать, что рассматриваемая прямая  $l$  — это экватор, то её эквидистантами будут параллели — окружности с центрами в северном и южном полюсе.

Второе замечание касается того, что будет «аналогом» точек  $A^*$  и  $B^*$  — точек, инверсных  $A$  и  $B$  относительно абсолюта модели Пуанкаре. Заметим, что точка  $A^*$  обладает тем свойством, что любая прямая, проходящая через  $A$  (которая является окружностью в модели Пуанкаре), проходит

также и через точку  $A^*$ . На сфере этим свойством обладает противоположная точка, т. е. точка, симметричная  $A$  относительно центра сферы.

Теперь сформулируем аналог теоремы 7 на сфере.

**ТЕОРЕМА 15 (ТЕОРЕМА ЛЕКСЕЛЯ).** Пусть  $A^*$  и  $B^*$  — это точки противоположных точек  $A$  и  $B$  относительно центра сферы. Пусть  $\omega$  — окружность, проходящая через точки  $A^*$  и  $B^*$ . Тогда для любой точки  $X$ , лежащей на окружности  $\omega$ , площадь треугольника  $ABX$  будет постоянной, т. е. не будет зависеть от точки  $X$ .

**ДОКАЗАТЕЛЬСТВО.** Существует очень простое и изящное доказательство того факта, что множество таких точек  $X$  является эквидистантой. Мы здесь приведем его набросок, подробное доказательство можно найти в [2] или [4].

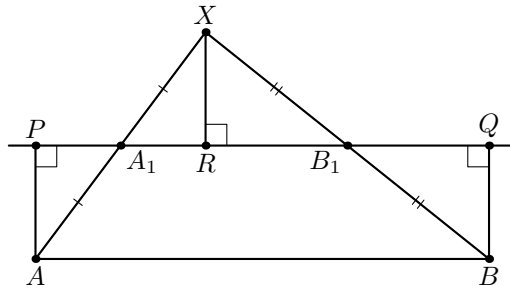


Рис. 5.

Пусть дан треугольник  $ABX$  (см. рис. 5). Обозначим через  $A_1$  и  $B_1$  середины сторон  $AX$  и  $BX$  соответственно, а прямую  $A_1B_1$  за  $l$ . Пусть  $P$ ,  $Q$  и  $R$  проекции точек  $A$ ,  $B$  и  $X$  на прямую  $l$ . Заметим, что прямоугольные треугольники  $APA_1$  и  $XRA_1$  равны по гипотенузе и острому углу. То же самое можно сказать и про треугольники  $BQB_1$  и  $XRB_1$ . Из этого можно сделать два вывода. Во-первых,

$$S_{\triangle A_1XB_1} = S_{\triangle APA_1} + S_{\triangle BQB_1}, \text{ поэтому } S_{\triangle AXB} = S_{\square APQB}.$$

А во-вторых,  $AP = BQ = XR$ , т. е. точки  $A$  и  $B$  лежат на эквидистанте прямой  $l$ , а на симметричной ей эквидистанте лежит точка  $X$ . Обозначим эти эквидистанты за  $\omega'$  и  $\omega$  соответственно.

Обратными рассуждениями можно показать, что для любой точки  $X'$  на окружности  $\omega$  площадь треугольника  $AX'B$  будет равна площади четырехугольника  $APQB$  (отметим, что равенство нужных нам прямоугольных треугольников  $APA'_1$  и  $X'R'A'_1$  будет выполняться по катету и острому углу).

Мы знаем, что на сфере для любой прямой  $l$  (большой окружности) эквидистанты — это обычные окружности, «параллельные» ей. Легко по-



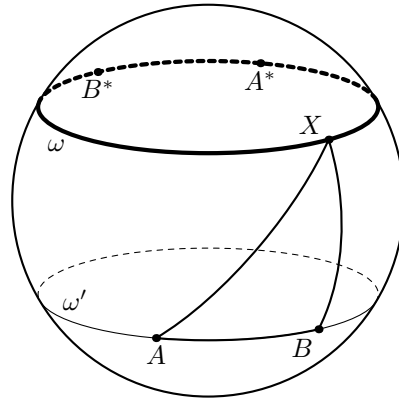


Рис. 6.

нять что  $\omega$  и  $\omega'$  симметричны не только относительно прямой  $l$ , но и также относительно центра сферы, поэтому  $\omega$  будет проходить через точки  $A^*$  и  $B^*$  (т. е. будет той самой окружностью из условия леммы).

Для геометрии Лобачевского последний кусок этого доказательства надо немножко изменить, поскольку неясно, что будет в ней «симметрией относительно центра». Заинтересовавшийся читатель может самостоятельно посмотреть, как преобразование относительно центра сферы выглядит на стереографической проекции, и обнаружить, что это — инверсия относительно круга радиуса  $i$ . Поэтому в модели Пуанкаре логично взять в качестве этого преобразования инверсию относительно абсолюта (окружности радиуса 1). Тогда мы увидим, что  $\omega$  и  $\omega'$  инверсны относительно абсолюта модели Пуанкаре. Особенно легко это понять, если рассмотреть модель в верхней полуплоскости и считать, что прямая  $l$  — это луч, перпендикулярный линии абсолюта.

### БЛАГОДАРНОСТЬ

Автор благодарен Фёдору Петрову и Михаилу Вялому за внимание к работе и полезные обсуждения.

### СПИСОК ЛИТЕРАТУРЫ

- [1] А. В. Акопян, А. А. Заславский. *Геометрические свойства кривых второго порядка*. М.: МЦМНО, 2007.
- [2] Л. С. Атанасян *Геометрия Лобачевского*. М.: Просвещение, 2001.

- [3] П. В. Бибииков. *Описанные циклические линии треугольника в геометрии Лобачевского* // Математическое просвещение. Сер. 3, вып. 13, 2009. С. 142–148.
- [4] П. В. Бибииков, И. В. Ткаченко. *О трисекции и бисекции треугольника на плоскости Лобачевского* // Математическое просвещение. Сер. 3, вып. 11, 2007. С. 113–126.
- [5] Л. А. Емельянов, Т. Л. Емельянова. *Семейство Фейербаха* // Математическое просвещение. Сер. 3, вып. 6, 2002. С. 78–92.
- [6] С. В. Маркелов. *Парабола как окружность* // Десятая летняя конференция Турнира городов. М.: МЦНМО, 1999. С. 36–42, 112–123.
- [7] Ф. В. Петров. *Вписанные четырёхугольники и трапеции в абсолютной геометрии* // Математическое просвещение. Сер. 3, вып. 13, 2009. С. 149–154.
- [8] В. В. Прасолов. *Геометрия Лобачевского*. М.: МЦМНО, 2004.
- [9] В. В. Прасолов. *Задачи по планиметрии*. М.: МЦМНО, 2006.
- [10] В. Ю. Протасов. *Вокруг теоремы Фейербаха* // Квант, №9, 1992. С. 51–58.
- [11] В. В. Прасолов, В. М. Тихомиров. *Геометрия*. М.: МЦНМО, 1998.
- [12] З. А. Скопец, В. А. Жаров. *Задачи и теоремы по геометрии. Планиметрия*. М.: Учпедгиз, 1962.
- [13] О. В. Шварцман. *Комментарий к статье П. В. Бибиикова и И. В. Ткаченко* // Математическое просвещение. Сер. 3, вып. 11, 2007. С. 127–130.

---

---

# Нам пишут

---

---

## Некоторые аддитивные соотношения в треугольнике Паскаля

А. В. Стояновский

Есть много известных аддитивных соотношений в треугольнике Паскаля (т. е. соотношений, включающих лишь суммы и разности чисел в треугольнике Паскаля). Некоторые из этих любопытных соотношений можно найти в книге Г. Эндрюса «Теория разбиений», М.: Наука, 1982. (Правда, в этой книге содержатся соотношения между так называемыми  $q$ -биномиальными коэффициентами, которые суть многочлены от переменной  $q$ . Чтобы получить из них соотношения между обычными биномиальными коэффициентами, нужно положить  $q = 1$ .)

В настоящей статье приводятся, по-видимому, новые любопытные соотношения, получаемые простейшим способом, а именно суммированием чисел в треугольнике Паскаля по вертикали до некоторого места. Чтобы показать, что такое суммирование до некоторой степени естественно, напомним некоторые хорошо известные соотношения, получающиеся суммированием по горизонталям и диагоналям. Те из них, которые читателю неизвестны, можно рассматривать в качестве упражнения.

Хорошо известно, что сумма чисел в треугольнике Паскаля,

$$\begin{array}{ccccccc} & & & & & & & 1 \\ & & & & & & & & 1 \\ & & & & & & & & & 1 \\ & & & & & & & & & & 1 \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & & & & & 1 \\ & 1 \\ & 1 \\ & \dots \end{array}$$

лежащих на  $n$ -й горизонтали, равна  $2^n$ ,

$$C_n^0 + C_n^1 + C_n^2 + \dots + C_n^n = 2^n. \quad (1)$$

Известно также, что сумма чисел  $k$ -й диагонали до некоторого места равняется очередному числу из следующей диагонали,

$$C_k^k + C_{k+1}^k + C_{k+2}^k + \dots + C_n^k = C_{n+1}^{k+1}. \quad (2)$$

Еще один известный факт состоит в том, что сумма чисел по  $n$ -й наклонной диагонали с наклоном  $1/3$  (следующее число диагонали получается из предыдущего «ходом коня») равняется  $n$ -му числу Фибоначчи,

$$C_n^0 + C_{n-1}^1 + C_{n-2}^2 + C_{n-3}^3 + \dots = u_n, \quad (3)$$

где

$$u_{n+2} = u_{n+1} + u_n, \quad u_0 = u_1 = 1 \quad (4)$$

суть числа Фибоначчи.

Цель этой заметки — вывести формулу для суммы чисел вдоль вертикальной прямой до некоторого места,

$$C_n^k + C_{n-2}^{k-1} + C_{n-4}^{k-2} + \dots \quad (5)$$

Вот эта формула.

**ТЕОРЕМА.** Сумма (5) равняется знакопеременной сумме по следующей диагонали с наклоном  $1/3$ , начиная с ближайшего числа, плюс возможно  $\pm 1$  в зависимости от конкретной вертикали:

$$\begin{aligned} & C_n^k + C_{n-2}^{k-1} + C_{n-4}^{k-2} + \dots = \\ & = C_{n+1}^{k+1} - C_n^{k+2} + C_{n-1}^{k+3} - \dots + \\ & + \begin{cases} 0, & n - 2k \leq 0, \\ 0, & n - 2k > 0, n - 2k = 6p, 6p + 3, \\ -1, & n - 2k > 0, n - 2k = 6p + 1, 6p + 2, \\ 1, & n - 2k > 0, n - 2k = 6p - 1, 6p - 2. \end{cases} \end{aligned} \quad (6)$$

Эту теорему нетрудно доказать, если заметить, что обе части равенства (6) удовлетворяют тому же рекуррентному соотношению, что и числа в треугольнике Паскаля: обозначим левую часть равенства (6) через  $LHS(n, k)$  (сокращение от *left hand side*), а правую часть через  $RHS(n, k)$  (сокращение от *right hand side*), тогда

$$\begin{aligned} LHS(n, k) &= LHS(n-1, k) + LHS(n-1, k-1), \\ RHS(n, k) &= RHS(n-1, k) + RHS(n-1, k-1). \end{aligned} \quad (7)$$

Более точно, соотношения (7) выполняются при всех  $n$  и  $k$ , не лежащих на вертикали  $n - 2k = 0$ . Для  $n$  и  $k$ , лежащих на этой вертикали, нужно

к правым частям обоих соотношений (7) прибавить единицу. Кроме того, числа  $LHS(n, k)$  и  $RHS(n, k)$  удовлетворяют одинаковым начальным условиям (при  $k = 0$  и  $k = n$ , т.е. на границе треугольника Паскаля). Отсюда следует утверждение теоремы.

## О жордановых многоугольниках

А. А. Азамов

Плоскую замкнутую ломаную без самопересечений назовем жордановым многоугольником. Слово «жорданов» связано с тем, что определенные выше многоугольники являются частными случаями жордановых кривых, разбивающих плоскость на две части — ограниченную (называемую внутренностью многоугольника) и внешнюю [1].

Каждый угол при вершине  $v$  жорданова многоугольника либо лежит в интервале  $(0, \pi)$ , либо в интервале  $(\pi, 2\pi)$ . В первом случае углу будем приписывать число нуль и писать  $\theta(v) = 0$ , во втором — единицу и писать  $\theta(v) = 1$ . Таким образом каждому жорданову  $n$ -угольнику сопоставляется  $n$ -мерный вектор из нулей и единиц. Из теоремы о сумме углов многоугольника вытекает, что такие векторы содержат не меньше трех нулей.

*УТВЕРЖДЕНИЕ. Любой  $n$ -мерный вектор, координаты которого нули и единицы, содержащий не менее трех нулей, является вектором, соответствующим некоторому жорданову многоугольнику.*

Пусть дан такой вектор, содержащий  $k$  нулей,  $k \geq 3$ . Соответствующий  $n$ -угольник можно построить, например, следующим образом: возьмем выпуклый  $k$ -угольник и каждую сторону заменим ломаной с необходимым числом звеньев выпуклостью вовнутрь. При этом получится такой многоугольник, что все лучи, исходящие из какого-то центра, пересекаются с многоугольником в точности в одной точке. Такие многоугольники называются звездными (относительно выбранного центра).

Назовем два жорданова  $n$ -угольника *изотопными*, если существует непрерывное отображение  $t \mapsto (\gamma_1(t), \dots, \gamma_n(t))$ ,  $0 \leq t \leq 1$ , переводящее вершины одного многоугольника в вершины другого так, что при любом  $t$  точки образуют жорданов  $n$ -угольник.

Нетрудно заметить, что все выпуклые  $n$ -угольники изотопны с правильным  $n$ -угольником, следовательно, и друг с другом.

Пусть дан жорданов  $n$ -угольник с последовательными вершинами  $(v_1, v_2, \dots, v_n)$ . Выбрав какую-то вершину  $v_i$ , сопоставим ей вектор  $\alpha_i = (\theta(v_i), \theta(v_{i+1}), \dots, \theta(v_n), \theta(v_1), \dots, \theta(v_{i-1}))$ , так что векторы  $\alpha_i$  получают один из другого циклической перестановкой компонент. Совокупность  $\sigma = (\alpha_1, \dots, \alpha_n)$  назовем *сигнатурой многоугольника*. Например, сигнатура выпуклого многоугольника состоит из одного нулевого вектора,

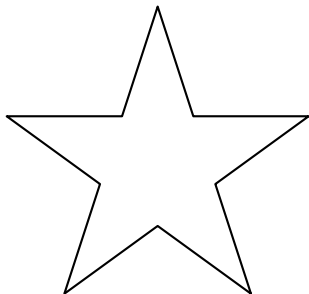


Рис. 1.

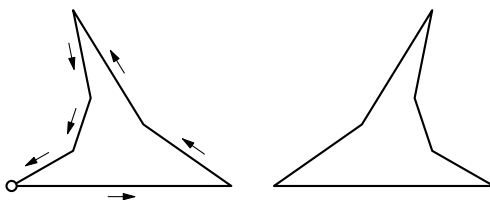


Рис. 2.

а сигнатура жордановой звезды (см. рис. 1) содержит всего два вектора:  $(0, 1, 0, 1, 0, 1, 0, 1, 0, 1)$  и  $(1, 0, 1, 0, 1, 0, 1, 0, 1, 0)$ .

**ТЕОРЕМА.** *Сигнатура является инвариантом изотопических многоугольников: если у двух многоугольников сигнатура одинакова, многоугольники изотопичны, если разная, то не изотопичны.*

Сперва докажем, что при изотопии сигнатура многоугольника не меняется. Пусть  $(\gamma_1, \gamma_2, \dots, \gamma_n)$  — изотопия  $n$ -угольника  $M_0$  в  $n$ -угольник  $M_1$ , а  $M_t$  —  $n$ -угольник с вершинами  $\gamma_i(t)$ , которые соединяются в таком же порядке, как вершины  $\gamma_i(0)$  в многоугольнике  $M_0$  ( $i = 1, 2, \dots, n$ ). Из сигнатуры  $\sigma_t$  многоугольника  $M_t$  выберем тот вектор  $\theta(t)$ , компоненты которого определяются, начиная с вершины  $\gamma_1(t)$  в направлении часовой стрелки. Ясно, что если значение  $t$  достаточно близко к 0, то  $\theta(t)$  совпадает с  $\theta(0)$  (при изотопии угол, меньше развернутого, в течение какого-то времени останется все еще меньше развернутого; то же самое верно и для угла больше развернутого).

Пусть  $s$  — наименьшее значение  $t$ , когда  $\theta(t)$  окажется отличным от  $\theta(0)$ . Без потери общности можно считать, что в момент  $t = s$  изменится первая компонента вектора (при этом не исключается, что в это же самое время у вектора еще какие-то другие компоненты также изменятся). Отсюда неумолимо следует, что при  $t = s$  угол  $n$ -угольника  $M_t$  с вершиной  $\gamma_1(t)$  становится развернутым (в частности, несколько соседних сторон, укорачиваясь, могут вырождаться в вершину). В любом случае около вершины  $\gamma_1(t)$  число сторон уменьшится. Поскольку в процессе изотопии  $M_t$  должен оставаться  $n$ -угольником, то в момент  $t = s$  где-то должны рождаться новые углы для компенсации потери сторон. Но угол многоугольника, существовавший в момент  $t = s$ , обязательно наличествовал и при значениях  $t$ , меньших, но достаточно близких к  $s$ . Это приводит к противоречию — получается, что для этих значений  $t$  число углов  $M_t$  было больше  $n$ .

Таким образом, сигнатура действительно не меняется при изотопии. Отсюда следует, например, что многоугольники, изображенные на рис. 2, не изотопны.

Далее докажем, что если два  $n$ -угольника имеют одну и ту же сигнатуру, то они изотопны.

Для этого достаточно убедиться, что любой  $n$ -угольник изотопен со звездным  $n$ -угольником. Это утверждение тривиально для  $n = 3, 4$ . Предположим, что оно верно для  $(n - 1)$ -угольников. Пусть  $M$  — произвольный  $n$ -угольник с некоторой сигнатурой  $\sigma$ . Будем считать, что *никакие три вершины  $M$  не лежат на одной прямой* — этого всегда можно добиться изотопией, мало меняющей положение вершин. (Такие изотопии называются «малым шевелением»; ясно, что они не меняют сигнатуру.)

**ЛЕММА.** *Существуют три последовательные вершины такие, что внутренность образуемого ими треугольника свободна от вершин  $M$ .*

В самом деле, всегда существует диагональ, делящая  $n$ -угольник на два многоугольника, естественно, с числом сторон, меньше  $n$  (см. [3, решение задачи 108]. Поэтому утверждение леммы легко выводится индукцией по числу сторон.

Возьмем треугольник  $v_1v_2v_3$ , внутренность которого свободна от других вершин. Внутренность отрезка  $v_1v_3$  также свободна от вершин в силу сделанного выше предположения. Поэтому если удалить из  $n$ -угольника  $M$  стороны  $v_1v_2$ ,  $v_2v_3$  и присоединить вместо них новую сторону  $v_1v_3$ , то получится жорданов  $(n - 1)$ -угольник  $M'$ . При этом вновь соединенная сторона не может оказаться продолжением соседних, так что в вершинах  $v_1$  и  $v_3$  образуется два новых угла, которые обозначим  $v'_1$  и  $v'_3$ . Все остальные углы  $M$  останутся неизменными.

Пусть  $M$  имеет сигнатуру  $\sigma$ , а  $M'$  — сигнатуру  $\sigma'$ . Возьмем в качестве представителя  $\sigma$  вектор  $\alpha = (\theta(v_1), \theta(v_2), \theta(v_3), \theta(v_4), \dots, \theta(v_n))$ , а в качестве представителя  $\sigma'$  — вектор  $\alpha' = (\theta(v'_1), \theta(v'_3), \theta(v_4), \dots, \theta(v_n))$ .

Легко заметить, что если координаты  $\theta(v_1)$ ,  $\theta(v_2)$ ,  $\theta(v_3)$  были равны 0, то и  $\theta(v'_1)$ ,  $\theta(v'_3)$  окажутся равными 0. Однако в других случаях возможны изменения. Например, из комбинации  $(1, 0, 0)$  может образоваться и  $(1, 0)$ , и  $(0, 0)$  в зависимости от того, по какую сторону от продолжения отрезка  $\mu v_1$  лежит сторона  $v_1v_n$  (рис. 3).

По предположению индукции  $M'$  изотопичен со звездным  $(n - 1)$ -угольником  $M''$  с вершинами  $v''_1, v''_3, v''_4, \dots, v''_n$ . Заменяя сторону  $v''_1v''_3$  этого многоугольника ломаной  $v''_1v''_2v''_3$ , получим  $n$ -угольник  $M'''$ . Основываясь на том, что при малом шевелении свойство звездности многоугольника сохраняется, такую замену можно осуществить так, что, во-первых,  $M'''$  также будет звездным, во-вторых, сигнатура  $M'''$  совпадет со сигнатурой  $\sigma$ . На



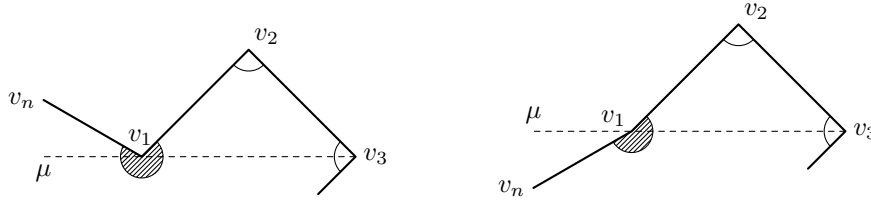


Рис. 3.

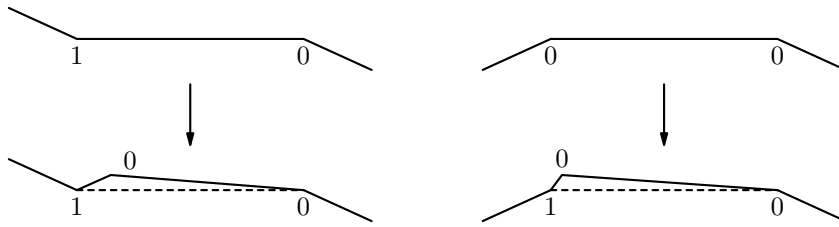


Рис. 4.

рис. 4 это показано для рассмотренных выше случаев (центр звездности находится на нижней стороне).

Ясно, что изотопия  $M' \rightarrow M''$  порождает изотопию  $M \rightarrow M'''$ . Непосредственно можно убедиться в том, что эти рассуждения справедливы и в других случаях редукции  $\theta(v_1), \theta(v_2), \theta(v_3) \rightarrow \theta(v'_1), \theta(v'_3)$ . Теорема доказана.

Число классов изотопии плоских  $n$ -угольников совпадает с числом сигнатур длины  $n$ . Обозначим это число  $s_n$ . Для него существует точная формула:

$$s_n = \frac{1}{n} \sum_{d|n} 2^d \varphi\left(\frac{n}{d}\right) - \left[\frac{n}{2}\right] - 2, \quad (1)$$

где  $\varphi(m)$  — количество чисел среди  $1, 2, \dots, m$ , взаимно простых с  $m$  (функция Эйлера), а суммирование в (1) производится по всем делителям  $d$  числа  $n$ . Например,  $s_4 = 2, s_5 = 4, s_6 = 9, s_7 = 15, s_8 = 29, s_9 = 54, s_{10} = 101$ . Равенство (1) получается из известной формулы для числа ожерелий ([2, формула (3.56)]).

Таким образом, число различных классов изотопных  $n$ -угольников быстро растет вместе с  $n$ , а точное значение сложным образом зависит от того, сколько и каких делителей имеет  $n$ .

**ЗАКЛЮЧЕНИЕ.** Посмотрим на вышеизложенное с «высоты птичьего полета». В соответствии с «Эрлангенской программой» Ф. Клейна, существуют разные «геометрии». Каждая такая «геометрия» состоит из

совокупности каких-то объектов и определенной группы преобразований. При этом если один объект можно преобразовать в другой, то они считаются эквивалентными. Согласно Клейну, цель соответствующей «геометрии» состоит в изучении свойств, общих для всех эквивалентных объектов, в том числе нахождений инвариантов группы преобразований. С этой точки зрения типичным является вопрос типа «Как узнать, что два объекта эквивалентные или, наоборот, не эквивалентные?» В одних случаях ответ получают сравнительно просто (например, если объекты — жордановы многоугольники, а эквивалентность означает конгруэнтность), в других случаях очень сложно (например, когда объекты — узлы в трехмерном пространстве, а эквивалентность означает возможность преобразования одного узла в другой), чаще всего получить полный ответ не представляется возможным.

То, что вопрос об изотопичности плоских  $n$ -угольников удалось разобрать полностью, означает, что многоугольники являются всё-таки очень простым объектом геометрии. Об этом свидетельствует и то, что единственный инвариант — сигнатура образует полную систему инвариантов для изотопии. Нам кажется, что именно из-за простоты «геометрия», состоящая из совокупности жордановых многоугольников с группой изотопий может служить удачным примером для демонстрации идей, лежащих в основе «Эрлангенской программы». Это и было основной целью заметки. Автор благодарит В. М. Тихомирова, А. В. Арутюнова и М. Вялого за внимание и содействие.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Болтянский В.Г., Ефремович В.А. *Наглядная топология*. М.: Наука, 1982.
- [2] Сачков В.Н. *Введение в комбинаторные методы дискретной математики*. М.: Наука, 1982.
- [3] Шклярский Д.О., Ченцов Н.Н., Яглом И.М. *Избранные задачи и теоремы планиметрии*. М.: Наука, 1967.

---

---

## Задачный раздел

---

---

В этом разделе вниманию читателей предлагается подборка задач разной степени сложности, в основном трудных. Составителям этой подборки кажется, что предлагаемые ниже задачи окажутся интересными как для сильных школьников, интересующихся математикой, так и для студентов-математиков.

Мы обращаемся с просьбой ко всем читателям, имеющим свои собственные подборки таких задач, присылать их в редакцию. И, разумеется, мы с удовольствием будем публиковать свежие авторские задачи.

В скобках после условия задачи приводится фамилия автора (уточнения со стороны читателей приветствуются). Если автор задачи нам неизвестен, то в скобках указывается «фольклор».

1. Треугольник  $ABC$  может быть покрыт тремя единичными кругами с центрами в его вершинах. Соответствующие стороны треугольника  $A'B'C'$  меньше соответствующих сторон треугольника  $ABC$ . Верно ли, что треугольник  $A'B'C'$  обладает тем же свойством?

(Фольклор)

2. К чему стремится объем  $n$ -мерного шара радиуса 2008 при  $n \rightarrow \infty$ ?

(А. Я. Белов)

3. Известно, что для любой последовательности  $\{a_i\}_{i=1}^{\infty}$  такой, что  $\sum_{i=1}^{\infty} a_i^2 < \infty$  ряд  $\sum_{i=1}^{\infty} a_i b_i < \infty$ . Докажите, что тогда  $\sum_{i=1}^{\infty} b_i^2 < \infty$ .

(Фольклор)

4. Для каких  $\lambda \in [0, 1]$  для любой непрерывной функции  $f: [0, 1] \rightarrow \mathbb{R}$  такой, что  $f(0) = f(1)$  обязательно найдется такое  $x \in [0, 1 - \lambda]$ , что  $f(x) = f(x + \lambda)$ ?

(Фольклор)

5. Известно, что в любом треугольнике расстояние между  $O$  и  $I$  описанной и вписанной окружностей выражается через их радиусы  $R$  и  $r$  с помощью формулы Эйлера:  $OI^2 = R^2 - 2Rr$ . Докажите обобщение этой формулы: если в треугольник вписан эллипс с фокусами  $F_1, F_2$  и малой осью  $l$ , то

$$R^2 l^2 = (R^2 - OF_1^2)(R^2 - OF_2^2).$$

(А. А. Заславский)

6. Докажите, что в алгебре матриц порядка  $n$  выполняются следующие тождества:

а) Тождество Размыслова:

$$\begin{aligned} n \cdot \text{Tr}(A) \cdot \sum_{\sigma \in S_{n^2}} (-1)^\sigma x_{\sigma(1)} y_1 x_{\sigma(2)} y_2 x_{\sigma(3)} \cdots x_{\sigma(n^2-1)} y_{n^2-1} x_{\sigma(n^2)} = \\ = \sum_{\sigma \in S_{n^2}} (-1)^\sigma A x_{\sigma(1)} y_1 x_{\sigma(2)} y_2 \cdots y_{n^2-1} x_{\sigma(n^2)} + \\ + \sum_{\sigma \in S_{n^2}} (-1)^\sigma x_{\sigma(1)} y_1 A x_{\sigma(2)} y_2 \cdots y_{n^2-1} x_{\sigma(n^2)} + \\ + \cdots + \sum_{\sigma \in S_{n^2}} (-1)^\sigma x_{\sigma(1)} y_1 x_{\sigma(2)} y_2 \cdots y_{n^2-1} A x_{\sigma(n^2)} \end{aligned}$$

б) Тождество Амицура – Левицкого:

$$\sum_{\sigma \in S_{2n}} (-1)^\sigma x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(2n)} = 0.$$

Здесь  $x_i, y_i, A$  – произвольные матрицы,  $S_k$  – группа перестановок из  $k$  элементов,  $(-1)^\sigma = +1$  для четных перестановок и  $(-1)^\sigma = -1$  для нечетных.

7. 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 еще через одну точку этого множества? (Фольклор)
8. Непрерывная функция  $f$  такова, что

$$\int_0^1 x^k f(x) dx = 1 \text{ для любого } k \in \{1, \dots, n-1\}.$$

Докажите, что

$$\int_0^1 (f(x))^2 dx \leq n^2.$$

(SEEMOUS 2008, Mircea Dan Rus)

9. В компании  $n$  человек. У каждого своя новость. Они перезваниваются, причем в каждом разговоре собеседники сообщают друг другу все известные им новости.
- а) Докажите, что понадобится не менее, чем  $2n - 4$  звонка, чтобы все узнали все новости.
- б) За один день каждый человек участвует не более, чем в одном разговоре. Какое минимальное количество дней необходимо, чтобы все узнали все новости? (А. В. Анджанс)

10.  $k$ -параллелепипедом называется прямоугольный параллелепипед, среди ребер которого имеется не более  $k$  различных. Докажите, что если параллелепипед  $P$  можно разрезать на  $k$ -параллелепипеды, то длины его ребер порождают векторное пространство размерности не выше  $k$  над  $\mathbb{Q}$ .

(Л. Радзивиловский, И. Фещенко, Д. Радченко, М. Танцура)

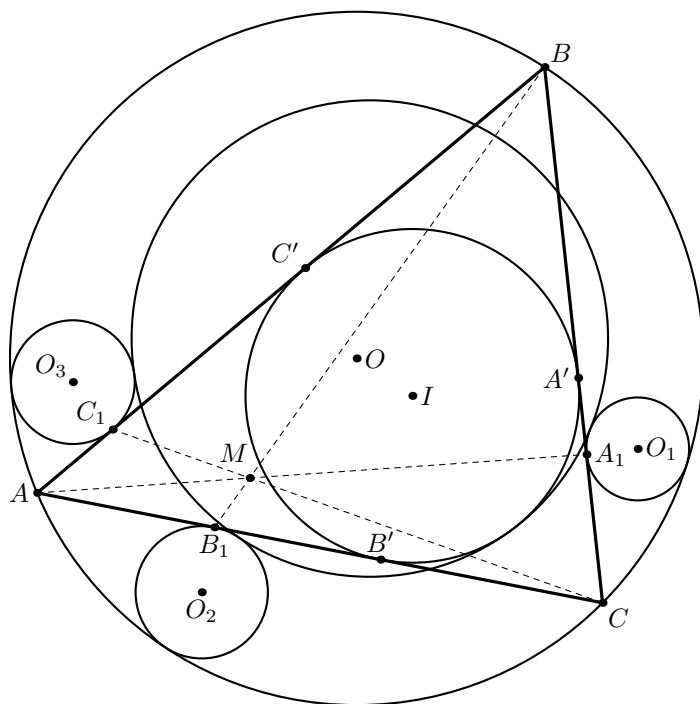
11. Докажите, что бесконечно много натуральных чисел не представимо в виде разности  $x^2 - y^3$ , ( $x, y$  – целые). (Фольклор)

12. Докажите, что следующие числа могут начинаться с любой комбинации цифр: а)  $2^{n^2}$  б)  $2^{2^n 3^k}$ .

в) Докажите, что множество  $A \subset \mathbb{R}$  чисел таких, что последовательность первых цифр  $c^{2^n}$  периодична, счетно, а множество  $B \subset \mathbb{R}$  чисел таких, что последовательность первых цифр  $c^{10^n}$  периодична, несчетно. (А. Канель)

## Решения задач из предыдущих выпусков

7.8. УСЛОВИЕ. Пусть  $ABC$  — произвольный треугольник, а  $M$  — точка внутри треугольника. Проведём через точку  $M$  три чевианы, основания которых —  $A_1, B_1, C_1$ . Построим вне треугольника три окружности, касающиеся сторон треугольника в основаниях чевиан и описанной окружности, и четвёртую, касающуюся этих трёх внешним образом. Тогда эта окружность касается вписанной окружности треугольника внутренним образом.



РЕШЕНИЕ.

Мы применим теорему Кэзи<sup>1)</sup>: четыре окружности  $s_i$  ( $i = 1, 2, 3, 4$ ) касаются окружности или прямой тогда и только тогда, когда

$$t_{12}t_{34} \pm t_{13}t_{42} \pm t_{14}t_{23} = 0,$$

<sup>1)</sup> Доказательство теоремы Кэзи см., например, в книге И. М. Яглома «Геометрические преобразования», М.: ГИТТЛ, 1956.

где  $t_{ij}$  — длина общей касательной к окружностям  $s_i$  и  $s_j$ .

В нашем случае окружности  $s_1, s_2, s_3$  — окружности с центрами  $O_1, O_2, O_3$  (см. рис.), а окружность  $s_4$  — это вписанная в треугольник  $ABC$  окружность. Докажем, что

$$t_{12}t_{34} - t_{13}t_{42} - t_{14}t_{23} = 0, \quad (1)$$

где  $t_{12}, t_{13}$  и  $t_{23}$  — длины общих внешних касательных, а  $t_{34}, t_{42}$  и  $t_{14}$  — длины общих внутренних касательных.

Будем рассматривать точки  $A, B, C$  как вырожденные окружности радиуса 0. Через  $t_{Ai}$  ( $t_{Bi}, t_{Ci}$ ) будем обозначать длины касательных из точки  $A$  ( $B, C$ ) к окружности  $s_i$ .

Запишем теорему Кэзи для окружностей  $A, B, s_1, C$  (они все касаются описанной вокруг треугольника  $ABC$  окружности):

$$\begin{aligned} t_{A1}t_{BC} &= t_{AB}t_{C1} + t_{AC}t_{1B}, \text{ т. е.} \\ t_{A1} \cdot BC &= AB \cdot CA_1 + AC \cdot A_1B. \end{aligned}$$

Отсюда получаем

$$t_{A1} = AB \cdot \frac{CA_1}{BC} + AC \cdot \frac{A_1B}{BC}. \quad (2)$$

Аналогично

$$t_{B2} = BA \cdot \frac{CB_1}{AC} + BC \cdot \frac{B_1A}{AC}, \quad (3)$$

$$t_{C3} = CB \cdot \frac{AC_1}{AB} + CA \cdot \frac{C_1B}{AB}. \quad (4)$$

Запишем теорему Кэзи для окружностей  $B, C, s_2$  и  $s_3$ :

$$\begin{aligned} t_{C3}t_{B2} &= t_{CB}t_{23} + t_{C2}t_{3B}, \text{ т. е.} \\ t_{C3}t_{B2} &= CB \cdot t_{23} + CB_1 \cdot C_1B, \end{aligned}$$

откуда, используя (3) и (4) получаем

$$t_{23} = \frac{AB \cdot B_1C \cdot C_1A + BC \cdot C_1A \cdot AB_1 + CA \cdot AB_1 \cdot BC_1}{AB \cdot CA}. \quad (5)$$

Аналогично

$$t_{12} = \frac{AB \cdot B_1C \cdot CA_1 + BC \cdot CA_1 \cdot AB_1 + CA \cdot AB_1 \cdot BC_1}{BC \cdot CA}, \quad (6)$$

$$t_{13} = \frac{AB \cdot BC_1 \cdot CA_1 + BC \cdot C_1A \cdot A_1B + CA \cdot A_1B \cdot BC_1}{BC \cdot AB}. \quad (7)$$

Теперь найдем  $t_{34}, t_{42}, t_{14}$ :

$$t_{34} = C_1C' = BC_1 - BC' = BC_1 - \frac{AB + BC - CA}{2}, \quad (8)$$

$$t_{42} = B'B_1 = AB' - AB_1 = \frac{CA + AB - BC}{2} - AB_1, \quad (9)$$

$$t_{14} = A_1A' = CA' - CA_1 = \frac{BC + CA - AB}{2} - CA_1, \quad (10)$$

здесь  $A', B', C'$  — точки касания сторон треугольника со вписанной окружностью.

Подставив (5)–(10) в (1) после упрощения получим

$$t_{12}t_{34} - t_{13}t_{42} - t_{14}t_{23} = \frac{AB_1 \cdot BC_1 \cdot CA_1 - A_1 \cdot B_1C \cdot C_1A}{2 \cdot AB \cdot BC \cdot CA} (2 \cdot AB \cdot BC + 2 \cdot BC \cdot CA + 2 \cdot CA \cdot AB - AB^2 - BC^2 - CA^2).$$

Это выражение равно 0, если

$$AB_1 \cdot BC_1 \cdot CA_1 - A_1 \cdot B_1C \cdot C_1A = 0, \quad (11)$$

так как

$$2 \cdot AB \cdot BC + 2 \cdot BC \cdot CA + 2 \cdot CA \cdot AB - AB^2 - BC^2 - CA^2 = 4(AB' \cdot CA' + BC' \cdot AB' + CA' \cdot BC') \neq 0.$$

Условие (11) равносильно условию конкурентности прямых  $AA_1, BB_1, CC_1$ , т. е.

$$\frac{AB_1}{B_1C} \cdot \frac{CA_1}{A_1B} \cdot \frac{BC_1}{C_1A} = 1.$$

(Л. А. Емельянов)

8.1. УСЛОВИЕ. Можно ли получить все возможные состояния кубика Рубика, последовательно выполняя некоторую комбинацию поворотов? (Учитываются только конечные, а не промежуточные состояния.)

РЕШЕНИЕ.

Предположим, что это возможно. Тогда существует такая комбинация поворотов  $K$ , что каждое состояние представимо степенью  $K$ . Но тогда и каждый поворот каждой грани кубика представим степенью  $K$ . Следовательно, повороты граней кубика должны быть перестановочны (коммутировать), но это неверно: повороты двух соседних граней неперестановочны. Полученное противоречие доказывает, что такой комбинации не существует. (А. К. Ковальджи)

9.2. УСЛОВИЕ. Дано  $n$  магнитофонных катушек, на которые намотаны ленты красными концами наружу, и 1 пустая катушка. Можно ли перемотать все ленты так, чтобы каждая оказалась на своей катушке, но красным концом внутрь? (Перематывать можно с любой катушки на пустую в данный момент катушку, при этом наружный конец становится внутренним, и наоборот.)

РЕШЕНИЕ. 1. Будем считать, что на пустой катушке тоже намотана виртуальная лента, а после перемотки две катушки обмениваются лентами, причем обе ленты меняют свои концы. Тогда ленты меняют концы парами.



Предположим, что  $n$  нечетно и нам удалось перемотать все начальные ленты наоборот, тогда в силу четности перемен концов и виртуальная лента поменяет конец. Это означает, что каждая лента сделала нечетное число перемоток, в том числе и виртуальная.

2. Если виртуальная лента вернулась на пустую катушку, то некоторые из остальных лент сделали циклическую перестановку. А поскольку виртуальная лента в итоге оказалась на своем месте, то среди остальных лент было сделано несколько циклических перестановок. Вместе с тем виртуальная лента делает на одну перемотку больше, чем длина цикла перестановки. Следовательно, число четных циклов нечетно.

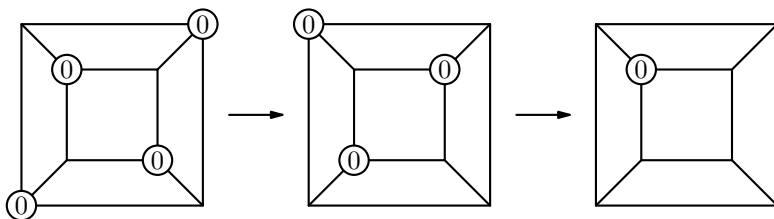
3. С другой стороны, чтобы вернуть все ленты на свои места необходимо четное число транспозиций, а нечетное число четных циклов требуют нечетного числа транспозиций. Полученное противоречие доказывает, что такая перемотка невозможна. (А. К. Ковальджи)

10.1. УСЛОВИЕ. В вершинах куба написаны числа. Вместо каждого числа записывают среднее арифметическое чисел, стоящих в трех соседних вершинах (числа заменяют одновременно). После 10 таких операций в вершинах оказались исходные числа. Найдите все такие числа.

РЕШЕНИЕ.

1. Если указанные 8 чисел уменьшить на одну и ту же величину, то указанное свойство их усреднения не изменится. Поэтому можно считать, что наименьшее число равно нулю, а остальные неотрицательны.

2. Поскольку нулевое число может получиться только усреднением нулей, то на предыдущем шаге (а процесс продолжается в обе стороны) три соседние числа равнялись нулю. А еще на предыдущем шаге уже 4 числа равнялись нулю (см. рис.). А еще на предыдущем шаге другие 4 числа должны равняться нулю. Возникает «мигалка», когда то одна, то другая четверка вершин обращается в нуль.



3. Рассмотрим четверку чисел, которая дополняет четверку нулей. Сделав один шаг, замечаем, что максимальное из чисел четверки не уменьшается только в том случае, если три наибольших числа равны. Это

необходимое условие заикливания. С таким же успехом и три минимальных числа в этой четверке равны. Следовательно, все числа равны.

4. Возвращаясь к исходным числам, получаем, что числа распались на две четверки одинаковых чисел, которые при каждом шаге меняются местами. Это и есть ответ. (А. К. Ковальджи)

11.1. УСЛОВИЕ. Вычислить

$$\int_0^{\infty} \frac{dx}{(1+x^2)(1+x^\pi)}.$$

РЕШЕНИЕ.

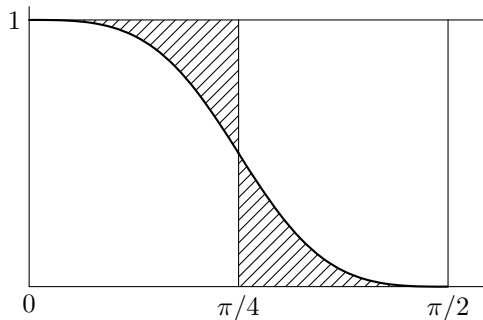
Сделаем замену  $t = \operatorname{arctg} x$ , тогда  $dt = \frac{dx}{1+x^2}$  и

$$\int_0^{\infty} \frac{dx}{(1+x^2)(1+x^\pi)} = \int_0^{\pi/2} \frac{dt}{1+\operatorname{tg}^\pi t}.$$

Вычислим площадь под графиком функции  $y(x) = \frac{1}{1+\operatorname{tg}^\pi x}$ ,  $x \in (0; \pi/2)$ . Для этого заметим, что график симметричен относительно точки  $(\pi/4, 1/2)$ :

$$y(\pi/2 - x) = \frac{1}{1+\operatorname{tg}^\pi(\pi/2 - x)} = \frac{1}{1+\frac{1}{\operatorname{tg}^\pi x}} = \frac{\operatorname{tg}^\pi x}{1+\operatorname{tg}^\pi x} = 1 - \frac{1}{1+\operatorname{tg}^\pi x} = 1 - y(x).$$

Значит, площадь под графиком функции  $y(x)$  при  $x \in (\pi/4; \pi/2)$  равна площади между прямой  $y = 1$  и графиком  $y(x)$  при  $x \in (0; \pi/4)$  (см. рис.). Тем самым, искомый интеграл равен площади прямоугольника со сторонами 1 и  $\pi/4$ .



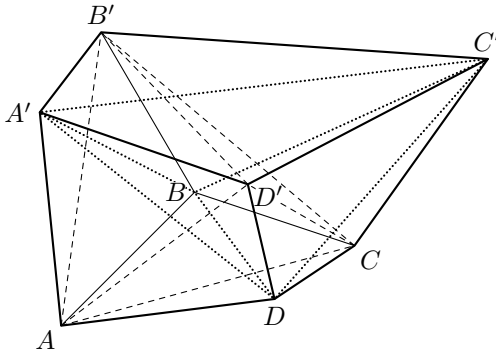
(М. Ю. Панов)

11.2. УСЛОВИЕ. Назовем *кубоидом* выпуклый многогранник в трехмерном пространстве, комбинаторно эквивалентный кубу (т. е. существует взаимно однозначное соответствие между их вершинами, ребрами и гранями, сохраняющее примыкание). Рассмотрим для каждой грани точку

пересечения отрезков, соединяющих середины противоположных сторон. Эти 6 точек являются вершинами некоторого октаэдра.

Какие значения может принимать отношение объема этого октаэдра к объему исходного кубоида?

РЕШЕНИЕ.



В любой кубоид можно канонически вписать два тетраэдра, все ребра которых являются диагоналями граней кубоида (см. рисунок). Обозначим через  $O$  объем октаэдра, заданного в условии задачи, через  $K$  — объем кубоида, через  $T_1$  и  $T_2$  — объемы канонически вписанных в кубоид тетраэдров. Ниже будет доказана формула:

$$O = \frac{1}{8} \left( K + \frac{T_1 + T_2}{2} \right). \quad (*)$$

Из нее следует очевидное двойное неравенство

$$\frac{1}{8} K < O < \frac{1}{4} K. \quad (**)$$

Станем плавно деформировать кубоид так, чтобы вершины одной из граней слились в одну точку (тогда кубоид превратится в пирамиду). При этом формула (\*) останется справедливой и для конечного положения, когда кубоид уже не будет являться кубоидом (правда, неравенства в (\*\*)) могут стать нестрогими). Очевидно, что в конечном положении оба тетраэдра вырождаются и имеет место формула  $O = \frac{1}{8} K$ .

В силу непрерывной зависимости объема от положения вершин при приближении тетраэдра к конечному положению  $O$  пробегает значения, как угодно близкие к  $\frac{1}{8} K$ .

Теперь произведем другую деформацию, при которой в конечном положении совпадут соответственно точки  $B, C$  с  $D, B'$  с  $C', D'$  с  $A'$ . Тогда кубоид превратится в тетраэдр, причем с этим тетраэдром совпадут и оба канонически вписанных в кубоид тетраэдра. Очевидно, в конечном

положении будет справедлива формула  $O = \frac{1}{4}K$ . Тем самым, получаем следующий

ОТВЕТ:  $O$  принимает все значения из интервала  $(\frac{1}{8}K; \frac{1}{4}K)$ .

ДОКАЗАТЕЛЬСТВО ФОРМУЛЫ (\*).

КОЕ-ЧТО О ВЫЧИСЛЕНИИ ОБЪЕМОВ. Известно, что если тройка векторов  $a_i, a_j, a_k$  имеет ориентацию по часовой стрелке относительно точки  $P$ , к которой приложены эти вектора, то смешанное произведение  $([a_i, a_j], a_k)$  равно ушестеренному объему тетраэдра, натянутого на эти вектора.

Этот факт допускает следующее обобщение. Пусть  $P$  — произвольная точка пространства,  $M$  — пространственное тело (не обязательно односвязное), поверхность которого разбита на треугольники с вершинами  $A_1, A_2, \dots, A_n$ . Пусть  $A_i, A_j, A_k$  — произвольный треугольник этого разбиения, причем его вершины перечислены в порядке, соответствующем обходу по часовой стрелке, если смотреть на треугольник со стороны внутренности тела. Сопоставим этому треугольнику значение смешанного произведения  $([\vec{PA}_i, \vec{PA}_j], \vec{PA}_k)$ , которое с точностью до знака равно ушестеренному объему тетраэдра  $PA_iA_jA_k$ . Сумма таких смешанных произведений по всем треугольникам разбиения поверхности тела  $M$  равна  $6 \cdot (\text{объем тела } M)$ . Из этой формулы можно получить такое

СЛЕДСТВИЕ. Пусть  $\omega_1, \omega_2, \omega_3$  — диагонали произвольного октаэдра, т. е. 3 вектора, каждый из которых соединяет пару противоположных вершин октаэдра. Тогда, с точностью до знака, их смешанное произведение равно  $6 \cdot (\text{объем октаэдра})$ .

Для доказательства достаточно представить каждую диагональ в виде разности векторов, проведенных к вершинам, и раскрыть скобки.

Далее обозначения  $\omega_1, \omega_2, \omega_3$  используются для октаэдра из условия задачи.

Вернемся к доказательству формулы (\*). Обозначим через  $d_1, d_2, d_3, d_4$  вектора главных диагоналей кубоида, направление которых выбрано так, чтобы они всегда начинались на одном из канонически вписанных тетраэдров, а заканчивались на другом:  $d_1 = AC'$ ,  $d_2 = B'D$ ,  $d_3 = CA'$ ,  $d_4 = D'B$ . Учитывая, что каждая вершина октаэдра (рассматриваемая, как вектор) представляет собой среднее арифметическое 4 вершин той грани, к которой она принадлежит, можно записать следующую формулу

$$\begin{aligned} \frac{1}{6}O = \omega_1\omega_2\omega_3 &= \frac{d_1 + d_2 - d_3 - d_4}{4} \cdot \frac{d_1 - d_2 + d_3 - d_4}{4} \cdot \frac{d_1 - d_2 - d_3 + d_4}{4} = \\ &= \frac{1}{16}(d_1d_2d_3 + d_1d_3d_4 + d_1d_4d_2 + d_4d_3d_2). \end{aligned}$$

Отметим, что величины  $d_1, d_2, d_3$  могут также рассматриваться как диагонали октаэдра  $ACDA'B'C'$ , представляющего собой кубоид без двух

«угловых» тетраэдров, «окружающих», соответственно, вершины  $B$  и  $D'$ , а именно,  $ABCB'$  и  $A'D'C'D$ . Аналогичные замечания можно сделать и для троек  $d_1, d_3, d_4$ ;  $d_1, d_4, d_2$ ;  $d_4, d_3, d_2$ . Отметим также, что каждый из двух канонических тетраэдров представляет собой разность кубоида и четырех из восьми «угловых» тетраэдров кубоида. С учетом этих наблюдений (допуская некоторую вольность обозначений) можно продолжить вышеприведенное равенство:

$$\begin{aligned} O &= \frac{1}{16}(4 \text{ октаэдра}) = \\ &= \frac{1}{16}(4(\text{кубоида} - 2 \text{ уг.тетраэдра})) = \frac{1}{16}(4 \text{ кубоида} - 8 \text{ уг.тетраэдра}) = \\ &= \frac{1}{16}(2 \text{ кубоида} + (\text{кубоид} - 4 \text{ уг.тетраэдра}) + (\text{кубоид} - 4 \text{ уг.тетраэдра})) = \\ &= \frac{1}{8}\left(K + \frac{T_1 + T_2}{2}\right). \end{aligned}$$

Формула (\*) доказана.

(Р. М. Травкин)

12.10. УСЛОВИЕ. Назовем множество  $C$  перестановок  $n$  элементов *хорошим*, если для любого ненулевого набора чисел  $v_1, \dots, v_n$  такого, что  $\sum_i v_i = 0$ , найдется такая перестановка  $\pi$  из множества  $C$ , что  $\sum_{i=1}^k v_{\pi(i)} > 0$  для всех  $k$  от 1 до  $n - 1$ . Если заменить строгое равенство на нестрогое, то получится определение *неплохого* множества. Какова мощность наименьшего хорошего (неплохого) множества? а) Мощность наименьшего неплохого множества равна  $n$ . б) Существует хорошее множество мощности  $n2^n$ . в) (открытая проблема) Доказать, что мощность хорошего множества экспоненциально велика по  $n$ .

РЕШЕНИЕ.

Покажем, что наименьший размер неплохого множества равен  $n$ , а наименьший размер хорошего множества равен  $n(n - 1)$ .

Сначала покажем, что каждое неплохое множество содержит хотя бы  $n$  перестановок, а каждое хорошее — хотя бы  $n(n - 1)$ .

В самом деле, рассматривая векторы, у которых  $v_k > 0$ , а остальные координаты отрицательны, получаем, что в неплохом множестве должна быть перестановка с первым элементом  $k$ . Отсюда получается, что перестановок в неплохом множестве должно быть хотя бы  $n$ .

Рассматривая для каждой пары  $1 \leq k, l \leq n$  различных номеров  $k$  и  $l$  вектор с  $v_k = 1$ ,  $v_l = -1$ ,  $v_j = 0$  при  $j \neq k, l$  выясняем, что в хорошем множестве должна найтись перестановка с первым элементом  $k$  и последним элементом  $l$ . Значит, перестановок в хорошем множестве должно быть хотя бы  $n(n - 1)$ .

Перейдем к построению примеров неплохих и хороших множеств.

Ключевой для нас является следующая (известная)

**ЛЕММА 1.** По кругу написано несколько вещественных чисел с неотрицательной суммой. Тогда найдется такое место, что все частичные суммы написанных чисел начиная с него по часовой стрелке неотрицательны.

**ДОКАЗАТЕЛЬСТВО.** Рассмотрим несколько подряд чисел так, чтобы их сумма была наибольшей возможной. Пусть эти числа (по часовой стрелке) суть  $x_1, x_2, \dots, x_k$ . Пронумеруем и оставшиеся числа — от  $x_{k+1}$  до  $x_n$ . Докажем, что в качестве требуемого числа подходит  $x_1$ . В самом деле, при  $j \leq k$  имеем  $x_1 + \dots + x_j = (x_1 + \dots + x_k) - (x_{j+1} + \dots + x_k) \geq 0$  в силу нашего выбора чисел от  $x_1$  до  $x_k$ . При  $j > k$  имеем

$$\begin{aligned} x_1 + \dots + x_j &= \\ &= (x_1 + \dots + x_n) + ((x_1 + \dots + x_k) - (x_{j+1} + \dots + x_n + x_1 + \dots + x_k)) \geq 0. \end{aligned}$$

Лемма доказана.

**СЛЕДСТВИЕ.** По кругу написано несколько вещественных чисел с положительной суммой. Тогда найдется такое место, что все частичные суммы начиная с него по часовой стрелке положительны.

**ДОКАЗАТЕЛЬСТВО.** Уменьшим немного каждое из чисел так, чтобы сумма осталась положительной и воспользуемся для нового набора чисел доказанной леммой.

Из леммы 1 сразу следует, что множество из  $n$  циклических перестановок чисел от 1 до  $n$  является неплохим.

Приведем пример хорошего множества из  $n(n-1)$  перестановок. Для каждого  $k = 1, 2, \dots, n$  рассмотрим  $n-1$  перестановку, в которых последним элементом является  $k$ , а остальные элементы переставляются по циклу. Например, при  $n = 4$  и  $k = 3$  это будут перестановки 1243, 2413, 4123. Докажем, что построенный набор  $n(n-1)$  перестановок является хорошим. Действительно, для ненулевого вектора  $(v_1, v_2, \dots, v_n)$  с нулевой суммой координат найдется такое  $k$ , что  $v_k < 0$ . Тогда сумма всех координат, кроме  $v_k$ , положительна, и среди перестановок нашего множества, в которых число  $k$  стоит на последнем месте, согласно следствию из леммы найдется требуемая.

**ЗАМЕЧАНИЕ.** Часто лемму формулируют так: На круговом шоссе расположено  $n$  бензоколонок, суммарного количества бензина в которых хватит, чтобы объехать все шоссе. Докажите, что машина с пустым баком может (заправляясь на бензоколонках) объехать все шоссе с некоторого места по часовой стрелке.

(Ф. В. Петров)



ОПЕЧАТКИ, ЗАМЕЧЕННЫЕ В №12

СТРАНИЦА,	СТРОКА	НАПЕЧАТАНО	СЛЕДУЕТ ЧИТАТЬ
89,	7 снизу	любая группа	любая простая группа
134,	2, 3 снизу	Написано с использованием текста И.Лукьянца и В.Соколова	Написано с использованием решения задачи 4d из <a href="http://www.turgor.ru/1ktg/2007/5/index.php">http://www.turgor.ru/1ktg/2007/5/index.php</a> , представленного Е.Лукьянцом, учеником ФМЛ 239 г. Санкт-Петербурга, и В.Соколовым, учеником гимназии 261 г. Санкт-Петербурга.

Подготовка оригинал-макета: L<sup>A</sup>T<sub>E</sub>X2<sub>ε</sub>,  
 МЕТАРОСТ, М. Н. Вялый

Издательство Московского Центра  
 непрерывного математического  
 образования  
 119002, Москва, Большой Власевский пер., 11. Тел. (499) 241 74 83

Отпечатано по StP-технологии в ОАО «Печатный двор» им. А. М. Горького.  
 197110, Санкт-Петербург, Чкаловский проспект, 15.

Подписано в печать 18.3.09. Формат 70×100/16. Бумага офсетная. Печать  
 офсетная. Печ. л. 12,0. Тираж 400 экз. Заказ №