В ЭТОМ ВЫПУСКЕ НАПЕЧАТАНО:

ПАМЯТИ А. Я. ХИНЧИНА:

Б. В. Гнеденко. Александр Яковлевич Хинчин.

А. Я. Хинчин. О воспитательном эффекте уроков математики.

А. Я. Хинчин. О так называемых «Задачах на соображение» в курсе арифметики.

ОБЗОРЫ И СТАТЬИ:

Р. Л. Добрушин. Математические методы в лингвистике.

И. М. Яглом. Комплексные числа и их применение в геометрии.

В. Г. Болтянский и В. А. Ефремович. Очерк основных идей топологии. *(Окончание)*

Ю. М. Безбородов и В. Б. Орлов. Машина играет в шахматы.

НАУЧНЫЕ И НАУЧНО-МЕТОДИЧЕСКИЕ СООБЩЕНИЯ.

НАУЧНАЯ И ПЕДАГОГИЧЕСКАЯ ХРОНИКА.

ЗАДАЧИ И ПРОБЛЕМЫ ПО МАТЕМАТИКЕ.

РЕШЕНИЯ ЗАДАЧ

И ДРУГИЕ МАТЕРИАЛЫ.

См. содержание на стр. 372.

МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ

МАТЕМАТИКА, ЕЕ ПРЕПОДАВАНИЕ, ПРИЛОЖЕНИЯ И ИСТОРИЯ

ВЫПУСК 6

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1961

«Математическое просвещение»
выпускается при редакционном участии
И. Н. БРОНШТЕЙНА, А. М. ЛОПШИЦА, А. А. ЛЯПУНОВА,
А. И. МАРКУШЕВИЧА, И. М. ЯГЛОМА

МАТЕМАТИЧЕСКОЕ ПРОСВЕЩЕНИЕ, вып. 6

Редакторы *И. Н. Бронштейн* и *Ф. Л. Варпаховский.*

Техн. редактор *В. Н. Крючкова.* Корректор *О. А. Сигал.*

Сдано в набор 5/IV 1961 г. Подписано к печати 3/VIII 1961 г. Бумага $60 \times 90^1/_{16}$. Физ. печ. л. 23,5 + 1 вкл. Услови. печ. л. 23,625. Уч.-изд. л. 23,63. Тираж 9 000 экз. Т-08722. Цена книги 71 коп. Заказ № 1667.

Государственное издательство физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Первая Образцовая типография имени А. А. Жданова
Московского городского совнархоза. Москва, Ж-54, Валовая, 28.

Александр Яковлевич Хинчин
(1894—1959)

ПАМЯТИ АЛЕКСАНДРА ЯКОВЛЕВИЧА ХИНЧИНА

АЛЕКСАНДР ЯКОВЛЕВИЧ ХИНЧИН

Б. В. Гнеденко

(Москва)

18 ноября 1959 года советская наука понесла тяжелую утрату — после длительной болезни скончался замечательный математик и педагог Александр Яковлевич Хинчин. В памяти тех, кто работал рядом с ним, учился у него или же пользовался его советами, навсегда останется обаятельный образ человека, преданного науке и делу, которому он посвятил свою жизнь. Скромный в обращении с людьми, он не добивался для себя преимуществ, а совершенно бескорыстно отдавал знания, идеи, мысли тем, в ком он замечал искреннее увлечение наукой и стремление настойчиво работать над нерешенными проблемами. Он умел деликатно поправить ученика, помочь ему выявить его заблуждения, направить его мысль на такой путь, который сулил успех. И это делалось так, чтобы не помешать самостоятельности другого лица, развитию его творческой индивидуальности. Одновременно он был исключительно строг к себе, к тому, что он делал, писал, говорил. Он не допускал в своих как устных, так и письменных выступлениях ни тени упрощенчества, искажения существа дела ради ложного популяризаторства и при этом добивался поразительной ясности и отточенности изложения.

Имя А. Я. Хинчина прочно вошло в историю советской математики. И это вызвано не только тем, что им получен ряд в высшей степени замечательных результатов во многих разделах науки, но и тем, что он явился создателем новых направлений исследований, воспитал большое число учеников, создал превосходные учебники и монографии. Его собственно математические работы относятся к теории функций действительного переменного, теории чисел, теории вероятностей и статистической механике. В каждой из этих столь различных по характеру областей математики он достиг поразительных по общности результатов и нашел глубокие объединяющие их связи. В теории вероятностей по всеобщему признанию он был одним из тех ученых, чьи исследования, идеи и начинания знаменовали новый этап в ее развитии.

Александр Яковлевич Хинчин был не только выдающимся математиком, но и замечательным педагогом. Он обладал исключительным даром не только понятно излагать материал, но и вскрывать при этом существо дела, глубину тех проблем, о которых шла речь. К каждому своему устному выступлению он тщательно готовился и всегда добивался того, что всё, что он рассказывал, излагалось исключительно ярким и впечатляющим образом. Было невозможно, прослушав даже его доклад о специальных исследованиях, уйти, не поняв во всех деталях идею работы, место результата в науке и его значение для дальнейшего прогресса соответствующей области математики. Его лекции высоко расценивались студентами. Они видели, что А. Я. Хинчин стремится донести до них не только и не столько формальный аппарат, не одни аналитические выкладки, а душу предмета, его идейное содержание. На этих лекциях они воспитывали в себе умение четко и лаконично рассуждать, за деталями вычислений постоянно видеть предмет рассуждений. Все эти качества сохранялись в его статьях и книгах. Однако они в далеко не полной мере передают очарование живых, образных, литературно и логически совершенных устных лекций А. Я. Хинчина.

Педагогические проблемы средней школы всерьез заинтересовали А. Я. Хинчина в 1938 году, когда он был приглашен в Научно-исследовательский институт школ для заведования кабинетом математики. Он сразу включился в живую и многообразную работу: посещал школы, давал советы сотрудникам, выступал с докладами перед педагогической общественностью, подверг тщательному анализу те недостатки, которые были свойственны советской средней школе того времени. Проведя такой анализ, он выступил с рядом статей, посвященных главным образом общим проблемам преподавания математики в средней школе.

Основные педагогические интересы А. Я. Хинчина касались не проблем частной методики, не разработки канонов изложения тех или иных частей школьного курса математики. Здесь он придерживался того мнения, что строгое регламентирование поведения педагога, навязывание ему системы и порядка изложения может только сковать инициативу учителя, лишить уроки вдохновения, уничтожить неповторимую ценность индивидуальности преподающего. Конечно, эту точку зрения он не доводил до крайности, и она совсем не означала, что не следует искать удачных методических решений для облегчения и улучшения системы изложения, для достижения лучших успехов в процессе обучения наиболее экономными средствами. Он считал, что необходимо проводить такие поиски систематически и поощрять самый широкий обмен опытом педагогического мастерства. Но он был яростным противником методического штампа, при котором лишь одна единственная система изложения признается правильной и законной, а все остальные считаются ошибочными и вредными.

Даже прекрасные педагогические образцы, повторенные другим лицом бездумно, без учета сложнейшей психологической особенности данного класса, как правило, обречены на провал. Педагогическая работа в значительной степени является искусством и потому не может терпеть шаблона. Для А. Я. Хинчина успех педагогического процесса состоит не в том, ка́к точно следует учитель методическим письмам, и не в числе хороших и отличных оценок, выставленных преподавателем в конце года, а в том, как глубоко поняты учащимися идеи математического метода, как далеко продвинулось воспитание навыков правильных логических рассуждений, насколько они привыкли улавливать диссонанс логических перескоков в рассуждениях, насколько самостоятельно они научились подходить к анализу и решению задач.

А. Я. Хинчин был убежден, что одним из самых тяжких недугов математического образования в средней школе является формализм знаний, при котором чисто внешнее усвоение математических понятий и результатов заслоняет основное их содержание. Поскольку успешное применение математических знаний может базироваться только на отчетливом понимании сути математических понятий и методов, а не на бездушном заучивании программного материала, то борьбу с формализмом в преподавании математики А. Я. Хинчин считал одной из основных задач советской школы.

В превосходной брошюре «Основные понятия математики и математические определения в средней школе» (1940) А. Я. Хинчин отметил два принципа, которых до́лжно придерживаться в школьном преподавании математики:

1) Учет возрастных особенностей учащихся может приводить к необходимости упрощенной трактовки понятий по сравнению с той, какая принята в современной науке. Однако «ни в одном случае школа не должна в целях упрощения искажать научную трактовку понятия, придавать ему черты, противоречащие научному его пониманию...».

2) «Замена отчетливых и точных определений, формулировок и рассуждений расплывчатыми, не имеющими точного смысла, и при последовательном использовании неизбежно приводящими к логическим неувязкам, ни в коем случае не может способствовать облегчению понимания, а, напротив, во всех случаях затрудняет его; мыслить расплывчато не может быть делом более легким, чем мыслить четко».

Сам А. Я. Хинчин придерживался очень строго этих простых принципов в своей практической работе. Более того, нечеткость и расплывчатость в рассуждениях ему была органически чужда. Эта черта его характера нашла отражение при переработке известного учебника арифметики А. П. Киселева. В результате вмешательства А. Я. Хинчина из учебника был изгнан ряд понятий, сохранявшихся в школьных учебниках лишь в силу слепой традиции. От этого книга выиграла сразу по меньшей мере в двух отношениях: она стала научнее и одновременно более доступной для понимания. Недаром

сам А. П. Киселев после завершения А. Я. Хинчиным указанной переработки прислал ему письмо с выражением самой искренней благодарности за существенное улучшение книги. В тесной связи с работой над учебником арифметики А. П. Киселева находится небольшая заметка А. Я. Хинчина «О понятии отношения двух чисел». Сейчас время убедительно показало, ка́к был прав А. Я. Хинчин, когда занес руку на понятие, абсолютно не нужное математической науке. Но в ту пору отстоять эту точку зрения стоило больших усилий.

Статья А. Я. Хинчина «О воспитательном эффекте уроков математики», предлагаемая (с некоторыми сокращениями) читателям настоящего сборника, была написана им еще в 1947 году в летнее каникулярное время. При разборе бумаг вдова покойного Н. А. Хинчина нашла эту рукопись, тщательно переписанную каллиграфическим почерком автора. В чем состоит причина того, что автор не передавал в течение двенадцати лет рукопись в печать, неизвестно. По-видимому, он испытывал какую-то неудовлетворенность написанным.

В рукописи поднимаются важные вопросы математического просвещения [1]). Они приобретают особый интерес в связи с той серьезной перестройкой, которую испытывает сейчас наша школа.

Заслуживает быть отмеченным то обстоятельство, что автор дважды упоминает в рукописи о задуманной им работе, посвященной вопросам преподавания в средней школе элементов математического анализа. Сейчас разработка этих вопросов исключительно актуальна, и можно только пожалеть, что мы ничего не знаем о тех соображениях, которые были на этот счет у А. Я. Хинчина. Ведь он заслуженно считался одним из лучших современных лекторов и в течение многих лет читал этот курс и в университете, и в педагогических институтах.

Вторая печатаемая здесь статья А. Я. Хинчина «О так называемых „Задачах на соображение“ в курсе арифметики» нигде не была напечатана. Рукопись удалось найти при разборе одной из папок с бумагами покойного. Поскольку эти вопросы интересовали А. Я. Хинчина в 1938—1939 гг., можно считать достаточно определенным, что именно тогда эта статья и была написана.

[1]) Не всегда их трактовка является бесспорной. Мне, например, кажется, что в параграфе, названном «Честность и правдивость», несколько недооцениваются возможности субъективных оценок и внутри самой математики. Для примера укажу, что нередко важность одной и той же области математики различные ее представители оценивают различно. Точно так же в заключительной части параграфа «Настойчивость и мужество», на мой взгляд, несколько преуменьшено творческое начало в работе ученика в нематематических областях знания.

О ВОСПИТАТЕЛЬНОМ ЭФФЕКТЕ УРОКОВ МАТЕМАТИКИ

А. Я. Хинчин

(Москва)

Математика, в отличие от большинства других преподаваемых в школе дисциплин, имеет предметом своего изучения не непосредственно вещи, составляющие окружающий нас внешний мир, а *количественные отношения и пространственные формы*, свойственные этим вещам. Этой особенностью математической науки в первую очередь объясняются те хорошо известные методические трудности, которые неизбежно встают перед преподавателем математики и которых почти не знают преподаватели других наук: перед учителем математики стоит нелегкая задача — преодолеть в сознании учеников возникающее со стихийной неизбежностью представление о «сухости», формальном характере, оторванности этой науки от жизни и практики. Об этом написано много ценного и полезного, и мы хорошо знаем, ка́к справляются с этой задачей лучшие мастера нашей школы.

Но этой же особенностью математической науки в значительной мере объясняется и специфика задач, встающих перед учителем математики, который хочет использовать преподавание своей науки *в воспитательных целях*. Ясно, что и здесь стоящая перед ним задача труднее, чем в случае большинства других наук. Ибо научная дисциплина, занятая изучением не самих вещей, а лишь отношений между ними, и потому необходимо требующая поднятия на некоторую ступень абстракции,— такая дисциплина, очевидно, лишь в редких случаях способна давать учителю повод к эффективному воздействию на формирование характера и мировоззрения учащихся, на регулирование их поведения. Этим, несомненно, объясняется то, что в исследованиях, посвященных вопросам воспитательной функции школьного обучения, об уроках математики обычно вовсе не говорится или говорится очень мало.

То немногое, что написано по этому поводу, в основном не вызывает возражений. Дело сводится обычно к двум рычагам воспитательного воздействия: с одной стороны, говорится, что специфическая для математики логическая строгость и стройность умозаключений призвана воспитывать в учащихся общую логическую культуру мышления; с другой — указывается, что предметно-содержательное

оснащение математических задач при надлежащем его выборе дает широкий простор для сообщения цифр и данных, способных значительно расширить кругозор учащихся, поднять их общий культурный уровень.

Всё это бесспорно верно. Я думаю, однако, что это далеко не всё. Прежде всего, здесь совершенно не затрагиваются важнейшие задачи морального воспитания, для которых, как мне кажется, уроки математики дают весьма ощутительные возможности. Далее, важная задача воспитания логической культуры мышления, которой обычно уделяется много внимания, тем не менее трактуется в большинстве случаев трафаретно, поверхностно и недостаточно расчлененно; приводимые примеры часто не выходят за рамки вульгарного шаблона и потому очень мало эффективны. Наконец, воспитывающее воздействие данных, приводимых в «текстовых» задачах, хотя и должно, конечно, всемерно быть использовано, но с математическим содержанием урока связано, очевидно, лишь весьма внешним образом; ясно, что здесь воспитывающее влияние призвана оказывать не сама математика, не ее законы и ее стиль, а те привязанные к ней чисто внешним образом данные, которые обрамляют собою «текстовые» задачи и которые без всякого изменения математического содержания задачи могли бы быть заменены любыми другими аналогичными данными. Ясно поэтому, что этот рычаг воспитывающего воздействия, будучи важным и действенным, не может считаться в прямом смысле принадлежащим самой преподаваемой в школе математической науке.

Все приведенные соображения показывают, что вопрос о воспитательном значении уроков математики у нас разработан еще далеко недостаточно. Предлагаемая статья ставит целью несколько продвинуть этот вопрос. Для этого я в дальнейшем кратко рассмотрю ряд моментов, которые, насколько я могу судить, при изыскании возможностей воспитательного влияния уроков математики до сих пор либо совсем оставлялись без внимания, либо рассматривались лишь весьма поверхностно.

I. КУЛЬТУРА МЫСЛИ

1. Правильность мышления. Роль и значение математики в воспитании навыков закономерного и безошибочного мышления в такой мере всеми признана, что нередко приходится встречаться с утверждениями, будто приучение к строгому в логическом отношении ходу мыслей есть первая и основная задача учителя математики, так что в сравнении с нею даже ознакомление учащихся с самим содержанием математической науки отодвигается на второй план (что несомненно следует признать уже вредным перегибом). Однако как раз потому, что эта воспитательная функция уроков математики приобрела характер банальности, именно в этом направлении мы слышим много вы-

сказываний, приводимых по готовому трафарету, без достаточного обдумывания. В результате внимание сосредоточивается на небольшом числе привычных (а подчас и набивших оскомину), хотя и важных, но по своему значению частных и узких вопросов, вроде, например, уже пресловутого различения между прямыми и обратными теоремами. Между тем оставляются в тени вопросы гораздо более общего принципиального значения.

Я думаю, что основным общим моментом воспитательной функции математического образования — моментом, который в значительной степени обусловливает собою всё остальное,— служит приучение воспитываемых к *полноценности аргументации*.

В обыденной жизни, даже в «любительских» (не строго научных) принципиальных спорах, мы, защищая какое-либо утверждение, довольствуемся обычно одним-двумя аргументами, говорящими в его пользу. Противник может привести в ответ несколько аргументов, говорящих против нашего утверждения. Однако обычно ни та, ни другая аргументация не бывает исчерпывающей; противники продолжают изыскивать новые аргументы, каждый в пользу своей точки зрения, и спор продолжается.

Примерно так же протекают и научные дискуссии в тех областях знания, которые не входят в число так называемых «точных» наук; конечно, аргументация здесь бывает уже, как правило, более полной, чем в обыденных спорах; но почти никогда не удается сделать ее исчерпывающей, не допускающей никаких возражений, и тем самым ликвидирующей самую дискуссию.

Иначе обстоит дело в математике. Здесь аргументация, не обладающая характером полной, абсолютной исчерпанности, оставляющая хотя бы малейшую возможность обоснованного возражения, беспощадно признается ошибочной и отбрасывается, как лишенная какой бы то ни было силы. В математике нет и не может быть «наполовину доказанных» и «почти доказанных» утверждений: либо полноценность аргументации такова, что никакие споры о правильности доказываемого утверждения более невозможны, либо аргументация вообще полностью отсутствует.

Изучая математику, школьник впервые в своей жизни встречает столь высокую требовательность к полноценности аргументации. Вначале она удивляет, отталкивает, пугает его, кажется ему излишней, сверхмерной, педантичной. Но постепенно, день за днем, он к ней привыкает. Хороший учитель много может сделать для того, чтобы этот процесс протекал и быстрее, и продуктивнее. Он приучит своих учеников к взаимной критике; когда один из них что-либо доказывает или решает какую-либо задачу перед всем классом, все остальные должны напряженно искать возможных возражений и немедленно их высказывать. Ученик, который «отобьется» от всех таких возражений, заставит умолкнуть всех своих критиков, неизбежно

испытает законную радость победы. Вместе с тем он ясно почувствует, что именно логическая полноценность аргументации была тем оружием, которое дало ему эту победу. А раз почувствовав это, он неизбежно научится уважать это оружие, постарается, чтобы оно всегда было при нем. И, конечно, не только в математических, но и в любых других дискуссиях он всё больше и настойчивее будет стремиться к полноценности аргументации. Каждый раз перед ним будет вставать задача — по возможности обезоружить своих противников, в полной мере используя весь запас аргументов, какие вообще возможны в данной ситуации.

Этот воспитывающий процесс имеет решающее значение для логической культуры мышления, в особенности если учесть, что учащийся привыкает быть беспощадно требовательным к полноценности аргументации не только в споре, но и в своем одиноком мышлении. Процесс этот протекает повседневно на наших глазах у многих тысяч школьников. Он неизбежно возникает и идет своим путем без нашего специального вмешательства, но это не значит, конечно, что мы вправе предоставить его такому самотеку; в нашей власти сделать его и более быстрым и более полным по богатству и прочности достижений; а раз мы можем, то мы, очевидно, и должны это делать; вопрос о том, какими приемами наиболее эффективно можно добиться этих целей, есть уже методическая задача, которую мы не можем здесь детально рассматривать.

Общий принцип борьбы за полноценность аргументации получает в ходе интеллектуального развития учащегося целый ряд типичных по своей форме конкретных разновидностей, важнейшие из которых мы теперь перечислим.

1) Борьба против незаконных обобщений. Натуралист, подметив наличие какого-либо свойства (признака) у ряда особей данного вида, с чистой научной совестью объявляет этот признак общим для всего рассматриваемого вида; и никто не упрекнет его за это,— такого рода индуктивные заключения представляют собою один из основных методологических стержней естественных наук. Конечно, и в этих науках координирующая и осмысливающая теоретическая мысль возможна и необходима; но как исходным пунктом, так и решающей проверкой всякого заключения здесь всегда остаются наблюдение или опыт, осуществляемые над отдельными экземплярами.

В математике дело обстоит принципиально иначе. Если мы проверили, что несколько десятков (или хотя бы и несколько миллионов) наудачу выбранных нами треугольников обладают каким-нибудь свойством, мы еще не вправе признать это свойство принадлежащим всем треугольникам. Такое заключение было бы не до конца

обоснованным, а в математической науке всё, что не обосновано до конца, расценивается как абсолютно необоснованное. Только исчерпывающее общее доказательство может дать уверенность в том, что данный признак действительно является общим свойством всех треугольников.

Чему же может и должна научить школьника та суровая критика по адресу не вполне обоснованных обобщений, с какою он встречается в математике? Конечно, он не должен стараться переносить такого рода требования на выводы других наук и, тем более, на практические жизненные ситуации. Требование абсолютной полноты индукции специфично для математического метода и совершенно невыполнимо ни в естественных науках, ни в практической жизни. Но привычка с критической тщательностью проверять законность всякого обобщения, привычка твердо помнить, что замеченное во многих случаях еще не обязано тем самым иметь место во всех случаях и что закономерности, установленные на основе (хотя бы и многих) единичных наблюдений и опытов, требуют поэтому всё новой и новой проверки,— все эти важнейшие методологические навыки, необходимые в любой научной и практической деятельности, в значительной степени воспитываются и укрепляются вместе с повышением математической культуры.

Это — процесс, который мы каждодневно видим происходящим на наших глазах.

2) Борьба против необоснованных аналогий. Заключения по аналогии служат обычным и законным приемом установления новых закономерностей как в эмпирических науках, так и в обыденной жизни. Если, допустим, естествоиспытатель помнит, что все встречавшиеся ему до сих пор виды, обладавшие признаками A и B, обладали также и признаком C, и если он нашел новый вид, у которого обнаружены признаки A и B, то он, естественно, заключит, что этот новый вид обладает также и признаком C. Такое заключение по аналогии значительно выигрывает в убедительности, если к чисто эмпирическим данным, описанным выше, присоединяются, как это часто бывает, какие-либо теоретические соображения, заставляющие предполагать, что совместное наличие признаков A, B и C является не случайным, а обосновано теми или другими общими принципиальными соображениями. Но только в математике возможно — и вместе с тем совершенно необходимо — требовать, чтобы эти принципиальные соображения были доведены до степени исчерпывающего доказательства. Либо мы со всей строгостью доказали, что из наличия признаков A и B с неизбежностью вытекает и наличие признака C; либо, если нам не удалось доказать этого с исчерпывающей полнотой, нам запрещается делать из наличия признаков A и B какие бы то ни было выводы относительно признака C. Но в первом случае (т. е. когда доказана теорема «из

A и B следует C») простое применение этой общей теоремы к конкретным частным случаям уже вряд ли может быть названо «заключением по аналогии». Будет, таким образом, правильно сказать, что в математике заключения по аналогии категорически запрещены (что не должно, конечно, умалять огромного эвристического значения заключений по аналогии), в то время как в эмпирических науках и практической деятельности заключениям по аналогии принадлежит почетная роль одного из основных приемов вывода новых закономерностей. Поэтому снова встает вопрос о том, что́ же в этом отношении могут дать уроки математики для воспитания общей культуры мышления? И снова приходится ответить на это то же, что и прежде: математическая вышколенность ума, привыкшего к тому, что заключение по аналогии может служить лишь эвристическим приемом, который сам по себе еще не имеет доказательной силы, неизбежно приучает прошедшего эту школу человека и во всех других областях мышления относиться к такого рода заключениям с большой осторожностью, памятуя, что во всех таких случаях нельзя без основательной проверки считать полученное заключение твердо установленным. Каждый из нас испытал в свое время на себе воспитывающее влияние этой особенности математического мышления, и каждодневно мы наблюдаем, ка́к влияние это содействует повышению мыслительной культуры наших воспитанников. Критическое отношение к заключениям по аналогии есть один из важнейших показателей, отличающих правильно воспитанное научное и практическое мышление от первобытного, обывательского, и занятия математикой всегда служат одним из основных средств воспитания этого важнейшего показателя.

3) Борьба за полноту дизъюнкций. Когда математик доказывает какое-либо общее свойство всех треугольников, то иногда ему приходится проводить доказательство отдельно для остроугольных, прямоугольных и тупоугольных треугольников. Известно, как часто в таких случаях начинающие делают ошибки, в особенности в тех случаях, когда рассуждение сопровождается ссылкой на чертеж; чертится, например, остроугольный треугольник, и рассуждение опирается на добавочные построения, которые либо невозможны, либо теряют доказательную силу, если выбранный треугольник имеет тупой угол. В математике такое рассуждение признается ошибочным, так как здесь нарушено основное требование полноты дизъюнкции: не предусмотрены все возможные разновидности данной ситуации, одна из них выпала из поля зрения.

В обыденных, не научных рассуждениях это требование нарушается на каждом шагу. Рассмотрев две-три наиболее часто встречающиеся или наиболее бросающиеся в глаза разновидности данной ситуации и убедившись, что в каждом из этих случаев мы неизбежно встречаемся с некоторым событием A, мы заключаем, что

это событие A сопутствует данной ситуации во всех случаях, хотя на самом деле данная ситуация может иметь, кроме двух-трех изученных нами, еще десяток других разновидностей, и среди этих разновидностей, скинутых нами со счета, могут быть и такие, в которых наступление события A вовсе не обязательно. Мы говорим, например, что ученика Иванова вообще нельзя дисциплинировать, потому что на него испытанным образом не действуют ни ласка, ни угрозы. Мы забываем при этом, что лаской и угрозами не исчерпываются еще все разновидности приемов дисциплинирующего воздействия, что существует еще, например, метод спокойного убеждения и что, стало быть, наша дизъюнкция страдает неполнотой. Мы часто наблюдаем, как начинающий, рассмотрев при исследовании какого-нибудь уравнения случай, когда некоторый данный коэффициент положителен, а затем — случай, когда этот коэффициент отрицателен, тем самым считает, что он провел исследование во всех случаях, забывая, что изучаемый коэффициент может оказаться равным нулю. Здесь также мы видим неполноту дизъюнкции, которая может привести и фактически приводит к тяжелым ошибкам в выводах.

В противоположность тем двум требованиям, которые мы рассматривали выше, требование полноты дизъюнкции, учета всех возможных разновидностей изучаемой ситуации, является необходимой принадлежностью не только математического, но и всякого правильного мышления. Аргументация, в которой не учтены все имеющиеся возможности, всегда оставляет место для законных возражений и потому не может быть признана полноценной. Военачальник, предпринимая какой-либо маневр, при учете его последствий должен предвидеть все возможные ответы врага; просмотр хотя бы одного из них может оказаться гибельным. Юридический кодекс в каждой статье обязательно должен охватывать все мыслимые разновидности данной ситуации, иначе он ставит судью перед необходимостью решать дела по своему произволу.

Но нигде требование безукоризненной чистоты дизъюнкций не выставляется так явно и категорически, как в математике; и никто не обрушивается с такой быстротой и беспощадностью на замеченный просмотр в дизъюнкции, как вышколенный математик. Вот почему уроки математики должны воспитывать и действительно воспитывают в мышлении учащихся этот важнейший закон правильного рассуждения в несравненно большей мере, чем занятия другими предметами.

4) Борьба за полноту и выдержанность классификации. Классифицирует не только ученый-теоретик в своем кабинете; классификацией приходится очень часто заниматься и практическому работнику: инженеру, врачу, учителю, статистику, агроному. Общеизвестно, что невышколенный ум склонен допускать, производя классификацию, ряд типических ошибок; наиболее распространенными из таких ошибок являются нарушение *полноты* клас-

сификации и нарушение ее *выдержанности*, единопринципности. Нарушение полноты классификации состоит в том, что остаются понятия, не входящие ни в один из названных классов, и что, стало быть, названы не все классы. Простые примеры: на вопрос «какие ты знаешь растения?» школьник отвечает «травы и деревья», забывая о кустарниках, лишаях и многих других типах; войсковые части делятся на сухопутные, водные и воздушные (упускаются интендантские, части связи и многие другие); натуральные числа делятся на простые и составные (упускается число 1); вещественные числа делятся на положительные и отрицательные (упускается нуль).

Требование полноты классификации формально аналогично рассмотренному нами выше требованию полноты дизъюнкции, но, конечно, отлично от него по содержанию. Там шла речь об обязательности охвата всех могущих возникнуть ситуаций, здесь же — о необходимости перечисления всех разновидностей некоторого понятия. Но здесь, как и там, явно и неукоснительно требование полноты классификации провозглашается в математике преимущественно перед всеми другими науками, и потому уроки математики более всех других воспитывают в школьнике этот обязательный элемент правильного мышления.

Требование выдержанности классификации состоит в том, чтобы она проводилась по единому принципу, по единому признаку. Это требование, при строго правильном мышлении совершенно обязательное, очень часто нарушается не только в обывательских рассуждениях, но и в серьезной практике. Вот простые примеры такой невыдержанной классификации: суда делятся на весельные, парусные, моторные и военные; очевидно, классификация начата по принципу различных движущих сил, и последняя рубрика этот принцип нарушает; другой пример: обувь подразделяют на кожаную, брезентовую, резиновую и модельную — та же картина. Конечно, подобного рода перечисления не всегда претендуют на роль классификации, и в таких случаях соблюдение единого принципа не обязательно (например, объявление: завод приглашает на работу плотников, штукатуров, женщин и подростков). Но во всех случаях, когда такому перечислению приписывается классифицирующая функция, невыдержанность разделяющего принципа вызывает такую неотчетливость всей схемы, которая может привести и к теоретическим смешениям, и к практической путанице. Поэтому логически вышколенный ум всегда ощущает недостаток выдержанности классификации как существенный дефект рассуждения. И снова наиболее чувствительна к этому дефекту математическая наука, и поэтому именно на уроках математики школьник преимущественно развивает в себе эту потребность видеть всякую классификацию выдержанной, построенной на едином классифицирующем принципе.

Я перечислил те моменты в борьбе за правильность мышления и полноценность аргументации, которые представляются мне наиболее важными. Как уже было сказано выше, я не могу входить в этой статье в обсуждение тех методических приемов, с помощью которых учитель математики может достигнуть наибольшего успеха в деле воспитания у своих учеников перечисленных мною моментов правильного мышления. Но я считаю необходимым сделать по этому вопросу одно методическое замечание общего характера (для опытного учителя, впрочем, совершенно очевидное): все те требования правильного мышления, о которых шла речь выше, должны воспитываться в учащихся исподволь, от случая к случаю, без излишнего педалирования; не может быть и речи о том, чтобы посвящать специальный урок, например, борьбе с незаконными аналогиями; такая постановка дела может только безнадежно погубить весь ожидаемый эффект. Надо, напротив, всемерно избегать во всем этом деле общих рассуждений и обращать внимание учащихся на тот или другой логический момент исключительно на базе ярко убедительного конкретного математического материала. Потребность в логической полноценности аргументации воспитывается не постоянным надоедающим напоминанием о необходимости этой полноценности, а показом на конкретных примерах (поводы к которым дает почти каждый урок), как несоблюдение этого требования ведет к ошибкам и неувязкам. Надо не отвлеченно проповедовать полноценность аргументации, а приучить учащегося к тому, что каждый пробел в аргументации немедленно вызывает придирчивый вопрос со стороны учителя или, что много лучше, со стороны товарищей.

Я не буду говорить здесь о том, что следует использовать уроки математики для правильного понимания различия между прямым и обратным утверждениями, а также и ряда других аналогичных различений. С одной стороны, об этом так много уже писалось, что вряд ли я смог бы прибавить здесь что-нибудь новое. С другой стороны, мне представляется, что этого рода моменты, будучи, конечно, обязательными для логически правильного мышления, всё же по своему частному, специальному характеру не имеют вне математики столь существенного значения, как те значительно более общие принципы, которые я перечислил выше.

2. Стиль мышления. Помимо специфических, особо строгих требований к логической правильности умозаключений, математика отличается от других преподаваемых в школе наук также и стилем своего мышления. Стиль этот хотя и претерпевает на протяжении веков, и даже десятилетий, довольно значительные изменения, всё же имеет некоторые общие для всех эпох непреходящие черты, заметно отличающие его от стилей, принятых в других науках.

Утвердившийся в той или другой науке стиль мышления не является, как можно было бы думать, только внешним и потому второстепенным фактором, имеющим лишь эстетическую ценность и не могущим поэтому существенно влиять на развитие данной науки. Напротив, стилем мышления в значительной степени определяется отчетливость теоретических связей, простота и ясность научных конструкций, наглядная конкретность понятий и многое другое, от чего в свою очередь зависят эффективность, плодотворность научных дискуссий и научного преподавания, а вместе с тем и темпы развития науки.

Среди тех особых черт, которые присущи стилю математического мышления, имеется ряд таких, которым свойственно весьма общее и широкое значение; такая черта, если она усваивается представителем какой-нибудь другой науки или практическим деятелем, оказывает нередко весьма существенные услуги как его собственному мышлению, так и усвоению его трудов учениками и последователями. Читая сочинения кого-либо из крупнейших классиков в другой научной области, математик подчас с некоторым удивлением восклицает: «Да ведь он мыслит совсем по-нашему!»,— удивление происходит оттого, что обычно в этой научной области принят совсем иной стиль мышления, имеющий очень мало общего с математическим.

Но если усвоение некоторых черт математического мышления способно облагородить мыслительный стиль и в других областях знания и практической деятельности, сделать этот стиль более мощным и продуктивным орудием мысли, то очевидно, что не следует пренебрегать использованием уроков математики для приучения молодых умов к постепенному усвоению этих черт, к тому, чтобы эти черты стали прочными навыками их мышления — сначала в пределах математики, а потом и за ее пределами. Для того чтобы это осуществить, надо в первую очередь постараться со всей тщательностью выявить те черты стиля математической мысли, о которых здесь идет речь.

В основе каждого правильно построенного хода мыслей, независимо от предметного содержания его, лежит такая формально-логическая схема, которая ощущается вышколенным умом как некий логический костяк, стройный и закономерный, обросший тем или другим конкретным содержанием. Независимо от стиля мышления эта логическая схема должна быть закономерной, лишенной пробелов,— без этого рассуждение становится недоброкачественным и должно быть отвергнуто.

Однако роль и положение этого логического скелета в данном ходе мыслей бывают весьма различны и существенным образом зависят именно от стиля мышления. В одних случаях логическая схема становится определяющим, руководящим моментом мышления, так что мыслящий всё время имеет ее перед глазами и сообразно с нею выбирает и направляет последовательные этапы рассуждения. В других, напротив, логический костяк остается затушеванным, мысль в гораздо большей степени направляется запросами конкретного содержания, роль логики сводится к последующему контролю, да и этот контроль в письменном или устном изложении часто только подразумевается и явно не проводится; логическая схема как целое остается вне поля зрения мыслящего. Разумеется, встречаются нередко и стили мышления, промежуточные между двумя указанными.

Для математики характерно доведенное до предела доминирование логической схемы рассуждения; математик, потерявший, хотя бы временно, из виду эту схему, вообще лишается возможности научно мыслить. Эта своеобразная черта стиля математического мышления, в столь полной мере не встречающаяся ни в одной другой науке, имеет в себе много ценного. Очевидно, что она в максимальной степени позволяет следить за правильностью течения мысли и гарантирует от ошибок; с другой стороны, она заставляет мыслящего при каждой дизъюнкции иметь перед глазами всю совокупность имеющихся возможностей и обязывает его учесть каждую из них, не пропуская ни одной (такого рода пропуски вполне возможны и фактически часто наблюдаются при других стилях мышления). Поэтому приобретенные на уроках математики стилистические навыки, связанные с указанной чертой, имеют существенное значение для повышения общей культуры мышления учащихся.

Очень интересным и ярким примером мышления в далекой от математики области, и тем не менее чрезвычайно насыщенного этой чертою, могут служить произведения Маркса. Читателя, который после изучения экономических трудов других ученых раскрывает «Капитал», с первых страниц поражает железная, непреклонная логика его строк. Логическая схема с ее неумолимыми требованиями не только определяет ход мысли автора, но и настойчиво убеждает читателя, который не может уйти от ее направляющего влияния. Этот необычный для экономического сочинения стиль, почти приближающийся к математическому, неизменно вызывает в читателе ощущение прочности, надежности, предельной убедительности и в то же время много помогает ему в усвоении читаемого.

Второй характерной чертой математического стиля мышления, о которой здесь должно быть упомянуто, является его *лаконизм*, сознательное стремление всегда находить кратчайший ведущий к данной цели логический путь, беспощадное отбрасывание всего, что не абсолютно необходимо для безупречной полноценности

аргументации. Математическое сочинение хорошего стиля не терпит никакой «воды», никаких украшающих, ослабляющих логическое напряжение разглагольствований, отвлечений в сторону; предельная скупость, суровая строгость мысли и ее изложения составляют неотъемлемую черту математического мышления. Черта эта имеет большую ценность не только для математического, но и для любого другого серьезного рассуждения; лаконизм, стремление не допускать ничего излишнего, помогает и самому мыслящему, и его читателю или слушателю полностью сосредоточиться на данном ходе мыслей, не отвлекаясь побочными представлениями и не теряя непосредственного контакта с основной линией рассуждения.

Корифеи науки, как правило, мыслят и выражаются лаконично во всех областях знания, даже тогда, когда мысль их создает и излагает принципиально новые идеи. Какое величественное впечатление производит, например, благородная скупость мысли и речи величайших творцов физики: Ньютона, Эйнштейна, Нильса Бора! Может быть, трудно найти более яркий пример того, какое глубокое воздействие может иметь на развитие науки именно *стиль* мышления ее творцов.

Для математики лаконизм мысли является непререкаемым, канонизированным веками законом. Всякая попытка обременить изложение не обязательно нужными (пусть даже приятными и увлекательными для слушателей) картинами, отвлечениями, разглагольствованиями заранее ставится под законное подозрение и автоматически вызывает критическую настороженность. И поэтому именно уроки математики призваны дать учащимся, предпочтительно перед другими предметами, навыки лаконического, прямого, не знающего отвлечений, не обремененного никакими излишними элементами мышления.

Далее, для стиля математического мышления характерна *четкая расчлененность* хода рассуждения. Если, например, при доказательстве какого-либо предложения мы должны рассмотреть четыре возможных случая, из которых каждый может разбиваться на то или другое число подслучаев, то в каждый момент рассуждения математик должен отчетливо помнить, в каком случае и подслучае его мысль сейчас обретается и какие случаи и подслучаи ему еще остается рассмотреть. При всякого рода разветвленных перечислениях математик должен в каждый момент отдавать себе отчет в том, для какого родового понятия он перечисляет составляющие его видовые понятия. В обыденном, не научном мышлении мы весьма часто наблюдаем в таких случаях смешения и перескоки, приводящие к путанице и ошибкам в рассуждении. Часто бывает, что человек начал

перечислять виды одного какого-нибудь рода, а потом незаметно для слушателей (а часто и для самого себя), пользуясь недостаточной логической отчетливостью рассуждения, перескочил в другой род и заканчивает заявлением, что теперь оба рода расклассифицированы; а слушатели или читатели не знают, где пролегает граница между видами первого и второго рода.

Для того чтобы сделать такие смешения и перескоки невозможными, математики издавна широко пользуются простыми внешними приемами нумерации понятий и суждений, иногда (но гораздо реже) применяемыми и в других науках. Те возможные случаи или те родовые понятия, которые надлежит рассмотреть в данном рассуждении, заранее перенумеровываются; внутри каждого такого случая те подлежащие рассмотрению подслучаи, которые он содержит, также перенумеровываются (иногда, для различения, с помощью какой-либо другой системы нумерации). Перед каждым абзацем, где начинается рассмотрение нового подслучая, ставится принятое для этого подслучая обозначение (например II 3,— это означает, что здесь начинается рассмотрение третьего подслучая второго случая, или описание третьего вида второго рода, если речь идет о классификации). И читатель знает, что до тех пор, покуда он не натолкнется на новую числовую рубрику, всё излагаемое относится только к этому случаю и подслучаю. Само собою разумеется, что такая нумерация служит лишь внешним приемом, очень полезным, но отнюдь не обязательным, и что суть дела — не в ней, а в той отчетливой расчлененности аргументации или классификации, которую она и стимулирует, и знаменует собою.

Наконец, следует упомянуть еще об одной чисто внешней традиции математического стиля, могущей при надлежащих условиях приобрести воспитательное значение, которым нельзя пренебрегать. Я имею в виду свойственную математике скрупулезную *точность символики*. Каждый математический символ имеет строго определенное значение; замена его другим символом или перестановка на другое место, как правило, влечет за собою искажение, а подчас и полное уничтожение смысла данного высказывания. Учащийся, не привыкший еще относиться с достаточной требовательностью к точности устной речи и письменного изложения, вначале может с некоторым легкомыслием отнестись к неуклонным и настойчивым приглашениям учителя математики — вести математическую запись с абсолютной точностью; эти требования могут даже показаться ему педантическими и вызвать насмешку. Однако он очень быстро убедится на собственном опыте, что несоблюдение безукоризненной точности символической записи в математике влечет за собою немедленную расплату: он сам теряет возможность понять смысл записанного, вынужден

гадать, угадывает неверно и либо получает неправильный ответ, либо вообще лишает себя возможности решить задачу. В лучшем случае ему ценою значительных усилий удастся восстановить правильную запись и шествовать дальше, отправляясь от нее.

Убедившись таким образом, что точность символической записи соответствует его собственным интересам, он начинает следить за собою в этом направлении, и постепенно строгая правильность математической символики становится его привычкой. Но такого рода привычка, приобретенная в какой-либо одной сфере мышления, неизбежно приводит к воспитанию и общего стиля мышления учащегося; он начинает точнее выражаться и в устной речи, и в письменном изложении; в частности, он уделяет больше внимания правописанию, орфографические ошибки переживаются им с такой же остротой и таким же беспокойством, как математические. Мы неизменно наблюдаем, что ученики, научившиеся требовательно относиться к точности математической символики, легче и быстрее перестают делать орфографические ошибки. И я не знаю, возможно ли окончить школу, обладая требуемой для аттестата зрелости математической культурой и не научившись в то же время писать совершенно безошибочно.

Заканчивая эту главу, посвященную вопросам воспитательного воздействия уроков математики на культуру мышления учащихся, я предвижу естественное и законное недоумение читателя по поводу того, что мною нигде даже не затронута проблема развития элементов диалектического мышления. Я считаю себя обязанным дать по этому вопросу краткое разъяснение.

Маркс и Энгельс с полным основанием утверждали, что математика не только дает для законов диалектического мышления богатейший иллюстративный материал, но систематически способствует развитию диалектических навыков мыслительного процесса. Однако, как это неоднократно отмечалось основоположниками марксизма, в полной мере это может быть отнесено лишь к так называемой «высшей» математике, т. е. к математике переменных величин. Именно здесь мы приучаемся к математическому исследованию явлений природы и процессов техники в их живой изменчивости, а не статической неподвижности. Именно здесь величины исследуются в их взаимной зависимости (понятие функции), а не в отрыве друг от друга. Нигде с такою наглядностью, как здесь, мы не видим в действии переход количества в качество, диалектический синтез первоначально антагонистических противоположностей и другие основные принципы диалектики. И это — одна из важнейших причин (впрочем, далеко не единственная), заставляющих нас признать абсолютно необходимым введение элементов высшей математики в курс средней школы.

Но пока мы только боремся за это. Что же касается преподаваемой в школе «элементарной» математики, то и она, конечно, как всякая подлинная и живая наука, не лишена диалектических элементов. Но здесь они выступают разрозненно и с малой мощностью, и говорить о них в статье, посвященной лишь основным рычагам воспитательного воздействия уроков математики, я не решился. Впрочем, я имею в виду в ближайшем будущем составление другой статьи, специально посвященной вопросу о необходимости введения в школьное преподавание элементов высшей математики; в этой статье я надеюсь дать развернутую и убедительную картину того, каким мощным орудием воспитания навыков диалектического мышления могли бы стать уроки математики переменных величин.

II. МОРАЛЬНЫЕ МОМЕНТЫ В ВОСПИТАНИИ

О роли и значении уроков математики в воспитании правильного и дисциплинированного мышления говорилось и писалось очень много. Напротив, о влиянии математических занятий на формирование характера и моральной личности учащегося не сказано почти ничего. Это вполне понятно: по абстрактности своего предмета математическая наука не может давать учащемуся тех непосредственно впечатляющих, этически воздействующих и формирующих характер образов, картин, эмоций, какими располагают, скажем, история или литература. Было бы, однако, весьма поверхностно делать отсюда тот вывод, что в деле формирования нравственной личности школьника уроки математики вообще должны быть скинуты со счетов. По моему многолетнему опыту работа над усвоением математической науки неизбежно воспитывает — исподволь и весьма постепенно — в молодом человеке целый ряд черт, имеющих яркую моральную окраску и способных в дальнейшем стать важнейшими моментами в его нравственном облике. Сделать этот процесс более активным и результаты его более прочными — достойная задача для учителя. Но прежде всего надо тщательно разобраться в том, что́ это за черты и какие особенности математической работы способны их воспитывать.

1. Честность и правдивость. В обывательских тяжбах всякого рода каждая из спорящих сторон исходит, как правило, из желательного ей, выгодного для нее решения вопроса и с большей или меньшей изобретательностью изыскивает возможно более убедительную аргументацию для решения вопроса в свою пользу. В зависимости от эпохи, среды и содержания спора стороны при этом апеллируют к тому или другому высшему авторитету — общечеловеческой морали, «естественному» праву, священному писанию, юридическому кодексу, действующим правилам внутреннего распорядка, а часто и к высказываниям отдельных авторитетных ученых или признанных политических руководителей. Все мы много раз наблюдали, с какою страст-

ностью ведутся подобного рода споры и какой убежденностью дышит по видимости аргументация каждой из сторон; можно подумать, что такой тяжущийся действительно обуреваем желанием найти и отстоять истинное, справедливое, отвечающее духу и букве признанного в качестве арбитра авторитетного источника решение.

Но хорошо известно, что эту картину мы часто наблюдаем не в одних только обывательских тяжбах. В точности те же черты являет подчас и научная дискуссия. Выводы, с полной убежденностью сделанные одним ученым, с такою же убежденностью оспариваются другими; завязывается полемика, в которой каждая из сторон находит всё новые и новые аргументы в пользу своей позиции — даже вновь поставленные опыты часто говорят каждому из спорящих как раз то, что ему желательно. В ходе полемики каждая из сторон не только стремится всё более и более усиливать свою собственную позицию, но и старается различными средствами дискредитировать позицию противной стороны, доходя иногда и до попыток персональной дискредитации. И лишь сравнительно редко бывает, чтобы в такой затянувшейся полемике одна из спорящих сторон нашла честность и мужество признать свою позицию ошибочной.

Субъективные основания такого рода явлений в жизни науки легко понятны: они ничем, к сожалению, не отличаются по своей неприглядности от субъективных оснований самых мелочных обывательских стычек. Что касается объективных оснований возможности подобного рода научных ситуаций, то и их нетрудно найти: в эмпирических науках всякая новая, еще не окончательно установленная закономерность фигурирует, по крайней мере временно, в качестве «рабочей гипотезы»; покуда вопрос не решен окончательно, имеются обычно как соображения (опытные и теоретические), говорящие в пользу этой гипотезы, так и такие, которые говорят против нее. Из двух ученых один может поставить своей задачей — собрать как можно больше аргументов, поддерживающих такую гипотезу, а другой — заняться собиранием фактов и соображений, способных вызвать к ней недоверие. Дело происходит как в уголовном процессе, где перед обвинителем и защитником ставятся задачи собрать, привести в порядок и изложить все аргументы, соответственно говорящие за и против виновности подсудимых.

Само собою разумеется, что так поставленная научная дискуссия сама по себе не содержит еще ничего морально одиозного: собрать с возможною полнотою все имеющиеся аргументы за и против данной «рабочей гипотезы» — это во всех случаях приносило пользу прогрессу науки; нет, очевидно, ничего предосудительного и в том, что сбор аргументов за и против гипотезы выполняется двумя различными учеными (или группами ученых), если только обе стороны подходят к своей задаче добросовестно, руководствуясь исключительно желанием способствовать отысканию объективной истины. Моральный одиум,

этическое неблагополучие начинаются там, где в своих выводах ученый перестает руководствоваться интересами объективной истины, а стремится поставить эти выводы — сознательно, полусознательно или бессознательно — на службу своим личным интересам — своему упрямству, своему честолюбию, своему корыстолюбию, когда аргументация приводится с пристрастием, «притягивается за волосы», необъективно акцентируется, точь-в-точь как в мелком сутяжничестве или в обывательских дрязгах. Такая деградация научного спора в иных случаях ложится мрачным пятном даже на крупнейших представителей научной мысли; среди же ученых меньшего ранга она представляет собою, к сожалению, довольно распространенное явление.

Одна только математическая наука полностью от всего этого избавлена. Она не знает «рабочих гипотез» — предложений, истинность которых может подлежать дискуссии. Пока предложение не доказано, оно вообще никак не входит в сокровищницу науки, никому не придет в голову его отстаивать; если же оно доказано, то истинность его никак не может быть подвергнута сомнению: оно является абсолютно общеобязательным. Никаких промежуточных ситуаций математика не знает. Полемизировать, например, в защиту неполноценного доказательства в математике может только неуч, шарлатан или душевно больной (все три категории действительно от времени до времени встречаются, достаточно вспомнить так называемых «фематистов», рыцарей квадратуры круга и трисекции угла); но такой «защитник» немедленно, единогласно и беспощадно разоблачается научным миром. Никакая аргументация с пристрастием или тенденцией, никакое «притягивание за волосы» ни при каких обстоятельствах не может в математике иметь успеха. Разумеется, это относится только к содержанию самой математической науки; в вопросах логического или философского обоснования математики дискуссии возможны и даже неизбежны; возможны (и к сожалению нередки) и споры персонального характера, связанные с развитием математики (например, по вопросам приоритета).

Каждый математик рано привыкает к тому, что в его науке всякая попытка по тем или иным мотивам действовать тенденциозно, заранее склоняясь к тому или другому решению вопроса и прислушиваясь только к аргументам, говорящим в пользу избранного решения,— всякая такая попытка заведомо обречена на неудачу, и ничего кроме разочарования пытающемуся принести не может. Такое положение, при котором неправильная или не до конца правильная аргументация могла бы оказаться выгодной для аргументирующего, здесь просто принципиально невозможно. Поэтому математик быстро привыкает к тому, что в его науке выгодна только правильная, объективная, лишенная всякой тенденциозности аргументация, что успех может принести только непредубежденное, беспристрастное напряжение мысли. И независимо от своего общего морального уровня он в своей

научной работе всегда руководствуется исключительно соображениями объективной истинности.

Но эту черту, естественно развивающуюся у математика-специалиста, в известной степени воспитывает в себе, занимаясь математикой, и каждый неспециалист, в частности каждый школьник. Ему хорошо известно, что «втереть очки» учителю математики невозможно, что никакой апломб и никакое красноречие не помогут ему выдать незнание за знание, неполноценную аргументацию за полноценную. И как бы лжив он ни был в других отношениях, в математике он остережется отстаивать неверное утверждение или неправильное доказательство.

Но и здесь, как это часто бывает, моральные навыки, приобретенные в какой-либо одной области, в известной мере переносятся и на другие сферы мышления и практической деятельности. Теоретическая честность, ставшая для математика непреложным законом его научного мышления и профессиональной (в частности, педагогической) деятельности, довлеет над ним во всех его жизненных функциях — от абстрактных рассуждений до практического поведения.

Я всегда интересовался этой чертою и много раз наблюдал, как она развивается в людях под влиянием серьезного научного общения, в частности под воздействием уроков математики. Это очень радостная и морально возвышающая картина, когда человек постепенно преодолевает в себе отвратительную мещанскую привычку — подчинять законы мышления своим личным, мелким, корыстным интересам, теоретически защищать всё то и только то, что ему практически выгодно; когда он научается уважать объективную правильность аргументации как высшую духовную и культурную ценность и всё чаще и со всё более легким сердцем жертвовать ради нее своими личными интересами. Доведенная до предела, эта черта представляет собою не что иное, как честность и правдивость — одно из лучших украшений нравственной личности человека.

2. Настойчивость и мужество. Добросовестная и серьезная работа над приобретением и укреплением знаний в любой научной области требует систематического напряжения умственных усилий, настойчивости в преодолении трудностей, мужественной встречи неудач; поэтому такая работа при правильном руководстве неизбежно воспитывает у учащегося соответственные черты характера: трудолюбие, усидчивость, упорство в преследовании намеченной цели, умение не останавливаться перед трудностями и не впадать в уныние при неудачах. Непосредственно ясно, какое решающее значение имеют все эти черты для развития морально и общественно полноценной человеческой личности, и с каким вниманием должен поэтому учитель следить за максимальным использованием своих уроков в целях воспитательного воздействия в указанном направлении. Те возможности,

которыми для этого располагают предметы школьного обучения, весьма многочисленны и многообразны, и нет такого предмета, в специфических чертах которого не было бы заложено особых, именно этому предмету свойственных движущих рычагов такого воспитательного воздействия. Наша задача здесь, естественно, должна состоять в указании тех черт математики как школьного предмета, которые, отличая ее от других предметов школьного преподавания, способствуют развитию в учащихся разумной настойчивости и сознательного мужества — этих неоценимых качеств будущего борца.

Прежде всего я хочу здесь отметить четкую определенность поставленной цели, желаемого и требуемого результата каждого математического задания. Если заданием служит сочинение исторического или литературного содержания, то нельзя указать момента, когда такое задание дефинитивно закончено выполнением — возможности дополнения и усовершенствования, систематических улучшений всякого рода здесь почти безграничны; с другой стороны, учащийся не чувствует себя здесь достаточно компетентным для авторитетной оценки своей работы: то, что ему представляется в его сочинении вполне удачным, может встретить совсем иную оценку со стороны учителя. Вся эта, по существу данного задания неизбежная неопределенность, расплывчатость в оценке законченности и качества проделанной работы должна, несомненно, оказывать некоторое расслабляющее влияние на волевое напряжение еще мало вышколенного молодого ума.

В математике дело обстоит иначе. Если заданием служит решение задачи или доказательство теоремы, то тем самым указывается с полной определенностью и тот момент, когда задание может считаться окончательно выполненным: когда решена задача или доказана теорема; всё остальное — изложение найденного решения, правильность и аккуратность записи и т. п. имеет и в глазах учителя и в глазах ученика лишь второстепенное, не решающее значение. Равным образом и *качество* работы здесь оценивается с однозначной определенностью: задача должна быть решена в е р н о, теорема должна быть доказана п р а в и л ь н о. Проверить отсутствие логических ошибок в своем рассуждении ученик может и должен уметь сам; в случае задачи он знает даже определенные приемы проверки решения. Легко понять, какое стимулирующее влияние на упорство, настойчивость в достижении цели может оказать и действительно оказывает эта четкая определенность показателей результата. Победа здесь так же непосредственно ощутительна, как в шахматной партии или спортивном состязании, и сам учащийся может с такой же уверенностью зафиксировать и оценить свое достижение, как и его авторитетный учитель.

Вторая, значительно более глубокая и важная черта математических заданий, которую я хочу здесь отметить, состоит в присущем им в значительном большинстве случаев *творческом* характере. В то

время как в большинстве других областей знания выполнение задания, за немногими исключениями, требует от учащегося лишь определенных знаний и навыков — в лучшем случае еще умения стройно и стилистически правильно излагать эти знания,— решение математической задачи, как правило, предполагает изобретение специального ведущего к поставленной цели рассуждения, и тем самым становится — пусть весьма скромным — творческим актом. Именно этот творческий, исследовательский характер математических заданий более чем что-либо другое влечет к себе молодые силы растущего и крепнущего интеллекта учащегося. Тот, кто раз изведал благородную радость творческого достижения, никогда уже не пожалеет усилий, чтобы вновь ее испытать. Никакие трудности его не остановят, сила его порыва и устремления, его усидчивость и выдержка в преодолении препятствий будут крепнуть с каждым новым достижением, а неудачи, ошибки, временные крушения и поражения он научится встречать, как подобает истинному борцу,— не опуская перед ними ру́ки, а черпая в них источник и стимул для все новых и новых напряжений мысли и воли.

ЗАКЛЮЧЕНИЕ

1. Я сознательно полностью оставил в стороне важнейший для темы настоящей статьи вопрос о значении уроков математики для формирования мировоззрения учащихся. Я сделал это по той же причине, по какой в свое время (в части I) отказался от рассмотрения вопроса об использовании уроков математики для воспитания навыков диалектического мышления: ознакомление с идеями и методами математической науки имеет фундаментальное мировоззренческое значение, но львиная доля воспитательного эффекта в этом направлении принадлежит математике переменных величин, так называемой высшей математике, с которой, по выражению Энгельса, в математику входит диалектика и которая, к сожалению, всё еще почти целиком остается за бортом наших школьных программ. Поистине не стоит при этих условиях говорить о влиянии на формирование мировоззрения той диалектически худосочной части математической науки, которая в настоящее время находит себе приют в нашем школьном преподавании. О тех же мощных сдвигах, направляющих и формирующих мировоззрение, которые может произвести и фактически часто производит изучение высшей математики, можно сказать очень много, и я собираюсь говорить о них в моей будущей статье, о которой я уже упоминал.

2. В этой статье я не касался методических вопросов. Я говорил только о том, какие особенности математической науки и для воспитания каких именно качеств интеллекта или моральной личности учащегося могут и должны быть использованы, но нигде не касался

вопроса о том, ка́к это может быть сделано. Мой опыт говорит мне, что это обстоятельство может вызвать недовольство в некоторых кругах читателей; вероятно, я не гарантирован от набивших уже оскомину упреков в том, что моя статья «ничего не дает учителю» и советов «повернуться лицом» и т. д. Поэтому я считаю необходимым сделать по этому поводу следующее краткое разъяснение.

1) Я считаю, что составление сколько-нибудь детальных методических указаний по рассматриваемым мною в настоящей статье вопросам действительно совершенно излишне. Сколько-нибудь заметный воспитательный эффект уроки математики (как и всякой другой науки) могут дать только при том условии, что учитель, во-первых, достаточно хорошо знает свою науку, ее методологию и ее историю, во-вторых, имеет достаточный педагогический такт и опыт и, наконец, в-третьих, сам обладает в достаточной мере всеми теми качествами, которые он собирается воспитывать в своих учениках. Учителю, который сам не умеет мыслить абсолютно отчетливо, никакие методические шпаргалки не помогут воспитать ясность мысли в учащихся.

И, напротив, если учитель стоит на высоте своей задачи, если он в полной мере обладает всеми перечисленными выше качествами, то никакие методические разработки по воспитательным вопросам ему заведомо не нужны — в каждом отдельном случае он с легкостью и непринужденностью сам найдет наиболее эффективный путь к поставленной цели. Навязывание ему определенной конкретной методики было бы для такого учителя только помехой в его работе.

2) Если, таким образом, я считаю составление детальных методических указаний по затронутым мною вопросам практически бесцельным, то может быть было бы полезно, всё же, дать по ним ряд общих методических советов; я думаю, что было бы хорошо, если бы моя статья побудила кого-либо из наших лучших учителей-методистов высказаться по этому вопросу и сделать достоянием младших товарищей некоторые общие выводы из своего опыта. В этом они во всяком случае компетентнее меня, и здесь я со всей необходимой скромностью должен уступить им слово.

3. Наконец, я хочу попытаться заранее оградить себя еще от одного рода упреков, которые я предвижу и которые обычно бывают основаны на недоразумении. Так как я говорил о воспитательном эффекте уроков математики, то мне, естественно, пришлось перечислять одну за другой именно те черты математической науки, которые в воспитательном отношении дают ей то или другое преимущество перед другими дисциплинами. В этом ведь и состояла моя задача. Но когда поступаешь таким образом, то у недостаточно вдумчивого читателя создается впечатление, будто бы ты поставил своей целью превознесение математики над всеми другими науками и вся твоя статья сплошь утверждает, что единственная подлинная наука

есть именно математика, все же другие дисциплины страдают теми или другими изъянами и науками могут быть названы лишь с известной оговоркой. Уважаемые коллеги — представители других наук — начинают чувствовать себя несправедливо обиженными и подвергают твою работу яростной критике, доказывая, что другие науки ничуть не хуже, и что всеми преимуществами, которыми в моем представлении монопольно обладает математика, на самом деле в такой же мере наделены и все прочие дисциплины.

С первым пунктом этого заявления я целиком и полностью согласен: другие науки действительно ничем не хуже математики; более того, представители этих других наук обычно вызывают во мне глубочайшее уважение тем, что творят великие ценности в таких областях, в которых творчество на мой взгляд неимоверно трудно, гораздо труднее, чем в математике. Но ведь в моей статье я нигде ни разу не называю математику лучшей из наук! Напротив, я несколько раз со всею скромностью подчеркиваю, что как орудие воспитания математика прежде всего отмечена такою особенностью, значение которой для данной цели очевидным образом отрицательно,— она абстрактна, предметом ее служат не сами вещи и явления реального мира, а лишь абстрагированные от них количественные отношения и пространственные формы. Это обстоятельство, как я несколько раз подчеркиваю в своей статье, делает для математики воспитательную задачу значительно труднее, чем для других школьных дисциплин. Но зато математика в некоторых других отношениях отмечена такими чертами, которые создают ей воспитательные возможности более значительные, чем у этих других дисциплин. И то, что я в своей работе, соответственно моей задаче, сосредоточиваю внимание читателя именно на этих чертах, никак не может, конечно, означать какого-либо гипостазирования математики, превознесения ее выше всех других наук.

Я отдаю математической науке лишь то, что ей принадлежит по праву, с полной откровенностью признавая и те ее черты, которые в отношении к данной цели составляют ее слабость. Но за те преимущества, которые я ей приписываю, я уж действительно готов драться до конца. И если уважаемые коллеги пожелают утверждать, что та или другая конкретная черта, по моему утверждению монопольно или преимущественно присущая математике, на самом деле является в такой же мере достоянием и всех других научных дисциплин, то здесь я где угодно и когда угодно готов держать ответ в полной уверенности, что смогу отстоять правильность моего утверждения перед любым компетентным и беспристрастным судом.

О ТАК НАЗЫВАЕМЫХ «ЗАДАЧАХ НА СООБРАЖЕНИЕ» В КУРСЕ АРИФМЕТИКИ

А. Я. Хинчин

(Москва)

Как-то мне пришлось спросить несколько опытных учителей пятых классов о том, какой примерно процент учащихся действительно научается решать арифметические задачи, не являющиеся простыми вычислительными примерами, т. е. такие, где способ решения, как бы прост он ни был, должен быть найден самим учащимся. Из всех опрошенных мною учителей только один утверждал, что этому искусству удается научить до 15% учащихся; все другие говорили, что лишь отдельные учащиеся овладевают этим искусством, а некоторые даже заявляли, что «этому вообще научить невозможно». Конечно, решив целый ряд совершенно однотипных задач, ученик без труда решит задачу в точности того же типа (этим объясняется отсутствие сплошных провалов на экзаменах и контрольных работах); но добиться, чтобы ученик самостоятельно нашел решение задачи нового, хотя бы и очень простого типа, — это, по единодушному мнению учителей, есть дело, удающееся только в самых исключительных случаях.

Таким образом, как раз то «развитие сообразительности», которое у нас любят выставлять как основную цель введения «трудных задач», оказывается, никак не удается даже у лучших учителей.

Сопоставим с этим другой факт: хорошо известно, что бóльшая часть наших ученых математиков, как правило, становится в тупик перед задачами элементарной арифметики. Я лично охотно признаюсь, что всякий раз, когда ученик пятого класса среди моих знакомых просил меня помочь решить арифметическую задачу, дело это для меня оказывалось весьма тяжелым, а подчас я терпел и полную неудачу. Я, как и большинство моих товарищей, легко решал, конечно, предложенную задачу естественным алгебраическим путем (т. е. составлением уравнения или системы уравнений); но ведь надо было во что бы то ни стало обойтись без алгебраического анализа! Обычно если мне в конце концов удавалось найти такое решение, оно всё

же оставляло меня неудовлетворенным: моя научная совесть неумолимо подсказывала мне, что тут остается какой-то туман, не всё ясно. В результате, как правило, и ученика мое решение не удовлетворяло, и он явно лишь из вежливости принимал его. Иногда в таких случаях я потом пытался узнать, как же объяснил решение задачи учитель? Должен признаться, что и в рассуждениях учителя для меня почти всегда оставался тяжкий элемент ненатуральности и искусственности. Я лично отказался бы преподнести классу такое рассуждение; всякий математик знает, какое мучительное чувство охватывает душу, когда подчас бываешь вынужден преподавать другим что-нибудь в такой форме, которая самого себя не вполне удовлетворяет: речь становится медленной и напряженной, слова выходят не те, какие нужно, и смущение твое быстро передается слушателям. При таких условиях вряд ли приходится удивляться тому, что большинство детей так и не научается решать арифметические задачи.

Описывая всю эту тяжелую ситуацию, я думаю, что не очень сгустил краски. Если в отдельных случаях дети всё же научаются решать задачи, интуитивно отличают правильное рассуждение от ложного, находят в этих упражнениях ума здоровое удовольствие и в конечном счете действительно развивают свою сообразительность, то такие исключения способны только подтвердить печальное общее правило. Кто же (или что же) несет ответственность за такое положение вещей? Ведь это не шутка, когда львиная доля огромного числа часов, уделяемого в пятых классах арифметике, тратится на дело, в отношении большинства учащихся заведомо безнадежное. Кстати, хорошо известно и многократно отмечалось, что, как правило, ни оканчивающие школу, ни студенты педвузов, ни начинающие учителя (ни, прибавим от себя, научные работники) не умеют решать арифметических задач, да и вряд ли на всем свете кто-нибудь умеет решать их, кроме учителей пятых классов.

Прежде чем попытаться ответить на этот вопрос, рассмотрим ряд примеров, анализ которых придаст больший вес нашим выводам.

Отец старше сына на 25 лет. Возраст отца относится к возрасту сына, как $3/2:2/3$. *Сколько лет отцу и сколько лет сыну?*

Сначала решим задачу алгебраически. Обозначим искомые числа через x и y, тогда

$$x-y=25,$$

$$\frac{y}{x}=\frac{2/3}{3/2}=\frac{4}{9},$$

откуда

$$y=\frac{4}{9}x,\quad x-\frac{4}{9}x=25,\quad \frac{5}{9}x=25,\quad x=45,\quad y=20.$$

Теперь попробуем решить задачу «арифметически». Конечно, здесь возможны разные варианты; я выбираю тот, который мне кажется простейшим.

1) Какую долю возраста отца составляет возраст сына?

$$\frac{2/3}{3/2}=\frac{4}{9}\ ^{1}).$$

2) Какую долю возраста отца составляет разность между возрастом отца и возрастом сына?

Так как возраст сына составляет $\frac{4}{9}$ возраста отца, то разность между ними составляет $1-\frac{4}{9}=\frac{5}{9}$ возраста отца.

Я хотел бы обратить внимание на то, что ответить на этот вопрос, а еще более — придумать, поставить этот вопрос есть дело, которое, на мой взгляд, самостоятельной изобретательности рядового ученика пятого класса недоступно. Если же вопрос и ведущее к ответу рассуждение ему будут навязаны, они непременно оставят в нем тяжелое ощущение искусственности, а может быть, даже и софистической подтасовки. Конечно, можно придумать сколько угодно более или менее пространных пояснений; однако, насколько я вижу, все они будут говорить с ребенком всё тем же, искусственно-надуманным, непривычным и не свойственным ему языком и потому оставят в нем то же тяжелое ощущение. В лучшем случае, если вопрос и ответ будут им поняты, у него все же останется недоумение, каким образом он сам мог бы догадаться, что именно этот странный вопрос здесь должен быть поставлен.

3) Сколько лет отцу?

Здесь все уже просто: если $\frac{5}{9}$ возраста отца составляют 25 лет, то этот возраст равен

$$25:\frac{5}{9}=45\ \text{(лет)};$$

разумеется, возраст сына теперь уже находится непосредственно.

Это пример очень легкой задачи. Прослеживая ее решение, мы прежде всего совершенно очевидным образом приходим к следующему

---

[1]) Е. С. Березанская обратила мое внимание на то, что опытный учитель дальше будет решать иначе, чем Александр Яковлевич, используя метод «разбиения на части». Дальнейшие рассуждения будут при этом таковы:

«Таким образом, возраст отца составляет 9 частей, а сына 4 части. Отец старше сына на 5 частей. На эти 5 частей приходится 25 лет. Значит, каждая часть составляет 5 лет. Отсюда — возраст отца $5\times9=45$ лет, возраст сына $5\times4=20$ лет».

По своему существу и этот прием является завуалированным алгебраическим.— *Б. Г.*

выводу: так называемое «арифметическое» решение задачи ничем с логической стороны не отличается от алгебраического. Мы отнюдь не заменяем алгебраического анализа каким-либо другим способом решения, а, напротив, от начала до конца этот анализ проводим, как и в случае алгебраического решения, с той только разницей, что вместо формул пользуемся тяжелыми и неуклюжими словесными формулировками.

Далее мы замечаем, что если с алгебраической точки зрения эта задача есть простой пример на составление системы двух уравнений первой степени с двумя неизвестными — пример, решаемый привычными, механическими приемами, не требующими ни малейшей изобретательности, то с точки зрения «арифметической», т. е. когда мы хотим тот же анализ провести без помощи привычного алгебраического аппарата, задача становится задачею «на соображение», т. е. требует вымысла, догадки, и притом — повторяю — догадки, на мой взгляд, искусственной и чуждой сознанию 12-летнего ребенка.

Рассмотрим теперь несколько более сложную (хотя тоже нетрудную) задачу.

Разделить число 22 *на три части при условии, что если прибавить к одному из полученных чисел* 0,5, *от другого отнять* 1,5, *третье умножить на* 2,5, *то получатся одинаковые результаты.*

Алгебраическое решение. Пусть x, y, z — искомые числа, тогда

$$x+y+z=22, \quad x+0{,}5=y-1{,}5=2{,}5\ z,$$

откуда

$$x=\frac{5z-1}{2}, \quad y=\frac{5z+3}{2},$$

$$\frac{5z-1}{2}+\frac{5z+3}{2}+z=22, \quad 6z+1=22,$$

$$z=3{,}5; \quad x=8{,}25; \quad y=10{,}25.$$

Теперь я снова предлагаю тот вариант «арифметического» решения, который представляется мне наиболее убедительным для ученика. Но здесь я с самого начала наталкиваюсь на трудность, которую я не в силах преодолеть: мне никак не удается найти сколько-нибудь разумной формулировки последовательных вопросов. Я поэтому предоставляю это дело (т. е. формулировку вопросов) более компетентным товарищам, а сам позволю себе обойтись вовсе без вопросов.

1) Так как, умножая третье число на 2,5 и прибавляя к первому числу 0,5, мы получаем одно и тоже число, то

первое число равно 2,5 *третьих чисел без* 0,5. (Я очень хотел бы знать, как товарищи учителя поставили бы вопрос к этому рассуждению; может быть, они сказали бы «чему равно первое число?»)

2) Так как тот же результат мы получили бы, отнимая 1,5 от второго числа, то

второе число равно 2,5 *третьих чисел плюс* 1,5.

3) Таким образом, сложить все три числа, это всё равно, что сложить

2,5 третьих чисел без 0,5;

2,5 третьих чисел да еще 1,5;

и третье число,

что в сумме дает шесть третьих чисел плюс 1. Но эта сумма по условию задачи равна 22. Значит шесть третьих чисел в сумме дают $22-1=21$, откуда третье число равно

$$21:6=3{,}5.$$

Первое и второе числа теперь уже легко находятся, и на этом можно не останавливаться [1]).

Неужели не ясно, что всё это рассуждение не представляет собой никакого своеобразного «арифметического метода», но есть дословный перевод вышеприведенного алгебраического решения с языка формул на язык слов? И неужели не видно, что этот последний язык в данном случае неуклюж, неточен, мало выразителен, производит впечатление надуманности, искусственности и что просто бессмысленно требовать от рядового ученика пятого класса, чтобы он сам всё это выдумал? Может быть я решил не так, как нужно? Пусть тогда мне укажут более естественное «арифметическое» рассуждение.

Я не подбирал примеров искусственно, а взял два первых попавшихся; положительно утверждаю, что почти все арифметические задачи, выходящие за пределы просто вычислительных примеров, носят тот же характер; это сплошь алгебраические задачи на составление уравнений и систем уравнений первой степени. Конечно, если это угодно, то можно всегда, ценою весьма неприятной искусственности и значительного затемнения метода, весь необходимый алгебраический анализ задачи провести словесно, без формул и буквенных обозначений; как всё это выглядит на практике, мы только

[1]) Е. С. Березанская предлагает иной ход рассуждений: по условию задачи первые два слагаемые дадут один и тот же результат (который мы назовем «паем», если к первому из них прибавить 0,5, а от другого отнять 1,5. Третье слагаемое по условию в 2,5 раза меньше одного пая и, значит, равно 2/5 пая. Таким образом,

1 пай + 1 пай + 2/5 пая = 2 2/5 пая;

эта сумма на единицу меньше, чем сумма первоначальных чисел, т. е. равна 21. Отсюда — один пай равен 8,75. Искомые числа, следовательно, равны $8{,}75-0{,}5=8{,}25$, $8{,}75+1{,}5=10{,}25$ и $8{,}75\cdot 2/5=3{,}5$.

Нет нужды говорить, что в этом несколько более простом рассуждении фактически также вводится неизвестное — пай. Весь ход рассуждений отличается от алгебраического только тем, что вместо символических рассуждений мы пользуемся словесными.—*Б. Г.*

что видели. Надеюсь, что я не одинок в резком чувстве отвращения к подобного рода «арифметическим» решениям.

Для чего всё это нужно? Какую «сообразительность», какие вообще ценные способности ума можно развить в ребенке, заставляя его проделывать такие противоестественные, инстинктивно отталкивающие его упражнения? В седьмом классе на уроках алгебры он научится решать те же задачи легко, естественно, почти механически. Не похоже ли это на то, как если бы солдата в течение первого года службы заставляли овладевать ружьями, скажем, допетровской Руси, а только потом дали ему в руки винтовку современного образца?

Много и настойчиво говорят у нас о трудностях, связанных с преподаванием арифметики в пятых классах. Однако при анализе такого рода жалоб почти всегда оказывается, что собственная задача арифметики — изучение чисел и действий над ними — никогда непреодолимых трудностей не вызывает; что остается подлинно непосильной задачей, это — научить детей решать задачи «на соображение», т. е. самостоятельно изобретать способ решения задач, не являющихся простыми примерами применения арифметических действий. Мы только что видели специфический характер этого рода задач и пытались выяснить всю естественность, даже неизбежность того, что решать их не могут не только учащиеся пятых классов, но и оканчивающие среднюю школу, студенты педвузов, а подчас и сами учителя (особенно начинающие). Надо тут же прибавить, что в старых задачниках попадались и «трудные» задачи иного рода (хотя и в значительно меньшем числе), где для решения требовался не элементарный алгебраический анализ, а изобретение специального, подчас весьма хитроумного рассуждения — изобретение, требующее специального конструкторского дарования, являющееся уже актом логического творчества; некоторые научные работники настаивали на включение в курс пятого класса именно такого рода задач, и притом даже в возможно большем числе. Не думаю, чтобы такое требование было целесообразно; массовая школа никак не может в своем основном преподавании ориентироваться на возможности и нужды будущих специалистов-математиков; требующее специальной одаренности научное творчество не может входить в круг ее основной, обязательной для всех учеников работы в качестве сколько-нибудь значительного элемента; в лучшем случае такого рода задачи могли бы входить в качестве факультативного элемента для желающих или для работы научного кружка; делать же ставку на овладение головоломными задачами всей основной массой учащихся значит, на мой взгляд, сознательно заниматься иллюзиями. Особенно досадно, когда сознаешь, что впустую потраченное на эти безнадежные попытки время могло бы быть с огромной выгодой использовано для лучшего овладения процентными вычислениями, действиями над десятичными дробями и т. п.; что неокрепшему еще в производстве этих элементарных операций детскому сознанию, вместо того чтобы по возможности расширить круг соответствующих упражнений, навязываются задачи-головоломки, как правило, ему недоступные и его бесцельно изнуряющие. Ни в одной другой школьной дисциплине ничего подобного не практикуется, и по вполне понятной причине: сколько бы мы ни говорили о «развитии сообразительности», остается несомненным фактом, что всякая попытка стимулировать научное творчество в области, которой учащийся сам еще твердо и полностью не овладел, способна вызвать либо верхоглядство, либо разочарование и недоверие к науке и ее преподаванию.

Итак, резюмируем. Те задачи в курсе арифметики пятого класса, которые называют «задачами на соображение» и на культивировании которых настаивают некоторые представители нашей математической общественности, представляют собой в подавляющем большинстве случаев алгебраические задачи на составление уравнений, которые лишь весьма искусственным путем могут быть решаемы без помощи алгебраического анализа, в некоторых же случаях — задачи-головоломки, требующие для своего решения прямого логического изобретательства и, значит, соответствующих специальных способностей.

По причинам, которые указаны выше, я считаю, что оба эти типа задач должны быть исключены из основного материала арифметики пятого класса. Программа этого курса достаточно напряженная, она богата новыми понятиями и арифметическими операциями. Надо, как всегда, ухватиться за главное звено и всё внимание направить на отнюдь не легкую задачу прочного усвоения этих понятий и этих операций, не отвлекая силы на чужеродные, второстепенные и — главное — вряд ли достижимые цели.

Те, кто настаивает на этого рода задачах, презрительно называют наши задачники «примерниками»; я предложил бы не пугаться этой клички; во всех школьных дисциплинах основною целью упражнений считается приобретение навыков в практическом применении усвоенного теоретического материала, и в этом не видно ничего зазорного; я не видел бы ничего зазорного и в том, чтобы арифметика следовала этому разумному и естественному обычаю.

Те, кто пропагандирует этого рода задачи, настаивают на том, что без них курс арифметики скучен и не занимателен. Для меня же непонятно, почему ребенку заниматься изобретением противоестественных, чуждых привычному ходу его мыслей рассуждений, вроде приведенных выше, или часами ломать голову над изобретением недоступного ему в большинстве случаев хитроумного приема есть дело более интересное, чем, скажем, могущее легко принять даже характер спорта соревнование в быстроте и безошибочности операций над дробями или процентных вычислений. Я боюсь, что «скучно» это не детям, а тем математикам, которые всем этим давно

и прочно овладели и для которых поэтому такие упражнения действительно носят характер чисто механического действия. Лично я много раз бывал свидетелем того, как ученик пятого класса, проделав ряд сложных вычислений и сверив свой результат с ответом, испытывал глубокое удовлетворение, и как тот же ученик, бесцельно продумав час над решением алгебраической задачи без помощи алгебры, с тупым и весьма пессимистическим равнодушием записывал потом это решение с чужих слов.

Будем помнить, что всякая арифметика, в том числе и школьная, есть всё-таки прежде всего учение о числах и операциях над ними. Дело это важное, ответственное и служит основою всего дальнейшего математического (и не только математического) образования. И при выполнении этого дела устремим все наши усилия и все усилия наших учеников в этом основном направлении, не уклоняясь от него и не подменивая его как угодно соблазнительными лозунгами вроде «развития сообразительности».

I. ОБЗОРЫ И СТАТЬИ

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ЛИНГВИСТИКЕ [1])

Р. Л. Добрушин

(Москва)

Ученые почти всегда вдохновляются в своей деятельности внутринаучными потребностями — такое положение является необходимым условием плодотворного развития науки. При этом иногда забывают о внешних приложениях науки (к ним относятся не только непосредственные ее применения к практической деятельности, но и косвенные приложения данной науки через посредство других наук). Научные работы по этим приложениям часто выглядят на фоне настоящих теоретических работ кустарными — они значительно меньше используют глубокий аппарат, уже созданный наукой. Отсюда, естественно, возникает отношение к таким работам по приложениям, как, может быть, нужным и полезным, но не столь важным для самой науки. Однако именно эти связи науки с внешним миром являются единственным оправданием самого существования науки и поэтому определяют в конечном итоге весь ход ее развития.

Это — общие очевидные истины, которые особенно актуальны в применении к современной лингвистике. Можно указать на два основных применения лингвистики: первое — приложения лингвистики к вопросам преподавания языка (родного, иностранного, литературного и т. д.; отсюда в конечном счете выросли и грамматика, и фонетика, и стилистика), и второе — приложения лингвистики к вопросам истории [2]). Эти приложения определили стиль и метод лингвистики — она не использует глубокого, если сравнить с другими науками, специфического аппарата — ни экспериментального, ни математического [3]). Этим также объясняется тот факт, что в целом

[1]) Статья представляет собой обработку докладов, читанных на первой всесоюзной конференции по машинному переводу в мае 1958 г. Основой этих докладов послужила в свою очередь лекция, прочитанная школьникам VIII—X классов в январе 1958 г. [Примечание при корректуре: статья была сдана в печать в июне 1958 г. и поэтому не отражает развития математической лингвистики за последние два с половиной года.]

[2]) Можно еще к этому добавить третью, также важную функцию лингвистики — ее приложения к вопросам психологии. Но эта функция не получила должного развития и определяющего влияния на развитие науки не имела.

[3]) Например, фонетика, по существу, вся посвящена изучению звука.

лингвистика довольствовалась до сих пор довольно туманными формулировками основных понятий своей науки.

Таково было положение лингвистики до самого последнего времени. Область приложения лингвистики была стабильна в течение столетий: приложения ее определились очень давно, и с тех пор новых приложений не появлялось. Этим можно объяснить относительную медленность развития лингвистики: конечно, в этой науке появлялись замечательные работы и новые идеи, но в ней не было открытий революционного значения, сыгравших определяющую роль в развитии человечества. Здесь нечего сравнить с открытием атомной энергии, с изобретением дифференциального и интегрального исчисления, с появлением понятий и представлений современной генетики, с переворотом, который произвели в общественных науках идеи Маркса [1]).

Но за последнее десятилетие положение коренным образом изменилось. Появились новые приложения лингвистики, настолько важные, что их можно сравнить с теми, которые определили ее предыдущее развитие. Они возникли потому, что теперь с языком приходится иметь дело не только человеку, но и машине, и, конечно, требования, которые предъявляют к лингвистике машины и люди, совершенно разные.

Укажем важнейшие из этих приложений. Это, во-первых, *машинный перевод с одного языка на другой* [2]) и примыкающая сюда проблема машинного хранения информации, заданной первоначально в языковой форме (проблема создания информационной машины) [3]). Тесная связь этих двух задач объясняется тем, что, по существу, во второй из них идет речь о переводе с человеческого языка на специально созданный машинный язык.

Во-вторых, это *передача информации, заданной в форме письменного и устного языка* [4]). Этот вопрос не нов: и раньше в письмах, телеграммах и по телефону информация передавалась в языковой форме. Однако быстрое усовершенствование техники связи, рост потребностей в передаче информации, «кризис эфира», в котором «не умещается» информация, передаваемая в форме электромагнитных волн,—

с точки зрения создания его ротовым аппаратом. Это произошло потому, что фонетисты ориентировались на приложение фонетики к обучению человека произношению звуков чужого языка. Однако есть и другие, не менее важные аспекты звука, например физическая сторона его: звук как колебание воздуха. Этим лингвисты до сих пор почти не интересовались.

[1]) Может быть единственными достижениями сравнимого масштаба были в лингвистике общие идеи о том, что слова поддаются грамматической классификации и что языки тоже поддаются классификации по степеням родства. Но эти идеи появились весьма давно.

[2]) Литературу о машинном переводе см. в конце статьи (стр. 59).

[3]) См. В. А. Успенский, К проблеме построения машинного языка для информационной машины, Сборник «Проблемы кибернетики», вып. 2, М., 1959, стр. 39—51.

[4]) См. книгу C. Cherry (см. стр. 59 настоящего выпуска).

ней это поставило очень остро проблему создания более экономных методов передачи информации. С подобными вопросами тесно связана возникшая очень недавно и бурно развивающаяся математическая наука — теория информации [1]); она позволяет четко поставить некоторые вопросы о свойствах речи, например вопрос о вычислении энтропии письменной речи [2]), и т. п. Ответы на эти вопросы в основном еще не получены.

В-третьих, это *машинный перевод устной речи в письменную и письменной в устную*. Речь идет о машинах, синтезирующих звуковую речь и, наоборот, анализирующих ее и переводящих в форму письменной (проблема ввода в машину и вывода из нее письменной речи не является принципиально трудной).

Это проблема совершенно новая; перспективы, открываемые ее решением, очевидны. Ясно также, что ее разработка тесно связана с изучением свойств языка.

Если вопрос о передаче информации не нов (и большая актуальность его в наши дни объясняется лишь необходимостью резко увеличить количество передаваемой информации), то машинный перевод и машины, синтезирующие и анализирующие речь, еще несколько десятков лет тому назад относились к области научной фантастики. Но сейчас это — весьма реальные и актуальные научные проблемы.

Новые приложения требуют от лингвистики введения новых методов — метода *математической абстракции* и метода *физического экспериментального изучения языка*. Роль каждого из этих методов будет, конечно, различной в разных разделах новой лингвистики. Так, например, в актуальной и мало разработанной проблеме установления физических характеристик звуковой речи нужна большая экспериментальная работа, а применения математики, по крайней мере в ближайшие годы, будут нужны лишь в той мере, в какой математика всегда применяется в описательных разделах науки, т. е. там, где знание предмета еще недостаточно для создания обобщающей теории.

Однако для многих других разделов лингвистики применение математики — задача актуальная. В первую очередь это относится к тем случаям, когда с языком приходится иметь дело машине. Грубо говоря, математические методы оказываются здесь необходимыми из-за того, что современные машины слишком примитивны, чтобы непосредственно «понимать» язык лингвистов, которые сами не всегда понимают друг друга, а если и понимают, то далеко не однозначно. Машина не может опираться на интуицию. С ней можно «говорить» лишь простым, четким и однозначным языком, а такой язык может

---

[1]) См., например, А. М. Яглом и И. М. Яглом, Вероятность и информация, изд. 2-е, М., 1960.

[2]) См. там же.

быть лишь у математики, которая, по существу, много проще лингвистики. Ведь математика отбирает самые простые и самые ясные понятия с тем, чтобы вывести из них нужные следствия. Поэтому приложения лингвистики к использованию машин невозможны без математической переформулировки понятий лингвистики.

С другой стороны, законы лингвистики в основном достаточно просты для того, чтобы допустить математическое описание,— это доказывает уже сама принципиальная возможность осуществления машинного перевода. Автоматизация процесса, не допускающего формального математического описания, невозможна, поэтому практически доказанная возможность машинного перевода доказывает также возможность математически сформулировать законы языка.

В лингвистических приложениях математики речь еще не идет о создании какой-либо глубокой и развитой математической теории. Нельзя сравнивать термин «математическая лингвистика» с аналогичным термином «математическая физика». Математическая физика — это особый раздел математики, нацеленный на специфические физические приложения; по своим методам он не менее сложен, чем любой другой раздел математики. В лингвистике же речь должна идти о первых шагах применения математики. Уровень предстоящей на первых порах математизации лингвистики можно сравнить с применением математики в биологии, химии и т. п., т. е. с применениями с точки зрения математики не очень глубокими. Но в то же время лингвистические применения математики имеют существенную специфику, отличающую их от большинства иных применений математики. Эта специфика связана с тем, что большинство явлений лингвистики имеет по существу своему дискретный характер, и значит, здесь нужно применять в первую очередь методы дискретной математики. Эти методы не сложнее методов непрерывной математики, т. е. методов дифференциального и интегрального исчисления, но они значительно менее популярны, чем методы математического анализа.

Каковы же конкретные области приложения математических методов в лингвистике? Мы будем говорить лишь об областях уже сложившихся, для которых ясно видны их большие перспективы.

Во-первых, это приложения математических методов к формальному *математическому описанию понятий грамматики* [1]).

Здесь речь идет о том, чтобы, введя некоторую систему аксиом, проанализировать и пересмотреть с точки зрения этих аксиом всю логическую структуру грамматики. По-видимому, окончательные выводы из такого пересмотра будут не слишком сильно отличаться от

[1]) См. книги J. Greenberg и J. Chomsky (стр. 59 настоящего выпуска).

того, что является теперь общепринятым в лингвистике: поколения лингвистов работали, конечно, не зря. Но в то же время использование математических методов даст возможность формулировать факты грамматики четко, просто и ясно, и отличие от классических формулировок будет существенным.

В последние годы делаются первые попытки математически формулировать основы грамматики [1]). Они носят еще характер первоначальных опытов, и дальнейшие исследования приведут безусловно к более удачным концепциям. В частности, основной недостаток сделанных построений состоит, по нашему мнению, в том, что они относятся к фиксированному языку, не касаясь его отношений с другими языками или с реальностью, которую язык отражает. В то же время классическая грамматика описывает не только формальную структуру самих языков, но и их отношение к отображаемой языком реальности (скажем, такой характер носят классические, хотя и очень туманные, утверждения о том, что «существительные — это наименования предметов, а глаголы — наименования действий» и т. п.). Но сама реальность по своей сложности с трудом поддается формализации; поэтому на первых порах более содержательными нам кажутся попытки дать математическое описание связей между структурами различных языков [2]).

Создание математической грамматики будет иметь очень большое значение для усовершенствования машинного перевода и для создания информационных машин. В связи с этим хочется решительно возразить против высказываемого иногда мнения о том, что при построении алгоритмов машинного перевода можно обойтись традиционными лингвистическими понятиями и методами, только несколько подлатав их и подштопав. Сделаем по этому поводу одно сравнение. Теперь уже никто не сомневается в очень большом значении математических методов для техники. Тем не менее и сейчас имеется много инженеров, часто талантливых и работающих очень продуктивно, которые даже с некоторой гордостью говорят о том, что из математики они используют в своей работе только логарифмическую линейку. Они забывают, что, во-первых, формулы, по которым они считают на логарифмической линейке, созданы с серьезным участием математики и, во-вторых, подобное пренебрежение к теории возможно лишь до поры до времени, в относительно устоявшихся разделах

[1]) Одна из таких попыток описана в приложении к этой статье (стр. 52—59).

[2]) В частности, именно на этом пути, по-видимому, можно четко понять различие между словарем и грамматикой (т. е., грубо говоря, ответить на вопрос, почему считают, что «волк» и «волку» — грамматические формы слова, а «волк» и «волчица» — разные слова). Это различие имеет не качественный, а лишь количественный характер. Естественно думать, что грамматика — это совокупность приемов для более простого описания алгоритма перевода с одного языка на другой (в частности, с языка реальности на язык речи).

техники. Аналогично обстоит дело и с машинным переводом. Как и во всякой практической деятельности, основанной не на теории, а на эмпирике (попытка исходить из неадекватной теории — это та же эмпирика), успех может быть лишь временным. В начальной стадии исследования такой путь может показаться даже более экономным: ведь разработка теории требует сил и времени, а эффект приходит не сразу. Но в конечном счете победа всегда остается за практикой, вооруженной адекватной теорией.

Вторая уже успешно разрабатываемая область применения математических методов в лингвистике — это *приложения теории информации*. Возникшая в связи с потребностями техники связи теория информации сразу поставила вопрос о нахождении изучаемых в этой теории статистических характеристик (энтропия, избыточность) для реальных языков. Актуальность лингвистических вопросов теории информации можно проиллюстрировать, например, тем фактом, что в 1957 г. в США создан новый научный журнал «Information and Control», в редакцию которого, наряду с основоположником теории информации Клодом Шенноном и основоположником кибернетики Норбертом Винером, входит известный американский лингвист Роман Якобсон. В этом журнале, посвященном широкому комплексу вопросов, связанных с теорией информации, опубликовано уже несколько интересных работ лингвистического содержания; в частности, там описаны эксперименты, проводившиеся для вычисления избыточности английской письменной речи. Для русской речи подобные эксперименты еще не проводились [1]).

Совсем неразработанным остается вопрос о вычислении информационных характеристик устной речи — вопрос, очень важный для технических проблем компрессии речи (т. е. более экономых методов ее передачи). Видимо, разумная постановка задачи связана здесь с введенным А. Н. Колмогоровым понятием энтропии при заданных условиях точности воспроизведения. Такие условия должны учитывать особенности слухового аппарата человека.

Наконец, третий важный раздел применения математики в лингвистике — это вопросы *лингвистической статистики* [2]). Вопрос о статистических методах в лингвистике совсем не нов — новые приложения только подчеркнули большую актуальность применения статистических методов. В самом деле, почти любая лингвистическая работа использует (хотя часто и помимо сознания автора) статистические методы потому, что в языке есть явления редкие и частые, и без выделения частых явлений невозможны какие-либо лингвистические выводы. Необходимость в количественных оценках частот лингвистических явлений не меньшая, чем необходимость количественного подхода в других естественных науках или в экономике. Язык цифр, которого, к сожалению, лингвисты еще боятся, может не только иллюстрировать выводы, но сделать их значительно более доказательными. Приведем простейший (нарочито угрубленный) пример, относящийся к статистике. Можно сказать: «А. Н. Толстой предпочитает более длинные фразы, а А. И. Куприн — более короткие». А можно сказать и так: «среднее число слов в фразе в произведении Толстого „Сестры“ равно 11,9, а среднее число слов в фразе в произведении Куприна „Поединок“ — 9,5» [1]). Разница будет примерно такая же, как если в одном случае ограничиться утверждением, что производство угля в Советском Союзе больше, чем в Англии, а в другом — привести конкретные цифры. Каждому ясна бо́льшая доказательность утверждения во втором случае. Но лингвисты обычно идут по первому пути.

Нам кажется, что простейшие графики и таблицы станут скоро главным методом получения лингвистических выводов [2]). Необходимость использования количественных методов постепенно осознается лингвистами: в 1956 г. появилась первая сводная монография G. Herdan по статистическим методам в лингвистике [3]) (впрочем, написанная не безупречным образом).

Для лингвистов важны не только простейшие вычисления средних и частот, но и более сложные вопросы об оценке случайных расхождений вычисленных частот и средних с истинными лингвистическими параметрами. А это требует уже более глубокого использования методов математической статистики. В качестве примера приведем график, изображенный на стр. 44 [4]). Исследовались частоты появления фраз с тем или иным числом слов в романе А. Н. Толстого «Сестры». Сплошная жирная линия указывает зависимость этой частоты от длины m (т. е. количества слов) фразы. На фоне плавного в целом поведения кривой бросается в глаза минимум в точке $m = 6$, который трудно было бы объяснить из качественных соображений. Положение проясняют нарисованные выше и ниже кривой линии так называемые «доверительные интервалы». Эти линии наглядно показывают пределы, в которых может варьироваться частота из-за случайностей в выборе фраз для подсчета частоты. Об истинной вероятности (определяемой в нашем случае, как частота появления фраз длины m во всем произведении) можно на основе наших подсчетов сказать

[1]) Примечание при корректуре: в 1960—61 годах такие эксперименты в обновленной методике были проведены А. Н. Колмогоровым.

[2]) Литературу см. на стр. 59.

[1]) Эти цифры получены в 1959 г. в математическом практикуме для специалистов по теории вероятностей в Московском университете.

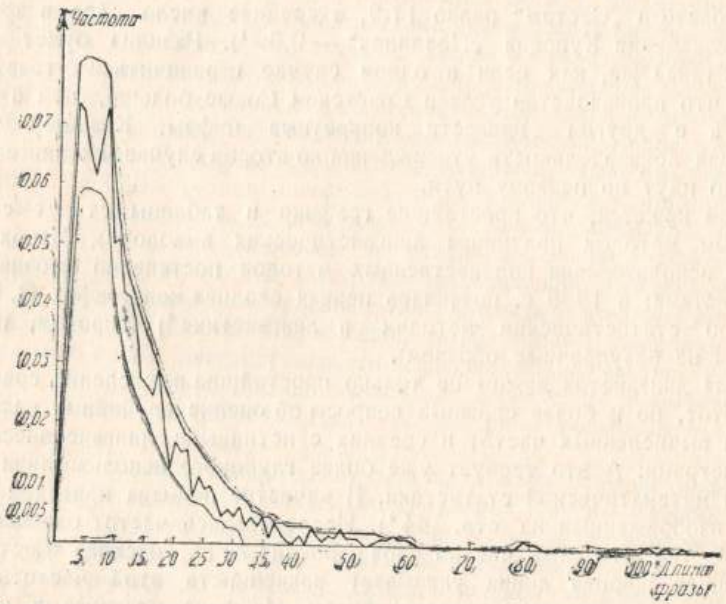
[2]) Прекрасным примером четкого применения статистических идей во вполне классической проблематике является работа А. Мельчука (Бюллетень объединения по проблемам машинного перевода, № 7 — см. стр. 59 наст. выпуска).

[3]) См. стр. 59.

[4]) Получен в том же математическом практикуме.

лишь то, что она лежит между этими линиями, и, таким образом, у нас нет никаких оснований говорить о том, что вероятность имеет минимум в точке $m = 6$. Не рассматривая доверительных интервалов, мы могли бы впасть в ошибку.

Лингвистическая статистика применяется пока лишь для количественного описания лингвистических явлений. Как и в любой науке, следующий, более сложный шаг должен состоять в построении теорий, позволяющих предсказывать количественно те или иные закономер-



ности, получать «уравнения лингвистики». Укажем, например, на следующий факт. Из известных положений теории вероятностей следует, что *число появлений заданного слова в тексте фиксированной длины должно иметь распределение Пуассона* [1]). Это было проверено экспериментально; получилось достаточно хорошее совпадение теории с опытом. Это обстоятельство может быть использовано при составлении частотных словарей языка. Однако для более сложных случаев создание таких теорий — дело будущего.

Очень остро стоит вопрос о создании фундаментальных статистико-лингвистических справочников, каталогизирующих языковый материал с количественной точки зрения. Такие справочники, содержащие частоты появления слов и тех или иных грамматических явле-

[1]) Т. е. вероятность того, что число появлений равно n, равняется $\frac{\lambda^n e^{-\lambda}}{n!}$, где λ — среднее число появлений этого слова.

ний, нужны не только специалистам по машинному переводу или теории информации, но и каждому лингвисту, составляющему учебник языка. Невнимание к частотным характеристикам языка (идущее от пренебрежения ко всему тому, что связано с математикой) приводит к тому, что многие элементарные учебники иностранного языка содержат на первых страницах очень редкие слова и не содержат широко распространенных и т. п. Частотные словари языка изданы в настоящее время для всех основных мировых языков, кроме русского [1]).

Существующие частотные словари языка далеки от совершенства: в них приводятся частоты слов, которые подвержены случайным отклонениям в зависимости от выбранного для подсчета материала, и не делается никаких оценок для их вероятностей, в то время как лингвистический смысл имеют именно вероятности, а не частоты.

Вероятность того или иного лингвистического явления приобретает четкий смысл лишь после того, как задан какой-либо *статистико-лингвистический ансамбль* — отсутствие четкого выбора этого ансамбля приводит к путанице. Рассмотрим этот вопрос подробнее.

Лингвистический ансамбль письменной речи удобно отождествлять с некоторым фиксированным множеством книг. Таким ансамблем может быть, например, множество изданных в 1958 г. газет на русском языке («ансамбль русской газетной речи 1958 г.»), некоторое конкретное издание собрания сочинений А. С. Пушкина («ансамбль языка Пушкина»), вся литература, изданная на русском языке с начала книгопечатания («ансамбль русской печатной речи»). Иногда необходимость фиксировать тот или иной ансамбль кажется лингвисту неприятной, так как тем самым он ограничивает область своих исследований, однако надо понимать, что в любом статистическом исследовании ансамбль всё равно фиксируется, но только чаще неявным, а потому совсем неразумным способом. После того как ансамбль фиксирован, для нахождения вероятности появления некоторого определенного варианта текста [2]) надо перечитать все книги нашего ансамбля и, сосчитав общее число текстов в них и число текстов заданного варианта, разделить второе число на первое. Это и будет, по определению, вероятностью появления заданного варианта текста в заданном статистическом ансамбле. Например, вероятность появления в современной газете текста «социализм» больше нуля, а в про-

[1]) Для русского языка существует небольшой словарь, изданный в США на основе произведения А. С. Пушкина «Капитанская дочка».

[2]) Например, при составлении частотного словаря под заданным вариантом текста надо понимать слово, вероятность которого мы хотим найти, а под общим числом текстов — общее число всех слов в ансамбле. Если мы хотим найти вероятность появления некоторого набора букв, например «ОГО», то под общим числом текстов надо понимать общее число наборов из трех последовательных букв в ансамбле и т. п.

изведениях Пушкина — нуль; следовательно, вероятность появления данного текста может существенно меняться в зависимости от того, по отношению к какому статистическому ансамблю мы его рассматриваем.

Заметим, что практически подсчитать вероятность появления данного текста так, как указано выше, нереально для любого достаточно большого ансамбля. Однако нам на помощь приходит так называемый *закон больших чисел*. А именно: выберем случайным образом из нашего ансамбля n текстов. Пусть интересующий нас вариант текста встретится среди этих n текстов m раз. Тогда, как утверждает закон больших чисел, при достаточно большом n отношение $\frac{m}{n}$ будет близко к искомой вероятности. Число n называют *объемом выборки*. Методы математической статистики позволяют оценить, сколь большим должен быть объем выборки для того, чтобы вероятность равнялась $\frac{m}{n}$ с заданной точностью.

Мы не будем здесь описывать эти методы, отослав читателя к специальной литературе [1]); отметим лишь, что требующийся для этого объем выборки n будет, конечно, увеличиваться с увеличением длины текста и усилением требований точности. В то же время объем выборки не зависит от того, как велик рассматриваемый нами статистический ансамбль. Поэтому правильно используемый выборочный метод позволяет определить вероятность того или иного текста (да и вообще любую лингвистическую вероятность) для любого статистического ансамбля. К сожалению, в лингвистической статистике используются в большинстве случаев лишь относительно узкие статистические ансамбли (например, определенная книга определенного автора), а подсчет производится сплошной. В результате выводы оказываются лишь частными, и при этом проделывается излишняя работа. Использование выборочного метода имеет поэтому серьезные преимущества, однако оно требует статистической обработки результатов подсчетов.

Для многих практических задач являются существенными лишь отдельные стороны такого сложного явления, каким является язык. Например, для проблем теории связи существенно лишь то, какова статистическая структура подлежащего передаче текста, и несущественен его смысл, а также несущественно, имеет ли он смысл вообще. В связи с этим встает вопрос о создании относительно простых моделей, адекватно отображающих с точки зрения каких-

[1]) См., например, A. Hald, Statistical theory with Engeneering application, N. Y. — London, 1952; русский перевод: А. Хальд, Математическая статистика с техническими приложениями, ИЛ, М., 1956.

либо заданных потребностей структуру языка. Простая модель для письменного языка была предложена Клодом Шенноном и проиллюстрирована на примере английского письменного языка. На языке современной теории вероятностей можно сказать, что Шеннон предложил моделировать язык в виде k-сложной стационарной цепи Маркова с теми же $(k+1)$-мерными распределениями, что и в реальном языке. Для читателя, не владеющего теорией вероятностей, эти слова требуют расшифровки — это делается ниже вместе с иллюстрациями из русского языка.

Пусть нам задан некоторый статистический лингвистический ансамбль, который мы будем представлять себе как некоторую совокупность книг [1]). Произведем мысленно следующий опыт. Расчленим все эти книги на отдельные буквы; будем считать за равноправную букву промежуток между словами (который будем обозначать точкой •), так что у нас русский алфавит будет содержать не 34, а 35 букв (существованием знаков препинания мы для простоты будем пренебрегать). Затем, сложив их все вместе и перетасовав, будем вытаскивать буквы одну за другой. В результате получится некоторый текст, вроде следующего:

Я•ЦЫНЬ•ВЕРС•ГНОД•БЕЗ•ЕШВЕШУ•ЬУДЕШЛОЛЙК (1)

Пока получилось нечто совсем не похожее на русский язык. Конечно, для того чтобы получить текст (1), мы на самом деле в нашем эксперименте выбирали случайно одну из книг ансамбля, затем выбирали в ней случайно страницу и, наконец, на этой странице также случайно выбирали одну из букв; затем мы вновь и вновь повторяли тот же процесс. Наглядно ясно (это может быть и строго обосновано), что описанная процедура будет приводить к последовательности букв с теми же статистическими свойствами, как текст (1), полученный первоначальной процедурой. На языке современной теории вероятностей можно сказать, что мы рассмотрели последовательность *независимых* испытаний («цепь Маркова порядка нуль»).

Представим себе далее, что тексты наших книг расчленены на пары букв [2]). Перетасовав все эти пары, извлечем одну из них (например, КА). Затем рассортируем все оставшиеся пары, включив в каждую группу все пары, начинающиеся с одной и той же буквы; мы получим 35 групп. Выберем группу, которая содержит пары, начинающиеся с буквы, на которую кончалась извлеченная пара

---

[1]) В приводимых ниже примерах в качестве статистического ансамбля был взят последний, 5-й том собрания сочинения А. И. Бунина (М., 1956).

[2]) Например, текст «Мой дядя самых честных правил» разбивается на следующие пары:

МО, Й·, ДЯ, ДЯ, ·С, АМ, ЫХ, ·Ч, ЕС, ТН, ЫХ, ·П, РА, ВИ, Л·

(в нашем случае группу пар, начинающихся на А), и извлечем из нее новую пару; пусть, например, извлечена пара АЧ.

Из соответствующей группы (в нашем случае из группы пар, начинающихся на Ч) извлечем некоторую пару и т. д. В результате возникает текст такого вида:

КАЧК•ВСВАННЫЙ•РОСЯ•НЫХ•КОВКРОВ, (2)

который уже несколько напоминает русский. Для реального составления текста (2) мы, взяв случайно одну из книг нашего ансамбля, выбрали в ней также случайно некоторую пару соседних букв (КА). Затем снова открывали некоторую другую книгу в некотором случайном месте и читали ее, начиная с этого места, пока впервые не встретили последнюю букву ранее полученного текста (А). Встретив эту букву, мы включали в текст следующую после нее букву. (Например, если впервые буква А встретилась в слове *врач*, то мы писали в нашем тексте: Ч.) Далее процедура продолжалась аналогичным образом (в нашем примере мы читали текст до тех пор, пока не встретили букву Ч; если она встретилась в слове *точка*, писали после нее К) и т. д. На языке теории вероятностей можно сказать, что мы построили последовательность букв, образующих обычную цепь Маркова (*цепь Маркова порядка* 1).

Обобщая предыдущее построение, разобьем теперь текст на тройки букв [1]). Затем выберем случайно одну из троек (например, ПОК). После этого рассортируем тройки на 35^2 групп троек, начинающихся одинаковыми парами букв. Будем теперь последовательно выбирать тройку, две первые буквы которой совпадают с последними двумя буквами уже написанного текста (так, в нашем примере мы берем группу троек, начинающихся на ОК и извлекаем, например, тройку ОКА; затем берем группу троек, начинающихся на КА, и вытаскиваем тройку КАК и т. п.). В результате получается текст:

ПОКАК•ПОСТИВЛЕННЫЙ•ПОТ•ДУРНОСКАКА•НАКОНЕПНО•
ЗНО•СТВОЛОВИЛ•СЕ•ТВОЙ•ОБНИЛЬ (3)

Для получения этого текста мы фактически использовали следующий метод. Открыв книгу на случайном месте, мы прочитали ПОК; после этого мы читали книгу, пока не встретили сочетание ОК в слове *востока*; написав теперь текст ПОКА, мы искали слог КА. Найдя слово *никак*, мы записали текст ПОКАК, после чего искали пару АК. Встретив слово *чердак*, мы поставили пробел •, после чего искали слог К• (то есть слово, кончающееся на К) и т. д. Описанный процесс дает нам *цепь Маркова порядка* 2.

[1]) Например: МОЙ, •ДЯ, ДЯ•, САМ, ЫХ•, ЧЕС, ТНЫ, Х•П, РАБ, ИЛ•

Сходство полученного текста (3) с русскими текстами бросается в глаза. Оно наглядно доказывает, что цепь Маркова порядка 2 может служить удачной моделью реального русского языка.

Теоретически можно представить себе и дальнейшие все более и более точные приближения к русскому языку: *цепи Маркова порядка* 3, *порядка* 4 и т. д. (т. е. разбиение на четверки, пятерки и т. п.).

Однако если пользоваться описанным выше методом построения текста, то работа по построению дальнейших приближений становится очень трудоемкой. Дело в том, что, как нетрудно показать, для того чтобы найти одну букву в k-й из наших моделей, надо в среднем прочитать текст в 35^{k-1} букв. Для $k=4$ это будет примерно 40 000 букв, т. е. около десятка страниц!

Предложены методы, облегчающие эту работу; правда, они связаны с некоторым искажением теоретической модели. Принцип этих методов напоминает принцип распространенной среди студентов и школьников игры в «балду». А именно [1]), для построения цепи Маркова порядка 3 мы сначала, так же как и раньше, выберем случайно четыре буквы (например, ВЕСЕ). Затем попросим кого-либо придумать слово, содержащее последние три буквы уже написанного текста, не показывая ему весь текст. (Так, в приводимом ниже примере мы просим испытуемого придумать слово, содержащее сочетание ЕСЕ.) Если он назвал слово *веселый*, мы записываем текст ВЕСЕЛ и просим другого испытуемого назвать слово, содержащее текст СЕЛ. Если он назвал слово *присел*, мы записываем текст ВЕСЕЛ• и т. п.

По нашему мнению, этот метод содержит слишком много субъективного произвола. Более удачным кажется следующий метод. Если нужное сочетание букв является относительно распространенным в русском языке, мы находим его так же, как это делалось раньше для цепей Маркова порядка 2. Это сделать практически нетрудно, так как, поскольку это сочетание частое, оно скоро встретится нам в тексте. Если же это сочетание является редким, то можно перечислить после некоторого размышления все слова русского языка, которые это сочетание содержат (ведь их будет немного). После этого мы выбираем случайно одно из этих слов. Если бы мы имели частотный словарь русского языка, то можно было бы достигнуть большей точности, не считая эти слова равновероятными, а приписав им вероятности, пропорциональные вероятностям, указанным в частотном словаре. Таким способом мы на практике получили текст:

ВЕСЕЛ•ВРАТЬСЯ•НЕ•СУХОМ•И•НЕПО (4)

С первого взгляда текст (4) производит впечатление даже более далекого от содержательной речи, чем (3). Причины этого чисто психологические. В тексте (3) большинство слов не было словами русского

[1]) См. E. G. Miller, IRE Trans. on Inform. Theory, 1956.

языка, но так как они напоминали слова русского языка, то читатель невольно, не задумываясь о смысле, воспринимает его как русский текст. В нашем же новом тексте почти все слова действительно являются словами русского языка, и поэтому бросается в глаза его бессмысленность.

В целом рассмотренные выше примеры показывают, что вероятностные модели письменной речи в виде сложных цепей Маркова хорошо описывают письменную речь. Сложнее обстоит дело с получением статистических моделей устной звуковой речи. Получение таких моделей важно для исследований, связанных с техникой связи. Вопрос упирается в недостаточную изученность статистических свойств устной речи [1]).

Остановимся, наконец, на вопросе о том, каковы взаимоотношения между традиционной, «классической» лингвистикой и вновь вторгающимися в лингвистику идеями. Большинство современных лингвистов полагает, что новыми приложениями, новыми задачами и методами могут заниматься математики, техники, физики — все, кто хочет, лишь бы только оставили в покое самих лингвистов и их науку.

Как нужно отнестись к этой точке зрения? Если определять лингвистику как науку об языке, то вне всяких сомнений все описанные выше проблемы относятся к лингвистике. Но это — не главное. Границы между науками создаются независимо от формальных определений науки; наоборот, эти определения возникают и меняются вместе с естественным образом создающимися границами наук. Если, с другой стороны, относить к лингвистике лишь то, чем занимались лингвисты до сих пор и то, чем они могли бы заниматься на основе уже имеющихся знаний, без освоения математики, физики, техники и т. п., то, конечно, при таком узком подходе описанные выше проблемы к лингвистике не относятся.

Обычно ученые какой-либо специальности проявляют энтузиазм по поводу появления новых приложений своей науки, так как это предвещает бурный революционный рост самой науки. В нашем случае произошло не так. Среди лингвистов имеются лишь отдельные горячие приверженцы новых идей; отношение же к этим идеям большинства лингвистов напоминает испуг. В результате новыми областями лингвистики в основном занимаются как побочным делом математики и

[1]) Иногда в прикладных расчетах предполагают, за отсутствием лучшего, что устная речь может быть описана как гауссовский случайный процесс. Несовершенство этого предположения приводит к расхождению теории с практикой. В докладе на Всесоюзной конференции по теории вероятностей и математической статистике в г. Ереване в 1958 г. автор предложил описывать устную речь как кусочно-гауссовский случайный процесс. Этот вопрос требует дальнейших исследований.

техники. Такое положение не может быть стабильным. Кроме непосредственных прикладных проблем, возникает много проблем теоретического характера; поэтому потребность в людях, целиком посвятивших себя теоретическому осмыслению свойств речи с вновь возникших точек зрения, очень велика и еще далеко не удовлетворена. Можно не сомневаться, что уже через несколько лет таких ученых будет немало.

Нет оснований надеяться, что в дальнейшем будут параллельно существовать две науки об языке: классическая лингвистика, использующая лишь качественные методы рассуждения, и новая лингвистика, основанная на глубоком количественном изучении речи и ориентированная на технические приложения. Единство предмета изучения — человеческой речи — приводит к многочисленным связям между новыми и старыми проблемами, относящимися к языку. Мы уже показывали на примерах, как новые методы могут изменить лицо старых областей. С другой стороны, любые математические схемы приобретают значение лишь после трудоемкой работы по наполнению их конкретным лингвистическим материалом. Здесь незаменима деятельность лингвистов, владеющих этим материалом. Подобные связи будут, безусловно, возникать и в будущем. Мы считаем, что трудоемкие исследования в области сравнительного языкознания, связанные с сопоставлением различных языков, могут поддаваться механизации при помощи современной вычислительной техники.

Попытки считать, что существуют две лингвистики или что есть лингвистика первого сорта — классическая, и второго сорта — прикладная, обречены на неудачу, в частности, потому, что техника в наш век развивается очень бурно. Если какая-то наука нужна для важных технических приложений, то можно не сомневаться, что она будет развиваться очень быстро. Поэтому если попытаться делить лингвистику на первосортную и второсортную, то через несколько десятков лет нельзя будет уже понять, какая из них первого сорта и какая — второго.

Мы на пороге создания единой лингвистики, в которую войдут как старые классические методы, так и новые, зарождающиеся. Это бесспорно.

Спорным и открытым является один вопрос: кто будет заниматься созданием этой науки? Примут ли в этом должное участие те, кто занимается сейчас лингвистикой, или их место придется занять людям, пришедшим со стороны?

Приложение

ОПЫТ ОПРЕДЕЛЕНИЯ ПОНЯТИЯ ЭЛЕМЕНТАРНОЙ ГРАММАТИЧЕСКОЙ КАТЕГОРИИ [1])

Здесь нам все время придется иметь дело с математическим понятием *множества*, или, что то же самое, *совокупности*. С этим понятием часто приходится иметь дело и нематематикам. Так, говорят о множестве людей, находящихся в данной комнате, о множестве всех книг в некоторой библиотеке; можно говорить о множестве всех слов, всех треугольников и т. п. Следует заметить, что математическое понимание термина «множества» отличается от житейского толкования этого слова в первую очередь тем, что совокупность объектов, или, как говорят в математике, *элементов*, образующих множество, должна быть точно очерчена. Нужно, чтобы способ задания множества был таков, чтобы о любом объекте можно было с полной определенностью сказать, принадлежит ли он к этому множеству или нет.

Для того чтобы указать, что какой-либо элемент x принадлежит множеству M, мы будем писать $x \in M$. Если M — совокупность всех книг некоторой библиотеки, а x — некоторая книга, то запись $x \in M$ будет означать, что книга x входит в эту библиотеку.

Часто приходится иметь дело с подмножествами какого-либо множества. Например, подмножеством множества всех учеников некоторой школы будут ученики всех седьмых классов, подмножеством множества всех слов современного русского языка будут слова, начинающиеся с буквы «А», или слова (и это уже другое подмножество), состоящие из 5 букв, или (и это уже третье подмножество) все трехсложные слова.

Общее математическое определение подмножества имеет следующий вид.

Определение. Если каждый элемент x множества N принадлежит множеству M, то множество N называется *подмножеством* множества M.

Этот факт записывается так: $N \subset M$, и читается: «N принадлежит M», или, иначе, «N включено в M». Подмножество множества M может иногда состоять только из одного элемента и даже не иметь ни одного элемента (быть пустым) или совпадать со всем множеством M.

Например, множество N всех книг русских авторов некоторой библиотеки есть подмножество множества M всех книг этой библиотеки; этот факт и находит отражение в записи $N \subset M$.

Пусть теперь в множестве M имеется *соотношение эквивалентности*, т. е. установлен некоторый признак, по которому относительно любой пары элементов, принадлежащих множеству M, можно сказать, эквивалентны они или нет (тот факт, что элементы x и y эквивалентны, будем записывать так: $x \sim y$), и пусть эта эквивалентность обладает следующими тремя свойствами:

1°. *Рефлексивность.* Каждый элемент эквивалентен сам себе.

2°. *Симметрия.* Если x эквивалентно y, то y эквивалентно x.

3°. *Транзитивность.* Если x эквивалентно y, а y эквивалентно z, то x эквивалентно z.

[1]) Основная идея наших построений связана с понятием эквивалентности слов, которое было введено независимо от автора настоящей статьи и несколько раньше О. С. Кулагиной в работе «Об одном способе определения грамматических понятий на базе множеств», сборник «Проблемы кибернетики», вып. 1, М., 1959, стр. 203—234. Однако дальнейшая разработка этой идеи проводится нами и О. С. Кулагиной в разных направлениях. Краткое изложение основной части данной работы опубликовано в 1959 г.: Р. Л. Добрушин, Элементарная грамматическая категория, Бюллетень объединения по проблемам машинного перевода, Изд-во 1-го Моск. гос. пед. ин-та иностр. языков.

В качестве примера рассмотрим опять множество M всех книг некоторой библиотеки. Будем считать книгу x эквивалентной книге y, если ее автор тот же, что и у книги y.

Выполнение свойства 1° очевидно. Ясно, далее, что если книга x принадлежит автору книги y, то книга y принадлежит автору книги x; следовательно, выполнено свойство 2°. И, наконец, если книга x принадлежит автору книги y, а книга y принадлежит автору книги z, то книга x принадлежит автору книги z; тем самым показано выполнение свойства 3°.

Мы будем говорить, что множество M разбито на *классы*, если в нем заданы фиксированные подмножества так, что каждый элемент $x \in M$ принадлежит какому-нибудь подмножеству, и только ему одному.

Будем относить, например, к одному подмножеству все слова, имеющие одинаковое число букв. Тогда возникает разбиение всех слов на классы. Ясно, что каждое слово принадлежит какому-нибудь классу и никакое слово не принадлежит двум классам одновременно.

В теории множеств имеет место следующая легко доказываемая теорема.

Всякое отношение эквивалентности, установленное между элементами множества M, удовлетворяющее свойствам рефлексивности, симметрии и транзитивности, определяет разбиение множества M на классы попарно эквивалентных между собой элементов [1]).

В нашем примере с библиотекой все книги, находящиеся в данной библиотеке, разобьются на классы, в каждый класс будут входить книги одного автора. Если в библиотеке имеется единственная книга Грибоедова «Горе от ума», то соответствующий класс будет состоять только из одного этого элемента. Если в библиотеке 3 комплекта полного собрания сочинений Достоевского, то все они войдут в один и тот же класс.

Возьмем другой пример. Пусть множество M составляют все треугольники плоскости. Будем считать эквивалентными подобные треугольники. Ясно, что свойства 1° — 3° будут иметь место, и в силу нашей теоремы всё это множество разобьется на классы подобных между собой треугольников.

В дальнейшем мы будем считать заданным некоторое множество, элементы которого будем называть *словами*. Для того чтобы развиваемая ниже теория приводила к разумным результатам, нужно, чтобы это множество представляло собой совокупность всех слов некоторого реального языка. Например, для русского языка сюда входят слова: маме, колесо, прыгаю, вертишь, чертенком и т. д. Следует заметить, что мы считаем разными слова мама и маме, вертеть и вертишь и т. п. Конечно, при любой попытке перечислить все слова реального языка возникнут спорные вопросы, такие, как вопрос о том, надо ли в эту совокупность включать собственные имена, вульгаризмы или устаревшие слова (так как неясно, какие именно слова считать уже устаревшими, а какие нет). Однако развиваемая теория в равной мере применима при любом способе решения этих спорных вопросов. Надо лишь, чтобы эти вопросы были решены с полной определенностью, хотя классификация, возникающая при применении развиваемой теории, может меняться в зависимости от того, как выбрано основное множество слов.

Фразой мы будем называть совокупность конечного числа слов, расставленных в определенном порядке. Например:

1. Я вчера клевал салфеткой.
2. Колеса вертится быстро.
3. Эта юноша молодая.
4. Колеса светят широко.
5. Я видел мост.
6. Нет собакой они которая.

[1]) Доказательство этой теоремы см., в книге: П. С. Александров, Введение в общую теорию множеств и функций, М.—Л., 1948, стр. 25.

Мы исходим из существования в реальном языке объективно заданного разделения фраз на *грамматически допустимые* и *грамматически недопустимые*, так что о каждой фразе мы можем сделать один из двух взаимно исключающих друг друга выводов: эта фраза допустима или недопустима.

Попытка перечислить грамматически допустимые фразы языка связана с еще бóльшими трудностями, чем попытка перечислить все слова. Однако и в этом случае развиваемая теория применима при любом способе выделения грамматически допустимых фраз, лишь бы он был вполне определен.

Если обратиться к приведенным выше примерам, то естественно считать фразы 1, 4, 5 грамматически допустимыми, а фразы 2, 3, 6 — грамматически недопустимыми. Подчеркнем, что здесь речь идёт именно о грамматической, а не смысловой допустимости. Вряд ли можно было бы построить четкую теорию, опирающуюся на понятие смысловой допустимости фразы, потому что критерий осмысленности является очень неопределенным. Часто для фразы, являющейся на первый взгляд бессмысленной, но грамматически допустимой, можно придумать контекст, в котором она станет осмысленной, подобно тому, как английский писатель Г. Х. Честертон написал серию «Охотничьих рассказов», в которых он пытался придать реальный смысл нелепым ситуациям, встречающимся в английских пословицах «съесть свою шляпу», «зажечь Темзу» и др.

Далее мы будем называть *фразой с многоточием* совокупность конечного числа слов и одного многоточия, расставленных в определенном порядке. Фразы с многоточием будем обозначать большими латинскими буквами. Примеры фраз с многоточием:

A. ...вертится быстро.
B. Я подошел к ...
C. Я видел ...
D. Мой ...большой.

Малыми греческими буквами мы будем обозначать слова. Если A — некоторая фраза с многоточием, а α — некоторое слово, то через Aα мы будем обозначать фразу, получающуюся при подстановке слова α в фразу A с многоточием. Например, если α — слово стол, то под Aα, Bα, Cα, Dα понимаются соответственно фразы:

Aα. Стол вертится быстро.
Bα. Я подошел к стол.
Cα. Я видел стол.
Dα. Мой стол большой.

После такой подстановки одни фразы, в данном случае Aα, Cα, Dα, оказываются грамматически допустимыми, другие же, как Bα, — недопустимыми. При подстановке в эти же фразы с многоточием слóва β — столу — допустимой окажется фраза Bβ, а фразы Aβ, Cβ, Dβ — недопустимыми. При подстановке туда же слóва γ — кенгуру — все фразы Aγ, Bγ, Cγ, Dγ становятся допустимыми, а при подстановке слова δ — коня — допустимой будет фраза Cδ, а фразы Aδ, Bδ, Dδ — недопустимыми.

Мы назовем два слова α и β *эквивалентными*, если для любой фразы с многоточием N или обе фразы Nα и Nβ недопустимы, или же обе фразы Nα и Nβ допустимы.

Слова α, β, γ, δ, приведенные выше, все попарно неэквивалентны, так как для пары слов α — стол и β — столу фраза Aα грамматически допустима, а Aβ — нет; для пары слов α — стол и γ — кенгуру фраза Bα недопустима, а Bγ допустима и т. д. Легко видеть, что слова стол и стул эквивалентны. Так же эквивалентны слова корову ~ забаву, смотрю ~ сижу.

Это соотношение эквивалентности обладает свойствами 1°—3° рефлексивности симметрии и транзитивности. Действительно, $\alpha \sim \alpha$, так как при любой фразе с многоточием A и любом слове α фразы Aα и Aα одновременно либо допустимы, либо нет; также очевидно, что из $\alpha \sim \beta$ следует $\beta \sim \alpha$ и, наконец, если $\alpha \sim \beta$, $\beta \sim \gamma$, то $\alpha \sim \gamma$, так как если при любой фразе A одновременно допустимы (недопустимы) пара Aα и Aβ и пара Aβ и Aγ, то допустима (недопустима) одновременно и пара фраз Aα, Aγ.

Но нами было указано, что в этом случае всё множество слов можно разбить на совокупность непересекающихся классов попарно эквивалентных между собой элементов. Будем обозначать эти классы большими готическими буквами $\mathfrak{N}$, $\mathfrak{V}$, $\mathfrak{C}$, $\mathfrak{D}$, ... Значит, если $\alpha \in \mathfrak{N}$, $\beta \in \mathfrak{N}$, то $\alpha \sim \beta$, а если $\alpha \in \mathfrak{N}$, $\beta \in \mathfrak{V}$, то α не эквивалентно β. Так, классами будут, например, совокупности слов:

$\mathfrak{N}$. стол, стул, дом, пол, стон, ...
$\mathfrak{V}$. смотрю, сижу, вижу, пилю, ...
$\mathfrak{C}$. видеть, смотреть, брить, ...
$\mathfrak{D}$. корове, забаве, надежде, ...

Для того чтобы полностью проверить эквивалентность какой-либо пары слов, нужно было бы вставлять их во всевозможные фразы с многоточием. Таких фраз бесконечное множество, так как мы в нашем определении не ограничивали длину фразы. По-видимому, наша конструкция не изменится, если мы ограничимся, например, фразами длиной меньше чем в 100 слов, а тогда (поскольку число всех слов русского языка конечно) для проверки существует алгоритм конечной длины [1]). Но всё равно и в этом случае полностью осуществить проверку было бы нельзя. Таким образом, может показаться, что критерий эквивалентности не может быть применен на практике.

Такой вывод неверен. Дело в том, что эта ситуация аналогична в главных чертах ситуации, возникающей при любой попытке осуществить физически, на практике математическую схему. Например, многие физические величины определяются при помощи «мысленного эксперимента». Этот мысленный эксперимент осуществляется на практике лишь приблизительно, и всегда есть некоторая вероятность ошибки; тем не менее такой способ определять физические величины остается вполне разумным. Так происходит и в нашем случае. Мы не можем перебрать все фразы, чтобы доказать эквивалентность двух слов. Однако, перебрав достаточное количество разнообразных фраз, мы можем прийти к практически обоснованной уверенности в том, что дальнейшая проверка ничего не изменит, и наши слова действительно эквивалентны.

Недостаток введенной только что классификации состоит в том, что слова, выполняющие несколько функций одновременно, оказываются включенными в особые классы. Например, имеются отдельные классы:

$\mathfrak{P}$. окно, весло, лето, ...
$\mathfrak{W}$. окну, веслу, лету, ...
$\mathfrak{M}$. метро, пальто, ...

Однако естественно считать, что слова класса $\mathfrak{N}$ должны одновременно входить как в класс $\mathfrak{P}$, так и в класс $\mathfrak{W}$. Поэтому введем следующее определение: мы будем говорить, что *слово* α *предшествует слову* β, а *слово* β

[1]) Т. е. прием, позволяющий в определенной последовательности перебрать все возможные фразы с многоточием и в каждую из этих фраз подставить оба слова данной пары, с тем чтобы убедиться, являются ли эти слова эквивалентными.

следует за словом α (обозначается $\alpha \longrightarrow \beta$ или $\beta \longleftarrow \alpha$), если для всякой фразы с многоточием А такой, что фраза $A\alpha$ грамматически допустима, допустима также и фраза $A\beta$. Например, окно $\longrightarrow$ метро, окну $\longrightarrow$ метро, окном $\longrightarrow$ метро.

Докажем следующую теорему.

Для любых несовпадающих классов $\mathfrak{A}$ и $\mathfrak{B}$ и любых слов α и β, принадлежащих соответственно этим классам ($\alpha \in \mathfrak{A}$, $\beta \in \mathfrak{B}$), имеет место одно из трех утверждений: 1) *либо* $\alpha \longrightarrow \beta$, 2) *либо* $\beta \longrightarrow \alpha$, 3) *либо ни одно из утверждений* $\alpha \longrightarrow \beta$, $\beta \longrightarrow \alpha$ *неверно*. При этом если одно из утверждений верно для некоторой пары α, β ($\alpha \in \mathfrak{A}$, $\beta \in \mathfrak{B}$), то оно верно и для всякой другой пары α', β' ($\alpha' \in \mathfrak{A}$, $\beta' \in \mathfrak{B}$).

В самом деле, если утверждение 1) имеет место для некоторой пары слов $\alpha \in \mathfrak{A}$ и $\beta \in \mathfrak{B}$, то оно также верно и для любой другой пары слов $\alpha' \in \mathfrak{A}$, $\beta' \in \mathfrak{B}$.

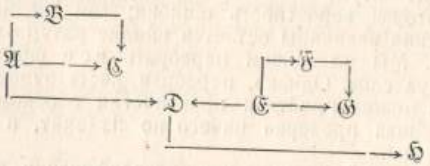
Действительно, так как $\alpha \in \mathfrak{A}$ и $\alpha' \in \mathfrak{A}$, то $\alpha \sim \alpha'$, фразы $A\alpha$ и $A\alpha'$ одновременно грамматически допустимы или недопустимы. Аналогично, одновременно грамматически допустимы или недопустимы фразы $A\beta$ и $A\beta'$. Но тогда если из грамматической допустимости фразы $A\alpha$ следует грамматическая допустимость фразы $A\beta$, то и из грамматической допустимости фразы $A\alpha'$ следует грамматическая допустимость фразы $A\beta'$, т. е. если $\alpha \longrightarrow \beta$, то и $\alpha' \longrightarrow \beta'$.

Аналогично, если утверждение 2) имеет место для некоторой пары слов $\alpha \in \mathfrak{A}$ и $\beta \in \mathfrak{B}$, то оно верно и для любой другой пары слов $\alpha' \in \mathfrak{A}$ и $\beta' \in \mathfrak{B}$.

Если же ни 1), ни 2) не имеют места для $\alpha \in \mathfrak{A}$ и $\beta \in \mathfrak{B}$, то отсюда вытекает 3). Теорема, таким образом, доказана.

На основе этой теоремы можно ввести следующие определения. Будем говорить, что *класс $\mathfrak{A}$ предшествует классу $\mathfrak{B}$, а класс $\mathfrak{B}$ следует за классом $\mathfrak{A}$* (обозначения: $\mathfrak{A} \longrightarrow \mathfrak{B}$ или $\mathfrak{B} \longleftarrow \mathfrak{A}$), если имеет место первое из наших утверждений. Так, в приведенном выше примере $\mathfrak{B} \longrightarrow \mathfrak{N}$, $\mathfrak{M} \longrightarrow \mathfrak{N}$.

Рассмотрим теперь систему из нескольких различных классов $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$, $\mathfrak{E}$, $\mathfrak{F}$, $\mathfrak{G}$, $\mathfrak{H}$, связанных друг с другом так, что их можно изобразить, например, в виде следующей схемы:



(1)

Такая схема всегда будет обладать двумя следующими свойствами:

1) *Если* $\mathfrak{L} \longrightarrow \mathfrak{M}$ *и* $\mathfrak{M} \longrightarrow \mathfrak{N}$, *то и* $\mathfrak{L} \longrightarrow \mathfrak{N}$ (т. е. на схеме $\mathfrak{L} \longrightarrow \mathfrak{M} \longrightarrow \mathfrak{N}$ можно провести дополнительную стрелку $\mathfrak{L} \longrightarrow \mathfrak{N}$).

2) $\mathfrak{M} \longrightarrow \mathfrak{N}$ и $\mathfrak{N} \longrightarrow \mathfrak{M}$ *одновременно не могут существовать* (т. е. схема $\mathfrak{M} \rightleftarrows \mathfrak{N}$ невозможна).

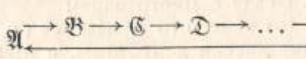
Докажем это.

1) Если $\mathfrak{L} \longrightarrow \mathfrak{M}$, то для любого $\lambda \in \mathfrak{L}$ и любого $\mu \in \mathfrak{M}$ из грамматической допустимости фразы $A\lambda$ следует грамматическая допустимость фразы $A\mu$. Если $\mathfrak{M} \longrightarrow \mathfrak{N}$, то для любого $\mu \in \mathfrak{M}$ и любого $\nu \in \mathfrak{N}$ из грамматической допустимости фразы $A\mu$ следует грамматическая допустимость фразы $A\nu$. Но если из грамматической допустимости фразы $A\lambda$ следует грамматическая допустимость фразы $A\nu$ для любых λ и ν, то это значит, что класс $\mathfrak{L}$ предшествует классу $\mathfrak{N}$: $\mathfrak{L} \longrightarrow \mathfrak{N}$.

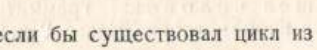
2) Пусть $\mathfrak{M} \longrightarrow \mathfrak{N}$ и $\mathfrak{N} \longrightarrow \mathfrak{M}$. Так как $\mathfrak{M} \longrightarrow \mathfrak{N}$, то если $\mu \in \mathfrak{M}$ и $\nu \in \mathfrak{N}$, то из грамматической допустимости $A\mu$ следует грамматическая допустимость $A\nu$, а так как $\mathfrak{N} \longrightarrow \mathfrak{M}$, то из грамматической допустимости $A\nu$ следует

грамматическая допустимость $A\mu$. Значит, фразы $A\nu$ и $A\mu$ одновременно либо допустимы, либо нет. Но тогда для любого $\mu \in \mathfrak{M}$ и любого $\nu \in \mathfrak{N}$ слова μ и ν эквивалентны и классы $\mathfrak{M}$ и $\mathfrak{N}$ должны совпадать, что противоречит тому, что они различны.

Положения 1) и 2) показывают, что на нашей схеме не может быть «замкнутых циклов», т. е. систем стрелочек вида

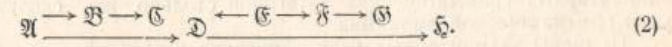


Действительно, если бы существовал цикл из трех классов вида



то по свойству 1) на этой схеме можно было бы провести дополнительную стрелку $\mathfrak{A} \longrightarrow \mathfrak{C}$. Но, по свойству 2), одновременно $\mathfrak{A} \longrightarrow \mathfrak{C}$ и $\mathfrak{C} \longrightarrow \mathfrak{A}$ существовать не могут. Аналогично, из невозможности существования цикла из трех классов доказывается невозможность существования цикла из четырех классов и т. д.

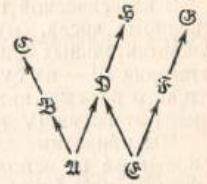
Будем теперь из нашей системы [например, (1)] отбрасывать те стрелки, которые получаются как следствия из остальных стрелок по свойству 1), — так, из схемы $\mathfrak{A} \longrightarrow \mathfrak{B} \longrightarrow \mathfrak{C}$ отбросим стрелку $\mathfrak{A} \longrightarrow \mathfrak{C}$, оставив лишь промежуточные стрелки. Такое выбрасывание мы будем производить до тех пор, пока совсем не останется стрелочек, подлежащих выбрасыванию. Оставшиеся стрелочки мы назовем *элементарными*. Так, схема (1) после указанных выбрасываний сведется к следующей схеме, содержащей только элементарные стрелочки:



(2)

Процесс выбрасывания стрелочек и сведения данной схемы типа (1) к схеме типа (2) обязательно должен окончиться, так как классов слов не больше, чем самих слов, а слов — конечное число; стрелочек же столько, сколько пар классов, а их — конечное число, так как самих классов конечное число.

Таким образом, если класс $\mathfrak{A}$ предшествует классу $\mathfrak{B}$ или, что то же самое, класс $\mathfrak{B}$ следует за классом $\mathfrak{A}$ ($\mathfrak{A} \longrightarrow \mathfrak{B}$), то всегда есть система элементарных стрелочек, ведущая из $\mathfrak{A}$ в $\mathfrak{B}$ (хотя сама стрелочка, ведущая из $\mathfrak{A}$ в $\mathfrak{B}$, могла оказаться не элементарной и быть выкинутой).



Систему элементарных стрелочек типа (2), остающуюся после окончания процесса выбрасывания, мы назовем *деревом грамматической омонимии* или, для краткости, просто *деревом*. Приводим здесь схематический рисунок дерева.

Назовем класс $\mathfrak{M}$ *корнем* дерева, если к этому классу не ведет никакая стрелка (т. е. не существует класса $\mathfrak{N}$ такого, что $\mathfrak{N} \longrightarrow \mathfrak{M}$) либо если существует такая фраза с многоточием А, что фраза $A\mu$ при $\mu \in \mathfrak{M}$ грамматически допустима, а фраза $A\nu$, где $\nu \in \mathfrak{N}$, грамматически недопустима при любом классе $\mathfrak{N} \longrightarrow \mathfrak{M}$.

Например, корнем является класс слов стол, стул, дом, пол, стон, ..., потому что к этому классу не ведет никакая стрелка. По этой же причине корнями будут классы слов

$\mathfrak{C}$. окно, весло, лето, ...

$\mathfrak{M}$. окну, веслу, лету, ...

Возьмем в качестве еще одного примера класс $\mathfrak{R}$, состоящий из двух слов

$\mathfrak{R}$. это, то.

К этому классу ведет стрелка из единственного класса $\mathfrak{L}$, состоящего из слов

красное, зеленое, широкое, ..., к которому не ведет никакая стрелка.

Если через A обозначить фразу с многоточием

A. ... шел человек, то ясно, что фразы «это шел человек», «то шел человек» грамматически допустимы, а при любом $\beta \in \mathfrak{L}$ фраза $A\beta$ (например, «красное шел человек») грамматически недопустима. Значит, класс слов $\mathfrak{R}$ является корнем дерева. Класс $\mathfrak{L}$ также является корнем дерева. Класс же

$\mathfrak{N}$. метро, пальто, кенгуру, ... корнем дерева не будет. Во первых, к нему ведут стрелки, например, из класса слов окно, весло, лето... Во-вторых, какую бы мы фразу с многоточием A ни взяли, всегда найдется такой класс $\mathfrak{P}(\mathfrak{P} \to \mathfrak{N})$, что фраза $A\pi$ будет допустимой при $\pi \in \mathfrak{P}$. Например, если A — фраза с многоточием

A. Я думал о ...,

то таким окажется класс

$\mathfrak{P}$. окне, весле, лете...

При другой фразе с многоточием это будет уже другой класс, но найдется он обязательно.

Сопоставим каждому корню дерева совокупность всех слов, входящих в этот класс и во все следующие за ним классы. Всю эту совокупность слов мы будем называть *элементарной грамматической категорией*. Значит, элементарных грамматических категорий столько же, сколько корней у дерева грамматической омонимии.

Примеры элементарных грамматических категорий:

I. Стол, стул, дом, пол, стон, ...
II. Окно, метро, солнце, пальто, лето, кенгуру, ...
III. Окну, метро, солнцу, пальто, лету, кенгуру, ...
IV. Это, то, красное, зеленое, широкое, ...
V. Это, то.
VI. Сижу, сторожу, режу, плаваю, ...

В классической грамматике категория I состоит из существительных единственного числа, мужского рода, именительного падежа; категория II — из существительных единственного числа, среднего рода, именительного падежа; категория III — из существительных единственного числа, среднего рода, дательного падежа; категория VI — из глаголов единственного числа, первого лица. Категории IV и V в классической грамматике аналогов не имеют.

Мы приводим для иллюстрации наших рассуждений только примеры, основанные на использовании слов русского языка. Мы должны предупредить читателя, что трудоемкая работа по выявлению эквивалентности или неэквивалентности данных слов или исследованию данных грамматических категорий была проведена нами далеко не полностью, и поэтому не исключено, что некоторые из приводимых примеров разобраны неверно, что, конечно, не отражается на не зависящих от этих примеров общих построениях. Кроме того, надо сказать, что русский язык является относительно мало благодарным материалом для примеров, так как высоко развитая система окончаний приводит к тому, что явление грамматической омонимии встречается в русском языке редко, и дерево омонимии русского языка является слабо развет-

вленным. По-видимому, его устройство окажется значительно более сложным для такого языка, как, например, английский.

Безусловно, что делаемые в настоящее время первые попытки теоретико-множественного подхода к построению грамматики, в том числе в предлагаемой работе, очень несовершенны. Однако их усовершенствование возможно лишь на основе большой экспериментальной работы, связанной с наполнением абстрактных математических схем конкретным языковым материалом из данных по грамматической структуре языков. В частности, кажется очень полезным отыскание элементарных грамматических категорий для разных языков и сравнение получившихся результатов с классификацией слов, принятой в традиционной грамматике. Особенно многообещающей кажется возможность приложения этой схемы к языкам не индоевропейской системы, для которых грамматика таких языков создавалась под невольным влиянием привычек, возникавших при изучении европейских языков. Подход на основе математических методов не несет в себе подобной европоцентристской предвзятости. Например, теоретико-множественный подход может помочь решению спорного вопроса о том, имеются ли в таких языках, как полинезийский, морфологические категории.

Подробные исследования на конкретном материале укажут, в каких именно улучшениях нуждается введенное в этой работе определение понятия элементарной грамматической категории.

В заключение укажем некоторую литературу по применению математических методов в лингвистике.

Следующие книги и статьи посвящены вопросам машинного перевода:

1. Machine Translation of languages, New York, 1955. (Русский перевод: Машинный перевод, ИЛ, 1957.)
2. Report of the eight annual round table meetings on linguistics and language studies. Research in Machine Translation, Washington, 1957.
3. A.D. Booth, L. Brandwood, I. R. Cleave, Mechanical Resolution of linguistic problems, London, 1958.
4. Д. Ю. Панов, А. А. Ляпунов, И. С. Мухин, Автоматизация перевода с одного языка на другой. Сборник «Сессия АН СССР по научным проблемам автоматизации производства», 15—20. X. 1956 г.».

Попыткам математического исследования грамматической структуры языка посвящены книги:

1. J. Greenberg, Essays in linguistics, N.-Y., 1957.
2. J. Chomsky, Syntactic structures, Grawenhage, 1957.

Общие вопросы лингвистики и теории связи обсуждаются в книге:

C. Cherry, On Human communication, A review, a survey and a criticism, N.-Y.— London, 1957.

Статистическим проблемам лингвистики посвящены книги:

1. G. Herdan, Language as Choice and Chance, Groninden, 1956.
2. Вопросы статистики речи (Материалы совещания), Изд-во Ленинградского ун-та, 1958.

Укажем также периодические издания и сборники, в которых систематически публикуются материалы по затронутым в статье проблемам:

1. Бюллетень объединения по проблемам машинного перевода, Изд-во 1-го Моск. гос. пед. ин-та иностр. языков (с 1957 г.).
2. «Проблемы кибернетики», Физматгиз, М. (с 1957 г.).
3. «Mechanical Translation» (с 1954 г.).
4. «Information and Control» (с 1957 г.).
5. «Language and Speech» (с 1958 г.).

Содержание настоящей статьи во многом определилось в дружеских беседах автора с В. А. Успенским, привлекшим его внимание к вопросам лингвистики, и В. В. Ивановым. Очень важными были также мысли, высказывавшиеся неоднократно А. Н. Колмогоровым. Большую помощь в написании статьи оказала И. С. Добрушина. Автор приносит им свою искреннюю благодарность.

КОМПЛЕКСНЫЕ ЧИСЛА И ИХ ПРИМЕНЕНИЕ В ГЕОМЕТРИИ

И. М. Яглом

(Москва)

Тема настоящей статьи касается одновременно и алгебры, и геометрии. Связи между этими двумя дисциплинами очень разнообразны и плодотворны для каждой из них. Многие применения алгебры к геометрии и геометрии к алгебре были известны уже в далекой древности: достаточно вспомнить про вывод формулы $(a+b)^2 = a^2+2ab+b^2$, опирающийся на разбиение квадрата со стороной $a+b$ на два меньших квадрата и два прямоугольника, или про алгебраический метод решения задач на построение. Ближе к нашему времени, уже в этом веке, в математике появились большие научные направления, в одинаковой степени относящиеся и к алгебре, и к геометрии; хорошими примерами здесь могут служить алгебраическая геометрия, а также активно развивающееся направление, которое, за неимением лучшего названия, пока приходится называть «линейная алгебра и проективная геометрия». Многими нитями связано с геометрией и возникшее в рамках алгебры учение о комплексных числах; в настоящей статье будет говориться только об одной из этих связей.

I. РАЗЛИЧНЫЕ ТИПЫ КОМПЛЕКСНЫХ ЧИСЕЛ

1. Что такое комплексные числа. С *комплексными числами* считает себя знакомым каждый, изучавший в средней школе математику. Так называют систему величин особого рода, представляемых в виде сумм $a+b\cdot i$ или $a+bi$, где a и b — обыкновенные (вещественные) числа, а i — так называемая *мнимая единица*. Сложение и вычитание комплексных чисел производится как сложение и вычитание сумм:

$$(a+bi)+(c+di)=a+c+bi+di=(a+c)+(b+d)i \quad (1а)$$

и

$$(a+bi)-(c+di)=(a-c)+(b-d)i. \quad (1б)$$

Из этого определения следует, что сложение и вычитание комплексных

чисел по своим свойствам весьма близки к сложению и вычитанию вещественных чисел; так, если $z_1 = a + bi$, $z_2 = c + di$ и $z_3 = e + fi$ — какие-то комплексные числа, то

$$\left.\begin{aligned} &z_1 + z_2 = z_2 + z_1 \ [= (a + c) + (b + d)\,i], \\ &(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3) \ [= (a + c + e) + (b + d + f)\,i] \end{aligned}\right\} \quad (2а)$$

и

$$z_1 - z_2 = z_1 + (-z_2), \quad (2б)$$

где $-z_2 = (-c) + (-d)\,i$ — такое число, что $z_2 + (-z_2) = 0 + 0i = 0$.

Определение умножения комплексных чисел уже требует некоторых дополнительных соглашений. Естественно считать, что

$$(a + bi)\cdot(c + di) = ac + adi + bci + bdi^2. \quad (3)$$

Однако для того, чтобы выяснить смысл последнего выражения, нам надо условиться, чему равна величина i^2. Соображения, связанные с вопросом о решении квадратных уравнений, делают естественным предположение: $i^2 = -1$. В таком случае мы будем иметь

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)\,i. \quad (4)$$

Таким образом, произведение двух комплексных чисел также является комплексным числом; при этом, как легко проверить, при любых комплексных z_1, z_2 и z_3

$$z_1 z_2 = z_2 z_1, \quad (z_1 z_2)\cdot z_3 = z_1 (z_2 z_3) \quad \text{и} \quad (z_1 + z_2) z_3 = z_1 z_3 + z_2 z_3. \quad (5)$$

Несколько сложнее обстоит дело с делением комплексных чисел. Очевидно, для того чтобы научиться делить комплексное число $z_1 = c + di$ на другое комплексное число $z = a + bi$, достаточно определить, чему равно частное $\frac{1}{z}$; после этого деление комплексных чисел сведется к известной уже нам операции умножения:

$$z_1 : z = z_1 \cdot \frac{1}{z}. \quad (6)$$

Заметим теперь, что существуют такие комплексные числа z, для которых определение величины $\frac{1}{z}$ не представляет никакого труда,— это числа $a + 0\cdot i$, которые естественно отождествить с вещественными числами: $a + 0\cdot i = a$. Отсюда мы заключаем, что при $a \neq 0$ $\frac{1}{a + 0i} = \frac{1}{a}$, и следовательно,

$$\frac{c + di}{a + 0i} = \frac{1}{a}(c + di) = \frac{c}{a} + \frac{d}{a}\,i. \quad (7)$$

Рассмотрим теперь произвольное комплексное число $z = a + bi$. Сопряженным к нему называется комплексное число $a - bi$; оно

обозначается через $\bar{z}$. Операция взятия сопряженного комплексного числа обладает многими замечательными свойствами; так,

$$\bar{\bar{z}} = z \quad (8)$$

(ибо если $z = a + bi$, то $\bar{z} = a - bi$ и $\overline{(\bar{z})} = a - (-b)\,i = a + bi = z$)

и

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2, \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2 \quad (9а)$$

(например, если $z_1 = a + bi$, $z_2 = c + di$, то $\overline{z_1 \cdot z_2} = (ac - bd) - (ad + bc)\,i = (a - bi)(c - di) = \bar{z}_1 \cdot \bar{z}_2$); ясно также, что комплексное число лишь тогда совпадает со своим сопряженным, если оно является вещественным, т. е. имеет форму $a + 0i$. Но наиболее важно для нас то, что *произведение двух сопряженных комплексных чисел всегда будет числом вещественным:* произведение $z \cdot \bar{z} = (a + bi)(a - bi) = a^2 + b^2$ есть квадрат числа $r = \sqrt{a^2 + b^2}$, называемого *модулем* z и обозначаемого через $|z|$. То, что произведение сопряженных комплексных чисел вещественно, позволяет свести деление произвольных комплексных чисел к разобранному уже выше случаю деления числа на вещественное:

$$\frac{z_1}{z} = \frac{z_1 \bar{z}}{z \bar{z}}, \quad (10а)$$

или, подробнее,

$$\frac{c + di}{a + bi} = \frac{(c + di)(a - bi)}{(a + bi)(a - bi)} = \frac{(ac + bd) + (ad - bc)\,i}{a^2 + b^2} = \frac{ac + bd}{a^2 + b^2} + \frac{ad - bc}{a^2 + b^2}\,i. \quad (10б)$$

Отметим кстати, что это определение частного двух комплексных чисел показывает, что в полном согласии с равенствами (9а)

$$\overline{z_1 : z_2} = \bar{z}_1 : \bar{z}_2. \quad (9б)$$

Впоследствии нами также будет использовано то обстоятельство, что сумма $z + \bar{z}$ двух сопряженных комплексных чисел всегда является числом вещественным $[(a + bi) + (a - bi) = 2a]$: она равна удвоенной *вещественной части* a числа z. Разность $z - \bar{z}$ двух сопряженных чисел является числом чисто мнимым (т. е. имеет вид $2bi$, где $2b$ — вещественно).

Итак, комплексные числа можно складывать, вычитать, умножать и делить, причем все законы, которым подчиняются эти действия, точно совпадают с законами действий над обыкновенными вещественными числами [1]). В частности, как и в случае вещественных чисел,

[1]) С алгебраической точки зрения это означает, что комплексные числа, как и числа вещественные, образуют *коммутативное поле*.

деление на комплексное число $z=a+bi$ возможно не всегда: для выполнимости деления необходимо, чтобы модуль $|z|=\sqrt{a^2+b^2}$ числа z был отличен от нуля, так что существует единственное комплексное число $0=0+0i$, деление на которое невозможно. В тех случаях, когда невозможность деления на нуль представляет неудобства, условливаются считать, что частное $1:0$ существует, но является числом особого рода, для которого вводится специальное обозначение ∞; другими словами *расширяют множество комплексных чисел*, вводя новое число ∞ («бесконечность»), по определению равное $\frac{1}{0}$ [1]). Правила действий над символом ∞ определяются следующим образом:

$$z+\infty=\infty, \quad z-\infty=\infty, \quad z\cdot\infty=\infty, \quad \frac{\infty}{z}=\infty, \quad \frac{z}{\infty}=0; \tag{11}$$

здесь z — произвольное число, причем в третьем равенстве $z\neq 0$, а во втором и в двух последних $z\neq\infty$; разность $\infty-\infty$, произведение $0\cdot\infty$ и отношение $\frac{\infty}{\infty}$ $\left(\text{а также и отношение } \frac{0}{0}=0\cdot\frac{1}{0}=0\cdot\infty\right)$ приходится считать, вообще говоря, не имеющими смысла, причем здесь уже ничем помочь нельзя [2]).

В ряде случаев оказывается более удобной иная форма записи комплексного числа $z=a+bi$, выдвигающая на первый план его модуль $|z|=\sqrt{a^2+b^2}$. Вынесем число $|z|$ за скобки:

$$z=a+bi=\sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}}+\frac{b}{\sqrt{a^2+b^2}}i\right).$$

Стоящие в скобках вещественные числа $\frac{a}{\sqrt{a^2+b^2}}$ и $\frac{b}{\sqrt{a^2+b^2}}$ обладают тем свойством, что сумма их квадратов равна 1; отсюда следует существование такого угла φ, что

$$\cos\varphi=\frac{a}{\sqrt{a^2+b^2}}, \quad \sin\varphi=\frac{b}{\sqrt{a^2+b^2}}. \tag{12}$$

[1]) Аналогичным образом строятся и сами комплексные числа: желание сделать все квадратные уравнения разрешимыми вынуждает считать, что корень $\sqrt{-1}$ существует, но является числом особого рода; это «число» и обозначают буквой i.

[2]) Отметим, впрочем, что отношение $\frac{a\cdot\infty+b}{c\cdot\infty+d}$, где a, b, c, d — произвольные комплексные числа, в силу тождеств $\frac{az+b}{cz+d}=\frac{a+b\cdot\frac{1}{z}}{c+d\cdot\frac{1}{z}}$ и $\frac{1}{\infty}=0$, следует считать имеющим вполне определенное значение, а именно $\frac{a}{c}$. Это замечание нам будет полезно впоследствии.

Если еще обозначить модуль $|z|=\sqrt{a^2+b^2}$ числа z одной буквой r, то мы будем иметь

$$z=r(\cos\varphi+i\sin\varphi), \tag{13}$$

где $r=\sqrt{a^2+b^2}$ и $\cos\varphi=\frac{a}{r}$, $\sin\varphi=\frac{b}{r}$. Угол φ (определяемый равенствами (12) с точностью до слагаемого, кратного 2π) называется *аргументом* числа z и обозначается через $\operatorname{Arg} z$. Если его ограничить, например, условием $-\pi<\varphi\leqslant\pi$, то для положительных вещественных чисел он будет равен 0, а для отрицательных — равен π; сопряженные числа будут иметь одинаковый модуль r и противоположные аргументы φ и $-\varphi$.

Форма (13) записи комплексных чисел называется *тригонометрической формой*. Она чрезвычайно удобна, когда приходится перемножать два или несколько комплексных чисел. В самом деле,

$$\begin{aligned} r(\cos\varphi+i\sin\varphi)\cdot r_1(\cos\varphi_1+i\sin\varphi_1)= \\ =rr_1[(\cos\varphi\cos\varphi_1-\sin\varphi\sin\varphi_1)+i(\cos\varphi\sin\varphi_1+\sin\varphi\cos\varphi_1)]= \\ =rr_1[\cos(\varphi+\varphi_1)+i\sin(\varphi+\varphi_1)]; \end{aligned} \tag{14}$$

таким образом, *модуль произведения двух комплексных чисел равен произведению модулей сомножителей, а аргумент произведения — сумме аргументов сомножителей* [ср. со значительно менее удобной формулой (4)]. Отсюда вытекает также, что *модуль частного двух комплексных чисел равен частному модулей этих чисел, а аргумент частного — разности соответствующих аргументов*:

$$\frac{z_1}{z}=\frac{r_1(\cos\varphi_1+i\sin\varphi_1)}{r(\cos\varphi+i\sin\varphi)}=\frac{r_1}{r}[\cos(\varphi_1-\varphi)+i\sin(\varphi_1-\varphi)]. \tag{15}$$

Из этих правил сразу выводятся законы, позволяющие возводить комплексное число z в любую степень и извлекать из него корень:

$$\begin{aligned} [r(\cos\varphi+i\sin\varphi)]^n=r^n(\cos n\varphi+i\sin n\varphi); \\ \sqrt[n]{r(\cos\varphi+i\sin\varphi)}=\sqrt[n]{r}\left(\cos\frac{\varphi}{n}+i\sin\frac{\varphi}{n}\right) \end{aligned} \tag{16}$$

(n различных значений корня n-й степени мы получим, выбрав в качестве φ в последней формуле n значений аргумента $\varphi=\varphi_0+2k\pi$, $k=0, 1, \ldots, n-1$, φ_0 — какое-то одно из возможных значений аргумента).

Сделаем теперь еще один шаг в сторону дальнейшего обобщения понятия комплексного числа. Выше мы говорили, что соглашение о том, что $i^2=-1$, диктуется соображениями, связанными с теорией квадратных уравнений. Но в настоящей статье мы совсем не будем интересоваться вопросом об уравнениях и их корнях. А если так, то является ли для нас это соглашение обязательным?

Спросим себя, *что изменится в наших рассуждениях, если мы положим величину i^2 равной не -1, а какому угодно другому комплексному числу $p+qi$?* Очевидно, что определение (1а,б) сложения и вычитания комплексных чисел мы сможем сохранить, свойства (2а,б) этих действий при этом не подвергнутся никаким изменениям. Также и определение (3) умножения комплексных чисел нам не надо будет менять. Правда, в силу равенства $i^2=p+qi$ вместо (4) мы теперь будем иметь

$$(a+bi)(c+di)=ac+adi+bci+bd(p+qi)=\\=(ac+bdp)+(ad+bc+bdq)i, \qquad (17)$$

но и теперь, как легко проверить, привычные свойства (5) умножения останутся неприкосновенными. Далее, для того чтобы определить частное $\frac{z_1}{z}=\frac{c+di}{a+bi}$ рассматриваемых здесь «общих» комплексных чисел, надо лишь суметь подобрать для каждого числа $z=a+bi$ такое число $\bar{z}=a_1+b_1i$, чтобы произведение $z\bar{z}$ имело вид $A+0\cdot i$, т. е. являлось вещественным; после этого мы сможем по-прежнему использовать формулы (7) и (10а). Но так как в силу (17)

$$(a+bi)(a_1+b_1i)=(aa_1+bb_1p)+(ab_1+ba_1+bb_1q)i, \qquad (18)$$

то мы можем положить, например, $b_1=-b$, $a_1=\frac{-ab_1-bb_1q}{b}=\\=a+bq$, т. е.

$$\bar{z}=(a+bq)-bi; \qquad (19)$$

при этом и деление «общих» комплексных чисел будет определяться по той же формуле (10а), что и для обыкновенных комплексных чисел.

Итак, мы видим, что *различных систем комплексных чисел имеется бесконечно много*; для того чтобы выбрать какую-нибудь одну из этих систем, достаточно задать (произвольным образом!) два (вещественных) числа p и q. Самыми важными из этих систем комплексных чисел являются следующие:

обыкновенные комплексные числа	$a+bi$,	$i^2=-1$	(здесь $p=-1$, $q=0$);
дуальные числа	$a+b\varepsilon$,	$\varepsilon^2=0$	(здесь $p=0$, $q=0$);
двойные числа[1])	$a+be$,	$e^2=1$	(здесь $p=1$, $q=0$).

[1]) Об этих числах см. статью Д. Д. Ивлева на стр. 197—203 настоящего выпуска.

Можно даже доказать, что все системы самых общих комплексных чисел в некотором смысле сводятся к этим трем системам[1]).

Обыкновенные комплексные числа тесно связаны с вопросом о решении уравнений второй и высших степеней; они играют основную роль в алгебре и во многих разделах математического анализа. Происхождение этих чисел проследить нелегко; считается, что впервые их стали употреблять итальянские математики XVI века Джироламо Кардано и Рафаэль Бомбелли, однако в неявном виде эти числа можно найти и в более ранних работах; с другой стороны, еще долго после Кардано и Бомбелли даже выдающиеся математики не имели отчетливого представления о комплексных числах. Дуальные же и двойные числа не имеют никакого отношения к теории квадратных уравнений с вещественными коэффициентами и вообще сравнительно мало связаны с алгеброй; основные применения эти числа находят в геометрии[2]). Дуальные числа, по-видимому, впервые рассматривал известный немецкий геометр конца прошлого и начала этого века Евгений Штуди; двойные числа были введены современником Штуди английским геометром Вильямом Клиффордом.

Основные применения двойных чисел относятся к неевклидовой геометрии Лобачевского[3]); поэтому в настоящей статье мы сосредоточим наше внимание в первую очередь на обыкновенных комплексных числах и на дуальных числах.

2. Дуальные числа. Сложение, вычитание и умножение дуальных чисел определяется формулами

$$(a+b\varepsilon)+(c+d\varepsilon)=(a+c)+(b+d)\varepsilon,\\(a+b\varepsilon)-(c+d\varepsilon)=(a-c)+(b-d)\varepsilon,\\(a+b\varepsilon)\cdot(c+d\varepsilon)=ac+(ad+bc)\varepsilon \qquad (20)$$

[ср. (1а,б) и (17)]. Из последней формулы, в частности, следует, что произведение дуального числа $z=a+b\varepsilon$ на сопряженное дуальное число $\bar{z}=a-b\varepsilon$ [ср. (19)] будет вещественным:

$$z\cdot\bar{z}=(a+b\varepsilon)(a-b\varepsilon)=a^2. \qquad (21)$$

Квадратный корень a из произведения $z\bar{z}$ (совпадающий с полусуммой $\frac{z+\bar{z}}{2}$ сопряженных чисел z и $\bar{z}$) называют *модулем* дуального числа z

[1]) Для читателя, знакомого с терминологией, принятой в современной алгебре, укажем, что *любая система комплексных чисел изоморфна системе обыкновенных комплексных чисел, дуальных чисел или двойных чисел.*

[2]) Некоторые применения эти системы комплексных чисел находят также в теории чисел.

[3]) И к некоторым другим геометриям, отличным от привычной геометрии Евклида (например, к так называемой *псевдоевклидовой геометрии*, играющей фундаментальную роль в теории относительности).

и обозначают через $|z|$ (отметим, что модуль дуального числа может быть и отрицательным!). Сумма $z+\bar{z}=2a$ двух сопряженных чисел является вещественной; разность $z-\bar{z}=2b\varepsilon$ является числом *чисто мнимым* (т. е. отличается от ε лишь вещественным множителем). Заметим еще, что — в полной аналогии с обыкновенными комплексными числами — дуальное число z тогда и только тогда совпадает со своим сопряженным $\bar{z}$, когда оно является вещественным; формулы (8) и (9а) также остаются справедливыми для дуальных чисел.

Правило деления на дуальное число $z=a+b\varepsilon$ мы теперь можем записать так:

$$\frac{c+d\varepsilon}{a+b\varepsilon}=\frac{(c+d\varepsilon)(a-b\varepsilon)}{(a+b\varepsilon)(a-b\varepsilon)}=\frac{ca+(-cb+da)\varepsilon}{a^2}=\frac{c}{a}+\frac{-cb+da}{a^2}\varepsilon \quad (22)$$

[см. (10б)]; из этого правила вытекает, что для дуальных чисел остается в силе и правило (9б). Отсюда же следует, что для возможности деления на дуальное число z необходимо, чтобы модуль $|z|=a$ этого числа был отличен от нуля; при этом, в противоположность обыкновенным комплексным числам, *дуальное число нулевого модуля само может быть отличным от нуля*. Поскольку невозможность деления на числа нулевого модуля является в ряде случаев ограничением стеснительным, мы будем считать, что частные $\frac{1}{\varepsilon}$ и $\frac{1}{0}$ являются числами новой природы, которые условимся обозначать через ω и ∞; введем также в рассмотрение всевозможные числа вида $c\omega$, где c вещественно. При этом любое дуальное число будет иметь обратное ему:

$$\frac{1}{b\varepsilon}=\frac{1}{b}\omega \quad \text{при} \quad b\neq 0; \quad \frac{1}{0}=\infty.$$

Правила действий над символом ∞ здесь определяются теми же формулами (11), что и выше (причем число z в этих формулах может быть и числом вида $c\omega$); правила действий над «числами» $a\omega$ определяются так [1]):

$$(a+b\varepsilon)+c\omega=c\omega, \quad (a+b\varepsilon)-c\omega=(-c)\omega, \quad (a+b\varepsilon)\cdot c\omega=(ac)\omega,$$
$$\frac{c\omega}{a+b\varepsilon}=\frac{c}{a}\omega, \quad \frac{a+b\varepsilon}{c\omega}=\frac{a}{c}\varepsilon, \quad c\omega\pm d\omega=(c\pm d)\omega, \quad c\omega\cdot d\omega=\infty. \quad (23)$$

[1]) Здесь мы исходим из того, что, например, $\frac{c\omega}{a+b\varepsilon}=\frac{\frac{c}{\varepsilon}}{a+b\varepsilon}$ естественно приравнять $\frac{c}{\varepsilon(a+b\varepsilon)}=\frac{c}{a\varepsilon}$, а $\frac{a+b\varepsilon}{c\omega}=\frac{a+b\varepsilon}{c/\varepsilon}$ считать равным $\frac{\varepsilon(a+b\varepsilon)}{c}=\frac{a}{c}\varepsilon$.

Не имеющими смысла остаются выражения $\frac{\infty}{\infty}$ и $0\cdot\infty$; впрочем, значение дроби $\frac{az+b}{cz+d}$ при $z=\infty$ можно определить так же, как мы это делали в случае обыкновенных комплексных чисел (см. сноску [2]) на стр. 64).

Заметим еще, что числа нулевого модуля характеризуются тем, что произведение такого числа $c\varepsilon$ на отличное от нуля дуальное число z может равняться нулю:

$$c\varepsilon\cdot d\varepsilon=(cd)\varepsilon^2=0; \quad (24)$$

поэтому эти числа называют *делителями нуля*.

Дуальные числа ненулевого модуля a можно также записать в форме, близкой к «тригонометрической форме» (13) комплексного числа:

$$a+b\varepsilon=a\left(1+\frac{b}{a}\varepsilon\right)=r(1+\varepsilon\varphi); \quad (25)$$

здесь, как прежде, $r=a$ обозначает модуль $|z|$ числа $z=a+b\varepsilon$, а отношение $\frac{b}{a}=\varphi$ называется *аргументом* этого числа и обозначается через $\operatorname{Arg} z$ (r может быть произвольным вещественным числом, отличным от нуля, φ — произвольным вещественным числом). Очевидно, что вещественные числа $a=a+0\cdot\varepsilon$ характеризуются равенством нулю их аргумента; сопряженные дуальные числа $z=a+b\varepsilon$ и $\bar{z}=a-b\varepsilon$ имеют одинаковый модуль r и противоположные аргументы φ и $-\varphi$.

Форма (25) записи дуальных чисел очень удобна в тех случаях, когда эти числа приходится перемножать или делить. Действительно,

$$r(1+\varepsilon\varphi)\cdot r_1(1+\varepsilon\varphi_1)=rr_1(1+\varepsilon\varphi+\varepsilon\varphi_1+\varepsilon^2\varphi\varphi_1)=$$
$$=rr_1[1+\varepsilon(\varphi+\varphi_1)]; \quad (26)$$

следовательно, *модуль произведения двух дуальных чисел равен произведению модулей сомножителей* [1]), *а аргумент произведения — сумме аргументов* (ср. выше, стр. 65). Отсюда вытекает, что модуль частного двух комплексных чисел равен частному модулей этих чисел, а аргумент частного — разности соответствующих аргументов:

$$\frac{z_1}{z}=\frac{r_1(1+\varepsilon\varphi_1)}{r(1+\varepsilon\varphi)}=\frac{r_1}{r}[1+\varepsilon(\varphi_1-\varphi)]. \quad (27)$$

Наконец, из этих правил выводятся также и законы, позволяющие возвышать дуальное число в любую степень и извлекать из него корень:

$$[r(1+\varepsilon\varphi)]^n=r^n(1+\varepsilon\cdot n\varphi); \quad \sqrt[n]{r(1+\varepsilon\varphi)}=\sqrt[n]{r}\left(1+\varepsilon\frac{\varphi}{n}\right) \quad (28)$$

[1]) Это утверждение остается в силе и в том случае, когда модуль одного или обоих сомножителей равен нулю [ибо если $|z|=0$, то и $|zz_1|=0$; так, например, $c\varepsilon\cdot(a+b\varepsilon)=(ac)\varepsilon$].

(из последней формулы вытекает, что корень нечетной степени из дуального числа при $r \neq 0$ определяется однозначно; корень же четной степени не существует, если $r < 0$, и имеет два значения, если $r > 0$ [1]).

II. ГЕОМЕТРИЧЕСКИЕ ИНТЕРПРЕТАЦИИ КОМПЛЕКСНЫХ ЧИСЕЛ

3. Обыкновенные комплексные числа как точки плоскости. Развитие теории комплексных чисел в значительной степени связано с геометрическим истолкованием обыкновенных комплексных чисел как точек плоскости, впервые отмеченным, по-видимому, датским математиком XVIII в. Карлом Весселем, но введенным в науку в первую очередь трудами Карла Фридриха Гаусса и Огюстена Коши. Это истолкование состоит в том, что точке плоскости с декартовыми прямоугольными координатами x, y или с полярными координатами r, φ сопоставляется комплексное число

$$z = x + iy = r(\cos\varphi + i\sin\varphi) \qquad (1)$$

(рис. 1). При этом, очевидно, вещественным числам $z = x + 0 \cdot y$ отвечают точки оси x (*вещественная ось* o); числам модуля $r = 1$ отвечают точки окружности с центром в начале координат O и радиусом 1 (*единичная окружность* S). Противоположным числам $z = x + iy$, $-z = -x - iy$ отвечают точки, симметричные относительно точки O; сопряженным комплексным числам $z = x + iy$ и $\bar{z} = x - iy$ отвечают точки, симметричные относительно прямой o. В дальнейшем точку, отвечающую комплексному числу z, мы часто будем обозначать той же буквой z; при этом равенства

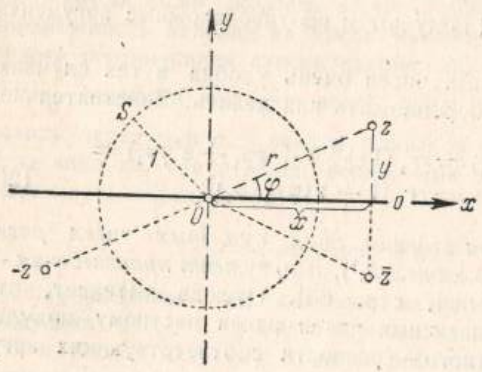


Рис. 1.

$$z' = -z \text{ (а) и } z' = \bar{z} \text{ (б)} \qquad (2)$$

можно понимать как записи определенных точечных преобразований, сопоставляющих каждой точке z новую точку z'. Эти преобразования представляют собой *симметрию относительно точки* O и *симметрию относительно прямой* o.

[1]) Нетрудно видеть, что корень целой степени $n > 1$ из дуального числа $z = b\varepsilon$, модуль $|z|$ которого равен нулю (т. е. из числа, являющегося делителем нуля), извлечь нельзя.

Зафиксируем определенные комплексные числа (точки) $q = a + ib$ и $p = t(\cos\alpha + i\sin\alpha)$. Равенство

$$z' = z + q \quad [(x' + iy') = (x + iy) + (a + ib)] \qquad (3)$$

означает, что $x' = x + a$, $y' = y + b$, т. е. что вектор $\overrightarrow{zz'}$ (вектор с началом в точке z и концом в точке z') совпадает с вектором $\overrightarrow{Oq}$ (геометрический смысл сложения комплексных чисел); поэтому равенство (3) определяет *параллельный перенос* плоскости на вектор $\overrightarrow{Oq}$ (рис. 2). Равенство

$$z' = pz \quad [r'(\cos\varphi' + i\sin\varphi') = t(\cos\alpha + i\sin\alpha)\cdot r(\cos\varphi + i\sin\varphi)] \qquad (4)$$

означает, что $r' = tr$, $\varphi' = \varphi + \alpha$, т. е. что $Oz' = t \cdot Oz$, $\angle\{Oz, Oz'\} = \alpha$ (геометрический смысл умножения комплексных чисел). При этом символ $\angle\{Oz, Oz'\}$ означает так называемый *ориентированный угол* между лучами Oz и Oz', т. е. угол, на который надо повернуть Oz против часовой стрелки, для того чтобы получить Oz' (если вращение происходит в направлении по часовой стрелке, то углу приписывается знак «минус»; ориентированный угол между двумя лучами определяется с точностью до слагаемого, кратного 2π). Таким образом, равенство (4) определяет так называемое *центрально-подобное вращение* плоскости, составляющееся из вращения вокруг O на угол α (в направлении против часовой стрелки) и *центрально-подобного преобразования* (*гомотетии*) с центром O и коэффициентом подобия t (рис. 3, *а*). В частности, если модуль $|p| = t$ комплексного числа p равен 1, преобразование (4) представляет собой *вращение* на угол α (рис. 3, *б*).

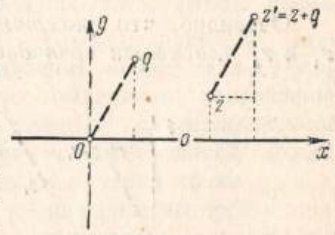


Рис. 2.

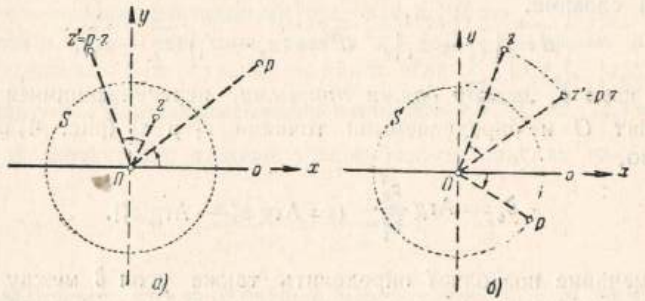


Рис. 3.

Каждое движение плоскости можно представить как вращение вокруг фиксированной точки O, сопровождаемое параллельным переносом, или как симметрию относительно фиксированной прямой o, сопровождаемую вращением вокруг выбранной точки O и параллельным переносом [1]). Отсюда следует, что *каждое движение плоскости можно записать в виде*

$$z' = pz + q \qquad (|p| = 1) \tag{5}$$

или

$$z' = p\bar{z} + q \qquad (|p| = 1)\ ^{2}). \tag{5а}$$

Очевидно, что *расстояние* $d = (z_1, z_2)$ *между двумя точками* z_1 *и* z_2 *плоскости совпадает с модулем* $|w| = |z_2 - z_1|$ *комплексного числа* $w = z_2 - z_1$ (ибо вектор $\overrightarrow{z_1 z_2}$ равен вектору $\overrightarrow{Ow}$, рис. 4, *а*); другими словами,

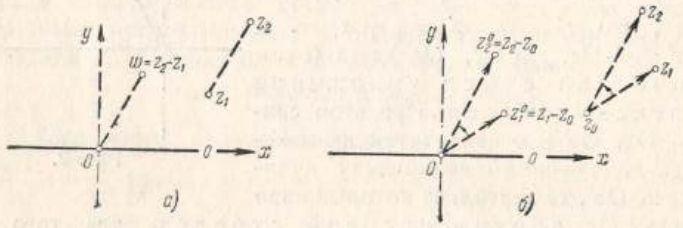


Рис. 4.

$$d = |z_2 - z_1|, \quad d^2 = (z_2 - z_1)(\bar{z}_2 - \bar{z}_1). \tag{6}$$

Далее, *угол* δ_0 *между двумя прямыми*, пересекающимися в начале координат O и определяемыми точками z_1^0 и z_2^0 (рис. 4, *б*), равен, очевидно,

$$\delta_0 = \operatorname{Arg}\frac{z_2^0}{z_1^0} \quad (= \operatorname{Arg} z_2^0 - \operatorname{Arg} z_1^0). \tag{7}$$

Это замечание позволяет определить также угол δ между прямыми, пересекающимися в произвольной точке z_0 и проходящими через точки z_1 и z_2. Простейшим движением, переводящим z_0 в O, является параллельный перенос

$$z' = z - z_0;$$

[1]) См., например, И. М. Яглом, Геометрические преобразования I, М., Гостехиздат, 1955.

[2]) Аналогично можно показать, что *каждое преобразование подобия плоскости можно записать в виде*

$$z' = pz + q \quad \text{или} \quad z' = p\bar{z} + q.$$

этот параллельный перенос переводит точки z_1 и z_2 в точки $z_1^0 = z_1 - z_0$ и $z_2^0 = z_2 - z_0$ (рис. 4, *б*). Отсюда следует, что угол δ равен углу δ_0 между прямыми, пересекающимися в начале координат и проходящими через z_1^0 и z_2^0, т. е.

$$\delta = \operatorname{Arg}\frac{z_2 - z_0}{z_1 - z_0} \qquad [= \operatorname{Arg}(z_2 - z_0) - \operatorname{Arg}(z_1 - z_0)]. \tag{8}$$

Заметим, что δ — это угол, на который надо повернуть против часовой стрелки луч $\overrightarrow{z_0 z_1}$ первой прямой для того, чтобы совместить его с лучом $\overrightarrow{z_0 z_2}$ второй прямой. Прямая, одно из двух направлений которой выделено (это направление можно задать, указав определенный луч прямой; на чертежах выделенное направление обыкновенно указывается стрелкой), называется *ориентированной прямой* или *осью*; выделенное направление ориентированной прямой часто называют *положительным*. Проходящую через точки z_1 и z_2 ориентированную прямую, положительное направление которой совпадает с направлением от z_1 к z_2, мы будем обозначать через $[z_1 z_2]$. *Ориентированный угол* $\angle\{l_1, l_2\}$ *между ориентированными прямыми* l_1 и l_2 определяется как угол, на который надо повернуть против часовой стрелки прямую l_1, для того чтобы ее положительное направление совпало с положительным направлением прямой l_2; этот угол задается с точностью до произвольного слагаемого, кратного 2π. Таким образом, δ есть ориентированный угол между ориентированными прямыми $[z_0 z_1]$ и $[z_0 z_2]$ ($\delta = \angle\{[z_0 z_1], [z_0 z_2]\}$).

Условием того, что три точки z_0, z_1 *и* z_2 *лежат на одной прямой, является* равенство нулю или π угла $\angle\{[z_1 z_2], [z_0 z_2]\}$ или, в силу формулы (8), *вещественность отношения* $V(z_0, z_1, z_2) = \dfrac{z_0 - z_2}{z_1 - z_2}$ *этих трех точек*; это условие можно также записать так:

$$\frac{z_0 - z_2}{z_1 - z_2} = \frac{\bar{z}_0 - \bar{z}_2}{\bar{z}_1 - \bar{z}_2}. \tag{9}$$

Отсюда вытекает, что проходящая через точки z_1 и z_2 *прямая* l — это геометрическое место таких точек z, для которых

$$\frac{z - z_2}{z_1 - z_2} = \frac{\bar{z} - \bar{z}_2}{\bar{z}_1 - \bar{z}_2}. \tag{10}$$

Другими словами, можно сказать, что уравнение этой прямой имеет вид (10) или вид

$$(\bar{z}_1 - \bar{z}_2)z - (z_1 - z_2)\bar{z} + (z_1\bar{z}_2 - \bar{z}_1 z_2) = 0. \tag{10а}$$

Таким образом, *уравнение каждой прямой плоскости можно*

записать в форме

$$Bz - \overline{B}\overline{z} + C = 0, \quad C - \text{чисто мнимое}^1). \tag{11}$$

Нетрудно показать, что и, обратно, *каждое уравнение вида* (11) *задает некоторую прямую линию* (проходящую через такие точки z_1 и z_2, что $z_1 - z_2 = \overline{B}$, $z_1\overline{z}_2 - \overline{z}_1 z_2 = C$).

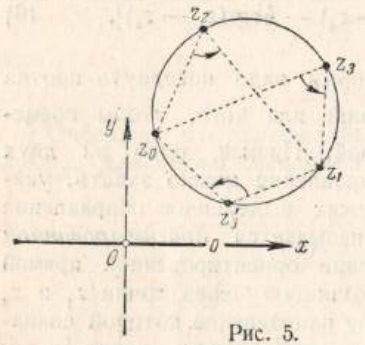


Рис. 5.

Условием того, что четыре точки z_0, z_1, z_2 и z_3 лежат на одной окружности (или прямой), является равенство нулю или π разности углов $\angle\{[z_0z_2], [z_1z_2]\} - \angle\{[z_0z_3], [z_1z_3]\}$ (рис. 5), или в силу (8) *вещественность «двойного отношения»* $W(z_0, z_1, z_2, z_3) = \frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3}$ *этих четырех точек*; это условие можно записать так:

$$\frac{z_0 - z_2}{z_1 - z_2} : \frac{z_0 - z_3}{z_1 - z_3} = \frac{\overline{z}_0 - \overline{z}_2}{\overline{z}_1 - \overline{z}_2} : \frac{\overline{z}_0 - \overline{z}_3}{\overline{z}_1 - \overline{z}_3}. \tag{12}$$

Отсюда вытекает, что уравнение окружности (или прямой) S, проходящей через точки z_1, z_2 и z_3, имеет вид

$$\frac{z - z_2}{z_1 - z_2} : \frac{z - z_3}{z_1 - z_3} = \frac{\overline{z} - \overline{z}_2}{\overline{z}_1 - \overline{z}_2} : \frac{\overline{z} - \overline{z}_3}{\overline{z}_1 - \overline{z}_3} \tag{13}$$

или

$$(z - z_2)(\overline{z} - \overline{z}_3)[(z_1 - z_3)(\overline{z}_1 - \overline{z}_2)] =$$
$$= (z - z_3)(\overline{z} - \overline{z}_2)[(z_1 - z_2)(\overline{z}_1 - \overline{z}_3)],$$

т. е.

$$Az\overline{z} + Bz - \overline{B}\overline{z} + C = 0,$$

где

$$A = (z_1 - z_3)(\overline{z}_1 - \overline{z}_2) - (z_1 - z_2)(\overline{z}_1 - \overline{z}_3),$$
$$B = -z_3(z_1 - z_3)(\overline{z}_1 - \overline{z}_2) + \overline{z}(z_1 - z_2)(\overline{z}_1 - \overline{z}_3),$$
$$C = z_2\overline{z}_3(z_1 - z_3)(\overline{z}_1 - \overline{z}_2) - z_3\overline{z}_2(z_1 - z_2)(\overline{z}_1 - \overline{z}).$$

[1]) Очевидно, что угол $\angle\{l, o\}$, который образует описываемая уравнением (11) прямая l с вещественной осью o, равен $\operatorname{Arg} B = \operatorname{Arg}(\overline{z}_1 - \overline{z}_2) = -\operatorname{Arg}(z_1 - z_2)$; далее, так как $|C| = |Bz - \overline{B}\overline{z}| \leqslant 2|B||z|$ ($|Bz - \overline{B}\overline{z}| = 2|B||z|$, если Bz чисто мнимое, т. е. если $\operatorname{Arg} B + \operatorname{Arg} z = \pm \operatorname{Arg} i$), то для точек z прямой l имеем $|z| \geqslant \frac{|C|}{2|B|}$, откуда следует, что $\frac{|C|}{2|B|}$ есть расстояние прямой l от начала координат.

Таким образом, *уравнение каждой окружности (или прямой) плоскости можно записать в следующей форме*[1]):

$$Az\overline{z} + Bz - \overline{B}\overline{z} + C = 0, \quad A \text{ и } C - \text{чисто мнимые}. \tag{14}$$

Обратно, *геометрическое место точек, удовлетворяющих любому уравнению вида* (14) (если только такие точки существуют[1])) *представляет собой окружность или прямую.*

Мы уже знаем, что *уравнение* (14) *выражает прямую при*

$$A = 0. \tag{15}$$

Изложенное выше позволяет использовать комплексные числа для доказательства многочисленных теорем, касающихся прямых и окружностей. Здесь мы ограничимся единственным примером такого рода.

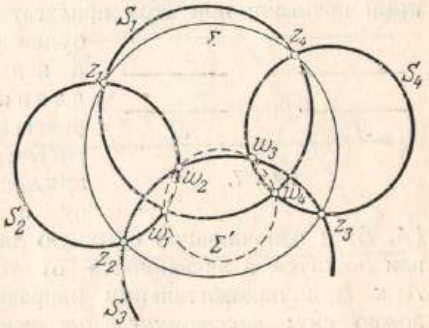


Рис. 6.

Т е о р е м а. *Пусть на плоскости даны четыре окружности S_1, S_2, S_3, S_4; z_1 и w_1 — точки пересечения S_1 и S_2; z_2 и w_2 — точки пересечения S_2 и S_3; z_3 и w_3 — точки пересечения S_3 и S_4; наконец, z_4 и w_4 — точки пересечения S_4 и S_1. Тогда, если точки z_1, z_2, z_3 и z_4 лежат на одной окружности или прямой Σ, то и точки w_1, w_2, w_3 и w_4 лежат на одной окружности или прямой Σ'* (рис. 6).

Д о к а з а т е л ь с т в о. Точки z_1, z_2, w_1 и w_2 лежат на окружности S_2; точки z_2, z_3, w_2 и w_3 — на окружности S_3; точки z_3, z_4, w_3 и w_4 — на окружности S_4; точки z_4, z_1, w_4 и w_1 — на окружности S_1. Отсюда вытекает, что вещественны следующие четыре двойных отношения:

$$W(z_1, w_2, z_2, w_1) = \frac{z_1 - z_2}{w_2 - z_2} : \frac{z_1 - w_1}{w_2 - w_1},$$
$$W(z_2, w_3, z_3, w_2) = \frac{z_2 - z_3}{w_3 - z_3} : \frac{z_2 - w_2}{w_3 - w_2},$$
$$W(z_3, w_4, z_4, w_3) = \frac{z_3 - z_4}{w_4 - z_4} : \frac{z_3 - w_3}{w_4 - w_3},$$
$$W(z_4, w_1, z_1, w_4) = \frac{z_4 - z_1}{w_1 - z_1} : \frac{z_4 - w_4}{w_1 - w_4}.$$

[1]) Предоставим читателю самому убедиться, что радиус r окружности (14) определяется из формулы $r^2 = -\frac{AC + B\overline{B}}{A^2}$, а ее центром служит точка $\frac{\overline{B}}{A}$. Отсюда следует, что при $AC + B\overline{B} = 0$ уравнению (14) удовлетворяет единственная точка $z = \frac{\overline{B}}{A}$ плоскости, а при $AC + B\overline{B} > 0$ ему не удовлетворяет ни одна точка.

Следовательно, вещественно и выражение

$$\frac{W(z_1, w_2, z_2, w_1)\,W(z_3, w_4, z_4, w_3)}{W(z_2, w_3, z_3, w_2)\,W(z_4, w_1, z_1, w_4)} = \left\{\frac{z_1 - z_2}{z_3 - z_2} : \frac{z_1 - z_4}{z_3 - z_4}\right\}\left\{\frac{w_1 - w_2}{w_3 - w_2} : \frac{w_1 - w_3}{w_3 - w_4}\right\} = W(z_1, z_3, z_2, z_4)\,W(w_1, w_2, w_2, w_3).$$

Поэтому из вещественности двойного отношения $W(z_1, z_3, z_2, z_4)$ вытекает также и вещественность двойного отношения $W(w_1, w_3, w_2, w_4)$, что и доказывает теорему.

4. Дуальные числа как ориентированные прямые плоскости. Ниже мы будем иметь дело почти исключительно с ориентированными прямыми; при этом прилагательное «ориентированная» мы часто будем опускать. *Углом* между прямыми a и b мы будем называть ориентированный угол $\angle\{a, b\}$ между ориентированными прямыми (см. выше, стр. 73); *расстоянием между двумя точками* A и B *прямой* l — ориентированную длину отрезка AB, обозначаемую через $\{A, B\}$ и означающую обычную длину, взятую со знаком «плюс» или «минус» в зависимости от того, совпадает ли направление от A к B с положительным направлением прямой l или противоположно ему; *расстоянием от точки* M *до* (ориентированной) *прямой* l — ориентированное расстояние $\{M, l\}$ от M до l, т. е. расстояние, взятое со знаком «плюс» или «минус» в зависимости от того, лежит ли M слева или справа от ориентированной прямой l. Две ориентированные прямые мы будем называть *параллельными* лишь в том случае, если они параллельны в обычном смысле и направления этих прямых совпадают (рис. 7, *а*); параллельные прямые противоположных направлений мы будем иногда называть *противопараллельными* (рис. 7, *б*). Под *расстоянием от прямой* a *до непересекающей ее прямой* b мы будем часто понимать ориентированное расстояние $\{a, b\}$ от a до b, т. е. ориентированное расстояние от произвольной точки прямой a до прямой b; очевидно, что $\{a, b\} = -\{b, a\}$, если a и b параллельны и $\{a, b\} = \{b, a\}$, если a и b противопараллельны.

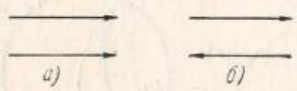


Рис. 7.

Вспомним теперь, что полярные координаты точек плоскости определяются заданием некоторой точки O (полюса системы координат) и проходящей через O (ориентированной) прямой o (полярной оси); координатами точки M служат здесь расстояние $r = OM$ этой точки от полюса и угол $\varphi = \angle\{o, m\}$, образованный с прямой o (ориентированной) прямой m, соединяющей O и M. Аналогично этому можно определить *полярные координаты* (ориентированных) *прямых* плоскости, для задания которых надо также указать некоторую (ориентированную) прямую o (*полярную ось*) и лежащую на o точку O (*полюс*); координатами прямой l служат угол $\theta = \angle\{o, l\}$, образованный l с полярной осью o и (ориентированное) расстояние $s = \{O, L\}$ от O до точки L пересечения l и o (рис. 8, *а*). Очевидно, что координата s ориентированной прямой l может иметь любое значение, заключенное между $+\infty$ и $-\infty$; координата θ — любое значение, заключенное между 0 и 2π. Естественно считать, что $\theta = 0$ для прямых, параллельных полярной оси o, а $\theta = \pi$ для прямых, противопараллельных o; координаты s прямая, не пересекающаяся с o, не имеет (можно считать, что в этом случае $s = \pm\infty$).

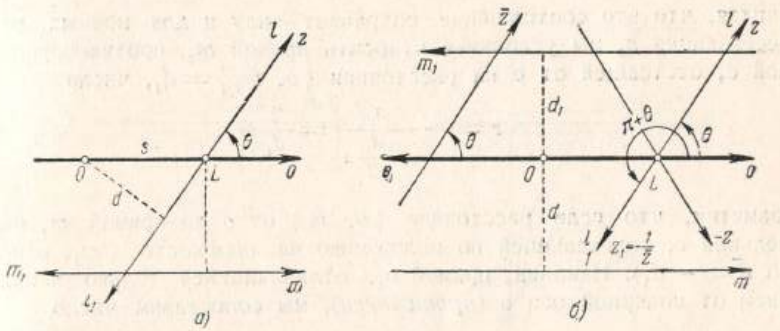


Рис. 8.

Условимся теперь сопоставлять (ориентированной) прямой l с полярными координатами θ и s дуальное число

$$z = \operatorname{tg}\frac{\theta}{2}(1 + \varepsilon s) = u + \varepsilon v, \quad u = \operatorname{tg}\frac{\theta}{2}, \quad v = \operatorname{tg}\frac{\theta}{2}\cdot s. \qquad (16)$$

При этом прямым, параллельным полярной оси o, для которых $\theta = 0$, естественно соотносить числа нулевого модуля, т. е. делителя нуля εv (см. выше, стр. 69).

Для того чтобы установить соответствие между прямыми, параллельными o, и делителями нуля, заметим, что расстояние $d = \{O, l\}$ прямой l, не параллельной o, от полюса O равно

$$d = s\sin\theta = s\frac{2\operatorname{tg}\frac{\theta}{2}}{1 + \operatorname{tg}^2\frac{\theta}{2}} = \frac{2s\operatorname{tg}\frac{\theta}{2}}{1 + \operatorname{tg}^2\frac{\theta}{2}} = \frac{2v}{1 + |z|^2}; \qquad (17)$$

если желать, чтобы эта формула сохранила силу и для прямых m, параллельных o, то такой прямой m, отстоящей от o на расстоя-

ние $\{o, m\}=d$, придется сопоставить число

$$z=\frac{d}{2}\varepsilon \quad \text{(т. е. } z=u+\varepsilon v, \text{ где } u=0 \text{ и } \frac{2v}{1+|z|^2}=2v=d\text{)}.$$

Далее, так как две прямые l и l_1, пересекающие o, отличающиеся только направлением, имеют полярные координаты (θ, s) и $(\pi+\theta, s)$, то им отвечают дуальные числа

$$z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s) \text{ и } z_1=\operatorname{tg}\frac{\pi+\theta}{2}(1+\varepsilon s)=-\operatorname{ctg}\frac{\theta}{2}(1+\varepsilon s)=$$

$$=-\frac{1}{\operatorname{tg}\frac{\theta}{2}(1-\varepsilon s)}=-\frac{1}{z};$$

считая, что это соотношение сохраняет силу и для прямых, не пересекающих o, мы условимся относить прямой m_1, противопараллельной o, отстоящей от o на расстоянии $\{o, m_1\}=d_1$, число

$$z=-\frac{1}{-\frac{d_1}{2}\varepsilon}=\frac{2}{d_1}\omega$$

(заметим, что если расстояние $\{o, m\}$ от o до прямой m, параллельной o, совпадающей по положению на плоскости с m_1, равно d, то $d=-d_1$). Наконец, прямой o_1, отличающейся только направлением от полярной оси o (*противооси*), мы сопоставим число

$$\frac{1}{0}=\infty.$$

Тем самым мы устанавливаем полное соответствие между ориентированными прямыми плоскости и дуальными числами, включая в число последних также числа $v\omega$ ($v\neq 0$ вещественно) и ∞.

Очевидно, что вещественным числам $z=u=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon 0)$ отвечают прямые, проходящие через полюс O; числам модуля 1 — прямые, перпендикулярные o (точнее — такие прямые l, что $\angle\{o, m\}=\frac{\pi}{2}$); вообще числам постоянного модуля u отвечают прямые l, образующие с o постоянный угол $\angle\{o, l\}=2\operatorname{arctg} u$; «чисто мнимым» числам $v\varepsilon$ (числам нулевого модуля) и «числам бесконечного модуля» $v\omega$ отвечают прямые, параллельные и противопараллельные o. Сопряженным числам $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$ и $\bar{z}=\operatorname{tg}\frac{\theta}{2}(1-\varepsilon s)$ отвечают прямые, симметричные относительно полюса O; противоположным числам $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$ и $-z=-\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)=\operatorname{tg}\frac{2\pi-\theta}{2}(1+\varepsilon s)$ — прямые, симметричные относительно полярной оси o (т. е. прямые,

пересекающие o в одной и той же точке L и образующие с o равные углы: $\angle\{o, z\}=\angle\{-z, o\}$ — см. рис. 8, *б*); числам z и $-\frac{1}{z}$ отвечают прямые, отличающиеся только направлением. Таким образом, равенства

$$z'=\bar{z} \text{ (а)}, \quad z'=-z \text{ (б)} \quad \text{и} \quad z'=-\frac{1}{z} \text{ (в)} \tag{18}$$

можно понимать как записи определенных преобразований в множестве ориентированных прямых плоскости: а) *симметрии относительно точки* O, б) *симметрии относительно прямой* o и в) *переориентации* (изменении направления всех прямых плоскости на противоположное).

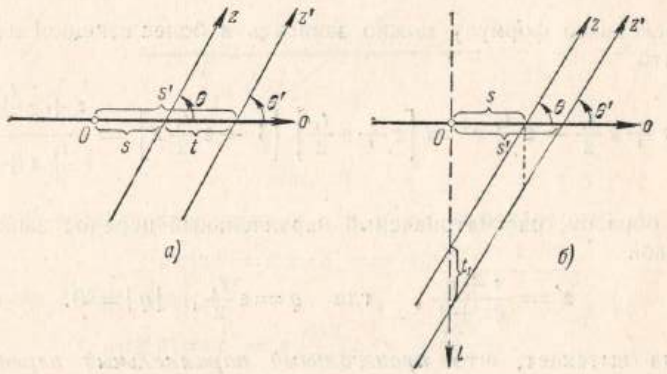


Рис. 9.

Выясним теперь, как записываются с помощью дуальных чисел произвольные *движения* (к числу которых мы будем относить и переориентацию, также не меняющую расстояний между точками плоскости). Прежде всего ясно, что *параллельный перенос вдоль* o на расстояние t переводит прямую, которой отвечает дуальное число $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$, в прямую, которой отвечает число

$$z'=\operatorname{tg}\frac{\theta'}{2}(1+\varepsilon s')=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon(s+t))$$

[рис. 9, *а*; впоследствии мы будем в таком случае кратко говорить: «переводит прямую $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$ в прямую $z'=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon(s+t))$»]. Отсюда вытекает, что этот параллельный перенос можно записать так:

$$z'=pz, \text{ где } p=1+\varepsilon t, \quad |p|=1 \tag{19}$$

(ибо $[1(1+\varepsilon t)]\cdot\left[\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)\right]=\operatorname{tg}\frac{\theta}{2}[1+\varepsilon(s+t)]$). Далее, *парал-*

лельный перенос на расстояние t_1 в направлении, перпендикулярном o (в направлении прямой l, такой, что $\angle\{l, o\}=\frac{\pi}{2}$), переводит прямую $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$ в прямую $z'=\operatorname{tg}\frac{\theta'}{2}(1+\varepsilon s')=$ $=\operatorname{tg}\frac{\theta}{2}[1+\varepsilon(s+t_1\operatorname{ctg}\theta)]$ (рис. 9, *б*). Но

$$z'=\operatorname{tg}\frac{\theta}{2}[1+\varepsilon(s+t_1\operatorname{ctg}\theta)]=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)+\varepsilon\operatorname{tg}\frac{\theta}{2}\frac{t_1\left(1-\operatorname{tg}^2\frac{\theta}{2}\right)}{2\operatorname{tg}\frac{\theta}{2}}=$$
$$=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)+\varepsilon\frac{t_1}{2}-\varepsilon\frac{t_1}{2}\operatorname{tg}^2\frac{\theta}{2}=z+\varepsilon\frac{t_1}{2}-\varepsilon\frac{t_1}{2}z^2.$$

Последнюю формулу можно записать в более изящном виде. Заметим, что

$$z+\varepsilon\frac{t_1}{2}-\varepsilon\frac{t_1}{2}z^2=\left[z+\varepsilon\frac{t_1}{2}\right]\left[1-\varepsilon\frac{t_1}{2}z\right]=\frac{z+\varepsilon\frac{t_1}{2}}{\varepsilon\frac{t_1}{2}z+1};$$

таким образом, рассматриваемый параллельный перенос записывается формулой

$$z'=\frac{z+q}{qz+1},\quad \text{где}\quad q=\varepsilon\frac{t_1}{2},\quad |q|=0. \tag{19a}$$

Отсюда вытекает, что *произвольный параллельный перенос*, т. е. перенос на расстояние t в направлении o и на расстояние t_1 в направлении l, записывается формулой

$$z'=\frac{(pz)+q}{q(pz)+1}\qquad \left(p=1+\varepsilon t,\ q=\varepsilon\frac{t_1}{2}\right)$$

или, если ввести обозначение $p=p_1^2\left(\text{т. е. } p_1=1+\varepsilon\frac{t}{2}\right)$ и воспользоваться тем, что $q=\varepsilon\frac{t_1}{2}=\varepsilon\frac{t_1}{2}\left(1+\varepsilon\frac{t}{2}\right)=qp_1$, $\bar{p}_1=1-\varepsilon\frac{t}{2}=$ $=\frac{1}{p_1}$, $\bar{q}=-q$, то

$$z'=\frac{p_1^2z+qp_1}{qp_1^2z+p_1\bar{p}_1}=\frac{p_1z+q}{qp_1z+\bar{p}_1}=\frac{p_1z+q}{-\bar{q}z+\bar{p}_1}, \tag{20}$$

где

$$p_1=1+\varepsilon\frac{t}{2},\quad q=\varepsilon\frac{t_1}{2};\quad |p_1|=1,\ |q|=0.$$

Перейдем теперь к *вращениям* плоскости. Очевидно, что поворот вокруг O на угол α переводит прямую $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$ в прямую $z'=\operatorname{tg}\frac{\theta'}{2}(1+\varepsilon s')$, где $\theta'=\theta+\alpha$ (рис. 10). Таким образом,

$$|z'|=\operatorname{tg}\frac{\theta+\alpha}{2}=\frac{\operatorname{tg}\frac{\theta}{2}+\operatorname{tg}\frac{\alpha}{2}}{1-\operatorname{tg}\frac{\theta}{2}\operatorname{tg}\frac{\alpha}{2}}=\frac{|z|+\operatorname{tg}\frac{\alpha}{2}}{-\operatorname{tg}\frac{\alpha}{2}|z|+1}=\left|\frac{z+\operatorname{tg}\frac{\alpha}{2}}{-\operatorname{tg}\frac{\alpha}{2}z+1}\right| \tag{21}$$

(здесь используется то, что если z_1 и z_2 — дуальные числа, то $|z_1\pm z_2|=|z_1|\pm|z_2|$, $|z_1z_2|=|z_1||z_2|$ и $\left|\frac{z_1}{z_2}\right|=\frac{|z_1|}{|z_2|}\Big)$. Далее,

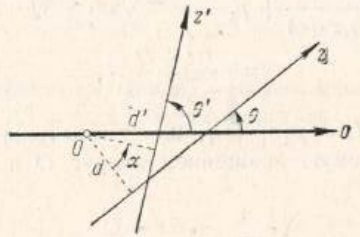


Рис. 10.

если d и d' — расстояния прямых z и z' от полюса O, то

$$s\sin\theta=d=d'=s'\sin\theta'$$

[ср. формулу (17)]; поэтому

$$\operatorname{Arg}z'=s'=s\frac{\sin\theta}{\sin\theta'}=s\frac{\sin\theta}{\sin(\theta+\alpha)}.$$

С другой стороны, поскольку $\operatorname{Arg}(u+\varepsilon v)=\frac{v}{u}$, то

$$\operatorname{Arg}\frac{z+\operatorname{tg}\frac{\alpha}{2}}{-\operatorname{tg}\frac{\alpha}{2}z+1}=\operatorname{Arg}\left(z+\operatorname{tg}\frac{\alpha}{2}\right)-\operatorname{Arg}\left(-\operatorname{tg}\frac{\alpha}{2}z+1\right)=$$
$$=\operatorname{Arg}\left[\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)+\operatorname{tg}\frac{\alpha}{2}\right]-\operatorname{Arg}\left[-\operatorname{tg}\frac{\alpha}{2}\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)+1\right]=$$
$$=\frac{s\operatorname{tg}\frac{\theta}{2}}{\operatorname{tg}\frac{\theta}{2}+\operatorname{tg}\frac{\alpha}{2}}-\frac{-s\operatorname{tg}\frac{\alpha}{2}\operatorname{tg}\frac{\theta}{2}}{1-\operatorname{tg}\frac{\alpha}{2}\operatorname{tg}\frac{\theta}{2}}=\frac{\operatorname{tg}\frac{\theta}{2}\left(\operatorname{tg}^2\frac{\alpha}{2}+1\right)}{\left(\operatorname{tg}\frac{\theta}{2}+\operatorname{tg}\frac{\alpha}{2}\right)\left(1-\operatorname{tg}\frac{\alpha}{2}\operatorname{tg}\frac{\theta}{2}\right)}s=$$
$$=\frac{\operatorname{tg}\frac{\theta}{2}\sec^2\frac{\alpha}{2}\left(\cos\frac{\theta}{2}\cos\frac{\alpha}{2}\right)^2}{\sin\frac{\alpha+\theta}{2}\cos\frac{\alpha+\theta}{2}}s=\frac{\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{\sin\frac{\alpha+\theta}{2}\cos\frac{\alpha+\theta}{2}}s=$$
$$=\frac{\sin\theta}{\sin(\alpha+\theta)}s=\operatorname{Arg}z'. \tag{22}$$

Из (21) и (22) следует, что наше вращение записывается формулой [1])

$$z' = \frac{z + q_1}{-\bar{q}_1 z + 1}, \quad \text{где} \quad q_1 = \bar{q}_1 = \operatorname{tg}\frac{\alpha}{2}, \quad \operatorname{Arg} q_1 = 0. \tag{23}$$

Наконец, *самое общее движение* представляет собой вращение (23) вокруг O на некоторый угол α, сопровождаемое параллельным переносом (20)

$$z' = \frac{p_1 \dfrac{z + q_1}{-\bar{q}_1 z + 1} + q}{-\bar{q}\dfrac{z + q_1}{-\bar{q}_1 z + 1} + \bar{p}_1} = \frac{(p_1 - q\bar{q}_1) z + (p_1 q_1 + q)}{-(\bar{p}_1 \bar{q}_1 + \bar{q}) z + (\bar{p}_1 - \bar{q} q_1)} =$$

$$= \frac{Pz + Q}{-\bar{Q} z + \bar{P}}, \tag{24а}$$

где $P = p_1 - q\bar{q}_1$, $Q = p_1 q_1 + q$, или симметрию (18б) относительно прямой, сопровождаемую вращением вокруг O и параллельным переносом (24а)

$$z' = \frac{-Pz + Q}{\bar{Q} z + \bar{P}}, \tag{24б}$$

или, наконец, переориентацию (18а), сопровождаемую одним из преобразований (24а) или (24б), т. е.

$$z' = \frac{-P\dfrac{1}{z} + Q}{\bar{Q}\dfrac{1}{z} + \bar{P}} = \frac{P_1 \bar{z} + Q_1}{-\bar{Q}_1 z + \bar{P}_1}, \quad \text{где} \quad P_1 = Q, \; Q_1 = -P \tag{24в}$$

или

$$z' = \frac{P\dfrac{1}{z} + Q}{-\bar{Q}\dfrac{1}{z} + \bar{P}} = \frac{-P_1 \bar{z} + Q_1}{\bar{Q}_1 \bar{z} + \bar{P}_1}, \quad \text{где} \quad P_1 = -Q, \; Q_1 = P. \tag{24г}$$

Очевидно, что (ориентированный) *угол* $\delta = \angle\{z_1, z_2\}$ *между прямыми* $z_1 = \operatorname{tg}\frac{\theta_1}{2}(1 + \varepsilon s_1)$ и $z_2 = \operatorname{tg}\frac{\theta_2}{2}(1 + \varepsilon s_2)$ равен $\theta_2 - \theta_1$

[1]) В этой формуле числу $q_1 = \operatorname{tg}\frac{\alpha}{2}$ можно также придать значение ∞, что отвечает вращению на угол 180°. Если $q_1 = \infty$, то имеем

$$z' = \frac{z + \infty}{-\infty \cdot z + 1} = -\frac{1}{z}$$

(см. сноску [2]) на стр. 64); это преобразование, очевидно, сводится к симметрии относительно точки O (18а) и последующей переориентации (18б) [отметим, что определенная выше симметрия (18а) относительно точки O не совпадает с вращением на 180° вокруг O].

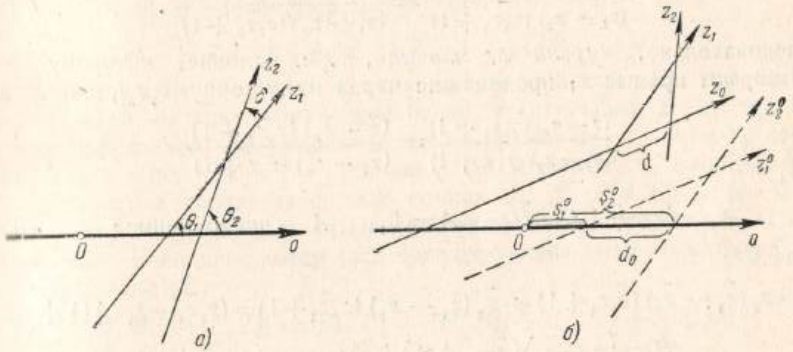
(рис. 11,*а*); это можно записать так:

$$\operatorname{tg}\frac{\delta}{2} = \frac{\operatorname{tg}\frac{\theta_2}{2} - \operatorname{tg}\frac{\theta_1}{2}}{1 + \operatorname{tg}\frac{\theta_2}{2}\operatorname{tg}\frac{\theta_1}{2}} = \frac{|z_2| - |z_1|}{1 + |z_2||z_1|} = \left|\frac{z_2 - z_1}{1 + z_1\bar{z}_2}\right|.$$

Полученный результат можно представить также в следующей симметричной форме:

$$\operatorname{tg}^2\frac{\delta}{2} = \frac{(z_2 - z_1)(\bar{z}_2 - \bar{z}_1)}{(1 + z_1\bar{z}_2)(1 + \bar{z}_1 z_2)}. \tag{25}$$

Найдем теперь (ориентированное) *расстояние* $d = \{[z_1 z_0], [z_2 z_0]\}$ *между точками* $[z_1 z_0]$ *и* $[z_2 z_0]$ *пересечения определенной прямой* z_0



а) *б*)

Рис. 11.

с двумя другими прямыми z_1 *и* z_2 (рис. 11,*б*). Очевидно, что расстояние d_0 между точками пересечения прямой o с прямыми $z_1^0 = \operatorname{tg}\frac{\theta_1^0}{2}(1 + \varepsilon s_1^0)$ и $z_2^0 = \operatorname{tg}\frac{\theta_2^0}{2}(1 + \varepsilon s_2^0)$ равно

$$d_0 = \operatorname{Arg}\frac{z_2^0}{z_1^0} \qquad (= \operatorname{Arg} z_2^0 - \operatorname{Arg} z_1^0 = s_2^0 - s_1^0).$$

Пример движения, переводящего данную прямую z_0 в прямую o, дается формулой

$$z' = \frac{z - z_0}{z_0 z + 1};$$

это движение переводит прямые z_1 и z_2 соответственно в прямые $z_1^0 = \frac{z_1 - z_0}{\bar{z}_0 z_1 + 1}$ и $z_2^0 = \frac{z_2 - z_0}{\bar{z}_0 z_2 + 1}$. Отсюда получаем [1]):

$$d = \operatorname{Arg} \frac{(z_2 - z_0):(\bar{z}_0 z_2 + 1)}{(z_1 - z_0):(\bar{z}_0 z_1 + 1)} = \operatorname{Arg}\left(\frac{z_2 - z_0}{z_1 - z_0} : \frac{\bar{z}_0 z_2 + 1}{\bar{z}_0 z_1 + 1}\right) \tag{26}$$

$$\left(= \operatorname{Arg} \frac{z_2 - z_0}{z_1 - z_0} - \operatorname{Arg} \frac{\bar{z}_0 z_2 + 1}{\bar{z}_0 z_1 + 1}\right).$$

Условием того, что три прямые z_0, z_1 и z_2 пересекаются в одной точке, является равенство нулю расстояния между точками пересечения z_1 и z_2 с z_0, т. е., в силу формулы (26), *вещественность отношения* $\frac{z_0 - z_2}{z_1 - z_2} : \frac{\bar{z}_0 z_0 + 1}{\bar{z}_2 z_1 + 1}$; это условие можно переписать так:

$$\frac{(z_0 - z_2)(\bar{z}_2 z_1 + 1)}{(z_1 - z_2)(\bar{z}_2 z_0 + 1)} = \frac{(\bar{z}_0 - \bar{z}_2)(z_2 \bar{z}_1 + 1)}{(\bar{z}_1 - \bar{z}_2)(z_2 \bar{z}_0 + 1)}. \tag{27}$$

Следовательно, «*уравнение точки*», т. е. условие, которому удовлетворяют прямые z, проходящие через одну точку $[z_1 z_2]$, *имеет вид*

$$\frac{(z - z_2)(\bar{z}_2 z_1 + 1)}{(z_1 - z_2)(\bar{z}_2 z + 1)} = \frac{(\bar{z} - \bar{z}_2)(z_2 \bar{z}_1 + 1)}{(\bar{z}_1 - \bar{z}_2)(z_2 \bar{z} + 1)}$$

или

$$Az\bar{z} + Bz - \bar{B}\bar{z} - A = 0, \qquad A \text{ — чисто мнимое} \tag{28}$$

[здесь

$$A = z_2(\bar{z}_1 - \bar{z}_2)(\bar{z}_2 z_1 + 1) - \bar{z}_2(z_1 - z_2)(z_2 \bar{z}_1 + 1) = (z_2 \bar{z}_1 - z_1 \bar{z}_2)(1 + z_2 \bar{z}_2),$$

$$B = (\bar{z}_1 - \bar{z}_2)(\bar{z}_2 z_1 + 1) + \bar{z}_2^2(z_1 - z_2)(z_2 \bar{z}_1 + 1)].$$

Обратно, нетрудно убедиться, что каждое уравнение вида (28) выражает точку.

Найдем теперь условие того, что четыре ориентированные прямые z_0, z_1, z_2 и z_3 принадлежат одной ориентированной окружности. При этом под *ориентированной окружностью* мы здесь понимаем совокупность («геометрическое место») всех ориентированных прямых l, ориентированное расстояние (O, l) которых от данной точки O (*центра* окружности) имеет фиксированное значение r. Число r называется *радиусом* окружности; таким образом, *радиус ориентированной окружности может быть как положительным, так и отрицательным*. (Если $r = 0$, то ориентированная окружность вырождается в точку, которая, таким образом, является частным случаем окружности.) На чертежах ориентированная окружность изображается как обычная окружность, снабженная стрелкой, указывающей определенное направление обхода, в каждой точке совпадающее с направлением касательной к окружности в этой точке (рис. 12). Из определения ориентированного расстояния (O, l) от точки O до прямой l (стр. 76) следует, что радиус ориентированной окружности будет положительным, если направление обхода противоположно направлению вращения часовой стрелки, и отрицательным — в противном случае.

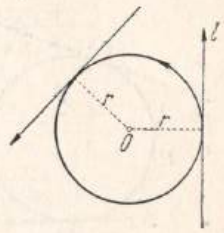


Рис. 12.

Можно показать, что *четыре* (ориентированные) *прямые z_0, z_1, z_2 и z_3 в том и только в том случае принадлежат одной* (ориентированной) *окружности или проходят через одну точку, если*

$$\{[z_0 z_2], [z_1 z_2]\} + \{[z_1 z_3], [z_0 z_3]\} = \{[z_3 z_0], [z_2 z_0]\} + \{[z_2 z_1], [z_3 z_1]\}^{1}). \tag{29}$$

Для того чтобы убедиться в этом, рассмотрим рис. 13, на котором изображены четыре (ориентированные) касательные z_0, z_1, z_2 и z_3 (ориентированной) окружности S (если прямые z_0, z_1, z_2 и z_3 проходят через одну точку, то условие (29), разумеется, выполняется!), касающиеся S соответственно в точках M, N, P и Q; точки $[z_0 z_2]$, $[z_1 z_2]$, $[z_1 z_3]$ и $[z_0 z_3]$ для краткости обозначены через A, B, C и D. При этом, очевидно, имеем (все фигурирующие ниже расстояния ориентированные!) [2]):

$$\{A, B\} + \{C, D\} = \{A, P\} + \{P, B\} + \{C, Q\} + \{Q, D\}$$

и

$$\{D, A\} + \{B, C\} = \{D, M\} + \{M, A\} + \{B, N\} + \{N, C\}.$$

Но так как, в силу известного свойства касательных к окружности, $\{A, P\} = \{M, A\}$, $\{P, B\} = \{B, N\}$, $\{C, Q\} = \{N, C\}$, $\{Q, D\} = \{D, M\}$, то во всех случаях выполняется условие (29):

$$\{A, B\} + \{C, D\} = \{D, A\} + \{B, C\}^{3}).$$

---

[1]) Если одна из прямых z_1, z_2 (или даже обе эти прямые) не пересекает z_0, то одно из чисел $(z_2 - z_0):(\bar{z}_0 z_2 + 1)$ и $(z_1 - z_0):(\bar{z}_0 z_1 + 1)$ (или оба эти числа) не приводится к форме (25) п. 2.

[2]) Здесь используется то, что если X, Y, Z — три точки ориентированной прямой, то в о в с е х с л у ч а я х

$$\{X, Y\} + \{Y, Z\} = \{X, Z\}.$$

[3]) Это есть точная формулировка известной теоремы школьного курса геометрии о равенстве сумм противоположных сторон описанного четырех-

Нетрудно убедиться, что и обратно, *если равенство* (29) *имеет место, то четыре прямые* z_0, z_1, z_2 *и* z_3 *принадлежат одной ориентированной окружности или проходят через одну точку* [1]).

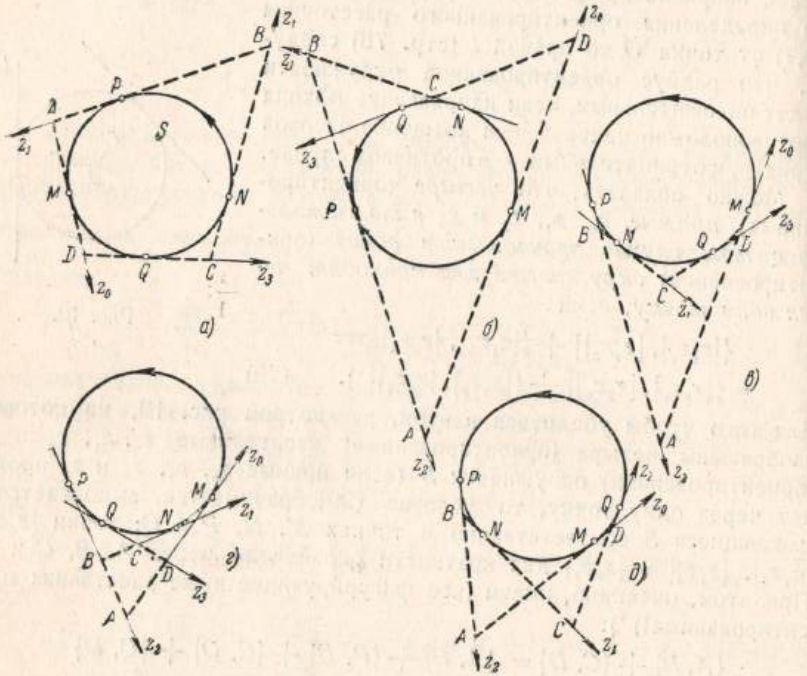


Рис. 13.

Воспользовавшись теперь формулой (26), мы можем переписать условие (29) следующим образом:

$$\operatorname{Arg}\frac{z_1-z_2}{z_0-z_2}-\operatorname{Arg}\frac{\bar z_2 z_1+1}{z_2 z_0+1}+\operatorname{Arg}\frac{z_0-z_3}{z_1-z_3}-\operatorname{Arg}\frac{\bar z_3 z_0+1}{z_3 z_1+1}=$$

$$=\operatorname{Arg}\frac{z_2-z_0}{z_3-z_0}-\operatorname{Arg}\frac{\bar z_0 z_2+1}{z_0 z_3+1}+\operatorname{Arg}\frac{z_2-z_1}{z_2-z_1}-\operatorname{Arg}\frac{\bar z_1 z_3+1}{z_1 z_2+1}.$$

угольника. Отметим, что отказ от ориентированных расстояний ведет к значительному усложнению формулировки этой теоремы: в случаях рис. 13, *а* и *б* мы приходим к равенству $AB+CD=AD+BC$ (расстояния не ориентированные!), а в случаях рис. 13, *в*, *г* и *д* — к равенству $AB-CD=AD-BC$.

[1]) Ср., например, Ж. Адамар, Элементарная геометрия, ч. 1, М., Учпедгиз, 1957, решение задачи 87.

Несколько упростив левую часть последнего равенства и преобразовав правую, будем иметь

$$\operatorname{Arg}\left(\frac{z_1-z_2}{z_0-z_2}:\frac{z_1-z_3}{z_0-z_3}\right)-\operatorname{Arg}\left(\frac{\bar z_2 z_1+1}{\bar z_2 z_0+1}:\frac{\bar z_3 z_1+1}{\bar z_3 z_0+1}\right)=$$

$$=\operatorname{Arg}\left(\frac{z_0-z_2}{z_1-z_2}:\frac{z_0-z_3}{z_1-z_3}\right)+\operatorname{Arg}\left(\frac{z_2\bar z_1+1}{z_2\bar z_0+1}:\frac{z_3\bar z_1+1}{z_3\bar z_0+1}\right).$$

Но

$$\operatorname{Arg}\left(\frac{z_1-z_2}{z_0-z_2}:\frac{z_1-z_3}{z_0-z_3}\right)=-\operatorname{Arg}\left(\frac{z_0-z_2}{z_1-z_2}:\frac{z_0-z_3}{z_1-z_3}\right)$$

$$\left[\text{ибо }\frac{z_1-z_2}{z_0-z_2}:\frac{z_1-z_3}{z_0-z_3}=1:\left(\frac{z_0-z_2}{z_1-z_2}:\frac{z_0-z_3}{z_1-z_3}\right)\right]$$

и

$$\operatorname{Arg}\left(\frac{\bar z_2 z_1+1}{\bar z_2 z_0+1}:\frac{\bar z_3 z_1+1}{\bar z_3 z_0+1}\right)=-\operatorname{Arg}\left(\frac{z_2\bar z_1+1}{z_2\bar z_0+1}:\frac{z_3\bar z_1+1}{z_3\bar z_0+1}\right)$$

$$\left[\text{ибо }\frac{\bar z_2 z_1+1}{\bar z_2 z_0+1}:\frac{\bar z_3 z_1+1}{\bar z_3 z_0+1}=\overline{\left(\frac{z_2\bar z_1+1}{z_2\bar z_0+1}:\frac{z_3\bar z_1+1}{z_3\bar z_0+1}\right)}\right].$$

Таким образом, равенство (29) можно переписать в следующей простой форме:

$$\operatorname{Arg}\left(\frac{z_0-z_2}{z_1-z_2}:\frac{z_0-z_3}{z_1-z_3}\right)=0. \tag{30}$$

Другими словами, *условием того, что четыре прямые* z_0, z_1, z_2 *и* z_3 *принадлежат одной* (ориентированной) *окружности* (ненулевого радиуса или окружности радиуса нуль — точке), *является вещественность двойного отношения* $W(z_0, z_1, z_2, z_3)=\dfrac{z_0-z_2}{z_1-z_2}:\dfrac{z_0-z_3}{z_1-z_3}$ *этих четырех прямых* (ср. выше, стр. 74).

Последнему условию можно также придать знакомую уже нам форму (12):

$$\frac{z_0-z_2}{z_1-z_2}:\frac{z_0-z_3}{z_1-z_3}=\frac{\bar z_0-\bar z_2}{\bar z_1-\bar z_2}:\frac{\bar z_0-\bar z_3}{\bar z_1-\bar z_3}, \tag{12}$$

откуда вытекает, что уравнение ориентированной окружности (которая в частном случае может оказаться и точкой), определяемой тремя данными прямыми z_1, z_2 и z_3, имеет вид (13) [1]):

$$\frac{z-z_2}{z_1-z_2}:\frac{z-z_3}{z_1-z_3}=\frac{\bar z-\bar z_2}{\bar z_1-\bar z_2}:\frac{\bar z-\bar z_3}{\bar z_1-\bar z_3}. \tag{13}$$

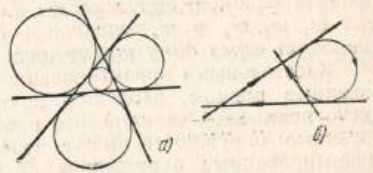


Рис. 14.

[1]) Заметим, что в то время как трех данных неориентированных прямых касаются, вообще говоря, четыре разные неориентированные окружности (рис. 14, *а*), три ориентированные прямые определяют единственную ориентированную окружность (рис. 14, *б*).

Таким образом, мы убеждаемся, что *уравнение каждой* (ориентированной) *окружности* (или точки) *плоскости можно записать в форме* (14):

$$Az\bar{z}+Bz-\bar{B}\bar{z}+C=0, \quad A \text{ и } C \text{ — чисто мнимые;} \qquad (14)$$

можно убедиться, что и обратно, уравнение (14) всегда выражает окружность (или точку).

Мы уже знаем, что прямую уравнение (14) выражает при

$$A+C=0. \qquad (31)$$

Развитая выше теория позволяет использовать дуальные числа для доказательства многочисленных теорем, относящихся к точкам, прямым и окружностям; при этом близость результатов настоящего пункта к результатам п. 3 позволяет зачастую использовать одну и ту же выкладку для доказательства двух различных предложений, считая в одном случае фигурирующие в рассуждении числа обыкновенными комплексными числами, а во втором случае — дуальными числами. Мы ограничимся здесь одним примером подобного рода.

Теорема. *Пусть на плоскости даны четыре* (ориентированные) *окружности* S_1, S_2, S_3 *и* S_4; z_1 *и* w_1 — *общие касательные* S_1 *и* S_2; z_2 *и* w_2 — *общие касательные* S_2 *и* S_3; z_3 *и* w_3 — *общие касательные* S_3 *и* S_4; z_4 *и* w_4 — *общие касательные* S_4 *и* S_1. *Тогда если прямые* z_1, z_2, z_3 *и* z_4 *касаются одной* (ориентированной) *окружности или проходят через одну точку, то и прямые* w_1, w_2, w_3 *и* w_4 *касаются одной* (ориентированной) *окружности или проходят через одну точку* (рис. 15).

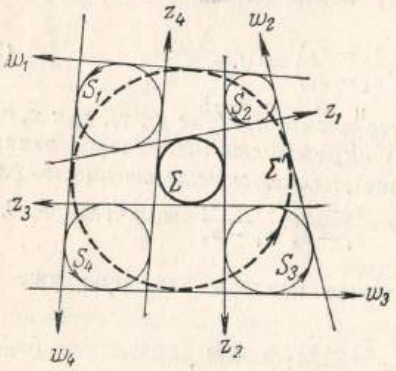


Рис. 15.

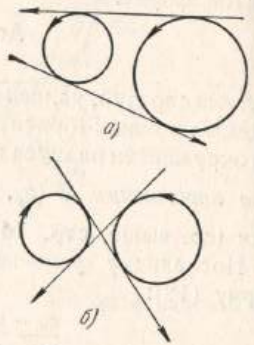


Рис. 16.

[Касательными ориентированной окружности S здесь считаются ориентированные прямые, входящие в состав S, понимаемой как «геометрическое место прямых». Очевидно, что в то время, как две обыкновенные (неориентированные) окружности имеют, вообще говоря, четыре общие касательные, две ориентированные окружности S_1 и S_2 не могут иметь больше двух общих касательных: это будут общие внешние касательные, если S_1 и S_2 имеют одно и то же направление, и общие внутренние касательные — в противном случае (рис. 16, *а, б*).]

Из условия теоремы вытекает, что вещественны двойные отношения $W(z_1, w_2, z_2, w_1)$, $W(z_2, w_3, z_3, w_2)$, $W(z_3, w_4, z_4, w_3)$ и $W(z_4, w_1, z_1, w_4)$, а также двойное отношение $W(z_1, z_3, z_2, z_4)$; требуется доказать вещественность двойного отношения $W(w_1, w_3, w_2, w_4)$ (см. выше стр. 75—76). Доказательство этого было дано в конце п. 3 (стр. 75); приведенные там рассуждения одновременно доказывают и теорему, сформулированную на стр. 75, и настоящую теорему.

III. КРУГОВЫЕ ПРЕОБРАЗОВАНИЯ

5. Обыкновенные круговые преобразования (преобразования Мёбиуса). В этом пункте рассмотрим произвольные дробно-линейные функции комплексного переменного и отвечающие им, в силу изложенного в п. **3**, *дробно-линейные преобразования плоскости*, т. е. преобразования, записываемые формулами [1])

$$z'=\frac{az+b}{cz+d} \qquad (1) \qquad \text{и} \quad z'=\frac{a\bar{z}+b}{c\bar{z}+d}. \qquad (1a)$$

Если $\Delta=\begin{vmatrix} a & b \\ c & d \end{vmatrix}=0$, т. е. $a=kc$, $b=kd$, функции (1) и (1а) сводятся к $z'=k$; таким образом, интерес представляет лишь случай, когда $\Delta=\begin{vmatrix} a & b \\ c & d \end{vmatrix}\neq 0$; ниже мы все время будем предполагать последнее условие выполненным.

При этом преобразования (1) и (1а) будут являться взаимно однозначными преобразованиями плоскости комплексного переменного, расширенной введением числа $\frac{1}{0}=\infty$, — каждому числу z соответствует единственное число z', определяемое формулой (1) или (1а) (в частности, числу $z=\infty$ отвечает значение $z'=\frac{a}{c}$ [2]), а числу $z=-\frac{d}{c}$, соответственно $z=-\frac{\bar{d}}{\bar{c}}$, такому, что $cz+d=0$ или $c\bar{z}+d=0$, отвечает $z'=\infty$), и каждому значению z' отвечает единственное значение z, которое легко найти из тех же формул (1) и (1а):

$$z=\frac{dz'-b}{-cz'+a} \quad \text{или} \quad z=\frac{\bar{d}\bar{z}'-\bar{b}}{-\bar{c}\bar{z}'+\bar{a}}. \qquad (2)$$

Частным случаем дробно-линейных преобразований являются *линейные преобразования*

$$z'=pz+q, \quad p\neq 0 \quad \text{или} \quad z'=p\bar{z}+q, \quad p\neq 0, \qquad (3)$$

[1]) Ср., например, А. И. Маркушевич, Элементы теории аналитических функций, М., 1944, гл. 5, § 3.

[2]) Ср. сноску [2]) на стр. 64.

к которым мы приходим, положив в формулах (1) и (1а) $c=0$ и полагая $p=\frac{a}{d}$, $q=\frac{b}{d}$. Иногда преобразования (1) называют *собственными дробно-линейными преобразованиями*, а преобразования (1а) — *зеркальными дробно-линейными преобразованиями*.

Выше мы уже видели, что геометрически линейное преобразование представляет собой *преобразование подобия*, складывающееся из центрально-подобного вращения (гомотетии и вращения с общим центром O), сопровождаемого параллельным переносом и, быть может, симметрией относительно прямой; в частности, при $|p|=1$ преобразование подобия является движением (см. п. **3**, стр. 72). Мы рассматривали также некоторые простейшие конкретные примеры линейных преобразований, как то преобразования

$$z'=-z \quad \text{(а)} \quad \text{и} \quad z'=\bar{z} \quad \text{(б)}, \tag{4}$$

представляющие собой *симметрию относительно точки O и симметрию относительно прямой o*, а также преобразования

$$z'=z+q \quad \text{(а)} \quad \text{и} \quad z'=pz \quad \text{(б)}, \tag{5}$$

т. е. *параллельный перенос и центрально-подобное вращение* (п. **3**, стр. 71). Здесь мы более подробно изучим геометрические свойства произвольных дробно-линейных преобразований.

Заметим прежде всего, что *дробно-линейные преобразования составляют группу* в том смысле, что произведение (результат последовательного осуществления) двух дробно-линейных преобразований также является дробно-линейным преобразованием; тождественное (или единичное) преобразование, оставляющее все точки плоскости на месте, является частным случаем дробно-линейного преобразования; преобразование, обратное дробно-линейному (т. е. переводящее каждую точку z' плоскости в ту точку z, из которой получалась z' в результате исходного преобразования), также является дробно-линейным. Действительно, если, например,

$$z_1=\frac{az+b}{cz+d} \quad \text{и} \quad z'=\frac{a_1z_1+b_1}{c_1z_1+d_1}, \tag{6а}$$

то

$$z'=\frac{a_1\dfrac{az+b}{cz+d}+b_1}{c_1\dfrac{az+b}{cz+d}+d_1}=\frac{(a_1a+b_1c)z+(a_1b+b_1d)}{(c_1a+d_1c)z+(c_1b+d_1d)}. \tag{6б}$$

Аналогично находится произведение собственного и зеркального дробно-линейных преобразований или двух зеркальных дробно-линейных преобразований; тождественное преобразование записывается формулой

$$z'=z, \tag{7}$$

являющейся частным случаем формулы (1) (при $b=c=0$, $a=d=1$); преобразование, обратное (1) или (1а), имеет вид

$$z'=\frac{dz-b}{-cz+a} \quad \text{или} \quad z'=\frac{\bar{d}\,\bar{z}-\bar{b}}{-\bar{c}\,\bar{z}+\bar{a}} \tag{8}$$

[ср. с формулами (2)].

Отметим теперь следующее фундаментальное свойство дробно-линейных преобразований: *если z_1', z_2', z_3' и z_4' суть четыре точки плоскости, в которые переводит дробно-линейное преобразование* (1) *или* (1а) *данные четыре точки z_1, z_2, z_3 и z_4, то*

$$W(z_1', z_2', z_3', z_4')=W(z_1, z_2, z_3, z_4)$$

или

$$W(z_1', z_2', z_3', z_4')=\overline{W(z_1, z_2, z_3, z_4)}, \tag{9}$$

где $W(z_1, z_2, z_3, z_4)=\frac{z_1-z_3}{z_2-z_3}:\frac{z_1-z_4}{z_2-z_4}$ — двойное отношение четырех точек (свойство инвариантности двойного отношения). Действительно, например, из формулы (1) получаем

$$W(z_1', z_2', z_3', z_4')=$$

$$=\frac{z_1'-z_3'}{z_2'-z_3'}:\frac{z_1'-z_4'}{z_2'-z_4'}=\frac{\dfrac{az_1+b}{cz_1+d}-\dfrac{az_3+b}{cz_3+d}}{\dfrac{az_2+b}{cz_2+d}-\dfrac{az_3+b}{cz_3+d}}:\frac{\dfrac{az_1+b}{cz_1+d}-\dfrac{az_4+b}{cz_4+d}}{\dfrac{az_2+b}{cz_2+d}-\dfrac{az_4+b}{cz_4+d}}=$$

$$=\frac{[(ad-bc)(z_1-z_3)]:(cz_1+d)}{[(ad-bc)(z_2-z_3)]:(cz_2+d)}:\frac{[(ad-bc)(z_1-z_4)]:(cz_1+d)}{[(ad-bc)(z_2-z_4)]:(cz_2+d)}=$$

$$=\frac{z_1-z_3}{z_2-z_3}:\frac{z_1-z_4}{z_2-z_4}=W(z_1, z_2, z_3, z_4).$$

Аналогично проверяется, что $W(z_1', z_2', z_3', z_4')=\overline{W(z_1, z_2, z_3, z_4)}$, если z и z' связаны формулой (1а).

Из свойства инвариантности двойного отношения прежде всего вытекает, что *дробно-линейное преобразование переводит четыре точки, принадлежащие одной окружности или прямой линии, в четыре точки, также принадлежащие окружности или прямой линии* (круговое свойство дробно-линейных преобразований); действительно, из вещественности двойного отношения $W(z_1, z_2, z_3, z_4)$ исходных точек вытекает также вещественность двойного отношения $W(z_1', z_2', z_3', z_4')$ преобразованных точек (ср. п. **3**, стр. 74). Отсюда в свою очередь вытекает, что *дробно-линейное преобразование переводит каждую окружность или прямую линию плоскости снова в окружность или прямую линию*. Это обстоятельство дает основание для того, чтобы называть дробно-

линейные преобразования плоскости *круговыми преобразованиями* [1]); так как эти преобразования впервые основательно изучались немецким геометром Августом Фердинандом Мёбиусом, то их часто называют круговыми преобразованиями Мёбиуса. Можно также говорить о *собственных* и *зеркальных* круговых преобразованиях, понимая под этим преобразования (1), соответственно (1а).

Нетрудно показать, что *существует единственное собственное дробно-линейное преобразование* (1), *переводящее три данные точки* z_1, z_2 и z_3 *в три другие заданные точки* w_1, w_2 *и* w_3. Действительно, если преобразование (1) переводит точки z_1, z_2, z_3 в точки w_1, w_2, w_3 и произвольную точку z плоскости в точку z', то в силу доказанного выше $W(z, z_1, z_2, z_3) = W(z', w_1, w_2, w_3)$, т. е.

$$\frac{z'-w_2}{w_1-w_2}:\frac{z'-w_3}{w_1-w_3}=\frac{z-z_2}{z_1-z_2}:\frac{z-z_3}{z_1-z_3}. \qquad (10)$$

Но равенство (10) определяет дробно-линейное преобразование: если выразить из него z' через z, то мы получим

$$z'=\frac{Az+B}{Cz+D},$$

где

$$A=(z_1-z_2)w_1w_2+(z_2-z_3)w_2w_3+(z_3-z_1)w_3w_1,$$
$$B=[(z_3-z_2)w_3w_2+(w_2-w_3)w_2z_2+w_1(z_2w_3-z_3w_2)](z_1-z_3),$$
$$C=(w_1z_2-w_2z_1)+(w_2z_3-w_3z_2)+(w_3z_1-w_1z_3),$$
$$D=(w_1-w_3)(z_1-z_3)(z_2-z_3).$$

Точно так же доказывается, что *существует единственное зеркальное дробно-линейное преобразование* (1а), *переводящее* z_1, z_2 *и* z_3 *в* w_1, w_2 *и* w_3; это преобразование задается формулой

$$W(z', w_1, w_2, w_3)=\overline{W(z, z_1, z_2, z_3)}$$

или

$$\frac{z'-w_2}{w_1-w_2}:\frac{z'-w_3}{w_1-w_3}=\frac{\bar{z}-\bar{z}_2}{\bar{z}_1-\bar{z}_2}:\frac{\bar{z}-\bar{z}_3}{\bar{z}_1-\bar{z}_3}. \qquad (10а)$$

Эти же рассуждения показывают, что *для того, чтобы четыре данные точки* z_1, z_2, z_3 *и* z_4 *можно было перевести круговым преобразованием в четыре другие точки* w_1, w_2, w_3 *и* w_4, *необходимо и достаточно, чтобы имело место одно из равенств*:

$$W(w_1, w_2, w_3, w_4)=\overline{W(z_1, z_2, z_3, z_4)}$$

или

$$W(w_1, w_2, w_3, w_4)=\overline{W(z_1, z_2, z_3, z_4)}.$$

[1]) Можно показать, что все преобразования плоскости, переводящие окружности или прямые снова в окружности или прямые, исчерпываются дробно-линейными преобразованиями (1) и (1а) (ср. И. М. Яглом, Геометрические преобразования II, М., 1956, § 4 гл. II, стр. 246—253).

Из доказанного следует, что *любую окружность или прямую можно перевести* (причем многими способами) *круговым преобразованием* (1) [или (1а)] *в любую другую заданную окружность или прямую*; действительно, для этого надо только позаботиться, чтобы какие-либо три точки первой окружности перешли в (какие угодно!) три точки второй окружности. В частности, *любую окружность можно бесчисленным множеством способов перевести в прямую линию* — обстоятельство, которое часто оказывается полезным. Таким образом, с точки зрения круговых преобразований все окружности и прямые являются совершенно равноправными; поэтому в вопросах, связанных с круговыми преобразованиями, обычно не различают между собой прямые линии и окружности, считая первые частными случаями последних («окружности бесконечно большого радиуса»). Также и мы в дальнейшем зачастую будем говорить просто «окружность» в тех случаях, когда правильнее было бы сказать «окружность или прямая».

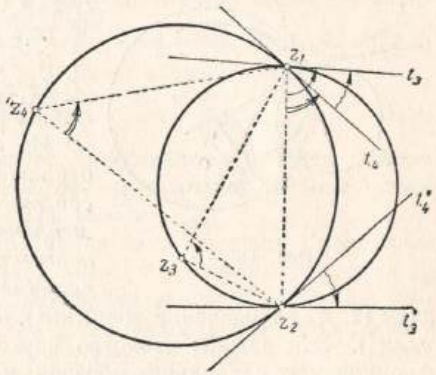


Рис. 17.

Рассмотрим теперь более подробно, какой геометрический смысл имеет двойное отношение $W(z_1, z_2, z_3, z_4)$ четырех точек z_1, z_2, z_3 и z_4. Мы уже знаем, что это двойное отношение будет вещественно, т. е. аргумент $\operatorname{Arg} W$ двойного отношения W будет равен нулю в том и только в том случае, если точки z_1, z_2, z_3 и z_4 лежат на одной окружности. В общем же случае, очевидно, имеем

$$\operatorname{Arg} W(z_1, z_2, z_3, z_4)=\operatorname{Arg}\left(\frac{z_1-z_3}{z_2-z_3}:\frac{z_1-z_4}{z_2-z_4}\right)=$$
$$=\operatorname{Arg}\frac{z_1-z_3}{z_2-z_3}-\operatorname{Arg}\frac{z_1-z_4}{z_2-z_4}=\angle\{z_3z_2,\ z_3z_1\}-\angle\{z_4z_2,\ z_4z_1\}$$

[см. формулу (8) на стр. 73]. Рассмотрим теперь две окружности (одна из них, или даже обе они, могут обратиться в прямую), проходящие через точки z_1, z_2, z_3 и z_1, z_2, z_4; эти окружности мы будем обозначать через $[z_1z_2z_3]$ и $[z_1z_2z_4]$ (рис. 17; вообще окружность, проходящую через три точки z, v и w, мы будем называть окружностью $[zvw]$). Проведем еще касательные t_3 и t_4 к этим окружностям в точке z_1. Из известных теорем о вписанных углах следует, что независимо от положения точек z_3 и z_4 на окружностях $[z_1z_2z_3]$ и $[z_1z_2z_4]$

$$\angle\{[z_3z_2],\ [z_3z_1]\}=\angle\{[z_1z_2],\ t_3\}\ \text{и}\ \angle\{[z_4z_2],\ [z_4z_1]\}=\angle\{[z_1z_2],\ t_4\};$$

поэтому имеем

$$\operatorname{Arg} W(z_1, z_2, z_3, z_4) = \angle\{[z_1z_2], t_3\} - \angle\{[z_1z_2], t_4\} = \angle\{t_4, t_3\}. \quad (11)$$

Угол между касательными к двум окружностям S_1 и S_2, проведенными в точке z их пересечения, называется *углом между окружностями* S_1 *и* S_2 и обозначается через $\angle(S_1, S_2)$; если рассматривается ориентированный угол между касательными, то говорят об *ориентированном угле* $\angle\{S_1zS_2\}$ *между окружностями* S_1 *и* S_2. Таким образом, мы видим, что *аргумент* $\operatorname{Arg} W(z_1, z_2, z_3, z_4)$ *двойного отношения* $W(z_1, z_2, z_3, z_4)$ *четырех точек* z_1, z_2, z_3 *и* z_4 *равен* (ориентированному) *углу* $\angle\{[z_1z_2z_4]z_1[z_1z_2z_3]\}$ *между окружностями* $[z_1z_2z_4]$ *и* $[z_1z_2z_3]$.

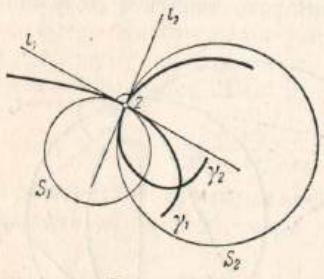


Рис. 18.

Из свойства инвариантности двойного отношения четырех точек вытекает, что *собственные круговые преобразования не меняют ориентированного угла между пересекающимися окружностями, а зеркальные круговые преобразования меняют знак* (т. е. направление вращения), *но не абсолютную величину этого угла*[1]). Это важное свойство круговых преобразований часто кратко формулируют следующим образом: *углы между окружностями сохраняются при круговых преобразованиях*[2]). В частности, *касающиеся окружности* (окружности, угол между которыми равен нулю) *переходят при круговом преобразовании в касающиеся окружности.*

---

[1]) Другими словами, если собственное (зеркальное) круговое преобразование переводит окружности S_1 и S_2 в окружности S_1' и S_2' и точку z_1 пересечения S_1 и S_2 в точку z_1', то $\angle\{S_1z_1S_2\} = \angle\{S_1'z_1'S_2'\}$, соответственно $\angle\{S_1z_1S_2\} = -\angle\{S_1'z_1'S_2'\}$. [Заметим, что если две окружности S_1 и S_2 пересекаются в точках z_1 и z_2, то $\angle\{S_1z_1S_2\} = -\angle\{S_1z_2S_2\}$ (ср. рис. 17); поэтому, говоря об ориентированном угле между двумя пересекающимися окружностями, необходимо указывать точку пересечения, в которой рассматривается этот угол (неориентированный угол между окружностями не зависит от выбора точки их пересечения).]

[2]) Так как угол между двумя любыми кривыми γ_1 и γ_2, пересекающимися в точке z_1 (по определению этот угол совпадает с углом между касательными к γ_1 и γ_2 в точке z_1), равен углу между проходящими через z_1 окружностями S_1 и S_2, касающимися наших кривых (рис. 18), и круговое преобразование, переводящее кривые γ_1 и γ_2 в новые кривые γ_1' и γ_2', переводит S_1 и S_2 в окружности S_1' и S_2', касающиеся γ_1' и γ_2', то отсюда следует, что *углы между произвольными кривыми сохраняются при круговых преобразованиях.* Все преобразования, обладающие этим последним свойством, называются конформными преобразованиями. Таким образом, *круговые преобразования являются конформными преобразованиями.*

Перейдем теперь к модулю $|W(z_1, z_2, z_3, z_4)|$ двойного отношения W четырех точек плоскости. В силу формулы (6) на стр. 72 имеем

$$|W(z_1, z_2, z_3, z_4)| = \left|\frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4}\right| = \frac{|z_1 - z_3|}{|z_2 - z_3|} : \frac{|z_1 - z_4|}{|z_2 - z_4|} = \frac{(z_1z_3)}{(z_2z_3)} : \frac{(z_1z_4)}{(z_2z_4)}, \quad (12)$$

где (z_1z_2), (z_2z_3) и т. д. — расстояния между точками z_1 и z_2, z_2 и z_3 и т. д. Вещественное число $\frac{(z_1z_3)}{(z_2z_3)} : \frac{(z_1z_4)}{(z_2z_4)}$ мы будем называть *двойным отношением расстояний между четырьмя точками* z_1, z_2, z_3 *и* z_4 и обозначать через $\bar{W}(z_1, z_2, z_3, z_4)$; таким образом, имеем

$$|W(z_1, z_2, z_3, z_4)| = \bar{W}(z_1, z_2, z_3, z_4). \quad (12a)$$

Из свойства инвариантности двойного отношения мы можем теперь заключить, что *круговые преобразования сохраняют двойные отношения расстояний между четверками точек.*

Теперь мы можем по-новому сформулировать *условия, необходимые и достаточные для того, чтобы четверку точек* z_1, z_2, z_3 *и* z_4 *можно было перевести круговым преобразованием в четверку точек* w_1, w_2, w_3 *и* w_4. А именно: для этого надо, чтобы *угол между окружностями* $[z_1z_2z_3]$ *и* $[z_1z_2z_4]$ *был равен углу между окружностями* $[w_1w_2w_3]$ *и* $[w_1w_2w_4]$ *и двойное отношение расстояний между точками* z_1, z_2, z_3 *и* z_4 *равнялось двойному отношению расстояний между точками* w_1, w_2, w_3 *и* w_4. Так как $\angle\{[z_1z_2z_3]z_1[z_1z_2z_4]\} = \angle\{[z_2z_2], [z_2z_1]\} - \angle\{[z_4z_2], [z_4z_1]\}$ и $W(z_1, z_2, z_3, z_4) = \frac{(z_1z_3)(z_2z_4)}{(z_2z_3)(z_1z_4)}$, то можно, в частности, сказать, что *для того, чтобы выпуклый четырехугольник* $z_1z_3z_2z_4$ *мог быть переведен круговым преобразованием в выпуклый четырехугольник* $w_1w_3w_2w_4$, *необходимо и достаточно, чтобы сумма противоположных углов* z_3 *и* z_4 *первого четырехугольника равнялась сумме противоположных углов* w_3 *и* w_4 *второго четырехугольника и отношение* $\frac{(z_1z_3)(z_2z_4)}{(z_2z_3)(z_1z_4)}$ *произведений противоположных сторон первого четырехугольника равнялось отношению* $\frac{(w_1w_3)(w_2w_4)}{(w_2w_3)(w_1w_4)}$ *произведений противоположных сторон второго четырехугольника* (заметим, что ориентированные углы $\angle\{[z_3z_2], [z_3z_1]\}$ и $\angle\{[z_4z_2], [z_4z_1]\}$ направлены в разные стороны, и поэтому их разность сводится к сумме положительных углов z_3 и z_4 четырехугольника, см. рис. 19)[1]).

---

[1]) Из доказанного вытекает, что *каждый выпуклый четырехугольник* $z_1z_3z_2z_4$ *можно круговым преобразованием перевести в параллелограмм* $z_1^0z_3^0z_2^0z_4^0$, *углы которого равны полусуммам противоположных углов* $z_1z_2z_3z_4$,

Коснемся еще вопроса о геометрическом описании произвольных круговых преобразований. Как мы знаем, линейные преобразования (3) представляют собой *преобразования подобия*. При этом и произведение двух или нескольких линейных преобразований снова является линейным преобразованием[1]); поэтому все круговые преобразования нельзя свести к одним лишь преобразованиям подобия.

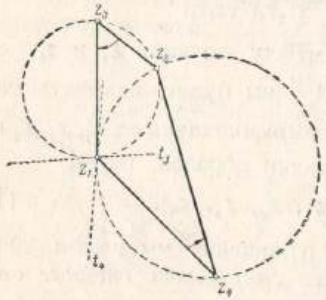


Рис. 19.

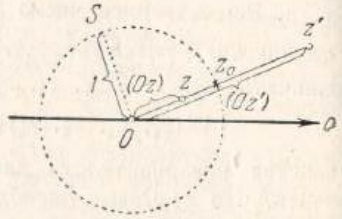


Рис. 20.

Простейшими круговыми преобразованиями, отличными от преобразований подобия, являются преобразования

$$z' = \frac{1}{z} \quad \text{(а)} \quad \text{и} \quad z' = \frac{1}{\bar{z}} \quad \text{(б)}, \tag{13}$$

которые можно также описать равенствами

$$\operatorname{Arg} z' = -\operatorname{Arg} z,\ |z'| = \frac{1}{|z|} \quad \text{(а)} \quad \text{и} \quad \operatorname{Arg} z' = \operatorname{Arg} z,\ |z'| = \frac{1}{|z|}. \quad \text{(б)} \tag{14}$$

Из этих двух преобразований более простой геометрический смысл имеет преобразование (13б), называемое *единичной инверсией*. При инверсии каждая точка z плоскости переходит в такую точку z' луча Oz (принадлежность z' лучу Oz вытекает из равенства $\operatorname{Arg} z' = \operatorname{Arg} z$), что

$$(Oz') = \frac{1}{(Oz)} \quad \text{или} \quad (Oz)\cdot(Oz') = 1 \tag{15}$$

(рис. 20). Инверсия является одним из самых простых круговых преобразований. Как и всякое круговое преобразование, каждую окружность или прямую она переводит снова в окружность или прямую;

а квадраты сторон — произведениям противоположных сторон z_1z_2, z_2z_3, z_3z_4. Отсюда можно вывести целый ряд разнообразных геометрических свойств четырехугольников.

[1]) Так, например, если $z_1 = az + b$ и $z' = a_1z_1 + b_1$, то $z' = Az + B$, где $A = a_1a$, $B = a_1b + b_1$.

углы между (пересекающимися) окружностями инверсия сохраняет (впрочем, знаки ориентированных углов меняются при инверсии на обратные — инверсия является *зеркальным* круговым преобразованием); двойные отношения расстояний четверок точек плоскости при инверсии также не меняются. Точки единичной окружности $z\bar{z} = 1$ переходят при инверсии сами в себя; кроме того, каждая внешняя по отношению к окружности $z\bar{z} = 1$ точка z переходит во внутреннюю точку z' (лежащую на луче Oz и такую, что радиус $(Oz_0) = 1$

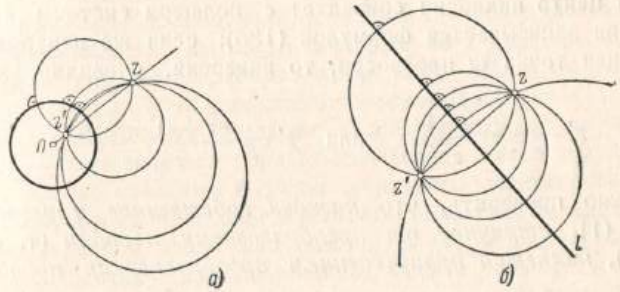


Рис. 21.

единичной окружности является средним пропорциональным между длинами отрезков (Oz) и (Oz'); см. рис. 20); точка z' переходит обратно в z. Все эти обстоятельства дают основание для того, чтобы употреблять наряду с названием единичная инверсия также и термин *симметрия относительно единичной окружности*[1]).

Преобразование (13а) не заслуживает специального рассмотрения — это есть, очевидно, произведение единичной инверсии (симметрии относительно единичной окружности) (13б) и симметрии (4б) относительно полярной оси o. Также сводится к рассмотренным ранее преобразованиям и *инверсия произвольной степени* k:

$$z' = \frac{k}{\bar{z}} \qquad (k \text{ вещественно}), \tag{16}$$

представляющая собой произведение единичной инверсии (13б) и гомотетии $z' = kz$ с коэффициентом k. И вообще, *каждое круговое преобразование представляет собой произведение преобразования*

[1]) Дополнительным стимулом здесь является то, что любая окружность, проходящая через отвечающие друг другу в единичной инверсии точки z и z', будет *перпендикулярна единичной окружности* $z\bar{z} = 1$ (рис. 21, *а*), подобно тому, как любая окружность, проходящая через две симметричные относительно прямой l точки z и z', будет перпендикулярна l (ср. рис. 21, *б*).

подобия и инверсии — доказательство этого предложения (основной теоремы о круговых преобразованиях) и составляет нашу основную цель.

Уточним прежде всего формулы, описывающие преобразование инверсии. Геометрически единичная инверсия определяется как преобразование, переводящее произвольную точку z плоскости в такую точку z' луча Oz (точка O фиксирована), что выполняется соотношение (15). Таким образом, в геометрическом описании инверсии важную роль играет точка O; эта точка называется *центром* инверсии. Если центр инверсии совпадает с полюсом системы координат, то инверсия записывается формулой (13б); если же центром служит произвольная точка w плоскости, то инверсия, очевидно, запишется так:

$$z'-w=\frac{1}{\overline{z-w}} \quad \text{или} \quad z'=\frac{w\bar{z}+(1-w\bar{w})}{\bar{z}-\bar{w}}. \tag{17}$$

Далее можно проверить, что *каждое собственное круговое преобразование* (1), *отличное от преобразования подобия* (т. е. такое, что $c\neq 0$), *является произведением преобразования подобия* (3б), *где*

$$p=-\frac{\bar{c}^2}{\bar{\Delta}}, \quad q=\frac{a\bar{c}\bar{d}-c\bar{c}\bar{d}-a\bar{c}\bar{b}}{c\bar{\Delta}}, \qquad \Delta=\begin{vmatrix} a & b \\ c & d \end{vmatrix}\neq 0,$$

и инверсии (17), *где*

$$w=\frac{a}{c},$$

а каждое зеркальное круговое преобразование (1а) — *произведением преобразования подобия* (3а), *где*

$$p=\frac{c^2}{\Delta}, \quad q=\frac{a\bar{a}\bar{d}-\bar{c}\bar{c}\bar{d}-\bar{a}\bar{c}\bar{b}}{\bar{c}\Delta}, \qquad \Delta\neq 0,$$

и той же инверсии, что и выше. В самом деле, если, например,

$$z_1=-\frac{\bar{c}^2}{\bar{\Delta}}\bar{z}+\frac{a\bar{a}\bar{d}-c\bar{c}\bar{d}-a\bar{c}\bar{b}}{c\bar{\Delta}} \quad \text{и} \quad z'=\frac{\frac{a}{c}\bar{z}_1+\left(1-\frac{a\bar{a}}{c\bar{c}}\right)}{\bar{z}_1-\frac{\bar{a}}{\bar{c}}},$$

то имеем

$$z'=\frac{\frac{a}{c}\left\{-\frac{c^2}{\Delta}z+\frac{a\bar{a}\bar{d}-c\bar{c}\bar{d}-\bar{a}\bar{c}\bar{b}}{c\bar{\Delta}}\right\}+\left(1-\frac{a\bar{a}}{c\bar{c}}\right)}{-\frac{c^2}{\Delta}z+\frac{a\bar{a}\bar{d}-c\bar{c}\bar{d}-\bar{a}\bar{c}\bar{b}}{c\bar{\Delta}}-\frac{\bar{a}}{\bar{c}}}=\frac{az+b}{cz+d}.$$

Таким образом, все отличные от преобразования подобия круговые преобразования плоскости сводятся в определенном смысле к единичной инверсии [1]).

6. Осевые круговые преобразования (преобразования Лагерра). Здесь мы рассмотрим дробно-линейные функции (1) и (1а) *дуального переменного*, где теперь придется требовать, чтобы определитель $\Delta=\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ не являлся делителем нуля, а также отвечающие этим функциям преобразования множества ориентированных прямых (осей) евклидовой плоскости, которые мы будем называть *дробно-линейными осевыми преобразованиями* [2]); при этом нам, как и выше, будет иногда удобнее называть преобразования (1) *собственными дробно-линейными осевыми преобразованиями*, а преобразования (1а) — *зеркальными дробно-линейными осевыми преобразованиями*. Дробно-линейные функции дуального переменного являются однозначными функциями, определенными на множестве всех дуальных чисел, расширенном введением чисел $c\omega$ (c вещественно) и ∞ [3]); соответственно этому дробно-линейные осевые преобразования являются взаимно однозначными преобразованиями множества всех осей (ориентированных прямых) плоскости. Частными случаями дробно-линейных осевых преобразований являются *симметрия относительно точки O, симметрия относительно прямой o и переориентация*

$$z'=\bar{z} \text{ (а)}, \quad z'=-z \text{ (б)} \quad \text{и} \quad z'=-\frac{1}{\bar{z}} \text{ (в)}, \tag{18}$$

а также произвольное *движение*

$$z'=\frac{pz+q}{-\bar{q}z+\bar{p}}, \quad z'=\frac{-pz+q}{\bar{q}z+\bar{p}}, \quad z'=\frac{p\bar{z}+q}{-\bar{q}\bar{z}+\bar{p}} \quad \text{или} \quad z'=\frac{-p\bar{z}+q}{\bar{q}\bar{z}+\bar{p}};$$
$$\Delta=p\bar{p}+q\bar{q}\neq 0. \tag{19}$$

В точности так же, как в п. **5**, показывается, что *дробно-линейные осевые преобразования составляют группу*. Особенно важным для нас здесь будет то обстоятельство, что *если* z_1', z_2', z_3' *и* z_4' *суть четыре* (ориентированные) *прямые плоскости, в которые переводит дробно-линейное осевое преобразование* (1) *или* (1а) *данные четыре*

[1]) Из доказанного следует также, что *каждое отличное от преобразования подобия круговое преобразование можно представить в виде произведения инверсии какой-то степени k и движения*; это предложение также можно назвать «основной теоремой о круговых преобразованиях».

[2]) В противоположность этим преобразованиям рассмотренные в предыдущем пункте дробно-линейные преобразования множества точек плоскости можно было бы назвать *точечными дробно-линейными преобразованиями*.

[3]) Ср. сноску [1]) на стр. 68.

прямые z_1, z_2, z_3 *и* z_4, *то имеет место соотношение* (9)

$$W(z_1', z_2', z_3', z_4') = W(z_1, z_2, z_3, z_4)$$

или

$$W(z_1', z_2', z_3', z_4') = \overline{W(z_1, z_2, z_3, z_4)} \qquad (9)$$

(свойство инвариантности двойного отношения); доказательство этого свойства также ничем не отличается от доказательства соответствующего предложения в п. **5**. Отсюда сразу вытекает, что *дробно-линейные осевые преобразования переводят четыре* (ориентированные) *прямые, принадлежащие одной* (ориентированной) *окружности или точке, в четыре прямые, также принадлежащие одной окружности или точке*; другими словами, *дробно-линейное осевое преобразование переводит каждую* (ориентированную) *окружность или точку снова в окружность или точку*. Это обстоятельство позволяет называть дробно-линейные осевые преобразования плоскости *осевыми круговыми преобразованиями*[1]) (причем преобразования (1) называются *собственными осевыми круговыми преобразованиями*, а преобразования (1а) — *зеркальными осевыми круговыми преобразованиями*). Так как осевые круговые преобразования впервые рассматривались французским математиком Эдмондом Лагерром, то их часто называют *преобразованиями Лагерра*.

Так же, как в п. **5**, показывается, что *существует единственное дробно-линейное осевое преобразование* (1) *и единственное зеркальное дробно-линейное осевое преобразование* (1а), *которые переводят три заданные* (ориентированные) *прямые* z_1, z_2 *и* z_3, никакие две из которых не параллельны, *в три другие заданные* (ориентированные) *прямые* w_1, w_2 *и* w_3, никакие две из которых также не параллельны — эти преобразования записываются фигурирующими в п. **5** равенствами (10) и (10а):

$$\frac{z'-w_2}{w_1-w_2}:\frac{z'-w_3}{w_1-w_3} = \frac{z-z_2}{z_1-z_2}:\frac{z-z_3}{z_1-z_3}, \qquad (10)$$

$$\frac{z'-w_2}{w_1-w_2}:\frac{z'-w_3}{w_1-w_3} = \frac{\bar{z}-\bar{z}_2}{\bar{z}_1-\bar{z}_2}:\frac{\bar{z}-\bar{z}_3}{\bar{z}_1-\bar{z}_3}. \qquad (10а)$$

С другой стороны, четыре (ориентированные) прямые z_1, z_2, z_3 и z_4 не всегда можно перевести осевым круговым преобразованием в четыре другие (ориентированные) прямые w_1, w_2, w_3 и w_4; для того, чтобы это было возможно, необходимо и достаточно, чтобы имело

[1]) Можно показать, что все преобразования множества ориентированных прямых плоскости, переводящие (ориентированные) окружности (к числу которых причисляются также и точки) снова в окружности, исчерпываются дробно-линейными преобразованиями (1) и (1а), сопровождаемыми еще, быть может, преобразованием подобия (ср. И. М. Яглом, Геометрические преобразования, II, § 5 гл. II, стр. 314—321).

место одно из равенств

$$W(w_1, w_2, w_3, w_4) = W(z_1, z_2, z_3, z_4)$$

или

$$W(w_1, w_2, w_3, w_4) = \overline{W(z_1, z_2, z_3, z_4)}.$$

Из доказанного следует, что *любую* (ориентированную) *окружность или точку можно* (и при этом многими способами) *перевести осевым круговым преобразованием в любую другую окружность или точку*; для этого надо лишь потребовать, чтобы три (ориентированные) касательные первой окружности перешли в три (какие угодно!) ориентированные касательные второй окружности. В частности, *любую* (ориентированную) *окружность можно осевым круговым преобразованием перевести в точку*, в силу чего в вопросах, связанных с этими преобразованиями, обычно не различают между собой точки и окружности, считая первые частным случаем последних («окружности нулевого радиуса»).

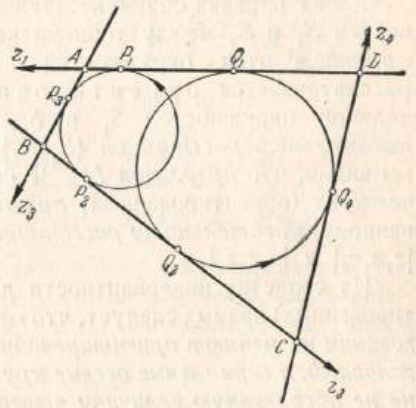


Рис. 22.

Выясним, какой геометрический смысл имеют аргумент Arg W и модуль $|W|$ двойного отношения $W(z_1, z_2, z_3, z_4) = \frac{z_1-z_3}{z_2-z_3}:\frac{z_1-z_4}{z_2-z_4}$ четырех (ориентированных) прямых z_1, z_2, z_3 и z_4. Выше (см. п. **4**) мы фактически доказали, что

$$\operatorname{Arg} W(z_1, z_2, z_3, z_4) = \operatorname{Arg}\frac{z_1-z_3}{z_2-z_3} - \operatorname{Arg}\frac{z_1-z_4}{z_2-z_4} =$$

$$= \frac{1}{2}(\{[z_1z_3], [z_2z_3]\} + \{[z_2z_4], [z_1z_4]\} - \{[z_4z_1], [z_3z_1]\} - \{[z_3z_2], [z_4z_2]\})$$

(ср. стр. 85 и след.). Рассмотрим теперь две (ориентированные) окружности, определяемые прямыми z_1, z_2, z_3 и z_1, z_2, z_4, или, как мы будем их обозначать впоследствии, окружности $[z_1z_2z_3]$ и $[z_1z_2z_4]$ (рис. 22). Обозначим точки касания окружностей S_1 и S_2 с прямыми z_1, z_2, z_3 и z_1, z_2, z_4 через P_1, P_2, P_3, соответственно Q_1, Q_2, Q_4; кроме того, условимся для простоты вместо $[z_1z_3]$, $[z_2z_3]$, $[z_2z_4]$ и $[z_1z_4]$ писать A, B, C и D. Так как, очевидно (ср. выше, стр. 85),

$$\frac{1}{2}(\{AB\} + \{CD\} - \{DA\} - \{BC\}) =$$

$$= \frac{1}{2}(\{AP_3\} + \{P_3B\} + \{CQ_4\} + \{Q_4D\} -$$

$$- \{DQ_1\} - \{Q_1P_1\} - \{P_1A\} - \{BP_2\} - \{P_2Q_2\} - \{Q_2C\})$$

и, по свойству касательных к окружности,

$$\{AP_3\}=\{P_1A\},\quad \{P_3B\}=\{BP_2\},\quad \{CQ_4\}=\{Q_2C\},\quad \{Q_4D\}=\{DQ_1\}$$

и

$$\{P_2Q_2\}=\{Q_1P_1\}=-\{P_1Q_1\},$$

то получим

$$\operatorname{Arg} W(z_1,\ z_2,\ z_3,\ z_4)=P_1Q_1. \tag{20}$$

Длина отрезка общей касательной z двух (ориентированных) окружностей S_1 и S_2 между точками касания называется *касательным расстоянием* этих окружностей и обозначается через $(S_1,\ S_2)$; если рассматривается ориентированная длина отрезка общей касательной окружностей S_1 и S_2, то говорят об *ориентированном касательном расстоянии* $\{S_1zS_2\}$ этих окружностей. Таким образом, мы видим, что *аргумент* Arg W *двойного отношения* $W(z_1,\ z_2,\ z_3,\ z_4)$ *четырех* (ориентированных) *прямых* z_1, z_2, z_3 *и* z_4 *равен* (ориентированному) *касательному расстоянию* $\{[z_1z_2z_3]\,z_1\,[z_1z_2z_4]\}$ *окружностей* $[z_1z_2z_3]$ *и* $[z_1z_2z_4]$.

Из свойства инвариантности двойного отношения четырех (ориентированных) прямых следует, что *собственные осевые круговые преобразования не меняют ориентированного касательного расстояния окружностей, а зеркальные осевые круговые преобразования меняют знак, но не абсолютную величину этого расстояния.* Это важное свойство осевых круговых преобразований обыкновенно формулируют так: *осевые круговые преобразования сохраняют касательные расстояния окружностей* [1]). В частности, *касающиеся* (ориентированные) *окружности* (окружности, касательное расстояние которых равно нулю) *переходят при осевых круговых преобразованиях в касающиеся окружности* [2]).

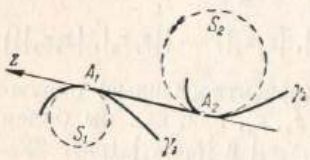


Рис. 23.

[1]) Пусть γ_1 и γ_2 — две произвольные кривые и z — их общая касательная (рис. 23); тогда расстояние $(A_1A_2)=(\gamma_1\gamma_2)$ между точками A_1 и A_2 соприкосновения γ_1 и γ_2 с прямой z называется *касательным расстоянием* γ_1 и γ_2. Так как, очевидно, $(\gamma_1\gamma_2)=(S_1S_2)$, где S_1 и S_2 — окружности, касающиеся кривых γ_1 и γ_2 в точках A_1 и A_2, и осевое круговое преобразование, переводящее кривые γ_1 и γ_2 в другие кривые γ_1' и γ_2', переводит S_1 и S_2 в окружности S_1' и S_2', касающиеся γ_1' и γ_2', то из доказанного вытекает, что *осевые круговые преобразования сохраняют касательные расстояния произвольных кривых.* Преобразования множества кривых плоскости, обладающие этим последним свойством, называются эквилонгальными преобразованиями; таким образом, *осевые круговые преобразования являются эквилонгальными преобразованиями.*

[2]) Заметим, что ориентированные окружности называются *касающимися* лишь в том случае, если они имеют единственную общую ориентированную касательную, т. е. если они касаются в обычном смысле и их направления в точке касания совпадают.

Перейдем теперь к модулю $|W|$ двойного отношения $W(z_1, z_2, z_3, z_4)$ четырех (ориентированных) прямых z_1, z_2, z_3 и z_4. Используя основную формулу (16), п. 4 (стр. 77), а также то обстоятельство, что модуль частного или разности двух дуальных чисел равен частному, соответственно разности модулей этих чисел, получим

$$|W(z_1,\ z_2,\ z_3,\ z_4)|=\frac{\operatorname{tg}\frac{\angle\{z_1o\}}{2}-\operatorname{tg}\frac{\angle\{z_3o\}}{2}}{\operatorname{tg}\frac{\angle\{z_2o\}}{2}-\operatorname{tg}\frac{\angle\{z_3o\}}{2}}:\frac{\operatorname{tg}\frac{\angle\{z_1o\}}{2}-\operatorname{tg}\frac{\angle\{z_4o\}}{2}}{\operatorname{tg}\frac{\angle\{z_2o\}}{2}-\operatorname{tg}\frac{\angle\{z_4o\}}{2}}.$$

Учитывая еще, что $\operatorname{tg}\alpha-\operatorname{tg}\beta=\frac{\sin(\alpha-\beta)}{\cos\alpha\cos\beta}$, будем иметь

$$|W(z_1,\ z_2,\ z_3,\ z_4)|=\frac{\sin\left(\frac{1}{2}\angle\{z_3z_1\}\right)}{\sin\left(\frac{1}{2}\angle\{z_3z_2\}\right)}:\frac{\sin\left(\frac{1}{2}\angle\{z_4z_1\}\right)}{\sin\left(\frac{1}{2}\angle\{z_4z_2\}\right)}.$$

Вещественное число, стоящее в правой части этого равенства, мы будем называть *двойным отношением углов между четырьмя* (ориентированными) *прямыми* z_1, z_2, z_3 *и* z_4 и обозначать через $\widehat{W}(z_1,z_2,\ z_3,\ z_4)$; таким образом,

$$|W(z_1,\ z_2,\ z_3,\ z_4)|=\widehat{W}(z_1,\ z_2,\ z_3,\ z_4). \tag{21}$$

Из инвариантности двойных отношений четверок (ориентированных) прямых при осевых круговых преобразованиях мы можем заключить, что *осевые круговые преобразования сохраняют двойные отношения углов между четверками* (ориентированных) *прямых.*

Из доказанного выше следует, что *четыре* (ориентированные) *прямые* z_1, z_2, z_3 *и* z_4 *можно в том и только в том случае перевести осевым круговым преобразованием в другие четыре прямые* w_1, w_2, w_3 *и* w_4, *если касательное расстояние окружностей* $[z_1z_2z_3]$ *и* $[z_1z_2z_4]$ *равно касательному расстоянию окружностей* $[w_1w_2w_3]$ *и* $[w_1w_2w_4]$ *и двойное отношение углов между* z_1, z_2, z_3 *и* z_4 *равно двойному отношению углов между* w_1, w_2, w_3 *и* w_4 [1]).

[1]) Отсюда легко вывести, что *для того, чтобы выпуклый четырехугольник* $z_1z_2z_3z_4$ (т. е. такой, который лежит по одну сторону — слева или справа — от каждой из своих сторон z_1, z_2, z_3 и z_4) *мог быть переведен осевым круговым преобразованием в другой выпуклый четырехугольник* $w_1w_2w_3w_4$, *необходимо и достаточно, чтобы разность* $(\{AB\}+\{CD\})-(\{DA\}+\{BC\})$ (где $A\equiv[z_1z_3]$, $B\equiv[z_2z_3]$, $C\equiv[z_2z_4]$, $D\equiv[z_1z_4]$) *между суммами противоположных сторон первого четырехугольника равнялась разности сумм противоположных сторон второго четырехугольника и отношение*

$$\frac{\sin\left(\frac{1}{2}\angle\{z_3z_1\}\right)\sin\left(\frac{1}{2}\angle\{z_4z_2\}\right)}{\sin\left(\frac{1}{2}\angle\{z_3z_2\}\right)\sin\left(\frac{1}{2}\angle\{z_4z_1\}\right)}$$

произведений синусов половин

В заключение коснемся вопроса о геометрическом описании осевых круговых преобразований. Одними из простейших таких преобразований, отличных от движений (19), являются преобразования

$$z' = \frac{k}{z} \quad \text{(а)} \quad \text{и} \quad z' = \frac{k}{\bar{z}} \quad \text{(б)} \quad (k \text{ вещественно}) \tag{22}$$

при k, отличном от ± 1 [1]); эти преобразования можно также переписать так:

$$\operatorname{Arg} z' = -\operatorname{Arg} z,\ |z'| = \frac{k}{|z|} \text{ (а) и } \operatorname{Arg} z' = \operatorname{Arg} z,\ |z'| = \frac{k}{|z|} \text{ (б).} \tag{23}$$

Более простым из этих двух преобразований является преобразование (22б), называемое *осевой инверсией степени k*; прямая o называется *осью* этой инверсии. При осевой инверсии степени k каждая (направленная) прямая z плоскости переходит в прямую z', пересекающую полярную ось o (ось инверсии) в той же точке M, что и z, и такую, что

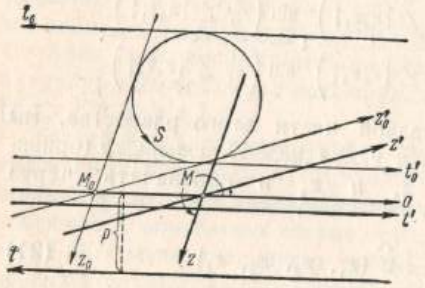


Рис. 24.

$$\operatorname{tg}\frac{\angle\{oz\}}{2}\operatorname{tg}\frac{\angle\{oz'\}}{2} = k \tag{24}$$

(рис. 24), а z' переходит в z; каждая же параллельная o прямая $t = \frac{1}{2}\varepsilon$, удаленная от o на (ориентированное) расстояние $\{ot\} = p$, переходит в прямую $t' = \frac{k}{\bar{t}} = -\frac{2k}{p}\omega$, противопараллельную o и удаленную от o на расстояние $\{ot'\} = -\frac{p}{k}$, а t' переходит в t. Что же касается до преобразования (22а), то оно складывается из осевой инверсии степени k (22б) и симметрии $z' = \bar{z}$ относительно полюса O.

Легко дать осевой инверсии и чисто геометрическое истолкование. Заметим прежде всего, что *каждая ориентированная окружность S*

$$Az\bar{z} + Bz - \bar{B}\bar{z} + C = 0 \quad (A \text{ и } C \text{ чисто мнимые}) \tag{25}$$

такая, что

$$\frac{C}{A} = k, \tag{26}$$

противоположных углов первого четырехугольника равнялось отношению произведений синусов половин противоположных углов второго четырехугольника. Из этого предложения в свою очередь вытекает возможность перевести осевым круговым преобразованием произвольный выпуклый четырехугольник в параллелограмм, что может быть использовано для доказательства целого ряда общих свойств четырехугольников.

[1]) При $k = \pm 1$ преобразования (22а,б) являются, очевидно, движениями.

переводится в себя осевой инверсией (22б) *степени k*. В самом деле, преобразование (22б) можно записать в виде $z = k/\bar{z}'$; оно переводит окружность (25) в окружность $A\frac{k}{\bar{z}'}\frac{k}{z'} + B\frac{k}{\bar{z}'} - \bar{B}\frac{k}{z'} + C = 0$ или

$$Cz'\bar{z}' + Bkz' - \bar{B}k\bar{z}' + Ak^2 = 0,$$

или, наконец, в окружность (напомним, что $C = Ak$, $Ak^2 = Ck$)

$$Az'\bar{z}' + Bz' - \bar{B}\bar{z}' + C = 0,$$

т. е. саму в себя. Отсюда вытекает для любых двух (ориентированных) касательных z_0 и z_0' окружности S, пересекающихся в точке M_0 оси o, $z_0' = k/\bar{z}_0$, $\bar{z}_0 z_0' = k$ [эти прямые отвечают друг другу в инверсии (22б)], т. е.

$$\operatorname{tg}\frac{\angle\{oz_0\}}{2}\operatorname{tg}\frac{\angle\{oz_0'\}}{2} = k;$$

параллельная же и противопараллельная o касательные t_0 и t_0' окружности S таковы, что $t_0\bar{t}_0' = k$ или $\{ot_0\} : \{ot_0'\} = -k$ (рис. 24).

Пусть теперь S — какая угодно окружность (25), такая, что $C/A = k$; эта окружность называется *направляющей окружностью* осевой инверсии. В таком случае для того, чтобы найти (ориентированную) прямую z', в которую переводит наша инверсия заданную прямую z, пересекающую o в точке M, надо провести касательную z_0 окружности S, параллельную z и пересекающую o в точке M_0, и вторую касательную z_0', проходящую через точку M_0; прямая z' будет параллельна z_0' и будет проходить через точку M (рис. 24). Прямая же t, не пересекающая o, перейдет в такую прямую t', что если t_0 и t_0' — не пересекающие o касательные S, то $t \parallel t_0$, $t' \parallel t_0'$ и $\{ot\} : \{ot'\} = \{ot_0\} : \{ot_0'\}$.

Наряду с движениями

$$z' = \frac{pz+q}{-\bar{q}z+\bar{p}} \quad z' = \frac{-pz+q}{\bar{q}z+\bar{p}}, \quad z' = \frac{p\bar{z}+q}{-\bar{q}\bar{z}+\bar{p}} \text{ или } z' = \frac{-p\bar{z}+q}{\bar{q}\bar{z}+\bar{p}} \tag{19}$$

и инверсией $z' = \frac{k}{\bar{z}}$ (22б) специального рассмотрения заслуживает еще одно интересное осевое круговое преобразование, а именно преобразование

$$z' = \frac{z+q}{-qz+1}, \tag{27}$$

где $q=\varepsilon\frac{t}{2}$, $|q|=0$ (ср. с параллельным переносом в направлении, перпендикулярном к полярной оси $z'=\frac{z+q}{qz+1}$, где $q=\varepsilon\frac{t}{2}$; см. формулу (19а) на стр. 80). Геометрический смысл этого преобразования весьма прост: каждую прямую $z=\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)$ оно переводит в прямую

$$z'=\operatorname{tg}\frac{\theta'}{2}(1+\varepsilon s')=\left[\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)+\varepsilon\frac{t}{2}\right]:\left[-\operatorname{tg}\frac{\theta}{2}(1+\varepsilon s)\,\varepsilon\frac{t}{2}+1\right]=$$

$$=\operatorname{tg}\frac{\theta}{2}\left[1+\varepsilon\left(s+\frac{t}{2\operatorname{tg}\frac{\theta}{2}}\right)\right]:\left(1-\varepsilon\frac{t}{2}\operatorname{tg}\frac{\theta}{2}\right)=$$

$$=\operatorname{tg}\frac{\theta}{2}\left(1+\varepsilon\left[s+\frac{t}{2}\left(\operatorname{tg}\frac{\theta}{2}+\operatorname{ctg}\frac{\theta}{2}\right)\right]\right)=\operatorname{tg}\frac{\theta}{2}\left[1+\varepsilon\left(s+\frac{t}{\sin\theta}\right)\right],$$

параллельную z (ибо $|z'|=|z|$) и такую, что расстояния p и p' прямых z и z' от полюса O полярной системы координат связаны соотношением [см. формулу (17) на стр. 77]

$$p'=s'\sin\theta'=\left(s+\frac{t}{\sin\theta}\right)\sin\theta=s\sin\theta+t=p+t. \qquad (28)$$

Другими словами, *прямая z' параллельна прямой z и расстояние $\{z, z'\}$ прямой z от z' равно t* (рис. 25). Преобразование (27) называется *расширением* на (положительную или отрицательную!) величину t. Очевидно, что (ориентированную) окружность S (положительного или отрицательного) радиуса r расширение переводит в концентрическую с S окружность S' радиуса $r+t$; в частности, точки оно переводит в окружности радиуса t, а окружности радиуса $(-t)$ — в точки.

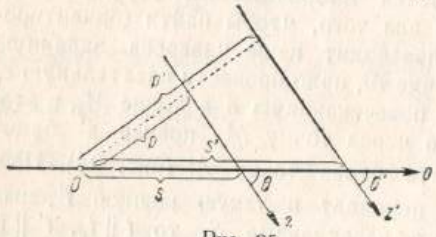


Рис. 25.

«Основная теорема об осевых круговых преобразованиях» утверждает, что *всякое отличное от движения* (19) *осевое круговое преобразование* (1) *или* (1а) *представляет собой произведение движения и осевой инверсии или произведение движения и расширения*. Доказательство этой теоремы можно получить прямым подсчетом, подобрав, подобно тому, как мы поступали в п. **5**, такие движение и осевую инверсию (или движение и расширение), произведение которых дает наперед заданное осевое круговое преобразование; мы предоставим это доказательство читателю.

ОЧЕРК ОСНОВНЫХ ИДЕЙ ТОПОЛОГИИ

В. Г. Болтянский и В. А. Ефремович

(Москва)

(Окончание)

8. Группы гомологий

В этой главе нашего очерка мы рассмотрим важные топологические инварианты — *группы гомологий*, по существу введенные уже Пуанкаре. Гомологические понятия находят применения не только в самой топологии, но и в других разделах математики: алгебре, анализе, геометрии и т. д. Для понимания последних глав от читателя требуется знание простейших сведений из теории групп[1]). Первый раздел носит характер введения и не требует предварительных знаний.

Введение — наглядное описание гомологий

Будем называть *циклом* на некоторой поверхности X ориентированную замкнутую линию[2]) на этой поверхности или несколько таких линий, вместе взятых. Например, ориентированные параллель и меридиан тора (рис. 137), вместе взятые, являются циклом; каждая из этих линий, взятая в отдельности, также представляет собой цикл.

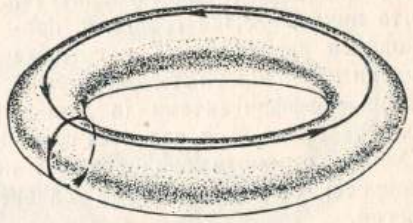


Рис. 137.

Циклы на поверхности мы будем разделять на такие, которые ограничивают некоторую область (иначе говоря, являются границей этой области), и на такие, которые никакой области не ограничивают. Эти понятия, кажущиеся на первый взгляд простыми

[1]) Эти сведения кратко изложены на стр. 51—52 4-го выпуска «Математического просвещения» и на стр. 136—138 настоящего выпуска.

[2]) То есть линию с установленным на ней направлением; на рисунках направление на линии отмечается стрелкой.

и ясными (например, на плоскости окружность или другая замкнутая линия без самопересечений, очевидно, «ограничивает» область, находящуюся внутри нее), на самом деле требуют уточнений — иначе могут появиться неясности.

Мы будем считать, что поверхность X двусторонняя. Выберем некоторый ее кусок U и окрасим его с одной стороны в какой-либо цвет. На линии z, являющейся краем этого куска, выберем такое направление, чтобы глядя (в пространстве) на кусок U с окрашенной стороны и двигаясь вдоль линии z в этом направлении (вдоль цикла z), мы слева видели точки, принадлежащие U, а справа — не принадлежащие. В этом случае мы и будем считать ориентированную линию z *границей* куска U (или ограничивающей кусок U). Ясно, что если бы мы окрасили кусок U с обратной стороны, то его границей была бы та же линия, но с противоположным направлением; этот цикл мы будем называть $-z$.

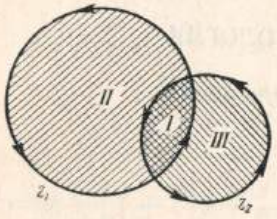


Рис. 138.

Понятие цикла, ограничивающего область, не столь тривиально, как это кажется с первого взгляда. На рис. 138 изображены на одной и той же плоскости два цикла z_1 и z_2, каждый из которых является границей области, находящейся внутри заштрихованного круга. Взятые вместе эти циклы также образуют *цикл-сумму* (его обозначают $z_1 + z_2$); он ограничивает область, состоящую из «дважды взятого» куска I и «один раз взятых» кусков II и III. Таким образом, чтобы убедиться, что цикл является границей, приходится иногда брать некоторые куски поверхности X с определенными коэффициентами (в данном случае цикл $z_1 + z_2$ является границей области $U \equiv 2 \cdot I + 1 \cdot II + 1 \cdot III$). Однако в нашем введении мы не будем останавливаться на этом вопросе: он прояснится в дальнейшем при изучении групп гомологий.

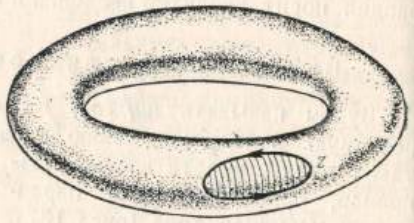


Рис. 139.

На торе цикл z, изображенный на рис. 139, ограничивает заштрихованную область, а меридиан или параллель тора, а также обе эти линии, взятые вместе, никакой области не ограничивают.

Очевидно, ограничивающие циклы можно провести на любой поверхности: всякая «маленькая окружность» на поверхности является ограничивающим циклом. Циклы же, которые не являются границей области, встречаются реже. Например, на торе имеются неограничивающие циклы, а на сфере их нет; отсюда можно заключить, что

сфера негомеоморфна тору. Изображенный на рис. 140 цикл на поверхности «кренделя» является ограничивающим (ограниченная им область заштрихована). Область, которую ограничивает некоторый цикл, будем называть *пленкой*, натянутой на этот цикл.

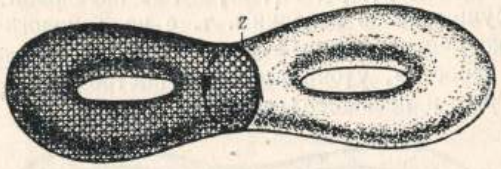
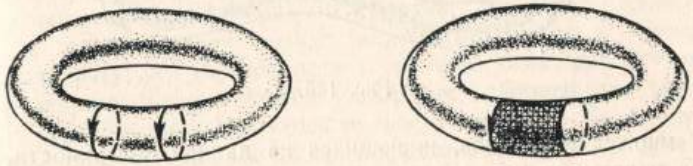


Рис. 140.



а) *б)*

Рис. 141.

Мы будем считать ограничивающие циклы «несущественными», а к «существенным» будем причислять те, которые никакой области не ограничивают. Будем также считать, что два различных цикла z_1 и z_2 «несущественно отличаются друг от друга» (в этом случае говорят также, что они *гомологичны* друг другу), если циклы z_1 и $-z_2$, взятые вместе, образуют несущественный цикл. Таковы, например, два меридиана z_1 и z_2 тора (рис. 141, *а*), имеющих сходную ориентацию; на рис. 141, *б* заштрихована пленка, которую ограничивают вместе взятые циклы z_1 и $-z_2$. Точно так же гомологичны друг другу циклы z_1 и z_2, изображенные на рис. 142,— здесь также z_1 и $-z_2$ ограничивают заштрихованную область (пленку).

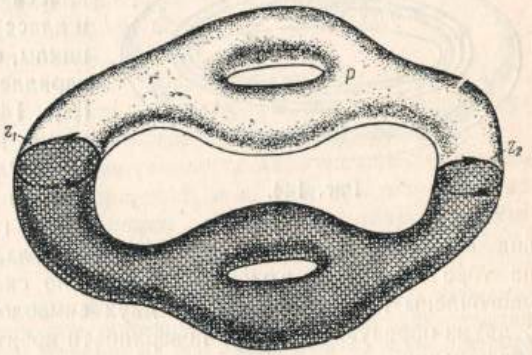


Рис. 142.

В дальнейшем нас будет интересовать вопрос — *сколько существенно различающихся друг от друга циклов можно провести на данной поверхности?* Очевидно, что «запас циклов» на торе больше, чем на сфере, а на кренделе — больше, чем на торе (на рис. 143 изображены четыре цикла p_1, p_2, m_1, m_2 — две параллели и два меридиана, все они существенно различны, т. е не гомологичны между собой). Несколько неопределенный термин «запас циклов, имеющихся на данной поверхности», уточнится после построения групп гомологий, изучению которых и посвящена настоящая глава.

Рис. 143.

Всё множество циклов, имеющихся на данной поверхности, разобьем на отдельные классы циклов (называемые *классами гомологий*), объединяя все гомологичные между собой циклы в один класс. Например, на сфере имеется только один класс гомологий, потому что все циклы гомологичны между собой (все они являются ограничивающими). На торе же имеется бесконечно много классов гомологий, но все они выражаются через два основных класса — через класс меридианов и класс параллелей. Например, все циклы, обходящие тор три раза по параллели и два раза по меридиану (рис. 144), составляют класс гомологий, который можно обозначить через $2m + 3p$, где m — меридиан, а p — параллель. Вообще, любой класс гомологий на торе имеет вид $\alpha m + \beta p$, где α и β — целые числа, т. е. классы гомологий на торе образуют группу (их можно складывать и вычитать как многочлены первой степени от двух символов m и p), причем группу с двумя образующими. На поверхности кренделя уже имеются четыре образующие: m_1, m_2, p_1, p_2 (см. рис. 143), т. е. любой класс гомологий записывается в виде $\alpha_1 m_1 + \alpha_2 m_2 + \beta_1 p_1 + \beta_2 p_2$. Таким образом, «группа гомологий» кренделя содержит «больше» элементов, чем группа гомологий тора, и поэтому эти поверхности негомеоморфны.

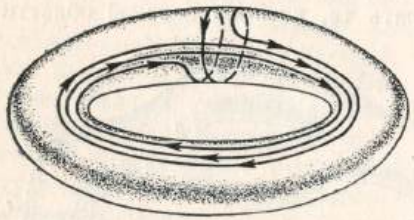


Рис. 144.

На этих примерах выясняется, что множество всех классов гомологий на некоторой поверхности представляет собой группу; эту

группу и называют *группой гомологий*. Например, группа гомологий сферы тривиальна, группа гомологий тора является свободной абелевой группой с двумя образующими, а группа гомологий поверхности кренделя — свободной абелевой группой с четырьмя образующими.

Как же строятся группы гомологий в общем случае? Для того чтобы разъяснить это, заметим, что все вообще циклы на рассматриваемой поверхности составляют абелеву группу — их можно складывать и вычитать. Все ограничивающие циклы также составляют группу — подгруппу группы циклов (ибо сумма и разность двух ограничивающих циклов снова являются ограничивающими циклами). Мы хотим «собрать» все гомологичные между собой циклы в один класс; в частности, в один класс должны быть собраны все ограничивающие циклы («гомологичные нулю»)[1]).

В заключение этого введения отметим, что в приведенных примерах мы рассматривали только «одномерные» циклы (линии) на двумерных фигурах (поверхностях). В топологии рассматриваются циклы различных размерностей (нульмерные, одномерные, двумерные и т. п.); ограниченные ими области могут иметь и более чем два измерения. Заметим еще, что «область», ограниченная некоторым несущественным циклом (т. е. пленка, натянутая на этот цикл), может располагаться «в толще» рассматриваемой фигуры. Пусть, например, X есть часть трехмерного пространства, ограниченная тором (т. е. X состоит из всех точек, лежащих на самом торе и внутри него). Тогда меридиан тора является несущественным циклом фигуры X, так как на него можно «натянуть пленку» в X (т. е. кусок поверхности, краем которого является рассматриваемый цикл); такой пленкой может быть сечение тора, изображенное на рис. 145. Напротив, параллель, изображенная на этом рисунке, является существенным циклом — она не ограничивает никакой пленки в X.

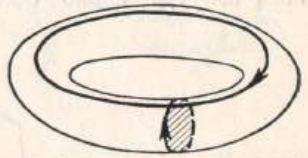


Рис. 145.

Очень важна также связь понятия гомологии двух циклов с понятием их *деформации*. Именно, *если два цикла z_1 и z_2 таковы, что один из них можно с помощью деформации перевести, «передвинуть» в другой* (такие циклы называются гомотопными, см. главу 7)[2]),

[1]) Алгебраически это соответствует тому, что мы должны взять *фактор-группу* группы циклов по подгруппе ограничивающих циклов. (О понятии фактор-группы см. стр. 137.) Действительно, два цикла z_1 и z_2 тогда и только тогда принадлежат одному элементу фактор-группы, когда их разность $z_1 - z_2$ принадлежит подгруппе ограничивающих циклов, т. е. когда они гомологичны между собой. Иначе говоря, элементами такой фактор-группы как раз и являются классы гомологий. Сама же фактор-группа и является группой гомологий рассматриваемой поверхности. Эти идеи будут более подробно разъяснены в настоящей главе.

[2]) См. «Математическое просвещение», вып. 4, стр. 28 и след.

то эти циклы гомологичны между собой. Это видно на рис. 141, *а* (стр. 109). Но обратное не имеет места: гомологичные циклы на рис. 142 (стр. 109) не гомотопны: перемещению цикла z_1 на поверхности P в цикл z_2 мешают «дырки», имеющиеся на поверхности между этими циклами. Таким образом, для гомологичности двух циклов достаточно (но не необходимо), чтобы они были гомотопны.

Ориентация. Коэффициенты примыкания

Мы уже говорили об ориентированных ребрах [1]). В топологии *ориентация* определяется не только для ребер, но и для двумерных клеток (граней), трехмерных клеток и т. д. Вершины же (нульмерные клетки) не ориентируются.

Ориентация двумерной клетки задается направлением обхода на ее контуре. Так как на контуре клетки можно задать два различных направления обхода («по» и «против» часовой стрелки), то каждая двумерная клетка имеет две различные ориентации. Например, если на контуре многоугольника задан обход против часовой стрелки, то это значит, что, подходя изнутри к какой-либо стороне этого многоугольника, мы видим на этой стороне направление движения справа налево (рис. 146). Иначе говоря, направление всех сторон многоугольника изнутри видно одинаковым (справа налево). При обратной

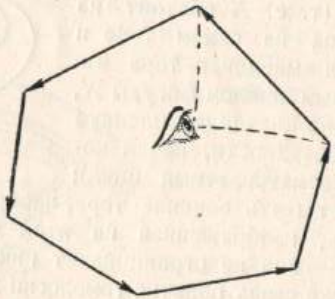


Рис. 146.

ориентации двумерной клетки (по часовой стрелке) все стороны кажутся изнутри также направленными одинаково (слева направо). Таким образом, *задать ориентацию многоугольника — это значит выбрать такие направления на его сторонах, которые изнутри наблюдаются как одинаковые.*

Аналогично определяется *ориентация трехмерной клетки.* Для того чтобы ориентировать многогранник, нужно для всех его граней выбрать такие ориентации, которые изнутри кажутся одинаковыми: либо все по часовой стрелке, либо все против часовой стрелки

[1]) «Математическое просвещение», вып. 3, стр. 37.

(рис. 147). Если мы видим (изнутри) все грани ориентированными по часовой стрелке, то направление луча зрения согласуется с этим обходом по *правилу буравчика* (рис. 148 [1]). Поэтому такую ориентацию трехмерного многогранника иногда называют *«правой»* («правый винт»), а обратную ориентацию — *левой.*

Введем теперь важное понятие *«коэффициента примыкания»* (или, как говорят топологи, *коэффициента инцидентности*) двух клеток.

Прежде всего определим коэффициент примыкания ориентированного ребра к его вершине. У такого ребра две вершины — начало и конец. Если ребро незамкнуто (т. е. его начало и конец не совпадают), то коэффициент примыкания этого ребра к началу считается равным «-1»,

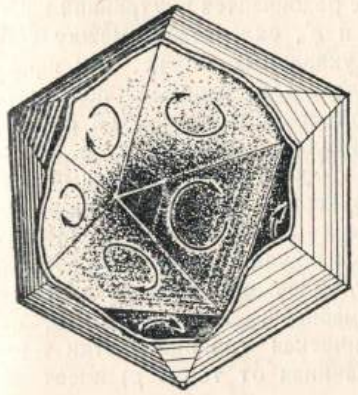


Рис. 147.

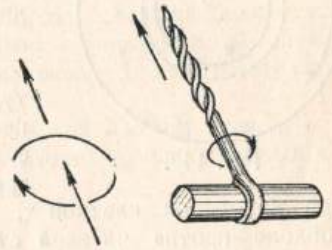


Рис. 148.

а коэффициент примыкания к концу — равным «$+1$». В случае замкнутого ребра (начало совпадает с концом) коэффициенты $+1$ и -1 взаимно уничтожаются и коэффициент примыкания к вершине считается равным нулю [2]).

Укажем теперь, как определяется коэффициент примыкания ориентированной грани к ориентированному ребру. Если при обходе по контуру грани мы проходим это ребро в направлении имеющейся на нем стрелки, то коэффициент примыкания считается равным «$+1$», если против направления стрелки, то «-1»; наконец, если ребро проходится несколько раз, то получающиеся коэффициенты алгебраически складываются. Из этого определения непосредственно следует, что *если составить гомотопическую границу* [3]) *грани* g (двигаясь в том направлении,

[1]) Ср. с рис. 125 на стр. 46 4-го выпуска «Математического просвещения».

[2]) На рис. 106, *а* («Математическое просвещение», вып. 4, стр. 37) коэффициент примыкания ребра к вершине *а* равен -1, а к вершине *b* равен $+1$; на рис. 106, *б* коэффициент примыкания ребра к вершине *а* равен нулю.

[3]) «Математическое просвещение», вып., 4. стр. 37—39.

которое указано на контуре ориентированной грани g), *то коэффициент примыкания этой грани к некоторому ориентированному ребру r равен сумме показателей степени, с которыми r входит в гомотопическую границу грани g.*

Рассмотрим, например, лист Мёбиуса. Мы знаем, что его можно представить себе в виде кругового кольца, у которого склеены диаметрально противоположные точки на внутреннем контуре [1]). Возьмем отрезок pq (рис. 149), соединяющий точку p внешнего контура с точкой q внутреннего, и пусть q' — точка, диаметрально противоположная q. Тогда при построении листа Мёбиуса две полуокружности, на которые разбивается внутренний контур точками q и q', склеиваются между собой (с учетом указанных на рис. 149 направлений стрелок). Обозначим эти полуокружности одной и той же буквой a (ибо в листе Мёбиуса они склеены, т. е. составляют одну линию), ориентированный отрезок pq — буквой b, а весь внешний контур — буквой c. Тогда лист Мёбиуса представится в виде клеточного разбиения с двумя нульмерными клетками p, $q=q'$, тремя одномерными клетками a, b, c и одной двумерной клеткой τ. Гомотопическая граница клетки τ (при ее обходе против часовой стрелки начиная от точки p) имеет вид:

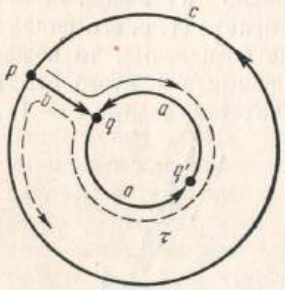


Рис. 149.

$$ba^{-1}a^{-1}b^{-1}c = ba^{-2}b^{-1}c;$$

поэтому коэффициенты примыкания клетки τ к a, b и c соответственно равны -2, 0, $+1$.

Рассматривая теперь снова примеры 4—7, приведенные в предыдущей главе [2]), мы получаем следующие факты. Проективная плоскость может быть представлена в виде клеточного разбиения, содержащего одну вершину a, одно ребро r и одну грань g; коэффициент примыкания грани к ребру равен ± 2 (в зависимости от того, как эта грань и это ребро ориентированы). Проективная плоскость по модулю m может быть представлена в виде клеточного разбиения с одной вершиной, одним ребром и одной гранью; коэффициент примыкания грани к ребру равен $\pm m$. Тор представляется в виде клеточного разбиения с одной вершиной, двумя ребрами r_1, r_2 и одной гранью g; коэффициент примыкания грани к каждому из ребер равен нулю. Ориентируемая поверхность P_m (сфера с m ручками) — в виде клеточного разбиения с одной вершиной, $2m$ ребрами и одной гранью; коэффициент примыкания грани к каждому из ребер равен нулю. Кроме того, во всех этих четырех примерах коэффициенты примыкания ребер к вершине равны нулю (ибо все ребра — замкнутые).

[1]) «Математическое просвещение», вып. 2, стр. 29.
[2]) «Математическое просвещение», вып. 4, стр. 41—45.

Цепи и их границы

Тот факт, что незамкнутое ориентированное ребро ab (будем его обозначать через r) примыкает к вершине b с коэффициентом $+1$, а к вершине a — с коэффициентом -1, условно выражают словами: «граница ориентированного ребра r равна точке b, взятой с коэффициентом $+1$, плюс точка a, взятая с коэффициентом -1». Записывают это так:

$$\Delta r = b - a.$$

Здесь Δ есть «символ границы» (выражение Δr читается: «граница клетки r»), а стоящая справа алгебраическая «сумма» двух точек имеет лишь следующий условный смысл: знак, стоящий перед вершиной, дает соответствующий коэффициент примыкания. В случае, если r обозначает замкнутую клетку, мы имеем: $\Delta r = 0$ (более подробно можно было бы записать: $\Delta r = 0a$).

Аналогично определяется граница двумерной клетки: если g есть двумерная ориентированная клетка, то ее граница Δg определяется как «сумма» входящих в контур клетки ребер

$$\Delta g = \sum_i \varepsilon_i r_i, \tag{1}$$

причем коэффициент ε_i, с которым берется ребро r_i, есть коэффициент примыкания грани g к ребру r_i. Например, в случае листа Мёбиуса, изображенного на рис. 149, мы имеем следующую границу клетки τ:

$$\Delta\tau = -2a + 0b + 1c = c - 2a.$$

Если мы условимся клетку g, ориентированную противоположным образом, обозначать через $-g$, то будем иметь

$$\Delta(-g) = \sum_i (-\varepsilon_i r_i) = -\sum_i \varepsilon_i r_i = -\Delta g,$$

ибо при перемене ориентации клетки g все коэффициенты примыкания меняют знак.

В правой части равенства (1) стоит сумма ориентированных ребер с некоторыми целочисленными коэффициентами. Сумма эта, конечно, только формальная, так как никакого «сложения» двух ребер мы произвести не можем,— мы можем только обозначить это сложение.

Всякую (формальную) сумму клеток, взятых с некоторыми коэффициентами, называют цепью. Цепи бывают *нульмерные, одномерные, двумерные* и т. д., в зависимости от того, какие клетки мы

складываем. Например, в правой части равенства (1) стоит одномерная цепь. Цепи можно складывать, вычитать и умножать на число. Складываются цепи как многочлены, причем при возникновении подобных членов они приводятся по обычным правилам (напомним, что при перемене ориентации клетки g мы получаем клетку, обозначаемую через $-g$, так что два члена, соответствующие двум различным ориентациям одной и той же клетки, являются подобными).

С помощью этой операции сложения множество всех двумерных цепей (а также одномерных или нульмерных) превращается в абелеву группу[1]); нулем этой группы является цепь, все коэффициенты которой равны нулю. Эта группа называется *группой цепей* (двумерных, соответственно одномерных, нульмерных). Мы будем обозначать группы цепей через $L_0(P)$, $L_1(P)$, $L_2(P)$, ... Здесь P означает то разбиение, в котором мы рассматриваем цепи, а индексы 0, 1, 2, ... указывают размерность рассматриваемых цепей.

Выше мы определили границу клетки; теперь же мы можем определить границу любой цепи. Именно, если $x_2 = \sum_i \alpha_i g_i$ есть двумерная цепь (α_i — коэффициенты, g_i — двумерные клетки), то ее граница Δx_2 определяется равенством

$$\Delta\left(\sum_i \alpha_i g_i\right) = \sum_i \alpha_i (\Delta g_i).$$

Так как границы Δg_i отдельных клеток мы уже умеем вычислять, то стоящую в правой части цепь нетрудно вычислить. Аналогично определяется граница одномерной (или трехмерной) цепи. Граница же любой нульмерной цепи считается равной нулю.

Нетрудно убедиться в том, что *граница алгебраической суммы двух (или нескольких) цепей равна алгебраической сумме границ*; например, если x_2' и x_2'' — две двумерные цепи, то

$$\Delta(x_2' \pm x_2'') = \Delta x_2' \pm \Delta x_2''. \tag{2}$$

Действительно, если $x_2' = \sum_i \alpha_i g_i$, $x_2'' = \sum_i \beta_i g_i$ (суммирование ведется по всем двумерным клеткам рассматриваемого разбиения, причем некоторые из коэффициентов α_i, β_i могут оказаться равными нулю), то мы имеем:

$$\Delta(x_2' + x_2'') = \Delta\left(\sum_i \alpha_i g_i + \sum_i \beta_i g_i\right) = \Delta \sum_i (\alpha_i + \beta_i) g_i = \sum_i (\alpha_i + \beta_i) \Delta g_i;$$

$$\Delta x_2' + \Delta x_2'' = \Delta \sum_i \alpha_i g_i + \Delta \sum_i \beta_i g_i =$$

$$= \sum_i \alpha_i \Delta g_i + \sum_i \beta_i \Delta g_i = \sum_i (\alpha_i + \beta_i) \Delta g_i.$$

[1]) Определение абелевой группы см. «Математическое просвещение», вып. 4, стр. 52.

Таким образом, операция Δ взятия границы (т. е. операция перехода от цепи x_2 к Δx_2) представляет собой отображение группы $L_2(P)$ в группу $L_1(P)$: если x_2 принадлежит $L_2(P)$, то Δx_2 принадлежит $L_1(P)$; при этом сумме $x_2' + x_2''$ отвечает сумма $\Delta x_2' + \Delta x_2''$, т. е. указанное отображение является *гомоморфным отображением* группы $L_2(P)$ в $L_1(P)$. [В общем случае Δ есть гомоморфное отображение группы $L_n(P)$ в группу $L_{n-1}(P)$].

Основное свойство границы

Операция взятия границы цепи обладает следующим основным свойством: *для любой цепи x имеем*

$$\Delta(\Delta x) = 0 \tag{3}$$

(«граница границы равна нулю»). Для одномерной цепи x это соотношение очевидно: Δx есть *нульмерная* цепь, и ее граница равна нулю по определению. Поясним соотношение (3) для двумерных цепей.

Рассмотрим сначала двумерную клетку g, являющуюся многоугольником, на границе которого не произведено никаких «склеиваний». Обойдем контур этой клетки в направлении, соответствующем

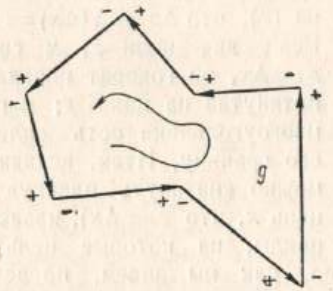


Рис. 150.

ее ориентации, и расставим на сторонах многоугольника стрелки, показывающие направление движения при таком обходе (рис. 150). Тогда все ребра, составляющие границу клетки g, получают определенную ориентацию, причем все коэффициенты примыкания равны $+1$. Следовательно, $\Delta g = r_1 + r_2 + \ldots + r_n$, где r_1, r_2, ..., r_n — все стороны многоугольника g, ориентированные указанным образом. Так как теперь в каждую вершину многоугольника g входит одно ребро, а выходит тоже одно ребро, то в цепь $\Delta(\Delta g) = \Delta r_1 + \Delta r_2 + \ldots + \Delta r_n$ каждая вершина войдет дважды: один раз с коэффициентом $+1$ (как конец входящего ребра) и один раз с коэффициентом -1 (как конец выходящего ребра). Следовательно, $\Delta(\Delta g) = 0$.

Аналогично дело обстоит и в случае более сложной двумерной клетки — многоугольника, на границе которого произведены склеивания (рис. 151): и в этом случае, обходя контур клетки, мы заметим, что для каждой вершины g число «входящих» ребер равно числу «выходящих» ребер, т. е. число «плюсов» равно числу «минусов», и потому каждая вершина входит в $\Delta(\Delta g)$ с коэффициентом нуль.

Мы пояснили соотношение (3) для одномерных и двумерных цепей. В топологии доказывается, что соотношение (3) имеет место для клеток и цепей любых размерностей.

Циклы и гомологии

Каждая цепь, граница которой равна нулю, называется *циклом*. В этом названии отражен тот факт, что если несколько направленных отрезков следуют друг за другом в «циклическом» порядке (например, так, как стороны многоугольника на рис. 150), то они образуют цепь, граница которой равна нулю (т. е. образуют «цикл»).

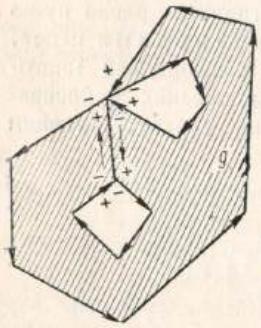


Рис. 151.

Из основного свойства (3) легко вытекает, что *граница любой цепи является циклом*. Действительно, взяв цепь $z=\Delta x$, являющуюся границей некоторой цепи x, мы легко получим из (3), что $\Delta z=\Delta(\Delta x)=0$, т. е. z есть цикл. Если две цепи z, x связаны соотношением $z=\Delta x$, то говорят иногда, что x есть *пленка*, натянутая на цикл z; например, внутренность многоугольника есть «пленка», натянутая на его границу. Итак, всякая цепь z, на которую можно «натянуть пленку» (т. е. найти такую цепь x, что $z=\Delta x$), является циклом. Однако существуют и такие циклы, на которые нельзя натянуть пленку. Например, тор можно, как мы знаем, представить в виде разбиения T, состоящего из четырех клеток: одной нульмерной, одной двумерной и двух одномерных клеток («меридиан» и «параллель»). При этом все коэффициенты примыкания равны нулю, т. е. граница любой клетки равна нулю. Отсюда следует, в частности, что граница любой двумерной цепи разбиения T равна нулю, и потому ни на какую отличную от нуля одномерную цепь в разбиении T натянуть пленку не удастся. В частности, и меридиан тора, и его параллель дают примеры циклов, на которые нельзя натянуть расположенную на торе пленку.

Таким образом, можно различать два вида циклов:

1) циклы «несущественные», т. е. циклы, на которые можно натянуть пленку (или иначе, циклы, которые служат границами некоторых цепей);

2) циклы «существенные», на которые нельзя натянуть пленку.

При этом если разность z_1-z_2 двух «существенных» циклов z_1, z_2 являются несущественным циклом (т. е. если найдется такая цепь x, что $z_1-z_2=\Delta x$), то циклы z_1 и z_2 считаются «несущественно отличающимися» друг от друга. В топологии говорят в этом случае, что циклы z_1 и z_2 *гомологичны* между собой.

Объединим теперь все циклы некоторого разбиения P в классы, считая два цикла принадлежащими одному классу, если они гомологичны (т. е. «несущественно» отличаются друг от друга). Мы получим некоторое количество *классов гомологий*; циклы, принадлежащие одному и тому же классу, гомологичны между собой; циклы, принадлежащие разным классам, не гомологичны. В частности, все циклы, на которые можно натянуть пленки, составляют один класс, а именно класс циклов, гомологичных нулевому циклу (ибо соотношение $z=\Delta x$ можно записать в виде $z-0=\Delta x$). Поэтому циклы, на которые можно натянуть пленки (т. е. «несущественные» циклы), называются также *циклами, гомологичными нулю*.

Если цикл z_1 гомологичен циклу z_1', а цикл z_2 гомологичен циклу z_2' (т. е. если $z_1-z_1'=\Delta x_1$, $z_2-z_2'=\Delta x_2$), то цикл z_1+z_2 гомологичен циклу $z_1'+z_2'$:

$$(z_1+z_2)-(z_1'+z_2')=(z_1-z_1')+(z_2-z_2')=\Delta x_1+\Delta x_2=\Delta(x_1+x_2).$$

Это дает возможность складывать между собой классы гомологий: если ξ_1 и ξ_2 — некоторые классы гомологий, а z_1 и z_2 — циклы, взятые из этих классов, то мы возьмем цикл z_1+z_2 и его класс гомологий обозначим через $\xi_1+\xi_2$. Из сказанного выше ясно, что если бы мы вместо z_1 и z_2 взяли другие циклы z_1' и z_2' из тех же классов ξ_1 и ξ_2, то снова получили бы цикл из класса $\xi_1+\xi_2$. Таким образом, класс гомологий, обозначенный через $\xi_1+\xi_2$, действительно зависит только от самих классов ξ_1 и ξ_2, а не от тех циклов z_1, z_2, которые мы в этих классах выбирали.

Определенное таким образом сложение превращает множество всех классов гомологий в группу. Нулем этой группы является класс циклов, гомологичных нулю. Построенная группа называется *группой гомологий* рассматриваемого разбиения P и обозначается через $H_r(P)$; число r показывает, какой размерности циклы нами рассматривались: при рассмотрении нульмерных циклов мы получаем *нульмерную группу* гомологий $H_0(P)$, при рассмотрении одномерных циклов — *одномерную группу* $H_1(P)$ и т. д.

Если мы обозначим через $Z_r(P)$ множество всех r-мерных циклов, а через $B_r(P)$ — множество всех r-мерных цепей, каждая из которых является границей некоторой $(r+1)$-мерной цепи, то заметим, что все множество $B_r(P)$ содержится в $Z_r(P)$ (так как граница любой цепи является циклом). При этом множества $Z_r(P)$ и $B_r(P)$ представляют

собой *группы*, ибо сумма и разность двух циклов являются циклами, а сумма и разность двух границ являются границами. Иначе говоря, $B_r(P)$ есть *подгруппа* группы $Z_r(P)$. Мы объединяем два цикла z_1, z_2 в один класс гомологий тогда и только тогда, когда разность $z_1 - z_2$ принадлежит подгруппе $B_r(P)$. Поэтому (см. стр. 137) *классы гомологий являются элементами фактор-группы* $Z_r(P)/B_r(P)$; сложение классов гомологий также совпадает со сложением, устанавливаемым в фактор-группе. Таким образом, поясненное выше определение групп гомологий можно сформулировать таким образом: *r-мерная группа гомологий* $H_r(P)$ *определяется как фактор-группа*

$$Z_r(P)/B_r(P). \tag{4}$$

Основным фактом теории гомологий являются (довольно сложно доказываемая) теорема о том, что *группы гомологий являются топологическими инвариантами*. Более полно: группа $H_r(P)$ не зависит от выбора разбиения полиэдра на клетки, а определяется самим этим полиэдром; гомеоморфные между собой полиэдры имеют одинаковые (изоморфные) группы гомологий. Отсюда, в частности, следует, что для вычисления групп гомологий некоторого полиэдра можно взять любое его разбиение на клетки; конечно, следует стараться найти возможно более простое разбиение, чтобы вычисление групп гомологий было наиболее легким.

Примеры

1. Вычислим группы гомологий (двумерной) сферы S^2. Самое простое разбиение сферы на клетки содержит только две клетки: одну вершину a и одну грань. Одномерных клеток это разбиение совсем не содержит; поэтому группа одномерных цепей $L_1(S^2)$ тривиальна (т. е. имеется лишь одна одномерная цепь — равная нулю). Поэтому лишь один нульмерный цикл, а именно равный нулю, является границей одномерной цепи. Но все нульмерные цепи являются циклами (по определению любая нульмерная цепь имеет нулевую границу), т. е. $Z_0(S^2) = L_0(S^2)$. Поскольку, кроме того, гомологичных нулю циклов (кроме нулевого) нет, то и группа гомологий $H_0(S^2)$ совпадает с $L_0(S^2)$. Это следует из того, что H_0 есть фактор-группа группы $Z_0 = L_0$ по подгруппе B_0, которая в нашем случае состоит лишь из одного элемента — нуля. Но каждая нульмерная цепь имеет вид αa, где α — целочисленный коэффициент. Поэтому группа $H_0(S^2)$ совпадает (изоморфна) с группой целых чисел (т. е. является свободной циклической группой).

Далее, из того, что $L_1(S^2) = 0$, вытекает, что все двумерные цепи являются циклами, т. е. $Z_2(S^2) = L_2(S^2)$; гомологичных же нулю двумерных цепей (кроме нулевой) нет (так как нет ни одной трехмерной клетки). Отсюда, как и выше, следует, что группа $H_2(S^2) = L_2(S^2)$ совпадает с группой целых чисел. Группа же $H_1(S^2)$ тривиальна, так как тривиальна группа цепей $L_1(S^2)$.

Условимся обозначать через Z группу целых чисел (свободную циклическую группу). Тогда мы можем написать

$$H_0(S^2) = H_2(S^2) = Z, \quad H_1(S^2) = 0. \tag{5}$$

[Совершенно так же доказывается, что гомологии n-мерной сферы S^n имеют вид $H_0(S^n) = H_n(S^n) = Z$ и $H_r(S^n) = 0$ при $r \neq 0$, $r \neq n$].

2. Вычислим теперь гомологии *тора* P_1. Используя рассмотренное ранее разбиение тора на клетки [1]), мы находим, что любая двумерная цепь имеет вид αg, любая одномерная — вид $\beta_1 r_1 + \beta_2 r_2$, а любая нульмерная — вид γa, где α, β_1, β_2, γ — целочисленные коэффициенты. Так как все коэффициенты примыкания равны нулю, то граница любой цепи равна нулю. Следовательно, все цепи являются циклами, а циклов, гомологичных нулю (кроме нулевого) — нет. Поэтому

$$H_0(P_1) = L_0(P_1), \quad H_1(P_1) = L_1(P_1), \quad H_2(P_1) = L_2(P_1),$$

так что $H_0(P_1)$ и $H_2(P_1)$ являются свободными циклическими группами, а группа $H_1(P_1)$ образована парами целых чисел, т. е. является *прямой суммой* двух групп Z (или, что то же самое, свободной абелевой группой с двумя образующими). Это можно записать следующим образом:

$$H_0(P_1) = H_2(P_1) = Z, \quad H_1(P_1) = Z + Z. \tag{6}$$

[Буквально так же, используя клеточное разбиение ориентируемой поверхности P_m [2]) (т. е. сферы с m ручками), мы получаем следующий результат:

$$H_0(P_m) = H_2(P_m) = Z, \quad H_1(P_m) = Z + Z + \ldots + Z (2m \text{ слагаемых}), \tag{7}$$

где $Z + Z + \ldots + Z$ обозначает прямую сумму $2m$ групп Z, т. е. свободную абелеву группу с $2m$ образующими.]

3. Изучим, наконец, гомологии проективной плоскости N_1. Как мы видели на стр. 114, существует разбиение проективной плоскости на три клетки a, r, g (размерностей 0, 1 и 2), причем коэффициент примыкания грани g к ребру r равен ± 2, а коэффициент примыкания ребра r к вершине a равен нулю. Мы выберем такие ориентации клеток, чтобы коэффициент примыкания грани g к ребру r был равен $+2$. Цепи размерностей 0, 1 и 2 в этом клеточном разбиении имеют вид

$$x_0 = \alpha a, \quad x_1 = \beta r, \quad x_2 = \gamma g,$$

где α, β, γ — целочисленные коэффициенты, а их границы легко вычислить, зная коэффициенты примыкания:

$$\Delta x_0 = 0, \quad \Delta x_1 = 0, \quad \Delta x_2 = 2\gamma r. \tag{8}$$

---

[1]) См. «Математическое просвещение», вып. 4, стр. 42.
[2]) См. там же, стр. 44.

Первые два из соотношений (8) показывают, что все нульмерные цепи являются циклами, а циклов, гомологичных нулю (кроме нулевого), нет. Таким образом, $H_0(N_1)=Z$. Последнее из соотношений (8) показывает, что если $x_2\neq 0$ (т. е. $\gamma\neq 0$), то и $\Delta x_2\neq 0$, т. е. никакая отличная от нуля двумерная цепь не является циклом, и потому $H_2(N_1)=0$. Остается вычислить одномерную группу гомологий. Второе из соотношений (8) показывает, что все одномерные цепи являются циклами; из них гомологичны нулю те, которые (как показывает третье из соотношений (8)) имеют четные коэффициенты. Таким образом, $H_1(N_1)$ есть фактор-группа группы Z целых чисел по подгруппе четных чисел, т. е. $H_1(N_1)$ есть *группа вычетов по модулю два*. Итак,

$$H_0(N_1)=Z,\quad H_1(N_1)=Z_2,\quad H_2(N_1)=0, \tag{9}$$

где Z_2 есть обозначение группы вычетов по модулю два.

Предоставляем читателю установить, что проективная плоскость по модулю m — обозначим ее через Q_m — имеет следующие гомологии:

$$H_0(Q_m)=Z,\quad H_1(Q_m)=Z_m,\quad H_2(Q_m)=0, \tag{10}$$

где Z_m — группа вычетов по модулю m.

Обратим внимание читателя на тот факт, что *нульмерная* группа гомологий была во всех рассмотренных выше примерах свободной циклической. Оказывается, что вообще для любого *связного* полиэдра нульмерная группа гомологий равна Z.

Если же полиэдр P не связен, т. е. состоит из нескольких *компонент*[1]), то группа $H_0(P)$ имеет вид $Z+Z+\ldots+Z$, причем число слагаемых в этой прямой сумме равно числу k компонент рассматриваемого полиэдра (т. е. $H_0(P)$ есть свободная абелева группа с k образующими). Предоставим читателю самостоятельно доказать эти утверждения.

Числа Бетти и эйлерова характеристика

В общем случае группа гомологий $H_r(P)$, где P — некоторое разбиение, имеет вид $H_r(P)=Z+Z+\ldots+Z+K$, где K — конечная группа. Число слагаемых Z в этой прямой сумме (т. е. *ранг* группы $H_r(P)$) называется *r-мерным числом Бетти*[2]) разбиения P и обозначается через p_r.

Например, числа Бетти ориентируемой поверхности P_m имеют следующие значения [см. (7)]:

$$p_0=1,\quad p_1=2m,\quad p_2=1. \tag{11}$$

[1]) См. «Математическое просвещение», вып. 2, стр. 11.

[2]) Термин *числа Бетти* был введен А. Пуанкаре, который приписал итальянскому математику Бетти свои замечательные открытия.

Далее, проективная плоскость имеет следующие числа Бетти [см. (9)]:

$$p_0=1,\quad p_1=p_2=0. \tag{12}$$

Как и группы гомологий, числа Бетти являются топологическими инвариантами рассматриваемого полиэдра и, в частности, не зависят от выбора клеточного разбиения.

В главе 3 мы рассмотрели важный топологический инвариант, называемый *эйлеровой характеристикой*. Для двумерного разбиения P эйлерова характеристика $\chi(P)$ определялась равенством

$$\chi(P)=\Gamma-P+B,$$

где Γ — число граней, P — число ребер, B — число вершин. В общем случае эйлерова характеристика определяется равенством

$$\chi(P)=\sum_r(-1)^r\alpha_r,$$

где α_r есть число r-мерных клеток разбиения P (так что $\alpha_0=B$, $\alpha_1=P$, $\alpha_2=\Gamma$). Оказывается, что *эйлерова характеристика может быть вычислена с помощью чисел Бетти; именно, имеет место формула*

$$\chi(P)=\sum_r(-1)^r p_r, \tag{13}$$

где p_r есть r-мерное число Бетти полиэдра P. Так как числа Бетти топологически инвариантны, то из формулы (13) следует, что *эйлерова характеристика также является топологическим инвариантом*. Отметим, что с помощью формулы (13) мы можем легко вычислить эйлеровы характеристики поверхностей P_m и N_1. Действительно, используя формулы (11), (12), получаем

$$\left.\begin{aligned}\chi(P_m)&=1-2m+1=2-2m,\\ \chi(N_1)&=1-0+0=1\end{aligned}\right\} \tag{14}$$

в полном соответствии с вычислениями, проведенными в главе 3.

§ 9. Некоторые приложения теории гомологий

Выше было показано, что эйлерова характеристика является понятием гомологическим. Другим важным гомологическим понятием является понятие степени отображения, о которой будет сказано в этой главе. В дальнейшем мы рассмотрим интересные геометрические приложения этих понятий.

Векторные поля на поверхности

Важным примером применения эйлеровой характеристики служит теорема Пуанкаре о векторных полях на поверхностях. Рассмотрим какую-либо ориентируемую поверхность P, гладко расположенную

в пространстве, т. е. такую поверхность, которая в каждой своей точке имеет касательную плоскость (или, иначе, не имеет линий и точек «излома» — таких, например, как ребра и вершины многогранников). Примерами гладких поверхностей могут служить обычная сфера, тор или «крендель» (рис. 137, 140 на стр. 107, 109). Поставим теперь следующую задачу: можно ли на всей поверхности P

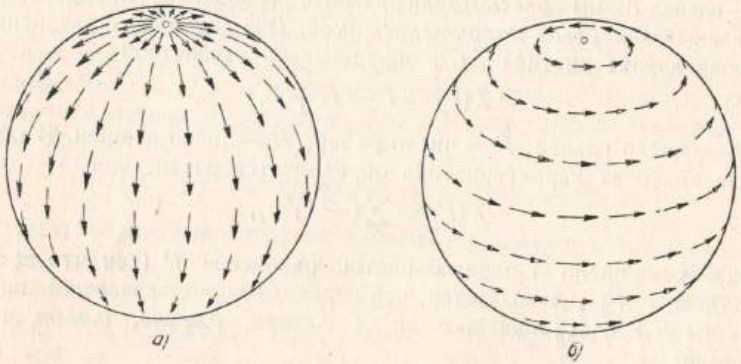


Рис. 152.

построить непрерывное поле направлений, т. е. выбрать в каждой точке поверхности такое касательное к поверхности направление (изображаемое стрелкой на этой поверхности), что при переходе от точки к точке направление стрелки меняется непрерывно? Если, например, на сфере мы возьмем стрелки, идущие в каждой точке с севера на юг (рис. 152, *а*), то всюду, кроме полюсов, мы получим непрерывное поле направлений, но в полюсах мы не сможем выбрать направления, непрерывно связанные с направлениями в близких точках. То же будет, если в каждой точке мы возьмем направление стрелки с запада на восток (рис. 152, *б*). Таким образом, построение непрерывного поля направлений *на всей сфере* встречает затруднения. На торе же поле направлений строится легко — см., например, рис. 153.

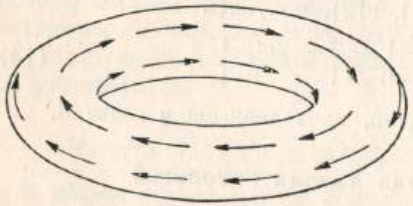


Рис. 153.

Вместо полей направлений можно рассматривать *поля касательных векторов*, сформулировав задачу следующим образом: можно ли в каждой точке поверхности P выбрать такой ненулевой вектор, касающийся поверхности в этой точке, что при переходе от точки к точке эти касательные векторы меняются непрерывно?

Решение этих задач было дано Пуанкаре: *на поверхности P тогда и только тогда можно выбрать непрерывное поле направлений (или непрерывное поле ненулевых касательных векторов), когда эйлерова характеристика поверхности равна нулю.* Из формулы (14) следует теперь, что *единственной замкнутой ориентируемой поверхностью, на которой можно задать непрерывное поле направлений, является поверхность P_1*, т. е. тор. На сфере же невозможно задать непрерывное поле направлений. Этот факт иногда формулируют в виде следующей «теоремы о еже»: если из каждой точки поверхности сферы растет «колючка» (ненулевой вектор, не обязательно касающийся сферы) и направления «колючек» от точки к точке меняются непрерывно, то найдется хотя бы одна «колючка», направленная перпендикулярно к сфере. Действительно, если бы все «колючки» были направлены не перпендикулярно к сфере, то, спроектировав каждую «колючку» MK на касательную плоскость, проведенную в той точке, из которой эта «колючка» растет (рис. 154), мы получили бы на всей сфере непрерывное поле ненулевых касательных векторов, а это невозможно.

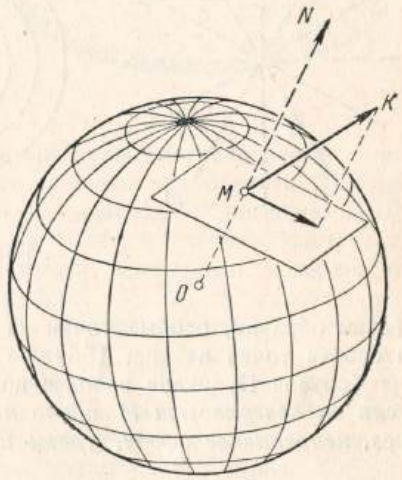


Рис. 154.

Как мы говорили выше, при попытке построения поля направлений *на всей сфере* возникают затруднения. Так, на рис. 152 показаны поля направлений, определенные всюду, кроме полюсов. В полюсах же выбрать определенное направление не удается, так как вблизи этих точек стрелки направлены в разные стороны; такие точки называются особыми точками поля направлений. Других особых точек (кроме полюсов) у полей направлений, изображенных на рис. 152, нет. На рис. 155, *а* и *б* показан вид этих полей направлений вблизи северного полюса, а на рис. 155, *в* изображен более сложный пример особой точки (так называемое седло). Проведем теперь маленькую окружность с центром в особой точке и посмотрим, как направлены векторы в точках этой окружности (рис. 156, *а*, *б*, *в*). Если мы один раз обойдем эту окружность против часовой стрелки, то заметим, что в случаях *а*, *б* направления векторов также совершат один поворот против часовой стрелки, а в случае *в* — один поворот, но уже по часовой стрелке. Вообще, если мы обойдем маленькую окружность с центром в особой точке, то

векторы, имеющиеся вдоль этой окружности, совершат некоторое целое число k оборотов (которое мы считаем положительным, если векторы совершают несколько оборотов в *ту же сторону*, в какую мы движемся по окружности, и отрицательным — в противном случае). Это число k называется *индексом* рассматриваемой особой точки.

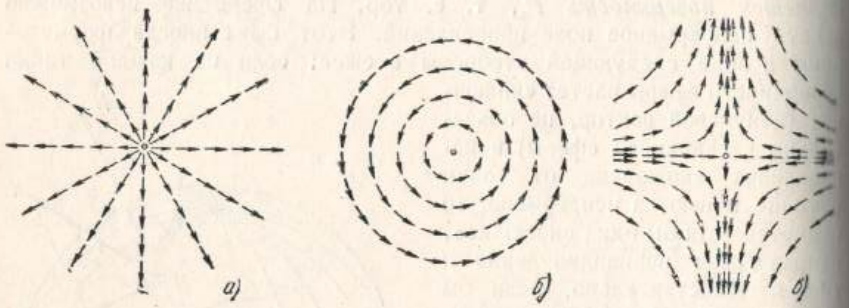


Рис. 155.

Таким образом, особые точки на рис. 156, *а* и *б* имеют индекс $+1$, а особая точка на рис. 156, *в* — индекс -1.

Теорема Пуанкаре в более полной формулировке утверждает, что *если на поверхности P задано поле ненулевых касательных векторов, непрерывное всюду, кроме конечного числа особых точек, то*

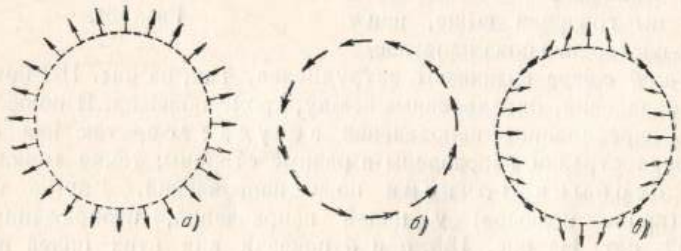


Рис. 156.

сумма индексов всех особых точек равна эйлеровой характеристике рассматриваемой поверхности.

Интересные и важные обобщения этой теоремы Пуанкаре были найдены за рубежом Хопфом, Штифелем, Уитни, Стинродом и другими математиками, а у нас — Л. С. Понтрягиным и его учениками М. М. Постниковым, В. А. Рохлиным, В. Г. Болтянским и др.).

Степень отображения и теорема Гаусса—Бонне

Пусть P и Q — две ориентируемые поверхности и f — непрерывное отображение поверхности P на поверхность Q. Отображение f мы не предполагаем гомеоморфным; можно представлять себе, что поверхность P как-либо «наложена» на поверхность Q, быть может, располагаясь на ней несколькими «слоями» и образуя складки. Мы будем предполагать, что поверхности P и Q *ориентированы*, т. е. в каждой точке поверхности выбрано направление обхода (задаваемое стрелкой на окружности и согласованное в близких точках поверхности, рис. 157). Если некоторый кусок поверхности P с помощью отображения f накладывается на кусок поверхности Q таким образом, что стрелки, соответствующие ориентации поверхности P, переходят в стрелки, соответствующие ориентации поверхности Q (рис. 158, *а*), то говорят, что на рассматриваемом куске поверхности P *отображение f положительно*. Если же получающиеся направления стрелок противоположны тем, которые имеются на поверхности Q (рис. 158, *б*), то на рассматриваемом куске поверхности отображение f *отрицательно*. Пусть, например, некоторый кусок поверхности P при отображении f образует складку (рис. 159); тогда на одной части этого куска отображение f положительно, а на другой — отрицательно. Возьмем теперь на поверхности Q некоторую точку y и маленький кружок K, содержащий точку y. Возьмем все те куски поверхности P, которые отображаются на кружок K, и предположим, что все они отображаются на K гомеоморфно, причем число тех кусков, на которых отображение f положительно, равно p, а число кусков, на которых отображение f отрицательно, равно n. В этом случае число $p-n$ называется степенью отображения f

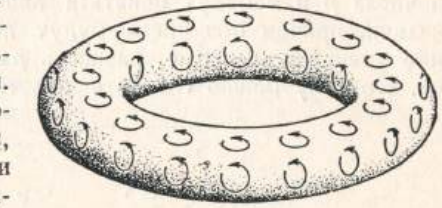


Рис. 157.

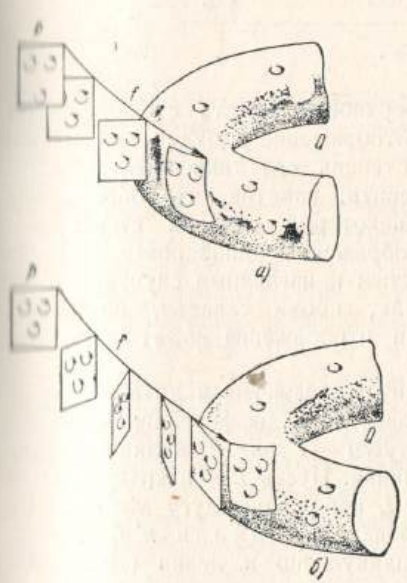


Рис. 158.

вблизи точки y. Иначе говоря, степень отображения равна числу слоев, положительно накладывающихся на точку y, минус число слоев, отрицательно покрывающих точку y. Нетрудно понять, что степень отображения f *одинакова* вблизи любой точки поверхности Q. Действительно, если мы будем перемещать точку y по поверхности Q, то числа p и n будут меняться лишь при прохождении через край складки, причем оба числа будут либо увеличиваться на единицу, либо уменьшаться, т. е. разность $p-n$ будет оставаться неизменной. Поэтому можно говорить просто о *степени отображения*, не

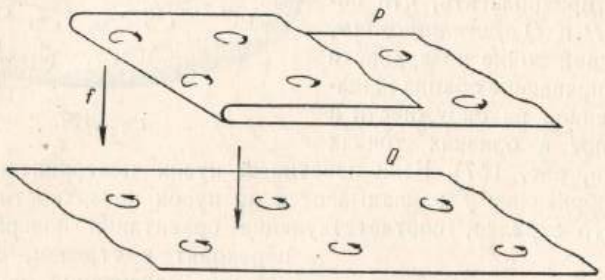


Рис. 159.

указывая, вблизи какой точки поверхности Q эта степень вычисляется. Заметим еще, что если отображение f будет непрерывно меняться, «деформироваться», то степень его также будет оставаться неизменной; это можно пояснить, заметив, что образование новых складок или расправление некоторых складок степени не меняет. (При описании степени отображения одной поверхности на другую мы ограничились самым простым и наглядным случаем, когда отображение никаких «особенностей», кроме складок, не имеет; общее и точное определение степени отображения может быть дано с помощью гомологий.)

Применим теперь понятие степени отображения к доказательству одной изящной формулы, которая известна в дифференциальной геометрии под названием *теоремы Гаусса — Бонне*. Вспомним прежде всего понятие кривизны кривой линии. Пусть L — некоторая линия и x — ее точка. Возьмем на линии L небольшую дугу $x'x''$, заключающую внутри себя точку x, и проведем нормали $x'y'$ и $x''y''$, т. е. векторы, направленные перпендикулярно к линии (рис. 160). Если α — угол между этими нормалями, а s — длина дуги $x'x''$, то отношение $\frac{\alpha}{s}$ выражает «среднюю скорость поворота» нормали на каждую единицу длины дуги; отношение $\frac{\alpha}{s}$ называют *средней кривизной* дуги $x'x''$ линии L. Если мы будем брать все меньшие и меньшие дуги линии L, заключающие точку x, то средние кривизны этих

дуг будут приближаться к некоторому значению; это значение (предел средней кривизны при условии, что длина дуги $x'x''$ приближается к нулю) называется *кривизной* линии L в точке x. Немного видоизменив это определение, можно описать кривизну следующим образом. Проведем окружность C радиуса 1 с центром в некоторой точке O. Каждой точке x линии L поставим в соответствие такую точку $z=f(x)$ окружности C, что отрезок Oz параллелен нормали к линии L в точке x. В результате каждой точке линии L поставлена в соответствие некоторая точка окружности, т. е. мы получаем некоторое *отображение* f линии L на окружность. Если z' и z'' — те точки

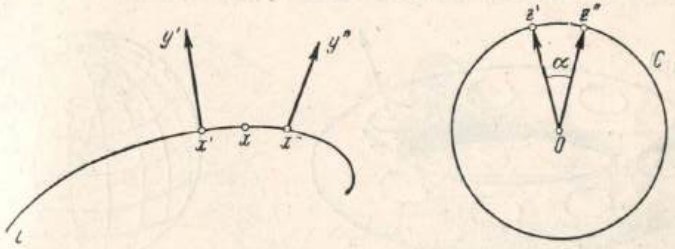


Рис. 160.

окружности C, которые соответствуют точкам x' и x'' линии L, то длина дуги $z'z''$ численно равна углу $z'Oz''$, т. е. углу α между нормалями $x'y'$ и $x''y''$ (рис. 160). Поэтому отношение $\frac{\alpha}{s}$ равно отношению длин дуг $z'z''$ и $x'x''$. Итак, если $x'x''$ — дуга линии L, содержащая точку x, а $z'z''=f(x'x'')$ ее образ при отображении f, то отношение длин дуг $z'z''$ и $x'x''$ стремится к кривизне линии L в точке x, если дуга $x'x''$ стягивается к точке x.

Аналогично определяется *кривизна поверхности*. Пусть Π — некоторая поверхность. Для того чтобы определить ее кривизну в некоторой точке x, мы рассмотрим *гауссово сферическое* отображение поверхности Π на сферу радиуса 1. Строится оно следующим образом. Выберем некоторую ориентацию поверхности Π и из каждой точки x этой поверхности проведем к ней нормаль $N(x)$, т. е. единичный вектор, направленный таким образом, что из его конца стрелка, соответствующая ориентации поверхности, наблюдается идущей против часовой стрелки (рис. 161). Возьмем теперь сферу S единичного радиуса и из ее центра проведем отрезок, равный и параллельный нормали $N(x)$. Конец этого отрезка (очевидно, лежащий на сфере S) обозначим через $f(x)$. Таким образом, каждой точке x поверхности Π поставлена в соответствие некоторая точка $f(x)$ сферы S, т. е. мы получаем определенное отображение (причем непрерывное) поверхности Π на сферу S. Оно и называется гауссовым сферическим отображением.

Пусть x — некоторая точка поверхности Π, M — некоторый кусок поверхности Π, содержащий точку x, а $f(M)$ — образ этого куска при сферическом отображении f. Обозначим через $S(M)$ и $S(f(M))$ площади кусков M и $f(M)$, причем площадь куска $f(M)$ будем брать со знаком: если отображение f положительно на куске M, то будем считать площадь куска $f(M)$ положительной, а если отображение f отрицательно, то площадь куска $f(M)$ будем считать отрицательной. Тогда отношение $\frac{S(f(M))}{S(M)}$ называется *средней кривизной* куска M поверхности Π. Когда кусок M стягивается к точке x, это отношение приближается к некоторому пределу, который и называется кривизной поверхности Π (*гауссовой кривизной*) в точке x. Таким образом, обозначая кривизну поверхности Π в точке x через $K(x)$, мы сможем написать

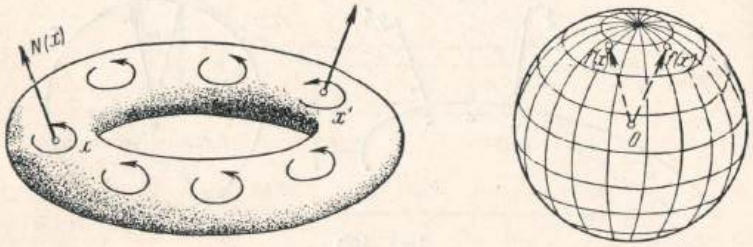


Рис. 161.

$$S(M)\,K(x) \approx S(f(M)),$$

причем это равенство тем более точно (т. е. отношение обеих частей тем ближе к единице), чем меньше по размерам кусок M, содержащий точку x.

Заметим, что если гауссова кривизна поверхности Π в точке x положительна, то вблизи точки x вся поверхность расположена по одну сторону от своей касательной плоскости, т. е. выпукла (рис. 162); такие точки поверхности называются *эллиптическими*. Если же кривизна отрицательна, то поверхность имеет вблизи точки x вид седла (рис. 163), пересекая свою касательную плоскость по двум линиям; такие точки называются *гиперболическими*. Существуют также *параболические* точки — точки, в которых гауссова кривизна равна нулю.

Итак, в каждой точке x поверхности Π определена гауссова кривизна $K(x)$, т. е. на поверхности Π задана функция $K(x)$. Эта функция непрерывна (предполагается, что поверхность задана функциями, которые имеют вторые непрерывные производные). Если поверхность Π является замкнутой (т. е. не имеет края) и ориентируемой, то можно рассматривать двойной интеграл $\iint_{\Pi} K\,ds$ от функции $K(x)$, взятый по всей поверхности Π (ds — элемент площади поверхности). Теорема Гаусса — Бонне утверждает, что *для любой замкнутой ориентируемой поверхности Π имеет место формула*[1])

$$\iint_{\Pi} K\,ds = 2\pi\chi,$$

где χ — эйлерова характеристика поверхности Π. Эта формула выражает замечательный факт геометрии: хотя значение гауссовой кривизны в каждой точке существенно зависит от строения (т. е. искривления формы) поверхности вблизи этой точки, однако интеграл гауссовой кривизны по всей поверхности не зависит от формы поверхности, не зависит от того, как она изогнута, расположена в пространстве и т. п., а зависит только от *топологического* строения этой поверхности.

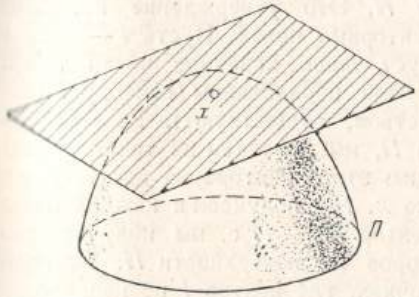


Рис. 162.

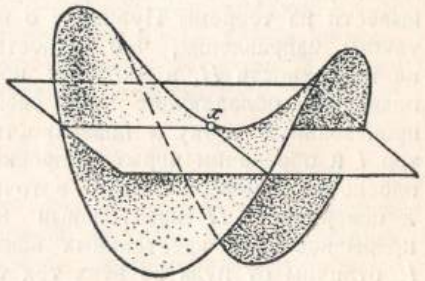


Рис. 163.

Дадим теперь другую интерпретацию формулы Гаусса — Бонне, связанную с понятием степени отображения. Для этого разобьем поверхность Π на мелкие куски $M_1, M_2, \ldots, M_n$ и внутри них выберем точки $x_1, x_2, \ldots, x_n$. Тогда

$$\sum_{i=1}^{n} S(M_i)\,K(x_i) \approx \sum_{i=1}^{n} S(f(M_i)). \qquad (*)$$

Слева стоит интегральная сумма, которая тем ближе к $\iint K\,ds$, чем мельче куски $M_1, M_2, \ldots, M_n$. Справа же суммируются площади кусков $f(M_i)$, покрывающих сферу столько раз, какова степень отображения f [напомним, что площадь $S(f(M_i))$ берется со знаком «$+$», если кусок M_i отображается на $f(M_i)$ положительно, и

[1]) Формула впервые опубликована Бонне в работе 1848 г., однако, по существу, она была известна еще Гауссу.

со знаком « — » в противном случае]. Таким образом, правая часть соотношения (*) равна $\gamma\sigma$, где γ — степень сферического отображения f, а σ — площадь сферы S, т. е. эта правая часть равна $4\pi\gamma$. Итак,

$$\iint_{\Pi} K\,ds = 4\pi\gamma.$$

Сравнивая это соотношение с формулой Гаусса — Бонне, мы видим, что для доказательства последней достаточно установить соотношение $\gamma = \frac{\chi}{2}$, т. е. установить, что *степень сферического отображения поверхности Π на сферу S равна половине эйлеровой характеристики поверхности Π*. Это утверждение нетрудно вывести из теоремы Пуанкаре о векторных полях. Пусть l — вектор такого направления, что существует лишь конечное число точек на поверхности Π, в которых нормали параллельны вектору l (направления, обладающие этим свойством, существуют). Взяв теперь произвольную точку x поверхности Π, мы проведем из точки x вектор l и обозначим через l_x проекцию этого вектора на касательную плоскость к поверхности Π в точке x. Мы получаем в каждой точке x поверхности Π касательный вектор l_x, т. е. мы получаем непрерывное поле касательных векторов на поверхности Π. Векторы l_x отличны от нуля во всех тех точках, где вектор l не параллелен нормали к поверхности (всюду, кроме конечного числа точек). Таким образом, построенное векторное поле имеет лишь конечное число особых точек. Построим точно таким же способом векторное поле l_x и на сфере S. Очевидно, на сфере мы получим поле, изображенное на рис. 152, *а* с двумя особыми точками в «полюсах» сферы. Очевидно также, что в точке x поверхности Π и в точке $f(x)$ сферы S касательные векторы l_x и $l_{f(x)}$ *параллельны*.

Пусть теперь $a_1, a_2, \ldots, a_q$ — все особые точки векторного поля l_x на поверхности Π. Возьмем точку a_i. Тогда $f(a_i)$ также есть особая точка векторного поля на сфере, т. е. либо ее северный, либо южный полюс. Если $K(a_i) > 0$, то сферическое отображение f вблизи точки a_i положительно. Поэтому если точка x обходит окружность с центром a_i против часовой стрелки, то точка $f(x)$ обходит окружность с центром $f(a_i)$ также против часовой стрелки. Векторы же l_x и $l_{f(x)}$ параллельны и потому при обходе этих окружностей поворачиваются одинаково. Таким образом, индексы особых точек a_i и $f(a_i)$ совпадают, т. е. индекс особой точки a_i равен $+1$. Если же $K(a_i) < 0$, то сферическое отображение f вблизи точки a_i отрицательно. Поэтому если точка x обходит окружность с центром a_i против часовой стрелки, то точка $f(x)$ обходит окружность с центром $f(a_i)$ по часовой стрелке. Векторы же l_x и $l_{f(x)}$ при этих обходах по-прежнему поворачиваются одинаково. Следовательно, если обе окружности [с центрами a_i и $f(a_i)$]

обходить против часовой стрелки, то число поворотов вектора l_x будет отличаться знаком от числа поворотов вектора $l_{f(x)}$. Таким образом, в этом случае индексы особых точек a_i и $f(a_i)$ отличаются знаком, т. е. индекс особой точки равен -1.

Итак, точки $a_1, \ldots, a_q$ — это все те точки, которые при сферическом отображении f отображаются в полюсы сферы, причем индексы тех точек, вблизи которых отображение f положительно, равны $+1$, а индексы точек, вблизи которых оно отрицательно, равны -1. Поэтому если p' и p'' — числа тех из точек $a_1, \ldots, a_q$, которые положительно отображаются в северный и южный полюсы соответственно, а n' и n'' — числа отрицательно отображающихся точек, то сумма индексов всех особых точек $a_1, \ldots, a_q$ равна $(p' + p'') - (n' + n'') = (p' - n') + (p'' - n'') = \gamma + \gamma = 2\gamma$ (ибо $p' - n'$ есть степень отображения f в северном полюсе, а $p'' - n''$ — в южном). Но по теореме Пуанкаре сумма индексов особых точек равна χ. Таким образом, $\chi = 2\gamma$; это завершает доказательство теоремы Гаусса—Бонне.

Основная теорема алгебры

С помощью понятия степени отображения можно дать изящное геометрическое доказательство основной теоремы алгебры. Будем изображать комплексное число $z = x + iy$ точкой в плоскости, считая x и y координатами этой точки (рис. 164). Возьмем сферу S, касающуюся плоскости в начале координат, и будем называть точку касания южным полюсом этой сферы, а противоположную точку N сферы — ее северным полюсом (рис. 165). Пусть теперь $z = x + iy$ — произвольное комплексное число, т. е. точка в плоскости. Отрезок Nz пересекает сферу S в некоторой точке, которую мы также будем обозначать через $z = x + iy$ и считать изображением комплексного числа z на сфере S. В результате все комплексные числа оказываются изображенными на сфере S. Обратно, имея на сфере точку a, легко узнать, какое комплексное число она изображает: надо провести прямую Na, которая при пересечении с плоскостью и даст искомое комплексное число. Однако одна точка сферы S, а именно ее северный полюс N, не изображает никакого комплексного числа. Мы условимся считать, что точка N изображает особое, «бесконечное» комплексное число, обозначаемое символом ∞. Поводом для такого соглашения служит то обстоятельство, что если точка z на плоскости неограниченно удаляется (в любую сторону) от начала координат, то изображающая ее точка на сфере приближается к северному полюсу N. Итак, каждая точка сферы S изображает некоторое

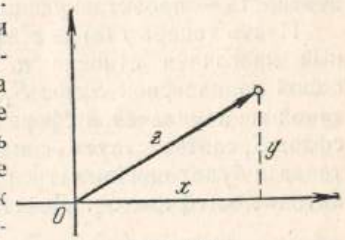


Рис. 164.

со знаком « — » в противном случае]. Таким образом, правая часть соотношения (*) равна $\gamma\sigma$, где γ — степень сферического отображения f, а σ — площадь сферы S, т. е. эта правая часть равна $4\pi\gamma$. Итак,

$$\iint_{\Pi} K\,ds = 4\pi\gamma.$$

Сравнивая это соотношение с формулой Гаусса — Бонне, мы видим, что для доказательства последней достаточно установить соотношение $\gamma = \frac{\chi}{2}$, т. е. установить, что *степень сферического отображения поверхности Π на сферу S равна половине эйлеровой характеристики поверхности Π*. Это утверждение нетрудно вывести из теоремы Пуанкаре о векторных полях. Пусть l — вектор такого направления, что существует лишь конечное число точек на поверхности Π, в которых нормали параллельны вектору l (направления, обладающие этим свойством, существуют). Взяв теперь произвольную точку x поверхности Π, мы проведем из точки x вектор l и обозначим через l_x проекцию этого вектора на касательную плоскость к поверхности Π в точке x. Мы получаем в каждой точке x поверхности Π касательный вектор l_x, т. е. мы получаем непрерывное поле касательных векторов на поверхности Π. Векторы l_x отличны от нуля во всех тех точках, где вектор l не параллелен нормали к поверхности (всюду, кроме конечного числа точек). Таким образом, построенное векторное поле имеет лишь конечное число особых точек. Построим точно таким же способом векторное поле l_x и на сфере S. Очевидно, на сфере мы получим поле, изображенное на рис. 152, *а* с двумя особыми точками в «полюсах» сферы. Очевидно также, что в точке x поверхности Π и в точке $f(x)$ сферы S касательные векторы l_x и $l_{f(x)}$ *параллельны*.

Пусть теперь $a_1, a_2, \ldots, a_q$ — все особые точки векторного поля l_x на поверхности Π. Возьмем точку a_i. Тогда $f(a_i)$ также есть особая точка векторного поля на сфере, т. е. либо ее северный, либо южный полюс. Если $K(a_i) > 0$, то сферическое отображение f вблизи точки a_i положительно. Поэтому если точка x обходит окружность с центром a_i против часовой стрелки, то точка $f(x)$ обходит окружность с центром $f(a_i)$ также против часовой стрелки. Векторы же l_x и $l_{f(x)}$ параллельны и потому при обходе этих окружностей поворачиваются одинаково. Таким образом, индексы особых точек a_i и $f(a_i)$ совпадают, т. е. индекс особой точки a_i равен $+1$. Если же $K(a_i) < 0$, то сферическое отображение f вблизи точки a_i отрицательно. Поэтому если точка x обходит окружность с центром a_i против часовой стрелки, то точка $f(x)$ обходит окружность с центром $f(a_i)$ по часовой стрелке. Векторы же l_x и $l_{f(x)}$ при этих обходах по-прежнему поворачиваются одинаково. Следовательно, если обе окружности [с центрами a_i и $f(a_i)$]

обходить против часовой стрелки, то число поворотов вектора l_x будет отличаться знаком от числа поворотов вектора $l_{f(x)}$. Таким образом, в этом случае индексы особых точек a_i и $f(a_i)$ отличаются знаком, т. е. индекс особой точки равен -1.

Итак, точки $a_1, \ldots, a_q$ — это все те точки, которые при сферическом отображении f отображаются в полюсы сферы, причем индексы тех точек, вблизи которых отображение f положительно, равны $+1$, а индексы точек, вблизи которых оно отрицательно, равны -1. Поэтому если p' и p'' — числа тех из точек $a_1, \ldots, a_q$, которые положительно отображаются в северный и южный полюсы соответственно, а n' и n'' — числа отрицательно отображающихся точек, то сумма индексов всех особых точек $a_1, \ldots, a_q$ равна $(p'+p'')-(n'+n'') = (p'-n')+(p''-n'') = \gamma+\gamma = 2\gamma$ (ибо $p'-n'$ есть степень отображения f в северном полюсе, а $p''-n''$ — в южном). Но по теореме Пуанкаре сумма индексов особых точек равна χ. Таким образом, $\chi = 2\gamma$; это завершает доказательство теоремы Гаусса—Бонне.

Основная теорема алгебры

С помощью понятия степени отображения можно дать изящное геометрическое доказательство основной теоремы алгебры. Будем изображать комплексное число $z = x + iy$ точкой в плоскости, считая x и y координатами этой точки (рис. 164). Возьмем сферу S, касающуюся плоскости в начале координат, и будем называть точку касания южным полюсом этой сферы, а противоположную точку N сферы — ее северным полюсом (рис. 165). Пусть теперь $z = x + iy$ — произвольное комплексное число, т. е. точка в плоскости. Отрезок Nz пересекает сферу S в некоторой точке, которую мы также будем обозначать через $z = x + iy$ и считать изображением комплексного числа z на сфере S. В результате все комплексные числа оказываются изображенными на сфере S. Обратно, имея на сфере точку a, легко узнать, какое комплексное число она изображает: надо провести прямую Na, которая при пересечении с плоскостью и даст искомое комплексное число. Однако одна точка сферы S, а именно ее северный полюс N, не изображает никакого комплексного числа. Мы условимся считать, что точка N изображает особое, «бесконечное» комплексное число, обозначаемое символом ∞. Поводом для такого соглашения служит то обстоятельство, что если точка z на плоскости неограниченно удаляется (в любую сторону) от начала координат, то изображающая ее точка на сфере приближается к северному полюсу N. Итак, каждая точка сферы S изображает некоторое

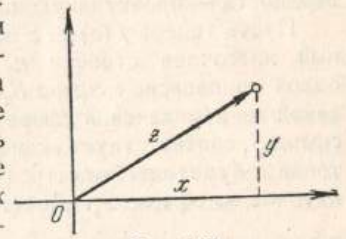


Рис. 164.

комплексное число, причем точка N изображает число ∞, а все остальные точки сферы изображают обычные («конечные») комплексные числа. Сфера S называется *комплексной сферой* или *сферой Римана*. Отметим, что сфера S получилась из плоскости добавлением *одной* бесконечно удаленной точки ∞. Как мы видели выше [1]), иное присоединение несобственных точек к плоскости дает другую поверхность — проективную плоскость.

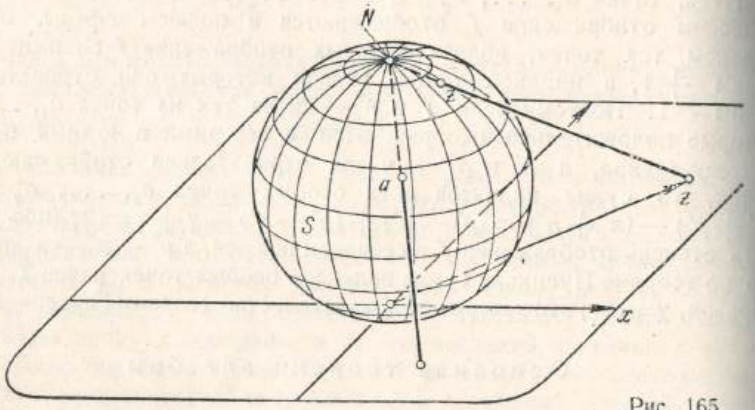


Рис. 165.

Пусть теперь $f(z)=z^n+a_1z^{n-1}+\ldots+a_{n-1}z+a_n$ — произвольный многочлен степени n. Мы будем изображать значения z на одной комплексной сфере S_1, а значения многочлена $f(z)$ — на другой такой же комплексной сфере S_2. Каждой «конечной» точке $z=x+iy$ сферы S_1 соответствует «конечная» точка $f(z)$ сферы S_2. При этом если точка z будет приближаться к ∞, то $f(z)$ также будет приближаться к точке ∞ сферы S_2. Действительно, записав многочлен $f(z)$ в виде

$$f(z)=z^n\left(1+\frac{a_1}{z}+\frac{a_2}{z^2}+\ldots+\frac{a_{n-1}}{z^{n-1}}+\frac{a_n}{z^n}\right),$$

мы легко заметим, что при $z\to\infty$ (т. е. при неограниченном увеличении числа $|z|$) выражение в скобках будет приближаться к единице, а множитель z^n неограниченно увеличивается. Таким образом, дополнив определение многочлена условием $f(\infty)=\infty$, мы получаем *непрерывное* отображение f всей сферы S_1 на сферу S_2.

Для того чтобы установить, что уравнение

$$z^n+a_1z^{n-1}+a_2z^{n-2}+\ldots+a_{n-1}z+a_n=0$$

имеет хотя бы одно решение (как утверждает основная теорема алгебры), нужно показать, что найдется такая точка z сферы S_1, для которой $f(z)=0$, т. е. что точка 0 сферы S_2 является образом хотя бы одной точки z сферы S_1. Если бы это было не так, т. е. точка 0 сферы S_2 не покрывалась образом $f(S_1)$ сферы S_1, то степень отображения f вблизи точки 0 сферы S_2 была бы равна нулю, а так как степень одинакова вблизи любой точки, то просто степень отображения f была бы равна нулю. Поэтому для доказательства основной теоремы алгебры достаточно установить, что степень отображения f отлична от нуля. Мы покажем, что эта степень отображения равна n (т. е. совпадает со степенью многочлена $f(z)$; это и послужило причиной введения термина «степень отображения»).

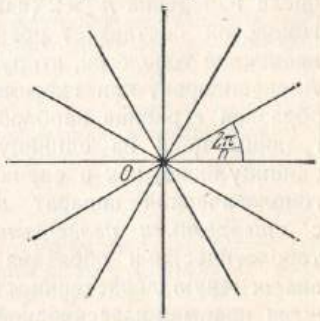


Рис. 166.

Будем изменять значения всех коэффициентов $a_1, a_2, \ldots, a_{n-1}, a_n$, приближая их к нулю. Тогда многочлен $f(z)$ будет меняться, отображение f сферы S_1 на S_2 будет непрерывно изменяться, деформироваться. (Непрерывность перемещения точек $f(z)$ при изменении коэффициентов очевидна для всех «конечных» точек z сферы S_1; что же касается точки ∞, то соотношение $f(\infty)=\infty$ все время сохраняется, так как старший коэффициент многочлена остается равным единице.) В результате такого изменения коэффициентов мы получим многочлен

$$f_1(z)=z^n.$$

Но так как при непрерывной деформации отображения степень его не меняется, то отображения f и f_1 сферы S_1 на S_2 имеют одинаковую степень. Степень же отображения f_1 очень легко подсчитать.

Разобьем плоскость прямыми, исходящими из точки 0 на n углов, каждый из которых равен $\frac{2\pi}{n}$ (рис. 166). Так как при возведении комплексного числа z в степень n его аргумент увеличивается в n раз, то каждый из этих углов с помощью отображения f_1 отображается на всю сферу S_2, причем внутренность угла отображается гомеоморфно. Таким образом, при отображении f_1 образ сферы S_1 покрывает n раз сферу S_2, причем отображение f_1 везде положительно. Отсюда следует, что степень отображения f_1 (а значит, и f) равна n.

В заключение отметим, что, не желая чрезмерно увеличивать объем статьи, мы были вынуждены оставить в стороне весьма многие важные и вместе с тем совершенно наглядные разделы топологии. Так, понятие фундаментальной группы весьма тесно связано с так называемыми *накрывающими пространствами*. Определенную связь

[1]) «Математическое просвещение», вып. 2, стр. 31 — 32.

с понятием фундаментальной группы имеет и упомянутая в нашем очерке общая *теория гомотопий*. Занимающее большое место в статье учение о поверхностях является частью общей теории многообразий, наиболее интересной (но увы! — и наиболее трудной) при числе измерений $n > 2$ (например, $n = 3$). Намеченная в главе 8 теория гомологий составляет ровно половину гомологического аппарата современной топологии, вторую половину дают «гомологии Колмогорова — Александера», или «когомологии», в основу которых положена своеобразная «граница наоборот» или «кограница», такая, что размерность границы цепи на единицу выше размерности самой цепи (а не на единицу ниже, как в случае обычной границы). Возможность строить гомологический аппарат двумя различными путями тесно связана с «*теоремами двойственности*», устанавливающими связь между топологическими образами разной размерности (в известном смысле аналогичную двойственности между нульмерными точками и одномерными прямыми классической проективной геометрии на плоскости); этот круг вопросов, идущий еще от Пуанкаре, получил глубокое развитие в работах советских топологов (Понтрягин, Колмогоров, Александров). Теоремы двойственности можно выводить, как показал Л. С. Понтрягин, из *теории зацеплений*, устанавливающей различие, скажем, между двумя незацепленными и двумя зацепленными замкнутыми кривыми трехмерного пространства. Наконец, в этой статье очень мало сказано про замечательный расцвет топологии в последние, уже послевоенные годы; расцвет этот обязан в первую очередь созданной трудами молодой французской школы топологов теории расслоенных пространств. Авторы надеются остановиться на всех этих вопросах при предполагаемой публикации их статьи отдельной книгой.

Добавление

Ниже излагаются некоторые важные понятия теории групп, используемые в главе 8 при изложении теории гомологий.

Будем рассматривать здесь только абелевы группы. Групповую операцию условимся обозначать как *сложение*, а не как умножение, т. е. результат применения групповой операции к элементам a и b будем обозначать через $a+b$. В соответствии с этим обратный к a элемент будет обозначаться через $-a$, а «единица» группы будет называться *нулем* и обозначаться символом 0, так что (согласно аксиомам группы) $a+0=a$ и $a+(-a)=0$ для любого элемента a группы. Условимся еще обозначать через Z свободную циклическую группу, а через Z_n — циклическую группу порядка n.

Пусть G — некоторая группа. Подмножество G' группы G называется *подгруппой*, если сумма любых двух элементов множества G' снова является элементом множества G' и если относительно именно этой операции сложения множество G' представляет собой группу. *Для того чтобы подмножество G' группы G являлось подгруппой, необходимо и достаточно, чтобы вместе с каждыми двумя элементами a, b множеству G' принадлежала и их сумма $a+b$, и вместе с каждым элементом a множеству G' принадлежал элемент $-a$.* Если, например, G есть группа целых чисел (включая и отрицательные) с обычной операцией сложения, то множество всех четных чисел (а также множество всех чисел, кратных данному числу m) является его подгруппой.

Пусть теперь G' — некоторая подгруппа (абелевой) группы G. Разобьем все элементы группы G на классы, считая два элемента a, b «эквивалентными», если разность $a-b$ является элементом подгруппы G' (т. е. если $a=b+g'$, где $g' \in G'$) и объединяя все эквивалентные между собой элементы в один класс. (Заметим, что два элемента, эквивалентные третьему, эквивалентны между собой.) В результате вся группа G разбивается на классы, не имеющие друг с другом общих элементов; эти классы называются *смежными классами* группы G по ее подгруппе G'. Сама подгруппа G' представляет собой один из смежных классов. Каждый элемент a, принадлежащий смежному классу A, называется *представителем* этого смежного класса. Пусть A и B — два смежных класса группы G по ее подгруппе G'; легко показать, что какие бы представители a и b этих классов ни выбирались, сумма $a+b$ всегда будет принадлежать одному определенному смежному классу C. Этот класс C называется *суммой* смежных классов A и B, так что мы будем писать $C=A+B$. Итак, в множестве всех смежных классов группы G по подгруппе G' определена операция сложения. Без особого труда проверяется, что эта операция сложения удовлетворяет всем аксиомам группы, так что *множество всех смежных классов (абелевой) группы G по ее подгруппе G' является группой (относительно введенной выше операции сложения смежных классов).* Эта группа называется *фактор-группой* группы G по ее подгруппе G'; она обозначается через G/G'. Заметим, что нулевым элементом фактор-группы G/G' является подгруппа G' (которая, как мы выше отмечали, является одним из смежных классов).

Рассмотрим в качестве примера случай, когда G есть группа всех целых чисел с обычной операцией сложения, а G' — подгруппа, состоящая из всех чисел, делящихся на m. Два элемента группы G, т. е. два целых числа a, b эквивалентны, если разность $a-b$ принадлежит подгруппе G', т. е. делится на m. Иначе говоря, эквивалентность чисел a и b означает, что они имеют одинаковые остатки от деления на m, и потому каждый смежный класс состоит из всех чисел, дающих при делении на m один и тот же остаток. Так как остатками при делении на m могут служить лишь числа $0, 1, 2, \ldots, m-1$, то существует ровно m смежных классов группы G по подгруппе G', которые мы обозначим через $0, A_1, \ldots, A_{m-1}$. Очевидно, что $2 \cdot A_1 = A_2$, $3 \cdot A_1 = A_3$, $\ldots$, $(m-1)A_1 = A_{m-1}$, $mA_1 = 0$. Таким образом, фактор-группа G/G' в данном случае является циклической группой Z_m порядка m. Это — так называемая *группа вычетов* по модулю m.

Пусть G — некоторая (абелева) группа, а $G_1, G_2, \ldots, G_k$ — ее подгруппы. Говорят, что группа G *распадается в прямую сумму* своих подгрупп $G_1, G_2, \ldots, G_k$, если каждый элемент группы G однозначно представляется в виде суммы $g_1+g_2+\ldots+g_k$, где g_i есть элемент подгруппы G_i, $(i = 1, 2, \ldots, k$ (некоторые из элементов g_i могут оказаться равными нулю). В этом случае пишут: $G = G_1 + G_2 + \ldots + G_k$. В качестве примера рассмотрим группу G, элементами которой являются всевозможные *пары* (m_1, m_2) целых чисел, а сложение определяется соотношением $(m_1, m_2)+(n_1, n_2) = (m_1+n_1, m_2+n_2)$. (Можно представлять себе группу G как множество всех векторов на плоскости, имеющих в заданной координатной системе целочисленные координаты, с обычной операцией сложения векторов.) Обозначим через G_1 множество всех пар, у которых равно нулю второе число [т. е. пар $(m_1, 0)$], а через G_2 — множество всех пар вида $(0, m_2)$ — пар, у которых равно нулю первое число. Тогда G_1 и G_2 являются подгруппами группы G, причем обе эти подгруппы являются свободными циклическими группами. Легко видеть, что каждый элемент (m_1, m_2) группы G однозначно представляется в виде суммы g_1+g_2 элементов, принадлежащих подгруппам G_1 и G_2:

$$(m_1, m_2) = (m_1, 0) + (0, m_2).$$

Таким образом, G есть прямая сумма своих подгрупп G_1 и G_2, т. е. $G=G_1+$ Обозначая каждую из подгрупп G_1, G_2 символом Z (ибо каждая из них лиется свободной циклической), мы можем написать: $G=Z+Z$. Каждую груп вида $Z+Z$, т. е. группу, распадающуюся в прямую сумму двух своих бодных циклических подгрупп, называют свободной *абелевой группой с дву образующими*. (Легко видеть, что элементы $e_1=(1, 0)$ и $e_2=(0, 1)$ являю образующими рассмотренной выше группы G, причем никаких соотношен между ними, кроме соотношения коммутативности $e_1+e_2=e_2+e_1$, нет.)

Аналогично, множество G, элементами которого являются всевозможные системы $(m_1, m_2, \ldots, m_k)$, состоящие из k целых чисел, а сложение — «по координатное»

$$(m_1, m_2, \ldots, m_k)+(n_1, n_2, \ldots, n_k)=(m_1+n_1, m_2+n_2, \ldots, m_k+n_k),$$

является абелевой группой. (Группа G может трактоваться как группа векторов k-мерного пространства, имеющих целые координаты.) Обозначим через G_1 ее подгруппу, состоящую из всех элементов вида $(m_1, 0, \ldots, 0)$, т. е. элементов, у которых на втором, третьем, ..., k-м месте стоят нули. Аналогично определяются подгруппы $G_2, \ldots, G_k$. Тогда каждая из подгрупп $G_1, G_2, \ldots, G_k$ является свободной циклической, и группа G распадается в прямую сумму своих подгрупп $G_1, G_2, \ldots, G_k$, так что мы можем написать

$$G=Z+Z+\ldots+Z \qquad (k \text{ слагаемых}).$$

Группа G называется свободной абелевой группой с k образующими.

В теории групп доказывается, что каждая абелева группа G с конечным числом образующих (т. е. группа, обладающая системой образующих, содержащей лишь конечное число элементов) представляется в виде

$$G=Z+Z+\ldots+Z+K,$$

где K — конечная группа (т. е. группа, состоящая из конечного числа элементов). Число «слагаемых» Z в этом прямом разложении называется *рангом* группы G.

Наконец, напомним еще понятие *гомоморфизма* и, в частности, *изоморфизма*. Пусть G_1 и G_2 — две группы. Всякое отображение группы G_1 в G_2, сохраняющее групповую операцию, называется гомоморфизмом группы G_1 в группу G_2. Иначе говоря, если мы условимся обозначать через $f(g_1)$ элемент группы G_2, являющийся образом элемента g_1 группы G_1, то требование «сохранения» групповой операции (т. е. определение гомоморфизма) запишется в виде

$$f(g_1+g_1')=f(g_1)+f(g_1')$$

для любых элементов g_1, g_1' группы G_1. Если гомоморфное отображение f является взаимно однозначным, то оно называется изоморфизмом. Например, все свободные циклические группы (или все циклические группы одного и того же порядка) изоморфны между собой. Две *изоморфные* группы, т. е. такие группы, что одну из них можно изоморфно отобразить на другую, ничем *алгебраически* друг от друга не отличаются, т. е. все утверждения о групповой операции в одной группе справедливы и для второй группы и наоборот. (Подобно этому, две *равные*, т. е. «совпадающие при наложении», фигуры мы считаем геометрически ничем друг от друга не отличающимися.) Например, утверждение о том, что группу всех пар (m_1, m_2) целых чисел «можно представлять себе» как множество всех векторов с целыми координатами на плоскости, в точной формулировке означает, что группа всех пар (m_1, m_2) *изоморфна* группе векторов с целыми координатами на плоскости.

МАШИНА ИГРАЕТ В ШАХМАТЫ

Ю. М. Безбородов и В. Б. Орлов

(Москва)

Настоящая статья ставит своей целью рассмотреть вопрос о применении электронных цифровых машин для игры в шахматы. При этом у читателя не предполагается сведений по этим вопросам, выходящих за рамки основных идей статей А. А. Ляпунова и Г. А. Шестопал, знакомых читателям «Математического просвещения» [1]). Что касается шахматной подготовки, то требуется лишь знать правила игры и записи ходов.

1. НЕМНОГО ИСТОРИИ

Литературная тема об автомате, играющем в шахматы и непобедимом для игроков-людей, разрабатывалась различными писателями задолго до появления действительных устройств, приспособленных к ведению игры. Вероятно, большую роль в популярности этого сюжета сыграла нашумевшая мистификация знаменитого венгерского механика XVIII века Фаркаша Кемпелена. «Шахматный автомат» Кемпелена представлял собой куклу в костюме турка [рис. 1 [2])], соединенную со шкафчиком, на котором располагались шахматы. Открывая дверцы шкафчика, любопытные могли видеть сложное переплетение шестеренок и рычажков. Через каждые 12 ходов изобретатель заводил механизм с помощью огромного ключа (это давало лишнее время для обдумывания ходов замаскированному механизмом человеку). Для управления «автоматом» привлекались крупные шахматисты (Альгайер, Льюис и другие), так что уровень игры был весьма высоким. Выдумка Кемпелена произвела настолько большой эффект

[1]) А. А. Ляпунов и Г. А. Шестопал, Начальные сведения о решении задач на электронных вычислительных машинах, «Математическое просвещение», вып. 1, 1957, стр. 57—74; А. А. Ляпунов и Г. А. Шестопал, Об алгоритмическом описании процессов управления, там же, вып. 2, 1957, стр. 81—95. См. также Г. Ф. Бонеблуст, Теория игр, там же, вып. 4, 1959, стр. 53—86.

[2]) Фотография взята из книги: Е. Гижицкий, С шахматами через века и страны, Warszawa, 1958.

в Европе, что Екатерина II собиралась даже купить машину. Появлявшиеся в дальнейшем аналогичные подделки уже не имели такого успеха.

Первым настоящим шахматным автоматом следует считать механизм, построенный испанцем Л. Торресом-и-Кеведо около полувека тому назад. Устройство было приспособлено для разыгрывания только одного простейшего вида эндшпиля — матования одинокого короля ладьей и королем. Зато эта задача выполнялась безошибочно. Предпосылкой создания такой машины является существование для данной задачи сравнительно простого алгоритма (четкой и исчерпывающей системы правил), ведущего к цели при любом начальном расположении фигур. Реализующий именно этот специфический алгоритм, автомат Кеведо не может быть использован для решения других шахматных задач.

Рис. 1. Обнаруженный в одном из парижских антикварных магазинов старый шахматный автомат после ремонта демонстрируется как произведение Кемпелена.

Переворот в решении общей проблемы автоматизации различных видов умственной деятельности (и, в частности, игры в шахматы) связан с появлением-универсальных программно-управляемых электронных цифровых машин. Задача составления программы для игры в шахматы привлекла внимание теоретиков машинной математики уже в первые годы развития нового могучего орудия человеческого разума. В 1950 г. один из создателей современной кибернетики Клод Шеннон опубликовал посвященные этому вопросу статьи [1]), которые явились источником идей для дальнейших работ.

[1]) C. E. Shannon, Programming a computer for playing chess, Philosophical Magazin, s. 7, 41; C. E. Shannon, A chess playing machine, Scientific American 182, 1950, № 2.

Честь первого практического осуществления машинной игры в шахматы принадлежит, по-видимому, известному английскому ученому-логику Алану Тьюрингу, который использовал вычислительную машину «МАДАМ» Манчестерского университета. В Советском Союзе программа для решения двух-трехходовых шахматных задач была составлена В. М. Курочкиным и с успехом испытана на машине БЭСМ. В более поздние годы ряд работ был проведен в США.

Следует отметить, что в практической игре машины пока еще очень слабы. Уровень игры кемпеленовского «турка» был несравненно выше, и шахматные мастера пока могут не бояться соперничества автоматов. В то же время в области решения шахматных задач (мат в 2—3 хода) результаты, достигаемые машинами, вполне удовлетворительны уже сейчас.

Однако интерес рассматриваемой нами темы не сводится к чисто шахматным устремлениям. Разберем этот вопрос подробнее.

2. В ЧЕМ ЗАКЛЮЧАЕТСЯ ИНТЕРЕС ЗАДАЧИ

Шахматы издавна рассматривались не только как игра, но и как своеобразный прототип умственной деятельности вообще. Разумеется, параллель между шахматистом и, скажем, переводчиком нельзя считать особенно близкой, хотя бы потому, что они имеют дело с различным конкретным материалом и различными по характеру целями. Но когда речь заходит об автоматизации решения выполняемых обоими задач, то сходство становится более заметным.

Шеннон приводит целый список задач, автоматизация решения которых в той или иной мере аналогична автоматизации шахматной игры. Среди них упоминаются: перевод с одного языка на другой, конструирование релейно-контактных схем [1]), оркестровка мелодий, военно-стратегические задачи и многие другие.

Какое же сходство обнаруживается при попытках решать упомянутые разнообразные проблемы с помощью машин? Можно указать на три основных обстоятельства.

Во-первых, все эти задачи (в отличие от обычных, вычислительных) носят «неарифметический», нечисловой характер. Поэтому относительно большую роль играет разумная кодировка исходных данных с помощью чисел. От этого может сильно зависеть эффективность решения.

Во-вторых, самый процесс решения имеет сходный характер. Он напоминает нечто вроде поиска, производимого путем «проб и ошибок».

[1]) Так называются схемы, составленные из особых элементов (реле) и используемые в устройствах управления (а также для выполнения математических операций над числами); электромагнитное реле под действием поля катушки с током замыкает или размыкает свои контакты.

В-третьих, во всех упоминаемых случаях нет строгого деления возможных решений только на две категории — правильных и неправильных. Существует постепенная градация от самых лучших до самых плохих решений. И задача, собственно, заключается не в отыскании самого лучшего решения (это слишком сложно), а в нахождении достаточно хорошего и одновременно достаточно простого.

Итак, мы имеем дело с целой группой более или менее сходных проблем, причем некоторые из них имеют существенное практическое значение. И шахматы могут послужить, по образному выражению Шеннона, «острием клина», пробивающим новые пути исследования.

Имеются серьезные основания для того, чтобы выбрать именно шахматы в качестве экспериментальной модели для обнаружения закономерностей, могущих оказаться полезными при решении более важных задач.

Прежде всего, здесь достаточно четко определена постановка задачи (правила игры).

Затем степень сложности оказывается как раз удачной. Задача не тривиальна, но и не чрезмерно трудна.

Наконец, дискретный (скачкообразный) характер шахматной игры удобен для реализации в цифровых машинах.

Возможно, некоторые читатели будут разочарованы тем, что имеется в виду отыскивать лишь «достаточно хорошее» решение вопроса, а задача о наилучшей, идеально безошибочной игре даже не ставится. Многим представляется, что это было бы не такой уж недостижимой задачей. К сожалению (или, вернее, к счастью), это не так — такая задача еще очень далека от решения[1]). Вопрос о машинных возможностях для безошибочной игры будет рассмотрен в настоящей статье после уточнения некоторых основных понятий.

3. ШАХМАТЫ И ТЕОРИЯ ИГР

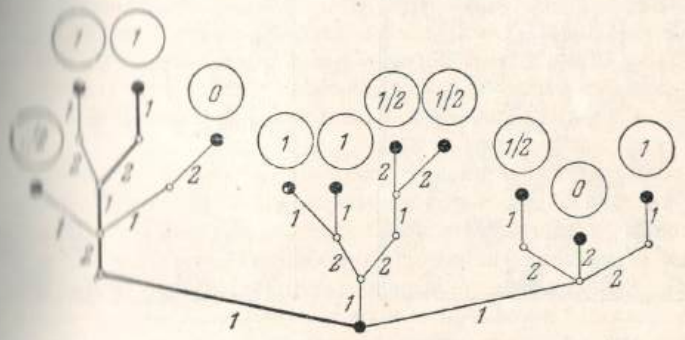
Шахматы принадлежат к так называемым *играм в развернутой форме*[2]). Эти игры характеризуются тем, что каждая конкретная партия представляет собой, вообще говоря, последовательность многих ходов, а не обязательно ограничивается однократным выбором какой-либо возможности каждым из партнеров.

Для игр в развернутой форме часто пользуются графическим представлением в виде *дерева*, аналогичного изображенному на рис. 2.

[1]) В 20-х гг. нашего века сложилась легенда о близости шахмат к исчерпанию, к «ничейной смерти», навеянная «непобедимостью» тогдашнего чемпиона мира Капабланки. Но мифу о ничейной смерти решительный удар был нанесен Алехиным, отыскавшим слабые места в капабланковской стратегии и открывшим новые возможности в шахматах.

[2]) Употребляются также термины: *позиционные игры, обобщенные игры, игры в расширенной форме.*

[illegible] соответствуют различным возможным ходам, концы отрезков — [illegible] положениям игры. При этом самая нижняя точка изображает [illegible] положение, а концевые точки — положения, в которых [illegible] (поставленные около них числа в кружочках соответствуют исходам игры: *1* — первый игрок выиграл, *1/2* — ничья, *0* — первый игрок проиграл). Цифры, стоящие около отрезков, означают номер игрока, которому принадлежит выбор хода. Каждая [illegible] партия графически представляется неразветвленной цепью [illegible] отрезков, соединяющей начальное положение с одним из [illegible] (например, жирная линия на рис. 2).


1
1
0
1/2
1/2
1
1
1/2
1
0
1
2
1
2
1
2

однажды разъединившись, уже не соединяются впоследствии. Это свойство (принадлежащее дереву по определению) означает, что положения игры считаются различными, если они получены из начального разными путями. Иными словами, в совокупность характеризующих позицию сведений должна быть включена информация о ее «предыстории».

Игры, в которых игрок располагает сведениями о всех предыдущих ходах, называются *играми с полной информацией* [1]). В обычной шахматной партии это условие соблюдено. Если же речь идет об

Рис. 3. Белые выигрывают при своем ходе, если их ладья и король еще не ходили.

Рис. 4. Ход белых. Они выигрывают, если предыдущий ход черных был b7—b5, и проигрывают в противном случае.

анализе какой-либо позиции, то приходится требовать, чтобы эта позиция была задана не только расположением фигур, но и некоторыми сведениями об ее предыстории. Иначе мы не всегда сумеем даже определить совокупность возможных в позиции ходов (примеры на рис. 3 и 4 показывают, что это может быть существенным), не говоря уже о применении ничейных правил. Легко видеть, что достаточно задать дополнительно: а) последовательность ходов, сделанных начиная с последнего движения пешки или взятия и б) сведения о возможности рокировок. Тогда совокупности ходов, возможных в данной и последующих позициях, будут точно определены, равно, как и последствия этих ходов.

В силу приведенных соображений мы должны различать в дальнейшем понятия *позиции* и *расположения фигур*. Легко видеть, что позиция в описанном смысле не может повторяться в одной и той же партии, тогда как расположение фигур может повторяться.

[1]) К *играм с неполной информацией* относится, например, домино, так как здесь игроки не располагают всеми сведениями о предыдущих ходах (к числу которых естественно отнести и выбор костей партнерами).

Таким образом, шахматы (с указанными оговорками) можно считать принадлежащими к классу конечных игр с полной информацией. Относительно таких игр доказана важная общая теорема; для ее формулировки введем некоторые новые понятия.

Будем говорить, что задана *чистая стратегия белых*, или просто *стратегия белых*, если каждой возможной позиции с ходом белых поставлен в соответствие о д и н из допустимых в этой позиции ходов. Аналогично определяются *стратегии черных*. Множество всех стратегий, очевидно, конечно.

Если заданы одновременно стратегия белых, стратегия черных и позиция, то продолжение игры (а значит, и ее исход) определяются однозначно. Обратно, всякая партия может быть представлена как результат выбора белыми и черными некоторых стратегий и применения их, начиная с исходной позиции (нужно учесть сделанное выше замечание о невозможности повторения позиции в одной и той же партии).

Это показывает, что шахматы теоретически могут быть сведены к игре, которая заключается в однократном выборе стратегии каждым из игроков. Аналогичное построение может быть легко выполнено для любой игры в развернутой форме (при соответствующем определении стратегий) и называется *приведением к нормальной форме*.

Если взять произвольную игру в нормальной форме, то известно [1]), что для игроков не всегда существуют оптимальные чистые стратегии. Это значит, что наилучший для первого игрока исход партии, который он может в ы н у д и т ь, не всегда совпадает с тем исходом, который может в ы н у д и т ь второй игрок, стремясь к наилучшему для себя результату. Поясним это примером игры, определенной следующей таблицей.

	1-я стратегия черных	2-я стратегия черных
1-я стратегия белых	выигрыш белых	ничья
2-я стратегия белых	ничья	выигрыш белых

Здесь белым г а р а н т и р о в а н а только ничья, так как после любого их выбора черные могут избежать поражения. В то же время у черных нет фиксированной стратегии, которая давала бы им ничью н е з а в и с и м о от способа игры белых.

Если оптимальные исходы, которые могут вынудить оба игрока, совпадают, то соответствующий результат партии называется *ценой игры*.

[1]) См. цитированную выше статью Г. Ф. Б о н е н б л у с т а.

Теорема, о которой мы говорили, может быть сформулирована так: *для нормальной формы конечной игры с полной информацией всегда существует цена игры* [1]). Отсюда вытекает для шахмат такое утверждение: *если в какой-либо позиции ни белые, ни черные не могут выиграть форсированно* (независимо от выбора стратегии противником), *то каждая сторона может вынудить ничью.*

Рекомендуем читателям самостоятельно доказать это. Нужно воспользоваться индукцией по числу ходов самого длинного продолжения игры, возможного в позиции.

4. ВОЗМОЖНА ЛИ БЕЗОШИБОЧНАЯ ИГРА НА МАШИНЕ?

В предыдущем изложении мы интересовались только теоретическими возможностями позиции, совершенно снимая вопрос о трудности их достижения. Благодаря этому оценка позиции получилась только трех видов (выигрышная, проигрышная и ничейная), соответственно принципам «выигрышную позицию улучшить невозможно» и «в проигрышной позиции ошибок не бывает». Рассмотрим теперь, какова сложность осуществления безошибочной игры в шахматы на машине.

Существуют игры, для которых вполне осуществим безошибочный расчет с помощью современных электронных устройств. Разберем алгоритмы, которые при этом могут быть использованы.

Наиболее примитивный способ связан с тем, что в запоминающем устройстве машины находится «словарь позиций», т. е. список всех возможных позиций с указанием для каждой из них хода, который надлежит сделать. Код позиции, подлежащей рассмотрению, последовательно сравнивается с кодами записанных в «словаре» позиций. При равенстве управление передается командам, выдающим ход, соответствующий данной позиции. Ясно, что по этому принципу могут работать устройства, реализующие лишь простейшие позиционные игры с полной информацией.

Неосуществимость метода «словаря позиций» для шахмат вытекает из слишком большого количества возможных шахматных позиций. В самом деле, число различных расположений 32 шахматных фигур на 64-клеточной доске равно $\frac{C_{64}^{32}}{(P_8)^2 (P_2)^6}$. Можно смело утверждать, что это число меньше числа всех возможных позиций, так как можно варьировать все остальные данные о позиции (очередь хода и т. д.), и это компенсирует с лихвой то обстоятельство, что некоторые расположения фигур противоречат правилам. К тому же мы не считали позиций, где часть фигур разменена. Порядок полученного числа

[1]) См. Д. Блекуэлл и М. А. Гиршик, Теория игр и статистических решений, ИЛ, 1958; или Дж. Мак-Кинси, Введение в теорию игр, М., Физматгиз, 1960.

примерно равен 10^{43}. Если бы даже удалось осуществить запись одной позиции в элементе объемом в кубический микрон, то и тогда запоминающее устройство машины, работающей по указанному методу, должно было бы значительно превосходить по размерам весь земной шар.

Усовершенствование этого способа может заключаться в том, чтобы разбить всё множество позиций на сравнительно небольшое количество типов и для каждого типа позиции указать вид рекомендуемого хода. Тогда работа машины могла бы, например, протекать по следующей схеме. Путем проверки ряда условий устанавливается тип позиции. Все возможные ходы в позиции последовательно проверяются, и выдается первый из ходов, удовлетворяющих условиям, рекомендованным для данного типа позиций.

Этот способ «типовых позиций и типовых ходов» вполне применим в некоторых эндшпилях. Например, в эндшпиле «король и пешка против короля» вполне возможно указать признаки позиций, в которых следует двигать пешку, занимать королем оппозицию и т. д. Однако никому еще не удалось построить исчерпывающую систему правил, позволяющую распространить такое положение на всю совокупность шахматных позиций. Существует множество позиционных правил типа «если есть открытая линия, то нужно занимать ее ладьей», но ни одно из этих правил не носит абсолютного характера. Всегда можно найти примеры, когда данному правилу не нужно следовать.

Безошибочную игру на машине можно осуществить и в тех играх, где существует точная формула для определения наилучшего хода в каждой возможной позиции. Такая формула известна, например, для игры в «Ним», и созданы специальные устройства и программы для цифровых программно-управляемых машин, реализующие эту формулу.

Если бы существовала такая формула и для шахмат (вроде якобы открытой героем фантастического рассказа А. Абрамова «Смерть шахмат»), то машина получила бы реальное преимущество в игре против человека ввиду своего быстродействия. Однако (к счастью для любителей шахмат) такой формулы не найдено.

Все указанные способы страдают тем недостатком, что они требуют предварительной аналитической работы по отысканию стратегии, которую нужно использовать. Этого недостатка лишен четвертый метод — метод «полного перебора вариантов». Здесь исследуются поочередно все возможные в данной позиции продолжения и производится их сравнительная оценка на основании исходов вариантов.

Процесс осуществляется с помощью характерной для теории игр процедуры попеременной максимизации и минимизации. На каждом этапе сравниваются два варианта, различающиеся лишь последним ходом. Если это — ход противника, то выбирается вариант, приводящий к худшему исходу. Если же последний ход свой, то выбирается вариант, приводящий к лучшему исходу. При одинаковом исходе

выбор безразличен и производится по какому-нибудь постороннему легко проверяемому признаку. По рассмотрении всей группы вариантов, отличающихся лишь последним ходом, последней общей для них позиции приписывается исход оставшегося («экстремального») варианта. Так постепенно уменьшается длина рассматриваемых вариантов. Этот

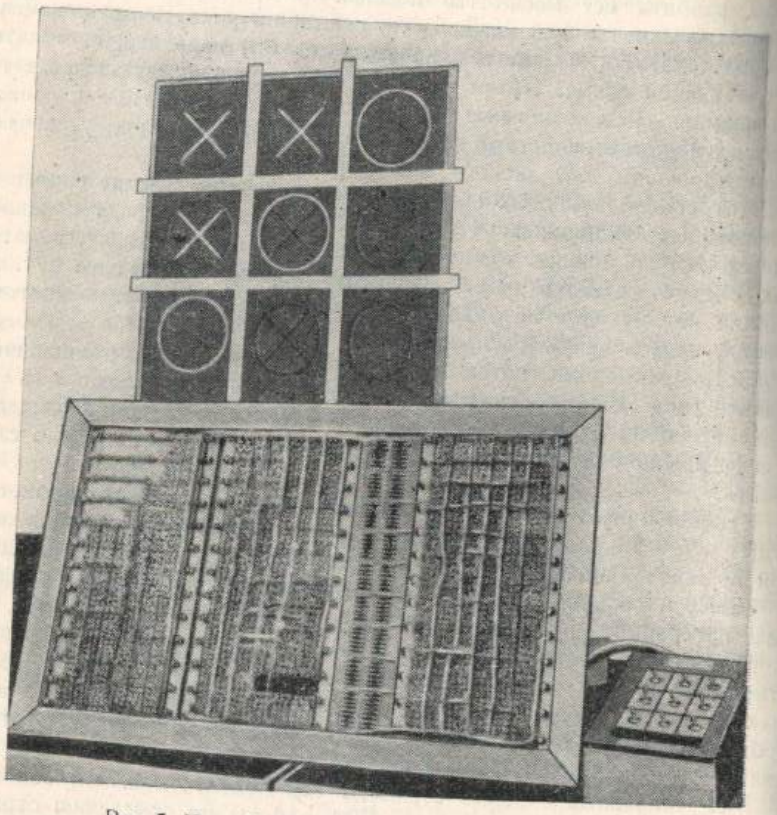


Рис. 5. Устройство для игры в «крестики-нулики».

процесс сводится в конце концов к выбору между вариантами длиной в один ход, т. е. к нахождению наилучшего хода.

Такой метод вполне приемлем с точки зрения требуемого объема памяти машины, так как варианты строятся и рассматриваются поочередно. Однако время исследования всей сети вариантов зависит от количества возможных вариантов. Попробуем оценить его.

Для игры в «крестики-нулики» на 9 клетках количество различных продолжений равнялось бы $2^9 \cdot 9! = 22\,498\,560$, если бы большая

часть вариантов не кончалась раньше девятого хода. Фактически вариантов во много раз меньше. Показанное на рис. 5 устройство (сконструированное Уоллесом Чейсом) анализирует более 633 000 продолжений, что оказывается достаточным для безошибочной игры.

В случае шахмат положение значительно сложнее. Количество ходов, возможных в начальной позиции, равно 20, и столько же возможно ответов. Это дает 400 вариантов уже на первом ходу. Ясно, что по мере развития фигур количество возможных ходов должно возрастать. Затем оно должно убывать в связи с разменами и уменьшением количества фигур. Подсчеты, произведенные де Гроотом [1])

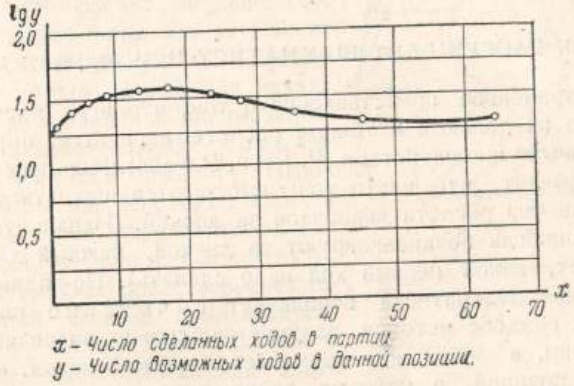


Рис. 6. График логарифма числа возможных в позиции ходов, полученных де Гроотом. По горизонтали отложено количество ходов, сделанных с начала партии.

на основании большого числа партий мастеров, дают в среднем картину, изображенную на графике (рис. 6) и подтверждающую сделанное предположение. Из графика видно, что на промежутке до 30 хода ($0 < x < 30$) можно считать среднее число возможных ходов в позиции близким к $y = \sqrt{1000}$ ($\lg y \approx 1{,}5$). Это значит, что увеличение длины расчета на 1 ход (белых и черных) вызывает увеличение количества вариантов в 1000 раз. Общее же количество вариантов до 30 хода получается порядка $1000^{30} = 10^{90}$. Если сравнить эту цифру с порядком скорости самых лучших современных машин (10^6 элементарных операций в секунду, 10^{13} — в год), то становится ясной полная безнадежность такого метода исследования позиции.

Конечно, нельзя утверждать, что не существует какой-либо другой алгоритм, также дающий точное решение проблемы шахматной игры и в то же время существенно менее трудоемкий. Однако пока

[1]) De Groot, Het Denken vom den Schaker, Amsterdam, 1946.

положение аналогично вопросу о распределении простых чисел, где имеется универсальное и вполне точное, но слишком трудоемкое «решето Эратосфена» (аналогичное в этом отношении алгоритму «перебора вариантов»), но не существует практичной формулы, дающей все простые числа.

Как мы видим, задача о совершенной машинной игре не может быть успешно решена даже с помощью «чудесных способностей» быстродействующих цифровых машин. Поэтому разумнее заняться другой (гораздо более интересной) задачей о «достаточно хорошей» игре. С этой целью надо прежде всего присмотреться к тому, как играет шахматист-человек.

5. КАК РАССУЖДАЕТ ШАХМАТИСТ ПРИ ВЫБОРЕ ХОДА

Распространенное представление о том, что игра мастеров основана только на далеком и точном расчете вариантов, опровергалось многими видными шахматистами (Р. Рети, Р. Файн). Ряд простых наблюдений показывает, что часто ходы выбираются шахматистами любой силы вообще без расчета вариантов за доской. (Начальную позицию, например, никогда не анализируют за доской, каждый для себя заранее решает, какой первый ход надо сделать). По-видимому, ясно, что мышление шахматиста использует *сочетание* описанных в предыдущем разделе методов, разумеется, сильно видоизмененных.

Рассмотрим, в каком виде применяется первый метод. «Словарик» конкретных позиций, в которых точно известно, как надо играть, хранится, пополняется и исправляется в памяти любого шахматиста. Но размеры его по необходимости невелики, и он почти исключительно ограничен дебютными позициями. Ведь только для позиций, близких к начальной, есть заметная вероятность точного повторения в нескольких партиях.

Гораздо большую роль в мышлении шахматиста играет обобщение того метода, при котором частные позиции заменяются типами, группами позиций, объединенных каким-нибудь характеризующим их признаком. Рассуждение происходит по схеме: «раз позиция обладает таким-то признаком, то надо делать такой-то ход». Например: «если есть открытая линия, то надо ставить на нее ладью». Разумеется, при таком обобщении метод теряет в надежности (не учитываются индивидуальные особенности позиции), но выигрывает в легкости использования. Впрочем, как мы уже отмечали, в теории эндшпилей существует ряд правил, каждое из которых вполне точно решает целую группу позиций.

Однако наибольшую роль всё же играет метод расчета, являющийся упрощением алгоритма перебора всех вариантов. При этом полностью сохраняется процедура попеременной минимизации и максимизации, однако неизмеримо сокращается рассматриваемая сеть

вариантов. Это происходит как за счет сокращения глубины расчета, так и за счет отбрасывания многих вариантов. В цитированной выше книге де Гроота приведены результаты экспериментального исследования, проведенного с мастерами. Результат одного из опытов изображен на рис. 7. Как видно, глубина расчета сравнительно невелика и притом крайне неравномерна для разных вариантов. Для семи начальных ходов рассматривалось только 16 вариантов и 47 позиций.

При этом возникают своеобразные и нелегкие проблемы. Одна из них связана с тем, что расчет не доводится до конца и, стало быть, окончательные исходы вариантов неизвестны. Поэтому нужна какая-то иная основа для начала процессов максимизации и минимизации. В качестве таковой служит *сравнительная оценка* позиций. Игрок должен хотя бы для некоторых позиций уметь без расчета определить, какая из них выгодней. Если это достигнуто, то нужно стараться довести расчет до таких позиций. Оценка производится, по существу, на основах, сходных с методом «типовых позиций» или «аналогий», и опирается на накопленный опыт.

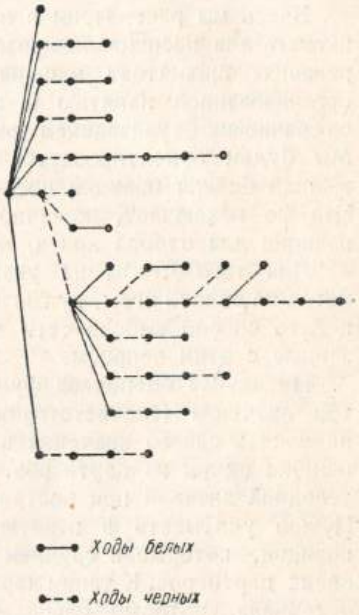


Рис. 7. Графическое изображение сети вариантов, рассмотренных мастером при анализе позиции в одном из опытов.

Другая проблема — отбор ходов для рассмотрения. По существу, это делается опять-таки на основе типовых признаков позиции, или хода, или того и другого. Примеры рассуждений: «надо посмотреть форсирующие ходы», «позиция близка к начальной, значит, надо рассмотреть развивающие ходы» и т. д.

Значение правильной оценки позиции для силы игры общеизвестно. Но не менее важен правильный выбор сети вариантов для рассмотрения. Многие шахматисты грешат тем, что отбирают для рассмотрения очень мало ходов противника, зачастую оставляя вне поля зрения самые сильные.

Следует заметить, что отбор сети вариантов для рассмотрения и расчет их — переплетающиеся процессы. Обычно шахматист сознательно или бессознательно сравнивает результаты расчета с имевшейся у него заранее интуитивной оценкой обдумываемой позиции. И если результаты его не удовлетворяют, он продолжает поиски, расширяя рассматриваемую сеть вариантов.

Большое влияние на всю схему расчета (выбор сети вариантов и оценку возникающих позиций) оказывает план, которым руководствуется игрок. В свою очередь расчет часто вносит коррективы в план, показывая степень его реальности.

6. ОСУЩЕСТВЛЕНИЕ ПРИБЛИЖЕННОЙ ИГРЫ НА МАШИНЕ

Здесь мы рассмотрим в общих чертах вопрос о программировании шахмат для неспециализированной электронной машины. Специализированная шахматная машина отличалась бы большой специально организованной памятью и развитой системой операций, связанных со сравнением, «узнаванием», выборкой кодов по определенным признакам. Мы будем ориентироваться на метод расчета вариантов и приближенной оценки позиций посредством вычисления некоторой функции. Это не исключает, конечно, использования анализа особенностей позиции для отбора ходов, подлежащих рассмотрению.

Для того чтобы на указанной основе построить определенный метод машинной игры, нужно вначале выбрать *оценочную функцию* $f(P)$ и дать способ выбора сети вариантов для расчета. Рассмотрим связанные с этим вопросы.

Идеальная оценочная функция должна была бы принимать только три значения (соответствующие выигрышу, проигрышу и ничьей) и не менять своего значения в течение партии, пока не будет сделана ошибка одним из партнеров. Это последнее свойство и является путеводной звездой при построении приближенных оценочных функций. Нужно учитывать в первую очередь те характеристики шахматной позиции, которые в среднем наиболее устойчивы при хорошей игре обоих партнеров. К таким характеристикам можно отнести соотношение материала, которым располагают игроки, а также плюсы и минусы пешечного костяка позиции. Но не следует забывать, что стремиться надо к устойчивости оценочной функции в целом, а не отдельных ее членов. В практической партии преимущество часто меняет свою форму, превращаясь из одного вида в другой. Можно пожертвовать пешку, получив взамен большую подвижность фигур или перевес в силах на важном участке доски. Поэтому такие скоропреходящие факторы тоже должны учитываться в оценочной функции.

Практически легче всего учесть материал. Соотношения силы различных фигур установлены длительным опытом и достаточно устойчивы. Поэтому основной частью всякой оценочной функции служит обычно предложенная еще Шенноном сумма $200(\text{Кр} - \text{Кр}') + 9(\text{Ф} - \text{Ф}') + 5(\text{Л} - \text{Л}') + 3(\text{С} - \text{С}' + \text{К} - \text{К}') + \text{П} - \text{П}'$. Здесь Кр, Ф, Л, С, К и П означают количество фигур данного типа у белых, а те же буквы со штрихами — количество тех же фигур у черных. Коэффициенты 9, 5, 3, 1 соответствуют относительной ценности ферзя, ладьи, легкой фигуры и пешки. Первый член этой суммы

равен нулю в течение всей партии, но определяет знак всей суммы после мата одному из королей или при попытках его «пожертвовать».

Дальнейшим шагом должно быть добавление со знаком «минус» членов, соответствующих изолированным, сдвоенным и отсталым пешкам и умноженных на некоторые небольшие коэффициенты. Со знаком «плюс» можно добавить член, учитывающий наличие проходных пешек.

Делались также попытки учесть подвижность фигур (количество возможных в позиции ходов). Это тоже оказалось полезным, хотя в экспериментах Улама и Стейна [1]) приводило к «шахобоязни» (при шахе резко ограничивается количество возможных ответов). В исследованиях А. Бернштейна и других [2]) функция $f(P)$, помимо подвижности и соотношения материала, учитывала также безопасность позиции короля и контролирование критических полей.

Следовало бы также учесть много других позиционных факторов, но это требует преодоления трудностей, связанных с их адекватным числовым выражением.

Что касается выбора сети вариантов, то простейшим решением вопроса является рассмотрение всех допустимых правилами игры вариантов определенной длины (скажем, в 2 или 3 хода белых и черных). Такую стратегию Шеннон называл стратегией типа A и указывал, что она весьма несовершенна по двум причинам. Во-первых, машина проделывает при этом колоссальное количество лишней работы, рассматривая все дальнейшие разветвления явно бессмысленных продолжений. Это должно привести к медленности игры. Во-вторых, в ряде случаев анализ необходимо было бы продолжить. Это имеет место при проведении серии разменов или многоходовой комбинации с жертвами, когда значение оценочной функции резко колеблется от хода к ходу.

В упомянутых выше опытах Улама и Стейна выбор стратегии типа A с глубиной расчета в 2 хода для каждого противника привел к тому, что машина МАНИАК тратила более часа на «обдумывание» одного хода (при скорости 6—8 тысяч элементарных операций в секунду!). В результате они перешли от обычных шахмат к упрощенным (с доской размером 6×6 и без слонов). Тогда на выбор хода стало затрачиваться около 10 минут.

Для улучшения стратегии Шеннон предлагал ввести две функции $g(P)$ и $h(P, M)$. Первая из них должна принимать только два значения, а именно 0 (в стабильной позиции P) и 1 (в случае нестабиль-

[1]) J. Kister, P. Stein, S. Ulam, W. Walden, M. Wells, Experiments in chess, Journal of the Association for Computing Machinery 4 (1957), № 2, стр. 143—147; S. Ulam, P. Stein, Experiments in chess on electronic computing machines, Computers and Automation 6 (1957), № 9, стр. 14—20.

[2]) A. Bernstein, M. Roberts, Computers v. chess-player, Scientific American 198 (1958), № 6, стр. 96—105.

ности P). Нестабильность понимается как наличие шаха, угроз материального выигрыша и других сильных угроз. Эта функция используется для определения глубины расчета. А именно, расчет продолжается до тех пор, пока $g(P)$ не обращается в нуль (но должны быть также даны безусловные нижняя и верхняя границы, — скажем, 2 и 10 ходов). Функция $h(P, M)$ определяет целесообразность рассмотрения хода M в позиции P. Предлагалось дать ей наиболее высокие значения при форсирующих и развивающих ходах, средние — при защитительных и более низкие — при остальных.

Внимательное рассмотрение этого предложения показывает, что в обоих случаях речь идет не столько о вычислительных операциях, сколько о проверке совокупности некоторых условий в позиции, что, конечно, также вполне доступно для машины.

Примитивный вариант функции $g(P)$ был использован в экспериментах на машине «МАДАМ». Рассчитывая обычные варианты всего лишь на один ход вперед, машина исследовала размены до их окончания.

Остановимся теперь на экспериментах с машиной ИБМ-704 [1]) как на иллюстрации идеи об использовании функции $h(P, M)$.

В каждой позиции машина отбирала лишь 7 ходов для образования вариантов. При глубине расчета в два хода это дает в общей сложности $7 + 7^2 + 7^3 + 7^4 = 2800$ позиций, что позволяло производить «обдумывание» хода в любой позиции за 8 минут.

Выбор ходов для рассмотрения при этом основан на том, что все ходы, возможные в данной позиции, располагаются в определенной последовательности. Рассчитываются 7 первых по счету ходов. Порядок расположения в этой последовательности определяется путем проверки ряда условий в позиции и выглядит следующим образом.

1) Защита от шаха (если он имеет место):
   а) взятием,
   б) перекрытием,
   в) уходом;
2) защита от угрозы материального проигрыша;
3) рокировка;
4) развитие легкой фигуры;
5) занятие открытой линии;
6) занятие критических полей, возникших в пешечном расположении;
7) остальные ходы пешками;
8) остальные ходы фигурами.

Несомненно, что такой порядок рассмотрения имеет ряд недостатков. Это обнаруживается и в приводимой ниже партии, где грубая ошибка (зевок ферзя) была сделана в пределах заданной глубины

[1]) См. указанную выше работу А. Бернштейна и М. Робертса.

расчета. Очевидно, соответствующий ход противника оказался за пределами группы из 7 первоочередных.

Машина — человек

1. e2 — e4	e7 — e5	12. d3:c4	e5 — e4
2. Cf1 — c4	b7 — b6	13. Kf3 — g5	Фc6 — g6
3. d2 — d3	Kg8 — f6	14. Kg5 — h3	e4 — e3
4. Cc1 — g5	Cc8 — b7	15. f2 — f3	Cf8 — c5
5. Cg5:f6	Фd8:f6	16. Лf1 — e1	0 — 0
6. Kg1 — f3	c7 — c6	17. Kb1 — c3	e3 — e2 +
7. 0 — 0	d7 — d5	18. Kh3 — f2	Cb7:f3
8. e4:d5	c6:d5	19. g2 — g3	e2:d1Ф
9. Cc4 — b5 +	Kb8 — c6	20. Kc3:d1	Фg6 — c2
10. c2 — c4	d5:c4	21. b2 — b3	Лa8 — d8
11. Cb5:c6 +	Фf6:c6	22. h2 — h4	Лd8:d1

Белые сдались.

Внешний вид шахматного «сражения» между машиной и человеком показан на рис. 8 [1]). Игрок располагает шахматной доской с фигурами. Машина фиксирует течение игры в своем запоминающем устройстве.

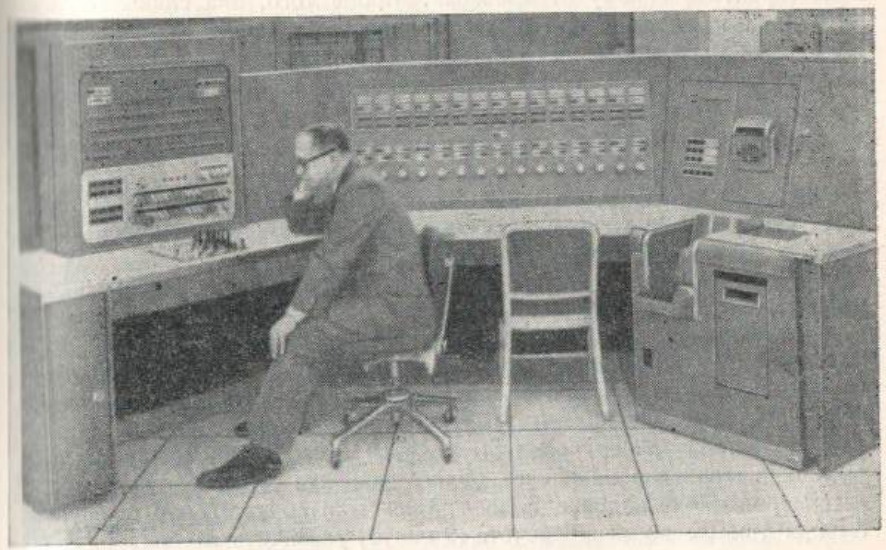


Рис. 8. Партия машины ИБМ-704 против А. Бернштейна.

Перед началом игры человеку предоставляется право выбора цвета фигур путем нажатия соответствующей кнопки. Над первым ходом, как и над всеми остальными, машина «думает» 8 минут, так как и

[1]) Фотография воспроизводится из журнала «Scientific American».

здесь она рассматривает 2800 вариантов с начальными ходами Ka3, Kc3, Kf3, Kh3, e3, e4, d4.

Человек «сообщает» машине свои ходы, пробивая соответствующую комбинацию отверстий на перфокарте или путем нажатия соответствующих клавиш на пульте управления. Машина отвечает, напечатая свой ход и получающуюся позицию. После ввода в машину очередного хода человека возникающая позиция также печатается для контроля. В случае, если человек ошибется и сделает ход, противоречащий правилам игры, будет напечатано вежливое «please check last move» (пожалуйста, проверьте последний ход). Вежливость вообще «свойственна» машине ИБМ-704; как бы ни окончилась игра, партнер получает благодарность «за интересную партию». (В скобках заметим, что в машину «МАДАМ» в свое время забыли «вложить вежливость». Она вела себя крайне «неспортивно», прекращая игру без всяких «объяснений» в случае безнадежного положения.)

Пока машине очень часто приходится благодарить за проигранные ею партии. Она играет подобно новичку, хотя и обладающему выдающимися способностями к расчету, но неумело пользующемуся ими и не гарантированному от «зевков». Кроме того, она слишком «догматична», и ее железная прямолинейность в проведении немногих вложенных в ее программу простейших принципов не всегда оказывается полезной. Можно ли существенно улучшить такое положение?

7. ВОЗМОЖНОСТИ УЛУЧШЕНИЯ ИГРЫ МАШИНЫ

Улучшение игры шахматных машин возможно за счет приближения метода нахождения «достаточно хорошего хода» машиной к методу, которым пользуется человек [1]).

На первом месте в этой связи стоит вопрос об уменьшении числа рассматриваемых вариантов. Только за счет уменьшения количества вариантов можно улучшить игру машин, увеличив глубину расчета. Конечно, отбор вариантов значительно увеличит объем некоторых частей программы и время расчета одного варианта, но сокращение количества рассматриваемых вариантов с запасом компенсирует это увеличение. Например, если бы удалось даже за счет двукратного увеличения времени на расчет одного варианта сократить количество рассматриваемых ходов с 7 до 4, (для ИБМ-704) то это позволило бы при том же времени расчета (8 минут) увеличить глубину расчета до трех ходов. (Заметим, что расчет на 3 хода вперед для 7 рассматриваемых ходов занял бы $6^1/_2$ часов.)

С другой стороны, должно быть ясно, что отбор в каждой позиции фиксированного числа ходов не имееет перспектив, так как даже

[1]) Очень ценные идеи в этом отношении содержатся в статье гроссмейстера М. Ботвинника «Люди и машины за шахматной доской» («Комсомольская правда» от 3 января 1961 г.).

из четырех отбираемых ходов (что не всегда достаточно, как видно из рис. 7) нужна скорость машины в несколько миллионов операций в секунду, чтобы в пределах нескольких минут рассчитать варианты на 4 хода вперед. Лишь построение хорошей функции $h(P, M)$ позволит сократить число рассматриваемых вариантов с тысяч до сотен или десятков и даст возможность рассчитывать «главные» варианты на глубину в 5—10 ходов.

Весьма перспективной идеей для поисков в этом направлении представляется алгоритмизация «игры по плану». Если в конкретной позиции разумно выбрать частную цель, к достижению которой следует стремиться, то можно резко сократить число вариантов, подлежащих рассмотрению. Попытка рассмотрения в первом приближении такого круга вопросов уже была сделана [1]). Наибольшие трудности, очевидно, представляет выбор плана. Он должен первоначально производиться на основе анализа особенностей позиции, а затем проверяться расчетом. Это потребует времени, но выбор плана должен делаться не на каждом ходу, а лишь в переломные моменты, при резком изменении типа позиции. Разумеется, осуществление всего этого потребовало бы огромной предварительной работы по выработке «системы игры», достаточно совершенной и не слишком громоздкой.

Осуществление указанной идеи имело бы значение и для решения второго важного вопроса — о наилучшем выборе оценочной функции $f(P)$. Оценка позиции станет более гибкой, если оценочная функция будет модифицироваться в зависимости от типа позиции и конкретной цели, которая имеется в виду для данного периода игры.

Как уже упоминалось выше, улучшения игры машины в дебюте можно добиться введением в ее память словаря дебютных вариантов. Машина выбирает ходы из этого словаря до тех пор, пока ее противник играет «по теории»; при отклонении от дебютных вариантов происходит переход машины на расчет вариантов. Реализация словаря дебютных вариантов на машине является принципиально несложным, но очень трудоемким делом.

Вполне реально также осуществление на машине системы правил для игры в эндшпилях (особенно в простейших). В случае обнаружения на доске признаков эндшпиля, машина определяет его тип и включает соответствующую программу, реализующую алгоритм для данного типа эндшпиля.

Очевидно, что добиться достаточно хорошей игры машины в дебюте и в эндшпиле можно было бы раньше и полнее, чем для середины игры. Но сейчас наибольшие усилия направлены на достижение хорошей игры машины в средней стадии, так как именно в этом

[1]) A. Newell, The chess machine; an example of dealing with a complex task by adaptation, Proceedings of the Western Joint Computer Conference, 1955, стр. 101 — 108.

заключается основной интерес задачи и именно здесь лежит ключ к решению других важных проблем.

Интересным, но пока нереализованным способом улучшения игры шахматной машины является метод обучения.

Под обучением машины понимается процесс, при котором меняется приспособления машины к некоторому классу условий, в которые ставится машина, улучшается. Для шахматной машины классом условий являются ее противники, а мерой приспособления — относительная частота выигрышей.

Были построены некоторые обучающиеся устройства и «обучающиеся» программы для цифровых программно-управляемых машин. В качестве примера обучающихся устройств и программ можно привести «мышь» К. Шеннона и программы А. Этингера. Обсуждалась возможность создания обучающейся программы и для шахматной машины [1]).

Предлагался такой принцип работы обучающейся шахматной машины. Вначале в машину вводят программу, содержащую правила игры, элементарную стратегию, а также методы улучшения этой стратегии.

Укажем на следующие возможные методы улучшения стратегии:

1. В ходе игры человек высказывает одобрение или неодобрение ходу, сделанному машиной в некоторой позиции, или, кроме того, указывает еще и на верный ход в позиции.

2. Машина анализирует ходы сильного противника в процессе игры и в дальнейшем старается подражать ему.

3. Во время или после игры машина сама анализирует игру для обнаружения своих ошибок и недопущения их в дальнейшем.

Все эти методы могут быть использованы машиной как для запоминания правил игры в типовых позициях, так и для выработки общих принципов шахматной игры.

Конечно, составление программы для обучающейся шахматной машины представляет дело огромной трудности, и пока практически можно говорить лишь о реализации на машине простейших элементов обучения. Например, можно поставить задачу об изменении машиной коэффициентов оценочной функции $f(P)$. Задав произвольные соотношения коэффициентов, отражающих относительную ценность фигур, интересно было бы проследить за тем, как они будут меняться машиной в зависимости от результатов ее игр. В дальнейшем можно будет перейти к изменению коэффициентов функций $g(P)$ и $h(P, M)$.

Упомянем еще об одном, легко осуществимом и полезном приеме — применении «стратегии поведения» [2]). Если машина пользуется чистой

[1]) C. E. Shannon, Computers and Automata, Proceedings of the IRE 41, № 10.

[2]) Считается, что задана *стратегия поведения* (в отличие от чистой стратегии, о которой шла речь выше), если каждой возможной позиции с ходом белых (черных) поставлен в соответствие не один из допустимых в этой позиции ходов, а распределение вероятностей допустимых ходов.

стратегией, то ее ход в каждой позиции определяется однозначно. Здесь кроется опасность, заключающаяся в том, что достаточно «сильный» противник, выиграв хотя бы один раз, будет в дальнейшем партиях повторять всю игру без изменения, выигрывая все время. Чтобы избежать этого, достаточно использовать стратегию поведения, приписывая вероятности, отличные от нуля, нескольким лучшим ходам, определяемым в результате расчета, и делая затем выбор между ними с помощью соответствующего случайного процесса.

Изложенные возможности улучшения игры машины и перспективы развития электронных цифровых машин заставляют согласиться с мнениями К. Шеннона и М. Ботвинника, что придет время, когда шахматная машина победит человека. Но, принимая во внимание трудности, которые предстоит преодолеть на этом пути, можно утверждать, что это произойдет нескоро.

8. ПРОГРАММИРОВАНИЕ ШАХМАТ ДЛЯ УВМ

В последней части статьи мы рассмотрим более подробно некоторые вопросы, связанные с непосредственным программированием шахматной игры для универсальной цифровой программно-управляемой машины. В качестве примера машины мы возьмем условную вычислительную машину (УВМ), подробно рассмотренную в статье А. А. Ляпунова и Г. А. Шестопал [1]).

Кодирование. Простейшим и наиболее естественным методом кодирования шахмат является метод, при котором каждому полю шахматной доски ставится в соответствие одна ячейка памяти УВМ. Если условиться, какое содержимое ячейки будет соответствовать данному типу фигуры, то по содержимому каждой из выбранных 64 ячеек памяти можно узнать, какая фигура находится на соответствующем поле.

Будем, например, изображать пешку единицей в 10-м разряде ячейки (т. е. в разряде единиц, соответствующем 2^0), коня — единицей в 9-м разряде ячейки (соответствующем 2^1), слона — единицей в 8-м разряде (2^2), ладью — единицей в 7-м разряде (2^3), ферзя — единицей в 6-м разряде (2^4) и короля — единицей в 5-м разряде (2^5). Для отличия белых фигур от черных в случае последних мы будем ставить единицу в первом разряде (разряде знака). При таком способе кодирования белым фигурам соответствуют положительные, а черным фигурам — отрицательные целые числа. На рис. 9 приведены коды ячеек, соответствующие некоторым фигурам.

Предположим далее для определенности, что «доска с фигурами» расположена в памяти в первых 64 ячейках с номерами N, равными 0, 1, 2, ..., 63, или в двоичной системе

$$0000000000,\ 0000000001,\ 0000000010,\ \ldots,\ 0000111111.$$

В дальнейшем мы будем несколько иначе записывать номера ячеек, перейдя к специальной, так называемой «двоично-восьмеричной» форме записи. При этом запись вида

$$N = \alpha_4\alpha_3\alpha_2\alpha_1$$

[1]) А. А. Ляпунов и Г. А. Шестопал, Начальные сведения о решении задач на электронных вычислительных машинах, «Математическое просвещение», вып. 1, 1957, стр. 57—74.

означает, что $N = \alpha_4 8^3 + \alpha_3 8^2 + \alpha_2 8 + \alpha_1$. Здесь α_1, α_2, α_3 могут принимать значения от 0 до 7, а α_4 — только два значения — 0 и 1.

От двоично-восьмеричной формы записи легко перейти к обычной двоичной и наоборот. При этом каждое α_i $(i = 1, 2, 3)$ заменяется своим двоичным

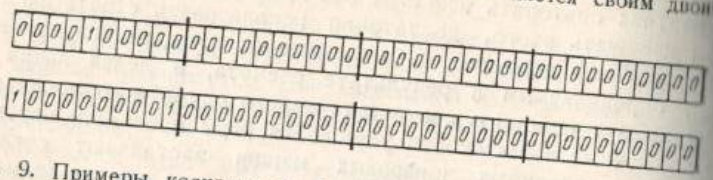


Рис. 9. Примеры кодирования фигур: *а*) белый король $(2^5 = 32)$, *б*) черная пешка $(-2^0 = -1)$.

(всегда состоящим из трех знаков) изображением. Перевод из одной системы в другую осуществляется по следующему «словарю»:

$0 \equiv 000$, $1 \equiv 001$, $2 \equiv 010$, $3 \equiv 011$, $4 \equiv 100$, $5 \equiv 101$, $6 \equiv 110$, $7 \equiv 111$.

Например,

$$1506 \equiv 1\,101\,000\,110,\ 0\,000\,111\,111 \equiv 0077$$

и т. д.

Приказ, приведенный на рис. 10, также легко перевести в двоично-восьмеричную запись:

$$0004\ 0015\ 0017\ 0017.$$

Теперь легко понять, что номера ячеек от 0 до 63 запишутся так:

$$0000,\ 0001,\ 0002,\ \ldots,\ 0077.$$

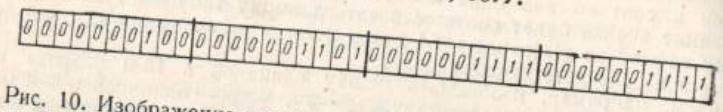


Рис. 10. Изображение приказа для УВМ: код, записанный в 13-й ячейке, разделить на код, записанный в 15-й ячейке, и результат записать снова в 15-ю ячейку.

Читатель легко поймет удобство и компактность новой записи, обратившись к рис. 11, *а*, на котором занумерованы поля шахматной доски (указаны лишь α_2, α_1; $\alpha_3 = \alpha_4 = 0$). Оказывается, что в новой записи цифра α_1 кодирует номер горизонтали, а α_2 — номер вертикали. На рис. 11, *б* приведены коды соответствующих ячеек при начальном расположении фигур.

«Ход» на нашей «доске» (со взятием или без него) будет заключаться в том, что содержимое двух ячеек памяти изменится — содержимое первой передается во вторую, а в первую записывается нуль.

Например, первый ход партии e2 — e4 осуществляется двумя приказами:

ПЧ 0041 0000 0043,
ПЧ 0154 0000 0041 [1])

(предполагается, что в ячейке 0154 находится вспомогательный код «нуль»; об этом см. ниже). Подобным же образом, но с некоторыми отличиями будут

[1]) Начиная с этого места мы не будем использовать жирный курсивный шрифт для обозначения чисел, записанных в двоично-восьмеричной системе.

осуществляться рокировка, превращение пешки, взятие на проходе. Например, h2 — h1 Ф производится приказами:

ПЧ 0101 0000 0070,
ПЧ 0154 0000 0071

при условии, что в ячейке 0101 находится вспомогательный код «—16», а в ячейке 0154 — «нуль»).

Заметим, что можно было бы предложить и другие способы кодирования шахматной доски. В частности, вместо 64 ячеек для кодирования доски можно было бы ограничиться 32 ячейками (по числу всевозможных фигур), которые

	a	b	c	d	e	f	g	h
8	07	17	27	37	47	57	67	77
7	06	16	26	36	46	56	66	76
6	05	15	25	35	45	55	65	75
5	04	14	24	34	44	54	64	74
4	03	13	23	33	43	53	63	73
3	02	12	22	32	42	52	62	72
2	01	11	21	31	41	51	61	71
1	00	10	20	30	40	50	60	70

а)

	a	b	c	d	e	f	g	h
8	-8	-2	-4	-16	-32	-4	-2	-8
7	-1	-1	-1	-1	-1	-1	-1	-1
6	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1
1	8	2	4	16	32	4	2	8

б)

Рис. 11. Кодирование шахматной доски: *а*) номера ячеек, соответствующих данному полю; *б*) содержимое соответствующих ячеек при начальном расположении фигур.

хранили бы информацию о поле, на котором стоит данная фигура. При уничтожении какой-нибудь из фигур соответствующая ячейка из дальнейшего рассмотрения выбывает. Ход любой фигурой заключается в изменении содержания (кода) соответствующей ячейки, хранящей информацию о положении этой фигуры.

После того как закодирована шахматная доска и расположение фигур на ней, можно перейти к вопросу о составлении программы, реализующей шахматную игру. Из всех блоков этой очень сложной программы мы сможем несколько подробнее рассмотреть лишь наиболее простые. Разберемся сначала в том, как машина осуществляет всевозможные ходы различных фигур, удовлетворяющие правилам игры.

Образование ходов фигур. По числу различных типов фигур естественно иметь в машине 6 отдельных программ (или, как говорят, подпрограмм), которые предназначаются для образования возможных ходов фигур [1]). При построении каждой из таких подпрограмм нужно учитывать следующие моменты.

Во-первых, каждая подпрограмма должна обеспечивать возможность образования максимального количества различных ходов, определяемого типом фигуры (для ферзя 27, для ладьи 14, для коня 8 и т. д.). Во-вторых, каждая

[1]) При желании количество таких подпрограмм можно сократить, заметив, что ходы ферзя получаются комбинацией ходов слона и ладьи, а ходы короля являются частным случаем всевозможных ходов ферзя.

подпрограмма должна учитывать ограничение количества возможных ходов в зависимости от положения фигуры на доске. (Например, конь в середине доски имеет 8 возможных ходов, а в углу доски лишь 2.) В-третьих, необходимо учитывать ограничение количества возможных ходов в зависимости от расположения на доске других фигур — как своих (занимающих необходимое поле или препятствующих движению в данном направлении), так и чужих (связывающих соответствующую фигуру или преграждающих ей путь).

Приведем пример простейшей подпрограммы для получения ходов коня, отвечающих правилам шахматной игры.

Выше каждому полю шахматной доски была поставлена в соответствие пара чисел a_2, a_1. Эти числа дают нам номер соответствующей вертикали (i) и горизонтали (j) шахматной доски. Если обозначить через i_0 и j_0 цифры, определяющие в некоторый момент положение коня на доске, то всевозможные ходы коня даются соотношениями $i = i_0 + a$, $j = j_0 + b$, где $a = \pm 1$ для $b = \pm 2$ и $a = \pm 2$ для $b = \pm 1$. При этом необходимо, чтобы $0 \leqslant i, j \leqslant 7$.

Восемь пар значений a и b можно свести в следующую таблицу (сверху дан номер k, соответствующий данной паре a и b):

k	0	1	2	3	4	5	6	7
a	1	1	−1	−1	2	2	−2	−2
b	2	−2	2	−2	1	−1	1	−1

Итак, пусть машина «знает» номер поля, на котором стоит конь. Подпрограмма должна работать таким образом, чтобы указать номера всех тех полей шахматной доски, куда этот конь может быть поставлен. Поэтому машина «перебирает» все значения k, приведенные в таблице, и для каждого из этих значений либо указывает номер поля, на которое может быть поставлен конь, либо дает информацию о том, что соответствующий ход конем невозможен. Пусть, например, конь стоит на поле f3 ($i_0 = 5$, $j_0 = 2$). Посмотрим, как машина «проверяет» очередное значение $k = 2$. Условимся прежде всего, что в ячейках 0102 и 0103 будут записаны номер поля f3 и значение $k = 2$ соответственно в виде:

(0102) 0000 0052 0000 0000 (0103) 0000 0002 0000 0000

Далее, в ячейках 0160—0167 специальным образом записаны значения a, соответствующие различным k, а в ячейках 0170—0177 — значения b:

k	№ ячейки	$a(k)$	№ ячейки	$b(k)$
0	(0160)	0000 0010 0000 0000	(0170)	0000 0002 0000 0000
1	(0161)	0000 0010 0000 0000	(0171)	1000 0002 0000 0000
2	(0162)	1000 0010 0000 0000	(0172)	0000 0002 0000 0000
3	(0163)	1000 0010 0000 0000	(0173)	1000 0002 0000 0000
4	(0164)	0000 0020 0000 0000	(0174)	0000 0001 0000 0000
5	(0165)	0000 0020 0000 0000	(0175)	1000 0001 0000 0000
6	(0166)	1000 0020 0000 0000	(0176)	0000 0001 0000 0000
7	(1067)	1000 0020 0000 0000	(0177)	1000 0001 0000 0000

Кроме того, в ячейках 0150—0154 записываются некоторые вспомогательные коды, назначение которых станет понятным из дальнейшего:

(0150) 0000 0070 0000 0000
(0151) 0000 0007 0000 0000
(0152) + 0160 0105 0107
(0153) + 0170 0106 0110
(0154) 0000 0000 0000 0000

Проверка значения $k = 2$ будет осуществляться с помощью следующей последовательности приказов:

(0200) ∧[1] 0102 0150 0105	Этот приказ с помощью вспомогательного кода из ячейки 0150 выявляет номер i_0 вертикали, на которой стоит конь. Результат записывается в ячейку 0105: (0105) 0000 0050 0000 0000
(0201) ∧ 0102 0151 0106	Этот приказ совершенно аналогично выявляет номер j_0 горизонтали и записывает его в ячейку 0106: (0106) 0000 0002 0000 0000
(0202) + 0103 0152 0203	Теперь нам нужно сложить $i_0 = 5$ и a, соответствующее $k = 2$. Вспомогательный код в ячейке 0152 мог бы служить для этой цели, если бы его первый адрес был не 0160, а 0162. Приказ 0202 каждый раз и меняет первый адрес приказа 0152 с 0160 на 0162 (т. е. на 0160 + k).
(0203) + 0162 0105 0107	Этот приказ осуществляет сложение $i_0 + a$ (2) и записывает результат i в ячейке 0107: (0107) 0000 0040 0000 0000
(0204) + 0103 0153 0205 (0205) + 0172 0106 0110	Эти два приказа совершенно аналогично складывают j_0 и b (2), причем результат j записывается в ячейке 0110: (0110) 0000 0004 0000 0000
(0206) УП$_1$ 0150 0107 p	Приказ сравнивает код i (ячейка 0107) с вспомогательным кодом 0150, проверяя условие $i \leqslant 7$. Поскольку это условие в нашем случае выполнено, дальше выполняется очередной приказ. Если бы оно было нарушено, то управление было бы передано в ячейку p управляющей программы, так как соответствующий ход был бы невозможен (выводил за пределы доски)
(0207) УП$_1$ 0107 0154 p	Здесь проверяется условие $i \geqslant 0$. Так как в нашем случае оно выполнено, то следующим исполняется очередной приказ.
(0210) УП$_1$ 0151 0110 p (0211) УП$_1$ 0110 0154 p	Эти приказы совершенно аналогично проверяют условия $j \leqslant 7$ и $j \geqslant 0$. (Код j записан в 0110.)
(0212) + 0107 0110 0104	После того как все необходимые условия проверены, этот приказ дает номер поля e5. Этот номер записывается в ячейке 0104: (0104) 0000 0044 0000 0000

[1]) Здесь для простоты вместо двоично-восьмеричного кода операции пишется ее условное обозначение (см. «Математическое просвещение», вып. 1, стр. 67. Символ ∧ указывает, что сравнивается содержимое двух ячеек и из каждой пары соответствующих цифр выбирается меньшая).

Таким образом, значение $k=2$ проверено — получен номер поля e5, на которое может быть поставлен конь с поля f3. Теперь нужно проверить и другие условия, в частности, узнать, не занято ли это поле своей фигурой. Поскольку коды белых и черных фигур отличаются знаком, то такая проверка всегда может быть проведена путем сравнения содержимого интересующей нас ячейки с нулем. Рекомендуем читателю самостоятельно найти последовательность приказов, осуществляющих указанную проверку.

Другие блоки программы. Помимо блока выработки удовлетворяющих правилам игры ходов различных фигур, программа, реализующая шахматную игру, содержит несколько других блоков. Упомянем здесь блок выработки всевозможных ходов в данной позиции, блок осуществления хода («мысленного» и фактического), блок вычисления значения оценочной функции $f(P)$, блок выбора хода, осуществляющий процесс максимизации и минимизации, и др. Должен существовать и некоторый ведущий блок для управления всеми указанными блоками. Рассмотрим принципы построения некоторых блоков.

Для получения всевозможных ходов стороны, которой предоставляется право хода, достаточно перебрать все ее фигуры, имеющиеся на доске, и для каждой из них найти все ходы, отвечающие правилам игры. Перебор всех фигур одной из сторон осуществляется сравнением содержимого ячеек, поставленных в соответствие полям шахматной доски, с некоторым числом. Например, если обнаруживается, что в ячейке стоит число, бо́льшее нуля, то это признак того, что найдена белая фигура. Какая из белых фигур найдена, узнается сравнением содержимого ячейки с последовательными степенями двойки до совпадения.

Получение различных вариантов и оценка их происходят последовательно, одно за другим, путем изменения в цепи ходов последнего хода. Изменение хода возможно образованием другого хода рассматриваемой фигуры или первого хода следующей обнаруженной фигуры. После перебора всех ответов черных на m-й ход белых этот последний меняется, и вновь образуются всевозможные ответы черных. После того как все m-е ходы белых будут рассмотрены, меняется $(m-1)$-й ход черных и вновь образуются всевозможные m-е ходы белых и черных и т. д. Такой процесс продолжается до тех пор, пока не будут образованы всевозможные первые ходы стороны, за которой очередь хода. Для управления процессом получения вариантов необходимо иметь в памяти несколько ячеек: а) ячейку для получения и хранения номера хода в цепи; б) ячейку для запоминания номера поля той фигуры, для которой вырабатываются ходы; в) ячейку для хранения номера хода данной фигуры.

Блок программы для вычисления значения оценочной функции $f(P)$ сам по себе достаточно прост, но требует предварительного подсчета количества фигур различных типов и обнаружения на доске определенных позиционных моментов, которые входят в оценку позиции.

Блок программы выбора хода осуществляет максимизацию и минимизацию функции $f(P)$ для вариантов, образуемых и оцениваемых последовательно один за другим. Количество ячеек, необходимых для хранения «экстремальных» значений функции $f(P)$ при работе программы, невелико, так как оно зависит не от числа вариантов в данной позиции, а лишь от глубины расчета вариантов. Результатом работы блока программы является информация о «лучшем» ходе в данной позиции (например, номер k хода фигуры, стоящей на поле i_0j_0).

Заметим еще в заключение, что удобно иметь в машине не одну «шахматную доску», а несколько и использовать «основную» доску для фактического осуществления выбранных ходов, а другие — для «мысленного» расчета вариантов, по аналогии с тем, как это делает человек при разборе партий.

II. НАУЧНЫЕ СООБЩЕНИЯ

ЗАДАЧА О КРУГЕ ЮНГА ДЛЯ СФЕРЫ

Ю. Д. Бураго

(Ленинград)

1. Будем рассматривать фигуры (произвольные множества точек) на сфере единичного радиуса. Расстоянием между двумя точками назовем длину кратчайшего пути по сфере от одной точки до другой. Точная верхняя граница расстояний (по сфере) между точками фигуры Ф называется диаметром этой фигуры. Вопрос, которым мы здесь займемся, формулируется так:

Каково наименьшее возможное R, такое, что кру́гом (точнее, круглой «шапочкой») [1] *радиуса R можно покрыть любую расположенную на сфере фигуру диаметра d?*

2. Аналогичный вопрос для плоскости был решен в 1901 г. английским математиком Юнгом [2]. В случае сферы вопрос значительно сложнее, и общее решение пока не найдено. Задача остается простой только до тех пор, пока фигура Ф лежит в одной полусфере.

Фигура Ф называется *связной*, если любые две ее точки можно соединить не выходящей из Ф кривой. Оказывается, что в отличие от плоскости при некоторых значениях d радиус R может быть существенно уменьшен, если ограничиться только связными фигурами.

Поэтому вопрос целесообразно решать отдельно для связных и для не обязательно связных фигур.

3. Сформулируем полученные результаты. Их доказательству будет посвящен п. 4.

---

[1]) Под «кругом» радиуса R с центром A на сфере мы понимаем геометрическое место точек сферы, удаленных от A не больше чем на R. При $R>\frac{\pi}{2}$ такой «круг» покрывает более полусферы.

[2]) Юнг доказал, что на плоскости $R=\frac{d}{\sqrt{3}}$. См., например, И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, М., 1951, задачи 16 и 67.

Теорема 1. *Если фигура* Ф *диаметра* d *лежит в одной полусфере, то* $R=\arcsin\left(\frac{2\sqrt{3}}{3}\sin\frac{d}{2}\right)$.

Теорема 2. *Если* $d<\frac{\pi}{2}+\arcsin\frac{1}{3}$ (т. е. меньше расстояния между вершинами вписанного в сферу правильного тетраэдра), *то фигура* Ф *целиком лежит в одной полусфере* [1]).

Теорема 3. *Если фигура* Ф *связная и ее диаметр* $d\leqslant\frac{2\pi}{3}$, *то* Ф *лежит в одной полусфере.*

Теорема 4. *Если* $d_n=\frac{2\pi n}{n+1}$, *где* n *— произвольное натуральное число, то* $R=d_n$.

Теоремы 1—4 исчерпывают всё, что мы можем сказать о величине R. Их содержание поясняется следующей таблицей, содержащей достаточное число нерешенных вопросов, которые можно рекомендовать вниманию читателя.

d	Произвольные фигуры	Связные фигуры
$0\leqslant d<\frac{\pi}{2}+\arcsin\frac{1}{3}$	$R=\arcsin\left(\frac{2\sqrt{3}}{3}\sin\frac{d}{2}\right)$	
$d=\frac{\pi}{2}+\arcsin\frac{1}{3}$	$R=d$	
$\frac{\pi}{2}+\arcsin\frac{1}{3}<d\leqslant\frac{2\pi}{3}$	?	
$\frac{2\pi n}{2n+1}<d<\frac{2\pi(n+1)}{2n+3}$, $n=1, 2, 3, \ldots$	?	?
$d=\frac{2\pi n}{2n+1}$, $n=1, 2, 3, \ldots$; $d=\pi$	$R=d$	

4. Доказательство теоремы 1 вполне аналогично доказательству теоремы Юнга на плоскости. Не представляет трудностей и доказательство теоремы 2, которое, так же как и доказательство теоремы 1, можно получить из рассмотрения (наименьшего) описанного круга фигуры Ф.

[1]) В то же время ясно, что если $d=\frac{\pi}{2}+\arcsin\frac{1}{3}$, то фигура Ф может не лежать в одной полусфере, и что в этом случае $R=d$; для доказательства достаточно принять за Ф четверку вершин правильного тетраэдра.

Для доказательства теоремы 4 достаточно рассмотреть множество Φ_n ($n=1, 2, 3, \ldots$), состоящее из точки O и концентрических окружностей (на сфере) с центрами в O и радиусами

$$r_k=\frac{2\pi k}{2n+1}, \qquad k=1, 2, \ldots$$

Наибольший интерес представляет теорема 3. Она показывает, что, в то время как для произвольных фигур значение R при переходе от значений меньших $\frac{\pi}{2}+\arcsin\frac{1}{3}$ к $d=\frac{\pi}{2}+\arcsin\frac{1}{3}$ претерпевает скачок, для связных фигур дело обстоит не так, и тем самым дает ответ на вопрос, предложенный на геометрическом семинаре Варшавского университета.

Доказательство теоремы 3. Рассмотрим минимальный замкнутый круг K, содержащий фигуру Ф (описанный круг фигуры Ф); в условиях теоремы 3 его радиус r заведомо не превосходит $\frac{2\pi}{3}$. Нетрудно доказать, что в случае замкнутой фигуры на границе минимального круга K может оказаться либо две точки фигуры (и тогда они делят окружность минимального круга пополам), либо больше (и тогда найдутся три точки, являющиеся вершинами остроугольного треугольника).

Для каждой точки O фигуры Ф построим открытый круг $K(O)$ радиуса $\frac{\pi}{3}$ с центром в точке O', диаметрально противоположной по отношению к точке O. Этот круг $K(O)$ мы назовем *кругом, соответствующим точке* O; он, очевидно, не содержит точек фигуры.

Идея доказательства заключается в следующем. Предположив, что Ф не умещается в полусфере (и следовательно, R заведомо больше $\frac{\pi}{2}$), мы сумеем разделить круг K на изолированные части кругами, соответствующими некоторым точкам фигуры Ф (т. е. кругами, не содержащими точек Ф), притом так, что в каждой части заведомо будут иметься точки нашей фигуры. Так как это невозможно для связной фигуры Ф, то мы заключим, что *связная фигура* Ф *диаметра* d *умещается в одной полусфере.*

Итак, мы считаем, что $\frac{\pi}{2}<r\leqslant\frac{2\pi}{3}$. Далее рассмотрим отдельно два случая.

Случай 1. Пусть $\frac{\pi}{2}<r<\frac{2\pi}{3}$. В этом случае на окружности описанного круга K не может быть двух противоположных точек, принадлежащих Ф (ибо расстояние между такими точками больше $\frac{2\pi}{3}$). Значит, на этой окружности найдутся три точки A, B, C фигуры Ф, являющиеся вершинами остроугольного треугольника.

Возьмем какие-то две из них, скажем A и B; они удалены друг от друга менее чем на $\frac{2\pi}{3}$. Заметим, что точка C лежит между концами A_1, B_1 диаметров AA_1 и BB_1 круга K. Круги $K(A)$ и $K(B)$ содержат точки A_1 и B_1, они пересекут круг K по разные стороны от точки C и пересекутся между собой внутри K (рис. 1).

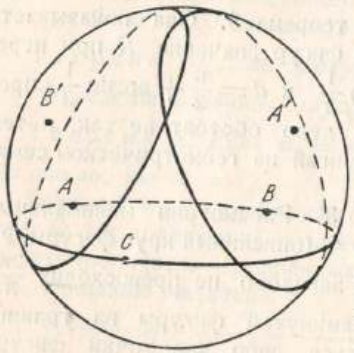


Рис. 1.

Таким образом, $K(A)$ и $K(B)$ разделят круг K на две изолированные области, причем, скажем, точки A и C лежат в разных областях. Но это противоречит связности Ф.

Случай 2. Пусть $r = \frac{2\pi}{3}$. Возьмем любую точку O фигуры Ф и построим круг $K(O)$. Дополнительный к нему круг K будет иметь радиус $\frac{2\pi}{3}$ и будет содержать Ф, т. е. он будет (минимальным) описанным кругом. Если бы на окружности круга K имелись три точки фигуры Ф, являющиеся вершинами остроугольного треугольника, то соответствующие им (открытые!) круги отделили бы точку O от остальных точек фигуры Ф (рис. 2), что противоречит связности Ф.

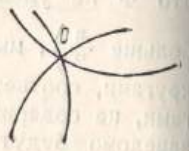


Рис. 2.

Остается предположить, что на окружности круга K лежат две точки A и B фигуры Ф, делящие эту окружность пополам. Построим круги $K(A)$ и $K(B)$. Три круга $K(A)$, $K(B)$ и $K(O)$ будут попарно касаться друг друга в точках O, B, A. Они разбивают оставшуюся часть сферы на две «треугольные» области U и V, общими точками которых являются лишь «вершины» A, B, O (рис. 3). Ясно, что каждая из этих областей целиком лежит в замкнутой полусфере (ограниченной большим кругом AOB). Докажем, что фигура Ф целиком лежит в одной из этих областей.

Предположим, что это не так. В таком случае могут иметь место три случая.

а) Фигура Ф содержит внутреннюю точку C_1 области U и внутреннюю точку C_2 области V. Точки C_1 и C_2 удалены от A, B, O менее чем на $\frac{2\pi}{3}$. Поэтому центры кругов $K(C_1)$ и $K(C_2)$ удалены от центров кругов $K(A)$, $K(B)$, $K(O)$ также менее чем на $\frac{2\pi}{3}$. Следовательно, круги $K(C_1)$ и $K(C_2)$ пересекут каждый из кругов $K(A)$, $K(B)$, $K(O)$. Поэтому, выбросив из областей U и V их пересечение с кругами $K(C_1)$ и $K(C_2)$, мы получим три изолированные части

сферы, в каждой из которых лежит одна из точек A, B, O фигуры Ф. Но это также противоречит связности Ф.

б) Фигура Ф содержит внутреннюю точку C_1 области U и граничную точку C_2 области V, скажем, точку C_2 дуги OB, отличную от O и от B (рис. 3). В таком случае пять кругов $K(A)$, $K(B)$, $K(O)$, $K(C_1)$ и $K(C_2)$ разделят дополнительную к ним часть сферы на две изолированные области, в одной из которых лежит точка A, а в другой точки O и B (рис. 4).

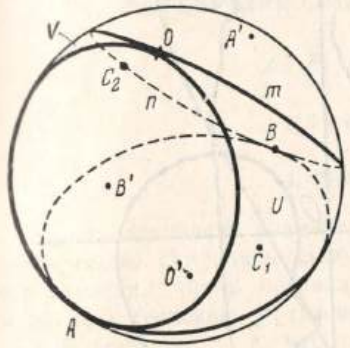


Рис. 3.

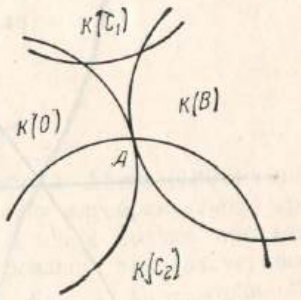
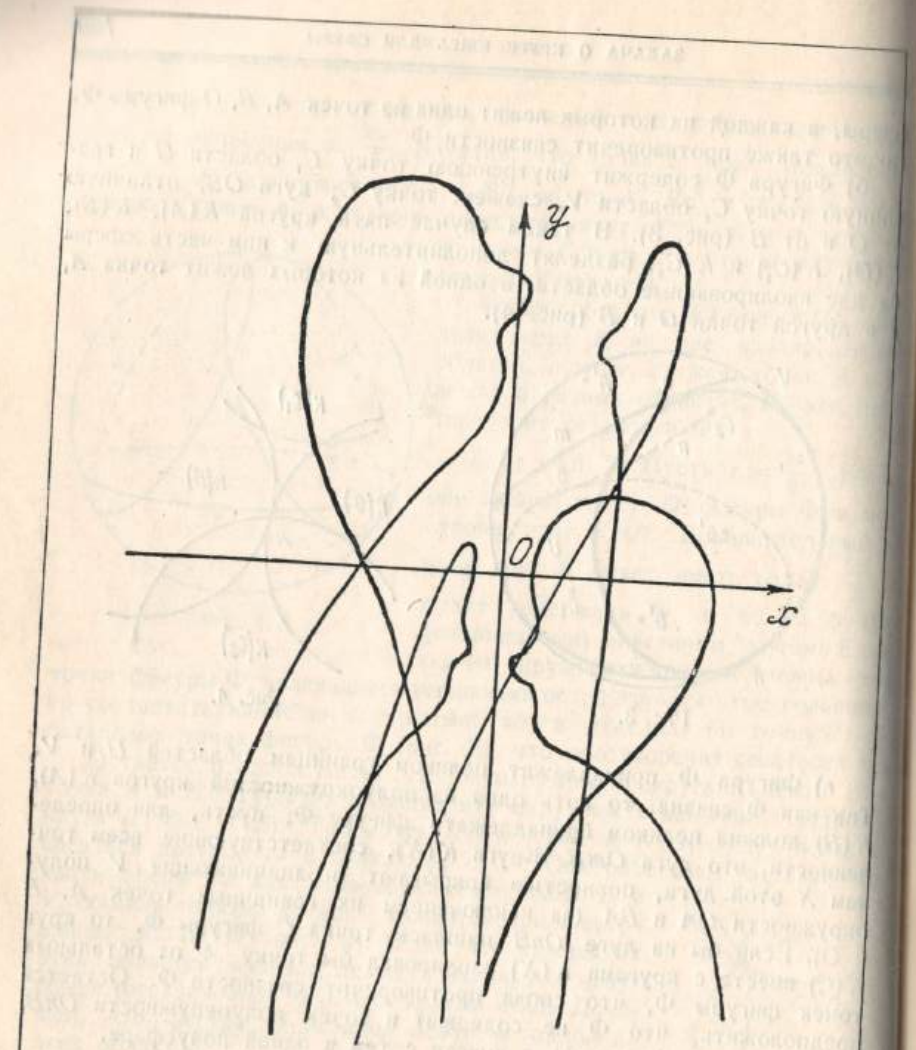


Рис. 4.

в) Фигура Ф принадлежит целиком границам областей U и V. Так как Ф связна, то хоть одна из полуокружностей кругов $K(A)$, $K(B)$ должна целиком принадлежать фигуре Ф; пусть, для определенности, это дуга OmB. Круги $K(X)$, соответствующие всем точкам X этой дуги, полностью покрывают ограничивающие V полуокружности OA и BA (за исключением их граничных точек A, B и O). Если бы на дуге OnB нашлась точка C фигуры Ф, то круг $K(C)$ вместе с кругами $K(X)$ изолировал бы точку A от остальных точек фигуры Ф, что снова противоречит связности Ф. Остается предположить, что Ф не содержит и точек полуокружности OnB, т. е. целиком лежит в U, а вместе с тем в одной полусфере.

Тем самым теорема 3 полностью доказана.



Семейство кривых

«Mathematics Magazine» 33, 1959 г.

РЕКУРРЕНТНЫЙ МЕТОД ВЫЧИСЛЕНИЯ КОРНЕЙ АЛГЕБРАИЧЕСКОГО УРАВНЕНИЯ [1])

А. Е. Гельман

(Ленинград)

I. ВВЕДЕНИЕ

1. До опубликования знаменитой работы Абеля (1826) основной задачей алгебры считалась задача решения алгебраического уравнения m-й степени. Абель показал, что с точки зрения предъявляющихся ранее к решению уравнения требований эта задача, вообще говоря, неразрешима, т. е. что если $x_1, x_2, \ldots, x_m$ — корни общего алгебраического уравнения степени $m \geqslant 5$

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m = 0, \qquad (1)$$

то невозможно с помощью конечного числа алгебраических действий (т. е. действий сложения, вычитания, умножения, деления и извлечения корня) построить функции $\varphi_1(a_0, a_1, \ldots, a_m)$, $\varphi_2(a_0, a_1, \ldots, a_m)$, ..., $\varphi_m(a_0, a_1, \ldots, a_m)$ коэффициентов $a_0, a_1, \ldots, a_m$ уравнения (1) такие, что

$$x_i = \varphi_i(a_0, a_1, \ldots, a_m).$$

Этот замечательный результат означал, что класс функций многих переменных, получаемых путем применения к аргументам функций конечного числа алгебраических действий, является слишком узким для задачи решения уравнений высоких степеней. В связи

[1]) Эта статья была напечатана в 1948 г. в машинописном сборнике студенческих научных работ Ленинградского университета. В 1949 г. в американском журнале American Mathematical Monthly была помещена задача В. Иванова (США), в которой требовалось доказать формулу, по существу совпадающую с формулой (17) настоящей статьи при $k = 1$ (т. 57 указанного журнала, № 2, стр. 111, задача 4331).

Решения этой задачи, принадлежащие Р. Лессару (Канада) и Р. Штейнбергу (США), были приведены в том же журнале за 1950 г. (т. 57, № 9, стр. 635—637); там же была указана некоторая относящаяся сюда литература [начинающаяся со статьи Даниила Бернулли в III томе журнала Commentarii (С.-Петербургской академии) за 1732 г.]. (*Прим. ред.*)

с этим представляется естественным расширить этот класс, отказавшись, например, от обременительного требования конечности числа операций.

В настоящей заметке показывается, как можно представить каждый вещественный корень произвольного алгебраического уравнения [1] явной функцией его коэффициентов, построенной применением к этим коэффициентам только рациональных действий (т. е. действий сложения, вычитания, умножения и деления), но, разумеется, вообще говоря, в бесконечном числе. Комплексный корень мы представляем как корень квадратного уравнения, коэффициенты которого также получаются применением рациональных действий к коэффициентам исходного алгебраического уравнения.

Полученные формулы удобообозримы: в них фигурируют определители неограниченно возрастающих порядков, элементами которых являются коэффициенты данного алгебраического уравнения, расположенные по весьма простому закону. При этом оказывается возможным установить рекуррентные формулы, позволяющие с большой легкостью последовательно вычислять эти определители; последнее обстоятельство открывает путь к практическому приложению полученных результатов, т. е. построению нового рекуррентного способа численного решения алгебраических уравнений. Достоинства этого способа (особенно ощутимые, если счет ведется с помощью вычислительных машин) состоят в следующем:

а) здесь употребляются только рациональные действия;

б) алгоритм вычисления корней совершенно однообразен и при любом числе пар комплексных корней не требует применения никаких искусственных приемов;

в) каждый шаг алгоритма требует всего четырех действий (два умножения, одно вычитание, одно деление), причем, что особенно важно, это число не зависит от степени уравнения (в отличие, например, от способа Данделена — Лобачевского — Греффе).

II. ФУНКЦИИ ε_k^n

2. Пусть

$$P_m(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m \tag{1}$$

— многочлен степени m с вещественными коэффициентами и $x_1, x_2, \ldots, x_m$ — корни этого многочлена, причем

$$|x_1|\leqslant|x_2|\leqslant\ldots\leqslant|x_m|.$$

[1] Из рассмотрения исключается узкий класс уравнений, у которых число корней, равных по модулю, но имеющих различные аргументы, больше двух.

В дальнейшем нам понадобятся следующие специальные симметрические функции корней:

$$\varepsilon_k^n=\sum x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_m^{\alpha_m},$$

где α_i — целые неположительные числа $(-n\leqslant\alpha_i\leqslant 0)$ и где $\alpha_1+\alpha_2+\ldots+\alpha_m=-kn$. Очевидно, что при любом n число k может принимать значения 0, 1, ..., m; так, например,

$$\varepsilon_0^1=1,\ \varepsilon_1^1=x_1^{-1}+x_2^{-1}+\ldots+x_m^{-1},$$
$$\varepsilon_2^1=x_1^{-1}x_2^{-1}+x_1^{-1}x_3^{-1}+\ldots+x_{m-1}^{-1}x_m^{-1},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\varepsilon_m^1=x_1^{-1}x_2^{-1}\ldots x_m^{-1}.$$

3. Согласно основной теореме о симметрических функциях, ε_k^n можно представить как рациональную функцию корней уравнения (1). Для того чтобы это сделать, рассмотрим тождество

$$B^n(x)\equiv\prod_{i=1}^{m}\frac{1-\left(\frac{x}{x_i}\right)^{n+1}}{1-\frac{x}{x_i}}=\frac{\left[\prod_{i=1}^{m}\left(1-\left(\frac{x}{x_i}\right)^{n+1}\right)\right]x_1x_2\ldots x_m}{(x_1-x)(x_2-x)\ldots(x_m-x)}=$$
$$=\frac{a_0}{P_m(x)}\prod_{i=1}^{m}\left(1-\left(\frac{x}{x_i}\right)^{n+1}\right)=b_0^n+b_1^nx+\ldots+b_{mn}^nx^{mn}. \tag{2}$$

С другой стороны, в силу определения функции $B^n(x)$ имеем

$$B^n(x)=b_0^n+b_1^nx+\ldots+b_{mn}^nx^{mn}=\prod_{i=1}^{m}\left(1+\frac{x}{x_i}+\frac{x^2}{x_i^2}+\ldots+\frac{x^n}{x_i^n}\right).$$

Приравнивая коэффициенты при одинаковых степенях x, получаем

$$b_\nu^n=\sum x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_m^{\alpha_m} \tag{3}$$
$$(\alpha_1+\alpha_2+\ldots+\alpha_m=-\nu,\quad 0\geqslant\alpha_i\geqslant-n).$$

Отсюда видно, что функции ε_k^n представляют собой частный случай коэффициентов b_ν^n:

$$\varepsilon_k^n=b_{kn}^n. \tag{4}$$

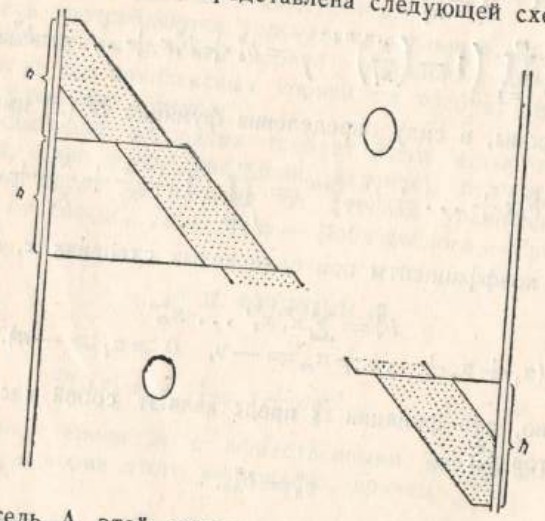
4. Коэффициенты b_ν^n многочлена $B^n(x)$ можно определить из тождества (2)

$$P_m(x)\cdot B^n(x)=a_0\prod_{i=1}^{m}\left[1-\left(\frac{x}{x_i}\right)^{n+1}\right]. \tag{2'}$$

Если произвести умножение в правой части последнего тождества, мы получим многочлен степени $m(n+1)$, ровно mn коэффициентов которого будут представлять собой тождественные нули. Приравнивая нулю соответствующие коэффициенты левой части, мы получим систему mn линейных уравнений для определения mn неизвестных $b_1^n, b_2^n, \ldots, b_{mn}^n$ ($b_0^n = 1$). Эта система имеет вид

$$\left.\begin{aligned}
&b_1^n a_0 \\
&b_1^n a_1 + b_2^n a_0 = -a_1, \\
&b_1^n a_2 + b_2^n a_1 + b_3^n a_0 = -a_2, \\
&b_1^n a_3 + b_2^n a_2 + b_3^n a_1 + b_4^n a_4 = -a_3, \\
&\cdots\cdots\cdots\cdots\cdots \\
&b_1^n a_{m-1} + b_2^n a_{m-2} + \ldots + b_m^n a_0 = -a_4, \\
&b_1^n a_m + b_2^n a_{m-1} + \ldots + b_m^n a_1 + b_{m+1}^n a_0 = -a_m, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&b_{nm-2}^n a_m + b_{mn-1}^n a_{m-1} + b_{mn}^n a_{m-2} = 0, \\
&b_{mn-1}^n a_m + b_{mn}^n a_{m-1} = 0.
\end{aligned}\right\} \quad (5)$$

5. Квадратная матрица порядка mn, отвечающая левой части системы (5), может быть представлена следующей схемой:



Определитель Δ этой матрицы легко вычисляется с помощью теоремы Лапласа. А именно:

$$\Delta = \prod_{k=0}^{m-1} D_k^n, \quad (6)$$

где D_k^n — определитель n-го порядка:

$$D_k^n = \begin{vmatrix}
a_k & a_{k+1} & a_{k+2} & \ldots & a_m & 0 & 0 & \ldots & \ldots & \ldots & 0 \\
a_{k-1} & a_k & a_{k+1} & \ldots & a_{m-1} & a_m & 0 & \ldots & \ldots & \ldots & 0 \\
a_{k-2} & a_{k-1} & a_k & \ldots & \ldots & \ldots & a_m & 0 & \ldots & \ldots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_0 & a_1 & a_2 & \ldots & & & & a_m & \ldots & 0 & \\
0 & a_0 & a_1 & \ldots & & & & & & & \\
0 & 0 & a_0 & \ldots & & & & & & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
 & & 0\,0\,a_0 & \ldots & a_{k-2} & a_{k-1} & a_k & a_{k+1} & & & \\
 & & 0\,0\,a_0 & \ldots & \ldots & a_{k-2} & a_{k-1} & a_k & & &
\end{vmatrix}. \quad (7)$$

Если kn-й столбец определителя Δ заменить на столбец, составляющий правую часть системы (5), то полученный определитель Δ_{kn} также легко вычисляется

$$\Delta_{kn} = (-1)^{kn} D_k^n \prod_{i=1}^{m-1} D_i^n. \quad (8)$$

Таким образом [1]),

$$\varepsilon_k^n = b_{kn}^n = \frac{\Delta_{kn}}{\Delta} = (-1)^{nk} \frac{D_k^n}{D_0^n}. \quad (9)$$

III. ВЫЧИСЛЕНИЕ ВЕЩЕСТВЕННОГО КОРНЯ

6. При вычислении корней уравнения (1) главную роль будут играть не функции ε_k^n, а образованные с помощью этих функций выражения

$$S_k^n = (x_1 x_2 \ldots x_k)^n \varepsilon_k^n, \quad (10)$$

уже не являющиеся симметрическими функциями корней и, следовательно, не выражаемые рационально через коэффициенты уравнения (1). В силу определения ε_k^n, очевидно, имеем

$$S_k^n = \sum \frac{x_1^{\beta_1} x_2^{\beta_2} \ldots x_k^{\beta_k}}{x_{k+1}^{\beta_{k+1}} \ldots x_m^{\beta_m}},$$

где

$$\beta_1 + \beta_2 + \ldots + \beta_k = \beta_{k+1} + \ldots + \beta_m, \quad 0 \leqslant \beta_i \leqslant n. \quad (11)$$

[1]) $\Delta \neq 0$, ибо, как легко видеть из нашей системы, можно при $a_0 \neq 0$ найти одно за другим все b_n^0.

7. Пусть теперь

$$|x_{k-1}|<|x_k|<|x_{k+1}|; \tag{12}$$

из этих неравенств, в частности, следует, что корень x_k является *вещественным*. Мы утверждаем, что в таком случае бесконечный ряд

$$S_k^\infty=\sum \frac{x_1^{\beta_1}x_2^{\beta_2}\dots x_k^{\beta_k}}{x_{k+1}^{\beta_{k+1}}\dots x_m^{\beta_m}}\ ^{1}), \tag{13}$$

где

$$\beta_1+\beta_2+\dots+\beta_k=\beta_{k+1}+\dots+\beta_m,$$

получающийся из S_k^n при $n\to\infty$, сходится абсолютно.

Действительно, всякий член этого ряда, для которого $\beta_1+\beta_2+\dots+\beta_k=\beta_{k+1}+\dots+\beta_m=N_0$, не превосходит по модулю $\left|\frac{x_k}{x_{k+1}}\right|^{N_0}$; число же таких членов равно

$$C_{A+N_0-1}^{N_0}=\frac{A(A+1)\dots(A+N_0-1)}{N_0!}\ ^{2}),$$

где $A=k(m-k)$. Таким образом, наш ряд мажорируется абсолютно сходящимся [в силу (12)!] рядом

$$\sum C_{A+N_0-1}^{N_0}\left|\frac{x_k}{x_{k+1}}\right|^{N_0}$$

и потому сам сходится абсолютно.

8. Из сходимости ряда $S_k^\infty=\lim\limits_{n\to\infty}S_k^n$ следует, что $\lim\limits_{n\to\infty}\frac{S_k^{n-1}}{S_k^n}=1$ или что

$$\lim_{n\to\infty}\frac{(x_1x_2\dots x_k)^n\varepsilon_k^n}{(x_1x_2\dots x_k)^{n-1}\varepsilon_k^{n-1}}=1. \tag{14}$$

Отсюда вытекает формула, являющаяся основным пунктом наших рассуждений:

$$x_1x_2\dots x_k=\lim_{n\to\infty}\frac{\varepsilon_k^{n-1}}{\varepsilon_k^n}. \tag{15}$$

[1]) Нетрудно проверить, что

$$S_k^\infty=\prod_{\substack{1\leqslant i\leqslant k\\ k+1\leqslant j\leqslant m}}\frac{1}{1-\frac{x_i}{x_j}}.$$

[2]) $C_{A+N_0-1}^{N_0}$ — число *сочетаний с повторениями* из A элементов (в нашем случае — элементов $\frac{x_i}{x_j}$, $1\leqslant i\leqslant k$, $k+1\leqslant j\leqslant m$) по N_0.

Разделив (15) почленно на аналогичное равенство

$$x_1x_2\dots x_{k-1}=\lim_{n\to\infty}\frac{\varepsilon_{k-1}^{n-1}}{\varepsilon_{k-1}^n}, \tag{15'}$$

получим

$$x_k=\lim_{n\to\infty}\frac{\varepsilon_k^{n-1}\varepsilon_{k-1}^n}{\varepsilon_k^n\varepsilon_{k-1}^{n-1}}. \tag{16}$$

Учитывая (9), имеем отсюда

$$\boxed{x_k=-\lim_{n\to\infty}\frac{D_k^{n-1}D_{k-1}^n}{D_k^nD_{k-1}^{n-1}}}. \tag{17}$$

Эта формула [в которой величины D_k^n определяются соотношением (7)] и дает искомое выражение корня x_k уравнения (1) через коэффициенты этого уравнения.

9. Если $x_k=x_{k+1}$, то, как легко видеть, хотя $S_k^n\to\infty$, но по-прежнему $\lim\limits_{n\to\infty}\frac{S_k^{n-1}}{S_k^n}=1$, и поэтому формула (17) сохраняет силу. Однако если $x_k\neq x_{k+1}$, но $|x_k|=|x_{k+1}|$, то $\frac{S_{n-1}}{S_n}$ не стремится ни к какому пределу, и формула (17) теряет смысл.

IV. СЛУЧАЙ КОМПЛЕКСНОГО КОРНЯ

10. В случае, когда

$$|x_{k-1}|<|x_k|=|x_{k+1}|<|x_{k+2}|, \tag{18}$$

но $x_k\neq x_{k+1}$ (например, когда корень x_k комплексный и $x_{k+1}=\bar{x}_k$), мы будем искать порознь величины $q=x_kx_{k+1}$ и $p=x_k+x_{k+1}$. Если p и q будут найдены, то, очевидно, корни x_k и x_{k+1} исходного уравнения (1) смогут быть определены как корни квадратного уравнения $x^2-px+q=0$.

Согласно формуле (15) имеем

$$x_1x_2\dots x_{k-1}=\lim_{n\to\infty}\frac{\varepsilon_{k-1}^{n-1}}{\varepsilon_{k-1}^n} \tag{19а}$$

и

$$x_1x_2\dots x_{k+1}=\lim_{n\to\infty}\frac{\varepsilon_{k+1}^{n-1}}{\varepsilon_{k+1}^n}, \tag{19б}$$

откуда

$$x_kx_{k+1}=\lim_{n\to\infty}\frac{\varepsilon_{k+1}^{n-1}\varepsilon_{k-1}^n}{\varepsilon_{k+1}^n\varepsilon_{k-1}^{n-1}} \tag{20}$$

или

$$q = x_k x_{k+1} = \lim_{n\to\infty} \frac{D_{k+1}^{n-1} D_{k-1}^{n}}{D_{k+1}^{n} D_{k-1}^{n-1}}. \tag{21}$$

11. Рассмотрим следующие функции корней полинома $P_m(x)$:

$$Q_k^n = \begin{vmatrix} \varepsilon_k^{n-1} & \varepsilon_k^{n+1} \\ \varepsilon_k^{n-2} & \varepsilon_k^{n} \end{vmatrix} \text{ и } R_k^n = \begin{vmatrix} \varepsilon_k^{n-1} & \varepsilon_k^{n} \\ \varepsilon_k^{n-2} & \varepsilon_k^{n-1} \end{vmatrix}. \tag{22}$$

Аналогично п. **8** можно показать, что

$$\lim_{n\to\infty} \frac{Q_k^n}{R_k^n} = \frac{x_k + x_{k+1}}{x_1 x_2 \ldots x_k x_{k+1}}. \tag{23}$$

Учитывая (15), заключаем отсюда:

$$x_k + x_{k+1} = \lim_{n\to\infty} \frac{\varepsilon_{k+1}^{n-1} Q_k^n}{\varepsilon_{k+1}^{n} R_k^n}, \tag{24}$$

или, в силу (22) и (9),

$$x_k + x_{k+1} = -\lim_{n\to\infty} \frac{D_{k+1}^{n-1} \begin{vmatrix} D_k^{n-1} & D_k^{n+1} \\ D_k^{n-2} & D_k^{n} \end{vmatrix}}{D_{k+1}^{n} \begin{vmatrix} D_k^{n-1} & D_k^{n} \\ D_k^{n-2} & D_k^{n-1} \end{vmatrix}}. \tag{25}$$

Так как из теоремы Лапласа легко вывести, что

$$\begin{vmatrix} D_k^{n-1} & D_k^{n} \\ D_k^{n-2} & D_k^{n-1} \end{vmatrix} = D_{k+1}^{n-1} D_{k-1}^{n-1}, \tag{26}$$

то окончательно имеем:

$$p = x_k + x_{k+1} = -\lim_{n\to\infty} \frac{\begin{vmatrix} D_k^{n-1} & D_k^{n+1} \\ D_k^{n-2} & D_k^{n} \end{vmatrix}}{D_{k+1}^{n} D_{k-1}^{n-1}}. \tag{27}$$

Формулы (21) и (27) полностью определяют корни x_k и x_{k+1} уравнения (1).

V. ПРАКТИЧЕСКИЕ ПРИМЕНЕНИЯ ПОЛУЧЕННЫХ ФОРМУЛ

12. Формулы (17), (21) и (27), определяющие вещественный или комплексно сопряженные корни уравнения (1), могут быть с успехом применены для практических вычислений. Действительно, имеет место следующее рекуррентное соотношение:

$$D_k^{n+1} = \frac{\begin{vmatrix} D_k^n & D_{k+1}^n \\ D_{k-1}^n & D_k^n \end{vmatrix}}{D_k^{n-1}} = \frac{(D_k^n)^2 - D_{k-1}^n D_{k+1}^n}{D_k^{n-1}}, \tag{28}$$

позволяющее последовательно вычислять определители D_k^n с большими значениями n (заметим, что при целых коэффициентах деление в формуле (28) происходит нацело). Эти вычисления заметно проще вычислений по способу Греффе, особенно для уравнений высоких степеней.

При последовательном вычислении D_k^n мы сразу сможем определить номера комплексных корней в ряду $|x_1| \leqslant |x_2| \leqslant \ldots \leqslant |x_m|$. Действительно, если x_k — вещественный корень уравнения (1), то

$$D_k^{2n} > 0, \quad \operatorname{sign} D_k^{2n+1} = \operatorname{sign} D_k^1 = \operatorname{sign} a_k.$$

Следовательно, знаки величин D_k^n, где k фиксировано, правильно чередуются (или сохраняются без изменения). Напротив, если x_k — комплексный корень уравнения, то знаки величин D_k^n изменяются совершенно неправильно. Это видно из того, что главная часть S_k^n для комплексного корня x_k имеет вид $\frac{\sin(n+1)\varphi}{\sin\varphi} z$, где $|z| = 1$.

13. Пример. Пусть

$$P_m(x) = -1 + 10x + 6x^2 + x^3.$$

В этом случае значения определителей D_k^n даются следующей таблицей:

n	D_3^n	D_2^n	D_1^n	D_0^n
1	1	6	10	-1
2	1	$6^2 - 10\cdot 1 = 26$	$10^2 - (-1)\cdot 6 = 106$	$+1$
3	1	$(26^2 - 106\cdot 1):6 = 95$	$(106^2 - 26\cdot 1):10 = 1121$	-1
4	1	$(95^2 - 1121):26 = 304$	$(1121^2 + 95):106 = 11\,856$	$+1$
5	1	848	125 392	-1
6	1	1953	1 326 177	$+1$
7	1	2934	14 025 978	-1
8	1	-2774	—	—

и согласно формулам (18), (21), (27)

$$x_1 = \frac{1\,326\,177}{14\,025\,978} = 0{,}09455148154\ldots,$$

$$x_2 x_3 = \frac{14\,025\,978}{1\,326\,177},$$

$$x_2 + x_3 = \frac{-2774\cdot 848 + 2934\cdot 1953}{1326177} = -\frac{8082454}{1326177}.$$

Точность вычислений — до 12-го знака после запятой.

МОДЕЛЬ ЕВКЛИДОВОЙ ГЕОМЕТРИИ НА ГИПЕРБОЛИЧЕСКОЙ ПЛОСКОСТИ [1])

Я. С. Дубнов

(Москва)

Хорошо известны различные модели гиперболической геометрии (геометрии Лобачевского) на евклидовой плоскости, например модели Клейна и Пуанкаре [2]). Цель настоящей заметки в известном смысле противоположна: здесь строится *модель евклидовой геометрии на гиперболической плоскости*. Таким образом, все геометрические термины и символы, написанные без кавычек, означают соответствующие понятия и объекты гиперболической плоскости (без несобственных элементов).

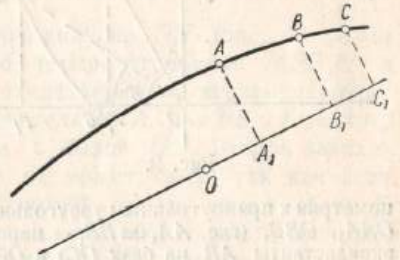


Рис. 1.

1. Выберем на плоскости определенную точку O и установим словарь.

1) «*точка*» — точка;

2) «*прямая*» — (полу)эквидистанта с базой, проходящей через O (включая сюда и прямые, проходящие через O);

3) «*инцидентность*» «точки» и «прямой» — инцидентность точки и эквидистанты;

4) «*между*» — для трех «коллинеарных» (т. е. «инцидентных» с одной «прямой») «точек» A, B и C — «точка» B лежит «между» A и C, если (см. рис. 1, где AA_1, BB_1, CC_1 перпендикулярны базе эквидистанты и, следовательно, не пересекаются) B_1 лежит между A_1 и C_1.

Из 4) и выполнения аксиомы Кантора для (гиперболической) прямой вытекает выполнение этой аксиомы и для «прямой» (вследствие взаимной однозначности соответствия $A \longleftrightarrow A_1$).

[1]) Текст этой статьи составлен И. М. Ягломом по записям, сохранившимся в литературном наследии Я. С. Дубнова.

[2]) См., например, И. М. Яглом, Геометрические преобразования, II, М., 1956, Приложения к гл. I и II.

Нам нужны будут также следующие вспомогательные понятия:

5) «*длина отрезка*». Если «точки» A, B и O не «коллинеарны», то «длину» отрезка AB «прямой» будем определять формулой (рис. 2)

$$«AB» = \operatorname{sh} OC,$$

где эквидистанта BC имеет базой OA (конечно, OC — гиперболическая длина; относительно существования точки C см. ниже, теорема 2).

Если же A, B, O «коллинеарны» (следовательно, коллинеарны), то по определению

$$«AB» = \begin{cases} |\operatorname{sh} OB - \operatorname{sh} OA|, & \text{когда } O \text{ не лежит между } A \text{ и } B, \\ \operatorname{sh} OB + \operatorname{sh} OA, & \text{когда } O \text{ лежит между } A \text{ и } B. \end{cases}$$

Примечание. Определению «длины» отрезка AB можно придать и другую форму. Восставим в точке O перпендикуляр к базе OC эквидистанты AB и опустим на него перпендикуляры AA' и BB' из точек A и B; длинам этих перпендикуляров припишем, как обычно, одинаковые или различные знаки, в зависимости от того, лежат ли точки A и B по одну или по разные стороны от $A'B'$. По определению примем

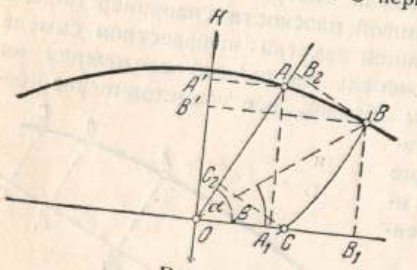


Рис. 2.

$$«AB» = |\operatorname{sh} BB' - \operatorname{sh} AA'|.$$

Для того чтобы убедиться в совпадении этого определения с прежним, достаточно заметить, что применение формул гиперболической тригонометрии к прямоугольным треугольникам OBB', OAA', а также к треугольникам OAA_1, OBB_1 (где $AA_1 = BB_1$ — перпендикуляры, опущенные из точек A и B эквидистанты AB на базу OC) и OCC_2, OBB_2 (где $CC_2 = BB_2$ — перпендикуляры, опущенные из точек C и B эквидистанты CB на базу OA) дает (мы ограничиваемся случаем, изображенным на рис. 2):

$$\text{из } \triangle OBB': \quad \operatorname{sh} BB' = \operatorname{sh} OB \cdot \sin(90^\circ - \beta) = \operatorname{sh} OB \cdot \cos\beta,$$

$$\text{из } \triangle OAA': \quad \operatorname{sh} AA' = \operatorname{sh} OA \cdot \sin(90^\circ - \alpha) = \operatorname{sh} OA \cdot \cos\alpha, \qquad (1)$$

$$\left.\begin{matrix} \text{из } \triangle OAA_1 \\ \text{и } \triangle OBB_1 \end{matrix}\right\} \quad \operatorname{sh} AA_1 = \operatorname{sh} BB_1 = \operatorname{sh} OA \cdot \sin\alpha = \operatorname{sh} OB \cdot \sin\beta, \qquad (2)$$

$$\left.\begin{matrix} \text{из } \triangle OCC_2, \\ \triangle OBB_2 \text{ и} \\ \text{из (1) и (3)} \end{matrix}\right\} \quad \begin{aligned} \operatorname{sh} CC_2 &= \operatorname{sh} BB_2 = \operatorname{sh} OC \cdot \sin\alpha = \operatorname{sh} OB \cdot \sin(\alpha - \beta) = \\ &= \operatorname{sh} OB \cdot \sin\alpha \cdot \cos\beta - \operatorname{sh} OB \cdot \sin\beta \cdot \cos\alpha = \\ &= \operatorname{sh} BB' \cdot \sin\alpha - \operatorname{sh} OA \cdot \sin\alpha. \end{aligned} \qquad (3)$$

Отсюда и из (2) следует

$$\operatorname{sh} OC = \operatorname{sh} BB' - \operatorname{sh} OA \cdot \cos\alpha = \operatorname{sh} BB' - \operatorname{sh} AA'.$$

Заметим еще, что последнее определение «длины» отрезка охватывает, очевидно, и случай «коллинеарности» точек A, B, O.

6) «*Мера угла*». Каждому «прямолинейному лучу» AB отвечает выходящий из O луч OB_1 базы эквидистанты AB, именно тот, на который проектируются точки «луча», достаточно удаленные от его

начала A (рис. 3). «Мерой» «угла» между двумя «лучами» назовем меру (конечно гиперболическую) угла между соответствующими лучами баз эквидистант.

7) «*Конгруентность отрезков*». Два отрезка назовем «конгруентными», если они имеют одинаковую «длину».

8) «*Конгруентность углов*». Два «угла» назовем «конгруентными», если они имеют одинаковую «меру».

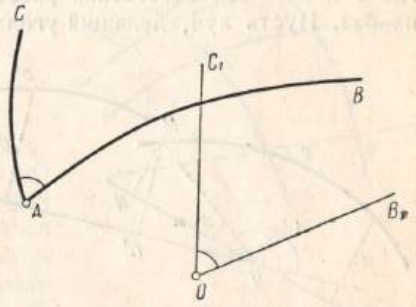


Рис. 3.

2. Теорема 1. *Если «точки» A и B различны, то существует одна и только одна «прямая», с которой они инцидентны.*

Доказательство (мы считаем точки A и B отличными от O, поскольку противный случай очевиден). Через середину M отрезка AB проведем перпендикуляр к AB и из O опустим на него перпендикуляр ON (рис. 4). Если AA_1 и BB_1 перпендикулярны ON, то четырехугольники MBB_1N и MAA_1N равны (перегибание по MN и единственность перпендикуляра, опущенного из точки на прямую); следовательно, $AA_1 = BB_1$, и значит, через A и B проходит эквидистанта с базой ON. Второй эквидистанты с базой, проходящей через O, не может быть, так как база проходящей через точки A и B эквидистанты обязательно перпендикулярна прямой, проведенной через середину хорды AB перпендикулярно AB.

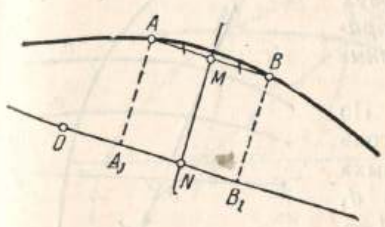


Рис. 4.

Теорема 2. *Две «прямые» пересекаются тогда и только тогда, когда базы соответствующих эквидистант различны.*

Две различные эквидистанты с одной и той же базой пересечься не могут, так как, по определению эквидистанты, расстояния их точек от базы различны. Поэтому остается рассмотреть случай, когда рассматриваемые эквидистанты имеют различные базы. Докажем, что у двух эквидистант a и a', базы которых b и b' проходят через точку O, обязательно существует общая точка.

Доказательство. Прямая b должна пересечь эквидистанту a, ибо расстояние между сторонами b и b' угла с вершиной O стремится к ∞ при неограниченном удалении точки от O; также и прямая b' должна пересечь эквидистанту a. Возьмем те лучи базы b (b'),

которые пересекаются с a' (с a); угол ω между этими лучами (рис. 5) разделим на (прилегающие к b и к b') части α и α' так, что

$$\frac{\sin\alpha}{\sin\alpha'} = \frac{\operatorname{sh} d}{\operatorname{sh} d'},$$

где d и d' — соответственно расстояния точек эквидистант a и a' от их баз. Пусть луч, делящий угол на части, пересекает a в M; такая точка существует, так как

$$\operatorname{sh} d = \operatorname{sh} MN = \operatorname{sh} OM \sin\alpha = \operatorname{sh} OM \frac{\operatorname{sh} d \sin\alpha'}{\operatorname{sh} d'},$$

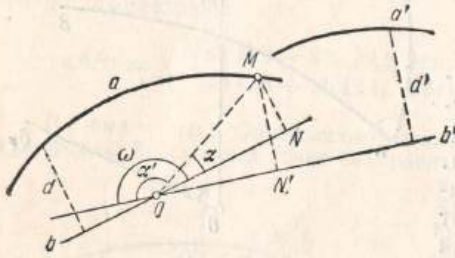


Рис. 5.

откуда

$$\operatorname{sh} d' = \operatorname{sh} OM \sin\alpha' = \operatorname{sh} MN',$$

и следовательно,

$$MN' = d',$$

т. е. точка M лежит и на эквидистанте a'.

Следствие (Евклид). *Через «точку» M, лежащую вне «прямой» a, можно провести одну и только одну «прямую», не пересекающую a, т. е. «параллельную» ей.*

Нетрудно теперь развить всю евклидову теорию параллельных. Мы ограничимся здесь двумя примерами.

Теорема 2а. *При пересечении двух «параллельных» «прямых» третьей «прямой» «соответственные углы» «равны» и т. д.*

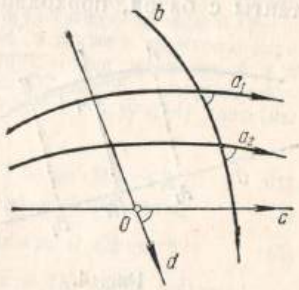


Рис. 6.

Доказательство (рис. 6). По определению «меры угла» каждый «угол» при пересечении «параллельных» «прямых» a_1 и a_2 с b равен углу между c и d, где c — база эквидистант a_1 и a_2, а d — база эквидистанты b.

Теорема 2б. *«Отрезки» «параллельных» «прямых», заключенные между «параллельными» «прямыми», «равны» между собой.*

Доказательство. По определению «длины» имеем (обозначения ясны из рис. 7)

$$«MB» = «NC» = «OQ»,\ «MA» = «ND» = «OP»,$$

откуда следует, что

$$«AB» = «DC»$$

(см. ниже теорему 3б).

3. Рассмотрим теперь подробнее установленные определениями 5 и 6 п. 1 «длины отрезков» и «меры углов».

Теорема 3. а) $«AB» = «BA»$; б) *если A, B, C коллинеарны, причем B лежит «между» A и C, то* $«AB» + «BC» = «AC»$.

Доказательство теорем а) и б) сводится к тому, чтобы показать, что в обозначениях, ясных из рис. 8, *а* и *б*,

а) $OC = OD$;

б) $\operatorname{sh} OD + \operatorname{sh} OE = \operatorname{sh} OF$;

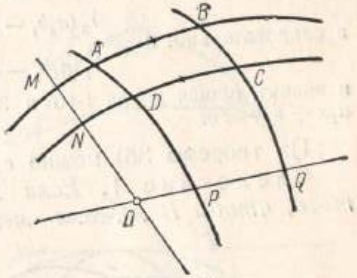


Рис. 7.

и то и другое нетрудно вывести из формул гиперболической тригонометрии (как?). Но еще проще воспользоваться вторым определением «длины», указанным в примечании на стр. 182; из этого определения утверждения а) и б) вытекают непосредственно.

Примечание. В интерпретации Кэли — Клейна теорема 3а) выражает следующий факт проективной геометрии: *если три конических сечения вписаны в четвертое коническое сечение* Σ (т. е. дважды касаются Σ) *так, что три хорды, соединяющие точки касания, проходят через одну точку, и если 1-е и 2-е конические сечения пересекаются на хорде, соединяющей точки касания с* Σ *3-го, а 2-е и 3-е — на хорде, соединяющей точки касания с* Σ *1-го, то 1-е и 3-е конические сечения пересекаются на хорде, соединяющей точки касания с* Σ *2-го* (*ср. рис. 9 с рис. 8, а*). Нетрудно дать

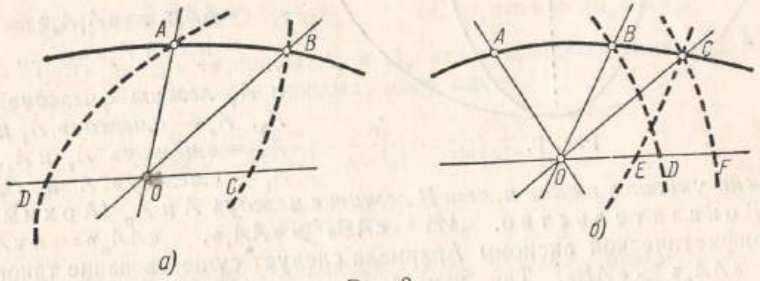


Рис. 8.

прямое аналитическое доказательство этой теоремы: если уравнение конического сечения Σ записать в виде $K = 0$, а уравнения хорд, соединяющих точки касаний Σ с нашими коническими сечениями, — в виде $a_1x + b_1y = 0$, $a_2x + b_2y = 0$, $a_3x + b_3y = 0$ (точку пересечения хорд мы принимаем за начало координат), то уравнения трех конических сечений будут иметь вид

$$K + \lambda_1(a_1x + b_1y)^2 = 0,\ K + \lambda_2(a_2x + b_2y)^2 = 0,\ K + \lambda_3(a_3x + b_3y)^2 = 0.$$

Общие точки первых двух конических сечений удовлетворяют уравнениям

$$\lambda_1(a_1x + b_1y)^2 - \lambda_2(a_2x + b_2y)^2 = 0$$

и (по условию)

откуда выводим $a_3x+b_3y=0$,

$$\lambda_1(a_1b_2-b_1a_2)^2=\lambda_2(a_2b_3-b_2a_3)^2.$$

Точно так же показывается, что

$$\lambda_2(a_2b_1-b_2a_1)^2=\lambda_3(a_2b_1-b_3a_1)^2,$$

а следовательно, и

$$\lambda_1(a_1b_2-b_1a_2)^2=\lambda_3(a_3b_2-b_3a_2)^2,$$

и значит, общая хорда 1-го и 3-го конических сечений совпадает с прямой $a_2x+b_2y=0$.

Из теоремы 3б) можно сделать следующие выводы.

Следствие 1. *Если A не лежит «между» B и C, то для того, чтобы B лежала «между» A и C, необходимо и достаточно, чтобы было* $«AB»<«AC»$.

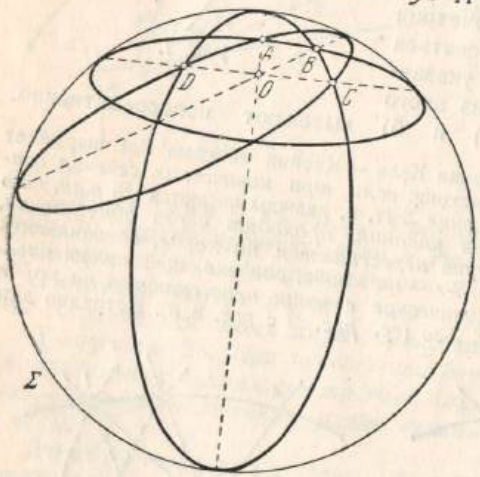


Рис. 9.

Действительно, если «точка» B лежит «между» A и C, то $«AC»=«AB»+«BC»>«AB»$; если C лежит «между» A и B, то $«AB»=«AC»+«CB»>«AC»$.

Следствие 2. *Если A, A_1, A_2, ..., A_n, ...; B — такие «точки» одной «прямой», что*

$$«AA_1»=«A_1A_2»=«A_2A_3»=\dots$$

и A_1 лежит «между» A и A_2, A_2 — «между» A_1 и A_3, A_3 — «между» A_2 и A_4, ...; A_1 — «между» A и B, то можно указать такое n, что B лежит «между» A и A_n (Архимед).

Доказательство. Из $«AB»>«AA_1»$, $«AA_n»=n«AA_1»$ и арифметической аксиомы Архимеда следует существование такого n, что $«AA_n»>«AB»$. Так как A не лежит «между» A_n и B, то остается применить предыдущее следствие.

4. Перейдем теперь к изучению «треугольников» рассматриваемой «геометрии».

Теорема 4. *Сумма внутренних «углов» «треугольника» равна $2d$.*

Доказательство. Каждый внутренний угол треугольника $A_1A_2A_3$ (рис. 10) равен углу между соответствующими лучами прямых *11*, *22*, *33* — баз эквидистант a_1, a_2, a_3. Ни один из этих лучей не повторяется дважды, т. е. среди трех углов с общей вершиной O

никакие два не имеют общей стороны (эти три из шести углов, образованных прямыми *11*, *22* и *33*, расположены «через один»). Заменяя один из этих углов вертикальным с ним, получим три угла, образующие в сумме развернутый угол.

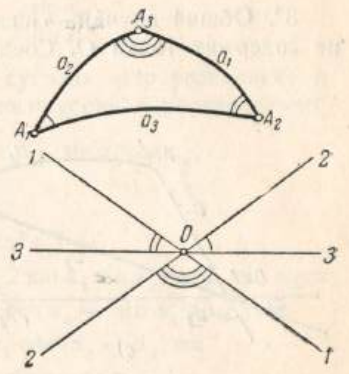


Рис. 10.

Теорема 5. *«Длины» сторон «треугольника» пропорциональны синусам противолежащих «углов».*

Доказательство. Через a_1, a_2, a_3 обозначим «длины сторон», через α_1, α_2, α_3 — «меры углов» «треугольника» $A_1A_2A_3$ (рис. 11).

1°. Сначала рассмотрим случай, когда вершина A_1 совпадает с точкой O (рис. 11, *а*). Проведя эквидистанту A_2A_2' с базой OA_3, имеем

$$«A_2A_3»=a_1=\operatorname{sh}OA_2',$$
$$a_2=\operatorname{sh}OA_3,\quad a_3=\operatorname{sh}OA_2.$$

Записывая теперь, что точки A_2 и A_3 эквидистанты A_2A_3 равноудалены от ее базы OA_2', точки A_2 и A_2' эквидистанты A_2A_2' равноудалены от ее базы OA_3, получим

$$\operatorname{sh}OA_2\sin\alpha_2=\operatorname{sh}OA_3\sin\alpha_3,\quad \text{т. е.}\quad a_3\sin\alpha_2=a_2\sin\alpha_3;$$
$$\operatorname{sh}OA_2\sin\alpha_1=\operatorname{sh}OA_2'\sin\alpha_3,\quad \text{т. е.}\quad a_3\sin\alpha_1=a_1\sin\alpha_3.$$

2°. Пусть теперь вершины A_2 и A_3 «треугольника» коллинеарны с O (рис. 11, *б*). Согласно предыдущему имеем

$$\frac{a_2}{«OA_1»}=\frac{\sin\omega}{\sin\alpha_3},\quad \frac{a_3}{«OA_1»}=\frac{\sin\omega}{\sin(\pi-\alpha_2)},\tag{1}$$

откуда

$$a_2\sin\alpha_2=a_3\sin\alpha_2.$$

Далее, из тех же треугольников OA_1A_3 и OA_1A_2 имеем

$$\frac{«OA_3»}{«OA_1»}=\frac{\sin(\pi-\omega-\alpha_3)}{\sin\alpha_3},\quad \frac{«OA_2»}{«OA_1»}=\frac{\sin(\alpha_2-\omega)}{\sin\alpha_2},$$

и следовательно

$$\frac{a_1}{«OA_1»}=\frac{«OA_3»-«OA_2»}{«OA_1»}=\frac{\sin\alpha_2\sin(\omega+\alpha_3)-\sin\alpha_3\sin(\alpha_2-\omega)}{\sin\alpha_2\sin\alpha_3}=$$
$$=\frac{\sin\alpha_2(\sin\omega\cos\alpha_3+\cos\omega\sin\alpha_3)-\sin\alpha_3(\sin\alpha_2\cos\omega-\cos\alpha_2\sin\omega)}{\sin\alpha_2\sin\alpha_3}=$$
$$=\frac{\sin\omega(\sin\alpha_2\cos\alpha_3+\cos\alpha_2\sin\alpha_3)}{\sin\alpha_2\sin\alpha_3}=\frac{\sin\omega\sin(\alpha_2+\alpha_3)}{\sin\alpha_2\sin\alpha_3}=\frac{\sin\omega\sin\alpha_1}{\sin\alpha_2\sin\alpha_3}.$$

Сравнивая этот результат с первым из равенств (1), получаем

$$\frac{a_1}{a_2} = \frac{\sin \alpha_1}{\sin \alpha_2}.$$

3°. Общий случай — ни одна из «сторон» «треугольника» $A_1A_2A_3$ не содержит точки O. Соединим O с A_3 (рис. 11, *в*) и предположим

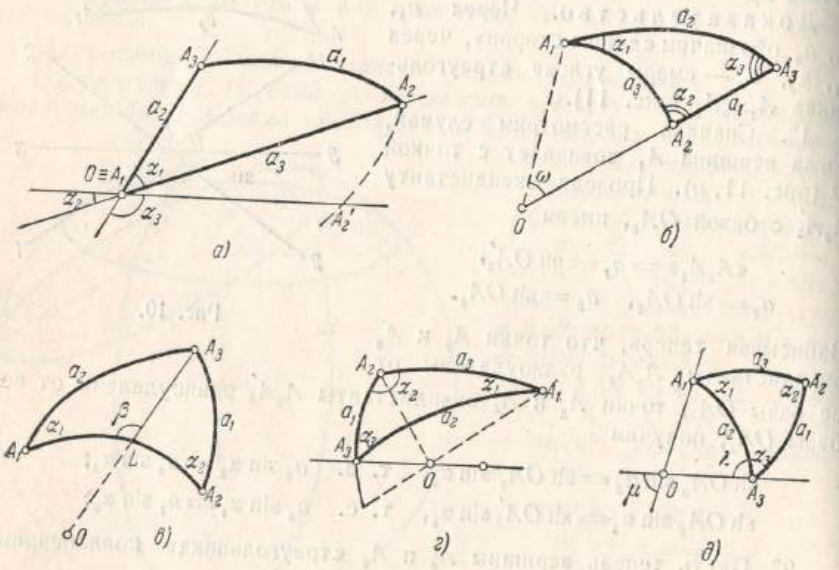


Рис. 11.

сначала, что «прямые» A_1A_2 и OA_3 не «параллельны» (пересекаются в точке L). В таком случае имеем (обозначения см. на рис. 11, *в*):

$$\frac{a_1}{\text{«}LA_2\text{»}} = \frac{\sin(\pi - \beta)}{\sin \alpha_2}, \quad \frac{a_2}{\text{«}LA_3\text{»}} = \frac{\sin \beta}{\sin \alpha_1},$$

откуда

$$\frac{a_1}{a_2} = \frac{\sin \alpha_1}{\sin \alpha_2}.$$

Если «сторона» A_1A_2 «параллельна» OA_3, но A_2A_3 не «параллельна» OA_1 и A_1A_3 не «параллельна» OA_2 (рис. 11, *г*), то

$$\frac{a_1}{a_3} = \frac{\sin \alpha_1}{\sin \alpha_3}, \quad \frac{a_2}{a_3} = \frac{\sin \alpha_2}{\sin \alpha_3},$$

откуда снова имеем

$$\frac{a_1}{a_2} = \frac{\sin \alpha_1}{\sin \alpha_2}.$$

Наконец, если A_1A_2 «параллельна» OA_3 и A_2A_3 «параллельна» OA_1 (рис. 11, *д*), то «OA_1» $= a_1$ и (по теореме 2а) $\lambda = \alpha_1$; кроме того, $\mu = \alpha_2$. А теперь из «треугольника» OA_1A_3 получаем (см. 1°)

$$\frac{\text{«}OA_1\text{»}}{a_2} = \frac{\sin \lambda}{\sin \mu} \quad \text{или} \quad \frac{a_1}{a_2} = \frac{\sin \alpha_1}{\sin \alpha_2}.$$

Следствие. Из теоремы о сумме «углов» «треугольника» и «теоремы синусов» получаем «теорему косинусов». Действительно,

$$\sin \alpha_1 = \sin(\alpha_2 + \alpha_3) = \sin \alpha_2 \cos \alpha_3 + \sin \alpha_3 \cos \alpha_2,$$

откуда следует

$$\sin^2 \alpha_1 = \sin^2 \alpha_2 (1 - \sin^2 \alpha_3) + \sin^2 \alpha_3 (1 - \sin^2 \alpha_2) + 2 \sin \alpha_2 \sin \alpha_3 \cos \alpha_2 \cos \alpha_3 =$$
$$= \sin^2 \alpha_2 + \sin^2 \alpha_3 + 2 \sin \alpha_2 \sin \alpha_3 (\cos \alpha_2 \cos \alpha_3 - \sin \alpha_2 \sin \alpha_3) =$$
$$= \sin^2 \alpha_2 + \sin^2 \alpha_3 + 2 \sin \alpha_2 \sin \alpha_3 \cos(\alpha_2 + \alpha_3) =$$
$$= \sin^2 \alpha_2 + \sin^2 \alpha_3 - 2 \sin \alpha_2 \sin \alpha_3 \cos \alpha_1.$$

Заменяя здесь синусы «углов» пропорциональными им «длинами сторон», получаем

$$a_1^2 = a_2^2 + a_3^2 - 2a_2a_3 \cos \alpha_1.$$

Следствие (важное). *«Длинами» двух «сторон» «треугольника» и «мерой угла», заключенного между ними, определяются однозначно «длина» третьей «стороны» и «меры» двух остальных «углов».*

Действительно, «сторона» a_1 определяется по a_2, a_3, α_1 с помощью «теоремы косинусов», а затем «углы» α_2 и α_3 (где $\alpha_2 + \alpha_3 = 2d - \alpha_1$) находятся из «теоремы синусов».

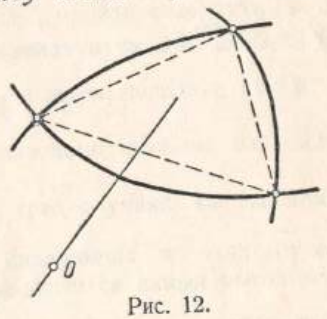


Рис. 12.

5. Всё сказанное позволяет утверждать, что *построенная нами «геометрия» полностью совпадает с обычной геометрией Евклида*. Нетрудно проверить выполнение в этой «геометрии» всех аксиом евклидовой планиметрии; в частности, следствие из теоремы 2 обеспечивает выполнение аксиомы параллельности, а следствие из «теоремы косинусов» — выполнение аксиомы о конгруентности треугольников. Выполнение линейных аксиом порядка вытекает из определения 4 п. 1 и из взаимной однозначности соответствия $A \longleftrightarrow A_1$; выполнение аксиомы Паша легко проверяется для того случая, когда секущая «прямая» проходит через O, т. е. является прямой, с помощью леммы, очевидно справедливой в гиперболической плоскости (рис. 12):

«прямая, пересекающая в одной точке дугу (хорду) эквидистанты, необходимо пересекает хорду (дугу)»; в общем же случае можно, опираясь на конгруентность треугольников, перенести всю фигуру так, чтобы секущая стала прямой.

Таким образом, мы построили модель евклидовой плоскости на гиперболической, позволяющую утверждать, что из непротиворечивости геометрии Лобачевского с необходимостью следует непротиворечивость геометрии Евклида.

КАК ВЫЙТИ ИЗ ЛЕСА?

(Об одной задаче Беллмана)

В. А. Залгаллер

(Ленинград)

В книге Р. Беллмана «Динамическое программирование»[1]) среди вопросов к главе IV автор, ссылаясь на О. Гросса, ставит перед читателями следующий вопрос.

Вы оказались внутри леса, о котором знаете только то, что он имеет вид бесконечной прямой полосы ширины l. По какой кривой следует идти, чтобы был возможно меньшим путь, после которого вы заведомо выйдете из леса, из какой бы точки вы ни отправлялись и в каком бы направлении ни пошли?

Можно, например, идти по окружности диаметра l (при этом мы заведомо выйдем из леса, пройдя путь $\pi l \approx 3{,}14\, l$) или пройти по прямой путь $l\sqrt{2}$ и затем, повернув на 90°, пройти еще путь $l\sqrt{2}$ (при этом мы заведомо выйдем из леса, пройдя путь $2\sqrt{2} \approx 2{,}82\, l$) или сначала пройти по прямой путь $l\frac{2}{\sqrt{3}}$ и затем, повернув на 120°, пройти еще путь $l\frac{2}{\sqrt{3}}$ (при этом мы заведомо выйдем из леса, пройдя путь $l\frac{4}{\sqrt{3}} \approx 2{,}31\, l$). Во всех этих трех случаях вы заведомо выйдете где-то на край леса; каждый следующий из этих путей выгоднее предыдущего, но и последний не является самым выгодным.

1. Сформулируем задачу так: *на плоскости надо найти возможно более короткую кривую, которая не может уместиться внутри бесконечной полосы P ширины l.* При преобразовании подобия длины всех кривых и ширина полосы изменяются пропорционально; поэтому достаточно решить задачу при $l = 1$.

Лемма 1. *Искомый минимум достигается хотя бы для одной выпуклой дуги L.*

[1]) Р. Беллман, Динамическое программирование, М, 1960. См. также Реферативный журнал «Математика», 1960, № 2, стр. 156.

Доказательство. Не умещающиеся внутри P кривые конечной длины существуют. Пусть σ — нижняя грань длин всех таких кривых и $\{L_i\}$ $(i=1, 2, \ldots)$ — последовательность таких кривых, длины которых стремятся к σ. Выберем некоторую последовательность $\{k_i\}$ чисел $k_i > 1$, такую, что $\lim\limits_{i\to\infty} k_i = 1$. Подобно увеличив каждую кривую L_i в k_i раз, получим последовательность кривых k_iL_i, не умещающихся в расширенных полосах k_iP; длины этих кривых также стремятся к σ. В каждую кривую k_iL_i впишем ломаную Λ_i, настолько близкую к k_iL_i, что Λ_i не умещается в P и длины Λ_i также стремятся к σ.

На каждой ломаной Λ_i фиксируем определенное направление обхода. Перенося теперь звенья Λ_i как векторы, переставим их в таком порядке, чтобы каждый следующий вектор поворачивал только вправо по отношению к предыдущему. Тем самым ломаные Λ_i удастся заменить выпуклыми ломаными Λ_i' той же длины. В любом направлении ширина Λ_i' будет не меньше, чем ширина Λ_i. Поэтому Λ_i' также не умещаются в P.

Из компактности расположенного в ограниченной части пространства множества кривых с ограниченными в совокупности длинами заключаем, что из последовательности выпуклых кривых Λ_i' можно выбрать сходящуюся к некоторой кривой L подпоследовательность Λ_{i_j}' [1]). Вместе с Λ_{i_j}' предельная кривая L будет выпуклой и не умещающейся внутри P. Кроме того, длина предельной кривой будет равна пределу длин ломаных Λ_{i_j}', т.е. σ. Поэтому на L достигается искомый минимум.

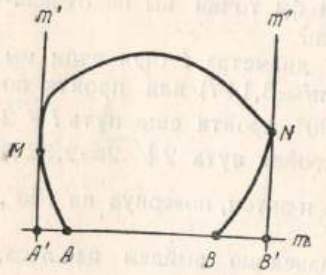


Рис. 1.

Лемма 2. *Концы A, B дуги L не совпадают друг с другом, и дуга L расположена между перпендикулярами к AB в точках A и B.*

Доказательство. Пусть m — прямая AB (или опорная прямая к L в точке A) [2]), если A совпадает с B). Проведем перпендикулярные m опорные прямые m', m'' к L. Если бы A, B не совпадали с точками пересечения m' и m, m'' и m (рис. 1), то, заменив у кривой L участки MA и NB на MA' и NB' соответственно, мы

[1]) Ср. И. М. Яглом и В. Г. Болтянский, Выпуклые фигуры, М.—Л., 1951, Дополнение 1.

[2]) Прямая m, проходящая через точку A кривой L, называется *опорной* по отношению к L, если все точки этой кривой расположены по одну сторону от m. Через любую точку выпуклой кривой может быть проведена опорная прямая.

укоротили бы L, сохранив невложимость L внутрь P. Это противоречит минимальности L и доказывает лемму 2.

Будем далее обозначать длину отрезка AB через x. Из леммы 2 следует, что $x \geqslant 1$ и A, B — точки пересечения m' и m, m'' и m (рис. 2). Проведем дуги $\overset{\frown}{CC'}$ и $\overset{\frown}{DD'}$ как дуги окружностей единичного радиуса с центрами A, B. Так как L не вмещается внутрь полосы ширины 1 с направлением AB, она хотя бы в одной точке достигает отрезка CD, отстоящего от AB на расстоянии 1. Кроме того, L на пути от A до отрезка CD не может пересечь сектора $\overset{\frown}{D'DB}$, а на пути от CD до B — сектора $\overset{\frown}{ACC'}$.

Поставим теперь вспомогательную задачу. Сначала при фиксированном $x \geqslant 1$ найдем кратчайшую из кривых $\bar{L}$, которые удовлетворяют следующим условиям:

1) соединяют точки A и B;

2) имеют хотя бы одну точку на отрезке CD;

3) на пути от A до CD не заходят внутрь области $\overset{\frown}{D'DB}$;

4) на пути от CD до B не заходят внутрь области $\overset{\frown}{ACC'}$.

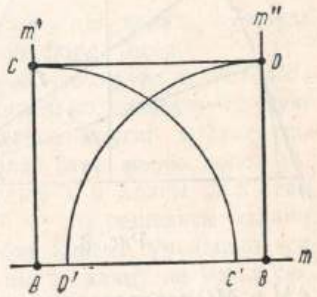


Рис. 2.

Все кривые, удовлетворяющие условиям 1) — 4), не могут умещаться внутри P; значит, все они не короче L. Кривая L — одна из таких кривых; для того чтобы ее найти, надо выяснить, какая из подобных кривых, отвечающих различным $x \geqslant 1$, является кратчайшей. Если кратчайшая линия будет единственной, то она будет единственной кривой L.

Пусть $\bar{L}$ достигает отрезка CD в точке M'. Кривая $\bar{L}$ разве лишь сократится, если ее участок $\overline{AM'}$ заменить кратчайшей линией AM', соединяющей точки A и M' и целиком лежащей в области $\overset{\frown}{ACDD'}$, а участок $\overline{M'B}$ — кратчайшей линией $M'B$, лежащей в области $\overset{\frown}{C'CDB}$. Кратчайшая AM' является либо просто отрезком прямой, либо состоит из касательной AE, проведенной из A к дуге $D'D$ некоторого участка EF этой дуги, и касательной $M'F$ к этой дуге из точки M'; аналогично выглядит и кратчайшая $M'B$. Линию, образованную двумя кратчайшими AM' и $M'B$, мы обозначим через $\bar{L}'$.

Если точка M' не является серединой M отрезка CD (но отлична от C и от D), то кривую L' можно еще сократить. Действительно, проведя близкую к CD прямую n, параллельную CD, мы отсечем от $\bar{L}'$ участок в виде двух боковых сторон треугольника с основанием, принадлежащим n, и вершиной M'. Если точка M' отлична от M, то этот треугольник будет неравнобедренным и кривую L' можно укоротить,

заменив этот участок кривой боковыми сторонами равнобедренного треугольника с тем же основанием и той же высотой. При этом выпуклость кривой будет нарушена, однако, переходом к выпуклой оболочке легко построить аналогичную выпуклую кривую не большей длины.

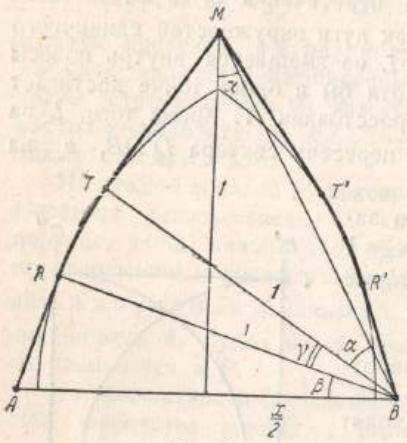


Рис. 3.

Нетрудно доказать, что кривую можно укоротить и в случае $M' = C$ или $M' = D$.

Итак, решением нашей вспомогательной задачи является *симметричная кривая* $L(x)$, *состоящая из кратчайшей линии* AM (*где* M — *середина* CD), *лежащей в области* $AC\overset{\frown}{DD'}$, *и из кратчайшей линии* MB, *лежащей в области* $BD\overset{\frown}{CC'}$. Теперь нам остается лишь выяснить, при каком $x \geqslant 1$ длина такой линии будет минимальной. При $x \geqslant \frac{2}{\sqrt{3}}$ кратчайшие линии AM и MB представляют собой отрезки прямых, длины которых убывают с уменьшением x. Поэтому нам остается найти минимум длины $L(x)$ при $1 \leqslant x \leqslant \frac{2}{\sqrt{3}}$.

Перейдем к решению этой последней задачи. Из рис. 3 находим

$$\alpha = \operatorname{arctg}\frac{x}{2}, \quad \beta = \arccos\frac{1}{x}, \quad \widehat{RT} = \gamma = \frac{\pi}{2} - 2\alpha - \beta;$$

$$L(x) = 2(AR + \widehat{RT} + TM) =$$

$$= 2\sqrt{x^2-1} + \pi - 4\operatorname{arctg}\frac{x}{2} - 2\arccos\frac{1}{x} + x.$$

Обычное исследование показывает, что производная

$$L'(x) = \frac{2\sqrt{x^2-1}}{x} - \frac{8}{4+x^2} + 1$$

непрерывна в промежутке $1 \leqslant x \leqslant \frac{2}{\sqrt{3}}$, отрицательна при $x = 1$, положительна при $x = \frac{2}{\sqrt{3}}$ и обращается в нуль в единственной точке внутри этого промежутка. Поэтому $L(x)$ достигает в этой точке минимума. Уравнение $L'(x) = 0$ преобразуется к виду

$$3x^6 + 36x^4 + 16x^2 - 64 = 0, \qquad (*)$$

его корень из интересующего нас промежутка $x \approx 1{,}0436$, чему отвечает $L(x) \approx 2{,}278$.

Мы получили следующий вывод.

Решением поставленной задачи при $l = 1$ *является изображенная на рис.* 3 *кривая* $L = ARTMT'R'B$ [*где* $x \approx 1{,}0436$ — *корень уравнения* (*)], *имеющая длину* $\approx 2{,}278$.

2. 1°. В п. 1 доказано, что построенная кривая L есть единственная минимизирующая среди выпуклых кривых. Мы утверждаем, что она вообще единственная кривая, на которой достигается минимум. Чтобы убедиться в этом, достаточно показать, что для каждой невыпуклой спрямляемой кривой ширины $\geqslant 1$ существует более короткая выпуклая кривая, также имеющая ширину $\geqslant 1$. Предоставим отыскание такого доказательства читателю.

2°. Р. Беллман ставит также вопрос о том, как искать лучшую кривую при произвольной, но точно известной форме леса.

Отметим два частных случая. Если лес имеет вид круга диаметра d, то лучший путь — отрезок длины d (действительно, всякую другую линию длины d можно уместить внутри такого круга: достаточно расположить ее так, чтобы центр круга делил эту линию пополам). Если лес имеет вид прямоугольника D ширины 1 и длины $\geqslant k$, где $k \approx 2{,}278$ — длина найденной выше кривой L, то решением задачи является та же кривая L. (Действительно, любая кривая, умещающаяся в полосе P и не умещающаяся в D, должна иметь длину не меньшую, чем длина D: поэтому такие кривые будут разве лишь длиннее L.) А какие кривые решают задачу для более коротких прямоугольников (в частности, для квадрата)?

3°. Для случая, когда надо выйти не из полосы, а из полуплоскости, направление края которой неизвестно идущему, но известно расстояние R исходной точки до края полуплоскости, задача о лучшем пути решена Исбеллом[1]): надо пройти прямо путь R, затем описать 3/4 окружности радиуса R вокруг исходной точки и снова пройти прямо путь R параллельно начальному участку пути.

4°. Выше шла речь о кривой, гарантирующей выход из полосы P при самых неблагоприятных для выбранной кривой положении начала пути и направлении движения из начальной точки. Естественнее, однако, ставить задачу Беллмана в ином смысле: какая кривая приведет к наименьшему среднему значению проходимого до выхода из P пути (скажем, при равномерном распределении исходных точек в P и равномерном распределении направлений движения в начальный момент)? Известно, что такая лучшая кривая существует, но как она выглядит?

[1]) См. Реферат. журнал «Математика», 1960, № 2, стр. 156.

Н. Н. ЛУЗИН О ТЕОРИИ ФУНКЦИЙ ДЕЙСТВИТЕЛЬНОГО ПЕРЕМЕННОГО

Отвечая на письмо О. С. Щербановского по поводу одной задачи из теории функций и обращая внимание автора письма на допущенную им ошибку, акад. Н. Н. Лузин пишет:

«Видите, Олег Сергеевич, каким осторожным должен быть всякий, кто приступает к работе по теории функций действительного переменного. Здесь всегда надо быть при выборе слов, образов, понятий и знаков столь осмотрительным, как если бы за всякое неудачное слово, определение знака полагалась смертная казнь. Чувство гибели всего предпринятого рассуждения должно все время висеть над головою. В этом отношении как предмет самодисциплины Теория Функций есть прелестнейшая из всех математических дисциплин, но и опаснейшая в то же самое время».

АВТОГРАФ ПОНСЕЛЕ

Я предполагаю, что эту заметку будут читать математики. Поэтому нет необходимости представлять им Жана Виктора Понселе — основателя проективной геометрии. Напомню только, что он участвовал в походе Наполеона в Россию. В сражении под Красным (14 августа н. ст. 1812 г.) Понселе был ранен и взят в плен. Два года он провел в плену в Саратове, усиленно занимаясь геометрией. За это время он разработал теорию свойств фигур, инвариантных относительно проектирования, т. е. проективную геометрию. В 1814 г. Понселе вернулся во Францию. С 1815 по 1825 г. работал в качестве военного инженера в арсенале города Метца. Несколько лет ушло на систематизацию и литера-

ОКОНЧАНИЕ НА СТР. 204

О ДВОЙНЫХ ЧИСЛАХ И ИХ ФУНКЦИЯХ [1])

Д. Д. Ивлев

(Москва)

1. Двойные числа. *Двойным числом* α назовем пару действительных чисел, взятых в определенном порядке: $\alpha=(a,\ b)$; при $b=0$ условимся обозначать: $(a,\ 0)=a$. *Сложение* и *умножение* двойных чисел определяется так:

$$\alpha+\beta=(a,\ b)+(c,\ d)=(a+c,\ b+d),$$
$$\alpha\cdot\beta=(a,\ b)\,(c,\ d)=(ac+bd,\ ad+bc).$$

При таких определениях сохраняются все законы арифметики.

Особую роль будет играть число $j=(0,\ 1)$. Очевидно,

$$j^2=(0,\ 1)\cdot(0,\ 1)=1,\quad j^{2n}=1,\quad j^{2n+1}=j.$$

Всякое двойное число можно представить в виде $\alpha=a+jb$:

$$\alpha=(a,\ b)=(a,\ 0)+(0,\ 1)\,(b,\ 0)=a+jb.$$

Число $\beta=\frac{1}{\alpha}=\frac{1}{a+jb}=\frac{a-jb}{a^2-b^2}$ назовем *обратным* числу α. Отсюда видно, что числа вида $a(1\pm j)$ не имеют обратных; их естественно называть *делителями нуля* (см. ниже, стр. 199).

Двойному числу $a+jb$ можно сопоставить точку плоскости с координатами a, b или вектор с теми же координатами. Числам $a(1\pm j)$ сопоставятся точки на двух пересекающихся прямых

[1]) Настоящая работа, выполненная автором еще на студенческой скамье, содержит обзор свойств так называемых двойных чисел, представляющих собой любопытную модификацию обыкновенных комплексных чисел (см. в настоящем выпуске статью И. М. Яглома «Комплексные числа и их применение в геометрии», стр. 61). С геометрическими применениями двойных чисел читатель может познакомиться по книге: Б. А. Розенфельд, Неевклидовы геометрии, гл. V, М., 1956; см. также И. М. Яглом, Проективные мероопределения на плоскости и комплексные числа, Труды семинара по векторному и тензорному анализу **7**, 1948, стр. 276—318. Обзор чисто алгебраических свойств двойных чисел содержится в статье М. С. Королева, К алгебре двойных чисел, Ученые записки Орехово-Зуевского пединститута **7**, вып. 2, 1957, стр. 113—136. (*Прим. ред.*)

$x \pm y = 0$, делящих плоскость на четыре четверти (*I*, *II*, *III*, *IV* на рис. 1).

Число $\bar{\alpha} = a - jb$ назовем *сопряженным* двойному числу $\alpha = a + jb$. *Модулем* двойного числа α назовем положительное число ρ, причем для I и III четвертей $\rho = \sqrt{a^2 - b^2}$, а для II и IV четвертей $\rho = \sqrt{a^2 + b^2}$. *Абсолютную величину* двойного числа

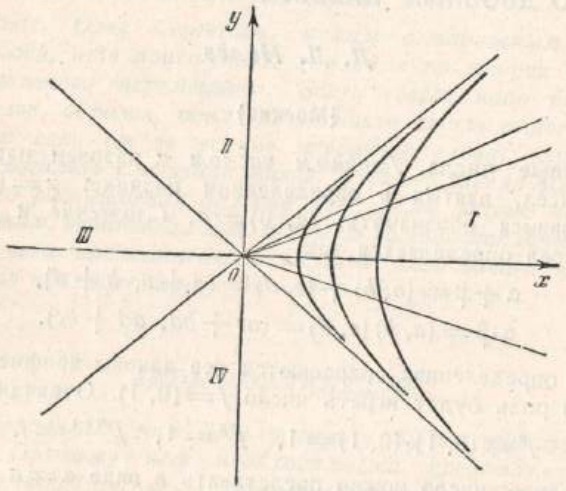


Рис. 1.

$\alpha = a + jb$ мы определим, как обычно: $|\alpha| = \sqrt{a^2 + b^2}$. Таким образом, для двойных чисел понятия модуля и абсолютной величины не совпадают, и мы будем обозначать модуль α символом $|\alpha|$ (жирные черточки!).

Определим теперь «тригонометрическую форму» двойного числа. Для чисел первой четверти

$$\alpha = a + jb = \sqrt{a^2 - b^2}\left(\frac{a}{\sqrt{a^2 - b^2}} + j\frac{b}{\sqrt{a^2 - b^2}}\right) = \rho(\operatorname{ch}\varphi + j\operatorname{sh}\varphi),$$

где φ — *аргумент* двойного числа;

$$\operatorname{tg}\varphi = \frac{b}{a}, \quad 0 < \rho < \infty, \quad -\infty < \varphi < \infty.$$

На (a, b)-плоскости линиями равного аргумента будут прямые, выходящие из начала координат, а линиями равного модуля — равносторонние гиперболы с асимптотами $x \pm y = 0$.

Для произведения и для частного двойных чисел будем иметь

$$\alpha_1 \cdot \alpha_2 = \rho_1 \cdot \rho_2 [\operatorname{ch}(\varphi_1 + \varphi_2) + j\operatorname{sh}(\varphi_1 + \varphi_2)],$$

$$\frac{\alpha_1}{\alpha_2} = \frac{\rho_1}{\rho_2}[\operatorname{ch}(\varphi_1 - \varphi_2) + j\operatorname{sh}(\varphi_1 - \varphi_2)].$$

Будет справедлива также «формула Муавра»:

$$[\rho(\operatorname{ch}\varphi + j\operatorname{sh}\varphi)]^n = \rho^n(\operatorname{ch} n\varphi + j\operatorname{sh} n\varphi), \quad \text{где } n > 0 \text{ — целое.}$$

Рассматривая корень n-й степени из двойного числа, мы находим, что имеется лишь одно значение корня, лежащее в I четверти:

$$[\rho(\operatorname{ch}\varphi + j\operatorname{sh}\varphi)]^{1/n} = \rho^{1/n}\left(\operatorname{ch}\frac{\varphi}{n} + j\operatorname{sh}\frac{\varphi}{n}\right).$$

Легко видеть, что сумма, произведение и частное двух двойных чисел I четверти сами принадлежат I четверти, а разность, вообще говоря, не принадлежит.

В приводимой таблице показано, какой четверти принадлежит произведение двойных чисел данных четвертей.

	I	II	III	IV
I	I	II	III	IV
II	II	I	IV	III
III	III	IV	I	II
IV	IV	III	II	I

Вопрос о возвышении в целую степень двойных чисел любой четверти решается очень просто: при возвышении в четную степень двойные числа будут лежать в I четверти, а в нечетную — в своей. Иначе обстоит дело с извлечением корня. Если $\sqrt{1}$ имеет четыре значения: ± 1, $\pm j$, то величины $\sqrt{j}$, $\sqrt{-1}$, $\sqrt{-j}$ вообще не будут иметь ни одного (двойного) значения. В самом деле, пусть, например, $\sqrt{j} = a + jb$; тогда $j = a^2 + b^2 + 2jab$, откуда $a^2 + b^2 = 0$, $2ab = 1$, что невозможно. Вообще, корень четной степени можно извлечь лишь из чисел I четверти, причем он имеет четыре значения, находящиеся в разных четвертях. Корень нечетной степени можно извлечь всегда, причем он имеет одно значение, принадлежащее той же четверти, что и исходное число.

Делители нуля, по определению, имеют модуль, равный нулю. Очевидно, что любое двойное число может быть представлено в виде суммы двух делителей нуля:

$$a + jb = \frac{a+b}{2}(1 + j) + \frac{a-b}{2}(1 - j).$$

Заметим, что произведение двух делителей нуля разного рода есть нуль, ибо $(1 + j)(1 - j) = 0$. Сокращать на нулевые числа нельзя, ибо, например, $(a + jb)(1 + j) = (a + b)(1 + j)$.

Аналогом числовой римановой сферы для двойных чисел служит «числовой гиперболоид», точки которого можно взаимно однозначно

отобразить на плоскость двойного переменного. Пусть уравнение гиперболоида имеет вид

$$\xi^2-\eta^2+\left(\zeta-\frac{1}{2}\right)^2=\frac{1}{4} \quad \text{или} \quad \xi^2-\eta^2=\zeta(1-\zeta).$$

Двойному числу $\alpha=x+jy$ сопоставим точку $M_1(\xi, \eta, \zeta)$ гиперболоида, принадлежащую прямой OM, где $O(0, 0, 1)$, $M(x, y)$ (стереографическая проекция). Отсюда легко получаем:

$$\zeta=\frac{x^2-y^2}{x^2-y^2+1}, \quad \xi=\frac{x}{x^2-y^2+1}, \quad \eta=\frac{y}{x^2-y^2+1}.$$

Введение числового гиперболоида наглядно поясняет понятие четвертей и делителей нуля. При этом равносторонняя гипербола плоскости xy: $A(x^2-y^2)+Bx+Cy+D=0$ (A, B, C, D — действительные числа) является проекцией плоского сечения гиперболоида $B\xi+C\eta+(A-D)\zeta+D=0$ и обратно.

Двойному числу $\alpha=a+jb$ сопоставим комплексное число $\alpha^*=a+ib$. Если обозначить $\Phi(\alpha)=\alpha^*$, то легко проверить, что

$$\Phi(C\alpha)=C\Phi(\alpha) \qquad (C \text{ — действительное число}),$$
$$\Phi(\alpha_1+\alpha_2)=\Phi(\alpha_1)+\Phi(\alpha_2).$$

Обратное преобразование Φ^{-1} обладает теми же свойствами. Аналогично, если положить $\Psi(\alpha)=\alpha^{**}$, где $\alpha=\rho(\operatorname{ch}\varphi+j\operatorname{sh}\varphi)$ — двойное число I четверти и $\alpha^{**}=\rho(\cos\varphi+i\sin\varphi)$ — комплексное число, то

$$\Psi(C\alpha)=C\cdot\Psi(\alpha) \qquad (C \text{ — действительное число}),$$
$$\Psi(\alpha_1\cdot\alpha_2)=\Psi(\alpha_1)\cdot\Psi(\alpha_2),$$

и для действительных чисел $\Psi(c_1+c_2)=\Psi(c_1)+\Psi(c_2)$. Очевидно, что обратное преобразование Ψ^{-1} обладает теми же свойствами.

Введенные преобразования могут быть полезными при доказательстве ряда теорем исчисления двойных чисел, так как они позволяют в некоторых случаях непосредственно переносить доказательства этих теорем из теории функций комплексного переменного.

Предел последовательности двойных чисел можно было бы определить, как обычно: z_0 является пределом последовательности двойных чисел, если, начиная с некоторого $n>N(\varepsilon)$, при любом $\varepsilon>0$ имеет место неравенство $|z_n-z_0|<\varepsilon$. Можно также дать определение предела последовательности, пользуясь понятием модуля двойного числа. Заметим, что из ограниченности двойного числа по модулю вовсе не следует ограниченность его по абсолютной величине. Ограниченность по абсолютной величине будет выполнена при $|z|<A$ и $|\varphi(z)|<B$, где $\varphi(z)=\arg z$, A и B — действительные положительные числа. Таким образом, можно назвать z_0 пределом

последовательности двойных чисел $z_1, z_2, \ldots, z_n, \ldots$, если при любом $\varepsilon>0$ найдется такое $N(\varepsilon)$, что при $n>N$ $|z_n-z_0|<\varepsilon$ и $|\varphi(z_n-z_0)|<B$.

Очевидно, что два приведенных определения предела не эквивалентны. Ограниченность по углу при втором определении исключает из рассмотрения числа, лежащие на прямых, параллельных прямым $x\pm y=0$ и проходящих через точку z_0. Отсюда вытекает, что последовательность, сходящаяся по первому определению, может расходиться по второму определению — первое определение сходимости является более широким.

2. Функции двойного переменного. *Функция двойного переменного* $w=f(z)$ имеет вид

$$w=u(x, y)+jv(x, y) \qquad (z=x+jy),$$

где $u(x, y)$, $v(x, y)$ — действительные функции двух переменных.

Понятие *непрерывности* функции $w=f(z)$ можно ввести по-разному: 1) функция $f(z)$ непрерывна в точке $z=z_0$, если для каждого $\varepsilon>0$ существует $\delta(\varepsilon)>0$ такое, что имеет место $|f(z)-f(z_0)|$ для всех z, удовлетворяющих равенству $|z-z_0|<\delta$; 2) функция $f(z)$ непрерывна в точке $z=z_0$, если для сколь угодно малого $\varepsilon>0$ существует $\delta(\varepsilon)>0$ такое, что выполняется неравенство $|f(z)-f(z_0)|<\varepsilon$ для всех z, удовлетворяющих неравенству $|z-z_0|<\delta$, причем $\operatorname{Arg}(f(z)-f(z_0))<B_1$, $\operatorname{Arg}(z-z_0)<B_2$, где $B_i=\text{const}$, $-\infty<B_i<\infty$.

Приведенные два определения непрерывности функции неэквивалентны: функция, непрерывная по второму определению, может быть разрывна по первому, так как ничего нельзя сказать о значениях функции на прямых, параллельных прямым $x\pm y=0$ и проходящих через точку $w_0=f(z_0)$ (такие прямые будем называть *характеристическими*). Функция, непрерывная по первому определению, может не удовлетворять условиям непрерывности по второму: в самом деле, точки плоскости $z=x+jy$, лежащие на каком-либо нехарактеристическом направлении, могут перейти в характеристические для $w=u+jv$, что невозможно в случае второго определения.

Переходя к *дифференцированию* функции двойного переменного, отметим, что формально можно определить понятие производной различными способами, соответственно разным определениям непрерывности в точке. Мы примем следующее определение: функция $f(z)$ *дифференцируема* в точке z_0 области G, если для каждого $\varepsilon>0$ можно найти такое $\delta>0$, что неравенство $\left|\frac{f(z)-f(z_0)}{z-z_0}-f'(z_0)\right|<\varepsilon$ выполняется, когда $|z-z_0|<\delta$ ($|z-z_0|\neq 0$). Здесь $f'(z_0)$ — вполне определенное двойное число, называемое *значением производной* функции от $f(z)$ в точке $z=z_0$.

Функцию $w=f(z)$ мы назовем *моногенной* в точке, если она в ней дифференцируема, и *аналитической*, если она дифференцируема при всех z, где $|z-z_0|<\varepsilon$, $|z-z_0|\neq 0$.

Условия Даламбера — Эйлера для функций двойного переменного запишутся в виде

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y},\quad \frac{\partial u}{\partial y}=\frac{\partial v}{\partial x},$$

так что

$$\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}=0,\quad \frac{\partial^2 v}{\partial x^2}-\frac{\partial^2 v}{\partial y^2}=0,\quad v=\int\frac{\partial u}{\partial y}\,dx+\frac{\partial u}{\partial x}\,dy+C.$$

Так как решения волнового уравнения могут быть написаны в общем виде, то легко найти

$$u=\Psi(x+y)+\theta(x-y),\quad v=\Psi(x+y)-\theta(x-y)$$

(постоянную в формуле для v мы отбрасываем).

Любую аналитическую функцию двойного переменного можно представить в виде суммы двух специальных функций:

$$w=u+jv=[\Psi(x+y)+\theta(x-y)]+j[\Psi(x+y)-\theta(x-y)]=$$
$$=\Psi(x+y)(1+j)+\theta(x-y)(1-j).$$

Очевидно, что если аналитическая функция двойного переменного определена на отрезке кривой $y=f(x)$, для которой $\left|\frac{dy}{dx}\right|<1$ или $\left|\frac{dy}{dx}\right|>1$, то функция $w=w(z)$ легко может быть определена в соответствующем прямоугольнике $ABCD$ (рис. 2).

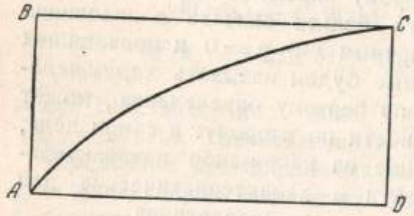


Рис. 2.

Ряды могут быть сходящимися в смысле первого или второго определения предела. Не останавливаясь на связанных с этим вопросах, мы коснемся лишь степенных рядов: $c_0+c_1z+c_2z^2+\ldots$ Если такой ряд сходится при $z=z_0$, то он сходится и при всяком z, для которого $|z|<|z_0|$, причем $\operatorname{Arg} z<B$. Определение *модуля сходимости* степенных рядов может быть дано на основе формулы Коши — Адамара $R=\lim\limits_{n\to\infty}\frac{1}{\sqrt[n]{|c_n|}}$. Из этой формулы следует, что ряды

$$1+\frac{z}{1!}+\frac{z^2}{2!}+\frac{z^3}{3!}+\ldots,\quad 1\pm\frac{z^2}{2!}+\frac{z^4}{4!}\pm\frac{z^6}{6!}+\ldots,$$
$$z\pm\frac{z^3}{3!}+\frac{z^5}{5!}\pm\frac{z^7}{7!}+\ldots$$

сходятся во всей плоскости двойного переменного. Если условимся обозначать функции, определяемые этими рядами, соответственно через e^z, $\operatorname{ch} z$, $\cos z$, $\operatorname{sh} z$, $\sin z$, то для этих функций, аналитических во всей плоскости, справедливы следующие соотношения:

$$\operatorname{ch} z=\frac{1}{2}(e^{jz}+e^{-jz}),\quad \operatorname{ch} z=\frac{1}{2}(e^{z}+e^{-z}),\quad \operatorname{sh} z=\frac{1}{2j}(e^{jz}-e^{-jz}),$$
$$\operatorname{sh} z=\frac{1}{2}(e^{z}-e^{-z}),\quad \operatorname{sh} jz=j\operatorname{sh} z,\quad \operatorname{ch} jz=\operatorname{ch} z,\quad \sin jz=j\sin z,$$
$$\cos jz=\cos z.$$

Отметим, что *некоторые соотношения теории относительности* удобно записать, пользуясь двойными числами. Для примера остановимся на преобразованиях Лоренца:

$$x^*=\frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}},\quad t^*=\frac{t-\frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Если положить $ct=t_1$, $ct^*=t_1^*$, то легко получить

$$x^*+jt_1^*=\frac{(x+jt_1)\left(1-\frac{v}{c}j\right)}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Полагая $z=x+jt_1$, $z^*=x^*+jt_1^*$, получим $z^*=ze^{j\varphi}$, где $\varphi=\operatorname{Arth}\frac{v}{c}$. Таким образом, преобразования Лоренца сводятся к гиперболическому повороту.

Понятие *интеграла* функции двойного переменного введем, как обычно. Составим сумму $\sum f(z_k)\Delta z_k$, где $\Delta z_k=z_{k+1}-z_k$ и устремим $|\Delta z_k|$ к нулю. Пределом составленной суммы и будет интеграл $\int f(z)\,dz$. Повторяя обычные в этих случаях рассуждения, можно показать, что

$$\int_\Gamma f(z)\,dz=\int_\Gamma(u\,dz+v\,dy)+j\int_\Gamma(u\,dy+v\,dx).$$

Легко доказать, что если функция $w=f(z)$ двойного переменного аналитическая в области G и имеет в каждой точке непрерывную производную, то интеграл вдоль замкнутого контура будет равен нулю (теорема Коши).

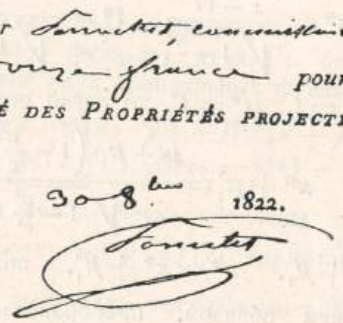
Пользуясь этой теоремой, можно показать, что

$$\int_{-\infty}^{\infty}e^{-\lambda x^2}\operatorname{ch}(2\lambda\alpha x)\,dx=\sqrt{\frac{\pi}{\lambda}e^{\lambda\alpha^2}}.$$

ОКОНЧАНИЕ ТЕКСТА СТР. 196

турное оформление результатов, полученных в Саратове. Наконец в 1822 г. они были опубликованы в виде книги «Traité des propriétés projectives des figures».

Читая эту книгу в библиотеке им. Ленина (в Москве), я обнаружил записку, приклеенную одним краем к странице книги. В конце ее — подпись чернилами: «Понселе». При чтении этой записи я испытал некоторое волнение, простительное мне как проективному геометру. Хочу извлечь на свет этот автограф великого человека, пролежавший в безвестности почти 140 лет. Вот фотокопия:



Даю русский перевод. Слова, выделенные разрядкой, написаны рукою Понселе, остальной текст напечатан:

«Получено от г-на (неразборчиво — фамилия и звание) сумма в двенадцать франков за один экземпляр „Трактата о проективных свойствах фигур".

Метц, 30 8-го, 1822.

Понселе».

Как видно, Понселе заготовил печатные бланки, на которых он покупателям своей книги давал расписку в получении денег. Первый владелец экземпляра, находящегося теперь в библиотеке, желая сохранить на память эту расписку, вклеил ее в книгу.

Н. М. Бескин

КРАТКИЕ СООБЩЕНИЯ

1. ***Р. Г. Бинг и Н. Д. Казаринов*** (R. H. Bing, N. D. Kazarinoff, *Мадисон, Висконсин, США*). **О конечности числа отражений, переводящих плоский невыпуклый многоугольник в выпуклый.** Пусть P — простой невыпуклый n-угольник (рис. 1), AB — сторона выпуклой оболочки H многоугольника P[1]), имеющая с P только две общие точки A и B. Часть $H-P$, прилежащую к AB (на рис. 1 заштрихована), отобразим симметрично от AB; в результате из исходного многоугольника путем замены в нем соответствующей части периметра ее зеркальным отображением, получим новый многоугольник $s(P)$, где s — примененная к P операция отражения. Ясно, что $s(P)$ есть k-угольник $(n-2 \leqslant k \leqslant n)$, стороны которого равны соответствующим сторонам P. Рассмотрим последовательность многоугольников $\{r_m(P)\}$, где $r_0(P)=P$, $r_m(P)=r_m(r_{m-1}(P))$ $(m>0)$ и все r_m при $m>0$ суть произвольно выбранные операции отражения s.

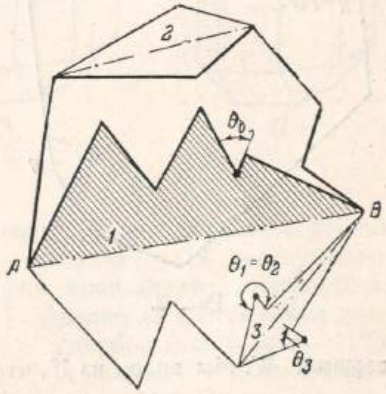


Рис. 1.

Теорема. *Последовательность $\{r_m(P)\}$ конечна, и ее последний член есть выпуклый k-угольник $(k \leqslant n)$*[2]).

Доказательство. Пусть V_0 — произвольная вершина P и $V_m=r_m(V_{m-1})$ $(m \geqslant 1)$. Допустим, что последовательность $\{r_m(P)\}$ бесконечна.

---

[1]) Выпуклой оболочкой многоугольника называется наименьший выпуклый многоугольник, содержащий данный. (*Все примечания к настоящей заметке — редакционные.*)

[2]) Эта теорема позволяет свести решение ряда экстремальных задач для простых многоугольников к (более простым!) задачам, касающимся выпуклых многоугольников. Несколько иное доказательство той же теоремы см. Ю. Г. Решетняк, Об одном приеме превращения невыпуклой ломаной в выпуклую, Успехи матем. наук **12**, № 3 (75), 1957, стр. 189—191.

1°. Докажем сначала, что $V_m \to V$ при $m \to \infty$; отсюда будет следовать, что последовательность $\{r_m(P)\}$ точечно сходится и имеет пределом некоторый предельный k-угольник, причем $k \leqslant n$. Для доказательства заметим, что если E, F, G — три внутренние точки H, не лежащие на одной прямой, то три последовательности расстояний $\{EV_m\}$, $\{FV_m\}$, $\{GV_m\}$ являются ограниченными последовательностями неубывающих положительных чисел[1]). Пусть $R_E = \lim\limits_{m\to\infty} EV_m$; аналогично определяются R_F и R_G. Три окружности с центрами в точках E, F и G и с радиусами R_E, R_F, R_G должны иметь по крайней мере одну общую точку, а поскольку точки E, F, G не лежат на одной прямой, то они имеют единственную общую точку V. Это доказывает 1°.

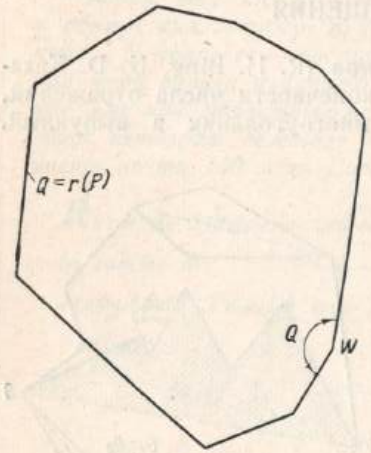


Рис. 2.

Пусть теперь W — любая вершина выпуклой оболочки Q предельного многоугольника $r(P)$ (рис. 2).

2°. Докажем, что любая вершина $r(P)$ получена в результате конечного числа N_W отражений. Чтобы доказать это, будем рассуждать следующим образом. Пусть θ_m есть угол многоугольника $r_m(P)$, соответствующий углу θ многоугольника $r(P)$ при вершине W. Мы знаем из 1°, что $\lim\limits_{m\to\infty} \theta_m = \theta$. Кроме того, $\theta < \pi$. Но если W_m не совпадает с W_{m+1} для некоторого m и если $\theta_m < \pi$, то $\theta_{m+1} = 2\pi - \theta_m > \pi$, а если $\theta_m > \pi$, то $\theta_{m+1} = 2\pi - \theta_m < \pi$ (см. рис. 1). Следовательно, $\lim\limits_{m\to\infty} \theta_m = \theta < \pi$ только в том случае, если все θ_m одинаковы, начиная с некоторого m, т. е. тогда и только тогда, когда W получится в результате конечного числа отражений.

3°. Пусть теперь $N = \max N_W$ для всех вершин N многоугольника Q. Тогда $r_N(P) = Q = r(P)$[2]).

Это очевидно! Доказательство завершено.

---

[1]) Ограниченность следует, например, из того, что все $r_m(P)$, как многоугольники одного и того же периметра p, имеют диаметр, меньший $\frac{p}{2}$, и, кроме того, имеют общую часть с H.

[2]) Q совпадает с $r_N(P)$, так как в противном случае существовала бы операция отражения, выводящая $r_N(P)$ за пределы Q, и Q не могло бы быть выпуклой оболочкой $r(P)$.

Предположение. При фиксированном n для всех n-угольников и при любом выборе операций обращения r_m число N ограничено величиной $2n$.

Замечание. Многоугольник P не обязательно должен быть простым.

Доказательство этой теоремы, данное Б. Секефальви-Надем [см. B. Sz.-Nady, Amer. Math. Monthly **46**, 1939, стр. 176—177], неверно.

2. ***А. Г. Дорфман*** (*Горький*). **Ромбокубооктаэдры и взаимные им многогранники.** Выпуклый многогранник называется *равноугольно-полуправильным* или *архимедовым*, если все его многогранные углы равны между собой, а все грани — правильные многоугольники нескольких типов. Простейшими примерами архимедовых многогранников являются призмы с правильными многоугольниками в основаниях и квадратными боковыми гранями и *скошенные призмы*, или *антипризмы* (основания — правильные многоугольники, боковые грани — равносторонние треугольники). Кроме двух бесконечных серий — призм и скошенных призм, — существует 14 равноугольно-полуправильных многогранников[1]). К ним принадлежит ромбокубооктаэдр (рис. 1, *а*) и обнаруженный В. Г. Ашкинузе[2]) *скошенный ромбокубооктаэдр* (рис. 1, *б*), остававшийся незамеченным около двух тысяч лет.

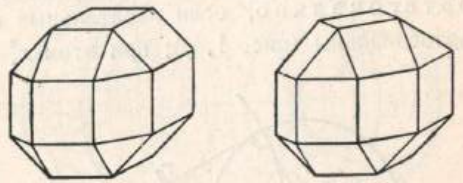


Рис. 1.

Выпуклый многогранник с одинаковыми гранями и правильными многогранными углами нескольких типов называется *равногранно-полуправильным*. Простейшими примерами равногранно-полуправильных многогранников являются *дипирамиды* с гранями, делящими пополам ребра призм Архимеда. Все остальные равноугольно- и равногранно-полуправильные многогранники находятся

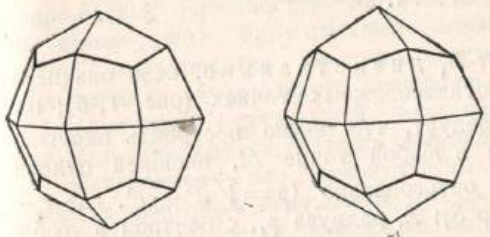


Рис. 2.

---

[1]) Они приводятся, например, в книге: Л. А. Люстерник, Выпуклые фигуры и многогранники, М., 1956, стр. 182—185.

[2]) В. Г. Ашкинузе, О числе полуправильных многогранников, Математическое просвещение, вып. 1, 1957.

в аналогичном соотношении: *середины ребер равноугольно-полуправильного многогранника принадлежат граням равногранно-полуправильного многогранника.* В частности, плоскости, делящие пополам ребра ромбокубооктаэдра и отсекающие равные пирамиды, образуют равногранно-полуправильный многогранник (рис. 2, *а*) — *икосотетраэдр* (по терминологии, принятой в кристаллографии). Он хорошо известен. Однако многогранник (рис. 2, *б*), соответствующий указанным образом скошенному ромбокубооктаэдру, приводится, по-видимому, впервые.

3. ***И. А. Каждан и С. С. Масько*** (*студентки 4-го курса МГПИ им. Ленина, Москва*). **Вариант понятия степени точки относительно окружности.** Окружность S_2 пересекает окружность S_1 ортогонально, если касательные к ним в точке их пересечения ортогональны (рис. 1, *а*); при этом $r_1^2+r_2^2=d^2$. Будем говорить,

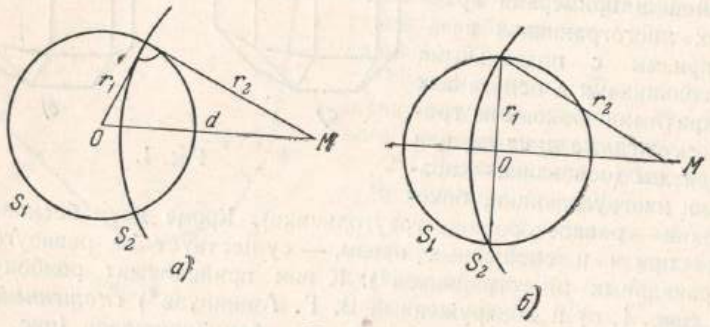


Рис. 1.

что окружность S_2 пересекает S_1 диаметрально, если она пересекает S_1 в диаметрально противоположных точках (рис. 1, *б*); при этом $r_2^2-r_1^2=d^2$. Отсюда следует, что можно построить окружности: Σ радиуса ρ, с центром в любой точке M, внешней относительно S, пересекающую S ортогонально ($\rho=\sqrt{d^2-r^2}$, где r — радиус S, $d=OM$, O — центр S); Σ_1 радиуса ρ_1, с центром в любой точке M, внутренней относительно S, которую S пересекает диаметрально ($\rho_2=\sqrt{r^2-d^2}$); Σ_2 радиуса ρ_2, с центром в любой точке M плоскости, пересекающую S диаметрально ($\rho_2=\sqrt{d^2+r^2}$).

Как известно, степень точки M относительно окружности S ($\mathrm{Ст}_S M$) равна d^2-r^2, т. е. $\mathrm{Ст}_S M=[r(\Sigma)]^2$ или $-[r(\Sigma)]^2$, где центр Σ совпадает с M, и окружность S пересекает Σ ортогонально в первом случае и диаметрально — во втором. Можно заменить это сложное определение следующим, сходным с ним: $\widetilde{\mathrm{Ст}}_S M=[r(\Sigma)]^2$, *где центр* Σ *совпадает с* M *и* Σ *пересекает* S *диаметрально.*

Величину $\widetilde{\mathrm{Ст}}_S M=d^2+r^2$ назовем *псевдостепенью* точки M относительно окружности S.

Нетрудно убедиться, что: а) геометрическое место (г. м.) точек M таких, что $\widetilde{\mathrm{Ст}}_S M=\mathrm{const}$ (где S — задана) есть окружность концентрическая с S; б) г. м. точек M таких, что $\widetilde{\mathrm{Ст}}_{S_1} M-\widetilde{\mathrm{Ст}}_{S_2} M=\mathrm{const}$ (где S_1 и S_2 заданы), есть прямая, перпендикулярная к линии центров S_1 и S_2; в) г. м. точек M таких, что $\widetilde{\mathrm{Ст}}_{S_1} M$: $\widetilde{\mathrm{Ст}}_{S_2} M=\mathrm{const}$, есть окружность с центром на O_1O_2. Все эти утверждения полностью совпадают с соответствующими теоремами о $\mathrm{Ст}_S M$.

Наиболее интересным из этих г.м. является г.м. точек M таких, что $\widetilde{\mathrm{Ст}}_{S_1} M=\widetilde{\mathrm{Ст}}_{S_2} M$. В силу б) это есть прямая, перпендикулярная к линии центров O_1O_2 окружностей S_1 и S_2; ее мы назовем *псевдорадикальной осью* двух окружностей. Нетрудно убедиться, что псевдорадикальная ось двух окружностей симметрична их радикальной оси относительно середины отрезка O_1O_2 [1]) и что

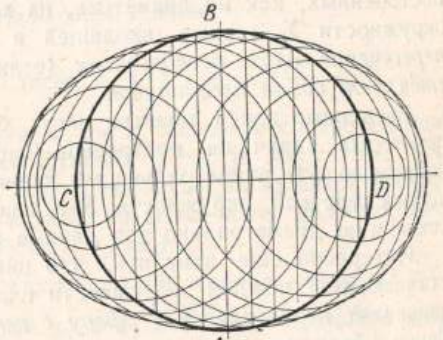


Рис. 2.

попарные псевдорадикальные оси трех окружностей или параллельны, или пересекаются в одной точке; эту точку назовем *псевдорадикальным центром* трех окружностей. Можно доказать, что *псевдорадикальный центр трех окружностей симметричен их радикальному центру относительно центра окружности, описанной вокруг треугольника, образованного центрами данных окружностей.*

Совокупность окружностей, имеющих попарно одну и ту же псевдорадикальную ось, назовем *псевдопучком*. Нетрудно убедиться, что псевдопучок — это совокупность окружностей, построенных на хордах некоторой окружности Σ (также входящей в псевдопучок), параллельных ее определенному диаметру AB (рис. 2); огибающей этих окружностей является эллипс, одна ось которого совпадает с AB, а другая равна $AB\sqrt{2}$.

Если некоторая окружность пересекает диаметрально две окружности псевдопучка, то ясно, что она диаметрально пересекает любую его окружность. Совокупность окружностей, пересекающих диаметрально все окружности псевдопучка, представляет собой (гиперболи-

[1]) Ср. И. М. Яглом, Геометрические преобразования II, М., 1956, стр. 227.

ческий) пучок окружностей, пересекающих окружность Σ псевдопучка в концах перпендикулярного AB диаметра (пучок окружностей, проходящих через точки C и D на рис. 2). Этот пучок назовем *сопряженным* исходному псевдопучку.

Псевдосвязку окружностей определим как совокупность окружностей, каждые три из которых имеют один и тот же псевдорадикальный центр (или как совокупность окружностей, относительно которых фиксированная точка O имеет одинаковую псевдостепень). Псевдосвязка — это совокупность всевозможных окружностей, построенных, как на диаметрах, на хордах некоторой фиксированной окружности Σ (также входящей в псевдосвязку). Очевидно, что *пересечение двух псевдосвязок* (если только оно существует) *представляет собой псевдопучок*.

Введенные здесь понятия могут быть использованы при решении некоторых задач на построение (например: «построить окружность, пересекающую диаметрально две данные окружности S_1 и S_2 и касающуюся третьей окружности S», или «построить окружность, пересекающую диаметрально три данные окружности»).

Интересно еще заметить, что циклографическая проекция, сопоставляющая каждой окружности плоскости точку пространства [1]), переводит псевдосвязку *в сферу с центром на основной плоскости*. Таким образом, псевдосвязки являются простейшим частным случаем так называемых *сферических систем* геометрии Лагерра [2]); они — единственные сферические системы, которые можно рассматривать, не вводя понятия о направленных окружностях.

4. ***А. В. Кужель*** (*Умань*). **Исследование корней уравнения $x^n+px^k+q=0$.** Рассмотрим уравнение

$$f(x)=x^n+px^k+q=0, \tag{1}$$

где n, k $(n>k)$ — натуральные нечетные, а p, q — вещественные не равные нулю числа, причем $p<0$. (Из дальнейшего видно, что случай $p\geqslant 0$ тривиален.) Положим

$$D=\left(\frac{q}{n-k}\right)^2-k^{\frac{2k}{n-k}}\left(\frac{|p|}{n}\right)^{\frac{2n}{n-k}}. \tag{2}$$

Тогда, как будет показано, относительно корней уравнения (1) имеют место следующие утверждения:

1°. *Если* $D>0$, то (1) *имеет один вещественный и* $n-1$ *невещественных корней.*

[1]) См. И. М. Яглом, Геометрические преобразования II, М., 1956, стр. 315.

[2]) W. Blaschke, Vorlesungen über Differentialgeometrie III, Berlin, 1929, стр. 152 и след.

2°. *Если* $D<0$, *то* (1) *имеет три различных вещественных корня и* $n-3$ *невещественных корня.*

3°. *Если* $D=0$, *то* (1) *имеет три вещественных корня, среди которых два равных; остальные* $n-3$ *корня — невещественные.*

Для уравнения $x^3+px+q=0$ $D=\left(\frac{q}{2}\right)^2+\left(\frac{p}{3}\right)^3$. Следовательно, в этом случае наши утверждения переходят в известные признаки, вытекающие из формулы Кардано.

Доказательство. Так как $f'(x)=nx^{n-1}+pkx^{k-1}$, то функция $f(x)$ имеет следующие экстремальные точки:

$$x_1=-\left(\frac{|p|k}{n}\right)^{\frac{1}{l}},\quad x_2=0 \text{ (если } k>1),\quad x_3=\left(\frac{|p|k}{n}\right)^{\frac{1}{l}},$$

где $l=n-k$. При этом точки x_1 и x_3 являются соответственно точками максимума и минимума функции $f(x)$ [в силу того, что $f(\pm\infty)=\pm\infty$]. Если $k=1$, то точка нуль не является экстремальной. Если же $k>1$ (т. е. $k\geqslant 3$), то точка нуль является точкой перегиба. Поэтому $f(x_1)>f(x_2)>f(x_3)$.

Положим теперь $D=\frac{1}{(n-k)^3}f(x_1)f(x_3)$. В случае, если $D>0$, значения функции $f(x)$ в точках x_1 и x_3 лежат одновременно либо в верхней, либо в нижней полуплоскостях. Следовательно, в этом случае график функции $f(x)$ может лишь в одной точке пересекать ось абсцисс. Так как эта точка не может совпадать с x_1, x_2 или x_3, то в этом случае уравнение (1) имеет один вещественный и $n-1$ невещественных корней. Аналогично рассматриваются случаи $D<0$ и $D=0$. Таким образом, для окончательного доказательства наших утверждений остается показать, что число D может быть записано в виде (2).

Так как $f(x)=x^k(x^l+p)+q$ и $x_1=-x_2$, то

$$f(x_1)=x_3^k|p|\frac{n-k}{n}+q,\quad f(x_3)=-x^k|p|\frac{n-k}{n}+q.$$

Отсюда

$$D=\frac{1}{(n-k)^2}f(x_1)f(x_2)=\left(\frac{q}{n-k}\right)^2-\left(\frac{|p|}{n}\right)^2\left(\frac{|p|k}{n}\right)^{\frac{2k}{l}}$$

или

$$D=\left(\frac{q}{n-k}\right)^2-k^{\frac{2k}{n-k}}\left(\frac{|p|}{n}\right)^{\frac{2k}{n-k}}.$$

Этим утверждения 1° — 3° доказаны.

5. ***З. А. Скопец*** (*Ярославль*). **Проективное обобщение теоремы Симсона.** Хорошо известная теорема элементарной геометрии, согласно которой *основания перпендикуляров, опущенных на*

стороны треугольника из любой точки описанной около него окружности, расположены на одной прямой — прямой Симсона (Robert Simson, 1687—1765), допускает следующее обобщение.

Пусть в кривую второго порядка Σ вписан треугольник ABC, стороны которого пересечены прямой g_1 в точках A_1, B_1, C_1 (рис. 1). Кроме того, на кривой даны две произвольные точки U и V, не совпадающие одна с другой и с вершинами треугольника. Из точки U проектируем точки A_1, B_1, C_1 на кривую Σ, затем полученные

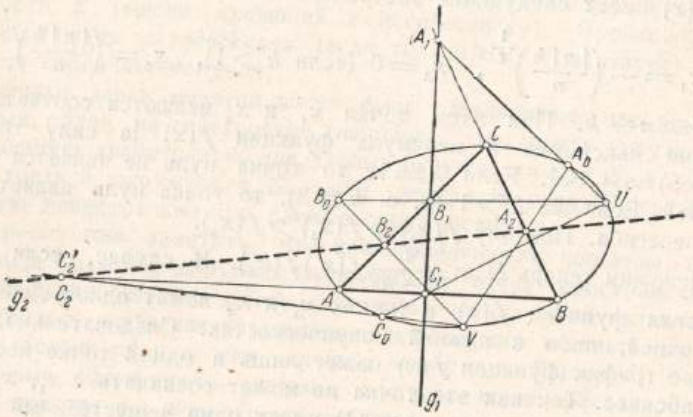


Рис. 1.

точки A_0, B_0, C_0 проектируем из точки V обратно на соответствующие стороны треугольника. Мы утверждаем, что в результате придем к точкам A_2, B_2, C_2, *также принадлежащим одной прямой* g_2.

Для того чтобы получить из этой теоремы теорему Симсона, достаточно считать, что кривая Σ есть окружность, прямая g_1 является несобственной, а точки U и V диаметрально противоположные; при этом мы получим, что *перпендикуляры, опущенные на стороны треугольника* из *точки* V, *лежат на одной прямой* (рис. 2). Если же Σ есть по-прежнему окружность, а g_1 — несобственная прямая, но точки U и V — не диаметрально противоположные, то получаем известное усиление теоремы Симсона для случая, когда из точки описанной около треугольника окружности к его сторонам проведены не перпендикуляры, а прямые одинакового наклона. И в этом случае *основания наклонных также лежат на одной прямой*, если считать углы наклонных со сторонами треугольника ориентированными, для чего надо надлежащим образом ориентировать стороны треугольника (предложенное ниже построение освобождает от рассмотрения ориентированных углов).

Для доказательства нашей теоремы достаточно заметить, что стороны треугольника и прямая g_1 образуют полный четырехсторонник, противоположные вершины которого, согласно второй теореме Дезарга [1]), проектируются из любой точки плоскости, отличной от его вершин, парами прямых, принадлежащих одной инволюции. Если за центр проектирования принять точку U, то на данной кривой получим три пары точек (A, A_0), (B, B_0), (C, C_0), также принадлежащих одной инволюции.

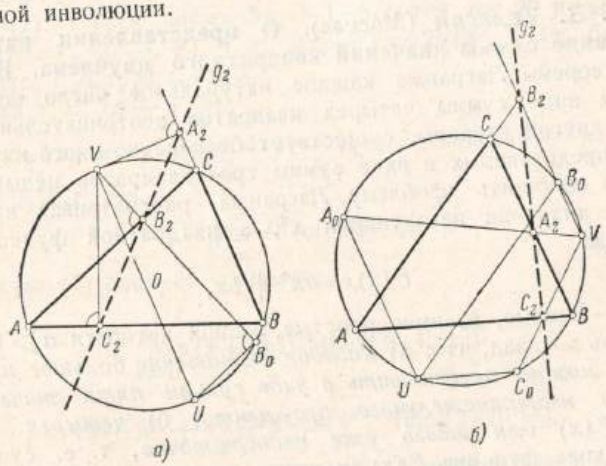


Рис. 2.

Пусть прямая A_2B_2 пересекает сторону AB в точке C_2'. Рассмотрим второй полный четырехсторонник, образованный сторонами данного треугольника и прямой A_2B_2. Проектируя его противоположные вершины на кривую из точки V, получим на кривой опять три пары точек, принадлежащих одной инволюции. Однако в полученных на кривой двух инволюциях имеются две общие пары (A, A_0), (B, B_0). Следовательно, обе инволюции совпадают, и поэтому третьи пары точек (C, C_0) и (C, C_0') должны также совпасть. Но это означает, что прямые VC_0 и VC_0' проектируют одну и ту же точку C_2, и поэтому точки C_2 и C_2' совпадают и три точки A_2, B_2, C_2 принадлежат одной прямой g_2 [2]).

Доказанная теорема остается справедливой и для распавшейся кривой второго порядка Σ, представляющей собой пару прямых. В этом случае точки U, V и одна вершина треугольника лежат на одной из двух прямых, а две другие вершины треугольника — на другой прямой.

[1]) См., например, Н. А. Глаголев, Проективная геометрия, М.-Л., 1936, стр. 118.

[2]) Если, в частности прямая g_1 проходит через U, то g_2 проходит через V.

Заметим, что соответствие прямых g_1 и g_2 является проективным. В этом проективном соответствии точке U соответствует точка V и стороны треугольника являются двойными прямыми преобразования. Если же кривая второго порядка распадается на две прямые, то проективное соответствие становится гомологией. Сторона треугольника, принадлежащая кривой, является осью гомологии, а противоположная вершина — центром гомологии.

6. *Ю. З. Телесин* (*Москва*). **О представлении натуральных чисел в виде суммы значений квадратного двучлена.** В силу известной теоремы Лагранжа каждое натуральное число можно представить в виде суммы четырех квадратов неотрицательных целых чисел. С другой стороны, существует бесконечно много натуральных чисел, непредставимых в виде суммы трех квадратов целых чисел [1]).

Можно обобщить проблему Лагранжа, рассматривая в качестве слагаемых значения не функции x^2, а квадратной функции более общего вида

$$F(x)=ax^2+bx,$$

где a и b — целые, взаимно простые, разной четности $a>0$ [2]).

Д. Паль доказал, что: а) *каждое достаточно большое натуральное число можно представить в виде суммы пяти значений $F(x)$ от целого неотрицательного аргумента*; б) *четырех значений функции $F(x)$ для этого уже недостаточно*, т. е. существуют такие «плохие» функции $F(x)$ (к числу которых, очевидно, не принадлежит функция x^2), что можно найти бесконечно много натуральных чисел, непредставимых в виде суммы четырех значений $F(x)$ от целого неотрицательного аргумента.

В сообщении о результате Паля [3]) приведено два примера «плохих» многочленов: $3x^2+2x$ и $3x^2+4x$. Первый многочлен представляет самостоятельный интерес.

Пусть $f(x)=3x^2+2x$. Докажем, что *существует бесконечно много натуральных чисел N, для которых уравнение*

$$N=f(x_1)+f(x_2)+f(x_3)+f(x_4) \tag{1}$$

неразрешимо в целых неотрицательных числах x_i $(i=1, 2, 3, 4)$ и даже не имеет целых решений $x_i>-k$, где k — фиксированное, сколь угодно большое положительное число [4]).

[1]) См., например, И. В. Арнольд, Теория чисел, § 32, М., 1939.

[2]) Если не налагать этих ограничений, то найдется бесконечно много натуральных чисел, непредставимых в виде суммы n значений от целочисленного аргумента, каково бы ни было целое n.

[3]) См. Bull. Am. math. soc. 36, стр. 862.

[4]) Иными словами, каково бы ни было k, найдется бесконечно много натуральных чисел N, которые непредставимы в виде (1) при $x_i>-k$.

Лемма. *Пусть $\varphi(x)=(3x+1)^2$. Если x_i $(i=1, 2, 3, 4)$ целое и $x_i>-\frac{2^{2n_0-1}+1}{3}$, то равенство*

$$\sum_{i=1}^{4}\varphi(x_i)=16^n \tag{2}$$

не имеет места ни при каком $n\geqslant n_0$.

Доказательство. Допустим, что при $n\geqslant n_0$ имеет место равенство (2). Пусть число $3x_i+1$ делится на 2^{t_i} и не делится на 2^{t_i+1} $(i=1, 2, 3, 4)$, т. е. $3x_i+1=2^{t_i}(2z_i+1)$; будем еще считать, что $t_1\leqslant t_2\leqslant t_3\leqslant t_4$.

1°. Если $t_1\leqslant 2n-2$, то

$$\varphi(x_i)=4^{t_i}[4z_i(z_i+1)+1]\equiv 4^{t_i}\pmod{2\cdot 4^{t_i+1}},$$

а так как $t_i\leqslant t_1$, то

$$\varphi(x_i)\equiv 4^{t_i}\pmod{2\cdot 4^{t_1+1}}.$$

Из равенства (2) следует сравнение

$$\sum_{i=1}^{4}\varphi(x_i)\equiv 16^n\pmod{2\cdot 4^{t_1+1}}$$

или

$$4^{t_1}+4^{t_2}+4^{t_3}+4^{t_4}\equiv 4^{2n}\pmod{2\cdot 4^{t_1+1}}.$$

Деля обе части последнего сравнения и модуль на 4^{t_1}, получаем в силу $t_1\leqslant 2n-2$,

$$1+4^{t_2-t_1}+4^{t_3-t_1}+4^{t_4-t_1}\equiv 0\pmod 8,$$

что невозможно.

2°. Если $t_1\geqslant 2n-1$, то $3x_1+1\equiv 0\pmod{2^{2n-1}}$, т. е.

$$3x_1+1=2^{2n-1}z. \tag{3}$$

Из (3) следует, что $z\equiv 2\pmod 3$ [1]).

а) Если $z\geqslant 2$, то из (3) вытекает

$$\varphi(x_1)=4^{2n-1}z^2\geqslant 4^{2n}=16^n.$$

Так как для всех целых x $(3x+1)^2>0$, то слагаемые $\varphi(x_2)$, $\varphi(x_3)$ и $\varphi(x_4)$ положительны и $\sum_{i=1}^{4}\varphi(x_i)>16^n$, что противоречит (2).

[1]) Последнее сравнение следует из (3) и из очевидного сравнения $2^{2n}-1\equiv 0\pmod 3$. В самом деле, из (3) имеем

$$z=\frac{3x_1+1}{2^{2n-1}}=\frac{3x_1+1+2^{2n}-2^{2n}}{2^{2n-1}}=2+\frac{3x_1-(2^{2n}-1)}{2^{2n-1}}=2+3K.$$

б) Если $z \leqslant -1$, то из (3) следует $3x_1 + 1 \leqslant -2^{2n-1} \leqslant -2^{2n_0-1}$ (так как $n \geqslant n_0$) и $x_1 \leqslant -\frac{2^{2n_0}+1}{3}$, что противоречит условию леммы.

Теорема. *Все натуральные числа N вида $\frac{16^n-4}{3}$, где $n \geqslant n_0$, непредставимы в виде*

$$N = \sum_{i=1}^{4} f(x_i), \tag{4}$$

где $x_i > -\frac{2^{2n_0+1}}{3}$.

Доказательство. Учитывая, что $N = \frac{16^n-4}{3}$, умножая обе части равенства (4) на 3 и прибавляя затем к обеим частям 4, мы получим равенство (2), невозможность которого доказана выше.

В заключение уместно указать на следующие интересные задачи:

1) Только ли для чисел $\frac{16^n-4}{3}$ уравнение (4) не имеет целых решений для бесконечного количества этих чисел?

2) Как описать все квадратные многочлены, обладающие тем же свойством, что и многочлен $f(x) = 3x^2 + 2x$?

3) Изменится ли результат, если область допустимых значений целочисленного аргумента x не ограничивать снизу?

III. НАУЧНО-МЕТОДИЧЕСКИЕ СООБЩЕНИЯ

(Опыт преподавания и педагогический эксперимент)

О СРЕДНИХ

И. И. Жогин

(Шадринск)

Во многих вопросах математики важную роль играет теорема о (взвешенных) степенных средних произвольного порядка. Теорема эта утверждает, что *непрерывная функция*

$$y = f(x) = \begin{cases} \left(\dfrac{p_1 a_1^x + \ldots + p_n a_n^x}{p_1 + \ldots + p_n}\right)^{\frac{1}{x}} & \text{при } x \neq 0, \\ (a_1^{p_1} \ldots a_n^{p_n})^{\frac{1}{p_1 + \ldots + p_n}} & \text{при } x = 0 \end{cases} \tag{1}$$

монотонно возрастает и ее значения заключены между числами $\underline{a} = \min(a_1, \ldots, a_n)$ *и* $\overline{a} = \max(a_1, \ldots, a_n)$:

$$\underline{a} < f(x) < \overline{a};$$

здесь $a_1, \ldots, a_n$ и $p_1, \ldots, p_n$ — две системы положительных чисел, причем в первой из систем по крайней мере два числа не равны между собой.

Функция $f(x)$ называется *взвешенным степенным средним порядка x чисел $a_1, \ldots, a_n$ с весами $p_1, \ldots, p_n$*.

В качестве следствий этой теоремы можно получить многие важные неравенства анализа, в частности неравенства Коши, Гёльдера, Минковского и др.

Так, например, полагая $x_1 = 0$, $x_2 = 1$ и учитывая, что

$$f(x_1) < f(x_2) \text{ (при } x_1 < x_2), \tag{2}$$

получаем теорему о среднем арифметическом и среднем геометрическом:

$$(a_1^{p_1} \ldots a_n^{p_n})^{\frac{1}{p_1 + \ldots + p_n}} < \frac{p_1 a_1 + \ldots + p_n a_n}{p_1 + \ldots + p_n}. \tag{2'}$$

В настоящем очерке мы поставили себе целью дать элементарное изложение основных понятий, относящихся к учению о средних. По нашему мнению, опыт такого изложения мог бы быть использован преподавателями педагогических институтов.

Доказательство основной теоремы начнем с рассмотрения следующей функции от t $(t>0)$:

$$\varphi(t)=(1+t)^{\alpha}\left(1+\frac{1}{t}\right)^{\beta}=\frac{(1+t)^{\alpha+\beta}}{t^{\beta}} \tag{3}$$

(α и β — произвольные положительные параметры).

Дифференцируя $\varphi(t)$, получим:

$$\varphi'(t)=\frac{\alpha(1+t)^{\alpha+\beta-1}}{t^{\beta+1}}\left(t-\frac{\beta}{\alpha}\right).$$

Легко видеть, что функция $\varphi(t)$ достигает минимума при $t=\frac{\beta}{\alpha}$:

$$\varphi\left(\frac{\beta}{\alpha}\right)\leqslant\varphi(t). \tag{4}$$

Знак равенства появляется лишь при $t=\frac{\beta}{\alpha}$. Полагая в (4) $t=\frac{y}{x}$ (x, y — произвольные положительные числа), получим

$$\left(\frac{\alpha+\beta}{x+y}\right)^{\alpha+\beta}\leqslant\left(\frac{\alpha}{x}\right)^{\alpha}\left(\frac{\beta}{y}\right)^{\beta};$$

равенство будет иметь место лишь при $\frac{\alpha}{x}=\frac{\beta}{y}$. Заменяя здесь β на $\beta+\gamma$, y на $y+z$, получим

$$\left(\frac{\alpha+\beta+\gamma}{x+y+z}\right)^{\alpha+\beta+\gamma}\leqslant\left(\frac{\alpha}{x}\right)^{\alpha}\left(\frac{\beta+\gamma}{y+z}\right)^{\beta+\gamma}\leqslant\left(\frac{\alpha}{x}\right)^{\alpha}\left(\frac{\beta}{y}\right)^{\beta}\left(\frac{\gamma}{z}\right)^{\gamma}.$$

Наконец, индукцией легко получить, что каковы бы ни были положительные числа

$$\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n,$$
$$\beta_1,\ \beta_2,\ \ldots,\ \beta_n,$$

справедливо следующее неравенство:

$$\left(\frac{\alpha_1+\ldots+\alpha_n}{\beta_1+\ldots+\beta_n}\right)^{\alpha_1+\ldots+\alpha_n}\leqslant\left(\frac{\alpha_1}{\beta_1}\right)^{\alpha_1}\left(\frac{\alpha_2}{\beta_2}\right)^{\alpha_2}\ldots\left(\frac{\alpha_n}{\beta_n}\right)^{\alpha_n}. \tag{5}$$

Равенство имеет место лишь в случае $\frac{\alpha_1}{\beta_1}=\ldots=\frac{\alpha_n}{\beta_n}$. Неравенство (5) можно представить и в другой, в ряде приложений более удобной форме:

$$\left(\frac{\beta_1+\ldots+\beta_n}{\alpha_1+\ldots+\alpha_n}\right)^{\alpha_1+\ldots+\alpha_n}\geqslant\left(\frac{\beta_1}{\alpha_1}\right)^{\alpha_1}\left(\frac{\beta_2}{\alpha_2}\right)^{\alpha_2}\ldots\left(\frac{\beta_n}{\alpha_n}\right)^{\alpha_n}. \tag{5'}$$

Подставляя в (5) вместо α_i и β_i p_i и p_ia_i соответственно, получим обычную, наиболее распространенную форму теоремы о среднем арифметическом и среднем геометрическом:

$$a_1^{p_1}\ldots a_n^{p_n}\leqslant\left(\frac{p_1a_1+\ldots+p_na_n}{p_1+\ldots+p_n}\right)^{p_1+\ldots+p_n}; \tag{6}$$

если в (6) все $p_i=1$ или $p_1+\ldots+p_n=1$ (в последнем случае вместо p пишут q), то соответственно

$$a_1\ldots a_n\leqslant\left(\frac{a_1+\ldots+a_n}{n}\right)^n, \tag{7}$$

$$a_1^{q_1}a_2^{q_2}\ldots a_n^{q_n}\leqslant q_1a_1+\ldots+q_na_n. \tag{8}$$

Во всех трех формулах (6), (7), (8) равенство возможно лишь в случае, когда все a_i одинаковы.

Отметим здесь еще неравенство

$$\sum_{\nu=1}^{n}a_\nu^{\alpha}b_\nu^{\beta}\ldots l_\nu^{\lambda}\leqslant\left(\sum_1^n a_\nu\right)^{\alpha}\left(\sum_1^n b_\nu\right)^{\beta}\ldots\left(\sum_1^n l_\nu\right)^{\lambda}, \tag{9}$$

верное для произвольных точек $A(a_\nu)$, $B(b_\nu)$, ..., $L(l_\nu)$ n-мерного пространства с положительными координатами в предположении, что сумма положительных α, β, ..., λ равна 1. Это неравенство легко получить из (8):

$$\frac{\sum a_\nu^{\alpha}b_\nu^{\beta}\ldots l_\nu^{\lambda}}{(\sum a_\nu)^{\alpha}(\sum b_\nu)^{\beta}\ldots(\sum l_\nu)^{\lambda}}=\sum\left(\frac{a_\nu}{\sum a_\nu}\right)^{\alpha}\left(\frac{b_\nu}{\sum b_\nu}\right)^{\beta}\ldots\left(\frac{l_\nu}{\sum l_\nu}\right)^{\lambda}\leqslant$$
$$\leqslant\sum\left(\alpha\frac{a_\nu}{\sum a_\nu}+\ldots+\lambda\frac{l_\nu}{\sum l_\nu}\right)=\alpha\sum\frac{a_\nu}{\sum a_\nu}+\ldots+\lambda\sum\frac{l_\nu}{\sum l_\nu}=$$
$$=\alpha+\ldots+\lambda=1.$$

Для двух точек $A(a_\nu^k)$ и $B(b_\nu^{k'})$, где $k>1$, $k'=\frac{k}{k-1}$ (т. е. $\frac{1}{k}+\frac{1}{k'}=1$), $\alpha=\frac{1}{k}$, $\beta=\frac{1}{k'}$, из (9) получаем теорему Гёльдера:

$$\sum_\nu a_\nu b_\nu\leqslant\left(\sum_\nu a_\nu^k\right)^{\frac{1}{k}}\left(\sum_\nu b_\nu^{k'}\right)^{\frac{1}{k'}} \tag{9'}$$

$\left(\text{равенство при }\frac{a_1^k}{b_1^k}=\ldots=\frac{a_n^k}{b_n^{k'}}\right)$, а из нее при $k=k'=2$ — неравенство Коши

$$\left(\sum_\nu a_\nu b_\nu\right)^2\leqslant\left(\sum_\nu a_\nu^2\right)\left(\sum_\nu b_\nu^2\right) \tag{10}$$

$\left(\text{равенство при }\frac{a_1}{b_1}=\ldots=\frac{a_n}{b_n}\right)$.

Покажем теперь, каким образом с помощью (5) можно доказать общую теорему о средних степенных, т. е. тот факт, что функция (1) при $-\infty < x < +\infty$ возрастает между $\underline{a}$ и $\bar{a}$. При этом, конечно, мы предполагаем, что не все (положительные) координаты точки $A(a_1, \ldots, a_n)$ одинаковы, так как иначе функция $f(x)$ сохранила бы постоянное значение a.

Убедимся, что $f(x)$ возрастает в каждой точке.

При $x \neq 0$ имеем:

$$\frac{f'(x)}{f(x)} = -\frac{1}{x^2}\ln\frac{p_1a_1^x+\ldots+p_na_n^x}{p_1+\ldots+p_n} + \frac{1}{x}\frac{p_1a_1^x\ln a_1+\ldots+p_na_n^x\ln a_n}{p_1a_1^x+\ldots+p_na_n^x};$$

$$x^2\frac{f'(x)}{f(x)} = -\ln\frac{p_1a_1^x+\ldots+p_na_n^x}{p_1+\ldots+p_n} + \frac{p_1a_1^x\ln a_1^x+\ldots+p_na_n^x\ln a_n^x}{p_1a_1^x+\ldots+p_na_n^x};$$

$$x^2(p_1a_1^x+\ldots+p_na_n^x)\frac{f'(x)}{f(x)} = -\ln\left(\frac{p_1a_1^x+\ldots+p_na_n^x}{p_1+\ldots+p_n}\right)^{p_1a_1^x+\ldots+p_na_n^x} +$$

$$+\ln\left((a_1^x)^{p_1a_1^x}\ldots(a_n^x)^{p_na_n^x}\right) = \ln\frac{(a_1^x)^{p_1a_1^x}\ldots(a_n^x)^{p_na_n^x}}{\left(\frac{p_1a_1^x+\ldots+p_na_n^x}{p_1+\ldots+p_n}\right)^{p_1a_1^x+\ldots+p_na_n^x}}.$$

Здесь под знаком логарифма, ввиду (5), при любом $x \neq 0$ стоит число, большее 1. Следовательно,

$$x^2(p_1a_1^x+\ldots+p_na_n^x)\frac{f'(x)}{f(x)} > 0, \text{ т. е. } f'(x) > 0.$$

Производная в точке $x=0$ также существует:

$$f'(0) = \lim_{h\to 0}\frac{1}{h}\left[\left(\frac{p_1a_1^h+\ldots+p_na_n^h}{p_1+\ldots+p_n}\right)^{\frac{1}{h}} - \left(a_1^{p_1}\ldots a_n^{p_n}\right)^{\frac{1}{p_1+\ldots+p_n}}\right] =$$

$$= \frac{1}{2}\left(a_1^{p_1}\ldots a_n^{p_n}\right)^{\frac{1}{p_1+\ldots+p_n}} \times$$

$$\times\left[\frac{(p_1\ln^2 a_1+\ldots+p_n\ln^2 a_n)(p_1+\ldots+p_n)-(p_1\ln a_1+\ldots+p_n\ln a_n)^2}{(p_1+\ldots+p_n)^2}\right];$$

ввиду неравенства Коши (10) $f'(0) > 0$.

Таким образом, функция $f(x)$ всюду дифференцируема и имеет положительную производную. То, что $\lim\limits_{x\to\infty} f(x) = \bar{a}$, $\lim\limits_{x\to-\infty} f(x) = \underline{a}$, где $\underline{a}$ и $\bar{a}$ обозначают соответственно наименьшее и наибольшее из чисел $a_1, \ldots, a_n$, очевидно. Теорема доказана.

Мы составили довольно ясное представление о течении кривой $y = f(x)$: она возрастает между своими асимптотами $y = \underline{a}$ и $y = \bar{a}$. Мы можем поэтому неравенство (2) дополнить следующей оценкой ($x_2 - x_1 > 0$):

$$0 < f(x_2) - f(x_1) < \bar{a} - \underline{a}. \tag{11}$$

Переходим к примерам.

1. Пусть $a_1, a_2, \ldots, a_n$ неотрицательны.

Имеем из (5)

$$\left(\frac{n}{n+\sum_1^n a_k}\right)^n = \left(\frac{1+\ldots+1}{(1+a_1)+\ldots+(1+a_n)}\right)^{1+\ldots+1} <$$

$$< \frac{1}{1+a_1}\frac{1}{1+a_2}\cdots\frac{1}{1+a_n},$$

откуда

$$\prod_{k=1}^n (1+a_k) < \left(1+\frac{\sum_1^n a_k}{n}\right)^n. \tag{*}$$

Точно так же, в предположении $1-a_k > 0$, получим

$$\frac{1}{\left(1+\frac{1}{n}\sum_1^n\frac{a_k}{1-a_k}\right)^n} = \left(\frac{n}{n+\sum_1^n\frac{a_k}{1-a_k}}\right)^n =$$

$$= \left(\frac{1+\ldots+1}{\left(1+\frac{a_1}{1-a_1}\right)+\ldots+\left(1+\frac{a_n}{1-a_n}\right)}\right)^{1+\ldots+1} < \prod_{k=1}^n (1-a_k).$$

Отсюда следует, что если ряд $\sum\limits_1^\infty a_k = S$ сходится (а вместе с ним тогда сходится и ряд $\sum\limits_1^\infty\frac{a_k}{1-a_k} = S'$ ввиду того, что, например, $\frac{a_k}{1-a_k} < 2a_k$ начиная с некоторого k), то возрастающая последовательность $\prod\limits_1^n(1+a_k)$ ограничена сверху числом e^S, а убывающая последовательность $\prod\limits_1^n(1-a_k)$ ограничена снизу числом $e^{-S'}$, так что

бесконечные произведения

$$\prod_1^\infty (1+a_k) \quad \text{и} \quad \prod_1^\infty (1-a_k)$$

сходятся.

Неравенство (*) можно дополнить:

$$\left(1-\frac{\sum_1^n \frac{a_k}{a_k+1}}{n}\right)^{-n} < \prod_{k=1}^n (1+a_k) < \left(1+\frac{\sum_1^n a_k}{n}\right)^n; \qquad (**)$$

соответствующую оценку легко получить по полной аналогии с тем, что было сделано выше.

В пределе, при $n \to \infty$, получим

$$e^{\sum_1^\infty \frac{a_k}{1+a_k}} \leqslant \prod_{k=1}^\infty (1+a_k) \leqslant e^{\sum_1^\infty a_k}.$$

2. Имеем

$$\left(1+\frac{1}{n}\right)^{n+1} = \left(\frac{n+1}{n}\right)^{n+1} = \left(\frac{2+1+\ldots+1}{1+1+\ldots+1}\right)^{2+1+\ldots+1} < 2^2 = 4.$$

Можно, конечно, более удобным разбиением $n+1$ на сумму n слагаемых (или разбиением n на сумму $n+1$ слагаемых) добиться понижения этой границы. Например:

$$\left(\frac{n+1}{n}\right)^{n+1} = \left(\frac{\frac{3}{2}+\frac{3}{2}+1+\ldots+1}{1+1+1+\ldots+1}\right)^{\frac{3}{2}+\frac{3}{2}+1+\ldots+1} <$$

$$< \left(\frac{3}{2}\right)^{\frac{3}{2}} \left(\frac{3}{2}\right)^{\frac{3}{2}} = 3\frac{3}{8}.$$

Этот способ доказательства ограниченности знаменитой последовательности следовало бы применять в преподавании. Возрастание другой последовательности, служащей для определения e, непосредственно вытекает из (5′):

$$\left(1+\frac{1}{n+1}\right)^{n+1} = \left(\frac{n+2}{n+1}\right)^{n+1} = \left(\frac{(n+1)+1}{n+1}\right)^{n+1} >$$

$$> \left(\frac{n+1}{n}\right)^n \left(\frac{1}{1}\right)^1 = \left(1+\frac{1}{n}\right)^n.$$

3. Пусть, как и в примере 1, $a_1, a_2, \ldots, a_k, \ldots$ — последовательность положительных чисел. Пусть для всех n, начиная с некоторого n_0, выполняется условие $na_{n+1} \leqslant \sum_1^n a_k$ (оно, например, выполняется для всякой невозрастающей последовательности). Тогда имеет место следующее неравенство:

$$\prod_1^n (1+a_n) < e^{\sum_1^n a_k}$$

или

$$\sum_1^n a_k > \ln \prod_1^n (1+a_k) \qquad (n \geqslant n_0).$$

Доказательство. В первом примере мы уже получили, что

$$\prod_1^n (1+a_k) < \left(1+\frac{\sum_1^n a_k}{n}\right)^n.$$

Здесь преобразование правой части продолжим следующим образом:

$$\left(1+\frac{\sum_1^n a_k}{n}\right)^n = \left(\left(1+\frac{1}{\frac{n}{\sum_1^n a_k}}\right)^{\frac{n}{\sum_1^n a_k}}\right)^{\sum_1^n a_k}.$$

По условию $na_{n+1} \leqslant \sum_1^n a_k$, т. е. $\frac{a_{n+1}}{\sum_1^n a_k} \leqslant \frac{1}{n}$, $1+\frac{a_{n+1}}{\sum_1^n a_k} \leqslant 1+\frac{1}{n}$,

$$\frac{\sum_1^{n+1} a_k}{\sum_1^n a_k} \leqslant \frac{n+1}{n}, \quad \frac{\frac{n+1}{\sum_1^{n+1} a_k}}{} \geqslant \frac{\frac{n}{\sum_1^n a_k}}{}.$$

Таким образом, последовательность $\frac{n}{\sum_1^n a_k}$ $(n=1, 2, \ldots)$ возрастает,

а вместе с ней, как известно, возрастает и последовательность

$$\left(1+\frac{1}{\frac{n}{\sum_1^n a_k}}\right)^{\frac{n}{\sum_1^n a_k}}.$$

В случае $\frac{n}{\sum_1^n a_k} \to \infty$ эта последовательность стремится к e, в случае же $\frac{n}{\sum_1^n a_k} \to A < \infty$ она стремится к $\left(1+\frac{1}{A}\right)^A < e$. В обоих случаях имеем

$$\prod_1^n (1+a_k) < e^{\sum_1^n a_k}.$$

Для примера возьмем $a_n = \frac{1}{n}$; условие $na_{n+1} < \sum_1^n a_k$ выполняется (начиная с $n_0 = 1$), поэтому для любого натурального n

$$\prod_1^n \left(1+\frac{1}{k}\right) = n+1 < e^{H_n}, \quad H_n > \ln(n+1)$$
$$\left(H_n = 1+\frac{1}{2}+\ldots+\frac{1}{n}\right),$$

что свидетельствует о *расходимости гармонического ряда*. Для последовательности чисел $\frac{1}{p_1}, \frac{1}{p_2}, \ldots, \frac{1}{p_n}, \ldots$, обратных простым, условие $\frac{n}{p_{n+1}} < \sum_1^n \frac{1}{p_k}$ также выполняется начиная уже с $n=1$, так что

$$\sum_1^n \frac{1}{p_k} > \ln \prod_1^n \left(1+\frac{1}{p_k}\right).$$

Так как $\left(\text{ввиду } 1+\frac{1}{p_k} > \frac{1}{2}\frac{1}{1-\frac{1}{p_k}},\ k=1, 2, \ldots\right)$

$$\prod_1^n \left(1+\frac{1}{p_k}\right) > \frac{1}{2}\prod_1^n \frac{1}{1-\frac{1}{p_k}} = \frac{1}{2}\prod_1^n \left(1+\frac{1}{p_k}+\frac{1}{p_k^2}+\ldots\right) >$$
$$> \frac{1}{2}\sum_1^n \frac{1}{k} = \frac{H_n}{2},$$

то

$$\sum_1^n \frac{1}{p_k} > \ln \frac{H_n}{2} > \ln \frac{\ln(n+1)}{2},$$

что свидетельствует о *расходимости ряда чисел, обратных простым.*

4. Возьмем [1]) для $0 < \alpha < 1$ и $x > -1$ точки плоскости

$$A(1-\alpha, \alpha), \quad B\left(\frac{1}{\alpha}-1, 1+x\right)$$

и подставим в (5′) их координаты (все они положительны):

$$\frac{1}{\alpha}+x \geqslant \left(\frac{\frac{1}{\alpha}-1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1+x}{\alpha}\right)^{\alpha} = \frac{1}{\alpha}(1+x)^{\alpha},$$

т. е.

$$(1+x)^{\alpha} \leqslant 1+\alpha x;$$

равенство возможно лишь при

$$\frac{1-\alpha}{\frac{1}{\alpha}-1} = \frac{\alpha}{1+x}, \quad \text{т. е. при } x=0.$$

Если $\alpha > 1$, то рассмотрим точки

$$A(\alpha-1, 1), \quad B\left(1-\frac{1}{\alpha}, \frac{1}{\alpha}+x\right) \qquad \left(\frac{1}{\alpha}+x > 0\right)$$

и из (5′) получим

$$\left(\frac{1+x}{\alpha}\right)^{\alpha} \geqslant \left(\frac{1-\frac{1}{\alpha}}{\alpha-1}\right)^{\alpha-1}\left(\frac{1}{\alpha}+x\right) = \frac{1}{\alpha^{\alpha}}(1+\alpha x),$$

т. е.

$$(1+x)^{\alpha} \geqslant 1+\alpha x \qquad \left(x > -\frac{1}{\alpha}\right).$$

В случае $-1 < x \leqslant -\frac{1}{\alpha}$ неравенство очевидно. Равенство опять возможно лишь при $x=0$.

Если, наконец, α отрицательно, то в зависимости от его абсолютной величины мы применим (5′) либо к координатам точек

$$A'(1+\alpha, -\alpha), \quad B'\left(-\frac{1}{\alpha}-1, 1+x\right) \qquad (|\alpha|<1),$$

либо к координатам точек

$$A''\left(1+\frac{1}{\alpha}, -\frac{1}{\alpha}\right), \quad B''(-\alpha-1, 1+x) \qquad (|\alpha|>1).$$

[1]) Ср. П. П. Коровкин, Неравенства, М., 1956.

В первом случае

$$\left(-\frac{1}{\alpha}+x\right)>\left(\frac{-\frac{1}{\alpha}-1}{1+\alpha}\right)^{1+\alpha}\left(\frac{1+x}{-\alpha}\right)^{-\alpha}=-\frac{1}{\alpha}(1+x)^{-\alpha};$$

$$1-\alpha x>(1+x)^{-\alpha},\quad (1+x)^{\alpha}>\frac{1}{1-\alpha x}>\frac{1-\alpha^2x^2}{1-\alpha x}=1+\alpha x \qquad (x\neq 0).$$

Во втором случае

$$(-\alpha+x)>\left(\frac{\alpha-1}{1+\frac{1}{\alpha}}\right)^{1+\frac{1}{\alpha}}\left(\frac{1+x}{-\frac{1}{\alpha}}\right)^{-\frac{1}{\alpha}}=-\alpha(1+x)^{-\frac{1}{\alpha}};$$

$$(1+x)^{-\frac{1}{\alpha}}<1-\frac{x}{\alpha},\quad (1+x)^{\frac{1}{\alpha}}>\frac{1}{1-\frac{x}{\alpha}}>1+\frac{x}{\alpha} \qquad (x\neq 0);$$

возводя здесь обе части в степень α^2 и принимая во внимание предыдущий результат для показателя, большего 1, получим

$$(1+x)^{\alpha}>\left(1+\frac{x}{\alpha}\right)^{\alpha^2}>1+\alpha x.$$

Таким образом, доказано следующее:

$$(1+x)^{\alpha}\leqslant 1+\alpha x \quad (0<\alpha<1,\ x>-1);$$
$$(1+x)^{\alpha}\geqslant 1+\alpha x \quad (\alpha>1 \text{ или } \alpha<0,\ x>-1).$$

Равенства возможны лишь при $x=0$.

КРИВЫЕ ПЕАНО

И. П. Натансон

(Ленинград)

1. В 1890 г. итальянский математик Д. Пеано (1858—1932) построил пример плоской непрерывной кривой, проходящей через все точки некоторого квадрата [1]). Естественно, что кривая с такими удивительными свойствами привлекла к себе всеобщее внимание, и вскоре были найдены и другие примеры таких кривых. Все они по имени автора первой из них получили название *кривых Пеано*. В частности, сравнительно простую кривую Пеано построил в 1891 г. немецкий ученый Д. Гильберт (1862—1943) [2]). Целью настоящей статьи является изложение конструкции Гильберта [3]).

2. Поскольку существуют различные определения термина «плоская непрерывная кривая», приведем то из них, которое имеется в виду в этой статье [4]).

Определение. *Плоской непрерывной кривой* называется множество M всех точек $(\varphi(t)),\ \psi(t))$, где $\varphi(t)$ и $\psi(t)$ — какие-нибудь две (фиксированные для данной кривой) функции, заданные и непрерывные на отрезке $[0,1]$.

Если истолковать параметр t как время, то приведенное определение, грубо говоря, будет означать, что плоская непрерывная кривая есть путь движущейся по плоскости точки. Равенства

$$x=\varphi(t),\ y=\psi(t) \qquad (0\leqslant t\leqslant 1)$$

являются уравнениями этого движения.

[1]) G. Peano, Sur une courbe, qui remplit toute une aire plane, Math. Annalen 36 (1890), 157—160.

[2]) D. Hilbert, Ueber die stetige Abbildung einer Linie auf ein Flächenstück, Math. Annalen 38 (1891), 459—460.

[3]) В нашей литературе пример Гильберта уже излагался: см. Н. Н. Лузин, Теория функций действительного переменного, Учпедгиз, 1940; А. С. Пархоменко, Что такое линия, М., 1954. Однако установление некоторых обстоятельств, принятых в этих доказательствах за очевидные, может затруднить малоопытного читателя. Я здесь вхожу в несколько бо́льшие подробности.

[4]) Это определение принадлежит французскому математику К. Жордану (Marie Ennemond Camille Jordan, 1838—1922). По поводу других определений см., например, литературу, указанную в предыдущей сноске.

Таким образом, задача построения кривой Пеано будет решена, если мы укажем такую пару непрерывных на $[0, 1]$ функций $\varphi(t)$ и $\psi(t)$, для которой упомянутое выше множество M совпадает с квадратом $Q(0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1)$.

3. Для построения нужной нам пары функций разработаем предварительно некоторый способ записи чисел, принадлежащих отрезку $T = [0, 1]$.

Начнем с того, что при помощи точек $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$ разложим T на четыре равных отрезка:

$$T_1^{(1)} = \left[0, \frac{1}{4}\right],\ T_2^{(1)} = \left[\frac{1}{4}, \frac{1}{2}\right],\ T_3^{(1)} = \left[\frac{1}{2}, \frac{3}{4}\right],\ T_4^{(1)} = \left[\frac{3}{4}, 1\right],$$

которые в дальнейшем будем называть *отрезками первого ранга*. Эти отрезки изображены на рис. 1. Каждый из отрезков первого

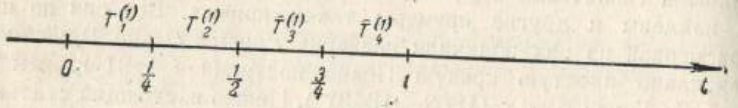


Рис. 1.

ранга в свою очередь мы разложим на четыре равных отрезка при помощи точек $\frac{i}{4^2}$ $(i = 1, 2, \ldots, 15)$ и введем для этих новых отрезков, которые будем называть *отрезками второго ранга*, обозначение

$$T_i^{(2)} = \left[\frac{i-1}{16}, \frac{i}{16}\right] \qquad (i = 1, 2, \ldots, 16).$$

Продолжая подобный процесс неограниченно, мы для каждого натурального k построим 4^k отрезков *k-го ранга*, для которых будем употреблять обозначение

$$T_i^{(k)} = \left[\frac{i-1}{4^k}, \frac{i}{4^k}\right] \qquad (i = 1, 2, \ldots, 4^k).$$

Полученная система отрезков и позволяет дать упомянутый выше способ записи чисел отрезка $[0, 1]$. Пусть $t_0 \in [0, 1]$, причем мы предположим сначала, что точка t_0 не совпадает ни с одной из введенных выше точек деления, т. е. ни для какого k

$$t_0 \neq \frac{i}{4^k} \qquad (i = 1, 2, \ldots, 4^k - 1).$$

Тогда существует только один отрезок $T_{i_1}^{(1)}$ первого ранга, содержащий точку t_0, только один отрезок $T_{i_2}^{(2)}$ второго ранга с

этим же свойством и т. д. Обозначим вообще через $T_{i_k}^{(k)}$ тот (единственный!) отрезок k-го ранга, который содержит t_0. Ясно, что

$$T_{i_1}^{(1)} \supset T_{i_2}^{(2)} \supset T_{i_3}^{(3)} \supset \ldots,$$

и потому

$$4i_k - 3 \leqslant i_{k+1} \leqslant 4i_k \qquad (k = 1, 2, 3, \ldots).$$

Условимся употреблять следующую запись:

$$t_0 = [i_1, i_2, i_3, \ldots].$$

Допустим теперь, что $t_0 \in [0, 1]$ совпадает с одной из введенных выше точек деления. Тогда

$$t_0 = \frac{m}{4^n} \qquad (1 \leqslant m \leqslant 4^n - 1),$$

причем m не делится на 4. Ясно, что имеется только по одному отрезку каждого из рангов $1, 2, \ldots, n-1$, содержащему t_0 (мы считаем $n > 1$; если $n = 1$, то рассуждение лишь упрощается). Что же касается рангов $n, n+1, n+2, \ldots$, то среди отрезков каждого из них будет по два, содержащих t_0, а именно:

$$T_m^{(n)} \text{ и } T_{m+1}^{(n)},\quad T_{4m}^{(n+1)} \text{ и } T_{4m+1}^{(n+1)}, \ldots$$

Пусть $k < n$ и $T_{i_k}^{(k)}$ есть тот единственный отрезок ранга k, который содержит [1]) t_0. Условимся употреблять для t_0 любую из записей [2]):

$$t_0 = [i_1, i_2, \ldots, i_{n-1}, m, 4m, 16m, \ldots],$$
$$t_0 = [i_1, i_2, \ldots, i_{n-1}, m+1, 4m+1, 16m+1, \ldots].$$

Таким образом, каждая из точек t_0 отрезка $[0, 1]$ допускает одну или две записи вида

$$t_0 = [i_1, i_2, i_3, \ldots],$$

где $1 \leqslant i_1 \leqslant 4$, $4i_k - 3 \leqslant i_{k+1} \leqslant 4i_k$.

Обратно, если $1 \leqslant p \leqslant 4^k$ и $4p - 3 \leqslant q \leqslant 4p$, то

$$T_p^{(k)} \supset T_q^{(k+1)},$$

и поэтому всякая бесконечная последовательность натуральных чисел $[i_1, i_2, i_3, \ldots]$, у которой $1 \leqslant i_1 \leqslant 4$, $4i_k - 3 \leqslant i_{k+1} \leqslant 4i_k$, определяет систему вложенных друг в друга отрезков

$$T_{i_1}^{(1)} \supset T_{i_2}^{(2)} \supset T_{i_3}^{(3)} \supset \ldots$$

---

[1]) Легко понять, что при $k = 1, 2, \ldots, n-2$
$$4i_k - 3 \leqslant i_{k+1} \leqslant 4i_k.$$

[2]) Пример: $\frac{5}{8} = [3, 10, 40, 160, \ldots] = [3, 11, 41, 161, \ldots]$.

и, стало быть, может рассматриваться как запись некоторой точки t_0 из $[0, 1]$.

4. Теперь мы построим нечто вроде такого же аппарата для записи точек квадрата $Q(0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1)$. Начнем с разложения Q при помощи прямых $x = \frac{1}{2}$ и $y = \frac{1}{2}$ на четыре равных квадрата, которые будем называть квадратами первого ранга и обозначать через $Q_1^{(1)}, Q_2^{(1)}, Q_3^{(1)}, Q_4^{(1)}$. Нижние индексы расставим так, как указано на рис. 2.

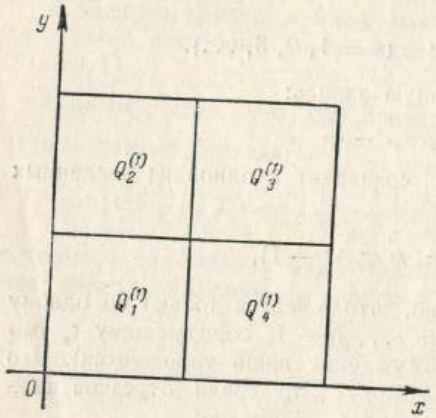


Рис. 2.

Проделав указанное построение, разложим каждый из квадратов $Q_i^{(1)}$ $(i = 1, 2, 3, 4)$ на четыре равных квадрата, которые назовем квадратами второго ранга и обозначим $Q_i^{(2)}$ $(i = 1, 2, \ldots, 16)$. Рис. 3 указывает расстановку нижних индексов i. Обратим внимание на то, что $Q_1^{(2)}$ содержит точку $O(0, 0)$ и что любая пара квадратов $Q_i^{(2)}$ и $Q_{i+1}^{(2)}$ $(i = 1, 2, \ldots, 15)$ имеет общую сторону.

Будем продолжать указанный процесс разложения построенных квадратов на четыре равных квадрата неограниченно. В результате для каждого натурального k будет получено 4^k квадратов ранга k, сторона каждого из которых равна $\frac{1}{2^k}$. Эти квадраты мы будем обозначать через $Q_i^{(k)}(i = 1, 2, \ldots, 4^k)$, озаботившись при расстановке нижних индексов i лишь тем, чтобы соблюдались три условия:

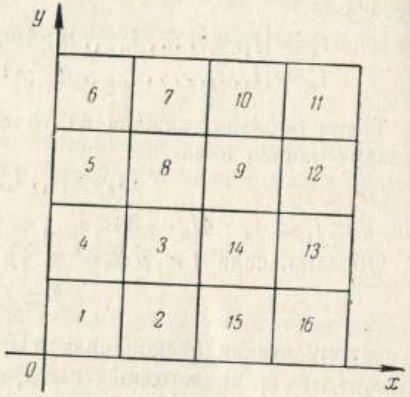


Рис. 3.

1) каждый квадрат $Q_1^{(k)}$ содержит точку $O(0, 0)$;

2) все четыре квадрата $Q_{4i-3}^{(k)}$, $Q_{4i-2}^{(k)}$, $Q_{4i-1}^{(k)}$, $Q_{4i}^{(k)}$ содержатся в $Q_i^{(k-1)}$ $(k > 1;\ i = 1, 2, \ldots, 4^{k-1})$;

3) любые два квадрата $Q_i^{(k)}$ и $Q_{i+1}^{(k)}$ $(i = 1, 2, \ldots, 4^k - 1)$ с соседними нижними индексами имеют общую сторону.

То обстоятельство, что первым двум условиям можно удовлетворять, не вызывает сомнения. Менее очевидно, что в каждой четверке квадратов $Q_{4i-3}^{(k)}$, $Q_{4i-2}^{(k)}$, $Q_{4i-1}^{(k)}$, $Q_{4i}^{(k)}$, полученной разложением одного и того же квадрата $Q_i^{(k-1)}$, можно так расставить нижние индексы, чтобы выполнялось и третье условие. Этот факт мы докажем индуктивно. Рис. 2 и 3 показывают, что нумерация квадратов первых двух рангов, удовлетворяющая всем трем поставленным условиям, нам удалась. Допустим, что нам так удалось перенумеровать все квадраты рангов $1, 2, \ldots, k-1$, что оказались соблюденными все три поставленных выше условия. Покажем, что и для квадратов ранга k можно установить требуемую нумерацию.

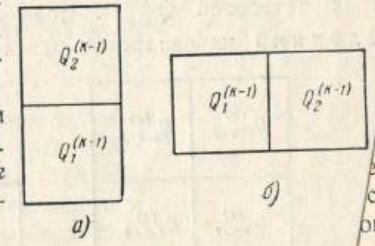


Рис. 4.

Для этого мы прежде всего, разбивая квадрат $Q_1^{(k-1)}$ на четыре равные части, обозначим через $Q_1^{(k)}$ ту из них, которая содерж точку O. Сделав это, мы различим два возможных взаимных ра ложения квадратов $Q_1^{(k-1)}$ и $Q_2^{(k-1)}$, указанных на рис. 4. В случа мы перенумеруем квадраты $Q_i^{(k)}$ для $i = 1, 2, 3, 4$ по обра указанному на рис. 5, *а*, а в случае *б*) — по образцу рис. 5, *б*).

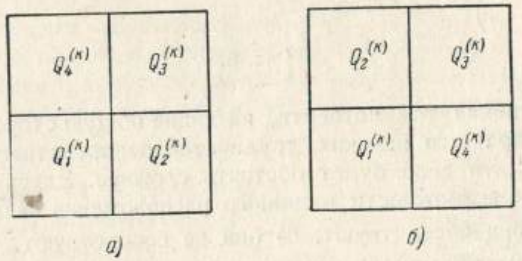


Рис. 5.

Ясно, что в обоих случаях любая пара квадратов $(i = 1, 2, 3)$ будет обладать общей стороной и чере обозначен квадрат, сторона которого является полови квадрата $Q_2^{(k-1)}$.

Допустим же, что для некоторого m $(1 \leqslant m \leqslant 4^{k-1}$ удалось перенумеровать квадраты

$$Q_1^{(k)}, Q_2^{(k)}, \ldots, Q_{4m}^{(k)}$$

так, что любые два из них с соседними нижними индексами имеют общую сторону, и через $Q_{4m}^{(k)}$ оказался обозначенным квадрат, сторона которого является половиной стороны квадрата $Q_{m+1}^{(k-1)}$. Тогда само собой определяется, какая именно из четвертей квадрата $Q_{m+1}^{(k-1)}$ должна получить обозначение $Q_{4m+1}^{(k)}$ — это будет четверть, которая имеет общую сторону с $Q_{4m}^{(k)}$. Остается указать нумерацию для остальных трех четвертей $Q_{m+1}^{(k-1)}$. Если $m+1=4^{k-1}$, т. е. $Q_{m+1}^{(k-1)}$ есть последний из квадратов ранга $k-1$, то надо позаботиться лишь о

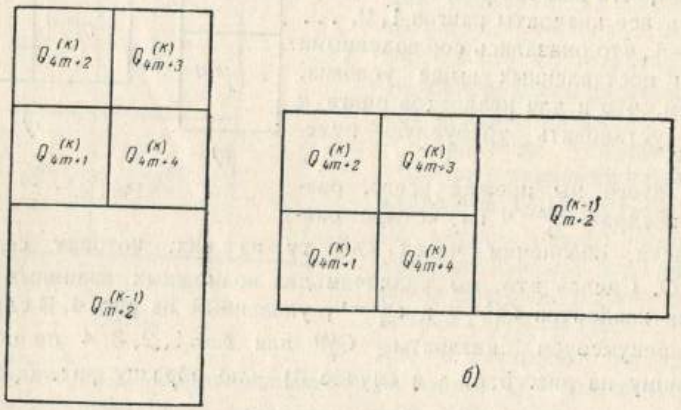


Рис. 6.

том, чтобы упомянутые четверти, имеющие общую сторону, получили соседние номера, что никаких трудностей не представляет. Если же $m+1<4^{k-1}$, то дело будет обстоять сложнее. Здесь нам придется различить две возможности взаимного расположения квадратов $Q_{4m+1}^{(k)}$ и $Q_{m+2}^{(k-1)}$: а) они соседствуют, б) они не соседствуют. Обе эти возможности схематически указаны на рис. 6, где одновременно с этим показана и та нумерация, которую следует применить.

Таким образом, всегда можно осуществить нумерацию, удовлетворяющую всем трем поставленным требованиям.

Как и в случае отрезков, рассмотренных в п. **3**, соотношения $1\leqslant p\leqslant 4^k$ и $4p-3\leqslant q\leqslant 4p$ имеют следствием включение

$$Q_p^{(k)}\supset Q_q^{(k+1)}.$$

Это непосредственно следует из свойства 2) нашей нумерации. Но тогда всякая бесконечная последовательность натуральных чисел

$$i_1,\ i_2,\ i_3,\ \ldots,$$

в которой $1\leqslant i_1\leqslant 4$ и $4i_k-3\leqslant i_{k+1}\leqslant 4i_k$, задает последовательность вложенных друг в друга квадратов

$$Q_{i_1}^{(1)}\supset Q_{i_2}^{(2)}\supset Q_{i_3}^{(3)}\supset\ldots$$

и тем самым определяет единственную точку (x_0, y_0) исходного квадрата Q.

Мы будем для этой точки применять запись

$$(x_0, y_0)=\langle i_1, i_2, i_3, \ldots\rangle,$$

используя ломаные скобки для отличия от ранее введенной записи $[i_1, i_2, i_3, \ldots]$, обозначающей точку отрезка $[0, 1]$.

Весьма существенно, что каждая точка $(x_0, y_0)\in Q$ допускает такую запись[1]), поскольку она является пересечением стягивающейся системы вложенных друг в друга квадратов возрастающих рангов.

5. Теперь мы можем построить интересующие нас функции $\varphi(t)$ и $\psi(t)$. Именно, взяв произвольное $t_0\in[0, 1]$, записываем это t_0 в форме

$$t_0=[i_1, i_2, i_3, \ldots]$$

и рассматриваем точку

$$(x_0, y_0)=\langle i_1, i_2, i_3, \ldots\rangle.$$

Решающим для дальнейшего является то обстоятельство, что *точка* (x_0, y_0) *полностью определяется точкой* t_0. Иными словами, если даже точка t_0 допускает не одну, а две записи вида $[i_1, i_2, i_3, \ldots]$, то обе эти записи приводят к одной и той же точке (x_0, y_0). В самом деле, две записи допускают точки вида $\frac{m}{4^n}$, где $1\leqslant m<4^n$ и m не делится на 4. Как мы видели выше, эти записи таковы: $[i_1, i_2, \ldots, i_{n-1}, m, 4m, 16m, \ldots]$ и $[i_1, i_2, \ldots, i_{n-1}, m+1, 4m+1, 16m+1, \ldots]$. Пусть же

$$\langle i_1, i_2, \ldots, i_{n-1}, m, 4m, 16m, \ldots\rangle=(x_1, y_1),$$
$$\langle i_1, i_2, \ldots, i_{n-1}, m+1, 4m+1, 16m+1, \ldots\rangle=(x_2, y_2).$$

Тогда при любом натуральном r будет

$$(x_1, y_1)\in Q_{4^r m}^{(n+r)},\quad (x_2, y_2)\in Q_{4^r m+1}^{(n+r)}.$$

---

[1]) Некоторые из точек (x_0, y_0) допускают только одну такую запись, а некоторые несколько. Например, точка $\left(\frac{1}{2}, \frac{1}{2}\right)$, очевидно, может быть записана четырьмя способами:

$$\left(\frac{1}{2}, \frac{1}{2}\right)=\langle 1, 3, \ldots\rangle=\langle 2, 8, \ldots\rangle=\langle 3, 9, \ldots\rangle=\langle 4, 14, \ldots\rangle.$$

У квадратов $Q_{4^r m}^{(n+r)}$ и $Q_{4^r m+1}^{(n+r)}$ соседние нижние индексы, и потому эти квадраты имеют общую сторону. Но длина стороны каждого из них есть $2^{-(n+r)}$. Отсюда

$$|x_2 - x_1| \leqslant 2^{1-(n+r)},$$

и ввиду произвольности r оказывается $x_2 = x_1$. Аналогично и $y_2 = y_1$. Стало быть, обе записи точки t_0 привели нас к одной и той же точке (x_0, y_0). Положив

$$\varphi(t_0) = x_0, \quad \psi(t_0) = y_0,$$

мы однозначно определяем функции $\varphi(t)$ и $\psi(t)$, причем само собою ясно, что множество M точек вида $(\varphi(t), \psi(t))$ тождественно с квадратом Q [1]).

6. Остается убедиться в непрерывности функций $\varphi(t)$ и $\psi(t)$. С этой целью рассмотрим две какие-либо точки t' и t'' из $[0, 1]$. Пусть

$$t' = [i_1, i_2, i_3, \ldots],$$
$$t'' = [j_1, j_2, j_3, \ldots].$$

Выберем и закрепим какое-нибудь натуральное k. Тогда

$$t' \in T_{i_k}^{(k)}, \quad t'' \in T_{j_k}^{(k)},$$

или, что то же самое,

$$\frac{i_k - 1}{4^k} \leqslant t' \leqslant \frac{i_k}{4^k}, \quad \frac{j_k - 1}{4^k} \leqslant t'' \leqslant \frac{j_k}{4^k}.$$

Отсюда

$$\frac{i_k - j_k - 1}{4^k} \leqslant t' - t'' \leqslant \frac{i_k - j_k + 1}{4^k},$$

и, стало быть,

$$\left| \frac{i_k - j_k}{4^k} - (t' - t'') \right| \leqslant \frac{1}{4^k}.$$

Но тогда

$$|i_k - j_k| \leqslant 1 + 4^k |t' - t''|.$$

Если

$$|t' - t''| < \frac{1}{4^k}, \qquad (*)$$

то

$$|i_k - j_k| < 2,$$

[1]) В самом деле, если $(x^*, y^*) \in Q$, то (x^*, y^*) допускает одну или несколько записей $\langle i_1, i_2, i_3, \ldots \rangle$. Выбрав одну из этих записей и положив $[i_1, i_2, i_3, \ldots] = t^*$, получим $\varphi(t^*) = x^*$, $\psi(t^*) = y^*$.

и поскольку оба числа i_k и j_k натуральны, ясно, что i_k и j_k либо совпадают, либо суть соседние натуральные числа. Но справедливы включения

$$(\varphi(t'), \psi(t')) \in Q_{i_k}^{(k)}, \quad (\varphi(t''), \psi(t'')) \in Q_{j_k}^{(k)}.$$

Квадраты $Q_{i_k}^{(k)}$ и $Q_{j_k}^{(k)}$ либо совпадают, либо имеют общую сторону. Длина же стороны каждого из них равна 2^{-k}. Стало быть, из неравенства $(*)$ вытекают неравенства

$$|\varphi(t') - \varphi(t'')| \leqslant 2^{1-k}, \quad |\psi(t') - \psi(t'')| \leqslant 2^{1-k}.$$

Теперь уже легко закончить рассуждение. В самом деле, взяв произвольное $\varepsilon > 0$, находим столь большое k, чтобы оказалось

$$2^{1-k} < \varepsilon.$$

Полагая $4^{-k} = \delta$, видим, что неравенство $|t' - t''| < \delta$ обеспечивает неравенства

$$|\varphi(t') - \varphi(t'')| < \varepsilon, \quad |\psi(t') - \psi(t'')| < \varepsilon,$$

чем и доказана непрерывность обеих функций φ и ψ.

7. В заключение мы предлагаем читателю несколько задач, связанных с изложенным материалом.

1. Точка (x_0, y_0) кривой $x = \varphi(t)$, $y = \psi(t)$ называется кратной, если существует больше одного значения t, для которого $\varphi(t) = x_0$, $\psi(t) = y_0$. Если таких значений t имеется ровно p, то точка (x_0, y_0) называется p-кратной. Доказать, что построенная выше кривая Пеано не имеет p-кратных точек при $p > 4$.

2. Доказать, что на построенной кривой имеются как некратные, так и p-кратные точки при $p = 2, 3, 4$. В частности, точка $\left(\frac{1}{2}, \frac{1}{2}\right)$ трехкратная [1]).

3. Доказать, что на всякой (а не только на построенной выше) кривой Пеано обязательно имеются кратные точки.

4. Построить пространственную кривую Пеано, т. е. задать на $[0, 1]$ такие три непрерывные функции $\varphi(t)$, $\psi(t)$, $\chi(t)$, чтобы множество всех точек вида $(\varphi(t), \psi(t), \chi(t))$ совпало с кубом $K(0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, 0 \leqslant z \leqslant 1)$.

[1]) Интересно, что сам Гильберт ошибочно утверждал (цит. соч., стр. 460), что на построенной им кривой имеются (помимо некратных) лишь двукратные и четырехкратные точки.

СТАРИННЫЕ МАТЕМАТИЧЕСКИЕ ЭНЦИКЛОПЕДИИ

Первые печатные математические энциклопедии появились во второй половине XVII века. По-видимому, их тиражи были невысокими: кроме того, эти книги постоянно находились в работе, не залеживаясь на книжных полках. Вот почему сравнительно мало экземпляров математических энциклопедий XVII века сохранилось до наших дней. Об их редкости можно судить хотя бы по тому, что известный историк математики Мориц Кантор, называя три древнейших математических словаря, не дает их подробного описания, так как ни одну из этих книг он сам не видел. Речь идет о следующих трудах:

1) Lexicon mathematicum, astronomicum et geometricum, составленный монахом-театинцем Иеронимом Витале (I. Vitale) и изданный в 1668 г. в Париже; 2) Лексикон Джозефа Моксона (J. Moxon),

Фронтиспис и титульный лист лексикона Витали

изданный в 1680 г. в Лондоне; 3) Dictionnaire mathematique Жака Озанама (J. Ozanam), напечатанный в 1691 г. в Париже (в том же году в Амстердаме).

Мы имеем под руками экземпляр первой (правда, в издании 1690 г.) и третьей из этих книг и рассчитываем рассказать о них на страницах «Математического просвещения». Но к ним нужно добавить еще наиболее обстоятельную математическую энциклопедию XVII века. Это «Двуглавая наука» (Mathesis Biceps) чешского епископа Карамуеля, изданная в Кампании (Италия) в двух больших томах in folio всего двумя годами позже книги Витале. Не вдаваясь в подробности, отметим, что книга Витале представляет, бействительно, лексикон, где в алфавитном порядке расположены

ОКОНЧАНИЕ НА СТР. 266

ВЫЧИСЛИТЕЛЬНЫЙ ЛАБОРАТОРНЫЙ ПРАКТИКУМ ПО КУРСУ ВЫСШЕЙ МАТЕМАТИКИ ВО ВТУЗАХ

Л. З. Румшиский

(Москва)

В инженерную практику всё больше и глубже входит вычислительная техника, и инженеру всё чаще приходится пользоваться методами вычислительной техники. Студенты младших курсов втузов, изучающие сейчас математику, через 3—5 лет придут инженерами на производство; если они к этому времени не будут владеть хотя бы основами вычислительной математики, то окажутся неполноценными работниками. Нужно подготовить студентов к этой стороне их будущей деятельности; этого можно достичь, во-первых, перестроив преподавание курса математики во втузах, а во-вторых, включив в учебные планы всех технических специальностей прохождение цикла лабораторных вычислительных работ с использованием хотя бы простейшей вычислительной техники.

Уже в 1960/61 учебном году во многих втузах введен вычислительный лабораторный практикум; для его проведения создаются математические лаборатории, которые должны быть обеспечены всем необходимым оборудованием, начиная с простейших вычислительных средств и кончая электрическими клавишными счетными машинами и малыми электронными математическими машинами [1]).

В Московском энергетическом институте (МЭИ) уже с 1950 г. при кафедре высшей математики работает математическая лаборатория, и все студенты на первых курсах выполняют от 5 до 8 лабораторных работ. Опираясь на опыт МЭИ, я остановлюсь здесь на некоторых вопросах преподавания вычислительной математики и проведения вычислительного практикума.

[1]) См. «Математическое просвещение», вып. 5, стр. 219—221.

1. ЧИСЛЕННЫЕ МЕТОДЫ В ОСНОВНОМ КУРСЕ ВЫСШЕЙ МАТЕМАТИКИ

В большинстве втузов нет специального курса вычислительной математики, и введение такого курса вряд ли целесообразно. Необходимые сведения о численных методах можно (и нужно!) сообщить студентам в основном курсе математики. Учитывая большую напряженность этого курса, мы вынуждены ограничиться лишь небольшим количеством методов, тесно связанных с курсом. Но каждый излагаемый метод надо «доводить до числа», обращая особое внимание на практическую сторону дела, подчеркивая такие вопросы, как оценку погрешности метода, возможность применения различных расчетных схем и вычислительных средств и т. п. Излагаемые методы должны быть достаточно просты, чтобы студенты могли их легко освоить; в то же время они должны быть и достаточно эффективны [1]).

Приведу несколько примеров.

1. Доказывая теорему «непрерывная функция $f(x)$, имеющая значения разных знаков на концах отрезка, должна обратиться в нуль где-то внутри отрезка», целесообразно тут же указать метод проб («вилки») решения уравнения $f(x)=0$ (его можно иллюстрировать, например, вычислением корня уравнения $x^5-x-1=0$, лежащего между $x=1$ и $x=2$, с точностью до 0,01). Далее, при изложении методов хорд и касательных полезно сравнить эти методы по их эффективности как с методом проб, так и друг с другом. При этом полезно указать различные оценки погрешностей методов, разъяснить вопросы быстроты сходимости, остановиться на вычислительных особенностях каждого метода, отметить удобство приближения к корню с двух сторон. Стоит рассказать также о модификациях метода Ньютона и метода секущих и пояснить критерии выбора метода при решении конкретного уравнения [2]).

2. В интегральном исчислении различные случаи интегрируемости в элементарных функциях и некоторые искусственные приемы интегрирования до сих пор занимают более значительное место, чем приближенные методы вычисления интеграла, хотя последние гораздо важнее для будущего инженера [3]). Это соотношение следует изменить в пользу приближенных и особенно численных методов. При этом нельзя ограничивать изложение только выводом формул (прямоугольников, трапеций и Симпсона) — надо излагать и соответствующие методы расчета, иллюстрируя их числовыми примерами, сравнивая их эффективность, указывая различные оценки соответствующих погрешностей. Важно, например, подчеркнуть, что хотя метод Симпсона не сложнее метода трапеций, он дает значительно более высокую точность. Учитывая важность формулы Симпсона при численном интегрировании дифференциальных уравнений методом Милна, не следует ограничиваться формой $\int_{x_0}^{x_2} y\,dx=\frac{h}{3}(y_0+4y_1+y_2)$, а следует привести также и другую форму:

$$\int_{x_0}^{x_2} y\,dx=h\left(2y_1+\frac{1}{3}\Delta^2 y_0\right)\qquad\left[h=\frac{x_2-x_0}{2}\right].$$

Надо дать указание о выборе шага h, о выборе схемы расчета для задач определенного и неопределенного интегрирования.

3. В разделе «Дифференциальные уравнения 1-го порядка» также слишком много времени уделяют типам уравнений, решаемых в квадратурах, и не указывают эффективных численных методов. Обычно ограничиваются изложением лишь наиболее простого численного метода — метода Эйлера. Следует сейчас же после этого метода (т. е. на одной из первых лекций по дифференциальным уравнениям) дать какой-либо достаточно эффективный метод численного интегрирования, пояснив простым примером недостаточную эффективность метода Эйлера. Как показывает опыт МЭИ, этим задачам лучше всего отвечает метод Милна [1]); он достаточно эффективен, являясь одним из лучших методов, по крайней мере, при ручном счете; с другой стороны, он очень прост и для лекционного изложения, где его легко увязать с формулой Симпсона, и для лабораторных расчетов.

4. В разделе «Ряды» изучению признаков сходимости уделяют гораздо больше внимания, чем приложению рядов к приближенным вычислениям, вопросам оценки остатка, быстроты сходимости. Следует переключить внимание именно на эти вопросы и, может быть, даже дать методы улучшения сходимости. Во всяком случае необходимо добиться того, чтобы студент мог с помощью рядов вычислить интеграл с заданной точностью.

2. ПРОГРАММА ВЫЧИСЛИТЕЛЬНОГО ПРАКТИКУМА

Эта программа должна содержать по крайней мере следующие 4 или 5 основных лабораторных работ: 1) *составление таблицы функций по заданной формуле* (*с заданной точностью*) *с применением*

[1]) При этом надо разъяснить студентам, что́ понимается под эффективностью метода.

[2]) Время на изложение этих вопросов можно выкроить, отказавшись от чего-либо менее существенного, например, от излишне детального исследования локального поведения функций.

[3]) Не говоря уже о том, что если инженеру понадобится нахождение первообразной функции в более или менее сложных случаях, он всегда предпочтет пользоваться справочником.

[1]) См. книгу, указанную в сноске на стр. 242.

простейших вычислительных средств, 2) *численное решение уравнения с одним неизвестным*, 3) *приближенное вычисление определенных интегралов*, 4) *численное интегрирование обыкновенных дифференциальных уравнений 1-го порядка*, и (во втузах с объемом курса математики не менее 360 часов) 5) *решение системы линейных уравнений*. Сверх этого обязательного минимума, в программу можно включить и дополнительные работы в зависимости от специальности (например, гармонический анализ с помощью шаблонов или приборов, обработка экспериментальных данных, употребление номограмм различных типов и т. д.).

В программу практикума должны, естественно, войти и *правила вычислений с приближенными числами*. Однако вряд ли целесообразно посвящать этим правилам специальные лабораторные занятия. Основные сведения о приближенных вычислениях можно сообщить и на практических занятиях по курсу математики, а закрепление соответствующих навыков достигается только непрерывным вниманием к этим вопросам при выполнении всех лабораторных работ [1]).

Общая программа вычислительного практикума (без специальных работ) выглядит в МЭИ примерно так (в скобках указано число двухчасовых лабораторных занятий):

1) правила приближенных вычислений (на одном практическом занятии);

2) составление таблицы функций с заданной точностью с помощью счетной линейки, арифмометра и математических таблиц (2—3);

3) решение уравнений способами хорд и касательных (2);

4) решение системы линейных уравнений методом Гаусса (1);

5) приближенное вычисление определенных интегралов по формуле Симпсона (1);

6) применение планиметра к вычислению интегралов (1);

7) численное интегрирование дифференциального уравнения 1-го порядка методом Милна (3);

8) вычисление значений функций и интегралов с помощью степенных рядов (на двух практических занятиях);

9) приближенный гармонический анализ с помощью шаблонов (1).

На кафедрах математики часто возникает вопрос о том, когда лучше проводить те или иные лабораторные работы. С точки зрения связи теории с практикой было бы хорошо, если бы выполнение каждой работы совпадало по времени с изучением соответствующего раздела курса. Но тогда работы будут распылены по всем семестрам; это мешает закреплению навыков студентов к расчетной работе. Опыт МЭИ показывает, что полезно группировать работы по родственным методам. Например, можно выделить в отдельный комплекс работы по интерполяции и составлению таблиц функций, по приближенному вычислению интегралов методом Симпсона и по численному интегрированию дифференциальных уравнений методом Милна. Это позволяет закрепить у студентов навыки по применению таблиц и табличных разностей, по численному интегрированию. Практически можно сделать так: сначала дать работу на составление таблицы функций $f(x)$ с некоторым шагом; затем предложить составить таблицу значений интеграла от этой же функции $\int_a^x f(x)\,dx$ с помощью формулы Симпсона (причем вопрос о шаге интегрирования решается уже в зависимости от величины разностей 4-го порядка таблицы $f(x)$); далее предложить вычислить какой-нибудь определенный интеграл от $f(x)$ с помощью интерполяции составленной таблицы неопределенного интеграла; наконец, дать работу на численное интегрирование дифференциального уравнения методом Милна (с использованием формулы Симпсона в качестве формулы уточнения и с применением формулы интерполяции на середину при уменьшении шага расчета).

[1]) Этой же цели может служить и специальная памятка по приближенным вычислениям, которая выдается каждому студенту при прохождении практикума.

3. ПРОВЕДЕНИЕ ЛАБОРАТОРНЫХ ЗАНЯТИЙ

Для того чтобы привить студентам навыки в организации и проведении вычислительной работы, развить у них «культуру счета» и обучить их работе на счетных машинах и приборах, нужна четкая организация лабораторных занятий. Студенту дается определенное задание и указывается метод расчета [1]). Он должен суметь составить схему расчета, установить порядок вычислений и порядок записи их результатов, быстро, экономно и правильно провести вычисления, оценить их точность.

Студенты должны выполнить вычислительную часть работы в помещении лаборатории под руководством опытного преподавателя. До лабораторного занятия студенты должны проделать самостоятельно всю подготовительную работу: ознакомиться с вычислительными особенностями метода расчета и с возможными схемами расчета, уяснить поставленную задачу и подготовить ее к расчету по выбранной (или заданной) схеме, подготовить расчетный бланк для работы, а если нужно, то и графическую часть работы (например, отделить

[1]) Мы пока не ставим задачу научить студента самостоятельно выбирать метод расчета — это уже дело *специального практикума* на старших курсах, если таковой будет создан.

искомый корень уравнения $f(x)=0$ графическим методом)[1]. Степень подготовленности студентов к занятию нужно обязательно проверять, например, путем коллоквиума.

Чтобы выполнить указанную подготовительную работу, студент должен ознакомиться с описаниями методов расчета и инструкциями по выполнению лабораторных работ. Всякая лабораторная работа должна быть обеспечена печатными инструкциями и тщательно выверенными заданиями[2]).

На самом лабораторном занятии основное внимание должно быть обращено на технику расчета, правильность записи промежуточных результатов и обеспечение требуемой точности. Остановлюсь подробнее на некоторых практических вопросах.

1. На первых лабораторных занятиях студенты, естественно, будут пользоваться готовыми расчетными бланками и заполнять их в соответствии с инструкцией. Как показывает опыт, больше всего неясностей возникает здесь в вопросах точности производимых вычислений. Надо *научить студентов свободно ориентироваться в правилах приближенных вычислений,* чтобы сохранять в каждом результате достаточное число знаков для обеспечения требуемой точности, но не сохранять лишних знаков, только усложняющих вычислительную работу. Этот вопрос требует пристального внимания преподавателя.

2. На первых же лабораторных занятиях проводится *обучение технике использования простейших вычислительных средств* (таблицы, линейка, счетные машины). Этой техникой студенты овладевают, как правило, довольно быстро в самом процессе выполнения работ. Достаточно дать лишь самые необходимые указания (и к тому же весьма краткие, чтобы лучше запоминались) и по мере надобности консультировать студента по отдельным техническим вопросам (что может с успехом делать и лаборант),— и через несколько часов студент будет уверенно работать на арифмометре или электрической счетной машине. Другое дело — *выяснение связи между выбором тех или иных вычислительных средств и требуемой точностью расчета, а также и применяемым методом вычислений.* На это нужно специально обращать внимание, указывая, например, в каких случаях счетную линейку следует предпочесть арифмометру или на каких машинах можно применять компактные схемы решения систем линейных уравнений.

---

[1]) Опыт МЭИ показывает, что на каждый час расчетной работы в лаборатории студент затрачивает примерно час на подготовку к ней.

[2]) По основным пяти лабораторным работам описания методов и инструкции имеются в книге: Л. З. Румшиский, Вычислительный лабораторный практикум по основному курсу высшей математики для втузов, Физматгиз, М., 1961.

3. При ознакомлении с расчетными схемами и порядком выполнения вычислений по сериям однотипных операций следует обращать внимание студентов на возможность *автоматизации вычислительного процесса.* Хотя в практикуме по основному курсу математики еще не ставится задача ознакомления с быстродействующими машинами с программным управлением, но практикум надо рассматривать и как необходимый элемент подготовки к такому ознакомлению: основные принципы программирования при ручном и при машинном счете одинаковы.

4. Надо научить студента *контролировать каждый этап расчета.* Схема контроля должна быть включена в схему расчета. Еще важнее убедить студента в необходимости систематического контроля вычислений. Наличие эффективного контроля позволяет ускорить все вычисления, а лишняя работа, затрачиваемая на самый контроль, вполне окупается гарантией верности результата. Надо добиться того, чтобы, проводя расчет, студент был уверен в правильности получаемого результата. Ведь не секрет, что студент часто сверяет решение задачи с ответом. Но в задачах, которые будет решать инженер, готовых ответов нет, а численные результаты не всегда можно проверять непосредственно. Очень важно, чтобы и преподаватель не сверял результат с ответом, а приобрел достаточный опыт для быстрого обнаружения ошибок в расчете (для этого необходимо, чтобы преподаватели сами просчитывали варианты работ, предлагаемых студентам). Обходя студентов на лабораторных занятиях и наблюдая за контрольными данными, опытный преподаватель сразу же укажет студенту на место ошибки и на ее причину. Если же проверяемая часть расчета верна, то желательно сделать об этом соответствующую пометку на расчетном бланке, чтобы студент увереннее и смелее проводил дальнейший расчет.

5. Несколько слов об *оформлении лабораторных работ.* Будущий инженер должен уметь оформить расчетную или расчетно-графическую работу. И, конечно, аккуратность в расчетной работе нужна более, чем в любой другой работе: расчетный бланк должен быть вычерчен по определенной схеме, каждое число должно быть вписано в соответствующую клетку этого бланка. Но ни в коем случае не надо требовать от студентов, чтобы они переписывали расчет «набело» ради «красивости» оформления. Больше того, надо *запретить такое переписывание,* подчеркнув, что всякое переписывание расчета не только отнимает лишнее время, но и является источником дополнительных ошибок.

6. И, наконец, *о проверке лабораторных работ.* Хотя это и трудно организовать, но надо стараться ввести защиту лабораторных работ, хотя бы наиболее трудных и наиболее важных для данной специальности.

4. НЕКОТОРЫЕ ОРГАНИЗАЦИОННЫЕ ВОПРОСЫ

Для успешного введения вычислительного лабораторного практикума кафедры математики втузов должны провести не только методическую, но и организационную подготовку.

Прежде всего, математическая лаборатория должна быть внесена в штатный список лабораторий на равных правах с другими лабораториями. Ей должны быть выделены помещения, штат лаборантов и необходимые ассигнования на приобретение сначала простейшего оборудования (счетных линеек, арифмометров и математических таблиц), а в дальнейшем и электрических клавишных счетных машин и электронных цифровых и аналоговых машин (снабжение которыми будет проводиться в централизованном порядке). По мере поступления электрических машин надо целиком переводить на них лабораторные работы, снимая с работ арифмометры «Феликс», отнимающие у студентов слишком много времени.

Лабораторные занятия должны быть введены как в учебный план по математике, так и в расписание учебных занятий на всех факультетах. Выполнение лабораторных работ входит в зачет по математике, который обязательно должен быть введен на соответствующих семестрах.

Наконец, при внутривтузовском распределении штатов кафедре математики необходимо выделить необходимые штаты для проведения лабораторных занятий [1]).

Для успешного введения вычислительного лабораторного практикума большое значение должен иметь обмен опытом как непосредственно между математическими лабораториями и кафедрами математики втузов, так и через печать, например через «Математическое просвещение».

---

[1]) В инструктивном письме Министерства высшего и среднего специального образования СССР («И-38») предусмотрено проведение лабораторных работ по полугруппам, чтобы преподаватель имел возможность уделить достаточно внимания каждому студенту.

К ВОПРОСУ ОБ ОПРЕДЕЛЕНИИ ТРИГОНОМЕТРИЧЕСКИХ ФУНКЦИЙ С ПОМОЩЬЮ ФУНКЦИОНАЛЬНЫХ УРАВНЕНИЙ

М. Я. Цыркин

(Армавир)

При аналитическом определении тригонометрических функций с помощью функциональных уравнений принято вводить косинус и синус совместно, рассматривая их как решения системы функциональных уравнений с двумя неизвестными функциями [1]), или одного функционального уравнения [2]). Ю. М. Гайдук [3]) показал, как можно определить косинус, пользуясь одним функциональным уравнением с одной неизвестной функцией.

В настоящей статье рассматриваются два функциональных уравнения:

$$g(x+y)+g(x-y)=2g(x)\,g(y) \tag{I}$$

и

$$f^2(x)-f^2(y)=f(x+y)\,f(x-y). \tag{II}$$

Первое, известное в литературе как уравнение Даламбера — Пуассона, было использовано Ю. М. Гайдуком для определения косинуса; второе мы используем для определения синуса. Мы рассмотрим сначала решения этих уравнений в непрерывных функциях действительной переменной, а затем покажем, как на основе этих уравнений можно определить тригонометрические функции.

1. Будем искать функции $g(x)$, определенные при всех действительных значениях аргумента и удовлетворяющие при любых действительных x и y уравнению (I). Исключая тривиальное решение

---

[1]) Ф. Франклин, Математический анализ, ч. 1, 1950; П. М. Котельников, О функциональных уравнениях, определяющих тригонометрические функции, «Математика в школе» (1951), № 2.

[2]) Waughan, American Mathematical Monthly, 62, № 10 (1955); стр. 707. С. И. Новоселов, Специальный курс тригонометрии, 1957.

[3]) Ю. М. Гайдук, К вопросу об аналитическом и геометрическом определении тригонометрических функций, «Математика в школе» (1953), № 4.

$g(x)\equiv 0$, выводим из этого уравнения, положив последовательно $y=0$, $x=0$, $y=x$ и $y=2x$:

$$\text{а) } g(0)=1,\ \text{б) } g(-y)=g(y),$$
$$\text{в) } g(2x)+1=2g^2(x),\ \text{г) } g(3x)+g(x)=2g(2x)\,g(x). \qquad (1)$$

Докажем, что решения уравнения (I) суть либо функции непрерывные при любом значении аргумента, либо терпящие в каждой точке разрыв, и притом такой, что функция не имеет ни правостороннего, ни левостороннего конечного предела.

Действительно, пусть существует конечный правосторонний предел функции $g(x)$ в точке $x=a$. Тогда из (1, в) следует существование правостороннего предела функции $g(x)$ при $x=2a$. Из $g(2a+x)+g(x)=2g(a+x)\,g(a)$ устанавливаем существование $\lim\limits_{x\to 0+} g(x)$. Обозначим его через α. Из (1, в, г) следует, что $\alpha+1=2\alpha^2$ и $2\alpha=2\alpha^2$, откуда $\alpha=1$. В силу четности $g(x)$ имеем $\lim\limits_{x\to 0-} g(x)=1$ и, следовательно, функция $g(x)$ непрерывна при $x=0$.

Докажем, что $g(x)$ непрерывна при любом x. Имеют место равенства:

$$\lim_{y\to 0}[g(x+y)+g(x-y)]=2g(x),$$
$$\lim_{y\to 0}[2g(x+y)\,g(x-y)]=\lim_{y\to 0}[g(2x)+g(2y)]=$$
$$=g(2x)+1=2g^2(x),$$

откуда

$$\lim_{y\to 0}[g(x+y)-g(x-y)]^2=0,\quad \lim_{y\to 0}[g(x+y)-g(x-y)]=0$$

и

$$\lim_{y\to 0} g(x+y)=\lim_{y\to 0} g(x-y)=g(x).$$

Существование разрывных решений уравнения (I) доказывается на основании базиса Хамеля всех действительных чисел с использованием аксиомы произвольного выбора. На этом вопросе мы задерживаться не будем [1]), ограничившись рассмотрением лишь непрерывных решений.

Из (1, в) следует, что $g(x)\geqslant -1$. Возможны два случая:

Случай 1. В какой-то точке x_0 $g(x_0)=-1$. Тогда, в силу (1, в), $g\left(\frac{x_0}{2}\right)=0$, т. е. существует нуль функции $g(x)$. Так как $g(0)=1$, то $x_0\neq 0$, а учитывая четность функции $g(x)$ можно принять, что $x_0>0$. Рассмотрим множество всех нулей функции $g(x)$, расположенных на оси OX справа от точки $x=0$. Оно ограничено снизу и, следовательно, имеет нижнюю грань, которую мы обозначим через $\frac{\lambda}{2}$. Очевидно, что в силу непрерывности $g(x)$, $g\left(\frac{\lambda}{2}\right)=0$ и в силу (1, а) $\lambda\neq 0$. Это значит, что существует наименьший положительный нуль функции $g(x)$. В интервале $\left(0,\ \frac{\lambda}{2}\right)$ функция $g(x)$ сохраняет свой знак, а так как $g(0)=1$, то в интервале $\left(0,\ \frac{\lambda}{2}\right)$ $g(x)>0$.

Из (1, в) находим

$$g\left(\frac{x}{2}\right)=\pm\sqrt{\frac{1+g(x)}{2}}, \qquad (2)$$

причем для $0<\frac{x}{2}<\frac{\lambda}{2}$ перед корнем надлежит ставить знак «плюс». Из (I) находим

$$g[(n+1)x]+g[(n-1)x]=2g(nx)\,g(x). \qquad (3)$$

Докажем, что

$$g(x)=\cos\frac{\pi x}{\lambda}. \qquad (4)$$

В самом деле, функция $\cos\frac{\pi x}{\lambda}$ удовлетворяет уравнению (I), и равенство (4) имеет место при $x=\frac{\lambda}{2}$. В силу (2) равенство (4) имеет место при $x=\frac{\lambda}{4}$, $\frac{\lambda}{8}$, ..., т. е. при $x=\frac{\lambda}{2^m}$ (m — любое натуральное число). В силу (3) равенство (4) имеет место при $x=\frac{n\lambda}{2m}$ (n, m — любые натуральные числа). Пусть x — произвольное положительное число. Перейдя в равенстве (4) к пределу при $\frac{n\lambda}{2^m}\to x$ и учитывая непрерывность $g(x)$ и $\cos\frac{\pi x}{\lambda}$, устанавливаем равенство при произвольном $x\geqslant 0$. Для $x<0$ оно следует из четности функции.

Итак, в рассматриваемом случае $g(x)=\cos\frac{\pi x}{\lambda}$.

Случай 2. $g(x)>-1$ при любом x. Тогда, как следует из (2), $g(x)\neq 0$, и в силу (1, а) и непрерывности $g(x)$ имеем $g(x)>0$. Применяя последовательно (1, в), находим

$$g(x)>\frac{1}{2}\sqrt{2},\quad g(x)>\frac{1}{2}\sqrt{2+\sqrt{2}},$$
$$g(x)>\frac{1}{2}\sqrt{2+\sqrt{2+\sqrt{2}}},\ \ldots$$

[1]) Об этом см.: Ю. М. Гайдук, Принцип полной математической индукции, его эквиваленты и обобщения, «Математика в школе» (1950), № 2.

и, следовательно,

$$g(x) \geqslant \lim \frac{\sqrt{2+\sqrt{2+\ldots+\sqrt{2}}}}{2} = 1.$$

Рассмотрим функцию $\operatorname{ch} ax = \frac{1}{2}(e^{ax} + e^{-ax})$ [1]), которая удовлетворяет (1). Задав в точке $x_0 \neq 0$ значение функции $g(x_0) \geqslant 1$, найдем такое значение параметра a, при котором $\operatorname{ch} ax_0 = g(x_0)$. В силу (2) и (3) имеем $\operatorname{ch} \frac{nax_0}{2^m} = g\left(\frac{nx_0}{2^m}\right)$ при любых натуральных m и n. Переходя к пределу в последнем равенстве при $\frac{nx_0}{2^m} \to x$ и учитывая непрерывность обеих функций, находим, что $g(x) = \operatorname{ch} ax$ при $x \geqslant 0$ и в силу четности функций также при $x < 0$. Итак, в рассматриваемом случае $g(x) \equiv \operatorname{ch} ax$.

Таким образом, уравнение (1) имеет непрерывные решения лишь вида:

$$g(x) \equiv 0, \quad g(x) \equiv 1, \quad g(x) = \cos ax, \quad g(x) = \operatorname{ch} ax \qquad (a \neq 0).$$

Ю. М. Гайдук использовал уравнение (1) для аналитического определения аналитического косинуса. Нам кажется целесообразным дать следующее определение.

Функцию $g(x)$, определенную на всей действительной оси, назовем аналитическим косинусом, если 1) *она удовлетворяет уравнению* (1) *при любых x и y,* 2) *в некотором интервале $0 < x < \frac{\lambda}{2}$ $g(x) > 0$,* 3) $g\left(\frac{\lambda}{2}\right) = 0$.

В этом определении условие непрерывности $g(x)$, имеющееся у Ю. М. Гайдука, заменено условием ее положительности в интервале $\left(0, \frac{\lambda}{2}\right)$, но, как легко доказать, они эквивалентны. В самом деле, выше мы видели, что из непрерывности функции $g(x)$ следует, что она положительна в интервале $\left(0, \frac{\lambda}{2}\right)$. С другой стороны, из условий, определяющих функцию $g(x)$, находим последовательно

$$g\left(x+\frac{\lambda}{2}\right) + g\left(x-\frac{\lambda}{2}\right) = 0, \quad g\left(x+\frac{\lambda}{2}\right) = -g\left(x-\frac{\lambda}{2}\right);$$

отсюда $g(x+\lambda) = -g(x)$. Далее,

$$g(x) - g(y) = g(x) + g(y+\lambda) = 2g\left(\frac{x+y}{2}+\frac{\lambda}{2}\right) g\left(\frac{x-y}{2}-\frac{\lambda}{2}\right) = -2g\left(\frac{\lambda}{2}-\frac{x+y}{2}\right) g\left(\frac{\lambda}{2}-\frac{x-y}{2}\right).$$

Пусть $0 \leqslant y < x \leqslant \frac{\lambda}{2}$. Тогда $0 \leqslant \frac{\lambda}{2} - \frac{x+y}{2} < \frac{\lambda}{2}$ и $0 \leqslant \frac{\lambda}{2} - \frac{x-y}{2} < \frac{\lambda}{2}$

[1]) При $a = 0$ имеем предельный случай $\operatorname{ch} ax \equiv 1$.

и в силу условий (1, 2 и 3) $g(x) - g(y) < 0$, т. е. $g(x)$ — функция, убывающая на сегменте $\left[0, \frac{\lambda}{2}\right]$. Отсюда следует существование одностороннего предела в точке $x = 0$ и, как выше доказано, непрерывность функции $g(x)$ при любом x [1]).

2. Легко видеть, что $\sin x$ является решением функционального уравнения (II). Наша цель — найти все непрерывные функции $f(x)$ действительного переменного, удовлетворяющие (II) при любых действительных x и y.

Уравнение (II) имеет тривиальное решение $f(x) \equiv 0$. Мы в дальнейшем примем, что $f(x) \not\equiv 0$.

Положив в (II) $x = y = 0$, находим $f(0) = 0$. Положив $x = 0$, находим $-f^2(y) = f(y) \cdot f(-y)$ или, заменяя y на $-y$, $-f^2(-y) = f(-y) \cdot f(y)$, откуда следует, что $f(y)$ и $f(-y)$ обращаются в нуль лишь одновременно и что $f(-y) = -f(y)$.

Докажем теперь, что решения уравнения (II) либо непрерывны на всем множестве действительных чисел, либо в каждой точке терпят разрыв и притом такой, что функция не имеет односторонних пределов.

Пусть существует конечный односторонний предел функции $f(x)$ в точке $x = a$, например правосторонний. Из $f^2(a+x) - f^2(x) = f(a+2x) \cdot f(a)$ заключаем о существовании правостороннего предела $f^2(x)$ в точке $x = 0$. Но из (II) следует $f^2(2x) - f^2(x) = f(3x) f(x)$. Возведя обе части в квадрат и переходя к пределу при $x \to 0^+$, находим, что $\lim\limits_{x \to 0+} f(x) = 0$. В силу нечетности функции $f(x)$ и $\lim\limits_{x \to 0-} f(x) = 0$; следовательно, $f(x)$ непрерывна в точке $x = 0$.

Докажем теперь, что $f(x)$ непрерывна при любом x. В самом деле, из (II) находим

$$\lim_{y \to 0} [f(x+y) \cdot f(x-y)] = f^2(x) \quad \text{и} \quad \lim_{y \to 0} [f^2(x+y) - f^2(x-y)] = 0,$$

откуда

$$\lim_{y \to 0} [f^2(x+y) + f^2(x-y)]^2 = 4f^4(x),$$

$$\lim_{y \to 0} [f^2(x+y) + f^2(x-y)] = 2f^2(x)$$

и

$$\lim_{y \to 0} f^2(x+y) = f^2(x).$$

[1]) Можно идти и другим путем: установить непрерывность $g(x)$ при $x = 0$ и ограниченность $g(x)$, а затем использовать вышеприведенное выражение для $g(x) - g(y)$.

Если $f(x)=0$ при каком-либо x, то $\lim\limits_{y\to 0} f^2(x+y)=0$ и $\lim\limits_{y\to 0} f(x+y)=0=f(x)$. При тех же значениях x, при которых $f(x)\neq 0$, из $f^2\left(x+\frac{y}{2}\right)-f^2\left(\frac{y}{2}\right)=f(x+y)\cdot f(x)$, переходя к пределу при $y\to 0$, находим

$$\lim_{y\to 0} f(x+y)=f(x).$$

Предложение доказано. В дальнейшем будем рассматривать лишь непрерывные решения уравнения (II).

Пусть α — такое число, что $f(\alpha)\neq 0$. Введем функцию

$$g(x)=\frac{f(x+\alpha)-f(x-\alpha)}{2f(\alpha)}. \tag{5}$$

Докажем, что

$$f(x+y)-f(x-y)=2f(y)\cdot g(x). \tag{6}$$

В самом деле, из (II) получаем: $f(y)[f(x+\alpha)-f(x-\alpha)]=$ $=f^2\left(\frac{x+y+\alpha}{2}\right)-f^2\left(\frac{x-y+\alpha}{2}\right)-f^2\left(\frac{x+y-\alpha}{2}\right)+f^2\left(\frac{x-y-\alpha}{2}\right)=$ $=f(\alpha)[f(x+y)-f(x-y)]$. Из (6) следует

$$f(x+y)+f(x-y)=2f(x)g(y). \tag{7}$$

Складывая (6) и (7), находим

$$f(x+y)=f(x)g(y)+g(x)f(y) \tag{8}$$

и, в частности, $f(2x)=2f(x)g(x)$. Применяя (6) и (5), находим

$$2f(\alpha)[g(x+y)+g(x-y)]=$$
$$=f(x+y+\alpha)-f(x+y-\alpha)+f(x-y+\alpha)-f(x-y-\alpha)=$$
$$=2f(y+\alpha)g(x)-2f(y-\alpha)g(x)=4f(\alpha)g(y)g(x),$$

откуда следует, что $g(x)$ удовлетворяет (I).

Функция $g(x)$ ввиду (5) непрерывна, и $g(0)=1$. Поэтому, как мы установили в п. **1**, $g(x)$ может быть только функцией вида

$$g(x)\equiv 1, \quad g(x)=\cos ax, \quad g(x)=\operatorname{ch} ax \qquad (a\neq 0).$$

Положив в (8) $g(x)=1$, находим, что $f(x)$ удовлетворяет уравнению $f(x+y)=f(x)+f(y)$, непрерывные решения которого, как известно, имеют вид $f(x)=Ax$ (A — произвольная постоянная; в данном случае $A\neq 0$). Функция $f(x)=Ax$, как легко проверить, удовлетворяет (II).

Чтобы найти остальные решения (II), заметим, что

$$f(\alpha)[g(\alpha+x)-g(\alpha-x)]=\frac{1}{2}[f(2\alpha+x)-2f(x)-f(2\alpha-x)]=$$
$$=f(x)g(2\alpha)-f(x),$$

откуда $f(x)=f(\alpha)\frac{g(\alpha+x)-g(\alpha-x)}{g(2\alpha)-1}$. Подставляя $g(x)=\cos ax$, находим $f(x)=\frac{f(\alpha)}{\sin a\alpha}\sin ax=A\sin ax$, а подставляя $g(x)=\operatorname{ch} ax$, находим $f(x)=\frac{f(\alpha)}{\operatorname{sh} a\alpha}\operatorname{sh} ax=A\operatorname{sh} ax$ (A — произвольная постоянная, отличная от нуля).

Проверкой убеждаемся, что функции $A\sin ax$ и $A\operatorname{sh} ax$ удовлетворяют (II). Итак, уравнение (II) имеет непрерывные решения только вида

$$f(x)=Ax, \quad f(x)=A\sin ax, \quad f(x)=A\operatorname{sh} ax,$$

где A и α — произвольные постоянные [1]).

Результаты показывают, что уравнение (II) может быть использовано для аналитического определения тригонометрического синуса, разумеется, при некоторых дополнительных условиях. Эти дополнительные условия могут быть сформулированы по-разному. Нам кажется целесообразным дать следующее определение.

Аналитическим синусом назовем функцию $f(x)$, определенную при всех действительных значениях аргумента и 1) *удовлетворяющую при всех действительных x и y уравнению* (II), 2) *положительную в некотором интервале $0<x<\frac{\lambda}{2}$ и* 3) $f(\lambda)=0$, $f\left(\frac{\lambda}{2}\right)=1$.

Функция $f(x)$ совпадает с $\sin\frac{\pi x}{\lambda}$.

Обоснование определения и утверждения, что $f(x)=\sin\frac{\pi x}{\lambda}$, может легко быть получено.

3. Рассмотрим теперь другие функциональные уравнения, доставляемые теоремами сложения тригонометрических и гиперболических синуса и косинуса:

$$f(x+y)-f(x-y)=2f(y)g(x), \tag{III}$$
$$f(x+y)+f(x-y)=2f(x)g(y), \tag{IV}$$
$$g(x+y)-g(x-y)=-2f(x)f(y), \tag{V}$$
$$g(x+y)-g(x-y)=2f(x)f(y). \tag{VI}$$

[1]) Здесь мы включаем также $f(x)\equiv 0$, что соответствует $A=0$ или $a=0$.

а) Уравнение (III) имеет очевидные решения: 1) $f(x)\equiv 0$, $g(x)$ — произвольная функция; 2) $g(x)\equiv 0$, $f(x)$ — произвольная постоянная.

Положим $f(x)\not\equiv 0$, $g(x)\not\equiv 0$. Из (III) легко выводим $f(-y)=-f(y)$, $g(-x)=g(x)$, $f(0)=0$, $g(0)=1$, $f(2x)=2f(x)g(x)$, откуда следует, что $f(x)$ и $g(x)$ удовлетворяют уравнению (IV).

Перемножив (III) и (IV), приходим к уравнению (II), в котором вместо x и y подставлено соответственно $(x+y)$ и $(x-y)$. Предполагая, что $f(x)$ имеет в какой-либо точке односторонний предел или что $\lim\limits_{x\to 0} f(x)=0$, находим решения (II):

$$f(x)=Ax,\quad f(x)=A\sin ax,\quad f(x)=A\operatorname{sh} ax,$$

где A и a — произвольные постоянные, отличные от нуля. Соответствующие значения $g(x)$ находим из (III):

$$g(x)=1,\quad g(x)=\cos ax,\quad g(x)=\operatorname{ch} ax.$$

Этим исчерпываются все непрерывные решения уравнения (III), отличные от тривиальных.

б) Уравнение (IV) имеет очевидное решение $f(x)\equiv 0$, $g(x)$ — произвольная функция.

Предполагая $f(x)\not\equiv 0$ [и в силу этого и $g(x)\not\equiv 0$], находим из (IV) $g(-y)=g(y)$.

Пусть α — такое значение x, что $f(\alpha)\neq 0$. Положив в (IV) $x=\alpha$, находим $g(y)=\frac{f(\alpha+y)+f(\alpha-y)}{2f(\alpha)}$. В силу этого

$$g(x+y)+g(x-y)=$$
$$=\frac{1}{2f(\alpha)}[f(\alpha+x+y)+f(\alpha-x-y)+f(\alpha+x-y)+$$
$$+f(\alpha-x+y)]=$$
$$=\frac{1}{2f(\alpha)}[2f(\alpha+x)g(y)+2f(\alpha-x)g(y)]=2g(x)g(y).$$

Полученное соотношение позволяет решить уравнение (IV), подставив в него вместо $g(x)$ решения функционального уравнения (II). Но в этом случае нам пришлось бы решать три уравнения (если ограничиться непрерывными функциями). Чтобы сократить выкладки, пойдем другим путем. Заметим, что если функции $f(x)$, $g(x)$ удовлетворяет уравнению (IV), то ему удовлетворяют и функции $f_1(x)=f(x)-f(0)g(x)$, $g(x)$, в чем убеждаемся непосредственной подстановкой. Но для функции $f_1(x)$ имеем $f_1(x)+f_1(-x)=f(x)+f(-x)-2f(0)g(x)=0$, т. е. $f_1(-x)=-f_1(x)$. Поэтому пара функций $f_1(x)$, $g(x)$ удовлетворяет также уравнению (III), решенному выше.

Предполагая дальше, что $f(x)$ и $g(x)$ [следовательно, и $f_1(x)$] имеют в какой-либо точке правосторонние (или левосторонние) пределы или же что $\lim\limits_{x\to 0} f(x)=f(0)$ и $\lim\limits_{x\to 0} g(x)=g(0)$, находим решения уравнений (IV), исключая тривиальные:

$$g(x)=1,\qquad f(x)=Ax+f(0),$$
$$g(x)=\cos ax,\qquad f(x)=A\sin ax+f(0)\cos ax,$$
$$g(x)=\operatorname{ch} ax,\qquad f(x)=A\operatorname{sh} ax+f(0)\operatorname{ch} ax,$$

где A, a, $f(0)$ — произвольные постоянные.

Очевидно, что уравнения

$$f(x+y)+f(x-y)=2f(x),\qquad f(x+y)+f(x+y)=2f(x)\cos ay,$$
$$f(x+y)+f(x-y)=2f(x)\operatorname{ch} ay,$$

имеют соответственно нетривиальные непрерывные решения:

$$f(x)=Ax+f(0),\quad f(x)=A\sin ax+f(0)\cos ax,$$
$$f(x)=A\operatorname{sh} ax+f(0)\operatorname{ch} ax\ [1]).$$

в) Рассмотрим теперь одновременно функциональные уравнения (V) и (VI) [2]). Оба уравнения имеют тривиальные решения: $f(x)\equiv 0$ $g(x)=\text{const}$.

Предположим $f(x)\not\equiv 0$. Из (V) и (VI) находим $f(0)=0$, $g(-y)=g(y)$ и $f(-y)=-f(y)$, а также

$$g(2x)=g(0)\mp 2f^2(x). \tag{9}$$

Используя эти соотношения, приводим уравнения (V) и (VI) к виду: $f^2\left(\frac{x+y}{2}\right)-f^2\left(\frac{x-y}{2}\right)=f(x)f(y)$ или (II).

Предполагая, что функция $f(x)$ имеет конечный односторонний предел в какой-либо точке или что $\lim\limits_{x\to 0} f(x)=0$, находим решения:

$$f(x)=Ax,\quad f(x)=A\sin ax,\quad f(x)=A\operatorname{sh} ax.$$

Подставляя найденные значения функции $f(x)$ в (9), находим соответствующие значения функции:

$$g(x)=g(0)\mp\frac{A^2x^2}{2},\quad g(x)=g(0)\mp A^2\pm A^2\cos ax,$$
$$g(x)=g(0)\pm A^2\mp A^2\operatorname{ch} ax,$$

причем верхние знаки для уравнения (V), а нижние для (VI).

---

[1]) Другой путь решения уравнения $f(x+y)+f(x-y)=2f(x)\cos ay$ приведен в статье: Я. Ацель, Некоторые общие методы в теории функциональных уравнений одной переменной, Успехи матем. наук (1956), т. XI, вып. 3.

[2]) Впрочем, каждое из них может быть приведено к другому подстановкой $g(x)=-g_1(x)$.

г) Очевидно, что уравнения (III), (IV) и (V) [а также (VI)] могут служить для аналитического определения тригонометрических функций.

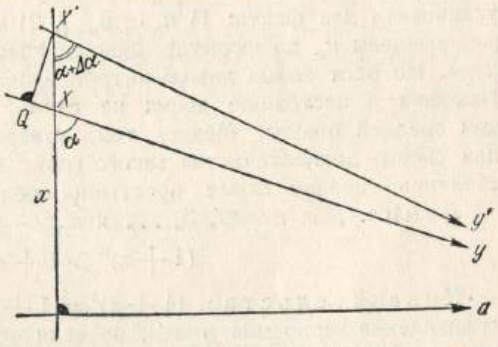
Приведем следующее определение.

Аналитическим синусом и аналитическим косинусом мы назовем функции $f(x)$ и $g(x)$, *определенные при всех действительных значениях аргумента и* 1) *удовлетворяющие уравнению* (III) *или* (IV) *или* (V) *при всех действительных значениях* x *и* y; 2) $f(x)$ $[g(x)]$ *положительна в некотором интервале*; 3) $f\left(\frac{\lambda}{2}\right)=1$; 4) $f(\lambda)=0$ $\left[g\left(\frac{\lambda}{2}\right)=0\right]$; 5) $f(0)=0$ [*при применении* (IV)], $g(0)=1$ [*при применении* (V)].

Функция $f(x)$ совпадает с $\sin\frac{\pi x}{\lambda}$, а $g(x)$ — с $\cos\frac{\pi x}{\lambda}$. Доказательство нетрудно провести.

КРАТКИЕ СООБЩЕНИЯ

1. ***В. А. Ефремович*** (*Москва*). **Наглядный вывод формулы Лобачевского для угла параллельности.** Приводимое здесь вычисление опирается на то свойство пространства Лобачевского, по которому малые фигуры в нем подчиняются соотношениям геометрии Евклида и при том тем точнее, чем меньше их линейные размеры (т. е. на свойство быть римановым многообразием). Кроме этого, будут привлечены лишь самые простые свойства параллелей (транзитивность параллелизма, монотонность искомой функции $\alpha(x)$, равенство $\alpha(0)=\frac{\pi}{2}$).



Рассмотрим в плоскости Лобачевского фиксированную ось (т. е. направленную прямую) a и точку X, скользящую по фиксированному перпендикуляру к a (см. рисунок). Обозначим через x расстояние от X до a и через $\alpha=\alpha(x)$ — угол параллельности, соответствующий точке $X=X(x)$. Проведем направленные прямые Xy и $X'y'$ параллельно a [здесь обозначено $X'=X(x+\Delta x)$], опустим перпендикуляр $X'Q$ на Xy. По транзитивности параллелизма из $X'y' \parallel a$, $Xy \parallel a$ следует $X'y' \parallel Xy$; значит, $\angle QX'y'=\alpha(QX')\approx\alpha(\Delta x\cdot\sin\alpha)$ [последнее равенство будет тем точнее, чем меньше Δx [1]]. Кроме того, $\angle QX'y'\approx$ $\approx(\alpha+\Delta\alpha)+\left(\frac{\pi}{2}-\alpha\right)=\frac{\pi}{2}+\Delta\alpha$ (опять по Евклиду!). Поэтому $\alpha(\Delta x\sin\alpha)\approx\frac{\pi}{2}+\Delta\alpha$; но $\frac{\pi}{2}=\alpha(0)$, следовательно, $\frac{\alpha(\Delta x\sin\alpha)-\alpha(0)}{\Delta x\sin\alpha}\approx$ $\approx\frac{\Delta\alpha}{\Delta x\sin\alpha}$, что после перехода к пределу [2] дает

$$\alpha'(0)=\frac{d\alpha}{dx}\frac{1}{\sin\alpha}.$$

[1] Для малого треугольника $XX'Q$ действует евклидова тригонометрия.

[2] Из предыдущей формулы следует, что если $\alpha(x)$ имеет конечную производную $\alpha'(x)$ в какой-нибудь точке, то это имеет место и в любой дру-

Обозначив отрицательную константу $\alpha'(0)$ через $-\frac{1}{k}$ и интегрируя последнее равенство, получим

$$\int \frac{d\alpha}{\sin\alpha} = -\frac{1}{k}x + C,$$

т. е. $\ln \operatorname{tg}\frac{\alpha}{2} = -\frac{x}{k}$ $\left[C=0 \text{ в силу равенства } \alpha(0)=\frac{\pi}{2}\right]$. Итак,

$$\operatorname{tg}\frac{\alpha(x)}{2} = e^{-\frac{x}{k}}.$$

2. ***В. И. Левин*** (*Москва*). **О числе *e*.** Для доказательства существования предела последовательности $u_n = \left(1+\frac{1}{n}\right)^n$ необходимо установить два факта: 1) $u_n < u_{n+1}$, 2) $u_n < A$. Это обычно делается разложением u_n по формуле бинома Ньютона с последующими оценками. Но этот вывод довольно громоздок; кроме того, формула бинома Ньютона в настоящее время не столь хорошо известна выпускникам средней школы. Между тем, утверждения 1) и 2) не требуют для своего доказательства такого громоздкого аппарата. Здесь можно обойтись гораздо более простыми средствами. Вполне достаточна

Лемма. Для $n=2, 3, \ldots$ и $\alpha > -1$, $\alpha \neq 0$

$$(1+\alpha)^n > 1+n\alpha.$$

Доказательство. $(1+\alpha)^2 = 1+2\alpha+\alpha^2 > 1+2\alpha$, так что утверждение верно для $n=2$; но если предположить, что для некоторого $n \geqslant 2$ $(1+\alpha)^n > 1+n\alpha$, то, умножая это неравенство на $1+\alpha > 0$, найдем, что и $(1+\alpha)^{n+1} > (1+n\alpha)(1+\alpha) = 1+(n+1)\alpha+n\alpha^2 > 1+(n+1)\alpha$.

Теперь доказательство утверждения 1), т. е. неравенства

$$u_{n+1} > u_n \quad \text{или} \quad \left(1+\frac{1}{n+1}\right)^{n+1} > \left(1+\frac{1}{n}\right)^n$$

$$\text{или} \quad \left(\frac{n+2}{n+1}\right)^{n+1} > \left(\frac{n+1}{n}\right)^n,$$

сводится к доказательству того, что $\left(\frac{n+2}{(n+1)^2}\right)^{n+1} > \frac{1}{(n+1)n^n}$; но по лемме

$$\left(\frac{n+2}{(n+1)^2}\right)^{n+1} = \frac{1}{n^{n+1}}\left[1-\frac{1}{(n+1)^2}\right]^{n+1} > \frac{1}{n^{n+1}}\left(1-(n+1)\frac{1}{(n+1)^2}\right) =$$

$$= \frac{1}{n^{n+1}}\left(1-\frac{1}{n+1}\right) = \frac{1}{n^n(n+1)}.$$

гой. Но по известной теореме Лебега монотонная функция дифференцируема почти всюду; значит, $\alpha'(x)$ конечна всюду.

Доказательство утверждения 2), как известно, вытекает из того, что последовательность чисел $v_n = \left(1+\frac{1}{n}\right)^{n+1}$ убывающая, откуда следует, что

$$u_n < v_n \leqslant v_1 = (1+1)^2 = 4$$

$$\left[\text{или, при } n \geqslant 5,\ u_n < v_n \leqslant v_5 = \left(1+\frac{1}{5}\right)^6 = \frac{46\,656}{15\,625} < 3\right].$$

Но неравенство

$$v_n > v_{n+1}$$

или

$$\left(1+\frac{1}{n}\right)^{n+1} > \left(1+\frac{1}{n+1}\right)^{n+2}$$

или

$$\left(\frac{n+1}{n}\right)^{n+1} > \left(\frac{n+2}{n+1}\right)^{n+2}$$

равносильно следующему:

$$\left(\frac{n+1}{n(n+2)}\right)^{n+2} > \frac{1}{n(n+1)^{n+1}},$$

а по лемме

$$\left(\frac{n+1}{n(n+2)}\right)^{n+2} = \frac{1}{(n+1)^{n+2}}\left(1+\frac{1}{n(n+2)}\right)^{n+2} >$$

$$\frac{1}{(n+1)^{n+2}}\left[1+(n+2)\frac{1}{n(n+2)}\right] = \frac{1}{(n+1)^{n+2}}\left(1+\frac{1}{n}\right) = \frac{1}{(n+1)^{n+1}n}.$$

3. ***Г. Мархасев*** (*Клязьма*). **Вычисление почти треугольных определителей.** *Почти треугольным* будем называть определитель n-го порядка $|a_{ij}|$, в котором $a_{ij}=0$, если $i > j+1$. У него (в отличие от треугольного определителя) помимо не равных нулю элементов, над главной диагональю и на ней имеются не равные нулю элементы в первой диагонали, прилегающей снизу к главной; назовем эту диагональ *поддиагональю*.

Для вычисления такого определителя можно дать правило, играющее в некотором смысле роль правила Саррюса для определителей 3-го порядка.

1. В «первую группу слагаемых» входит единственный член $a_{1n}\prod_{i=1}^{n-1} a_{i+1\,i}$. Ниже приводится определитель, в котором выделены

составляющие это произведение элементы:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1\,n-1} & \boxed{a_{1n}} \\ \boxed{a_{21}} & a_{22} & \dots & a_{2\,n-1} & a_{2n} \\ 0 & \boxed{a_{32}} & \dots & a_{3\,n-1} & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \boxed{a_{n\,n-1}} & a_{nn} \end{vmatrix}$$

2. «Вторую группу слагаемых» образуют произведения всех элементов поддиагонали, кроме k-го элемента $a_{k+1,k}$, $(k=1, 2, \dots, n-1)$, который заменяется k-м элементом первой строки a_{1k} и $(k+1)$-м элементом последнего столбца a_{k+1n}. Ниже выделены элементы, произведение которых дает k-е слагаемое второй группы

$$\begin{vmatrix} a_{11} & a_{12} & \dots & \boxed{a_{1k}} & \dots a_{1\,n-1} & a_{1n} \\ \boxed{a_{21}} & a_{22} & \dots & a_{2k} & \dots a_{2\,n-1} & a_{2n} \\ 0 & \boxed{a_{32}} & \dots & a_{3k} & \dots a_{3\,n-1} & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{k+1\,k} \dots a_{k+1\,n-1} & \boxed{a_{k+1\,n}} & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \dots & \boxed{a_{n\,n-1}} & a_{nn} \end{vmatrix}$$

3. Каждое слагаемое «третьей группы слагаемых» есть произведение всех элементов поддиагонали, кроме k-го $a_{k+1,k}$ $(k=1, 2, \dots, n-2)$ и l-го $a_{l+1,l}$ $(l=k+1, \dots, n-1)$. Они заменяются k-м элементом первой строки (a_{1k}), элементом, стоящим на пересечении $(k+1)$-й строки и l-го столбца $(a_{k+1,l})$ и $(l+1)$-м элементом последнего столбца $(a_{l+1,n})$. Ниже выделелы элементы, произведение которых дает слагаемое «третьей группы»:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & \boxed{a_{1k}} & \dots & a_{1l} & \dots & a_{1\,n-1} & \dots & a_{1n} \\ \boxed{a_{21}} & a_{22} & \dots & a_{2k} & \dots & a_{2l} & \dots & a_{2\,n-1} & \dots & a_{2n} \\ 0 & \boxed{a_{32}} & \dots & a_{3k} & \dots & a_{3l} & \dots & a_{3\,n-1} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{k+1\,k} & \dots & \boxed{a_{k+1\,l}} & \dots & a_{k+1\,n-1} & \dots & a_{k+1\,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & a_{l+1\,l} & \dots & a_{l+1\,n-1} & \dots & \boxed{a_{l+1\,n}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & \boxed{a_{n\,n-1}} & \dots & a_{nn} \end{vmatrix}$$

Аналогично получают следующие группы слагаемых. Последняя группа содержит единственное слагаемое — произведение элементов главной диагонали.

4. Правило знаков. Слагаемые s-й группы берутся со знаком плюс, если число $n-s$ четно и со знаком минус — в противном случае.

Справедливость указанного правила вычисления почти треугольного определения непосредственно следует из общего правила вычисления определителей n-го порядка.

Сами вычисления иногда удобнее производить в обратном порядке. Предоставляем читателю убедиться в этом на примере вычисления определителя

$$\Delta = \begin{vmatrix} x & a & 0 & \dots & 0 & 0 \\ -a & x & a & \dots & 0 & 0 \\ 0 & -a & x & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x & a \\ 0 & 0 & 0 & \dots & -a & x \end{vmatrix}$$

К вычислению почти треугольных определителей сводится решение системы линейных уравнений вида

$$\begin{aligned} b_{11}x_1 + b_{12}x_2 + \dots + b_{1k}x_k + \dots + b_{1n}x_n &= c_1, \\ b_{22}x_2 + \dots + b_{2k}x_k + \dots + b_{2n}x_n &= c_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots & \\ b_{kk}x_k + \dots + b_{kn}x_n &= c_k, \\ \dots\dots\dots\dots\dots\dots & \\ b_{nn}x_n &= c_n. \end{aligned}$$

Здесь

$$x_k=(-1)^{h-k-1}\frac{\Delta_{k+1}}{\Delta_n}\begin{vmatrix} b_{k\,k+1} & \dots & b_{kn} & c_k \\ b_{k+1\,k+1} & \dots & b_{k+1\,n} & c_{k+1} \\ 0 & \dots & b_{k+2\,n} & c_{k+2} \\ \dots & \dots & \dots & \dots \\ 0 & \dots & b_{nn} & c_n \end{vmatrix},$$

где

$$\Delta_i=\begin{vmatrix} b_{11} & b_{12} & \dots & b_{1i} \\ 0 & b_{22} & \dots & b_{2i} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{ii} \end{vmatrix}.$$

К почти треугольным определителям относятся континуанты [1]).

Через почти треугольные определители выражаются коэффициенты c_k частного $\sum c_i(z-z_0)^i$ от деления ряда $\sum a_i(z-z_0)^i$ на ряд $\sum b_i(z-z_0)^i$ (разумеется, предполагается сходимость двух последних рядов и отсутствие нулей у второго из них). В частности, соответствующие выражения можно найти и для чисел Бернулли [2]).

4. ***Е. Онофраш*** (E. Onofras, *Яссы, Румыния*). **Одно обобщение теоремы об углах равнобедренного треугольника.** **1.** Хорошо известно, что если в $\triangle ABC$ $\angle A=\angle B$, то и противолежащие этим углам стороны равны: $a=b$. Здесь мы докажем, что *если углы $\triangle ABC$ связаны отношением $A=nB$, где n — целое число, то стороны треугольника связаны зависимостью*

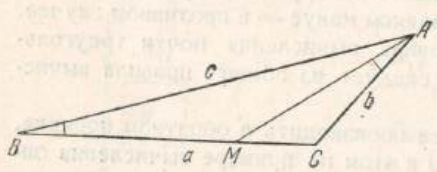


Рис. 1.

$$f_n(a, b, c)=0, \qquad (1)$$

где f_n — некоторый многочлен (порядка не выше чем $2n$), *который может быть определен для каждого n.* При $n=1$ (т. е. когда $A=B$) соотношение (1) имеет вид $a-b=0$.

Пусть $n=2$, т. е. $A=2B$ (рис. 1). Возьмем на стороне BC такую точку M, что $\angle CAM=\angle MBA$ (т. е. AM — биссектриса угла A). $\triangle ABC\backsim\triangle MAC$; следовательно, $\frac{c}{AM}=\frac{b}{MC}=\frac{a}{b}$, откуда $AM=\frac{bc}{a}$ и $MC=\frac{b^2}{a}$. А так как $\triangle AMB$ равнобедренный, то $BM-AM=0$, т. е. $\frac{a^2-b^2}{a}-\frac{bc}{a}=0$ или

$$f_2(a, b, c)=bc-a^2+b^2=0.$$

[1]) G. Kowalewski, Einführung in die Determinaten theorie, Lpz., 1909, стр. 152.

[2]) См. А. И. Маркушевич, Теория аналитических функций, М.—Л., Гостехиздат, 1950, стр. 257.

Это — искомое соотношение для $n=2$.

Предположим теперь, что соотношение $f_{n-1}(a, b, c)=0$, связывающее стороны $\triangle ABC$ при $A=(n-1)B$, нам известно, и пусть в изображенном на рис. 1 треугольнике $A=nB$ и по-прежнему $\angle CAM=\angle MBA$. В таком случае

$$\angle MAB=\angle A-\angle CAM=\angle A-\angle B=(n-1)\angle B.$$

По предположению, в треугольнике MAB имеет место известное нам соотношение $f_{n-1}(BM, AM, C)=0$. А так как $\triangle ABC\backsim\triangle MAC$, то снова $AM=\frac{bc}{a}$, $MC=\frac{b^2}{a}$ и $BM=a-MC=\frac{a^2-b^2}{a}$; наше соотношение переходит в

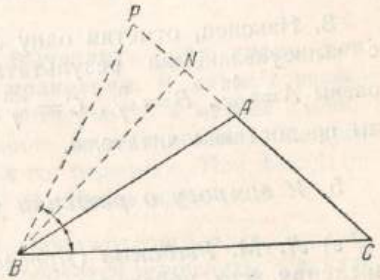


Рис. 2.

$$f_{n-1}\left(\frac{a^2-b^2}{a}, \frac{bc}{a}, c\right)=0.$$

Вследствие того, что многочлен $f_{n-1}(a, b, c)$ однородный [1]), можно написать

$$f_n(a, b, c)=f_{n-1}(a^2-b^2, bc, ac). \qquad (2)$$

Формула (2) позволяет рекуррентно вычислить многочлены f_n:

n	$f(a, b, c)$
1	$a-b$
2	$bc-(a^2-b^2)$
3	$bc^2-(a^2-b^2)(a-b)$
4	$a^2bc-(a^2-b^2-bc)^2(a^2-b^2+bc)$
5	$a^3bc^4-[(a+b)(a-b)^2][(a-b)(a+b)^2+bc^2]$
. . .	. . . . .

2. Уместно также поставить вопрос о нахождении отношения, существующего между сторонами треугольника, углы которого связаны соотношением вида $A=nB+mC$, где m и n — целые числа. Рассмотрим здесь для примера случай $n=2$, $m=1$.

Пусть в $\triangle ABC$ имеем $A=2B+C$ (рис. 2). Так как, кроме того, $A+B+C=180°$, то $A=90°+\frac{B}{2}>90°$.

Пусть N — проекция B на прямую AC. Возьмем на продолжении отрезка CA точку P так, что $NP=NA$. Тогда $\angle APB=\angle PAB=$

[1]) См., например, Д. И. Перепелкин, Курс элементарной геометрии, ч. 1, М.—Л., 1948, § 47.

$= 180° - \angle A$, и значит, $\angle A - \angle APB = 2\left(90° + \frac{B}{2}\right) - 180° = B$, т. е. BA есть биссектриса угла PBC. Отсюда имеем $\frac{PA}{AC} = \frac{PB}{BC}$, $\frac{PA}{AC} = \frac{AB}{BC}$, $\frac{2AN}{b} = \frac{c}{a}$, или $2a\,AN = bc$. Но так как, кроме того $a^2 = b^2 + c^2 + 2b\,AN$, то, в силу последнего равенства,

$$a(b^2 + c^2 - a^2) + b^2c = 0. \qquad (3)$$

Это и есть искомое соотношение.

3. Наконец, отметим одну любопытную теорему, тесно связанную с вышеуказанными результатами: *если углы треугольника* ABC равны $A = \frac{4\pi}{7}$, $B = \frac{\pi}{7}$, $C = \frac{2\pi}{7}$, то $\frac{1}{a} + \frac{1}{c} = \frac{1}{b}$. Доказательство ее мы предоставим читателю.

5. *К вопросу о решении уравнений методом итераций.*

а) ***Л. М. Рыбаков*** (*Ярославль*). **Метод последовательного вычисления всех действительных корней уравнения.** В этом методе заданное уравнение $f(x) = 0$ записывается в форме $x = \varphi(x)$, где обычно в качестве $\varphi(x)$ берут $x + \frac{f(x)}{M}$, выбирая M таким образом, чтобы обеспечить сходимость итерационного процесса к отделенному предварительно корню. Ниже показывается, что если в приведенной итерационной формуле заменить $f(x)$ на $|f(x)|$, то можно обеспечить выбор M таким образом, чтобы полученная итерационная формула позволяла вычислить автоматически последовательно все корни $f(x)$, находящиеся в заданном интервале $[a, b]$ в порядке их возрастания, исходя из начального значения $x_0 = 0$.

Пусть дано уравнение $f(x) = 0$, где $f(x)$ — функция действительного аргумента x, существенно от него зависящая и обладающая непрерывной производной. Пусть, далее, $[a, b]$ — любой отрезок, содержащий по крайней мере один корень $f(x)$, и M — действительное число, выбранное таким образом, чтобы для

$$a \leqslant x \leqslant b \quad \text{было} \quad M > |f'(x)|.$$

Рассмотрим итерационную формулу $x = \varphi(x)$, полагая $\varphi(x) = x + \frac{|f(x)|}{M}$. При таком выборе M и определении $\varphi(x)$ для всех $a \leqslant x \leqslant b$, кроме корней $f(x)$,

$$\varphi(x) > x \ \text{(а)} \quad \text{и} \quad \varphi'(x) > 0 \ \text{(б)}.$$

Пусть теперь $f(\bar{x}) = 0$, $a < \bar{x} \leqslant b$ и $a \leqslant x_1 < \bar{x}$, причем $f(x) \neq 0$,

если $x_1 < x < \bar{x}$. Тогда, принимая x_1 за приближенное значение $\bar{x}$, вычислим по формуле $x = \varphi(x)$ значение $x_2 = \varphi(x_1)$. Согласно (а), $x_2 > x_1$. С другой стороны, по (б), $x_2 = \varphi(x_1) < \varphi(\bar{x}) = \bar{x}$, т. е. $x_1 < x_2 < \bar{x}$. Формула $x = \varphi(x)$, таким образом, позволяет строить путем последовательных итераций, исходя из любой точки x_1 отрезка, последовательность, сходящуюся к ближайшему справа корню $f(x)$ этого отрезка,

Если $\bar{x} - x_k < \varepsilon$, где ε — требуемая точность, то число $x_k + \varepsilon > \bar{x}$ будет служить начальным значением для вычисления следующего за $\bar{x}$ (в порядке возрастания) корня.

б) ***В. Л. Загускин*** (*Ярославль*). **Вычислительная схема для ускоренного вычисления всех корней уравнения.** В связи с развитием вычислительной техники возникла потребность в методе вычисления корней, обладающем однообразной вычислительной схемой. Здесь излагается такого рода «метод k-го порядка». При $k = 1$ он совпадает с методом, предложенным Л. М. Рыбаковым.

1. Пусть задано уравнение $f(x) = 0$, корни которого $\alpha_1, \alpha_2, \ldots, \alpha_n$ (предполагаем, что их конечное число) требуется вычислить. Сначала следует определить [например, с помощью метода Маклорена, если уравнение алгебраическое[1])] границы корней $a \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_n \leqslant b$. Если это затруднительно, то можно применить k-метод дважды — для отыскания на отрезке $[-1, +1]$ всех корней функций $f(x)$ и $f\left(\frac{1}{x}\right)$, что определит, очевидно, все корни данного уравнения. Кроме того, следует, выбрав k, подобрать число C_k, удовлетворяющее условию

$$C_k \geqslant \frac{1}{k!} \max |f^{(k)}(x)|, \quad a \leqslant x \leqslant b.$$

Следует отметить, что неточность этой оценки, так же как и неточность в определении границ a и b, влияет на объем последующих вычислений в k-методе при $k \geqslant 2$ не очень сильно.

2. Вычисления в k-методе состоят в построении некоторой монотонной последовательности чисел $x_1 < x_2 < \ldots < x_m$; построение начинается с $x_1 = a$ и заканчивается при получении $x_m \geqslant b$.

Пусть x_i уже вычислено. Для нахождения x_{i+1} составим вспомогательное уравнение

$$F(\lambda) = f(x_i) + \frac{f'(x_i)}{1!}\lambda + \ldots + \frac{f^{(k-1)}(x_i)}{(k-1)!}\lambda^{k-1} - [\operatorname{sign} f(x_i)]\, C_k \lambda^k = 0. \quad (1)$$

Вычислив наименьший положительный корень λ_i этого уравнения,

[1]) См. В. Л. Загускин, Справочник по численным методам решения алгебраических и трансцендентных уравнений, М., 1960, стр. 37.

найдем x_{i+1} по формуле $x_{i+1} = x_i + \lambda_i + \varepsilon$, где ε — заданная точность, с которой следует найти корни уравнения $f(x) = 0$ [1]).

Для практического применения больше всего подходит, по-видимому, случай $k = 2$. Тогда уравнение (1) имеет только один положительный корень (так как знак коэффициента при λ^2 противоположен знаку свободного члена), и мы получаем следующую вычислительную формулу:

$$x_{i+1} = x_i + \frac{(\operatorname{sign} a_i)\, b_i + \sqrt{b_i^2 + |a_i|\, C_2}}{C_2} + \varepsilon, \qquad (2)$$

где

$$a_i = f(x_i), \quad b_i = \frac{1}{2} f'(x_i).$$

3. Из монотонности (и неограниченности) последовательности $\{x_i\}$ следует, что для любого корня α_s найдутся такие два последовательных значения x_i и x_{i+1}, что $x_i < \alpha_s \leqslant x_{i+1}$. Предположим для определенности, что $f(x_i) > 0$. Рассмотрим тогда функцию $\varphi(x) = f(x) - F(x - x_i)$. В силу определения $F(\lambda)$ функция $\varphi(x)$ обладает тем свойством, что на отрезке $[x_i, b]$ ее k-я производная $\varphi^{(k)}(x)$ неотрицательна. Следовательно, $\varphi^{(k-1)}(x)$ есть монотонно возрастающая функция на этом отрезке. Но учитывая, что $\varphi^{(k-1)}(x_i) = 0$, приходим к выводу, что и $\varphi^{(k-1)}(x)$ неотрицательна на отрезке $[x_i, b]$. Применяя затем метод индукции, мы получим, что и сама функция $\varphi(x)$ неотрицательна на рассматриваемом отрезке [2]). Отсюда, очевидно, следует, что $x_i + \lambda_i \leqslant \alpha_s$, что вместе с неравенством $x_i < \alpha_s \leqslant x_{i+1}$ приводит нас к выводу, что x_{i+1} либо совпадает с α_s, либо отличается от него не больше чем на ε.

Таким образом, в последовательности $\{x_i\}$ содержатся (с точностью до ε) все искомые корни данного уравнения. Определение того, какие из чисел x_i являются корнями, а какие нет, несложно, так как для каждого x_i вычисляется $f(x_i)$.

4. Пример. Вычислить с точностью $\varepsilon = 0{,}0001$ корни уравнения

$$f(x) = \sin x - 0{,}5x^2 - 0{,}1 = 0. \qquad (3)$$

Простое исследование этого уравнения показывает, что корни его лежат на отрезке [0; 1,35], а максимум второй производной не превышает 2, так что можно взять $C_2 = 1$.

Взяв $x_0 = 0$, найдем по формуле (2) $a_0 = -1$; $b_0 = 0{,}5$ откуда $\lambda_0 = (-1)\cdot 0{,}5 + \sqrt{0{,}25 + 0{,}1} = 0{,}0916$; $x_1 = x_0 + \lambda + \varepsilon = 0{,}0917$. Дальнейшие вычисления дают: $x_2 = 0{,}1056$, $x_3 = 0{,}1059$, $x_4 = 0{,}9945$, $x_5 = 1{,}2755$, $x_6 = 1{,}3126$, $x_7 = 1{,}3172$, $x_8 = 1{,}3177$, $x_9 = 1{,}5407 > b$, и вычисления заканчиваются. При этом $f(x_8) = 0{,}0001$, $f(x_8) = -0{,}0001$, что свидетельствует о том, что x_2 и x_8 являются корнями уравнения (3). Интересно заметить, что разности $x_4 - x_3 = 0{,}8886$ и $x_9 - x_8 = 0{,}2230$ значительно превосходят разности $x_3 - x_2 = 0{,}0003$ и $x_8 - x_7 = 0{,}0005$, что вообще характерно для процесса второго порядка, обладающего способностью быстро проходить свободные от корней интервалы. Убывание разностей:

$$x_2 - x_0 = 0{,}1059, \quad x_3 - x_1 = 0{,}0142, \quad x_3 - x_2 = 0{,}0003$$

и

$$x_4 - x_4 = 0{,}3232, \quad x_8 - x_5 = 0{,}0422, \quad x_8 - x_6 = 0{,}0051, \quad x_8 - x_7 = 0{,}0005$$

свидетельствует о быстрой сходимости метода.

5. В порядке эксперимента проводились вычисления на машине «Урал», работающей со скоростью 100 операций в секунду. Было задано уравнение

$$x^4 - 0{,}234375\, x^2 - 0{,}01953125\, x + 0{,}005859375 = 0$$

и значения $a = -0{,}75$, $b = 1$, $C = 8$, $\varepsilon = 2^{-20}$. В течение $2\frac{1}{2}$ минут машина прошла заданный интервал и выдала следующие значения:

$$x_1 = -0{,}374999998, \quad x_2 = -0{,}249999994,$$
$$x_3 = 0{,}125000002, \quad x_4 = 0{,}500000002$$

(точные значения: $-0{,}375$; $-0{,}25$, $0{,}125$, $0{,}5$). Программа получилась достаточно простой.

Вполне возможно создание стандартной подпрограммы для вычисления всех вещественных корней многочленов методом второго порядка.

---

[1]) Геометрически это сводится к построению параболы k-го порядка, имеющей с графиком функции $f(x)$ касание $(k-1)$-го порядка. В качестве x_{i+1} берется точка, расположенная вправо на расстоянии ε от первой (считая вправо от x_i) точки пересечения параболы с осью Ox.

[2]) Геометрически это означает, что график апроксимирующей параболы $F(x - x_i)$ повсюду в пределах отрезка $[x_i, b]$ проходит не выше графика $f(x)$.

От редакции. В 4 выпуске «Математического просвещения» была опубликована в разделе кратких научно-методических сообщений заметка А. В. Кужеля «Исследование корней кубического уравнения без использования формул Кардано». Б. А. Ларионов обратил внимание редакции на то, что аналогичный результат содержится в его статье (Б. А. Ларионов, Геометрический смысл дискриминанта алгебраического уравнения n-й степени с реальными коэффициентами, Труды Института математики и механики, АН УЗССР **18** (1956), стр. 115—122).

ОКОНЧАНИЕ ТЕКСТА СТР. 230

более или менее обстоятельные толкования латинских терминов, относящихся к традиционному квадривиуму: арифметике, геометрии, астрономии и музыке. На фронтисписе так и указано: «Научный квадривиум или Математический лексикон». Формат ее — малый, in 4°, в ней 1070 страниц текста с весьма немногими чертежами.

Книги Карамуеля и Озанама построены тематически — по отделам. Наиболее богато содержанием сочинение Карамуеля: в нем 40 отделов, начиная с арифметики и кончая «планетарной гипотезой» (которая охватывает три отдела). Второй том целиком посвящен «новой науке», куда входят логарифмика, тригонометрия и многие другие. Упомянем среди них отдел «Kybeja de ludis» (т. е. об играх).

Фронтиспис и титульный лист «Двуглавой науки» Карамуеля

«Двуглавая наука» содержит 137 математических и иных таблиц (среди них логарифмические) и множество искусно награвированных на меди чертежей, составляющих атлас на 32 листах. Хотя в этой книге много курьезного и наивного, подчеркивающего диллетантский облик автора, всё же это настоящая энциклопедия точных наук и их важнейших приложений, содержащая огромный фактический и справочный материал.

Книга Озанама напоминает нам книгу Карамуеля (хотя мы и не можем утверждать здесь, что Озанам ее использовал); она меньше по объему и в ней нет таких курьезов. Любопытно отметить, что если Карамуель подчеркивает разделение науки на старую и новую, то Озанам исходит из подразделения на чистую и прикладную. Точнее говоря, он различает простую математику (арифметика, алгебра, геометрия), где речь идет о количествах, абстрагированных от чувственных объектов, и смешанную математику (космография, механика, оптика, музыка), изучающую количества в связи с чувственными объектами.

А. Маркушевич

IV. НАУЧНАЯ И ПЕДАГОГИЧЕСКАЯ ХРОНИКА

КРУЖОК ПО МАТЕМАТИЧЕСКОМУ АНАЛИЗУ В ГРОДНЕНСКОМ ПЕДАГОГИЧЕСКОМ ИНСТИТУТЕ

Я. И. Ривкинд

(Гродно)

Кружок по математическому анализу в Гродненском пединституте существует уже больше десяти лет: он был основан в 1949 г. В настоящей заметке будет говориться только о тематике этого кружка, так как именно она определила работу кружка в целом. В тематике отразились вкусы и интересы членов кружка — студентов института, наиболее одаренных и интересующихся научной работой.

С темами происходил своеобразный «естественный отбор». Некоторые темы оказались жизнеспособными и успешно выдерживали испытание времени (хотя иногда и не было особых успехов в их разработке), другие же быстро отмирали. Эта участь в первую очередь постигла темы, мало связанные с обиходным математическим материалом пединститута и требующие знакомства с обширным новым материалом. (Такие темы не будут приведены в этой нашей статье.)

Большую роль в работе кружка сыграла вышедшая в 1949 г. монография И. П. Натансона по конструктивной теории функций [1]). Интересно и доступно написанная, она вместе с такими классическими произведениями, как «Задачи и теоремы из анализа» Полиа и Сеге, «Неравенства» Харди, Литльвуда и Полиа, «Расходящиеся ряды» Харди, стала настольной книгой для членов кружка. Из некоторых разделов этих книг образовался определенный минимум сведений, которым должен был овладеть каждый член кружка, независимо от проводимой им самостоятельной работы.

Первое время в работе кружка доминировала тематика по *теории приближения функций*; затем, вполне естественно, появилось стремление к расширению тематики. Это расширение тематики имело своеобразную окраску — проявился отход от специализированных

[1]) И. П. Натансон, Конструктивная теория функций, М.—Л., 1949.

математических дисциплин, преобладающими стали темы по общему анализу, примерно такие темы, которые в Реферативном журнале «Математика» фигурируют в разделе «Анализ (другие вопросы)».

Каждый член кружка работал по индивидуальному плану, причем ему предоставлялась полная свобода действий. На заседаниях кружка заслушивались сообщения о полученных результатах или обсуждались новые темы. Иногда члены кружка или его руководитель читали лекции по разным вопросам, в большей или меньшей мере связанным с тематикой работ кружка.

Ниже приводится ряд задач, над которыми работали участники кружка за 1949—1957 гг. Некоторые из них были успешно решены и опубликованы в «Ученых записках» института [1]), печатались стеклографически или оставались в рукописи. Некоторые задачи не поддались до сих пор решению, они оказались непосильными для студентов, которые их выбрали.

Темы работ выдвигались руководителем кружка; весьма возможно, что среди приведенных ниже задач имеются и такие, которые уже были рассмотрены раньше.

Задачи приводятся по темам. Это иногда нарушает хронологический порядок, но более удобно для обозрения. Заметим здесь только, что задачи раздела III — первые, которые рассматривались в кружке.

I. ТЕОРИЯ ИТЕРАЦИЙ

С тематикой по теории итераций мы столкнулись совершенно естественно при рассмотрении некоторых элементарных вопросов анализа.

Уже в самых началах анализа некоторые классы функций определяются при помощи простых функциональных уравнений; так, например, класс периодических функций с периодом a определяется функциональным уравнением $f(x+a)=f(x)$, а уравнение $f(-x)=f(x)$ определяет класс четных функций. Оба уравнения являются весьма частными случаями уравнения

$$f(\varphi(x))=f(x), \tag{1}$$

в котором $\varphi(x)$ — заданная монотонная функция вещественного переменного x. Таким образом возникла следующая задача.

Задача 1. *Выяснить, каким условиям должна удовлетворять функция $\varphi(x)$ для того, чтобы уравнение* (1) *имело нетривиальные* (т. е. отличные от постоянной) *решения, и изучить те классы функций, которые могут быть решениями этого уравнения.* Эту задачу решила Е. Ю. Захарчук [1]). Она показала, что основное значение в теории уравнения (1) имеет поведение бесконечной в обе стороны последовательности итераций функции $\varphi(x)$

$$\varphi_{-1}\varphi_{-1}(x),\ \varphi_{-1}(x),\ \varphi(x),\ \varphi\varphi(x),\ \varphi\varphi\varphi(x),\ \ldots, \tag{2}$$

где $\varphi_{-1}(x)$ — функция, обратная к функции $\varphi(x)$.

Рассмотрение последовательности (2) позволило, например, выяснить, что в случае возрастающей функции $\varphi(x)$, для которой уравнение $\varphi(x)=x$ имеет действительные корни, уравнение (1) имеет только тривиальные решения; справедливо также и обратное утверждение.

Изучение последовательности (2) итераций, естественно, наталкивало на мысль обобщить итерационный процесс и рассмотреть итерации, построенные на основе последовательного применения n функций.

Задача 2. *Пусть дано n непрерывных и монотонных функций*

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x), \tag{3}$$

определенных на всей действительной оси. Для каждой точки x_0 действительной оси образуем следующую последовательность точек:

$$\left.\begin{aligned} &x_1=\varphi_1(x_0),\ x_2=\varphi_2(x_1),\ \ldots,\ x_n=\varphi_n(x_{n-1}),\ x_{n+1}=\varphi_1(x_n),\\ &x_{n+2}=\varphi_2(x_{n+1}),\ \ldots,\ x_{sn+k}=\varphi_k(x_{sn+k-1})\ (k\leqslant n). \end{aligned}\right\} \tag{4}$$

Требуется изучить последовательность итераций

$$x_0,\ x_1,\ x_2,\ \ldots,\ x_n,\ x_{n+1},\ \ldots \tag{5}$$

Этой задаче посвящена работа Н. С. Щедровой «Последовательности итераций из n функций» [2]). Последовательность точек $\{x_k\}$ $(k=0,\ 1,\ \ldots,\ n-1)$ естественно назвать *циклом 1-го порядка* системы функций (3), если $x_n=x_0$. Аналогично этому последовательность точек $x_0,\ x_1,\ \ldots,\ x_n,\ \ldots,\ x_{2n-1}$ можно назвать *циклом 2-го порядка* системы (3), если $x_n\neq x_0$, но $x_{2n}=x_0$. Казалось бы, можно аналогично определить циклы 3-го и высших порядков. Однако на самом деле *система функций* (3) *не может иметь существенных циклов порядка выше чем 2* [3]).

Н. С. Щедрова установила ряд теорем, из которых выясняется, при каких условиях последовательность (5) имеет предельные точки; оказалось, что если эта последовательность имеет предельные точки,

---

[1]) «Ученые записки Гродненского педагогического института» (в дальнейшем цитируются: У. З.), вып. 1 — 1955 г., вып. 2 — 1957 г., вып. 3 — 1958 г.

[1]) Е. Ю. Захарчук, О решениях функционального уравнения $f(\varphi(x))=f(x)$, У. З., вып. 1.

[2]) У. З., вып. 2.

[3]) По существу это утверждение относится к свойствам одной монотонной функции $\varphi(x)=\varphi_n(\varphi_{n-1}(\ldots\varphi_1(x)\ldots))$.

то ими являются элементы некоторых циклов 1-го или 2-го порядка системы функций (3).

В задачах 1 и 2 рассматривались итерации функций одной переменной. Заслуживает внимания изучение итерационного процесса для функции двух или большего числа переменных. Для случая функции двух переменных задача сводится к рассмотрению двумерной последовательности итераций (x_n, y_n), где

$$x_{n+1} = \varphi(x_n, y_n), \quad y_{n+1} = \psi(x_n, y_n),$$

причем функции $\varphi(x, y)$ и $\psi(x, y)$ удовлетворяют некоторым условиям.

Однако в такой общей постановке задача, по-видимому, довольно трудна. В плане изучения двумерных итераций рассматривался следующий частный случай.

Задача 3. *Изучить двумерный итерационный процесс, заданный условиями:*

$$x_{n+1} = a_{11}x_n + a_{12}y_n + a_{13},$$
$$y_{n+1} = a_{21}x_n + a_{22}y_n + a_{23},$$

причем

$$\frac{a_{11}-1}{a_{21}} = \frac{a_{12}}{a_{22}-1} = \frac{a_{13}}{a_{23}}.$$

Здесь мы имеем дело с итерацией перспективно-аффинных преобразований.

Этот частный случай двумерной последовательности итерации также был подвергнут подробному изучению Н. С. Щедровой.

II. ТЕОРИЯ РЯДОВ И ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Тематика теории рядов и последовательностей всегда была привлекательна для студентов. Ниже указаны некоторые примыкающие сюда задачи, рассмотренные в кружке.

1. Ряды. Задача 4. *Обозначим через $P[\{c_n\}]$ — признак сходимости Куммера*[1]. *Признак этот зависит от счетного множества параметров $c_1, c_2, \ldots, c_n, \ldots$. Требуется при помощи надлежащего выбора параметров построить на основе признака Куммера «цепочку» признаков сходимости, т. е. последовательность признаков, в которой каждый следующий признак сходимости сильнее предыдущего.*

[1]) См. Г. М. Фихтенгольц, Курс дифференциального и интегрального исчисления, т. II, М.—Л., 1948, стр. 327.

Эта задача была решена Э. В. Стрелецким[1]). Положим

$$s_n^{(1)} = \sum_{k=1}^{n} \frac{1}{c_k}, \quad s_n^{(2)} = \sum_{k=1}^{n} \frac{1}{c_k s_k^{(1)}}, \quad s_n^{(3)} = \sum_{k=1}^{n} \frac{1}{c_k s_k^{(2)}}, \ldots;$$

тогда последовательность признаков сходимости

$$P[\{c_n\}],\ P[\{c_n s_n^{(1)}\}],\ P[\{c_n s_n^{(1)} s_n^{(2)}\}],\ \ldots,\ P[\{c_n s_n^{(1)} \ldots s_n^{(k)}\}] \ldots \quad (6)$$

является «цепочкой» критериев сходимости, удовлетворяющей поставленным условиям. В частности, если положить $c_n = 1$, то мы получим «цепочку», в которой на первом месте стоит признак Даламбера, на втором признак Раабе, а дальше следуют признаки сходимости, эквивалентные логарифмическим признакам сходимости рядов.

Задача 5. В теории рядов с положительными членами большую роль играет известная теорема о рядах с асимптотически одинаковыми членами[2]). Однако, как легко убедиться на простых примерах, эта теорема перестает быть верной при отказе от требования положительности членов ряда. Требуется так сузить понятие рядов с асимптотически одинаковыми членами, чтобы аналогичная теорема оставалась в силе и для рядов с произвольными знаками членов.

Этот вопрос был подвергнут изучению К. К. Искрой, который доказал следующую теорему[3]).

Если для рядов $\sum_{n=1}^{\infty} u_n$ и $\sum_{n=1}^{\infty} v_n$ выполняются условия

$$\sum_{n=1}^{\infty} \left| \frac{u_n}{v_n} - \frac{u_{n+1}}{v_{n+1}} \right| < \infty \quad \text{и} \quad \sum_{n=1}^{\infty} \left| \frac{v_n}{u_n} - \frac{v_{n+1}}{u_{n+1}} \right| < \infty, \quad (7)$$

то оба ряда или одновременно сходятся или одновременно расходятся.

Условия (7) являются безусловно более сильными, чем условие

$$\lim_{n\to\infty} \frac{u_n}{v_n} = k \qquad (k \neq 0, \quad k \neq \infty), \quad (8)$$

так как из них не только следует существование отличного от 0 и от ∞ предела $\lim_{n\to\infty} \frac{u_n}{v_n} = k$, но накладываются некоторые ограничения на скорость сходимости последовательности $\left\{\frac{u_n}{v_n}\right\}$ к пределу k.

[1]) Э. В. Стрелецкий, Цепь признаков сходимости для рядов с положительными членами, У. З., вып. 1.
[2]) Г. М. Фихтенгольц, цит. книга, стр. 313, теорема 2.
[3]) К. К. Искра, Некоторые теоремы о рядах, У. З., вып. 1.

В этой же работе интегральный признак сходимости рядов был перенесен и на ряды с немонотонно стремящимися к нулю членами.

Задача 6. Дан ряд с положительными членами $\sum_{k=1}^{\infty} a_k$. Положим

$$v_k^{(0)} = a_k, \quad v_n^{(s)} = \sum_{k=n+1}^{\infty} v_k^{(s-1)}, \qquad s = 1, 2, \ldots$$

Требуется:

а) *изучить класс рядов, для которых*

$$\sum_{k=1}^{\infty} v_k^{(s)} < \infty \quad \text{при} \quad s \leqslant l \quad \text{и} \quad \sum_{k=1}^{\infty} v_k^{(s)} = \infty \quad \text{при} \quad s > l,$$

и б) *выяснить, для каких рядов* $\sum_{k=1}^{n} v_k^{(s)} < \infty$ *при всяком* s. Это исследование осталось незавершенным.

Укажем еще на относящуюся к теории рядов работу Э. В. Стрелецкого [1]), в которой был построен пример одного интересного ряда.

2. Равномерно распределенные последовательности. К теории равномерно распределенных последовательностей [2]) относятся следующие две задачи.

Задача 7. Бесконечная треугольная матрица (M), n-я строчка которой удовлетворяет условию

$$0 \leqslant x_1^{(n)} \leqslant x_2^{(n)} \leqslant \ldots \leqslant x_{n-1}^{(n)} \leqslant x_n^{(n)} \leqslant 1,$$

называется *равномерно распределенной* на отрезке $[0, 1]$, если для всякой функции $f(x)$, непрерывной на отрезке $[0, 1]$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(x_k^{(n)}) = \int_0^1 f(x)\, dx.$$

Требуется найти достаточные условия для того, чтобы матрица была равномерно распределенной.

[1]) Э. В. Стрелецкий, Пример ряда суммируемого (A) и несуммируемого (C, p), $p > 0$, У. З., вып. 1.

[2]) Последовательность $\{x_i\}$ называется *равномерно распределенной на отрезке* $[a, b]$, если для любого отрезка длины l, лежащего внутри $[a, b]$, имеет место соотношение

$$\frac{1}{l} \lim_{N\to\infty} \frac{k}{N} = \text{const},$$

где k — число членов последовательности с номерами, не превосходящими N, лежащих внутри отрезка длины l.

Н. П. Виноградовым доказана следующая теорема: *если*

$$\delta_n = \sum_{k=1}^{n} \left(\Delta x_k^{(n)} - \frac{1}{n} \right)^2 = o\left(\frac{1}{n}\right),$$

то матрица (M) *равномерно распределена.*

Аппарат равномерно распределенных матриц может быть успешно использован для изучения равномерно распределенных последовательностей на отрезке $[0, 1]$. Действительно, с каждой равномерно распределенной последовательностью $\{x_i\}$ на отрезке $[0, 1]$ связана некоторая равномерно распределенная матрица, n-я строчка которой получается, если элементы последовательности $x_1, x_2, \ldots, x_n$ упорядочить по их величине.

Н. П. Виноградов изучал некоторые свойства равномерно распределенных последовательностей на основе изучения связанных с ними матриц; это исследование, однако, осталось незавершенным.

Задача 8. Рассмотрим следующее обобщение понятия последовательности, равномерно распределенной на отрезке $[0, 1]$. Последовательность $\{x_n\}$ назовем *равномерно распределенной в среднем относительно класса функций* K, если для всякой функции $f(x) \in K$, причем $K \subset L_2$,

$$\lim_{n\to\infty} \left\{ \int_0^1 \left[\frac{1}{n} \sum_{k=1}^{n} f(x_k + t) - \int_0^1 f(t)\, dt \right]^2 dt \right\}^{\frac{1}{2}} = 0.$$

Если $K = L^2$ [1]), то мы говорим просто о равномерной распределенности в среднем.

Спрашивается, *каким условиям должен удовлетворять класс* K, *чтобы из равномерной распределенности в среднем последовательности* $\{x_k\}$ *относительно класса* K *следовала ее равномерная распределенность на отрезке* $[0, 1]$.

На этот вопрос ответ дал Н. И. Бриш. Выяснилось, что для этого достаточно, чтобы класс K был плотен в L_2 [2]).

Понятие равномерного распределения в среднем является содержательным и заслуживает изучения.

[1]) Множество функций $x(t)$, определенных и измеримых на отрезке $[0, 1]$, для которых

$$\int_0^1 |x(t)|^2\, dt < \infty,$$

(интеграл понимается в смысле Лебега), называется *пространством Гильберта* L_2.

[2]) Н. И. Бриш, Последовательности, равномерно распределенные в среднем на отрезке $[0, 1]$, У. З., вып. 1.

III. ТЕОРИЯ ПРИБЛИЖЕНИЯ ФУНКЦИЙ

1. Полиномы С. Н. Бернштейна. По теории этих полиномов изучались следующие два вопроса.

Задача 9. Рассматривается класс функций $f(x)$, имеющих на отрезке $[0, 1]$ первую производную, удовлетворяющую условию Липшица [1]). Производную $f'(x)$ можно с любой степенью точности аппроксимировать полиномами С. Н. Бернштейна

$$B_n(f', x) = \sum_{k=0}^{n} f'\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k},$$

а также, как это следует из известных теорем [2]), производными $B_n'(f, x)$ полиномов

$$B_n(f, x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}.$$

Положим

$$\lambda_n^{(1)} = \sup_{0 \leqslant x \leqslant 1} |B_n(f', x) - f'(x)| \quad \text{и} \quad \tilde{\lambda}_n^{(1)} = \sup_{0 \leqslant x \leqslant 1} |B_n'(f, x) - f'(x)|.$$

Естественно возникает вопрос: существуют ли такие функции $f(x)$, непрерывные на отрезке $[0, 1]$ и обладающие производной $f'(x)$ такой, что $f'(x) \in \mathrm{Lip}_\alpha M$, для которых имеет место неравенство

$$\tilde{\lambda}_n^{(1)} < \lambda_n^{(1)}. \tag{9}$$

Иначе говоря, *существуют ли такие функции в классе* $\mathrm{Lip}_\alpha M$, *которые при данном* n *точнее аппроксимируются производными полиномов С. Н. Бернштейна их примитивных, чем полиномами С. Н. Бернштейна самих этих функций.*

Для величины $\lambda_n^{(1)}$ имеет место неулучшаемая в смысле порядка оценка [3])

$$\lambda_n^{(1)} \leqslant \frac{3M}{2} \cdot \frac{1}{n^{\alpha/2}}; \tag{10}$$

поэтому для решения поставленной задачи усилия были направлены на получение возможно более точной оценки величин $\tilde{\lambda}_n^{(1)}$. Такую оценку получил Г. В. Жидков [4]), доказав, что

$$\tilde{\lambda}_n^{(1)} \leqslant \left(\frac{3}{4}\right)^{\alpha/2} M \frac{1}{n^{\alpha/2}}. \tag{11}$$

[1]) См. указанную на стр. 267 книгу И. П. Натансона, стр. 22.
[2]) См. там же, стр. 250, 251.
[3]) См. там же, стр. 246.
[4]) Г. В. Жидков, Замечание о полиномах акад. С. Н. Бернштейна, У. З., вып. 1.

Так как оценка (10) не может быть улучшена и

$$\left(\frac{3}{4}\right)^{\alpha/2} < \frac{3}{2} \quad \text{при} \quad 0 < \alpha < 1,$$

то вероятно, что ответ на поставленный выше вопрос является положительным.

Затем намечалось изучение класса тех функций, для которых выполняется неравенство (9). Эта задача оказалась, однако, превосходящей возможности студентов. Трудным оказалось также обобщение неравенства (11) для величин

$$\tilde{\lambda}_n^{(s)} = \sup_{0 \leqslant x \leqslant 1} |B_n^{(s)}(f, x) - f^{(s)}(x)|,$$

где функция $f(x)$ имеет s-ю производную $f^{(s)}(x)$, принадлежащую классу $\mathrm{Lip}_1 M$.

В этом направлении были сделаны некоторые заслуживающие внимания попытки, но всё же обозримые результаты до сих пор не получены.

Задача 10. Рассматриваются сумматорные полиномы

$$B_n^*(f, x) = \sum_{k=0}^{n} f(\xi_n^{(k)}) C_n^k x^k (1-x)^{n-k}, \tag{12}$$

причем

$$0 = \xi_n^0 < \xi_n^{(1)} < \ldots < \xi_n^{(n-1)} < \xi_n^{(n)} = 1. \tag{13}$$

Эти полиномы являются обобщением полиномов С. Н. Бернштейна. Естественно изучить, при каких распределениях узлов (13) полиномы (12) обладают некоторыми основными свойствами, присущими полиномам С. Н. Бернштейна. В частности, имелось в виду ответить на следующие вопросы:

α) *При каких распределениях узлов* (13) *имеет место равенство*

$$\lim_{n \to \infty} B_n^*(f, x) = f(x) \tag{14}$$

равномерно на отрезке $[0, 1]$ *для всякой функции* $f(x)$, *непрерывной на этом отрезке.*

β) *При каких распределениях узлов* (13)

$$\lim_{n \to \infty} B_n^{*'}(f, x) = f'(x) \tag{15}$$

равномерно на отрезке $[0, 1]$ *для всякой функции* $f(x)$, *имеющей на отрезке* $[0, 1]$ *непрерывную первую производную.*

Заслуживает рассмотрения также вопрос о том, существуют ли такие распределения узлов (13), для которых выполняется условие (14) и не выполняется условие (15).

При рассмотрении вопросов α) и β) была введена как мера отклонения распределения узлов (13) от равномерно распределенных бернштейновских узлов $\left\{\frac{k}{n}\right\}$ величина

$$\delta_n = \sum_{k=0}^{n} \left(\xi_n^{(k)} - \frac{k}{n} \right)^2.$$

При этом удалось доказать (правда с некоторыми пробелами, которые, однако, легко могут быть восполнены), что если $\delta_n = o(n)$, то условие (14) выполняется.

Вопрос β) не был рассмотрен; осталось также неясным, насколько оценка $\delta_n = o(n)$ существенна и не может быть ослаблена.

2. Некоторые виды сингулярных интегралов. Как обобщение теории особых интегралов Фейера и Джексона [1]) была сформулирована следующая задача.

Задача 11. *Изучить интегралы вида*

$$f_n(x) = \lambda_n \int_{-\pi}^{\pi} f(x+t) \frac{\varphi(nt)}{\varphi(t)}\, dt, \tag{16}$$

где

$$\lambda_n = \left[\int_{-\pi}^{\pi} \frac{\varphi(nt)}{\varphi(t)}\, dt \right]^{-1}$$

и функция $\varphi(t)$ удовлетворяет следующим условиям:

1) $\varphi(t) \geqslant 0$, $\varphi(0) = 0$, $\varphi(t) \leqslant 1$ для $t \in [-\pi, \pi]$;
2) $\varphi(-t) = \varphi(t)$;
3) $\varphi(t + 2\pi) = \varphi(t)$;
4) *функция $\varphi(t)$ возрастает на отрезке* $[0, \pi]$;
5) $\int_0^{\pi} \frac{dt}{\varphi(t)} = \infty$.

---

[1]) Интегралом Фейера называется интеграл

$$\delta_n(f, x) = \frac{1}{2\pi n} \int_{-\pi}^{\pi} f(x+t) \left[\frac{\sin \frac{nt}{2}}{\sin \frac{t}{2}} \right] dt.$$

Исследованию интеграла (16) были посвящены работы студентов Н. Я. Шадриной и Х. С. Шустера [1]). Н. Я. Шадрина доказала, что *для всякой функции* $f(x) \in C_{2\pi}$ [2])

$$\lim_{n \to \infty} f_n(x) = f(x)$$

равномерно на отрезке $[-\pi, \pi]$. Существенную роль в доказательстве этой теоремы играет лемма, согласно которой

$$\lim_{n \to \infty} \left[\int_0^{\pi} \frac{\varphi(nt)}{\varphi(t)}\, dt \right] = \infty. \tag{17}$$

Х. С. Шустер уточнил эту лемму и одновременно получил более общую ее формулировку. В частности, накладывая на функцию $\varphi(t)$ некоторые дополнительные условия, удалось доказать, что

$$\lim_{n \to \infty} \left[\int_0^{\pi} \frac{\varphi(nt)}{\varphi(t)}\, dt - \int_{\pi/n}^{\pi} \frac{dt}{\varphi(t)} \right] = 0. \tag{18}$$

Полученная таким образом оценка (18) позволяет изучить скорость сходимости интеграла (16) к функции $f(x)$ в точках, в которых функция $f(x)$ имеет не равные между собой правую и левую производные. Затем удается доказать для интегралов (16) некоторые теоремы типа теорем Вороновской [3]).

Если теория интегралов Фейера и Джексона привела к рассмотрению задачи 11, то теория интегралов Валле-Пуссена и Ландау

$$\frac{(2n)!!}{(2n-1)!!} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \cos^{2n} \left(\frac{t-x}{2} \right) dt \quad \text{и} \quad \frac{1}{2} \cdot \frac{(2n+1)!!}{(2n)!!} \int_0^1 f(t)[1 - (x-t)^2]^n\, dt.$$

подсказала следующую постановку вопроса.

---

[1]) Н. Я. Шадрина, Об одном виде особых интегралов, У. З., вып. 1; Х. С. Шустер, Об особом интеграле вида $f_n(x) = \lambda_n \int_{-\pi}^{\pi} f(x+t) \frac{\varphi(nt)}{\varphi(t)}\, dt$, там же.

[2]) $C_{2\pi}$ означает класс всех непрерывных 2π-периодических функций действительной оси.

[3]) Теорема Вороновской утверждает, что

$$B_n(x) = f(x) + \frac{f''(x)}{2n} x(1-x) + o\left(\frac{1}{n}\right).$$

Задача 12. *Изучить интегралы вида*

$$f_n(x)=\lambda_n\int_{-1}^{1} f(x+t)\,\varphi^n(t)\,dt,$$

где

$$\lambda_n=\Big[\int_{-1}^{1}\varphi^n(t)\,dt\Big]^{-1}$$

и функция $\varphi(t)$ *удовлетворяет следующим условиям:*

1) $\varphi(t)$ *непрерывна и неотрицательная на отрезке* $[-1,\ 1]$;

2) $\max\limits_{-1\leqslant x\leqslant 1}\varphi(x)=\varphi(0)=1$; $\varphi(t)<1$ *при* $x\in[-1,\ 0)$ *и* $x\in(0,\ 1]$;

3) *существует такое положительное число* m, *что*

$$\lim_{x\to 0}\frac{1-\varphi(x)}{x^{2m}}=a>0.$$

Н. С. Шедров получил следующую теорему: если $f(x)$ принадлежит к классу $\mathrm{Lip}_1 M$, то

$$|f_n(x)-f(x)|<M\cdot\delta(n),$$

где

$$\lim_{n\to\infty} n^{1/2\,m}\delta(n)=\frac{1}{2}\,a^{-1/2\,m}e^{-C/2\,m}$$

(C — постоянная Эйлера).

Кроме того, в этой же работе было доказано, что для всякой функции $f(x)$, непрерывной на отрезке $[-1,\ 1]$,

$$\lim_{n\to\infty} f_n(x)=f(x)$$

равномерно на отрезке $[-1,\ 1]$.

К работам по теории приближения функций следует также отнести работу Ш. М. Гойдо[1]), в которой изучался двукратный аналог интеграла Валле-Пуссена, а также незавершенное исследование о приближении непрерывных функций на отрезке $[-\pi,\ \pi]$ выражениями вида

$$T_{m,\,n}(x)=\sum_{k=1}^{n}\{P_{m,\,k}(x)\cos kx+Q_{m,\,k}(x)\sin kx\},$$

где $P_{m,\,k}(x)$ и $Q_{m,\,k}(x)$ — алгебраические полиномы степени, не высшей чем m (студент Матеевский).

[1]) Ш. М. Гойдо, Об одном двукратном интеграле, У. З., вып. 1.

IV. НЕКОТОРЫЕ ДРУГИЕ ВОПРОСЫ АНАЛИЗА

Задача 13. Обозначим через Ω класс всех функций, определенных на вещественной оси, а через A и B — два подкласса в Ω. Арифметической суммой двух классов функции $A\,«+»\,B$ назовем класс C всевозможных сумм вида $\varphi(x)+\psi(x)$, где $\varphi(x)\in A$ и $\psi(x)\in B$. Знаком $+$ будем, как обычно, обозначать теоретико-множественное сложение классов.

Определим понятие дополнительных классов. Два класса A и B назовем *дополнительными*, если:

1) $A\,«+»\,A=A$ и $B\,«+»\,B=B$,

2) $A\,«+»\,B=\Omega$,

3) $A+B\neq\Omega$.

Примерами дополнительных классов могут служить классы четных и нечетных функций, неположительных и неотрицательных функций и т. д.

Возникает вопрос: *каким условиям должен удовлетворять класс функций* A, *чтобы для него существовал дополнительный класс* B?

Несмотря на некоторые усилия, эта задача осталась нерешенной.

Задача 14. *Если* $\varphi(x)$ — *выпуклая вниз функция*[1]) *и*

$$\delta(a,\ b)=\frac{\varphi(a)+\varphi(b)}{2}-\varphi\left(\frac{a+b}{2}\right),$$

то для всякого x, $a\leqslant x\leqslant b$, *имеет место следующее неравенство:*

$$\delta(a,\ x)+\delta(x,\ b)\leqslant\delta(a,\ b).$$

Это неравенство доказал студент Радюк.

Задача 15. Пусть дан треугольник ABC. На стороне BC выберем точку A_1, на стороне AC — точку B_1 и на стороне AB точку C_1.

Затем передвигаем точку A_1 в положение A_2[2]) так, чтобы было $A_2B_1=A_2C_1$, точку B_1 — в положение B_2 так, чтобы было $B_2A_2=B_2C_1$, и точку C_1 — в положение C_2 так, чтобы $C_2B_2=C_2A_2$, и этот процесс продолжаем неограниченно.

Получаем последовательность треугольников

$$\triangle_1,\ \triangle_2,\ \ldots,\ \triangle_n,\ \ldots,$$

где

$$\triangle_k=A_kB_kC_k.$$

[1]) Функция называется *выпуклой вниз*, если $f\left(\frac{x+y}{2}\right)\leqslant\frac{f(x)+f(y)}{2}$ для всех x, y, принадлежащих области определения функции.

[2]) Точка A_2 лежит на прямой BC.

Изучить: а) *имеет ли последовательность* $\triangle_k$ *предел, и какой будет предельный треугольник* (что понимать под пределом последовательности $\triangle_k$ очевидно); б) *в какой мере предел зависит от исходного треугольников* ABC и $A_1B_1C_1$.

При рассмотрении этого вопроса сразу же появляется необходимость рассмотреть совокупность правильных треугольников, которые могут быть вписаны в основной треугольник ABC. Этот вопрос и сам по себе представляет определенный интерес.

Можно вписывать треугольники $A_nB_nC_n$ и не в треугольник, а в другие фигуры, например в окружность или, в общем случае, в выпуклую замкнутую кривую.

ПРЕПОДАВАНИЕ ЭЛЕМЕНТАРНОЙ МАТЕМАТИКИ В США [1])

Е. П. Розенбаум

(США)

Если у вас в семье есть школьник, которому время от времени приходится помогать по математике, то сами задачи покажутся вам удивительно знакомыми, хотя, быть может, вы и не вспомните их решений. Да, это — та же самая старая математика, которой училось прошлое поколение, которой учились еще двести лет назад. Всё осталось прежним — и методы, и содержание. Однако революция назревает. Несколько влиятельных групп преподавателей, глубоко не удовлетворенных современным уровнем преподавания математики, проводят эксперименты по коренному изменению методов изложения материала.

Уже сейчас есть школы, где старшеклассникам предлагают задачи вроде следующей. Докажите, что $[A, B|l]$ в том и только в том случае, когда $A \notin l$, $B \notin l$ и $\overline{AB} \cap l = \varnothing$. (Перевод этой записи: «Докажите, что точки A и B лежат по одну сторону прямой l в том и только в том случае, когда A не принадлежит l, B не принадлежит l и отрезок AB не пересекается с l».)

Много спорят о том, насколько полезны подобные вещи и как далеко нужно идти в этом направлении, но в одном согласны почти все: существующий курс математики не годится. В нем не раскрывается само содержание предмета, он не дает учащимся понимания основных принципов математики. Курс настолько отстал от жизни, что в нем отсутствуют практически все новые идеи и открытия последних ста лет. И, самое главное, существующие школьные программы привели к тому, что математика стала одной из самых непопулярных отраслей знания, о полном невежестве в которой даже образованные люди заявляют с вызовом, а иногда и не без гордости.

Надо что-то сделать, чтобы курс математики стал более содержательным и более привлекательным. В настоящее время предприняты

[1]) E. P. Rosenbaum, The Teaching of Elementary Mathematics, Scientific American, май 1958, стр. 64—73. Автор — член редакционного совета этого журнала, в течение ряда лет преподававший математику и физику в школе (в Милфорде).

по крайней мере три важные попытки в этом направлении. Ниже я попытаюсь коротко их обрисовать.

Основное в этих попытках — стремление в той или иной степени модернизировать содержание курса математики.

Под модернизацией понимают в основном две вещи: устранение из программ устаревшего материала и введение в курс тех наиболее важных современных идей, которые за последнее столетие придали общность и сделали более содержательными все традиционные разделы математики.

В качестве примера устаревшего материала можно указать на методы решения треугольников с помощью логарифмов, занимающие немало места в учебниках по тригонометрии. Сто лет тому назад эти методы были единственными, имевшимися в распоряжении геодезистов и мореплавателей; сейчас такие задачи решаются на вычислительных машинах. Устарели также некоторые из классических евклидовых «доказательств» в геометрии, которые, по сути дела, вовсе не являются доказательствами. Современный математик подходит к таким задачам с совершенно иных позиций и получает действительно строгие доказательства [1]).

Математика, как ее сейчас преподают в школах и даже в колледжах, представляет собой набор отдельных предметов, каждый из которых имеет свои, представляющиеся произвольными, правила, заучиваемые наизусть. Однако работа в области оснований математики, проведенная за последнее столетие, показала, что все разделы математики могут быть сведены к чисто абстрактным понятиям, обладающим общими свойствами. Как числа являются элементами алгебры, так точки и прямые являются «первоначальными» элементами геометрии, и к каждому из этих множеств мы можем применять одни и те же логические операции. Правила математической логики носят универсальный характер, и сторонники модернизации считают, что основные ее идеи и операции могут быть усвоены даже школьниками.

После этого математика станет более понятной и более содержательной для учащихся. В то же время они получат хотя бы некоторое представление о современном математическом мышлении.

ПРОГРАММА СОВЕТА КОЛЛЕДЖЕЙ

Среди проектов обновления школьного курса математики выделяется проект Комиссии по математике Совета по вступительным экзаменам в колледжи. Эта комиссия, созданная в 1955 г., проводит серьезный пересмотр школьных курсов. Так как в комиссии представлены самые разнообразные точки зрения и так как она надеется оказать почти немедленное влияние на каждую школу в стране, ее программа сравнительно консервативна. Комиссия предлагает изменить дух и методику преподавания курса алгебры, без значительного изменения его содержания, но зато подвергает серьезному пересмотру курс геометрии, вносит некоторые изменения в курс тригонометрии и вводит в учебный план школ несколько совершенно новых курсов.

[1]) См. статью П. К. Рашевского «Геометрия и ее аксиоматика», «Математическое просвещение», вып. 5, стр. 73—98 (*Прим. ред.*)

Алгебру, как и геометрию, можно рассматривать как абстрактную дедуктивную систему. Она строится на базе некоторого множества неопределяемых первоначальных понятий и некоторого числа принимаемых без доказательства аксиом. Все остальные положения и правила алгебры могут быть получены отсюда путем логических рассуждений.

Комиссия не предлагает строить школьный курс алгебры чисто дедуктивно, но она считает, что учащихся надо подвести к пониманию дедуктивного характера алгебры, что они должны получить представление о том, что́ такое аксиомы алгебры и ка́к они используются для доказательства некоторых положений.

Если будут поняты основные идеи, лежащие в основе алгебраических операций, то легче будут приобретаться и необходимые навыки в оперировании с алгебраическими выражениями.

Каковы же «первоначальные» понятия алгебры? Это — понятие *числа* и понятие *операций*, таких, как *сложение* и *умножение*. Аксиомы также настолько просты для понимания, что кажется, что их не стоило бы и упоминать. Если a и b — числа, то $a+b$ — число. Если $a=b$ и $c=d$, то $a+c=b+d$ (суммы равных чисел равны). Кроме того, существуют коммутативные законы операций [$a+b=b+a$ и $a\times b=b\times a$], ассоциативные законы [например, $(a+b)+c=a+(b+c)$] и дистрибутивный закон [$a(b+c)=ab+ac$]. Эти положения, хотя они и могут показаться совершенно очевидными и не заслуживающими того, чтобы их специально формулировали, образуют математическую систему, которая лежит в основе всех общеизвестных процедур элементарной алгебры и из которой могут быть выведены все ее теоремы.

Рассмотрим, например, теорему, утверждающую, что сумма двух четных чисел четна. В начале доказательства дается следующее определение четности: число n четно тогда и только тогда, когда существует другое число p, такое, что $n=2p$.

Задача такова: дано, что a и b — четные числа, требуется доказать, что $a+b$ четно. По определению четности мы можем написать $a=2x$ и $b=2y$. По аксиоме о сложении равных чисел $a+b=2x+2y$. По аксиоме дистрибутивности $2x+2y=2(x+y)$. Так как x и y — числа, $x+y$ — число. Следовательно $2(x+y)$ — четное число и $a+b$ — четно.

Неспециалисту может показаться, что столь громоздкое доказательство почти очевидного положения не будет интересным и ничего не даст учащемуся средней школы. Подобной точки зрения придерживаются и некоторые математики. Но сторонники реформы утверждают, что при хорошо поставленном преподавании, упражнения подобного рода могут стать для учащихся захватывающим путешествием в область математической строгости.

ТЕОРИЯ МНОЖЕСТВ

Основное, что отличает «современную» алгебру, это — широкое использование теории множеств, одного из наиболее мощных орудий современной математики. *Множество* — это просто собрание или совокупность каких-либо элементов. Книги на полке, люди в комнате, буквы алфавита, числа 1, 2, 3, 4 и 5 — всё это примеры множеств. Множество может содержать только один элемент или совсем не иметь элементов («пустое» множество). Оно также может быть бесконечным — например, все точки внутри окружности или все целые положительные числа. В математике множество обычно обозначают написанием его элементов внутри скобок, например $\{1, 2, 3, 4, 5\}$.

Для записи высказываний о множествах и их элементах существуют общепринятые сокращения. Например, высказывание «4 является элементом множества A» записывается $4 \in A$, высказывание «6 не принадлежит A» записывается $6 \notin A$. Высказывание «A есть подмножество B» [т. е. множество, все элементы которого входят в множество B] записывается $A \subseteq B$. Пустое множество обозначается $\varnothing$. Буквой U обозначается *универсальное множество*, содержащее все объекты, рассматриваемые в данной области. В планиметрии, например, универсальное множество содержит все точки плоскости.

Может показаться, что этот необычный язык представляет ненужную трудность для начинающих изучать математику, но на деле учащиеся легко овладевают им после небольшой практики. Изучить его значительно легче, чем, скажем, стенографию.

Из действий над множествами я для иллюстрации упомяну только три. *Объединением* двух множеств A и B (записывается $A \cup B$) называется множество, состоящее из всех элементов A и всех элементов B. Таким образом, объединением множеств $\{1, 2, 3, 4\}$ и $\{2, 3, 4, 5\}$ является множество $\{1, 2, 3, 4, 5\}$. *Пересечением* двух множеств ($A \cap B$) называется множество всех элементов, общих обоим множествам. Таким образом, пересечением множеств $\{1, 2, 3, 4\}$ и $\{2, 3, 4, 5\}$ будет множество $\{2, 3, 4\}$. Наконец, *дополнением* множества A (записывается $\bar{A}$) является множество всех элементов универсального множества, которые не являются элементами данного множества. Например, если универсальным множеством является $\{1, 2, 3, 4, 5\}$, то дополнением множества $\{2, 3\}$ будет множество $\{1, 4, 5\}$. Эти

понятия могут быть проиллюстрированы графически с помощью так называемых *фигур Венна* (рис. 1).

Теория множеств позволяет раскрыть единство математики. И алгебра, и геометрия имеют дело с множествами: алгебра — с множествами чисел, геометрия — с множествами точек. Операции, изучаемые в этих двух разделах математики, могут рассматриваться как примеры общих операций над множествами — объединения, пересечения и т. д.

Посмотрим, как понятие множества может помочь разъяснить понятия *переменного* и *уравнения* на уроках алгебры. Комиссия Совета колледжей рекомендует, чтобы учащимся в первую очередь объяснили,

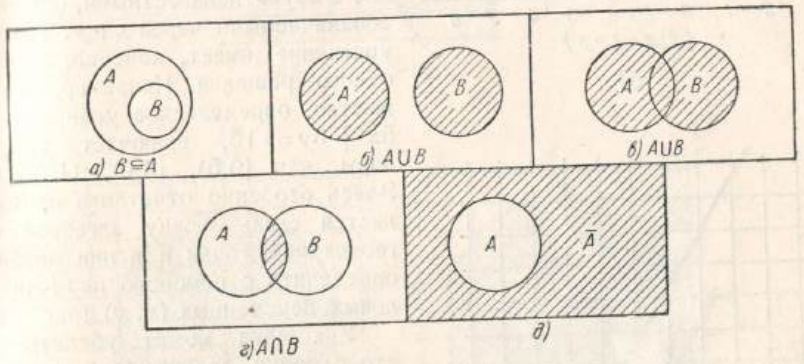


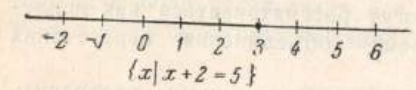
Рис. 1. Фигуры Венна наглядно изображают соотношения между множествами.

а) Включение: B является подмножеством *A*. *б)* и *в)* Штриховкой показано *объединение A* и *B*. *г)* Заштриховано *пересечение* множеств. *д)* Заштриховано *дополнение* множества *A*.

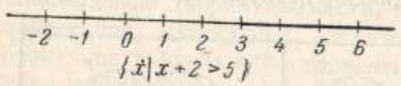
что алгебра занимается множествами чисел и соотношениями между ними. После этого они узнают, что переменное — это просто общее наименование (обозначаемое буквой) для элементов множества. Например, если множеством является ряд целых чисел, то x может быть любым целым числом, а $x+1$ может быть $1+1$ или $2+1$ или $3+1$ и т. д. Далее, уравнение — это высказывание об отношении между элементами множества. Само высказывание может быть истинным и ложным: $3+2=5$ — истинно, $3+2=6$ — ложно, но и то и другое — высказывания. Если высказывание содержит буквы, то его истинность не определена: $x+2=5$ — ни истинно и ни ложно до тех пор, пока вместо x не подставлен некоторый элемент подходящего множества. В начале курса учащихся знакомят также с неравенствами и символами $>$ (больше чем) и $<$ (меньше чем).

После этого каждое соотношение может рассматриваться как определение, выделяющее некоторое подмножество из рассматриваемого

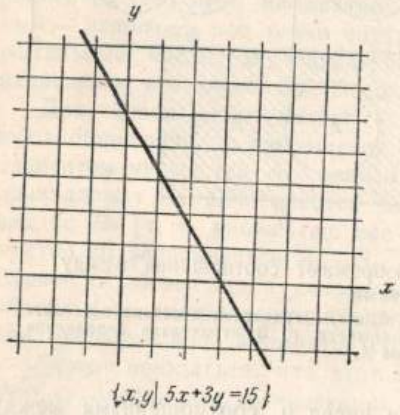
универсального числового множества. В терминах теории множеств оно может быть названо *определением* множества. Так, если в качестве универсального множества взято множество всех действительных чисел, то соотношение $x+2=5$ определяет множество $\{3\}$, а $x+2>5$ определяет множество, состоящее из всех чисел, больших 3. Множество, определяемое некоторым соотношением, называется множеством его решений.



а)



б)



в)

Рис. 2. Геометрическое и алгебраическое изображение множеств.

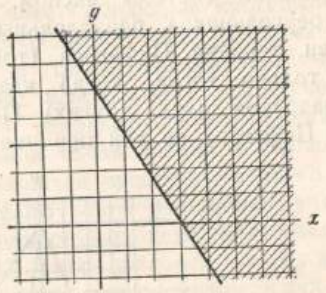
а) Точка на числовой прямой изображает множество решений уравнения $x+2=5$. *б)* Неравенства также определяют множества. Утолщенная часть линии изображает бесконечное множество, выделяемое неравенством $x+2>5$. *в)* Выражение, содержащее две переменные, определяет множество пар чисел. Совокупность точек этой прямой дает множество решений уравнения $5x+3y=15$.

Тот же подход годится и для пар чисел, и мы можем подобным же образом рассматривать уравнение с двумя неизвестными, обычно обозначаемыми через x и y. Такое уравнение имеет, конечно, более одного решения. Например, множество, определяемое уравнением $5x+3y=15$, включает такие пары, как (0,5), (3,0), $(1,3^1/_3)$. Здесь особенно отчетливо проявляется связь между алгеброй и геометрией: точки и линии можно определять с помощью пары числовых переменных (x, y) (рис. 2).

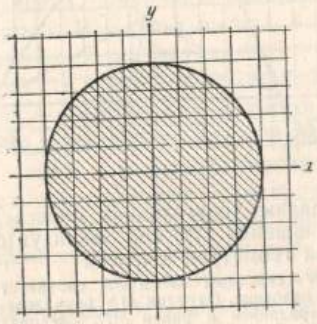
Учащийся может убедиться, что с точки зрения теории множеств нахождение решения системы двух алгебраических уравнений сводится к отысканию пересечения двух множеств решений: решениями будут пары чисел, общие обоим множествам.

Понятие множества особенно помогает при рассмотрении неравенств. Значение записи типа $5x+3y>15$ становится более ясным, когда ее рассматривают как «определяющую» для тех пар чисел, которым соответствуют точки, лежащие выше линии $5x+3y=15$. Соотношение $x^2+y^2<16$ выделяет точки внутри окружности $x^2+y^2=16$. Решение системы этих двух неравенств опять-таки сводится к нахождению пересечения множеств их решений (рис. 3).

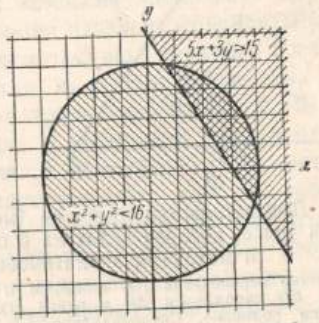
Приведенные здесь примеры иллюстрируют подход комиссии Совета колледжей к предлагаемой реформе курса алгебры. Если не считать некоторого расширения раздела неравенств, комиссия не вносит существенных изменений в содержание курса, но зато предлагает решительно изменить характер изложения материала.



а)



б)



в)

Рис. 3. Неравенства легко интерпретировать с помощью понятия множества.

а) Заштрихованная область соответствует множеству, выделяемому неравенством, стоящим в скобках. *б)* Пары чисел, удовлетворяющие неравенству, стоящему в скобках, соответствуют множеству точек, находящихся внутри окружности. *в)* Пересечение множеств, определяемых двумя неравенствами, выделено на рисунке густой штриховкой.

НОВАЯ ГЕОМЕТРИЯ

Комиссия Совета колледжей предлагает коренным образом изменить школьный курс геометрии. В первую очередь должно быть значительно сокращено число теорем, знание доказательств которых требуется от учащихся. Делается это на том основании, что, во-первых, учащиеся уже получают навыки в проведении дедуктивных рассуждений при изучении курса алгебры и, во-вторых, потому, что многие положения легче доказывать не по Евклиду, а другими

способами. В программе Совета колледжей число доказываемых теорем сокращено до 12, тогда как в существующих учебниках их более ста. Эти двенадцать теорем являются примерами дедукции геометрических положений из принятой системы аксиом. В курсе выводятся теоремы, касающиеся треугольников, параллельных линий, подобных треугольников и, наконец, теорема Пифагора (треугольник является прямоугольным тогда и только тогда, когда квадрат одной из его сторон равен сумме квадратов двух других). Первая часть курса заканчивается теоремой Пифагора по той причине, что она позволяет

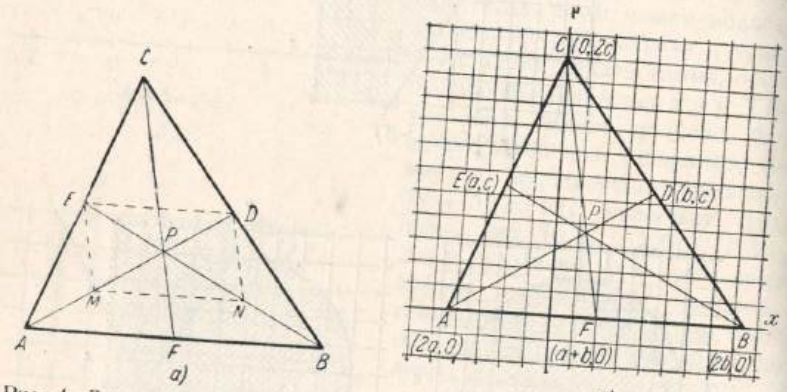


Рис. 4. Доказательство теоремы (медианы треугольника пересекаются в одной точке, которая делит их в отношении 2:1) «по Эвклиду» (слева) и методами аналитической геометрии (справа).

а) P — точка пересечения AD и BE. M и N — середины AP и BP соответственно. Отрезок ED параллелен AB и равен его половине. Отрезок MN параллелен AB и равен его половине. Следовательно, ED параллелен и равен MN. Следовательно, $MP=DP$ и $EP=NP$. Но $AM=MP$ и $BN=NP$. Следовательно, P делит отрезки AD и BE в отношении 2:1. Подобным же рассуждением доказывается, что точка пересечения AD и CF делит AD в отношении 2:1. Следовательно, медианы пересекаются в одной точке, которая делит их в отношении 2:1, ч. и т. д.

б) По правилам написания уравнений прямых, уравнения, определяющие AD, BE и CF, будут соответственно:

$$cx+(2a-b)\,y=2ac,$$
$$cx+(2b-a)\,y=2bc,$$
$$2cx+(a+b)\,y=2ac+2bc.$$

Эти уравнения имеют общее решение

$$x=\frac{2}{3}(a+b), \qquad y=\frac{2}{3}c.$$

Следовательно, медианы пересекаются в одной точке. Подстановка соответствующих координат в формулу расстояний показывает, что $AP=2PD$, $BP=2PE$ и $CP=2PF$, ч. и т. д.

перейти к аналитической геометрии, т. е. к решению геометрических задач алгебраическим методом (рис. 4). При помощи аналитического метода демонстрируется применение в геометрии понятия множества решений.

Комиссия Совета колледжей не предлагает, однако, полностью отбросить классические евклидовы доказательства — ведь в ряде случаев доказательства «по Евклиду» проще доказательств аналитических. Учащихся следует поощрять находить наиболее эффективный подход к каждой отдельной задаче.

Рекомендуемая более компактная программа позволяет значительно сократить время, необходимое для усвоения планиметрии, и проходить ее менее чем за год. Освободившееся время комиссия Совета колледжей предлагает использовать для изучения основ стереометрии. В этой части курса основные положения рассматриваются главным образом с интуитивной точки зрения, и первостепенной задачей является развитие у учащихся пространственного мышления.

На третьем году обучения математике учащиеся продолжают изучение алгебры и приобретают некоторые сведения из тригонометрии. При этом основной упор делается не на решение треугольников, а на изучение поведения тригонометрических функций и применение их к изучению векторов.

Часть учащихся продолжают изучать математику и на четвертом году обучения. Для этих учащихся комиссия предлагает два полугодовых курса. В первом из них, носящем название «элементы анализа», подвергаются более детальному рассмотрению понятия соотношения и функции. В курсе рассматриваются свойства полиномов, логарифмической, показательной и тригонометрических функций. В нем неформально вводится понятие предела и учащимся прививаются некоторые навыки в дифференцировании и интегрировании многочленов.

Во второй половине года предлагается курс теории вероятностей и математической статистики. Здесь учащиеся получают представление о применении статистических методов к обработке результатов измерений и о математическом рассмотрении случайных процессов. В курс включены и некоторые методы оценки надежности выборочной программы испытаний и методы определения статистической значимости результатов.

Авторы проекта указывают, что этот материал теснее всех остальных изучаемых в школе разделов математики связан с повседневной жизнью, и отмечают всё возрастающее значение статистики и теории вероятностей в науке и технике.

Так как бóльшая часть курса теории вероятностей никогда не преподавалась в средней школе, комиссия подготовила по этому предмету новый учебник[1]). В учебнике материал излагается сначала на интуитивной основе, а затем с более формальных позиций. В курсе широко используется теория множеств. Лично мне этот учебник показался интересным и легко читаемым, а его содержание не более трудным, чем содержание многих разделов математики, обычно включаемых в школьные программы повышенного типа.

[1]) Рецензию на этот учебник см. в настоящем выпуске «Математического просвещения», стр. 355—360.

В эти последние программы комиссия рекомендует ввести материал по анализу и аналитической геометрии в том объеме, в котором они обычно преподаются на первом курсе колледжей.

Таков вкратце проект реформы преподавания математики в средних школах, предлагаемый комиссией Совета колледжей. Комиссия обращается с призывом ко всем школам и преподавателям полностью или частично испробовать ее предложения на практике с тем, чтобы выяснить, что́ в них годится, а что́ нет.

Комиссия считает, что математика обретет настоящую жизнь в школе только тогда, когда перед учащимися научатся раскрывать красоту стиля и методов современной математики.

Хотя предлагаемая программа является лишь рекомендацией, комиссия, без сомнения, может оказать значительное влияние на школы посредством организуемых ею экзаменов и благодаря своей тесной связи с ведущими колледжами. Налаживая контакт с учительской общественностью страны, комиссия выпустила брошюры, рассказывающие об отдельных этапах ее работы. Этой осенью она намеревается выпустить подробный доклад.

Весьма возможно, что предложения комиссии приведут к созданию в ближайшие один-два года совершенно новых учебников по математике.

ИЛЛИНОЙССКАЯ ПРОГРАММА

А теперь рассмотрим другой проект реформы школьного курса математики, зародившийся в Иллинойсском университете. Этому проекту уделялось некоторое внимание в прессе. Его инициаторами являются Макс Беберман — преподаватель иллинойсской средней школы и Герберт И. Воэн — математик из Иллинойсского университета. Программа, разработанная ими, идет значительно дальше, чем программа комиссии Совета колледжей.

Осуществление иллинойсской программы на практике было начато в 1952 г. в средней школе при Иллинойсском университете. Эта программа, рассчитанная на четыре года, испытывается сейчас уже примерно в двенадцати школах в штатах Иллинойс, Миссури и Массачузетс, а также на курсах для служащих компании «Поляроид» в Кембридже (штат Массачузетс). Иллинойсская группа выпустила полный комплект учебников и методических руководств по предлагаемой ею программе. Она организовала при университете годичные курсы для преподавателей, имеющие целью познакомить их с программой и опытом ее преподавания. Средства, необходимые для всего этого, были получены от корпорации Карнеджи.

Характерной чертой иллинойсской программы является то, что в ней абстрактные обобщения не служат лишь упражнениями в логическом мышлении или выборочными иллюстрациями основ математики,

а, наоборот, лежат в основе всего курса. В девятом классе курс алгебры начинается с изложения системы аксиом, из которой в дальнейшем учащиеся выводят все необходимые им положения. Аксиомы, носящие название «принципы арифметики», включают, например, определение нуля (всякое число, умноженное на нуль, равно нулю), определение единицы (любое число, умноженное на единицу, равно самому себе) и, конечно, коммутативный, ассоциативный и дистрибутивный законы.

К математической строгости детей приучают постепенно, прибегая иногда к форме занимательного рассказа. Так, например, необходимость различения числа и обозначающего его символа поясняется на примере «переписки» между Эдом Брауном из Забрачбургской школы и Полем Мором — его другом из Аляски. Эти два школьника, как это ни странно, в своих письмах обсуждают проблемы арифметики. Поль в своем письме утверждает, что если от 21 отнять 2, то мы получим 1. Эд удивлен: «Что это — шутка, или в рассуждениях Поля действительно что-то есть?» В тексте учебника всё разъясняется. Оказывается, «21» — это лишь обозначение числа, его имя. Число 21 может быть обозначено и другими способами, например, «$20+1$» или «7×3». Затем учащиеся узнают, что число может быть обозначено буквой, и что последней можно придавать различные значения (т. е. буква оказывается «переменной» или «заменителем» числа). Для обозначения неизвестных в уравнении учитель пользуется пустыми квадратами. Доведению всего этого до сознания учащихся посвящена значительная часть первого года изучения математики. В учебнике отмечается сходная роль букв в математических выражениях и местоимений в обычном языке: выражение «$x+2=6$» подобно предложению «он — президент Соединенных Штатов» — ни то, ни другое не ложно и не истинно до тех пор, пока вместо буквы не поставлено число, а вместо местоимения — имя собственное. На этом этапе учитель вводит термин «прочисло» («pronumeral») для обозначения букв в алгебраических выражениях и пользуется этим термином на протяжении всего курса.

Значительная часть первого года посвящается также введению и исследованию понятия множества.

Поскольку понятиям «прочисла» и множества уделяется много времени, курс первого года содержит меньше материала, чем обычный курс элементарной алгебры.

На втором году модернизация идет еще дальше. Курс представляет собой «построение евклидовой геометрии не менее строгое, чем, например, у Гильберта, но которое, по нашему мнению, доступно для учащихся, усвоивших материал первого года».

В курсе предпринята попытка обнажить чисто дедуктивную сущность геометрии, т. е. изучать геометрию (до введения конкретных интерпретаций) как некую лишенную содержания логическую структуру.

Рассмотрим, например, следующие три постулата, необходимые для построения евклидовой геометрии:

1. Каждая прямая содержит по крайней мере две точки.
2. Существуют по крайней мере три точки, не принадлежащие ни к какой одной прямой.
3. Любые две точки принадлежат по крайней мере одной прямой.

Учитель указывает, что слова «точка» и «прямая» могут быть интерпретированы различными способами. В качестве основной модели в курсе принята «числовая плоскость», на которой точки определяются как пары чисел (x, y). Все теоремы планиметрии выводятся с помощью этой модели. Но для того, чтобы подчеркнуть, что дедуктивная система может иметь различные конкретные интерпретации, учитель рассматривает несколько моделей, включая и такую, в которой «точка» обозначает предпринимателя, а «прямая» — соучастие в делах. Перечисленные выше постулаты с таким же успехом могут быть приложимы и в этом случае. Предположим, что имеются три предпринимателя A, B и C, каждый из которых связан соучастием в делах поровну с каждым из двух других. Каждое соучастие (AB, AC и BC) является, таким образом, множеством предпринимателей, содержащим по крайней мере два элемента (постулат 1). Ни одно из соучастий не включает всех трех предпринимателей (постулат 2). Каждые два предпринимателя объединены по крайней мере одним соучастием (постулат 3).

Для иллюстрации методов преподавания я приведу отрывок из записи обзорного занятия по системе постулатов, которое Беберман проводил в классе повышенного типа.

«Учитель: С каким объектом связано происхождение этих постулатов?

Джордж: С числовой плоскостью.

Учитель: Что ты под этим подразумеваешь?

Джордж: Эти постулаты выражают свойства числовой плоскости.

Учитель: Если мы говорим о числовой плоскости и утверждаем, например, что каждая прямая является множеством точек, содержащим по крайней мере две точки, будет ли это верно?

Джордж: Будет.

Учитель: Но предположим, что мы не говорим о числовой плоскости. Что ты можешь тогда сказать об этих постулатах?

Джордж: Они будут ложными.

Голоса: Неверно!

Джим: Нужно придать конкретное значение терминам «линия» и «точка». Только тогда мы можем говорить об истинности или ложности постулатов.

Учитель: Что такое «модель» системы постулатов?

Джейн: Это нечто, обладающее свойствами, выраженными в постулатах, т. е. удовлетворяющее постулатам.

Учитель: Пусть существует несколько интерпретаций постулатов. Например, мы можем говорить о числовой плоскости, о «предпринимателях» или о «классных старостах» и «комитетах». Что мы тогда можем сказать?

Джордж: Если мы хотим доказать на основании постулатов некоторую теорему, то иногда мы можем убедиться в невозможности это сделать на основании того, что эта теорема неверна для модели».

Если способности учащихся выше средних, и если руководителем является преподаватель вроде Бебермана, то подобные дискуссии проходят живо и интересно.

В выпущенной иллинойсской группой брошюре утверждается:

«Школьники проявляют глубокий интерес к идеям. Им нравится оперировать с отвлеченными понятиями... Модная сейчас тенденция подчеркивать практическую полезность математики для различных профессий не находит отклика у большинства из них. Идея практической пользы слишком далека для ученика девятого класса. Он хочет знать, как математика входит в его собственный мир. А этот мир, к счастью для нас, полон фантазии и абстракций. Математика вводит школьников в увлекательный мир захватывающих приключений и заинтересовывает их».

Конечно, иллинойсский курс носит экспериментальный характер. Часть материала оказалась слишком трудной или требующей длительного времени для усвоения. В настоящее время иллинойсская группа рассматривает целесообразность перенесения некоторых вопросов строгого обоснования математики на четвертый год обучения. Однако исходная позиция Бебермана такова: не попробовав, не узнаешь, чему можно научить школьников.

В этом году только первая группа учащихся заканчивает четырехлетний курс, и еще трудно говорить о результатах, но посещение средней школы при университете убедило меня, что учащиеся выполняют свою отнюдь не легкую задачу с радостью и даже с энтузиазмом. Всё дело в том, может ли быть осуществлена такая программа в широком масштабе. Некоторые критики считают, что она может оказаться полезной для способных учеников и при наличии талантливых преподавателей, но выражают сильное сомнение в ее успехе в среднем классе со средним преподавателем.

КРИТИКИ

Некоторые педагоги-математики считают, что сам подход модернистов в корне неверен. Наиболее явно выразил эту точку зрения Морис Клайн — профессор Института математических наук Нью-Йоркского университета, известный своими популярными работами по современным достижениям математики. Клайн выражает сомнение в том, что абстрактный подход способен вызвать у подростков интерес к математике.

Клайн утверждает, что нельзя абстрагировать, пока вы не знакомы с тем, от чего вы абстрагируетесь. Профессиональные математики приходят к пониманию общих абстрактных построений математики

лишь только после знакомства со многими ее конкретными разделами. Ка́к же можно требовать бо́льшего от школьника? Кроме того, Клайн считает, что теории множеств не место в элементарной математике, так как в этой области она может найти лишь тривиальные приложения, да и вообще, по его мнению, значение теории множеств для высшей математики значительно переоценивается.

Клайн и другие считают, что интерес к математике можно пробудить, если идти не по пути абстракций, а, наоборот, по пути конкретизации, по пути более тесной связи между математикой и жизнью. Клайн предпочитает разъяснять математические идеи с помощью простых физических опытов и на примерах, взятых, например, из музыки. Активно поддерживает эту точку зрения английский математик и педагог Сойер, работающий сейчас в Иллинойсском университете.

Но, по всей видимости, антимодернисты терпят поражение. Они не организованы, и у них нет какой-либо определенной программы, а движение модернистов быстро расширяется. И это касается не только школ. Ряд колледжей уже принял новую программу для первокурсников, разработанную комитетом Американской математической ассоциации. Этой программой предусматривается полугодовой курс анализа и полугодовой курс по новым разделам математики (эти новые разделы условно называют «конечной» или «дискретной» математикой). В последнем курсе рассматриваются дискретные множества, находящие всё более широкое применение, особенно в общественных науках, имеющих дело с группами лиц, партиями изделий, состоящими из целого числа единиц и т. п. Даже физика дискретна (если считать атом элементарной частицей), а в основе квантовой теории лежит матричная алгебра — один из разделов дискретной математики. Задача нового курса — познакомить с некоторыми понятиями современной математики студентов, изучающих литературу и искусство, и привить некоторые математические навыки студентам, специализирующимся в области общественных наук. Наиболее широко распространенный сейчас вариант программы курса включает символическую логику, теорию вероятностей, матричную и векторную алгебру, теорию игр и линейное программирование.

ПРЕПОДАВАТЕЛИ

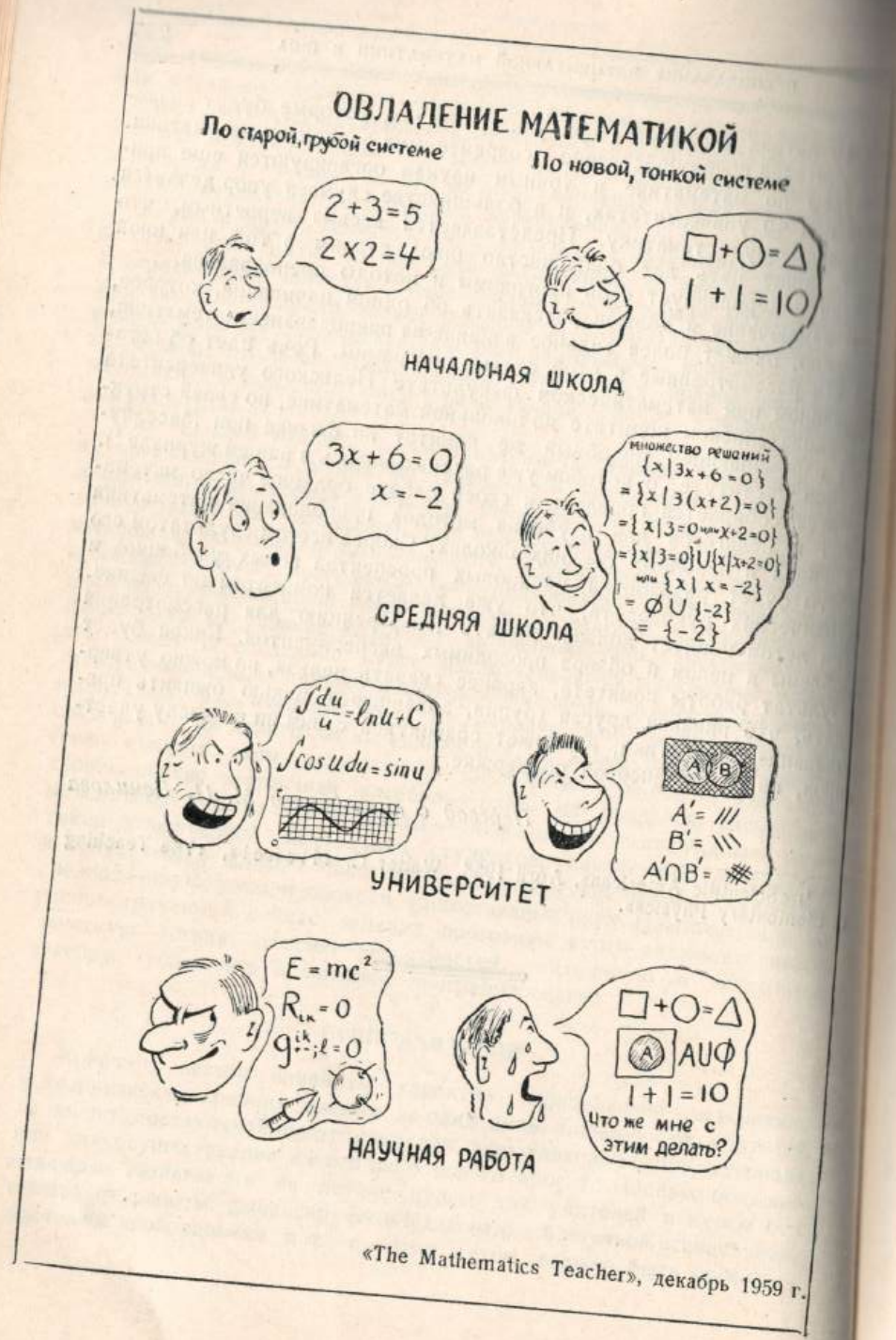
Конечно, нельзя изменить характер преподавания математики в американских средних школах за один день. Большинство учителей не имеют достаточной подготовки для преподавания нового материала или для осуществления какого-либо нового проекта. Однако большие надежды возлагаются на летние курсы для учителей и курсы без отрыва от работы, финансируемые Национальной научной ассоциацией, частными корпорациями и т. п. Этим летом должно быть организовано 10 летних институтов при университетах, которые будут знакомить учителей преимущественно с современными разделами математики.

Курсы по математике и точным наукам организуются еще примерно при 40 университетах, и в большинстве случаев упор делается на современную тематику. Представляется весьма вероятным, что в ближайшие пять лет большинство школ страны в той или иной степени модернизирует свои программы и методы преподавания.

В заключение я должен рассказать об одном начинании, которое, возможно, окажет более сильное влияние на преподавание математики, чем все рассмотренные в этой статье программы. Речь идет об организованном при математическом факультете Йельского университета Исследовательском комитете по школьной математике, по своей структуре напоминающем подобный же комитет по физике при Массачусетском институте, о котором уже рассказывалось в нашем журнале [1]). Комитет собирается привлечь к своей работе большое число математиков и педагогов для изучения методов преподавания математики в средних и неполных средних школах. Скорее всего результатом его работы будет разработка подробных проспектов новых учебников и методических руководств, что уже делается комитетом по физике. Этим летом комитет должен созвать конференцию для рассмотрения проблемы в целом и обзора проводимых экспериментов. Каков будет результат работы комитета, заранее сказать нельзя, но можно утверждать, что никакая другая группа, задавшаяся целью оживить преподавание математики, не может сравниться с ним ни по числу участников, ни по финансовой поддержке.

Перевод с английского М. П. Данилова

[1]) «Scientific American», April 1958, Walter C. Michels, «The Teaching of Elementary Physics».



«The Mathematics Teacher», декабрь 1959 г.

ПОДГОТОВКА УЧИТЕЛЕЙ МАТЕМАТИКИ ДЛЯ СРЕДНЕЙ ШКОЛЫ В НОРВЕГИИ

И. Я. Депман

(Ленинград)

Неоднократно констатировался факт, что подготовка абитуриентов средней школы по математике выше всего стоит в малых странах севера Европы (Скандинавских, Голландии, Бельгии, Люксембурге [1]). Поэтому представляет интерес своеобразная постановка подготовки учителей математики средней школы в Норвегии.

В Норвегии обязательной для всех детей до 14-летнего возраста является лишь *семилетняя начальная школа*. Около половины детей в дальнейшем продолжает обучение: небольшой процент — в *реальной школе*, которая дает общую подготовку, недостаточную для высших учебных заведений, а остальные — в *гимназиях*, которые одни, собственно, и соответствуют средней школе других стран.

Во всех гимназиях изучаются, кроме родного языка, языки английский, немецкий и французский, история, гражданствоведение, география, математика, биология. Дается, кроме того, художественное и физическое воспитание.

Каждый учащийся гимназии избирает себе один из пяти основных предметов: 1) латинский язык, 2) английский язык, 3) математику с физикой, 4) древненорвежский язык и 5) естественные науки. После пяти лет обучения учащийся, в 19- или 20-летнем возрасте, подвергается выпускному экзамену (examen artium), организуемому национальным (государственным) советом по среднему образованию. Этот экзамен является необходимым и достаточным для поступления в университет.

Уровень преподавания математики в гимназиях достаточно высок [2] и требует от учителей глубоких математических знаний.

[1]) См. И. Я. Депман, Вопросы преподавания элементарной математики на последнем международном конгрессе математиков, «Математика в школе», 1958, № 3, стр. 77—83.

[2]) По материалам статьи Ф. Фера (президента Национального совета подготовки учителей математики в США и руководителя отделения подготовки учителей в Колумбийском университете) и К. Пине (руководителя и ректора

Приводим некоторые сведения о подготовке учителей математики для норвежской средней школы [1]).

Окончивший университет и желающий получить право преподавания в I и II классах гимназии и в III классе реальной школы должен пройти полугодовую подготовку в педагогической семинарии, где он изучает основы педагогических дисциплин (истории, философии, психологии, гигиены, школьной администрации, методики преподавания) и знакомится с употребительными в школе учебниками. Одновременно кандидат преподает в школе. Примерно в 23-летнем возрасте он получает возможность приступить к педагогической деятельности.

Чтобы получить право преподавания математики в III, IV и V классах гимназии, кандидат на должность учителя должен после педагогического семинария вернуться на 1,5 — 2 года в университет для продолжения своего математического образования. В эти годы изучаются теория функций действительного и комплексного переменного, теория чисел, абстрактная алгебра, аналитическая геометрия пространства и некоторые специальные области — теория вероятностей, топология, исчисление конечных разностей и т. п. Этот период работы в университете является существенно важным — многие учителя математики в норвежских гимназиях при четырехлетнем прохождении основного университетского курса имели главным предметом изучения не математику, а физику, географию или иногда, хотя и редко, другие дисциплины.

После дополнительных занятий математикой в университете, сдачи экзаменов и представления диссертации соискатель получает право преподавания математики в гимназии.

ПРИЛОЖЕНИЯ

I. Вариант письменных вопросов для кончавших среднюю школу в 1956 году

1. а) Вывести формулу для корней уравнения

$$x^3 + ax^2 + bx + c = 0.$$

б) На основании этой формулы доказать, что если три корня уравнения составляют геометрическую прогрессию, то один из корней всегда равен $-\sqrt[3]{c}$ и тот же корень также равен $-\frac{b}{a}$.

в) В качестве примера п. б) взять уравнение, в котором c имеет значение -64. Найти остальные два корня уравнения, которое приведено в пункте б), выразив их через a, и найти значения a, при которых эти два корня будут действительными числами.

2. В полушар радиуса r вписана треугольная пирамида $ABCT$. Основание пирамиды есть равнобедренный треугольник ABC; основанием этого треугольника, вписанного в круговое основание полусферы, является BC. Вершина пирамиды T лежит на перпендикуляре к основанию пирамиды над серединой отрезка BC.

а) Предполагая расстояние точки A от BC равным x, вычислить объем пирамиды через r и x.

б) Если радиус r постоянный, то пирамида имеет максимальный объем для некоторого значения x. Найти максимальное значение объема и соответствующее значение x (с доказательством).

в) Пусть $x = 3/2\ r$ и проведены две плоскости через перпендикулярный к основанию полусферы радиус, причем одна плоскость содержит точку B, другая — точку C. Кроме того, проведена еще плоскость BCT. Вычислить через r площадь части поверхности полусферы, вырезанной этими тремя плоскостями.

3. а) Даны эллипс $4x^2 + 9y^2 = 36$ и прямая $x = 3$. Некоторая точка P на этой прямой имеет ординату q. Написать уравнение поляры точки P относительно эллипса. Показать, что поляра всегда проходит через постоянную точку A при движении P по прямой и найти координаты точки A.

б) Дана еще точка M с координатами $(m, 0)$. Найти уравнение геометрического места точек пересечения поляры точки P и прямой, проходящей через точки M и P, когда P движется по прямой $x = 3$. Полученные кривые при различных значениях m всегда проходят через A и M. Объяснить, как можно было бы предвидеть наперед, что это должно иметь место.

в) Геометрическое место, указанное в п. б) имеет уравнение

$$4x^2 + 3(3 - m)y^2 - 4(m + 3)x + 12m = 0$$

Описать конические сечения, которые это уравнение представляет при различных конечных значениях m (указать только типы кривых без нахождения центра, оси, длин осей и т. д.). Какую геометрическую интерпретацию имеет это уравнение при m, равном 3?

II. Задачи математического конкурса для учащихся скандинавских стран в 1957 г. [1])

1. P — точка на эллипсе, Q — проекция центра эллипса на касательную, проведенную к эллипсу в точке P. Показать, что максимальное значение расстояния PQ равно половине разности длин большой и малой осей эллипса.

2. Дан треугольник ABC и произвольная точка P в пространстве. Доказать, что

$$3(PA^2 + PB^2 + PC^2) \geqslant AB^2 + BC^2 + CA^2.$$

При каком положении точки P имеет место знак равенства?

3. Дан треугольник ABC. На сторонах BC, AC и AB взяты точки A_1, B_1 и C_1 и введены обозначения

$$\frac{BA_1}{A_1C} = \alpha; \quad \frac{CB_1}{B_1A} = \beta; \quad \frac{AC_1}{C_1B} = \gamma.$$

педагогического института в Осло): H. F. Fehr, K. Piene, Mathematics instruction in the secondary schools of Norway, журнал School Science and Mathematics, ноябрь 1956 г.

[1]) В приложениях к настоящей заметке приведены варианты письменных работ по математике для оканчивающих среднюю школу, а также задачи математического конкурса для учащихся скандинавских стран. Эти материалы ярко характеризуют уровень знаний по математике в норвежских средних школах.

[1]) В соревновании участвовали учащиеся Дании, Исландии, Норвегии, Финляндии и Швеции. Каждый участник конкурса присылал свои решения вместе с подпиской о том, что решения выполнены им без посторонней помощи.

Отрезки AA_1, BB_1 и CC_1 образуют внутри данного треугольника другой меньший треугольник. Найти отношение площадей двух треугольников через α, β, γ. Результат должен быть максимально упрощен.

4. Доказать, что функция

$$y=\frac{1}{\sin x}-\frac{1}{x}$$

возрастает с возрастанием x в интервале $0<x\leqslant\frac{\pi}{2}$.

Найти $\lim y$, когда $x\longrightarrow 0$.

5. Дана последовательность чисел

$$a_1=1,\quad a_2=1,\quad a_3=2,\quad a_4=3,\quad a_5=5,\ldots,\quad a_{n+2}=a_{n+1}+a_n,\ldots$$

(название «последовательность Фибоначчи» не употреблено). Обозначим сумму n первых членов ее через S_n.

а) Выразить S_n через один из членов последовательности.

б) Показать, что

$$a_n\cdot a_{n+2}-a_{n+1}^2=(-1)^{n+1}.$$

в) Найти

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}.$$

6. Дан выпуклый многоугольник с n сторонами, причем а) никакие две диагонали не параллельны; б) никакие три диагонали не проходят через одну точку (не считая вершины). Сколько точек пересечения диагоналей можно найти внутри многоугольника и сколько вне его?

XXI ШКОЛЬНАЯ МАТЕМАТИЧЕСКАЯ ОЛИМПИАДА В МОСКВЕ

В. Г. Болтянский и Э. Р. Розендорн

(Москва)

Ежегодное проведение математических олимпиад московских школьников стало прекрасной традицией. Олимпиады для учащихся VII-X классов подготовляет и проводит Механико-математический факультет МГУ совместно с Мосгороно и Московским математическим обществом. В апреле 1958 г. состоялась 21-я олимпиада.

Помимо подготовки и проведения самой олимпиады, в Московском университете ежегодно работают школьные математические кружки и читаются лекции по математике, что также вошло в традицию.

Чтение лекций проводится по воскресеньям (раз в две недели) профессорами и преподавателями МГУ и других московских вузов, причем, как правило, читаются две лекции одновременно: одна для учащихся VII—VIII классов и другая — для учащихся IX—X классов. В 1957—58 учебном году были прочитаны следующие лекции:

Для школьников VII—VIII классов

1. «Деление отрезка в данном отношении» (*Н. М. Бескин*).
2. «Геометрическая алгебра» (*А. М. Лопшиц*).
3. «Системы счисления» (*Н. Я. Виленкин*).
4. «Бесконечнозначные числа» (*В. Г. Болтянский*).
5. «Арифметика вычетов и ее применение в технике» (*А. Н. Колмогоров*).
6. Равновеликость и равносоставленность» (*А. И. Савин*).
7. «Думающие машины» (*А. А. Ляпунов*).
8. «Суммирование последовательностей» (*П. Л. Ульянов*).

Для школьников IX—X классов

1. «Теорема Эйлера о многогранниках и топология поверхностей» (*В. Г. Болтянский*).
2. «Основные понятия современной алгебры» (*Л. А. Скорняков*).
3. «Задача Дирихле» (*Е. М. Ландис*).
4. «Натуральные логарифмы» (*А. Д. Соловьев*).
5. «Комплексные числа и их применение в геометрии» (*И. М. Яглом*).
6. «Приближенные вычисления с помощью непрерывных дробей» (*Л. А. Люстерник*).

7. «Задача Дирихле» (продолжение) (*Е. М. Ландис*).
8. «Как строить график» (*Г. Е. Шилов*).
9. «О «производной» (*Г. Е. Шилов*).
10. «Математические методы в лингвистике» (*Р. Л. Добрушин*).
11. «Геометрические задачи теории приближения функций» (*С. Б. Стечкин*).

Занятия кружка проводились еженедельно с 20 сентября до окончания олимпиады. Работа кружка проводилась по секциям под общим руководством Э. Р. Розендорна. Всего работало 11 секций: 5 секций для X класса, 3 — для IX класса, 2 — для VIII класса и 1 — для VII класса.

В руководстве секциями участвовало 40 студентов и аспирантов механико-математического факультета МГУ.

Старшие руководители секций: Брюно (I курс), Емельянов, Маркова, Произволов (II курс), Вайнберг, Грушин, Олевский (III курс), Кристи, Розендорн, Смоляк, Хазен (IV курс).

Одна из секций кружка для 10 класса (под руководством студента Смоляка) занималась только алгеброй; остальные секции не были специализированными: они занимались разными вопросами алгебры, геометрии, теории чисел, теории вероятностей и т. д. Все занятия секций кружка проводились в старом здании МГУ.

Многолетний опыт показывает, что участие в работе секций кружка значительно повышает общий математический кругозор учащихся, помогает выработать навыки в решении трудных математических задач, развивает сообразительность. Большинство учащихся, регулярно посещавших занятия кружка, впоследствии становятся хорошими студентами МГУ, педагогических институтов, технических вузов. Подавляющее большинство кружковцев принимает участие и в олимпиаде. Многие из них оказываются в числе победителей.

Однако не следует думать, что именно кружковцы составляют основную массу участников или победителей олимпиады. Достаточно сказать, что даже в начале работы кружков на их занятия приходило около 600 учащихся VII — X классов, в то время как общее число участников олимпиады составило 1225 человек. Нередко школьник, впервые участвующий в олимпиаде, добивается в ней хороших результатов и, почувствовав интерес к математике, становится впоследствии активным участником кружка.

Организацией лекций, работой кружков, подготовкой и проведением олимпиады руководил оргкомитет из 14 человек (председатель В. Г. Болтянский). Оргкомитет поддерживал контакт с Московским математическим обществом и общественными организациями механико-математического факультета МГУ. Впервые вся олимпиада проводилась в новом здании МГУ, что оказалось очень удобным и, можно надеяться, войдет в традицию.

Сборник задач для подготовки к олимпиаде, составленный по материалам работы школьного кружка, был выпущен тиражом

6000 экземпляров. Однако ввиду того, что он вышел перед самой олимпиадой, для подготовки главным образом использовался прошлогодний сборник. Часть тиража нового сборника оставлена для работы кружка в 1958/59 учебном году.

При составлении задач для олимпиады члены оргкомитета стремились к тому, чтобы избежать дублирования одних и тех же задач для учащихся разных классов. Хотя в целом совокупность задач была удачной, распределение задач между I и II турами было неправильным. Так, например, задача 4, предложенная на II туре учащимся X класса, была, несомненно, чересчур легкой, а задача 1, предложенная учащимся X класса на I туре, была слишком трудной, тем более для I тура. Такое неправильное распределение произошло потому, что оргкомитет не имел полного списка задач перед началом I тура.

Большинство задач для II тура было составлено между I и II турами. Бо́льшая часть задач была составлена В. Г. Болтянским, студентами Смоляком, Ветухновским, Шуром, Кирилловым и другими членами оргкомитета и руководителями кружков.

Все задачи обсуждались на заседаниях оргкомитета. В результате этих обсуждений нередко бывало, что задачи изменялись до неузнаваемости. Задачи олимпиады в целом являются, таким образом, плодом коллективного творчества оргкомитета.

На I тур олимпиады, который состоялся в воскресенье 6 апреля 1958 г., пришло 1190 учащихся. Разбор задач I тура проводился в следующее воскресенье 13 апреля. Во II туре (20 апреля) приняло участие 349 школьников, в том числе 35 человек, не участвовавших в I туре, но допущенных к участию в олимпиаде по ходатайству учителей и руководителей кружков. 27 апреля состоялся разбор задач II тура и было проведено заключительное заседание олимпиады с вручением премий и грамот. Количество участников олимпиады по классам указано в следующей таблице:

Классы	Участвовали в I туре	Участвовали во II туре	Успешно прошли олимпиаду	Награждены	
				похвальными отзывами	премиями
VII	100	41	22	5	4
VIII	125	21	11	4	4
IX	235	83	12	2	1
X	730	204	56	19	7
Всего	1190	349	101	30	16

В проверке работ олимпиады участвовало около 50 студентов механико-математического факультета, которые были разбиты на отдельные «кусты». Первоначально проверка и обсуждение работ

проводились по «кустам», во главе каждого из которых стоял выделенный оргкомитетом «кустарь». Результаты этой предварительной проверки обсуждались затем на общем собрании оргкомитета, причем ряд работ, в том числе все работы, заслуживающие премий или отзывов, перед обсуждением зачитывались вслух.

Согласно традиции оргкомитетом была принята следующая система оценок:

$+$, если решение верное;
$\pm$ » » содержит легко исправимые ошибки;
$\mp$ » » неверное, но содержит верные идеи;
$-$ » » в принципе неверное;
0 » задача не решалась.

Кроме того, в некоторых исключительных случаях применялась оценка 1/2 (за 1 принимался $+$). Особенно выдающиеся решения выделялись знаком «!».

Всего был признан успешно прошедшим олимпиаду 101 учащийся (количество по классам было указано в таблице). Список успешно прошедших был оглашен на заключительном заседании. Из этого общего числа успешно прошедших 16 учащихся были награждены первыми, вторыми или третьими премиями и 30 человек получили похвальные отзывы.

Классы	Первые премии	Вторые премии	Третьи премии
VII	—	Гельфанд С. И., 151 шк. Чуркин В. П., 646 шк.	Гурин Г. Л., 310 шк. Дунаевский Л. В., 59 шк.
VIII	Ильяшенко Ю., 59 шк.	—	Кронрод В. А., 7 шк. Соколовская Т. В., 275 шк. Четаев А. Н., 5 шк.
IX	—	—	Тоом А. Л., 69 шк.
X	Леонтович А. М., Хотьково, 1 шк. Хазен М. М., 665 шк.	Кондорская Е. Е., 131 шк.	Алексеев И. Я., 279 шк. Голо В. Л., 510 шк. Леман А. А., 103 шк. Пресман Э. Л., 124 шк.

Как видно из приведенной таблицы, оргкомитет не придерживался того мнения, что по каждому классу должна быть выдана одна первая, одна вторая и одна третья премии. Судить о том, почему, например, лучшие из работ семиклассников должны получить вторую, а не первую премию, оргкомитет мог как в результате сравнения с работами учащихся других классов, так и на основании опыта предыдущих олимпиад.

Отметим, что в порядке исключения двум ученикам X классов, не прошедшим успешно II тура были выданы похвальные отзывы за особо интересное решение задачи 5 I тура.

Наилучшие результаты в олимпиаде показали учащиеся школы № 1 Октябрьской ж. д., школ №: 1, 7, 24, 50, 56, 59, 69, 103, 144, 167, 221, 231, 475, 556, 618, 646, 665 города Москвы.

За успехи, достигнутые их учениками на олимпиаде, грамоты получили преподаватели математики Бабичева Г. М. (59 шк.), Глаголев Н. С. (231 шк.), Ефремов В. А. (7 шк.), Морозкин И. В. (59 шк.), Морозов М. А. (1 шк.), Невельская И. М. (1 шк. Окт. ж. д.), Николаева М. О. (646 шк.), Остроумова М. А. (310 шк.), Покровская Н. А. (59 шк.), Токарь Н. А. (103 шк.), Фиделли Т. Н. (59 шк.), Шемякин А. И. (59 шк.), Шершевский А. А. (69 шк.).

Прежде чем закончить свою работу, оргкомитет по традиции дал рекомендации относительно состава оргкомитета для проведения очередной, 22-й по счету, математической олимпиады московских школьников.

Задачи, предложенные на XXI олимпиаде

I. Для VII класса

Первый тур

1. Имеется система уравнений

$$*x + *y + *z = 0,$$
$$*x + *y + *z = 0,$$
$$*x + *y + *z = 0.$$

Два человека поочередно вписывают вместо звездочек числа. Доказать, что начинающий всегда может добиться того, чтобы система имела ненулевое решение.

2. В круге проведены два диаметра AB и CD. Доказать, что если M — произвольная точка окружности, а P и Q — ее проекции на диаметры AB и CD, то длина отрезка PQ не зависит от выбора точки M.

3. Сколько существует четырехзначных номеров (от 0001 до 9999), у которых сумма двух первых цифр равна сумме двух последних цифр?

4. На плоскости даны точки A и B. Построить такой квадрат, чтобы точки A и B лежали на его границе и сумма расстояний от точки A до вершин квадрата была наименьшей.

5. Дана следующая треугольная таблица чисел:

$$\begin{array}{ccccccccccccc} 0 & & 1 & & 2 & & \cdots & & \cdots & & 1957 & & 1958 \\ & 1 & & 3 & & \cdots & & \cdots & & 3915 & & & \\ & & \cdots & & \cdots & & \cdots & & & & & & \\ & & & & \cdots & & \cdots & & & & & & \end{array}$$

Каждое число (кроме чисел верхней строчки) равно сумме двух ближайших чисел предыдущей строчки. Доказать, что число, стоящее в самой нижней строчке, делится на 1958.

Второй тур

1. Доказать, что на плоскости нельзя расположить больше четырех выпуклых многоугольников так, чтобы каждые два из них имели общую сторону.

2. Имеются два набора из $+1$ и -1, в каждом по 1958 чисел. Доказать, что за некоторое число шагов можно превратить первый набор во второй, если на каждом шагу разрешается одновременно изменить знак у любых 11 чисел первого набора. (Два набора считаются одинаковыми, если у них на одинаковых местах стоят одинаковые числа).

3. Каждая грань куба заклеивается двумя равными прямоугольными треугольниками с общей гипотенузой, один из которых белый, другой — черный. Можно ли эти треугольники расположить так, чтобы при каждой вершине куба сумма белых углов была равна сумме черных углов?

4. Доказать, что если целое $n > 2$, то

$$(1 \cdot 2 \ldots n)^2 > n^n.$$

5. Сторона клетки клетчатой бумаги равна 1. По линиям сетки построен прямоугольник со сторонами m и n. Можно ли в прямоугольнике провести по линиям сетки замкнутую ломаную, которая ровно один раз проходила бы через каждый узел сетки, расположенный внутри или на границе прямоугольника? Если можно, то какова ее длина?

II. Для VIII класса

Первый тур

1. Внутри треугольника ABC взята точка O. На лучах OA, OB и OC построены векторы единичной длины. Доказать, что сумма этих векторов имеет длину меньшую единицы.

2. Доказать, что если уравнения с целыми коэффициентами

$$x^2 + p_1 x + q_1 = 0,$$
$$x^2 + p_2 x + q_2 = 0$$

имеют общий нецелый корень, то $p_1 = p_2$ и $q_1 = q_2$.

3. На круглой поляне радиуса R растут три круглые сосны одинакового диаметра. Центры их стволов находятся на расстоянии $\frac{R}{2}$ от центра поляны в вершинах равностороннего треугольника. Два человека, выйдя одновременно из диаметрально противоположных точек поляны, обходят поляну по краю с одинаковой скоростью и в одном направлении и всё время не видят друг друга. Увидят ли друг друга три человека, если они так же будут обходить поляну, выйдя из точек, находящихся в вершинах вписанного в поляну правильного треугольника?

4. Решить в целых положительных числах уравнение

$$1 - \cfrac{1}{2 + \cfrac{1}{3 + \cdots + \cfrac{1}{1958}}} = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cdots + \cfrac{1}{x_{1958}}}}.$$

5. Проекции многоугольника на ось OX, биссектрису 1-го и 3-го координатных углов, ось OY и биссектрису 2-го и 4-го координатных углов равны соответственно 4, $3\sqrt{2}$, 5, $4\sqrt{2}$. Площадь многоугольника — S. Доказать, что $S \leqslant 17{,}5$.

Второй тур

1. Из бумаги вырезан многоугольник. Две точки его границы соединяются отрезком, по которому многоугольник складывается. Доказать, что периметр многоугольника, получающегося после складывания, меньше периметра исходного многоугольника.

2. Для любых чисел a_1 и a_2, удовлетворяющих условиям $a_1 \geqslant 0$, $a_2 \geqslant 0$, $a_1 + a_2 = 1$, можно найти такие числа b_1 и b_2, что $b_1 \geqslant 0$, $b_2 \geqslant 0$, $b_1 + b_2 = 1$, $\left(\frac{5}{4} - a_1\right) b_1 + 3\left(\frac{5}{4} - a_2\right) b_2 > 1$.

Доказать.

3. Внутри $\angle AOB$ взята точка C. Из нее опущены перпендикуляры: CD на сторону OA; CE — на сторону OB. Из точек D и E опущены перпендикуляры: EM на сторону OA; DN — на сторону OB. Доказать, что $OC \perp MN$.

4. Доказать, что если целое $n > 0$, то

$$1^1 \cdot 2^2 \cdot 3^3 \ldots n^n < n^{\frac{n(n+1)}{2}}.$$

5. Обозначим через a наибольшее число непересекающихся кругов диаметра 1, центры которых лежат внутри многоугольника M, через b — наименьшее число кругов радиуса 1, которыми можно покрыть весь многоугольник M.

Какое число больше: a или b?

III. Для IX класса

Первый тур

1. Бесконечная плоская ломаная $A_0 A_1 \ldots A_n \ldots$, все углы которой прямые, начинается в точке A_0 с координатами $x = 0$, $y = 1$ и обходит начало координат O по часовой стрелке. Первое звено ломаной имеет длину 2 и параллельно биссектрисе 4-го координатного угла. Каждое из следующих звеньев пересекает одну из координатных осей и имеет наименьшую возможную при этом целочисленную длину. Расстояние $OA_n = l_n$. Сумма длин первых n звеньев ломаной равна s_n. Доказать, что найдется n, для которого $\frac{s_n}{l_n} > 1958$.

2. Доказать, что если $|ax^2 - bx + c| < 1$ при любом x, $|x| \leqslant 1$, то и $|(a+b)x^2 + c| < 1$ при $|x| \leqslant 1$.

3. Какое наименьшее число осей симметрии может иметь пространственная фигура, состоящая из трех прямых, из которых никакие две не параллельны и не совпадают?

4. Решить в целых положительных числах уравнение

$$1-\cfrac{1}{2+\cfrac{1}{3+\cfrac{1}{\ddots+\cfrac{1}{n}}}}=\cfrac{1}{x_1+\cfrac{1}{x_2+\cfrac{1}{\ddots+\cfrac{1}{x_n}}}}.$$

5. Отрезок длиной 3^n разбивается на три равные части. Первая и третья из них называются отмеченными. Каждый из отмеченных отрезков разбивается на три части, из которых первая и третья снова называются отмеченными и т. д. до тех пор, пока не получатся отрезки длиной 1. Концы всех отмеченных отрезков называются отмеченными точками. Доказать, что для любого целого k $(1 \leqslant k \leqslant 3^n)$ можно найти две отмеченные точки, расстояние между которыми равно k.

Второй тур

1. Решить в целых положительных числах уравнение

$$x^{2y-1}+(x+1)^{2y-1}=(x+2)^{2y-1}.$$

2. Провести из точки O n лучей на плоскости так, чтобы сумма всех попарных углов между ними была наибольшей. (Рассматриваются только углы, не превышающие 180°.)

3. Игральная доска имеет форму ромба с углом 60°. Каждая сторона ромба разделена на 9 частей. Через точки деления проведены прямые, параллельные сторонам и малой диагонали ромба, разбивающие доску на треугольные клетки. Если на некоторой клетке поставлена фишка, проведем через эту клетку 3 прямые, параллельные сторонам и малой диагонали ромба. Клетки, которые они пересекут, будут считаться побитыми фишкой. Каким наименьшим числом фишек можно побить все клетки доски?

4. Обозначим через a наименьшее число кругов радиуса 1, которыми можно полностью покрыть заданный многоугольник M, через b — наибольшее число непересекающихся кругов радиуса 1 с центрами внутри многоугольника M. Какое из чисел больше, a или b?

5. Между зажимами A и B включено несколько сопротивлений. Каждое сопротивление имеет входной и выходной зажимы. Какое наименьшее число сопротивлений необходимо иметь и какова может быть схема их соединения, чтобы при порче любых 9 сопротивлений цепь оставалась соединяющей зажимы A и B, но не было короткого замыкания? (Порча сопротивления: короткое замыкание или обрыв.)

IV. Для X класса

Первый тур

1. Проекции плоского выпуклого многоугольника на ось OX, биссектрису 1-го и 3-го координатных углов, ось OY и биссектрису 2-го и 4-го координатных углов соответственно равны 4, $3\sqrt{2}$, 5, $4\sqrt{2}$. Площадь многоугольника равна S. Доказать, что $S \geqslant 10$.

2. Доказать, что

$$1155^{1958}+34^{1958} \neq n^2,$$

где n — целое.

3. Какое наибольшее число осей симметрии может иметь пространственная фигура, состоящая из трех прямых, из которых никакие две не параллельны и не совпадают?

4. На стол кладут правильный 100-угольник, в вершинах которого написаны числа 1, 2, ... 100. Затем эти числа переписывают в порядке удаления от переднего края стола. Если две вершины находятся на равном расстоянии от края, сначала выписывается левое число, затем правое. Выписаны всевозможные различные наборы чисел, соответствующие разным положениям 100-угольника. Вычислить сумму чисел, стоящих в этих наборах на 13-х местах слева.

5. Из четырех прямых на плоскости никакие две не параллельны, никакие три не пересекаются в одной точке. По каждой прямой с постоянной скоростью идет пешеход. Известно, что 1-й встречается со 2-м, с 3-м и с 4-м, а 2-й встречается с 3-м и с 4-м. Доказать, что 3-й пешеход встретится с 4-м.

Второй тур

1. Решить в целых положительных числах уравнение

$$x^{2y}+(x+1)^{2y}=(x+2)^{2y}.$$

2. В многоугольнике существуют такие точки A и B, что любая соединяющая их ломаная, проходящая внутри или по границе многоугольника, имеет длину больше 1. Доказать, что периметр многоугольника больше 2.

3. В школе изучают $2n$ предметов. Все ученики учатся на 4 и 5. Никакие два ученика не учатся одинаково, ни про каких двух нельзя сказать, что один из них учится лучше другого. Доказать, что число учеников в школе $\leqslant C_{2n}^{n}$.

4. Стороны параллелограмма равны a и b. Найти отношение объемов тел, полученных при вращении параллелограмма вокруг стороны a и вокруг стороны b.

5. На n карточках написаны с разных сторон числа — на 1-й: 0 и 1; на 2-й: 1 и 2; ...; на n-й: $n-1$ и n.

Один человек берет из стопки несколько карточек и показывает второму одну сторону каждой из них. Затем берет из стопки еще одну карточку и тоже показывает одну сторону.

Указать все случаи, в которых второй может определить число, написанное на обороте последней показанной ему карточки.

ИНДУКТИВНАЯ ИГРА

Все распространенные настольные игры (начиная от шахмат и кончая домино) основаны на дедуктивном процессе мышления: ***игрок должен выбрать наиболее рациональную стратегию на основе знания ситуации и правил игры.*** *В журнале «Scientific American» за июнь 1959 г. опубликована статья, в которой описана новая игра, изобретенная молодым нью-йоркским писателем Робертом Эббетом и названная им «Элеузис». Правила этой игры моделируют индуктивный процесс мышления:* ***игрок должен открыть неизвестные ему правила игры на основе наблюдения и сознательного эксперимента.*** *Здесь игрок находится не в положении математика, выводящего следствия из аксиом, а в положении физика, раскрывающего закон природы. Мы приведем правила несколько упрощенного варианта этой игры. Материалом для игры может служить любой набор предметов, допускающих естественную классификацию по нескольким признакам (кости домино, картинки детского лото и т. п.).*

Для наглядности предположим, что в распоряжении играющих есть полная колода из 52 игральных карт. Мы опишем правила игры для трех играющих, один из которых (попеременно в каждом туре игры) является «водящим». Читатель легко додумает правила игры и для большего числа играющих.

«Водящий» раздает все карты поровну двум игрокам и затем задумывает и фиксирует на бумаге (втайне от других игроков) некоторое правило, по которому должна выкладываться колода карт. Правила должны быть простыми, но могут быть весьма разнообразными. Например: 1° каждая следующая карта должна совпадать с предшествующей по масти или значению; 2° остатки от деления значения карты на три (при этом валет считается за 11, дама — за 12, король — за 13) должны меняться циклически (0, 1, 2, 0, 1, 2); 3° после карты черной масти должна идти карта с номером от 1 до 7, а после карты красной масти должна идти карта с номером от 8 до 13; 4° после карты красной масти должна идти карта черной масти той же четности, а после карты черной масти — карта красной масти противоположной четности и т. д. Правило не должно быть слишком стеснительным. Все карты во время игры держатся открытыми (единственный секрет — задуманное правило игры).

Первый из играющих вынимает произвольную карту, начиная ею последовательность. Затем игроки поочередно выбирают одну

ОКОНЧАНИЕ ТЕКСТА НА СТР. 328

СООБЩЕНИЯ О ШКОЛЬНЫХ МАТЕМАТИЧЕСКИХ ОЛИМПИАДАХ

В редакцию «Математического просвещения» продолжают поступать сообщения о математических олимпиадах, проводящихся для школьников в различных местах Советского Союза. В настоящем выпуске приводятся (в сокращенном виде) три таких сообщения.

1. МОЛДАВСКАЯ ССР

В 1956 г. в Молдавской ССР была организована и проведена первая республиканская математическая олимпиада школьников. Она вызвала большой интерес у учащихся — достаточно сказать, что в первом туре приняло участие свыше 1300 школьников VIII — X классов. Олимпиаде предшествовала большая работа кафедры математического анализа Кишиневского университета. Начиная с 1948 г. при физико-математическом факультете университета работает математический кружок для школьников, где решаются разные математические задачи и читаются лекции по математике, физике и механике. Так, в 1957/58 учебном году были прочитаны лекции на темы: искусственные спутники Земли, понятие о производной, необходимый и достаточный признаки, метод полной индукции, исследование функций, понятие о теории множеств, понятие о максимуме и минимуме и др.

В течение 7 лет (1949—1956) олимпиады проводились только для учащихся г. Кишинева; начиная с 1956 г. они проводятся в республиканском масштабе. Олимпиады организует Министерство просвещения республики совместно с Кишиневскими пединститутом и университетом.

Каждая олимпиада проводится в два тура. Первый тур проводится для школьников одного из районов или городов в одной из местных школ под руководством рай(гор)оргкомитета. Задачи для первого тура присылаются республиканским оргкомитетом — 4 задачи и примера, для решения которых отводится не более 5 часов. При проверке работ засчитывается: за верное решение 2 очка, за верное, но неполное — 1 очко, за неправильное решение — 0 очков. Решения могут быть написаны на русском, молдавском или на украинском языке.

После окончания первого тура рай(гор)оргкомитет присылает республиканскому оргкомитету те работы, за которые засчитано не менее 4 очков, и сообщает общее число участников олимпиады по каждому классу. Республиканский оргкомитет вновь просматривает эти работы и отбирает победителей первого тура для участия

во втором туре, который проводится в Кишиневе. Каждому его участнику также предлагаются 4 задачи, но уже повышенной трудности. Приводим результаты олимпиад за 1956/57 и 1957/58 гг.:

Год	Допущено ко II туру	Из них по классам			Явилось на II тур	Из них по классам			Из них премировано			
		VIII	IX	X		VIII	IX	X	Всего	VIII кл.	IX кл.	X кл.
1957	201	16	92	93	152	10	60	82	27	3	5	19
1958	180	24	51	105	171	15	51	105	47	8	16	23

При подведении итогов второго тура олимпиады проводится беседа с ее участниками и оглашаются результаты соревнования. Эта беседа носит торжественный характер. Министерство просвещения республики выдает победителям почетные грамоты и премирует их книгами.

Задачи, предлагавшиеся на втором туре

Для VIII класса

1. Доказать, что если $\sqrt[3]{a}+\sqrt[3]{b}+\sqrt[3]{c}=0$, то $(a+b+c)^3=27abc$.
2. Построить треугольник по двум сторонам b и c и биссектрисе l_A угла между ними.
3. Решить уравнение $\sqrt[3]{x+45}-\sqrt[3]{x-16}=1$.
4. На данной прямой OM найти точку, равноотстоящую от другой данной прямой OA и от данной точки S.

Для IX класса

1. Решить уравнение $\sqrt[m]{x(x+a)^2}+2\sqrt[m]{x^3}=3\sqrt[m]{x^2(x+a)}$ $(a>0)$.
2. При каких соотношениях между a и b, c и d имеет место равенство
$$(\sqrt{(a-b)^2}+\sqrt{(c-d)^2})^2=(a-b-c+d)^2.$$
3. $ABCDEFGHIK$ — правильный 10-угольник, вписанный в окружность с центром O; AD — его диагональ. Доказать, что $AD=OA+AB$.
4. Доказать, что если между углами треугольника существует соотношение $\sin C=\cos A+\cos B$, то треугольник прямоугольный.

Для X класса

1. Доказать, что уравнение $x^2+px+q=0$ не имеет рациональных корней, если p и q — нечетные числа.
2. Доказать, что между углами и сторонами треугольника имеет место соотношение
$$\frac{\sin(A-B)}{\sin(A+B)}=\frac{a^2-b^2}{c^2}.$$
3. Решить уравнение $|z|+3z=14-12i$, где $|z|$ — модуль числа z.
4. Доказать, что в любом треугольнике имеет место соотношение
$$2S_\Delta=\frac{a^2-b^2}{\operatorname{ctg}B-\operatorname{ctg}A}.$$

К. Спатар (Кишинев)

2. БЕЛГОРОДСКАЯ ОБЛАСТЬ

В 1959 и 1960 гг. впервые в Белгородской области были проведены математические олимпиады для школьников старших классов. Инициатором явилась кафедра математики педагогического института, переведенного в 1957 г. в областной центр из г. Старый Оскол.

Задачи первого тура публиковались в областной молодежной газете «Ленинская смена». Одновременно областной отдел народного образования направил во все школы письма, разъясняющие положение об олимпиаде. Такая организация обеспечила участие сельских школьников. Для старшеклассников г. Белгорода первый тур проводился непосредственно в школах.

Олимпиадам предшествовала большая работа преподавателей института, учителей и студентов физико-математического факультета института, проходивших педагогическую практику во всех школах Белгорода и в ряде сельских школ. Для учащихся были прочитаны лекции: «Что такое топология» (Е. С. Тихомирова), «Кибернетика и современные вычислительные машины» (Ю. И. Соколовский), «Решение задач на максимум и минимум» (И. А. Абрамов) и др.

В обеих олимпиадах приняло участие 260 учащихся из 56 школ, в том числе 196 из школ области. Жюри олимпиады допустило ко второму туру 167 участников, решивших не менее 3 задач из 5.

Заключительная часть олимпиад проводилась в Белгороде, куда съезжались победители первого тура. Победители олимпиад награждены почетными грамотами и памятными подарками. Особые призы вручены участникам, представившим лучшие решения задач.

Наряду с традиционными задачами, выясняющими общее математическое развитие учеников, предлагались и такие, которые требовали практической сметки. Предлагалось, например, определить радиус материального шара при помощи циркуля и линейки.

Олимпиады выявили ряд способных учащихся. Но они показали и наличие пробелов по таким темам, как решение, тригонометрические функции «абстрактного» аргумента, определение логарифма, исследование квадратного трехчлена, вычисление с помощью линейки. Например, предложенное во втором туре для X класса уравнение $\cos(\cos x)=0$ не смог решить ни один участник. Характерно, что примеры, требовавшие главным образом умения выполнять тождественные преобразования, даже сложные, не вызывали особых затруднений. Наибольшую трудность представили задачи хотя и несложные, но к решению которых нужно было искать новый подход.

Несмотря на некоторые недочеты, первый опыт проведения олимпиад в Белгородской области можно считать удавшимся. Соревнования юных математиков решено сделать ежегодными.

Ю. К. Василенко (Белгород)

3. ГОР. ЛЬВОВ

После пятилетнего перерыва (с 1956 г.) возобновились математические олимпиады школьников в Львове. Весной 1960 г. состоялась одиннадцатая олимпиада, организованная Львовским университетом с привлечением студентов 4-го курса, проходящих в школах педагогическую практику.

Первый тур олимпиады проводился заочно в виде домашних заданий, разосланных по школам. Прошедшие успешно первый тур были допущены ко второму, который состоялся 15 мая в университете.

Олимпиада состояла из двух секций: VII—VIII классы и IX—X классы. На втором туре на каждой секции было предложено по 10 задач: 5 сравнительно простых и 5 повышенной трудности. Участники должны были выбрать по одной задаче из каждой категории.

В конкурсе приняло участие 66 учащихся из 13 школ. Результаты конкурса оказались следующими:

Классы	Число участников	Полностью (2 задачи)	Неполностью	Не решили	Получили грамоты
VII—VIII	27	14	8	5	4
IX—X	39	23	6	10	9

Ниже приводятся образцы задач, предлагавшихся на олимпиаде (во втором туре).

Для VII—VIII классов

1. Доказать, что p^q+q^p делится на $p+q$, если p и q являются простыми числами — близнецами.

2. Решить уравнение $(x+2)!-(x+1)!-x!=x^4+x^2$.

3. Доказать, что заключенный между боковыми сторонами трапеции отрезок прямой, проходящей через точку пересечения диагоналей параллельно основаниям трапеции, равен среднему гармоническому оснований.

Для IX—X классов

1. Вычислить следующие суммы:

$$C_n^0+C_n^4+C_n^8+\ldots, \qquad C_n^1+C_n^5+C_n^9+\ldots$$

$$C_n^2+C_n^6+C_n^{10}+\ldots, \qquad C_n^3+C_n^7+C_n^{11}\ \ldots$$

2. Проверить (без помощи таблиц) равенство

$$\cos^2 73^\circ+\cos^2 47^\circ+\cos 73^\circ\cdot\cos 47^\circ=\frac{3}{4}.$$

А. С. Кованько (Львов)

НОВОСТИ МАТЕМАТИЧЕСКОЙ НАУКИ

1. **П. Эрдёш, П. Туран и др.** *Некоторые новые результаты в области аддитивной теории чисел.* Будем рассматривать такие всевозможные наборы различных целых положительных чисел, не превосходящих n, что суммы любых двух чисел набора все различны между собой. Пусть m — наибольшее возможное число таких чисел в наборе; тогда $m=f(n)$, т. е. мы определили некоторую теоретико-числовую функцию. Английский математик Сидон (Sidon) поставил задачу о вычислении значений этой функции. Легко, например, видеть, что $m=3$, если $n=4$. В самом деле, набор $\{1, 2, 3\}$ удовлетворяет требуемым условиям, а для единственного возможного набора из четырех чисел $\{1, 2, 3, 4\}$ имеем $1+4=2+3$.

Туран и автор [1] показали, что

$$f(n)<\sqrt{n}+\sqrt[4]{n}, \tag{1}$$

а американскому математику Зингеру [2] удалось доказать, что для бесконечно многих n имеет место неравенство

$$f(n)>\sqrt{n}. \tag{2}$$

Возможно, что

$$f(n)=\sqrt{n}+O(1)\ ^1).$$

Для доказательства Зингер вводит понятие полного множества разностей: если для множества 0, 1, 2, ..., $n-1$ вычетов по модулю n можно указать такое его подмножество $a_1, a_2, \ldots, a_k$, что любое m $(0<m\leqslant n-1)$ однозначно представляется либо в виде a_j-a_i, либо в виде a_j-a_i+n, то найденное подмножество называют *полным множеством разностей*. Поскольку различных разностей a_j-a_i насчитывается $k(k-1)$, то из существования полного множества разностей вытекает, что $n-1=k(k-1)$ или $n=k^2-k+1$. Для множества 0, 1, 2, 3, 4, 5, 6 всех вычетов по модулю 7 можно, например, указать подмножество 1, 2, 6, которое, как нетрудно проверить, является полным множеством разностей (здесь $n=7$, $k=3$). Зингер доказал, что для существования

[1]) Напомним, что символ $O(1)$ означает некоторую ограниченную величину. (*Прим. ред.*)

полного множества разностей достаточно, чтобы k было степенью простого числа. Согласно предположению Зингера, это условие также и необходимо. В некоторых частных случаях это последнее предположение было доказано Бруком и Ризером [3]. В числе неразобранных остается, например, случай $k=10$.

Из (1) и результата Зингера можно вывести, что

$$f(n)=(1+o(1))\sqrt{n}. \tag{3}$$

Если обозначить теперь через $f_3(n)$ *наибольшее число целых положительных чисел* $a_1<a_2<a_3<\dots$, *не превосходящих* n *и таких, что различными являются все суммы*

$$a_i+a_j+a_l,$$

то естественно предположить, что аналогично (3)

$$f_3(n)=(1+o(1))\sqrt[3]{n}. \tag{4}$$

Однако доказательство этого соотношения представляется достаточно трудной задачей, поскольку метод, при помощи которого удалось доказать соотношение (3), здесь отказывается служить.

Сидон поставил также вопрос о том, что можно сказать о бесконечной последовательности целых чисел

$$a_1<a_2<a_3<\dots, \tag{5}$$

такой, что *все суммы* a_i+a_j *являются различными*. Туран и автор [1] доказали, что для такой последовательности

$$\overline{\lim}\,\frac{a_k}{k^2}=\infty\ ^{1}); \tag{6}$$

однако нам удалось построить последовательность (5), для которой

$$\underline{\lim}\,\frac{a_k}{k^2}<\infty.$$

Можно показать, что существует последовательность (5), для которой при всех k

$$a_k<ck^3. \tag{7}$$

Между результатами (6) и (7) имеется значительный пробел, который пока еще никак не заполнен.

Реньи и автор [4] с помощью теоретико-вероятностных соображений доказали, что для каждого $\varepsilon>0$ можно найти такое число

[1]) Напомним, что символы $\overline{\lim}\,u_k$ и $\underline{\lim}\,u_k$ означают соответственно наибольший и наименьший частичные пределы последовательности $\{u_k\}$. (*Прим. ред.*)

$l=l(\varepsilon)$ и такую последовательность $\{a_k\}$ целых положительных чисел, что

$$a_k<k^{2+\varepsilon}$$

и число $g(n)$ решений уравнения

$$n=a_i+a_j \tag{8}$$

для каждого n будет меньше l.

Другая проблема Сидона ставит вопрос о *существовании такой последовательности* $\{a_k\}$ *целых положительных чисел, для которой по каждому* $\varepsilon>0$ *можно найти такое число* n_0, *что для всех* $n>n_0$

$$0<g(n)<n^{\varepsilon}. \tag{9}$$

Идеи, заимствованные из теории вероятностей, позволили автору [5] доказать более сильное предложение о существовании такой последовательности, для которой

$$c_1\log n<g(n)<c_2\log n. \tag{10}$$

Вопрос о существовании такой последовательности $\{a_k\}$, для которой $\lim\frac{g(n)}{\log n}=c>0$, остается открытым. Старое предположение Турана и автора заключается в том, что если $g(n)>0$ для всех $n>n_0$, то $\overline{\lim}\,g(n)=\infty$ (возможно даже, что в случае произвольной последовательности $\{a_k\}$ $g(n)>c\log n$ для бесконечно многих n, откуда следовало бы, что оценка (10) является наилучшей). Доказательство справедливости этого предположения представляется достаточно трудным.

Туран и автор предположили также, что для бесконечной последовательности $\{a_k\}$ целых чисел не может иметь место оценка

$$\sum_{k=1}^{n}g(k)=cn+O(1). \tag{11}$$

Фукс и автор [6] доказали следующую более сильную теорему: *при* $c>0$ *не может иметь место оценка*

$$\sum_{k=1}^{n}g(k)=cn+o\left(\frac{n^{\frac{1}{4}}}{(\log n)^{1/2}}\right). \tag{12}$$

Для последовательности квадратов целых чисел

$$a_1=1,\quad a_2=4,\quad a_3=9,\ \dots$$

Харди и Ландау [7] доказали, что

$$\sum_{k=1}^{n}g(k)-\frac{\pi}{4}n\neq o\left((n\log n)^{\frac{1}{4}}\right). \tag{13}$$

Случай $a_k = k^2$ является важным, потому что в этом случае число $g(n)$ решений уравнения (8) тесно связано с числом точек с целочисленными координатами, заключающихся в заданном круге (если в первом квадрате провести дугу окружности радиуса $\sqrt{n}$, то $g(n)$ есть число точек этой дуги с целочисленными координатами); оценка величины $\sum_{k=1}^{n} g(n)$ сводится, таким образом, к классической задаче об определении числа целочисленных точек в круге радиуса $\sqrt{n}$. Предполагают, что в этом случае для некоторого $\varepsilon > 0$

$$\sum_{k=1}^{n} g(k) = \frac{\pi}{4} n + o(n^{\frac{1}{4}+\varepsilon}); \qquad (14)$$

эта оценка является весьма глубокой.

Относительно рассмотренных здесь вопросов см. также статью Штёра [8], содержащую много других задач того же рода. Вот одна из них, поставленная Рорбахом (Rohrbach): *каково наименьшее число $k(n)$ целых положительных чисел*

$$a_1 < a_2 < \ldots < a_{k(n)}$$

таких, что каждое целое положительное число, не превосходящее n, можно представить в виде суммы $a_i + a_j$. Ясно, что число $\frac{k(k-1)}{2}$ всевозможных сумм должно быть не меньше n, откуда следует, что

$$k(n) > \sqrt{2n},$$

Рорбах улучшил эту оценку:

$$k(n) > (1+\varepsilon)\sqrt{2n},$$

где $\varepsilon > 0$ — некоторое фиксированное число; недавно Мозер [9] улучшил оценку для величины ε. Легко видеть, что

$$k(n) \leqslant 2\sqrt{n};$$

возможно, что

$$k(n) = \sqrt{2n} + O(1).$$

Литература

1. P. Erdös, P. Turán, On the problem of Sidon in additive number theory and on some related problems, Journ. London Math. Soc. **16** (1941), стр. 212—215.
2. J. Singer, A theorem in finite projective geometry and some applications to number theory, Trans. Amer. Math. Soc. **43** (1938), стр. 347—385.
3. R. H. Bruck, H. J. Reiser, The nonexistence of certain finite projective planes, Canad. Journ. of Math. **1** (1949), стр. 88—93.

4. P. Erdös, A. Renyi, Additive properties of random sequences of positive integers, Acta Arithmetica **6** (1960), стр. 83—110.
5. P. Erdös, On a problem of Sidon in additive number theory, Acta Szeged (1954), стр. 255—259.
6. P. Erdös, W. H. J. Fuchs, On a problem of additive number theory, Journ. London Math. Soc. **31** (1956), стр. 67—73.
7. E. Landau, Vorlesungen über Zahlentheorie, т. II.
8. A. Stöhr, Gelöste und ungelöste Fragen über Basen der natärlichen Zahlenreihe, Journ. fur reine und angew. Math. **194** (1955), стр. 40—65 и 111—140.
9. Moser, Acta Arithmetica **6** (1960).

П. Эрдёш (Венгрия)

2. **Е. В. Бет, А. Тарский.** *Равносторонний треугольник — основное понятие геометрии в пространстве.*

В основе каждого дедуктивного построения геометрии лежат некоторые первичные неопределяемые понятия и неопределяемые же отношения между ними; эти понятия и отношения должны удовлетворять системе аксиом, которую можно, если угодно, принять за косвенное определение основных понятий и отношений (см., например, [5], [6]). При этом сами исходные неопределяемые понятия могут выбираться по-разному даже при построении одной и той же геометрической системы, например, обыкновенной геометрии Евклида. Так, в аксиоматике Гильберта [5] основными понятиями являются *точки, прямые* и *плоскости,* а основными отношениями между ними — отношения: *принадлежит* (отношение, связывающее точку и прямую или точку и плоскость), *между* (отношение, связывающее тройку точек, принадлежащих одной прямой) и *конгруентен* [отношение, связывающее два отрезка или два угла [1])]. Популярен также вариант аксиоматики Гильберта (идущий от Шура), в котором вместо отношения «конгруентен» введено новое неопределяемое понятие — *движение*; можно также вместо произвольных движений ограничиться одними *симметриями* плоскости относительно прямой, считая их неопределяемыми понятиями (это впервые было отмечено Виллерсом) и т. д.

После установления в конце прошлого века принципиальной возможности строго дедуктивного построения геометрии на аксиоматической основе, дальнейшая работа в области оснований геометрии сводилась главным образом к попыткам упрощения и усовершенствования существующих систем аксиом. Особое внимание, естественно, привлекал вопрос о возможностях упрощения аксиоматики за счет сокращения списка вводимых аксиом; представляет также интерес и вопрос о возможном уменьшении числа исходных понятий и отношений. Этим последним вопросом много занимался, в частности, известный американский логик А. Тарский, его ученики и сотрудники.

[1]) Понятия *отрезок* и *угол* могут быть независимо от понятия «конгруентен» определены через основные понятия и отношения с использованием относящихся к ним аксиом.

При построении евклидовой геометрии естественно принять в качестве единственного неопределяемого понятия *точку*; при этом прямые и плоскости определяются как некоторые геометрические места точек. Хочется одновременно потребовать, чтобы и число неопределяемых отношений между точками было возможно меньшим; лучше всего если удастся обойтись единственным подобным отношением.

Серьезные успехи в этом отношении были достигнуты уже давно. В работе А. Линденбаума и А. Тарского [3], опубликованной в 1936 г., было показано, что одного бинарного отношения между точками (т. е. отношения связывающего две точки) никак не достаточно для полного построения всей геометрии. Что же касается до тернарных отношений (отношений между тремя точками), то одно такое отношение было указано еще в 1908 г. одним из создателей современного аксиоматического метода в геометрии — итальянским математиком М. Пиери [4]: это — отношение $J(A, B, C)$, означающее, что треугольник ABC равнобедренный с вершиной A (может быть вырожденный), т. е. что $AB = AC$.

Приведенное здесь описание отношения $J(A, B, C)$ показывает, как можно определить это отношение при помощи основных понятий и отношений аксиоматики Гильберта. Но замечательно, что и, обратно, *все* основные понятия и отношения гильбертовской аксиоматики можно выразить через понятие *точка* и отношение $J(A, B, C)$; это и доказывает, что одного неопределяемого понятия *точка* и одного тернарного отношения $J(A, B, C)$ достаточно для построения всей евклидовой геометрии. В самом деле, *плоскость* можно определить как совокупность (геометрическое место) точек M, удовлетворяющих отношению $J(M, A, B)$, где A и B — две какие-то фиксированные точки, выбор которых определяет плоскость, а *прямую* — как совокупность точек M таких, что выполняются сразу два отношения $J(M, A, B)$ и $J(M, B, C)$; эти определения снимают также вопрос о каком-то специальном определении отношения *принадлежит*. Далее понятие *между*, которое должно иметь смысл лишь для трех точек A, B, C, принадлежащих одной прямой, можно определить, например, так: «точка B лежит *между* точками A и C, если существуют такие точки M и N, что $J(M, A, B)$ и $J(C, A, M)$, $J(N, B, C)$ и $J(A, C, N)$» (рис. 1). Наконец, два *отрезка* AB и CD мы назовем *конгруентными*, если существуют какие-то точки $M_1, M_2, \ldots, M_k$ такие, что выполняются $J(A, B, M_1)$, $J(M_1, A, M_2)$, $J(M_2, M_1, M_3)$, $\ldots$, $J(M_{k-1}, M_{k-2}, M_k)$, $J(M_k, C, M_{k-1})$, $J(C, D, M_k)$. Аналогично можно определить и *конгруентность углов* [исходя из того, что $\angle ABC$ конгруентен $\angle ACB$, если $J(A, B, C)$].

Некоторым недостатком отношения $J(A, B, C)$ является его несимметричный характер: из $J(A, B, C)$ не вытекает $J(B, A, C)$. Сравнительно недавно (в 1956 г.) Е. В. Бет и А. Тарский [1] заметили,

что в *случае геометрии в пространстве* [но не на плоскости!] *можно вместо отношения* $J(A, B, C)$ (несимметричного) *использовать отношение* $E(A, B, C)$ (симметричное), *означающее, что точки* A, B, C (различные или совпадающие) *являются вершинами равностороннего треугольника*, т. е. что $J(A, B, C)$ и $J(B, A, C)$.

Доказательство того, что в геометрии в пространстве отношения $J(A, B, C)$ и $E(A, B, C)$ в определенном смысле «равносильны», чрезвычайно просто. Выше мы уже определили $E(A, B, C)$ через $J(A, B, C)$; определим теперь $J(A, B, C)$ через $E(A, B, C)$, например, следующим образом: $J(A, B, C)$ означает существование таких двух точек M и N, что $E(A, B, M)$, $E(A, M, N)$ и $E(A, N, C)$ (см. рис. 2, где треугольники ABM, AMN, ANC равносторонние). Однако легко видеть, что на плоскости нельзя сопоставить каждому равнобедренному треугольнику ABC подобную цепь равносторонних треугольников ABM, AMN, ANC; таким образом, наше доказательство равносильности отношений $J(A, B, C)$ и $E(A, B, C)$ не сохраняет силу в случае плоскости.

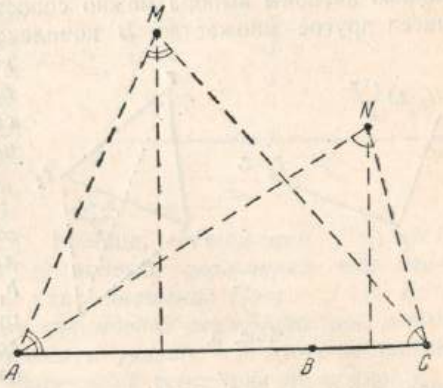


Рис. 1.

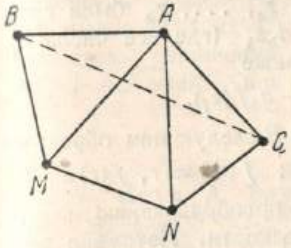


Рис. 2.

Из этого, однако, еще не следует, что никаким другим способом нельзя в геометрии на плоскости определить отношение $J(A, B, C)$ через отношение $E(A, B, C)$. Невозможность этого Е. В. Бет и А. Тарский доказывают довольно тонким построением. А именно, они показывают *существование такого точечного взаимно однозначного преобразования плоскости, которое не является движением* [и, следовательно, не сохраняет евклидову геометрию плоскости; в частности, три точки A, B, C такие, что $J(A, B, C)$, могут перейти при этом в три точки A', B', C', для которых отношение $J(A', B', C')$ не имеет места], *но которое переводит каждый равносторонний треугольник снова в равносторонний*, т. е. сохраняет отношение $E(A, B, C)$.

Описание этого преобразования $A' = f(A)$ в работе [1] существенно неэффективно и опирается на так называемую аксиому

выбора Е. Цермело [1]). А именно: отождествим совокупность точек плоскости с множеством комплексных чисел и выделим из этого множества подмножество A всех комплексных чисел вида $p+q\sqrt{3}$, где p и q — рациональные комплексные числа (т. е. числа вида r_1+ir_2, где r_1, r_2 — рациональные вещественные числа). С помощью аксиомы выбора можно сопоставить множеству A комплексных чисел другое множество B комплексных чисел такое, что *всякое комплексное число z может быть представлено (и притом единственным образом) в виде конечной суммы*

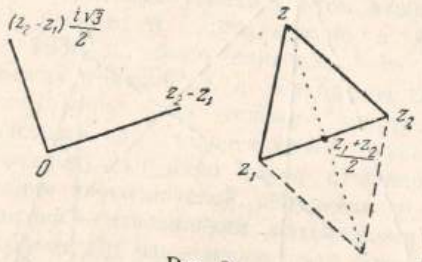


Рис. 3.

$$z=a_1b_1+a_2b_2+\ldots+a_kb_k,$$

где числа $a_1, a_2, \ldots, a_k$ принадлежат множеству A, а числа $b_1, b_2, \ldots, b_k$ — множеству B; при этом множество B заведомо содержит число 1, а также хоть одно отличное от 1 вещественное число r и какое-нибудь невещественное число m (ср. [7]). Пусть теперь преобразование f оставляет на месте все числа a и как-то переставляет числа b; исходя отсюда, уже можно распространить f на все множество комплексных чисел при помощи равенства

$$f(z)=a_1f(b_1)+a_2f(b_2)+\ldots+a_kf(b_k).$$

В таком случае, если какие-либо числа $z, z_1, z_2, \ldots, z_k$ были связаны соотношением $z=a_1z_1+a_2z_2+\ldots+a_kz_k$ (где все числа a_i принадлежат множеству A), то, очевидно, также

$$f(z)=a_1f(z_1)+a_2f(z_2)+\ldots+a_kf(z_i).$$

Определим теперь преобразование f множества B следующим образом: $f(b)=b$, *если число b отлично от r и от m;* $f(m)=r$, $f(r)=m$.

Тем самым мы приходим к определенному преобразованию всего множества комплексных чисел, т. е. всей плоскости. Нетрудно видеть, что это преобразование переводит любые три точки (три комплексных числа) z, z_1, z_2, являющиеся вершинами равностороннего треугольника, снова в вершины равностороннего треугольника — это следует из того, что, как легко проверить, наши три точки являются вершинами равностороннего треугольника тогда и только тогда, когда (рис. 3) $z=\frac{z_1+z_2}{2}\pm(z_2-z_1)\,i\,\frac{\sqrt{3}}{2}$, т. е. $z=z_1\alpha+z_2\beta i$ или $z=z_1\beta+z_2\alpha i$, где числа $\alpha=\frac{1}{2}+\frac{\sqrt{3}}{2}i$ и $\beta=\frac{1}{2}-\frac{\sqrt{3}}{2}i$ принадлежат

[1]) Об аксиоме Цермело см., например, П. С. Александров, Введение в общую теорию множеств и функций, М.— Л., 1948.

множеству A. Еще проще убедиться в том, что найдется равнобедренный треугольник, который наше преобразование переводит в неравнобедренный треугольник. Это вытекает, например, из существования такой точки z, что $J(z, 0, 1)$ и $J(z, 0, m)$, и отсутствия такой точки $f(z)$, для которой $J(f(z), f(0), f(1))$ и $J(f(z), f(0), f(m))$ или иначе $J(f(z), 0, 1)$ и $J(f(z), 0, r)$ (рис. 4).

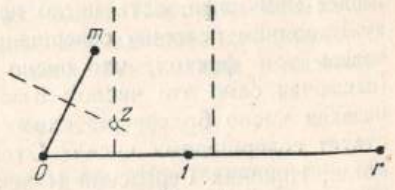


Рис. 4.

Это рассуждение и доказывает, что отношение $E(A, B, C)$ недостаточно для построения геометрии на плоскости; в частности, посредством него невозможно определить отношение $J(A, B, C)$. Вообще, *отношение $E(A, B, C)$ достаточно для обоснования евклидовой геометрии при числе измерений $n\geqslant 3$*, подобно тому как *отношение Пиери $J(A, B, C)$ может быть положено в основу евклидовой геометрии при любом числе измерений $n\geqslant 2$*. (А. Линдебаум указал, что этого отношения недостаточно для обоснования евклидовой геометрии на прямой.)

Укажем еще, что в том же журнале, где напечатана статья Е. В. Бета и А. Тарского, опубликована также тесно связанная с ней работа другого американского логика — Д. Скотта [2], в которой построено *симметричное тернарное отношение между точками, которое может быть положено в основу евклидовой геометрии при любом числе измерений $n\geqslant 2$* (т. е. пригодно и в случае плоскости) — это есть отношение $R(A, B, C)$, означающее, что (различные) точки A, B, C являются вершинами прямоугольного треугольника (т. е. что $AB^2+AC^2=BC^2$ или $AB^2+BC^2=AC^2$ или $AC^2+BC^2=AB^2$).

Литература

1. E. W. Beth, A. Tarski, Equilaterality as the only primitive notion of Euclidean geometry, Proc. Konikl. nederl. akad. Wetensch. **A. 59,** № 4, стр. 462—467; Indagationes math. **18,** № 4 (1956), стр. 462—467.
2. D. Scott, Symmetric primitive of Euclidean geometry, там же, стр. 456—461.
3. A. Lindebaum, A. Tarski, Über die Beschränkheit der Ausdrucksmittel deductive Theorien, Ergebnisse eines Mathematischen Kolloquims **7** (1936), стр. 15—22.
4. M. Pieri, La geometrie elementare institute sulla nozioni dil «puto» e «sfera», Memorie dil Mathematica e di Fisica delle Societa Italiana della Schieze, Ser. 3. **15** (1908), стр. 345—450.
5. Д. Гильберт, Основания геометрии, М.— Л., 1948.
6. П. К. Рашевский, Геометрия и ее аксиоматика, «Математическое просвещение», вып. 5 (1960), стр. 73—98.
7. Ю. М. Гайдук, Принцип полной математической индукции, его эквиваленты и обобщения, Математика в школе, 1950, № 2, стр. 1 — 20.

И. М. Яглом

3. *Применение электронных счетных машин для отыскания совершенных чисел.*

Появилось сообщение, что найдено восемнадцатое совершенное число. Для этого нужно было доказать, что число $2^{23217}-1$, содержащее 969 цифр, есть число простое.

Напомним историю совершенных чисел. Древние греки интересовались тем фактом, что число **6** равно сумме всех его делителей (исключая само это число): $6=1+2+3$. Из-за этого свойства они назвали число 6 «совершенным» и поставили вопрос — сколько существует совершенных чисел? Ответа на этот вопрос греки дать не могли; Никомах Геразский (I век нашей эры) писал:

> «Совершенные числа красивы. Но известно, что красивые вещи редки и немногочисленны, безобразные же встречаются в изобилии. Избыточными и недостаточными (несовершенными) являются почти все числа, в то время как совершенных чисел немного».

Легко было обнаружено проверкой, что вторым совершенным числом является **28:**

$$1+2+4+7+14=28.$$

Евклид доказал, *что всякое число, которое может быть представлено в виде произведения* $2^{n-1}(2^n-1)$, *где* (2^n-1) *есть простое число, является совершенным числом.* В случаях $n=2$ и $n=3$ числа $2^2-1=3$ и $2^3-1=7$ простые, и поэтому $2^1(2^2-1)=6$ и $2^2(2^3-1)=28$ — совершенные числа. Формула Евклида обнаружила еще два совершенных числа:

3-е: $n=5$, $2^5-1=31$, $2^4(2^5-1)=16\cdot 31=$ **496,**
4-е: $n=7$, $2^7-1=127$, $2^6(2^7-1)=64\cdot 127=$ **8128.**

Но отыскание дальнейших совершенных чисел этим путем оказалось делом весьма трудным. В течение многих столетий авторы, писавшие о совершенных числах, интересовались более суевериями и фантазиями, связанными с этими числами, чем их математической природой [1]). Если первые четыре совершенных числа были известны уже в глубокой древности, то пятое совершенное число

5-е: $n=13$, $2^{13}-1=8191$, $2^{12}(2^{13}-1)=$ **33 550 336**

было обнаружено лишь в XV веке — свыше чем через полторы тысячи лет после Эвклида.

В 1644 г. французский математик Марин Мерсенн объявил (не приводя доказательства), что первыми одиннадцатью совершенными числами вида $2^{n-1}(2^n-1)$ являются числа, отвечающие следующим значениям n:

2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257.

Математикам того времени было очевидно, что Мерсенн не мог проверить непосредственным вычислением простоты чисел 2^n-1 при всех указанных значениях n. Непосредственно удалось проверить только, что первые три из указанных Мерсенном шести новых совершенных чисел действительно являются совершенными. Вот эти числа:

6-е: $n=17$, $2^{17}-1=131\,071$, $2^{16}(2^{17}-1)=$ **8 589 869 056,**
7-е: $n=19$, $2^{19}-1=524\,287$, $2^{18}(2^{19}-1)=$ **137 438 691 328,**
8-е: $n=31$, $2^{31}-1=2\,147\,483\,647$, $2^{30}(2^{31}-1)=$
$=$ **2 305 843 008 139 952 128.**

6-е и 7-е числа указал еще Катальди в XVI веке; простоту чисел $2^{17}-1$, $2^{19}-1$, $2^{31}-1$ доказал Эйлер. Он же доказал, что *все четные совершенные числа имеют указанный Евклидом вид* $2^{n-1}(2^n-1)$, где (2^n-1) — простое (ясно, что для этого и n должно быть простым). *Ни одного нечетного совершенного числа обнаружить пока не удалось*; однако *не доказано и отсутствие таких чисел*, доказано лишь, что если нечетные совершенные числа и существуют, то они весьма велики.

Э. Ландау (1877—1938) писал в своей «Элементарной теории чисел»:

> «Древняя идея о совершенном числе и связанные с этим вопросы не имеют особенного значения, и я касаюсь этой материи лишь в связи с тем, чтобы указать на две проблемы, остающиеся нерешенными до сих пор:
> „Имеется ли бесконечное множество четных совершенных чисел?"
> — Не знаю.
> „Имеется ли бесконечное множество нечетных совершенных чисел?".
> — Я даже не знаю, существует ли одно такое число».

Более чем сто лет наибольшим известным совершенным числом оставалось число $2^{30}(2^{31}-1)$ с наибольшим известным простым числом 2 147 483 647. В 1876 г. французский математик Э. Люка указал метод, позволяющий проверить простоту числа без выполнения деления на всевозможные его простые делители. А именно, Люка доказал, что *число* 2^n-1 *будет простым тогда и только тогда,*

[1]) В диалогах Платона число 6 занимает особое место. У римлян на пирах самым почетным местом было шестое, на котором, как пишет Гораций, возлежал его благодетель Меценат. В Риме в 1917 г. при подземных работах была обнаружена постройка — общий зал с 28 кельями вокруг него. Оказалось, что это — помещение неопифагорейской академии, в которой полагалось 28 членов. Столько же членов полагалось в некоторых ученых обществах до последнего времени.

По религиозным преданиям мир был создан в 6 дней, потому что 6 — совершенное число. Английский богослов VIII века Алкуин учил, что человечество, происшедшее после потопа от 8 лиц, бывших в ковчеге Ноя, менее совершенно, чем до потопа, так как 8 — число несовершенное. Еще в XII веке церковники рекомендовали изучение совершенных чисел для спасения души.

если на него делится $(n-1)$-й *член рекуррентной последовательности чисел* (последовательность Люка):

4, 14, 194, 37 634, ..., где $u_1 = 4$, $u_{n+1} = u_n^2 - 2$ (см., например, [3]). Так, простота чисел $7 = 2^3 - 1$, $31 = 2^5 - 1$ вытекает из того, что на них делятся соответственно числа $u_2 = 4^2 - 2 = 14$ и $u_5 = 37\,634$. С помощью своей теоремы Люка установил, что число $2^{127} - 1$ является простым; для этого ему пришлось разделить 126-й член последовательности Люка на число

$$2^{127} - 1 = 170\,141\,183\,460\,469\,231\,731\,687\,303\,715\,884\,105\,727.$$

Этот последний результат был правильно предсказан Мерсенном, однако в других случаях он ошибся. Так, было установлено, что показатели $n = 67$ и $n = 257$, вопреки указанию Мерсенна, не дают совершенных чисел, но таковые дают не указанные Мерсенном показатели 61, 89 и 107.

Таким образом, список совершенных чисел в 1914 г. заключал в себе 12 чисел и оставался таким до 1952 г. Последние четыре числа этого списка следующие:

9-е: $n = 61$; $2^{60}(2^{61} - 1) = \mathbf{2^{60} \cdot 2\,305\,843\,009\,213\,693\,951}$
[37 цифр, И. М. Первушин [1]), 1833],

10-е: $n = 89$; $2^{88}(2^{89} - 1) = \mathbf{2^{88} \cdot 618\,970\,019\,642\,690\,137\,449\,562\,111}$
[54 цифры, Тарри, Поуэрс, 1911],

11-е: $n = 107$; $2^{106}(2^{107} - 1) =$
$= \mathbf{2^{106} \cdot 162\,259\,276\,829\,213\,363\,391\,578\,010\,288\,127}$
[65 цифр, Фокамбер, Поуэрс, 1914],

12-е: $n = 127$; $\mathbf{2^{126}(2^{127} - 1)}$ (77 цифр, Люка, 1876, Фокамберг, 1914).

Последнее 39-значное число $2^{127} - 1$ являлось с 1914 до 1952 г. наибольшим известным простым числом.

30 января 1952 г. американский математик Р. М. Робинзон использовал электронную счетную машину SWAC Калифорнийского университета для изучения простоты чисел вида $2^n - 1$ (см. [1]).

Первый опыт с машиной был сделан над определением простоты числа $(2^{257} - 1)$ — наибольшего из чисел вида $2^n - 1$, которое Мерсенн считал простым. В 1932 г. Д. Г. Лемер установил, что это число — составное, для чего ему при пользовании тогдашними счетными приборами пришлось работать целый год. Присутствуя в 1952 г. при проверке, он мог увидеть, как машина выполнила его годовую работу в 18 секунд, придя точно к тому же самому результату.

[1]) Иван Михеевич Первушин (1827—1900), священник из Перми, много и успешно занимавшийся теоретико-числовыми задачами, требующими колоссальных вычислений.

Машина продолжала дальнейшие поиски простых чисел. Мерсенн в свое время говорил, что вечности не хватит для проверки простоты числа, имеющего 15—20 десятичных знаков. Машина быстрее чем за два часа проверила 42 числа, наименьшее из которых имело свыше 80 цифр! Все эти числа оказались составными. К вечеру машина обнаружила новое совершенное число

13-е: $n = 521$, $\mathbf{2^{520}(2^{521} - 1)}$ (314 цифр),

а к полуночи, проверив еще 13 чисел, машина нашла следующее число

14-е: $n = 607$, $\mathbf{2^{606}(2^{607} - 1)}$ (366 цифр);

в десятичной системе простое число $2^{607} - 1$ имеет 183 цифры.

В свободное время от своих обычных более серьезных дел машина продолжала поиски больших простых чисел вида $2^n - 1$. В июне 1952 г. была доказана простота чисел $2^{1279} - 1$, а в октябре — простота чисел $2^{2203} - 1$ и $2^{2281} - 1$. Таким образом, было обнаружено еще три совершенных числа:

15-е: $n = 1279$, $\mathbf{2^{1278}(2^{1279} - 1)}$ (770 цифр),

16-е: $n = 2203$, $\mathbf{2^{2202}(2^{2203} - 1)}$ (1327 цифр),

17-е: $n = 2281$, $\mathbf{2^{2280}(2^{2281} - 1)}$ (1373 цифры).

Наконец, в сентябре 1957 г. шведский математик Г. Ризель (H. Riesel) при помощи электронной машины BESK (Стокгольм) установил простоту числа $2^{3217} - 1$ (см. [2]), состоящего из 969 цифр, и получил новое совершенное число,

18-е: $n = 3217$, $\mathbf{2^{3216}(2^{3217} - 1)}$ (около 2000 цифр).

Оно является пока наибольшим известным совершенным числом, число $2^{3217} - 1$ наибольшим известным простым числом.

Литература

1. H. S. Uhler, A Brief History of the Investigations on Mersenne Numbers and the Latest Immense Primes, Scripta Math. **18**, № 2 (1952), стр. 122—131.
2. J. Anderson, Le plus grand nombre premier, Mathesis. **66**, № 7—9 (1957), стр. 316.
3. Э. Трост, Простые числа, М., 1959.
4. И. В. Арнольд, Теория чисел, М., 1959.

И. Я. Депман

ОКОНЧАНИЕ ТЕКСТА СТР. 310

из своих карт и пытаются приложить ее к концу начатой последовательности. Если выбранная карта удовлетворяет правилам, «водящий» говорит «правильно», и оставляет карту; в противном случае «водящий» говорит «неправильно», и игрок откладывает карту в сторону.

После того как перебраны все 52 карты, начинается вторая половина игры. Игроки получают отложенные ими карты, и игра этими картами продолжается по прежнему правилу с единственным отличием, что неудачно выложенная карта теперь не откладывается в сторону, а возвращается игроку для дальнейших попыток. Тур игры заканчивается, когда у одного из игроков не останется карт или когда «водящий» объявит, что правило игры не позволяет далее ни одному из игроков выложить ни одной карты. Выигравшему игроку (тому, у которого в конце тура осталось меньше карт) засчитывается число очков, равное разности между числом карт противника и числом его собственных карт; если выигравший выложил все карты, ему добавляется премия в шесть очков.

Ясно, что игрок, первым разгадавшим правило, получает большое преимущество перед своим противником. В случае наличия сомнительной гипотезы игрок может делать сознательный эксперимент, выбирая карту так, чтобы ответ «водящего» подтвердил или опроверг его гипотезу. Существенным может оказаться даже частичное знание правила. Так, если задано упомянутое выше правило 4°, то игрок, первым подметивший чередование мастей, будет ошибаться, но в среднем вдвое реже, чем его противник. «Водящему» засчитывается в качестве его выигрыша абсолютное значение разности между числом отложенных карт двух игроков при окончании первой половины игры. Таким образом, «водящий» заинтересован в том, чтобы правило было быстро разгадано ровно одним игроком, т. е. он должен выбрать правило не слишком простым и не слишком сложным по отношению к умению игроков.

Читатель, попытавшийся играть в Элеузис, будет скорее всего поражен стандартностью своего мышления, не позволяющей ему выдвинуть никаких правдоподобных гипотез о последовательности карт, удовлетворяющей на самом деле очень простому правилу. Такова, например, следующая последовательность: 9 пик, 10 червей, 12 треф, 10 бубен, 5 треф, 12 червей, 7 треф, 8 пик, 10 треф, 1 бубен, 11 червей, 3 пик, 6 пик, 11 треф, 2 червей, 7 бубен, 5 пик, 13 червей, 9 червей, 4 пик, 3 треф, 7 пик, 1 пик, 7 червей, 13 пик. (Ответ будет дан в следующем выпуске.) Поэтому для начинающих игроков надо выбирать очень простые правила.

Р. Л. Добрушин

V. ЗАДАЧИ

Под редакцией И. М. Яглома

ЗАДАЧИ [1])

1. Задачи по элементарной математике

А. Задачи средней трудности

48. Легко видеть, что числа

$$2^{(2^1)}-2^{(2^0)}+1=3,\quad 2^{(2^2)}-2^{(2^1)}-1=11 \text{ и } 2^{(2^2)}-2^{(2^1)}+1=13$$

являются простыми; также и число $2^{(2^1)}-2^{(2^0)}-1=1$ не может быть разложено на множители. А не являются ли числа

$$2^{(2^n)}-2^{(2^{n-1})}\pm 1$$

простыми и при всех $n>2$?

В. И. Левин (Москва)

49. Доказать, что при любом n число $\underbrace{111\ldots11}_{2n \text{ единиц}}-\underbrace{22\ldots2}_{n \text{ двоек}}$ является полным квадратом.

В. В. Никитин (Тамбов)

50. Имеется два набора чисел $a_1, a_2, \ldots, a_n$ и $b_1, b_2, \ldots, b_n$; пусть A и B — наибольшая и наименьшая из сумм $a_1+b_1, a_2+b_2, \ldots, a_n+b_n$. Переставим теперь первые n чисел в порядке $a_{i_1}\geqslant a_{i_2}\geqslant \ldots \geqslant a_{i_n}$ и вторые n чисел в порядке $b_{j_1}\leqslant b_{j_2}\leqslant \ldots \leqslant b_{j_n}$, и пусть $\tilde{A}$ и $\tilde{B}$ — наибольшая и наименьшая из сумм $a_{i_1}+b_{j_1}, a_{i_2}+b_{j_2}, \ldots, a_{i_n}+b_{j_n}$. Докажите, что

$$A\geqslant \tilde{A}\geqslant \tilde{B}\geqslant B.$$

51. Какое наибольшее число подмножеств можно выделить из множества из n элементов так, чтобы ни одно из выделенных подмножеств не входило целиком в другое?

[1]) Все неподписанные задачи в этом выпуске «Математического просвещения» заимствованы из Студенческого научного журнала, выпущенного в 1958 г. НСО мех.-мат. ф-та Ленинградского университета.

52. Доказать, что если $x_0^n + a_1 x_0^{n-1} + \ldots + a_{n-1} x_0 + a_n = 0$, то

$$|x_0|^2 + \frac{|a_n|}{|x_0|^2} \leqslant 1 + |a_1|^2 + |a_2|^2 + \ldots + |a_n|^2.$$

53. Доказать, что если (нанесенная на глобусе) карта такова, что можно, двигаясь по границам карты, обойти все вершины (точки, в которых сходятся три или больше стран), исходя из какой-либо одной, и вернуться в исходную вершину, не побывав ни в одной промежуточной вершине дважды, то все страны на карте можно закрасить четырьмя различными красками так, чтобы никакие две страны, имеющие общую границу, не были закрашены одинаково.

54. Известно, что все отрезки, соединяющие вершины выпуклого пятиугольника с серединами противоположных сторон, пересекаются в одной точке; что три из них перпендикулярны противоположным сторонам; что пятиугольник имеет ось симметрии. Следует ли отсюда, что этот пятиугольник правильный?

В. В. Никитин (Томск)

55. Ясно, что момент инерции материальной окружности относительно оси, лежащей в плоскости окружности и проходящей через ее центр, не зависит от направления оси. Покажите, что тем же свойством обладает система материальных точек равной массы, расположенных в вершинах (вписанного в окружность) правильного многоугольника.

Л. Н. Бескин (Москва)

Б. Задачи повышенной трудности

30. Расстояние между двумя [выпуклыми [1])] фигурами Φ и Ψ можно определить разными способами [2]):

а) *Линейным расстоянием* от Φ до Ψ называется расстояние до Ψ от самой далекой от Ψ точки Φ; расстоянием же от точки A до Ψ называется расстояние от A до самой близкой к A точки Ψ (рис. *а*). Иногда под линейным расстоянием между Φ и Ψ понимается также наибольшее из линейных расстояний от Φ до Ψ и от Ψ до Φ (эти расстояния не обязательно равны).

б) *Расстоянием по периметру* между Φ и Ψ называется разность периметров объединения и пересечения фигур Φ и Ψ (рис. *б*).

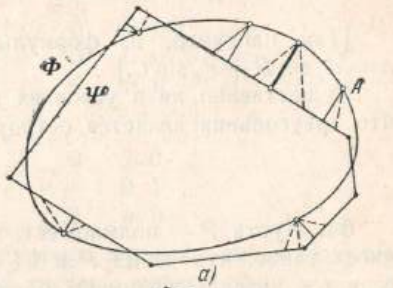
[1]) Фигура называется *выпуклой*, если через каждую граничную точку фигуры можно провести не пересекающую фигуру прямую.

[2]) См., например, Л. Фейеш Тот, Расположения на плоскости, на сфере и в пространстве, М., 1958, стр. 58, 61.

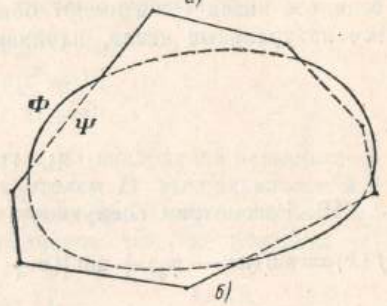
в) *Расстоянием по площади между* Φ и Ψ называется разность площадей объединения и пересечения Φ и Ψ (рис. *в*).

Из всех n-угольников найти тот, который будет ближе всего к заданному единичному кругу K (причем расстояние между многоугольником M и кругом K может пониматься в каждом из перечисленных выше трех смыслов).

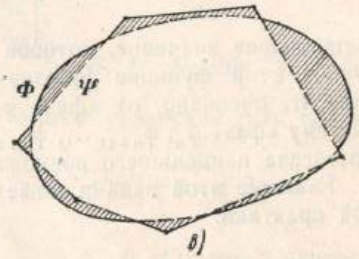
Е. И. Файнберг (Москва)



а)



б)



в)

31. В три сосуда, вместимостью m литров каждый, налито m литров жидкости так, что в каждом сосуде имеется целое число литров. Доказать, что последовательным переливанием жидкости, при котором каждый раз в один из сосудов доливается из другого сосуда такое же количество жидкости, какое в первом сосуде было раньше, можно опорожнить один сосуд.

А. И. Ширшов (Москва)

32. На прямой l даны пять точек A, B, A_1, B_1, O такие, что $OA \cdot OB = OA_1 \cdot OB_1$. Через данную точку M провести прямую, параллельную l, с помощью одной линейки.

З. А. Скопец (Ярославль)

33. Пусть A, B, C — углы произвольного треугольника T; a, b, c — противолежащие им стороны; h_a, h_b, h_c — соответствующие высоты; d_a, d_b, d_c — расстояния от ортоцентра (точки пересечения высот) треугольника до его вершин; D — диаметр описанной окружности. Доказать, что:

а) если для каждого остроугольного треугольника T имеет место равенство $f(a, b, c, D) = 0$, то и $f(a, d_c, d_b, D) = f(d_c, b, d_a, D) = f(d_b, d_a, c, D) = 0$.

[Так, например, из соотношения $D^2 = \dfrac{4a^2b^2c^2}{(a^2+b^2+c^2)^2 - 2(a^4+b^4+c^4)}$ следует, что $D^2 = \dfrac{4a^2d_c^2d_b^2}{(a^2+d_c^2+d_b^2)^2 - 2(a^4+d_c^4+d_b^4)} = \ldots$].

б) если для каждого остроугольного треугольника T имеет место равенство $f(a, b, c, h_a, D, A, B, C)=0$, то и

$$f(a, d_c, d_l, h_a-d_a, D, \pi-A, \tfrac{\pi}{2}-C, \tfrac{\pi}{2}-B)=0.$$

[Так, например, из формулы $a=b\cos C+c\cos B$ следует, что $a=d_c\sin B+d_b\sin C$.]

Существенно ли в условиях задач а) и б) предположение о том, что треугольник является остроугольным?

А. М. Рубинов (Ленинград)

34. Пусть P — подмножество множества целых неотрицательных чисел такое, что из $a\in P$ и $b\in P$ следует $a+b\in P$. Доказать, что если все числа P не имеют общего делителя $d>1$, то P содержит все натуральные числа, начиная с некоторого.

Ю. Л. Шмульян (Житомир)

В. Проблемы

13. Рассмотрим следующую функцию x:

$$f(x)=\sin(nx-\varphi_n)+\sin[(n+1)x-\varphi_{n+1}]+\ldots \\ \ldots+\sin[(n+k)x-\varphi_{n+k}].$$

Наибольшее значение, которое может принимать абсолютная величина $|f(x)|$ этой функции [обозначим его через $F(\varphi_n, \varphi_{n+1}, \ldots, \varphi_{n+k})$] зависит, очевидно, от «фаз» $\varphi_n, \varphi_{n+1}, \ldots, \varphi_{n+k}$. Как выбрать такую систему «фаз» $\varphi_n, \varphi_{n+1}, \ldots, \varphi_{n+k}$, чтобы величина $F(\varphi_n, \varphi_{n+1}, \ldots, \varphi_{n+k})$ достигала наименьшего возможного значения?

Решение этой задачи может представлять интерес для инженерной практики.

А. И. Цукублин (Москва)

14. Нетрудно составить набор сопротивлений такой, чтобы, соединяя эти сопротивления последовательно и параллельно различным образом, можно было получить (с точностью до 1 *ом*) любое сопротивление до *M ом* включительно. Требуется оценить наименьшее число сопротивлений, которое может включать такой набор.

15. Известно, что все хорды, делящие пополам площадь данной плоской фигуры Ф (ограниченной простой замкнутой кривой), имеют длину, не превосходящую 1. Оценить сверху площадь Ф; указать фигуры наибольшей площади, удовлетворяющие условию задачи.

2. Задачи по высшей математике

А. Задачи средней трудности

36. Привести к нормальной форме матрицу $(n-1)$-го порядка

$$\begin{pmatrix} 0 & n & 0 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 0 & n-1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 2 & 0 & n-2 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & \ldots & n-1 & 0 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & n & 0 \end{pmatrix}.$$

37. Найти необходимые и достаточные условия существования нетривиальных вещественных решений системы уравнений

$$a_1x^2-2b_1xy+c_1y^2=0,$$
$$a_2y^2-2b_2yz+c_2z^2=0,$$
$$a_3z^2-2b_3zx+c_3x^2=0.$$

38. Пусть A_{n-1} — матрица, полученная из квадратной невырожденной матрицы A_n порядка n с определителем D вычеркиванием k-й строки и l-го столбца; D_{ij} — алгебраическое дополнение элемента a_{ij} матрицы A_n, d_{ij} — алгебраическое дополнение того же элемента матрицы A_{n-1}. Доказать, что

$$d_{ij}=\frac{1}{D}\begin{vmatrix} D_{ij} & D_{it} \\ D_{kj} & D_{kt} \end{vmatrix}.$$

Л. Н. Бескин (Москва)

39. При каком условии сумма двух корней уравнения (над полем комплексных чисел)

$$x^6+ax^5+bx^4-(a^2-2a-2b+2)x^3+bx^2+ax+1=0$$

будет равна 1?

Г. Б. Гуревич (Москва)

40. Пусть O — точка внутри (гипер)сферы Σ n-мерного евклидова пространства; $A_1, A_2, \ldots, A_n$ — такие точки (гипер)сферы, что прямые $OA_1, OA_2, \ldots, OA_n$ попарно перпендикулярны. Найти геометрическое место центроидов всевозможных симплексов $OA_1A_2\ldots A_n$ (точка O фиксирована).

З. А. Скопец (Ярославль)

41. На двух перпендикулярных прямых a и b взяты точки A и B; через точку A проведены прямые a_1 и a_2, образующие с a угол α,

а через B — прямые b_1 и b_2, образующие угол α с прямой b. Доказать, что

а) Четыре (отличные от A и B) точки пересечения прямых a_1, a_2 и b_1, b_2 принадлежат одной окружности.

б) Семейство указанных в задаче а) окружностей, отвечающих всевозможным значениям α, принадлежит одному пучку.

в) Фигурирующие в задаче а) четыре точки при изменении α описывают кривую четвертого порядка, распадающуюся на окружность и равностороннюю гиперболу.

У. Асекритов, О. Котий (Ярославль)

42. а) Доказать, что каждый гиперболический параболоид имеет две и только две прямолинейные образующие, такие, что нормали к поверхности вдоль образующей сами образуют линейчатую поверхность второго порядка (являющуюся также гиперболическим параболоидом).

б) Результат задачи а) позволяет связать с каждым гиперболическим параболоидом два новых гиперболических параболоида; с ними в свою очередь можно таким же способом связать еще четыре гиперболических параболоида и т. д. При каких условиях хотя бы один из полученных на этом пути гиперболических параболоидов совпадает с исходным?

Ш. С. Симонян (Ереван)

Б. Задачи повышенной трудности

29. Пусть функции f и g определены на отрезке $[a, b]$. Мы будем говорить, что функция g возрастает медленнее функции f, если для любых x_1, x_2 таких, что $a \leqslant x_1 \leqslant x_2 \leqslant b$, имеет место неравенство $g(x_2) - g(x_1) \leqslant f(x_2) - f(x_1)$.

Для всяких ли двух непрерывных и строго возрастающих на $[a, b]$ функций существует отличная от константы неубывающая функция g, определенная на $[a, b]$ и возрастающая медленнее, чем f_1 и чем f_2?

30. Пусть $\mathfrak{f}$ — множество таких функций $f(x)$, что

1°. $f(x) \geqslant 0$.

2°. $\int\limits_{-\infty}^{+\infty} f(x)\,dx = 1$.

3°. Множество максимумов $f(x)$ есть точка или отрезок.

Для каких из функций $\varphi(x) \in \mathfrak{f}$ «свертка» $\tilde{f}(x) = \int\limits_{-\infty}^{+\infty} f(x-t)\,\varphi(t)\,dt$ функции f с функцией φ будет принадлежать множеству $\mathfrak{f}$ одновременно с функцией f?

31. Пусть функция $f(x)$ определена на всей вещественной оси, трижды дифференцируема и ее вторая и третья производные неотрицательны. Обозначим через $X(a, b)$ среднее арифметическое всех корней уравнения $f(x) = ax + b$ и определим функцию $B(x, y)$ условием

$$X(a, B(a, x)) = x,$$

а функцию $A(x)$ — условием

$$f(x) = A(x)\,x + B(A(x), x).$$

Доказать, что $A(x) = f'(x)$.

32. а) На отрезке $[a, b]$ задана несчетная система непрерывных функций. Доказать, что из этой системы можно выбрать равномерно сходящуюся последовательность.

б) Доказать, что утверждение задачи а) теряет силу, если от функций требовать не непрерывности, а лишь чтобы каждая из них являлась пределом сходящейся последовательности непрерывных функций (т. е. принадлежала к 1-му классу Бэра).

в) Построить несчетную систему функций последовательности, сходящейся в каждой точке, из которой нельзя выбрать никакой.

Е. М. Ландис (Москва)

В. Проблемы

12. На плоскости задана (вообще говоря, не центрально симметричная) выпуклая [1]) кривая Λ и точка O внутри нее. Кривая Λ является «единичной окружностью метрики Минковского»; другими словами, длина отрезка AB плоскости принимается равной отношению AB к «единичному» отрезку OM того же направления, где M — точка кривой Λ (отрезки AB и OM параллельны и одинаково направлены; так как кривая Λ не центрально симметрична, то длины отрезков AB и BA могут быть различными). Длину ориентированной кривой Λ (т. е. кривой Λ, на которой задано определенное «положительное» направление обхода) можно определить обычным образом как предел длин вписанных в Λ многоугольников, длины всех сторон которых неограниченно убывают (или предел длин описанных вокруг Λ многоугольников); эту длину мы обозначим через $2\pi_\Delta$. Спрашивается, для каких кривых Λ величина π_Δ достигает наибольшего и наименьшего значения?

[Ограничиваясь случаем центрально-симметричных кривых Λ, Ю. Г. Решетняк доказал, что $3 \leqslant \pi_\Delta \leqslant 4$; $\pi_\Delta = 3$, если Λ — правильный шестиугольник; $\pi_\Delta = 4$, если Λ — квадрат [2]).]

[1]) См. сноску [1]) на стр. 330.

[2]) См. Ю. Г. Решетняк, Одна экстремальная задача из теории выпуклых кривых, Успехи математических наук **8**, № 6, 1953, стр. 125—126.

Исправления:

Формулировка задачи **24** по высшей математике повышенной трудности («Математическое просвещение», вып. 5, стр. 258) должна начинаться так:

«Пусть Φ — произвольное непрерывное взаимно однозначное преобразование окружности единичной длины: $\alpha' = \Phi(\alpha) = \alpha + f(\alpha)$, где углы (дуги) α и α' определяют соответственно положения исходной и преобразованной точек на окружности, а $f(\alpha)$ — непрерывная, периодическая (с периодом 1) функция. Пусть далее α — угол, соответствующий произвольной точке M, $\alpha_1 = \Phi(\alpha)$, $\alpha_2 = \Phi(\alpha_1)$, $\alpha_3 = \Phi(\alpha_2)$,... — углы, соответствующие точкам $M_1 = \Phi(M)$, $M_2 = \Phi(M_1)$, $M_3 = \Phi(M_2)$ и т. д. Доказать, что...»

В статье П. К. Рашевского «Геометрия и ее аксиоматика» («Математическое просвещение», вып. 5, страница 73, строка 7 сверху) допущена искажающая смысл опечатка.

Напечатано: прозрачны

Должно быть: призрачны

РЕШЕНИЯ ЗАДАЧ

1. Задачи по элементарной математике

А. Задачи средней трудности

3. Доказать, что если луч отражается по одному разу от каждого из трех попарно перпендикулярных зеркал, то он меняет свое направление на обратное. Обратно, если произвольный луч после трехкратного отражения от трех плоских зеркал меняет свое направление на обратное, то плоскости этих зеркал попарно перпендикулярны. [Примечание. При отражении от плоского зеркала падающий и отраженный лучи лежат в одной плоскости с перпендикуляром к плоскости зеркала, восставленным в точке падения, и образуют с этим перпендикуляром равные углы.]

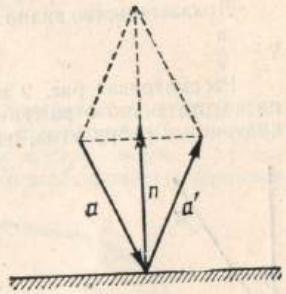


Рис. 1.

Первое решение. 1°. Если луч с направляющим вектором $\boldsymbol{a}$ падает на зеркало с нормальным вектором $\boldsymbol{n}$, то отраженный луч имеет направляющий вектор $\boldsymbol{a}'$, определяемый равенством (рис. 1; векторы $\boldsymbol{a}$, $\boldsymbol{n}$ и $\boldsymbol{a}'$ — единичные)

$$\boldsymbol{a}' = \boldsymbol{a} - 2(\boldsymbol{an})\boldsymbol{n}. \tag{1}$$

Если луч $\boldsymbol{a}$ последовательно отражается от зеркал с нормальными векторами $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$, то направляющие векторы отраженных лучей определяются по формуле (1):

$$\begin{aligned}
\boldsymbol{a}_1 &= \boldsymbol{a} - 2(\boldsymbol{an}_1)\boldsymbol{n}_1,\\
\boldsymbol{a}_2 &= \boldsymbol{a}_1 - 2(\boldsymbol{a}_1\boldsymbol{n}_2)\boldsymbol{n}_2 = \boldsymbol{a} - 2(\boldsymbol{an}_1)\boldsymbol{n}_1 - 2(\boldsymbol{an}_2)\boldsymbol{n}_2 + 4(\boldsymbol{an}_1)(\boldsymbol{n}_1\boldsymbol{n}_2)\boldsymbol{n}_2,\\
\boldsymbol{a}_3 &= \boldsymbol{a}_2 - 2(\boldsymbol{a}_2\boldsymbol{n}_3)\boldsymbol{n}_3 = \boldsymbol{a} - 2(\boldsymbol{an}_1)\boldsymbol{n}_1 - 2(\boldsymbol{an}_2)\boldsymbol{n}_2 - 2(\boldsymbol{an}_3)\boldsymbol{n}_3 +\\
&\quad + 4(\boldsymbol{an}_1)(\boldsymbol{n}_1\boldsymbol{n}_2)\boldsymbol{n}_2 + 4(\boldsymbol{an}_1)(\boldsymbol{n}_1\boldsymbol{n}_3)\boldsymbol{n}_3 + 4(\boldsymbol{an}_3)(\boldsymbol{n}_2\boldsymbol{n}_3)\boldsymbol{n}_3 -\\
&\quad - 8(\boldsymbol{an}_1)(\boldsymbol{n}_1\boldsymbol{n}_2)(\boldsymbol{n}_2\boldsymbol{n}_3)\boldsymbol{n}_3.
\end{aligned} \tag{2}$$

Если векторы $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$ попарно перпендикулярны, то (2) переходит в

$$\boldsymbol{a}_3 = \boldsymbol{a} - 2(\boldsymbol{an}_1)\boldsymbol{n}_1 - 2(\boldsymbol{an}_2)\boldsymbol{n}_2 - 2(\boldsymbol{an}_3)\boldsymbol{n}_3. \tag{3}$$

Положив $\boldsymbol{a} = x\boldsymbol{n}_1 + y\boldsymbol{n}_2 + z\boldsymbol{n}_3$, будем иметь:

$$x = \boldsymbol{an}_1,\ y = \boldsymbol{an}_2,\ z = \boldsymbol{an}_3$$

и

$$\boldsymbol{a}_3 = \boldsymbol{a} - 2x\boldsymbol{n}_1 - 2y\boldsymbol{n}_2 - 2z\boldsymbol{n}_3 = \boldsymbol{a} - 2\boldsymbol{a} = -\boldsymbol{a}. \tag{4}$$

2°. Обратно: пусть при любом векторе $\boldsymbol{a}$ имеет место $\boldsymbol{a}_3 = -\boldsymbol{a}$. Заметим, что векторы $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$ не компланарны (иначе условие $\boldsymbol{a}_3 = -\boldsymbol{a}$ нарушалось бы уже для лучей $\boldsymbol{a}$, компланарных с $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$). Пусть $\boldsymbol{a} = x\boldsymbol{n}_1 + y\boldsymbol{n}_2 + z\boldsymbol{n}_3$. Тогда, подставив в тождество $\boldsymbol{a} + \boldsymbol{a}_3 = 0$ выражение (2) и приравняв нулю коэффициенты при $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$, получим

$$2x - 2(\boldsymbol{a}\boldsymbol{n}_1) = 0, \quad 2y - 2(\boldsymbol{a}\boldsymbol{n}_2) + 4(\boldsymbol{a}\boldsymbol{n}_1)(\boldsymbol{n}_1\boldsymbol{n}_2) = 0,$$
$$2z - 2(\boldsymbol{a}\boldsymbol{n}_3) + 4(\boldsymbol{a}\boldsymbol{n}_1)(\boldsymbol{n}_1\boldsymbol{n}_3) + 4(\boldsymbol{a}\boldsymbol{n}_2)(\boldsymbol{n}_2\boldsymbol{n}_3) - 8(\boldsymbol{a}\boldsymbol{n}_1)(\boldsymbol{n}_1\boldsymbol{n}_2)(\boldsymbol{n}_2\boldsymbol{n}_3) = 0. \quad (6)$$

Подставив сюда снова $\boldsymbol{a} = x\boldsymbol{n}_1 + y\boldsymbol{n}_2 + z\boldsymbol{n}_3$ и учитывая, что $\boldsymbol{n}_1^2 = \boldsymbol{n}_2^2 = \boldsymbol{n}_3^2 = 1$, получим из равенств (5)

$$\boldsymbol{n}_1\boldsymbol{n}_2 = \boldsymbol{n}_2\boldsymbol{n}_3 = \boldsymbol{n}_1\boldsymbol{n}_3 = 0 \quad (7)$$

[равенство (6) в силу (7) превращается в тождество]. Итак, векторы $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, $\boldsymbol{n}_3$ попарно перпендикулярны.

Л. Н. Бескин (Москва)

Второе решение. Докажем сначала три леммы.

Лемма 1. *Если прямая a двукратно отражается от сторон угла φ, то отраженная прямая a_2 будет параллельна a в том и только в том случае, когда угол φ прямой.*

Доказательство видно из рис. 2: $\alpha + \delta = (\pi - 2\beta) + (\pi - 2\gamma) = \pi$, $\beta + \gamma = \frac{\pi}{2}$, $\varphi = \frac{\pi}{2}$.

Рассматривая рис. 2 как ортогональную проекцию двугранного угла и последовательно отраженных плоскостей a, a_1, a_2 в направлении, перпендикулярном к ребру угла, получаем аналогичное утверждение относительно двукратного отражения плоскости от сторон двугранного угла.

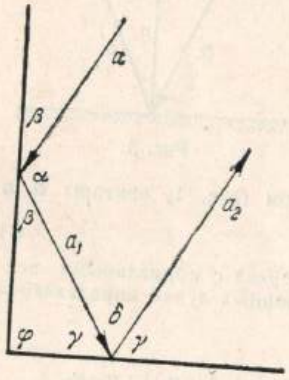


Рис. 2.

Лемма 2. Пусть прямые a и a_1, лежащие в плоскости α, ортогонально проектируются на плоскость α' в прямые a' и a_1' (рис. 3). Проведем через точки $M = a \times a_1$ и $M' = a' \times a_1'$ плоскость μ, перпендикулярную к α и α', и обозначим прямые $\alpha \times \mu$ и $\alpha' \times \mu$ соответственно через m и m'. Тогда a' и a_1' *будут взаимно отраженными* (т. е. симметричными) *прямыми относительно прямой m' (и плоскости μ) в том и только в том случае, когда a и a_1 — взаимно отраженные прямые относительно прямой m (и плоскости μ).*

Доказательство. Если A и A_1 — точки, лежащие соответственно на прямых a и a_1 и одинаково удаленные от плоскости μ, — проектируются на плоскость α' в точки A' и A_1', то равнобедренный треугольник AMA_1 проектируется в равнобедренный же треугольник $A'M'A_1'$. Действительно, если $AN = A_1N$, то $A'N' = A_1'N_1'$ ($MN \perp \mu$, $M'N' \perp \mu$). Если же $AN \neq A_1N$, то и $A'N' \neq A_1'N'$.

Лемма 3. *Пусть прямые a и a_1 и плоскости α и α_1 взаимно отражены от плоскости μ. Тогда из того, что a принадлежит α, следует, что и a_1 принадлежит α_1.*

Доказательство следует из определения взаимно отраженных (т. е. симметричных) прямых и плоскостей.

Переходим к доказательству прямой и обратной теорем, сформулированных в условии задачи.

Доказательство прямой теоремы. Пусть α, β, γ — попарно перпендикулярные плоскости, $OA = \beta \times \gamma$, $OB = \gamma \times \alpha$, $OC = \alpha \times \beta$, ломаная $aa_1a_2a_3$ — путь луча, последовательно отраженного от этих плоскостей (рис. 4). Спроектируем эту ломаную на плоскость α. Тогда проекции прямых

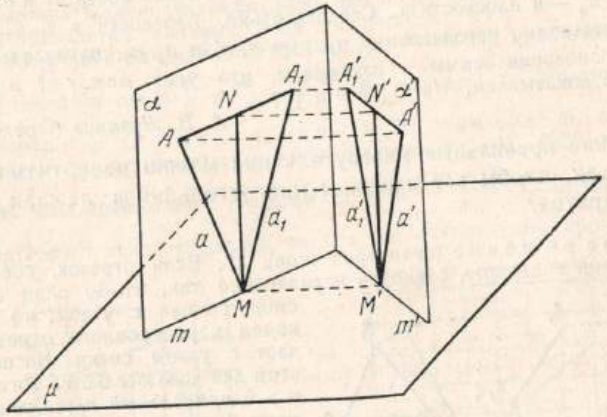


Рис. 3.

a и a_1 будут продолжением одна другой, проекции прямых a_1 и a_2 будут (по лемме 2) взаимно отраженными от OC и проекции прямых a_2 и a_3 (по той же лемме) — взаимно отраженными от OB. Поэтому в силу леммы 1 проекция a будет параллельна проекции a_3. Заменив в этом рассуждении плоскость α

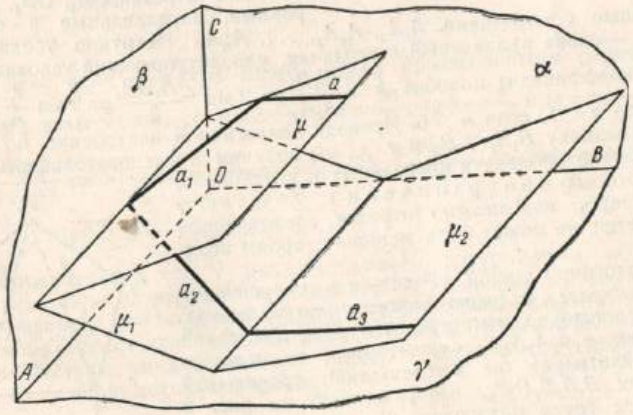


Рис. 4.

плоскостью γ, получим, что проекции прямых a и a_3 на плоскость γ также будут параллельны. Поэтому $a_3 \parallel a$.

Доказательство обратной теоремы. Пусть при последовательном отражении от трех плоскостей α, β, γ падающий и отраженный лучи a и a_3 всегда параллельны (рис. 4); требуется доказать, что α, β и γ попарно перпендикулярны. Проведем через a и a_1 плоскость μ и отразим ее после-

довательно от плоскостей β и γ, получим плоскости μ_1 и μ_2. В силу леммы 3 μ_1 содержит a_2, μ_2 содержит a_3; заменив луч a иным лучом a', также лежащим в плоскости μ, мы получим отраженный от α луч a_1', лежащий в той же плоскости μ, затем отраженный от β луч a_2' — в плоскости μ_1 и отраженный от γ луч a_3' — в плоскости μ_2. Следовательно, плоскости μ и μ_2, содержащие соответственно параллельные прямые a, a' и a_3, a_3', параллельны между собой. На основании леммы 1 заключаем, что угол между β и γ прямой. Аналогично доказываем, что $\alpha \perp \beta$ и $\alpha \perp \gamma$.

А. В. Лузанов (Орехово-Зуево)

11. Какие правильные многоугольники можно начертить на бумаге в клетку так, чтобы все вершины многоугольников лежали в узлах сетки квадратов?

Первое решение (геометрическое). 1°. Если отрезок, соединяющий два узла сетки квадратов, сдвинуть параллельно так, чтобы один его конец снова совпал с узлом, то и другой конец перенесенного отрезка совпадает с узлом сетки. Воспользуемся этим для доказательства того, что *при $n > 6$ правильный n-угольник не может быть начерчен с соблюдением условий задачи.*

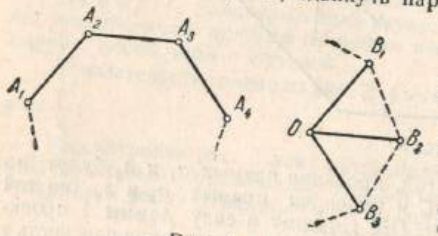


Рис. 5.

Допустим противное, и пусть $A_1A_2\ldots A_n$ — правильный n-угольник с вершинами в узлах сетки квадратов (рис. 5). Из произвольного узла O сетки отложим отрезки OB_1, OB_2, ..., OB_n, равные, параллельные и одинаково направленные с отрезками A_1A_2, A_2A_3, ..., A_nA_1. Нетрудно убедиться, что $B_1B_2\ldots B_n$ — снова правильный n-угольник, удовлетворяющий условиям задачи. При этом коэффициент подобия $k = \dfrac{B_1B_2}{A_1A_2} = 2\sin\dfrac{\angle B_1OB_2}{2} = 2\sin\dfrac{\pi}{n}$, откуда следует, что $k < 1$, если $n > 6$. Проводя аналогичное построение по отношению к n-угольнику $B_1B_2\ldots B_n$ и т. д., мы получим последовательность удовлетворяющих условию задачи правильных n-угольников, стороны которых неограниченно убывают, что, разумеется, невозможно (отрезок, соединяющий два узла сетки, не может быть меньше стороны квадрата сетки).

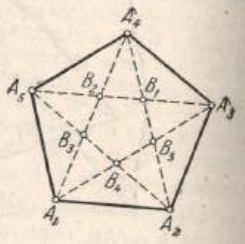


Рис. 6.

2°. Аналогично можно убедиться в *отсутствии удовлетворяющего условию задачи пятиугольника*: для этого достаточно заметить, что если бы такой пятиугольник $A_1A_2A_3A_4A_5$ существовал, то условию задачи удовлетворял бы и (меньший!) правильный пятиугольник $B_1B_2B_3B_4B_5$, изображенный на рис. 6 (B_1 есть узел сетки квадратов, ибо отрезки A_1A_2 и A_5B_1 равны, параллельны и одинаково направлены, поскольку $A_1A_5B_1A_2$ — параллелограмм: $A_1A_2 \parallel A_5B_1$ и $A_1A_5 \parallel A_2B_1$; нетрудно подсчитать, что в этом случае коэффициент подобия $k = \dfrac{B_1B_2}{A_2A_1} = \dfrac{A_4B_2}{A_4A_1} = \dfrac{3-\sqrt{5}}{2} < 1$).

3°. То, что *на сетке квадратов нельзя расположить правильный треугольник так, чтобы вершины его совпали с узлами сетки*, хорошо известно

(см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы элементарной математики, ч. 2, М., 1952, решение задачи 37 а). Отсюда вытекает, что и *при $n = 6$ построение удовлетворяющего условию задачи n-угольника невозможно.*

Итак, *задача имеет решения лишь при $n = 4$* (причем стороны квадрата могут и не совпадать с линиями сетки; см. рис. 7).

Второе решение (аналитическое). Известно, что *если угол φ соизмерим с π и отличен от $n\dfrac{\pi}{2}$ и $\dfrac{\pi}{4} + n\dfrac{\pi}{2}$ (где n целое)*, то $\operatorname{tg}\varphi$ *иррационален* (см., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избр. задачи и теоремы элементарной математики, ч. 1, М., 1959, задача 239б). Внутренний угол правильного n-угольника $\varphi_n = \dfrac{\pi(n-2)}{n}$. Если такой n-угольник удовлетворяет условию задачи, то $\operatorname{tg}\varphi_n$ не может быть иррациональным:

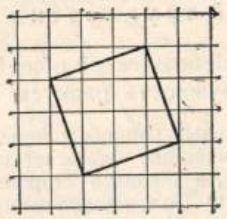


Рис. 7.

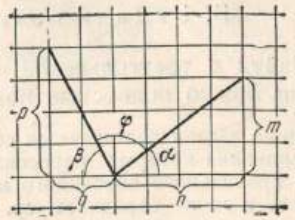


Рис. 8.

в самом деле, если φ — угол между двумя произвольными отрезками, соединяющими узлы сетки, то (рис. 8) $\operatorname{tg}\alpha = \dfrac{m}{n}$, $\operatorname{tg}\beta = \dfrac{p}{q}$ и $\operatorname{tg}\varphi = -\operatorname{tg}(\alpha+\beta) = \dfrac{\operatorname{tg}\alpha + \operatorname{tg}\beta}{1 - \operatorname{tg}\alpha\operatorname{tg}\beta}$ рациональны. Таким образом, $\dfrac{\pi(n-2)}{n}$ может равняться лишь $\dfrac{\pi}{2}$ или $\dfrac{3\pi}{4}$, откуда $n = 4$ или $n = 8$.

Заметим теперь, что у равнобедренного треугольника, образованного двумя соседними сторонами правильного n-угольника, удовлетворяющего условию задачи, и соединяющей их концы диагональю, квадраты всех сторон целые (они являются гипотенузами прямоугольных треугольников с целочисленными катетами; здесь за единицу мы принимаем сторону квадрата сетки). Отсюда по теореме косинусов имеем

$$b^2 = a^2 + a^2 - 2a^2\cos\varphi_n \quad \text{или} \quad \cos\varphi_n = \frac{b^2 - 2a^2}{2a^2},$$

где b^2 и a^2 — целые числа; таким образом, $\cos\varphi_n$ *должен быть рационален.* Но значение $n = 8$ не удовлетворяет этому условию; таким образом, $n = 4$ *является единственным возможным значением n.*

В. Л. Рабинович (Петропавловск-Казахстанский)

Дополнительный список лиц, приславших решения задач

Задача 1 — *У. Асекритов* (Ярославль) — геометрическое решение.
Задача 8 — *Э. Г. Кирьяцкий* (Вильнюс).
Задача 12 — *Е. Н. Ламбина* (Минск), *В. В. Малинин* (Ленинград).
Задача 14 — *С. Б. Аблялимов* (Уральск), *Е. Н. Ламбина* (Минск), *В. В. Малинин* (Ленинград), *З. А. Чантурия* (Тбилиси), *С. Шушбаев* (Ташкент).
Задача 15 — *С. Б. Аблялимов* (Уральск), *Е. Н. Ламбина* (Минск), *В. В. Малинин* (Ленинград), *З. А. Чантурия* (Тбилиси).
Задача 16 — *Ю. И. Кузьмин* (Москва), *Е. Н. Ламбина* (Минск), *П. С. Марголите* (Ярославль), *С. Шушбаев* (Ташкент).
Задача 17 — *Н. В. Аршава* (Харьковская обл.), *Е. Н. Ламбина* (Минск), *В. В. Малинин* (Ленинград), *З. А. Чантурия* (Тбилиси).
Задача 18 — *П. С. Марголите* (Ярославль), *Н. Малютина* (Чита).

Б. Задачи повышенной трудности

1. Найти n треугольников наименьшей возможной общей площади, которыми можно полностью покрыть окружность радиуса 1.

Решение задачи доставляется системой равных равнобедренных треугольников, основания которых пересекают ограниченный окружностью круг по сторонам правильного вписанного n-угольника, а боковые стороны касаются окружности в своих серединах (ср. Л. Фейеш Тот, Расположения на плоскости, на сфере и в пространстве, М., 1958, стр. 81 и примечание [43]).

3. Оптический прибор, называемый *трипельпризмой*, состоит из трех плоских зеркал, образующих прямой трехгранный угол. Этот набор зеркал отражает каждый падающий на него луч света в направлении, строго противоположном первоначальному, что позволяет, например, заменить в некоторых случаях осветительную систему маяка трипельпризмой (которую в этом случае требуется отыскать прожектором с судна). Найти все возможные системы плоских зеркал, отражающие каждый падающий луч в строго противоположном направлении.

Подобные системы зеркал были в 1953 г. перечислены И. Б. Кеплером (США); ниже излагаются его рассуждения.

1°. Условимся прежде всего каждый луч l заменять такой точкой L единичной сферы Σ с центром O, что направление луча задается вектором $\overrightarrow{LO}$. Плоскому зеркалу π будет отвечать большой круг p сферы, плоскость которого параллельна π; лучам l, падающим на зеркало с отражающей стороны, отвечают точки L, заполняющие одну из ограниченных p полусфер, которую будем называть «отмеченной». Если луч l отражается от зеркала π, то отраженному лучу l_1 отвечает точка L_1 сферы, симметричная точке L относительно большого круга p.

2°. Пусть π и ρ — два плоских зеркала нашей системы зеркал, которым отвечают большие круги p и r сферы; через D обозначим «двуугольник», образованный пересечением «отмеченных» полусфер, отвечающих зеркалам π и ρ (рис. 9). Пусть отраженный от π луч l_1 падает на ρ с отражающей стороны; при этом, выбрав l так, чтобы точка его падения на π была достаточно

близка к линии s пересечения π и ρ, мы можем добиться, чтобы луч l_1 не был задержан до падения на ρ никаким другим зеркалом нашей системы. Если отраженный от ρ луч l_2 падает на π с отражающей стороны, то мы также можем добиться, чтобы он был отражен именно этим зеркалом (заметим, что если точка падения l на π очень близка к s, то и точка падения l_1 на ρ будет близка к s). Таким образом, мы можем найти луч данного направления $\overrightarrow{LO}$ (где L — произвольная точка двуугольника D), который будет последовательно отражаться от зеркал π и ρ, пока не окажется падающим на эти зеркала

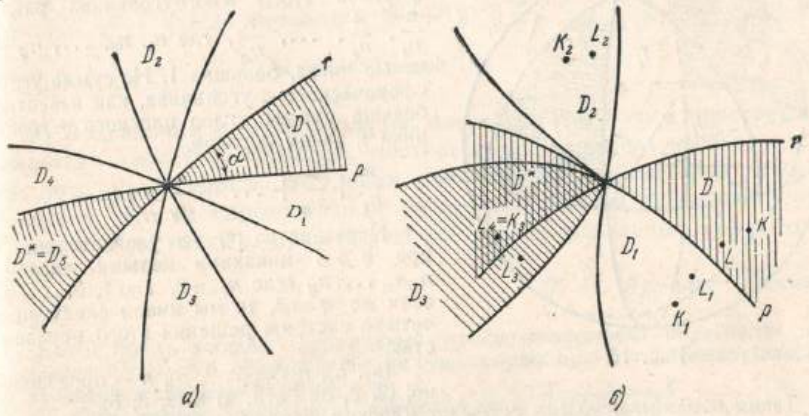


Рис. 9.

с неотражающей стороны, т. е. пока отвечающая последнему лучу l_n точка L_n (полученная из точки L n-кратным отражением от p и r), не попадет в симметричный D двуугольник D^*.

Заметим теперь, что если точка L принадлежит двуугольнику D, то точка L_1, отвечающая отраженному от π лучу l_1, принадлежит двуугольнику D_1, полученному из D отражением от стороны p двуугольника; точка L_2, отвечающая лучу l_2, полученному из l_1 отражением от ρ, принадлежит двуугольнику D_2, симметричному D_1 относительно r; точка L_3, отвечающая лучу l_3, полученному из l_2 отражением от π, принадлежит двуугольнику D_3, симметричному D_2 относительно p и т. д. Рассмотрим все получающиеся таким образом двуугольники $D_1, D_2, D_3, \ldots$ и пусть D_n — первый из них, имеющий общие точки с D^*. A priori возможны два случая:

а) Двуугольник D_n совпадает с D^*. При этом, как нетрудно видеть (см. рис. 9, *а*), *угол α двуугольника равен целой части* π $\left(\text{т. е. } \alpha=\frac{\pi}{n}\right)$.

б) Пересекающийся с D^* двуугольник D_n не совпадает с D^* (рис. 9, *б*). Покажем, что если система зеркал удовлетворяет условиям задачи, то *случай* б) *не может иметь места*. В самом деле, пусть этот случай имеет место и пусть L_n — точка двуугольника D_n, не принадлежащая двуугольнику D^*, но весьма близкая к его стороне; в таком случае еще одно отражение переведет ее в точку $L_{n+1}=K_n$, принадлежащую пересечению D_n с D^*. Таким образом, найдутся два таких луча k и l направлений $\overrightarrow{KO}$ и $\overrightarrow{LO}$, что n-кратное отражение одного из них и $(n+1)$-кратное отражение второго приводят к лучам k_n и l_{n+1} одного направления. Если теперь последующие m отражений от зеркал нашей системы (теперь уже не обязательно лишь от зеркал π и ρ) переводят луч k_n в луч, по направлению противо-

положный k (так что луч k после $n+m$ отражений меняет направление на противоположное), то эти же m отражений переведут луч l_{n+1} в луч, по направлению противоположный направлению k, что противоречит нашему условию, согласно которому каждый луч должен отражаться системой зеркал в противоположном ему направлении.

3°. Пусть наша система зеркал содержит q зеркал. Им отвечают q «отмеченных» полусфер, в пересечении которых образуется сферический q-угольник M. Согласно предыдущему пункту, q углов этого многоугольника равны $\frac{\pi}{n_1}, \frac{\pi}{n_2}, \ldots, \frac{\pi}{n_q}$, где $n_1, n_2, \ldots, n_q$ — целые числа, большие 1. Но сумма углов сферического q-угольника, как известно, больше суммы углов плоского q-угольника [1]); таким образом,

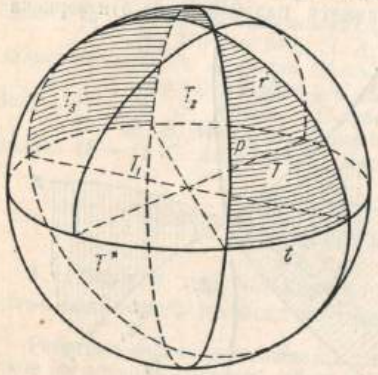


Рис. 10.

$$\frac{\pi}{n_1}+\frac{\pi}{n_2}+\ldots+\frac{\pi}{n_q}>\pi(q-2). \quad (*)$$

Неравенство (*) не удовлетворяется при $q>3$ никакими целыми числами $n_1, n_2, \ldots, n_q$ (где $n_i \geqslant 2$, $i=1, 2, \ldots, q$); если же $q=3$, то мы имеем следующие четыре системы решений этого неравенства:

$(n_1, n_2, n_3)=(2, 2, n)$, n — произвольно; (2, 3, 3); (2, 3, 4) и (2, 3, 5).

Таким образом, *искомая система плоских зеркал может лишь образовывать трехгранный угол, двугранные углы которого равны*

$$\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{n}\right), \left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{3}\right), \left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4}\right) \text{ или } \left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5}\right).$$

4°. Рассмотрим равнобедренный сферический треугольник T, углы которого равны $\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{n}$ (рис. 10). Последовательно отражая этот треугольник n раз относительно равных его сторон p и r, мы придем к противоположному T треугольнику T_n «отмеченной» полусферы, задаваемой третьей стороной t; отразив затем T_n относительно t, мы получим треугольник $T_{n+1}=T^*$, противоположный T (симметричный T относительно центра O сферы). Однако если n нечетно, то сторона p треугольника T после $n+1$ отражений совместится со стороной r^* треугольника T^*, симметричной стороне r; таким образом, точка L треугольника T в результате этих $n+1$ отражений, вообще говоря, не перейдет в диаметрально противоположную ей точку L^* сферы и, значит, отвечающий L луч l после $n+1$ отражений не изменит направления на противоположное. Напротив, если n четно, то треугольник T в результате $n+1$ отражений (выполненных в любом порядке, однако так, чтобы каждое отражение от p, r или t переводило треугольник с отмеченной плоскости, отвечающей этому большому кругу, на неотмеченную) совместится с противоположным ему треугольником T^* так, что каждая точка T перейдет в диаметрально ей противоположную. Таким образом, система трех плоских зеркал, образующих двугранные углы $\frac{\pi}{2}, \frac{\pi}{2}$ и $\frac{\pi}{2m}$, удовлетворяет условию задачи,

[1]) См., например, Ж. Адамар, Элементарная геометрия, ч. 2, М., Учпедгиз, 1948.

в то время как система зеркал, образующих углы $\frac{\pi}{2}, \frac{\pi}{2}$ и $\frac{\pi}{2m+1}$, этому условию не удовлетворяет.

Подобным же образом можно доказать, что система зеркал, образующих углы $\frac{\pi}{2}, \frac{\pi}{3}$ и $\frac{\pi}{3}$, не удовлетворяет условию задачи; две же другие системы зеркал ему удовлетворяют.

Итак, *все решения задачи даются трехгранными «зеркальными» углами с двугранными углами:* $\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2m}\right)$ (число отражений $2m+1$); $\left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{4}\right)$ (число отражений 9) и $\left(\frac{\pi}{2}, \frac{\pi}{3}, \frac{\pi}{5}\right)$ (число отражений 15).

9. Правильным n-угольником (на плоскости или в пространстве) называется замкнутая ломаная, состоящая из n равных звеньев и такая, что углы между соседними звеньями все равны между собой. Известно, что на плоскости правильные n-угольники существуют при любом n. При каких n существуют неплоские правильные многоугольники?

В. И. Арнольд (Москва)

Имеет место теорема: *правильный пространственный n-угольник существует при $n=4$ и при $n \geqslant 6$ и не существует при $n=5$.* Доказательство ее проведем в несколько шагов.

1°. *Существует правильный пространственный четырехугольник* — такой четырехугольник можно получить перегибом ромба по меньшей диагонали.

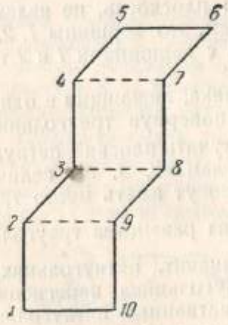


Рис. 11.

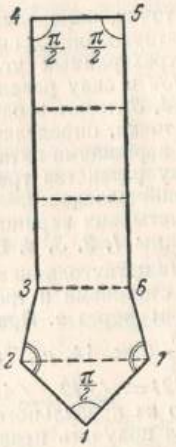


Рис. 12.

2°. *Существует правильный пространственный $2k$-угольник, где $k \geqslant 3$,* — такой $2k$-угольник (все углы которого прямые) можно получить, перегнув $k-2$ раз плоский прямоугольник, бо́льшая сторона которого в $k-1$ раз больше меньшей (рис. 11).

3°. *Существует правильный $(2k+1)$-угольник, где $k \geqslant 3$.* Такой многоугольник (все углы которого также прямые) можно получить из изображенного на рис. 12 плоского семиугольника *1234567* с тремя прямыми углами

$\left(\angle 1=\angle 4=\angle 5=\frac{\pi}{2}\right)$, пятью равными сторонами ($\overline{12}=\overline{23}=\overline{45}=\overline{67}=\overline{71}$) и двумя сторонами, в $k-2$ раз бóльшими остальных [$\overline{34}=\overline{56}=(k-2)\,\overline{12}$], перегнув этот семиугольник $k-1$ раз по пунктирным линиям.

4°. Так как совершенно ясно, что нет неплоского треугольника, то остается показать, что *не существует неплоского правильного пятиугольника.*

Допустим, что такой пятиугольник *12345* существует. Докажем, что в таком случае *четыре его вершины* (или *три последовательные стороны*) *лежат в одной плоскости.* Действительно, из пяти точек *1, 2, 3, 4, 5,* не лежащих в одной плоскости, всегда найдутся три такие, что остальные две точки лежат по одну сторону от задаваемых этими тремя точками плоскости. При этом могут иметь место два случая.

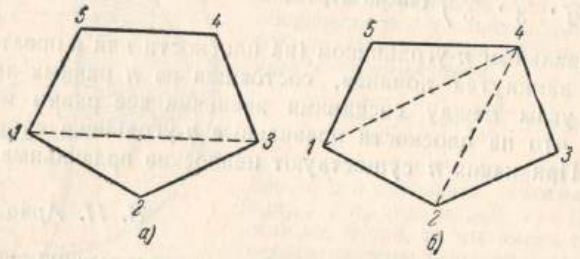


Рис. 13.

а) Три точки, определяющие указанную плоскость, суть последовательные вершины пятиугольника, скажем, вершины *1, 2, 3,* (рис. 13, *а*). Тогда в силу равенства трехгранных углов с вершинами *1* и *3* (по трем плоским углам; $\angle 315=\angle 134$ в силу равенства по трем сторонам треугольников *315* и *134*) точки *1, 5, 4, 3* лежат в одной плоскости.

б) Три точки, определяющие указанную плоскость, не являются последовательными вершинами пятиугольника; скажем, это вершины *1, 2, 4* (рис. 13, *б*). Тогда в силу равенства трехгранных углов с вершинами *1* и *2* точки *1, 2, 3, 5* лежат в одной плоскости.

Пусть четырьмя вершинами пятиугольника, лежащими в одной плоскости, будут вершины *1, 2, 3, 4.* В таком случае, повернув треугольник *145* вокруг диагонали $\overline{14}$ пятиугольника, мы сможем получить плоский пятиугольник *12345* с равными сторонами и тремя равными углами *2, 3, 5′*, величину которых мы обозначим через α. При этом a priori могут иметь место три случая.

а) $\alpha>\frac{\pi}{2}$ (рис. 14, *а*). Так как $\overline{24}=\overline{14}$ (из равенства треугольников *15′4* и *234*), то $\angle 421=\angle 412$, $\angle 423=\angle 415'$ и, значит, пятиугольник *12345′* правильный. Но из правильного плоского пятиугольника перегибом по диагонали *14* нельзя получить правильный пространственный пятиугольник (не будут равными углы между сторонами). Значит, этот случай отпадает.

б) $\alpha=\frac{\pi}{2}$. Здесь треугольник *15′4* должен быть одновременно равносторонним (ибо $\overline{14}=\overline{23}=\overline{15'}$) и прямоугольным; поэтому случай б) также отпадает.

в) $\alpha<\frac{\pi}{2}$. Здесь пятиугольник *12345* может иметь такой вид, как это изображено на рис. 14 *б, в*. Однако в случае рис. 14, *б* мы в точности так же, как и выше, убеждаемся, что $\angle 1=\angle 4=\alpha$, т. е. что α — это угол правильного пятиугольника, и следовательно, т у п о й. В случае же рис. 14, *в* также

имеем $\overline{24}=\overline{14}$ $(=\overline{13})$, при этом $\overline{23}\parallel\overline{41}$, $\angle 421=\angle 412=\angle 123=\alpha$ и (из $\triangle$ *423*) $\alpha+2\alpha+2\alpha=\pi$, $\alpha=\frac{\pi}{5}$; $\angle 245'\,(=\angle 315')=2\alpha+\alpha+2\alpha=\pi$, т. е. *235′* — это треугольник с углами $2\alpha=\frac{2\pi}{5}$, $\frac{2\pi}{5}$, $\frac{\pi}{5}$, а *4* и *1* — две точки на его равных сторонах, такие, что $\overline{5'4}=\overline{5'1}=\overline{23}$. Но из такого треугольника

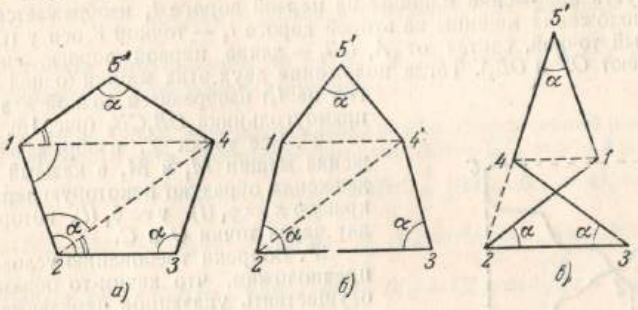


Рис. 14.

перегибом вокруг $\overline{14}$ также нельзя получить неплоский правильный пятиугольник: для того чтобы углы *4* и *1* стали равны величине $\alpha=\frac{\pi}{5}$, надо повернуть $\triangle$ *145′* вокруг $\overline{14}$ на 180°, что снова приводит нас к п л о с к о м у (звездчатому) правильному пятиугольнику.

А. П. Гарбер, В. И. Гарвацкий, В. Я. Ярмоленко (Винница)

Несколько более сложно решение автора задачи *Л. Н. Бескина* (Москва), содержащее, однако, одно простое соображение, позволяющее с самого начала исключить случай, приводящий в решении А. П. Гарбера, В. И. Гарвацкого и В. А. Ярмоленко к случаю рис. 14, *в*. А именно, В. И. Арнольд и Л. Н. Бескин замечают, что если у правильного пространственного пятиугольника *12345* диагонали $\overline{14}$ или $\overline{24}$ короче сторон (именно этот случай и соответствует рис. 14, *в*), то диагонали этого пятиугольника образуют также правильный пятиугольник *13524*, у которого уже диагональ $\overline{12}$ д л и н н е е стороны (и мы возвращаемся к случаю рис. 14, *а*).

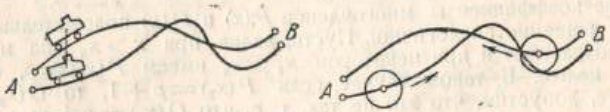


Рис. 15.

12. Из города A в город B ведут две дороги, каждая из которых не имеет самопересечений. Докажите, что если две машины M_1 и M_2 могут выехать одновременно из A по этим дорогам и проехать в B так, что расстояние между машинами ни в какой момент пути не будет превосходить 20 *м* (рис. 15, слева), то две круглые

платформы радиусом в 11 *м* не смогут выехать одновременно из A в B и из B в A и проехать по этим дорогам, не столкнувшись (рис. 15, справа) [1]).

Н. Н. Константинов (Москва)

Приводим решение автора задачи.

1°. Пусть положение машины на первой дороге l_1 изображается точкой X оси x, а положение машины на второй дороге l_2 — точкой Y оси y (OX — путь, пройденный точкой, считая от A, OB_1 — длина первой дороги; аналогичный смысл имеют OY и OB_2). Тогда положение двух этих машин (одной на l_1, другой на l_2) изобразится точкой Z замкнутого прямоугольника OB_1CB_2 (рис. 16).

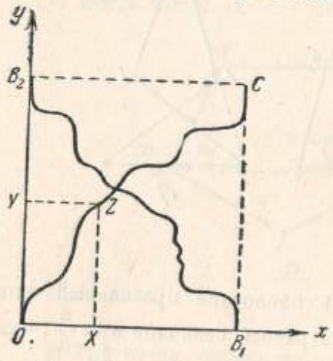


Рис. 16.

2°. Все точки Z, изображающие положения машин M_1 и M_2 в каждый момент t движения, образуют некоторую непрерывную кривую $x=\varphi_1(t)$, $y=\phi_1(t)$, которая проходит через точки O и C.

3°. Вопреки требованию условия задачи, предположим, что каким-то образом можно осуществить указанное передвижение платформ. Тогда аналогично получим кривую $x=\varphi_2(t)$, $y=\phi_2(t)$, соединяющую B_2 и B_1. Но две непрерывные кривые OC и B_1B_2, соединяющие противоположные вершины прямоугольника OB_1CB_2, обязательно должны пересечься в некоторой точке Z (это — непосредственное следствие непрерывности наших кривых). Точке Z отвечают такие точки на дорогах l_1 и l_2, что расстояние между этими точками не превосходит 20 *м* (ибо эта точка принадлежит линии OC, изображающей движение машин). Но это значит, что еще до того, как платформы (движение которых изображается линией B_1B_2) достигнут положения, изображаемого точкой Z, они неминуемо столкнутся (ибо расстояние между центрами платформ должно превышать 22 *м*).

13. Если многочлены $P(x)$ и $Q(x)$ (с вещественными коэффициентами) принимают целые значения в одних и тех же точках, то либо сумма $P(x)+Q(x)$, либо разность $P(x)-Q(x)$ постоянна.

Пусть, для определенности, $\lim\limits_{x\to\infty} P(x)=+\infty$ и $\lim\limits_{x\to\infty} Q(x)=+\infty$, т. е. старшие коэффициенты многочленов $P(x)$ и $Q(x)$ положительны (в других случаях решение аналогично). Пусть, далее, при $x>x_0$ оба многочлена монотонно возрастают и при некотором $x_1>x_0$ имеем $P(x_1)=p$, $Q(x_1)=q$, где p и q — целые. В таком случае, если $P(x_2)=p+1$, то $Q(x_2)=q+1$. В самом деле, допустим, что это не так, т. е. что $Q(x_2)=q+m$, где $m>1$. Тогда для некоторого $\tilde{x}$, где $x_1<\tilde{x}<x_2$, будем иметь $Q(\tilde{x})=q+1$. Но в этом случае $p<P(\tilde{x})<p+1$, что противоречит тому, что и многочлен $P(x)$ должен при $x=\tilde{x}$ иметь целочисленное значение.

Аналогично доказывается, что если $P(x_i)=p+i-1$ (где $x_i>x_0$), то $Q(x_i)=q+i-1$. Следовательно, многочлены $P(x)-P(x_1)$ и $Q(x)-Q(x_1)$

[1]) Условие этой задачи, приведенное в вып. 2 «Математического просвещения», содержало опечатку.

принимают одинаковые значения при бесчисленном множестве значений x_i (где $i=1, 2, \ldots$) и, значит,

$$P(x)-P(x_1)\equiv Q(x)-Q(x_1)$$

или

$$P(x)-Q(x)\equiv \text{const}.$$

Ф. Л. Варпаховский (Москва)

Близко к этому и решение *Л. Н. Бескина* (Москва).

14. Построить треугольник по высоте h_a, разности $b-c$ боковых сторон и радиусу r вписанной окружности.

Первое решение. Рассмотрим вписанный и вневписанный (касающийся стороны длины a) круги. Легко показать, что $BM=CL_1=p-b$ (обозначения см. на рис. 17); действительно, $BK=BM=\frac{1}{2}(BA+BC-AK-CM)=$ $=\frac{1}{2}(BA+BC-AN-CN)=\frac{1}{2}(c+a-b)$; $AQ=AP=p$, ибо $AQ=AP=$ $=\frac{1}{2}(AB+AC+BP+CQ)=\frac{1}{2}(AB+AC+BL_1+CL_1)=\frac{1}{2}(a+b+c)$; $CL_1=CQ=AQ-AC=p-b$. Поэтому

$$ML_1=a-2(p-b)=a-(a+c-b)=b-c,$$

и, стало быть, отрезок ML_1 нам известен. Отрезок $ML=2r$ также известен. Точка A является центром гомотетии рассматриваемых кругов; точки L и L_1, как точки касания кругов с параллельными касательными, отвечают друг другу в этой гомотетии; поэтому LL_1 проходит через A.

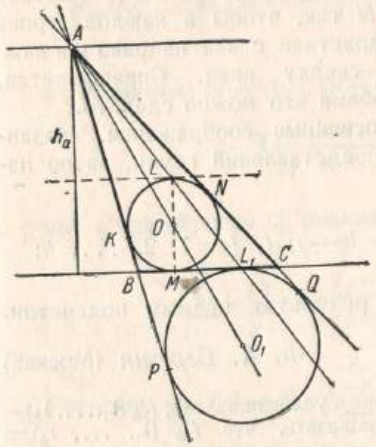


Рис. 17.

Построение. Строим $\triangle LML_1$ по катетам $2r$ и $b-c$; LL_1 продолжаем до пересечения в точке A с прямой, параллельной ML_1 и удаленной от ML_1 на расстояние h_a; из точки A проводим касательные AB и AC к кругу с диаметром ML.

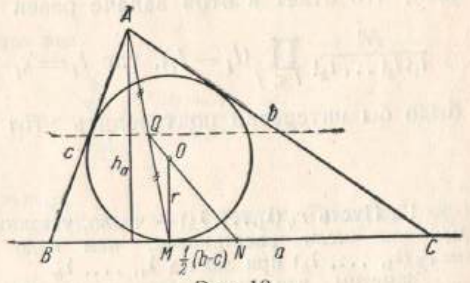


Рис. 18.

Второе решение. Расстояние точки M до середины N стороны BC (рис. 18) равно $BN-BM=\frac{1}{2}a-\frac{1}{2}(a+c-b)=\frac{1}{2}(b-c)$. Далее, для каждого описанного четырехугольника (в том числе и для вырожденного) имеет место теорема Ньютона: *центр вписанной окружности лежит на*

прямой, соединяющей середины диагоналей [1]). Применив эту теорему к (вырожденному) четырехугольнику $ABMC$, мы найдем, что середина Q отрезка AM лежит на прямой NO.

Построение. По катетам $\frac{1}{2}(b-c)$ и r строим треугольник OMN; на расстоянии $\frac{h_a}{2}$ от MN проводим среднюю линию треугольника и получаем в пересечении ее с прямой NO точку Q; на прямой MQ находим вершину A ($QA=QM$), из которой проводим две касательные к окружности с центром O и радиусом r.

З. А. Скопец (Ярославль)

Алгебраические решения задачи прислали *Л. Н. Бескин* (Москва), *Э. Г. Готман* (Печера), *В. Л. Рабинович* (Петропавловск-Казахстанский).

В. Проблемы

3 [2]). Рассмотрим схему типа изображенной на рис. 19, состоящую из произвольного числа k строк, содержащих соответственно $\lambda_1, \lambda_2, \ldots, \lambda_k$ клеток; $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k$ (на рис. 19 $k=6$ и $\lambda_1=7$, $\lambda_2=\lambda_3=5$, $\lambda_4=\lambda_5=4$, $\lambda_6=2$). Обозначим общее число клеток схемы (равное $\lambda_1+\lambda_2+\ldots+\lambda_k$) через N и зададимся задачей расставить в этих N клетках числа от 1 до N так, чтобы в каждой строке схемы числа возрастали слева направо и в каждом столбце — сверху вниз. Спрашивается, сколькими способами это можно сделать.

1	2	4	7	11	15	21
3	5	6	16	22		
8	9	12	17	26		
10	13	19	23			
14	18	25	27			
20	24					

Рис. 19.

Используя косвенные соображения, связанные с теорией представлений групп, автор нашел, что ответ в этой задаче равен

$$\frac{N!}{l_1!\,l_2!\ldots l_k!}\prod_{i<j}(l_i-l_j),\quad \text{где } l_i=\lambda_i+k-i;\ i,\ j=1,\ 2,\ \ldots,\ k;$$

было бы интересно подтвердить этот результат прямым подсчетом.

Ф. А. Березин (Москва)

1°. Пусть $r_k(\lambda_1,\ldots,\lambda_k)$ — число, указанное в условии задачи; $R_k(\lambda_1,\ldots,\lambda_k)$ — искомое число расстановок; нам надо доказать, что $R_k(\lambda_1,\ldots,\lambda_k)=r_k(\lambda_1,\ldots,\lambda_k)$ при любых $\lambda_1,\ldots,\lambda_k$.

Заметим, что выражение r_k определено для любых неотрицательных $\lambda_1,\ldots,\lambda_k$, а выражение R_k — только в случае $\lambda_1\geqslant\lambda_2\geqslant\ldots\geqslant\lambda_k>0$. Расширим область определения R_k.

[1]) См., например, Д. О. Шклярский, Н. Н. Ченцов, И. М. Яглом, Избранные задачи и теоремы элементарной математики, ч. 2, М., 1952, задача 115.

[2]) Условие задачи, напечатанное в вып. 2 «Математического просвещения», содержало ошибку.

а) Пусть $\lambda_1\geqslant\lambda_2\geqslant\ldots\geqslant\lambda_a>0$ и $\lambda_{a+1}=0,\ \cdots,\ \lambda_k=0$. Положим в этом случае

$$R_k(\lambda_1,\ \ldots,\ \lambda_k)=R_a(\lambda_1,\ \ldots,\ \lambda_a).$$

б) Если для некоторого i $\lambda_i=\lambda_{i+1}-1$, то положим $R_k(\lambda_1,\ \ldots,\ \lambda_k)=0$. (Легко видеть, что в этом случае r_k также равно нулю.)

2°. Доказательство проведем индукцией по k и по N. При $k=1$ и любом λ_1 $R_1(\lambda_1)=r_1(\lambda_1)$ $(=1)$. Допустим, что равенство $R_k=r_k$ справедливо для всех $k\leqslant k_0-1$. Докажем, что оно остается верным и для $k=k_0$.

При $N=k_0$ $R_{k_0}=r_{k_0}$ $(=1)$. Пусть теперь равенство $R_{k_0}=r_{k_0}$ справедливо для всех $N\leqslant N_0-1$. Нужно показать, что $R_{k_0}=r_{k_0}$ и при $N=N_0$.

3°. Пусть $\lambda_1\geqslant\ldots\geqslant\lambda_{k_0}$ и $\lambda_1+\ldots+\lambda_{k_0}=N_0$. Число N_0 может занимать в каждой строке лишь крайнюю правую клетку (причем последующие строки должны содержать меньше клеток). Если N_0 занимает последнюю клетку i-й строки, то общее число таких (допустимых) размещений равно $R_{k_0}(\lambda_1,\ \ldots,\ \lambda_i-1,\ \ldots,\ \lambda_{k_0})$. Значит, общее число всех размещений

$$R_{k_0}(\lambda_1,\ \ldots,\ \lambda_{k_0})=\sum_{i=1}^{k_0}R_{k_0}(\lambda_1,\ \ldots,\ \lambda_i-1,\ \ldots,\ \lambda_{k_0}), \tag{1}$$

или, в силу предположения индукции,

$$R_{k_0}(\lambda_1,\ \ldots,\lambda_{k_0})=\frac{(N_0-1)!\prod_{i<j}(l_i-l_j)}{l_1!\,l_2!\ldots l_{k_0}!}\sum_{i=1}^{k_0}l_i\prod_{j\neq i}\frac{l_i-l_j-1}{l_i-l_j}. \tag{2}$$

4°. Обозначим $(x-l_1)\ldots(x-l_{k_0})$ через $F(x)$. Тогда

$$\prod_{j\neq i}\frac{l_i-l_j-1}{l_i-l_j}=-\frac{F(l_i-1)}{F'(l_i)},$$

и сумма в правой части (2) принимает вид

$$\sum_{i=1}^{k_0}l_i\prod_{j\neq i}\frac{l_i-l_j-1}{l_i-l_j}=-\sum_{i=1}^{k_0}l_i\frac{F(l_i-1)}{F'(l_i)}. \tag{3}$$

Учитывая, что $F(l_i-1)=\sum_{p=1}^{k_0}\frac{F^{(p)}(l_i)}{p!}(-1)^p$ (по формуле Тэйлора), преобразуем правую часть (3) к виду

$$-\sum_{i=1}^{k_0}l_i\frac{F(l_i-1)}{F'(l_i)}=-\sum_{i=1}^{k_0}\sum_{p=1}^{k_0}(-1)^p l_i\frac{F^{(p)}(l_i)}{p!\,F'(l_i)}=$$

$$=\sum_{i=1}^{k_0}l_i-\sum_{i=1}^{k_0}l_i\frac{F''(l_i)}{F'(l_i)\,2!}-\sum_{p=3}^{k_0}\sum_{i=1}^{k_0}(-1)^p l_i\frac{F^{(p)}(l_i)}{p!\,F'(l_i)}. \tag{4}$$

Воспользуемся теперь формулой Эйлера [1]), согласно которой для любого многочлена $F(x)$ k-й степени с различными корнями $l_1, \ldots, l_k$

$$\sum_{i=1}^{k} \frac{l_i^p}{F'(l_i)} = \begin{cases} 0 & \text{при } p < k-1, \\ 1 & \text{при } p = k-1. \end{cases}$$

Отсюда, если $P(x) = a_0 x^n + \ldots + a_n$, то

$$\sum_{i=1}^{k} \frac{P(l_i)}{F'(l_i)} = \begin{cases} 0 & \text{при } n < k-1, \\ a_0 & \text{при } n = k-1. \end{cases}$$

Теперь легко преобразовать выражение (4) к окончательному виду

$$\sum_{i=1}^{k_0} l_i - \frac{k_0(k_0-1)}{2} = \sum_{i=1}^{k_0} \lambda_i = N_0,$$

откуда, наконец, следует [см. (1), (2)], что

$$R_{k_0}(\lambda_1, \ldots, \lambda_{k_0}) = \frac{N_0! \prod_{i<j} (l_i - l_j)}{l_1! \ldots l_{k_0}!}.$$

Л. А. Гутник (Москва)

М. Тер-Мкртачан (Ереван), также решивший эту задачу, указывает наряду с решением, опирающимся на рекуррентное соотношение (1), также и другое решение, исходящее из эквивалентности поставленной задачи следующей: *из урны, содержащей N шаров k разных цветов, последовательно извлекаются все шары; найти вероятность того, что в каждый момент число извлеченных шаров i-го цвета (где $i = 1, 2, \ldots, k-1$) будет не меньше числа извлеченных шаров $(i+1)$-го цвета.* Впрочем, это решение, подробно проведенное Тер-Мкртачаном для случая $k = 2$, в общем случае приводит к весьма громоздким выкладкам и поэтому уступает приведенному выше.

М. Тер-Мкртачан указывает также, что решение более общей задачи о вероятности такого извлечения из урны N шаров k разных цветов, при котором число извлеченных шаров i-го цвета все время превосходило бы число извлеченных шаров $(i+1)$-го цвета не меньше чем на a_i (где $a_1, a_2, \ldots, a_{k-1}$ — заданные неотрицательные числа) представляло бы известный интерес для статистики; он же ставит задачу об определении всех вообще функций $R(\lambda_1, \lambda_2, \ldots, \lambda_k)$ от k целочисленных аргументов, удовлетворяющих уравнению (1).

В проблеме 4 предлагалось найти такое число k, что любое число n попарно пересекающихся выпуклых фигур «можно проткнуть k иглами»; другими словами, число k такое, что можно найти k точек, хоть одну из которых задевает каждая из n фигур. При этом отдельно предлагалось решить задачу для окружностей; для произвольных равных выпуклых фигур; для каких

[1]) См., например, В. А. Кречмар, Задачник по алгебре, М., 1959, задача 79.

угодно выпуклых фигур; в последнем случае оговаривалась необходимость ограничить как-то изменение диаметра и ширины фигур. Подобная оговорка явно необходима и в случае *равных* фигур; в самом деле, достаточно рассмотреть сетку из n^2 узких и длинных прямоугольников (рис. 20), для которых потребуется n «игл», чтобы убедиться в необходимости ограничить отношение диаметра фигуры к ее ширине (от этого отношения, по-видимому, должно зависеть число k).

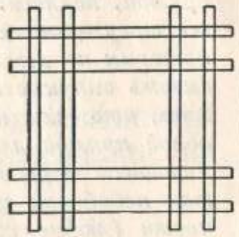


Рис. 20.

Редакция получила несколько писем, указывающих на это обстоятельство. В частности, *А. Хеппес* (Будапешт, Венгрия) считает, что эту задачу вообще нельзя перенести с кругов на произвольные выпуклые фигуры, поскольку здесь существенную роль играет то обстоятельство, что *круг остается связным множеством, если из него удалить пересечение с любым другим кругом*, а для произвольных выпуклых фигур это уже не будет верно. В этой связи он указывает на следующее предложение, в условии которого оказывается существенной оговорка подобного рода: *если A_1, A_2, A_3, A_4, A_5 — такие попарно пересекающиеся связные множества на плоскости, что удаление из каждого из них его пересечения с любым другим оставляет его связным, то найдутся три из этих множеств, имеющие общую точку* (три множества, которые «можно проткнуть одной иглой»).

ОДНА НЕРЕШЕННАЯ ПРОБЛЕМА П. ЭРДЁША

Совершенно ясно, что три точки, не лежащие на одной прямой, являются вершинами выпуклого треугольника (ибо все треугольники выпуклые). Четыре точки, никакие три из которых не лежат на одной прямой, могут уже не образовывать выпуклого четырехугольника; однако нетрудно показать, что если никакие три из пяти точек не лежат на одной прямой, то из них всегда можно выбрать четыре, являющиеся вершинами выпуклого четырехугольника. Можно так подобрать пять, шесть, семь или восемь точек на плоскости (любые три из которых по-прежнему не лежат на одной прямой), что никакие пять из них не являются вершинами выпуклого пятиугольника; однако если число точек равно девяти (или, тем более, больше девяти), то такие пять точек обязательно найдутся (доказательство последнего утверждения составляет отнюдь не простую задачу).

Известный венгерский математик П. Эрдёш доказал, что наименьшее число точек (никакие три из которых не лежат на одной прямой), из числа которых всегда можно выбрать шесть, являющиеся вершинами выпуклого шестиугольника, равно семнадцати (но это уже трудная теорема!).

Эрдёш заметил, что

$$3=2^1+1,\quad 5=2^2+1,\quad 9=2^3+1,\quad 17=2^4+1;$$

исходя из этого, он предположил, что наименьшее число точек плоскости, никакие три из которых не лежат на одной прямой, из числа которых всегда можно выбрать n точек, являющихся вершинами выпуклого n-угольника, равно

$$2^{n-2}+1.$$

Задача доказательства или опровержения гипотезы Эрдёша по форме совершенно элементарна, однако на сегодняшний день мы не имеем никаких надежд на ее скорое решение!

И. Я.

VI. МАТЕМАТИЧЕСКАЯ ЛИТЕРАТУРА

ЭКСПЕРИМЕНТАЛЬНЫЙ УЧЕБНИК ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ ДЛЯ АМЕРИКАНСКИХ СРЕДНИХ ШКОЛ

А. М. Яглом

(Москва)

Introductory Probability and Statistical Inference. An experimental course. Revised preliminary edition, prepared for the Commission on Mathematics, College Entrance Examination Board, New York, 1959.

Teachers' Notes and Answer Guide (Supplementary material for the revised preliminary edition of «Introductory Probability and Statistical Inference»), New York, 1959.

Развитие теории вероятностей с момента зарождения этой науки и до настоящего времени было несколько своеобразным. На самом первом этапе истории этой науки — в XVII в. — теория вероятностей рассматривалась как занимательный пустячок, как собрание курьезных задач, связанных в первую очередь с азартными играми в кости и карты. В следующем, XVIII в. впервые появились серьезные применения вероятностных расчетов к статистике народонаселения, страховому делу и некоторым другим частным вопросам; в начале XIX в. на базе теории вероятностей возникла весьма важная для всех естественных наук теория ошибок и обработки наблюдений. Вслед за этим, однако, наступил длительный период относительного застоя, в течение которого учение о вероятностях рассматривалось как специальный раздел математики со сравнительно узким кругом приложений; большие успехи в развитии математической теории вероятностей в конце XIX — начале XX вв. и широкое проникновение примерно в то же время статистических соображений в физику мало что изменили в этом отношении, так как вначале оба процесса развивались изолированно друг от друга. Лишь в последние десятилетия со всей очевидностью выяснилось, что понятие вероятности играет крайне важную роль буквально во всех областях науки и практической деятельности; в настоящее время многие крупные ученые считают

даже, что в основе всех вообще закономерностей реального мира лежат определенные статистические законы.

Естественно, что в последнее время во всём мире много говорится о том, что первоначальное знакомство с теорией вероятностей должно быть обязательным для каждого образованного человека, и, следовательно, теория вероятностей должна занять свое место в школьном курсе математики; показательно обсуждение этого вопроса на секции преподавания математики последнего Международного математического конгресса в Эдинбурге [1]). В настоящее время элементы теории вероятностей включены уже в программы средних школ Японии и Югославии; в значительном числе других стран горячо обсуждается вопрос о возможности изучения элементов этой теории в старших классах школы и проводятся первоначальные эксперименты в этом направлении. Недавно в Москве был получен экземпляр пересмотренного предварительного издания экспериментального курса теории вероятностей и математической статистики для американской high school (соответствующей старшим классам нашей средней школы), подготовленного группой авторов — ученых и педагогов (E. C. Douglas, F. Mosteller, R. S. Pieters, D. E. Richmond, R. E. K. Rourke, G. B. Thomas и S. S. Wilks). В предисловии сами авторы отмечают, что эта книга должна рассматриваться не как окончательный учебник, а лишь как его первоначальный набросок, который будет еще много раз исправляться и переделываться и по содержанию и по форме, в первую очередь, на основании непосредственного опыта преподавания. К этому надо добавить еще, что и объем книги (VIII + 240 стр.) заметно превосходит наибольший объем, который был бы допустим для учебника по сравнительно специальному предмету в школе. Тем не менее книга и в этом ее виде весьма любопытна и во многом поучительна: в ней нашли свое отражение как некоторые своеобразные стороны развития теории вероятностей в США, так и некоторые интересные тенденции преподавания математики в этой стране.

Книга открывается небольшой вступительной главой I (объемом около 2,5 стр.), носящей название «Что такое теория вероятностей и математическая статистика». Здесь на одной странице даются весьма скупые сведения об истории теории вероятностей (из ученых названы лишь Паскаль и Ферма), а затем приводятся конкретные примеры задач из теории вероятностей и математической статистики и делается вывод о том, что теория вероятностей — это наука, позволяющая выводить из закономерностей, относящихся ко всей массе однородных предметов [т. е. к генеральной совокупности (population); этот термин вводится тут же], некоторые заключения о составе выборок из этой совокупности, а математическая статистика — наука,

[1]) См. «Математическое просвещение», вып. 5, стр. 223.

позволяющая по данным о составе выборки из генеральной совокупности делать некоторые заключения о самой этой совокупности. Такое определение теории вероятностей и математической статистики звучит очень привлекательно и вполне доступно для школьников; к сожалению, его вряд ли можно признать правильным (во всяком случае в той части, которая относится к теории вероятностей). Вводя такое определение, авторы книги с самого начала неявно отождествляют вероятность с частотой данного признака в некоторой обширной (но, по-видимому, всё же конечной) совокупности; это может быть и не так страшно, если теорию вероятностей рассматривать лишь как рабочий аппарат математической статистики, но никак недопустимо, если стремиться выработать у учащихся правильные представления о математическом понятии вероятности, о статистических закономерностях и их значении.

Вслед за вводной идут две главы (40 стр. текста), фактически не относящиеся к теории вероятностей (что отмечают и авторы книги); эти главы называются «Приведение в систему и методы записи данных; частотные распределения» и «Подведение итогов ряда измерений; среднее значение и стандартное уклонение». В первой из этих глав живо и доходчиво рассказывается о различных формах табличного и графического представления совокупности данных; изложение сопровождается хорошо подобранными примерами и конкретными практическими указаниями. Основную роль в главе играют понятия «диаграммы накопленных сумм» (т. е. фактически эмпирической функции распределения) и частотной гистограммы; попутно определяются также понятия медианы, квартилей и квантилей (в советской литературе эти последние понятия можно найти лишь в специальных книгах по статистике). В главе III обсуждается вопрос об основных числовых характеристиках множества однородных эмпирических данных, и учащиеся естественно подводятся к понятию среднего арифметического и (что значительно сложнее) стандартного (т. е. среднеквадратичного) уклонения. Введение этих понятий также сопровождается значительным числом примеров и рядом практических указаний относительно методов вычисления соответствующих величин. Здесь же дается доказательство общего неравенства Чебышева, рассматриваемого как оценка минимальной доли общего числа данных, содержащейся внутри определенного симметричного интервала, окружающего среднее арифметическое значение. Попутно указывается, что на практике эта доля обычно оказывается много больше своей оценки по неравенству Чебышева, и в качестве эмпирического правила, значительные отклонения от которого встречаются крайне редко, приводятся цифры, отвечающие распределению Гаусса. Следует отметить также, что и в этих главах и во всех последующих буквально каждый шаг в изложении сопровождается рядом хорошо подобранных упражнений для учащихся; значительная

часть этих упражнений содержит примеры, заимствованные из реальной жизни и хорошо поясняющие практическое значение рассматриваемого материала.

В следующей главе «Интуитивное понятие о вероятности» впервые появляется само слово *вероятность*. В начале этой главы вкратце обсуждается вопрос о связи математики и действительной жизни (здесь затрагиваются вопросы о математической идеализации и о реальном смысле математических постулатов); вслед за этим авторы переходят к «классическому определению вероятности» (как отношения общего числа равноправных исходов опыта к числу благоприятных исходов) и к простейшим примерам на это определение. Почти сразу за классическим определением вводится понятие «выборочного пространства» (пространства, точками которого являются всевозможные «элементарные исходы» рассматриваемого опыта); случайные события определяются в этой связи как множества точек выборочного пространства. Далее вводятся понятия «несовместимых событий», «дополнительных событий» и «независимых событий» (последнее чисто формально — при помощи равенства $P(A \text{ и } B) = P(A) \cdot P(B)$ без всяких словесных разъяснений) и приводятся теоремы сложения и умножения вероятностей (включая и определение условных вероятностей). Общее понятие вероятности (в случае, когда классическое определение неприменимо) иллюстрируется на примере «неправильной монеты» (изогнутой и с острием с одной стороны); при этом утверждается только, что и для такой «монеты» существует определенная вероятность падения той или иной стороной; оценкой этой вероятности служит относительная частота в длинной серии выпадений (но никак не объясняется, что́ же в этом случае значит само слово вероятность). Вслед за этим приводится ряд практических примеров (не всегда очень важных), в которых требуется придать некоторой совокупности событий одинаковые вероятности, и в этой связи объясняется, что́ такое таблицы случайных чисел, приводится образец такой таблицы и указывается, ка́к этой таблицей пользоваться. Попутно на одном элементарном примере дается представление о «методе Монте-Карло», характеризуемом, как «метод решения обычных (не вероятностных) задач с использованием понятия вероятности». Любопытно отметить, что, несмотря на стремление к чрезмерной сжатости, авторы всё же сочли возможным потратить на разъяснение вопроса о «таблицах случайных чисел» больше 10 страниц текста.

Глава V посвящена «Формальному подходу к понятию вероятности». Фактически авторы понимают под этим аксиоматический подход: утверждается, что в общем случае вероятность определяется заданием некоторого «выборочного пространства» (состоящего из отдельных точек; впрочем, всё изложение строится так, чтобы в будущем от этого предположения было очень легко отказаться), каждой точке которого произвольным образом приписывается определенная

вероятность с тем лишь ограничением, что сумма всех этих вероятностей равна единице. Затем снова напоминается, что событие — это некоторое множество точек выборочного пространства, и еще раз доказываются все теоремы, приведенные в главе IV (с добавлением еще теоремы о вероятности суммы совместимых событий), с использованием формальных обозначений теории множеств ($A \cup B$, $A \cap B$ и т. д.) и чертежей, наглядно иллюстрирующих вводимые соотношения. В заключение главы сказано несколько слов о понятии математического ожидания (без употребления термина «случайная величина»).

В следующей VI главе рассматривается последовательность одинаковых независимых испытаний и дается вывод биномиального распределения для этого случая. Особенностью, отличающей эту главу от соответствующих мест советских учебников по теории вероятностей, является подробное описание «Таблиц биномиального распределения» (содержащих $\sum_{k=r}^{n} C_n^k p^k (1-p)^{n-k}$ при различных n, p и r); подобные таблицы для $n = 5$, 10, 15, 20 и 25 и $p = 0{,}01$, 0,05, 0,1, 0,2 ..., 0,8, 0,9, 0,95 и 0,99 приведены в книге вместе с рядом примеров на их использование (некоторые из этих примеров требуют также интерполяции внутри таблицы или искусственного перехода к дополнительным значениям n).

Глава VII посвящена статистическим задачам о приемочном контроле и о проверке гипотез. В начале главы подробно рассказывается о практическом значении приемочного контроля (или, общée, статистического контроля качества продукции) для промышленности, и в этой связи приводятся некоторые выразительные цифры: во время второй мировой войны Управление военной промышленностью США лишь за период между 1943 и 1945 гг. 33 раза организовало краткосрочные курсы для статистических контролеров качества; сразу после войны было организовано Американское общество по контролю качества продукции, которое в 1957 г. насчитывало уже 9671 членов; выгода, полученная за 1957 г. американской промышленностью от применения статистических методов контроля, была оценена в 4 миллиарда долларов. Вслед за этим перечисляется большое число практически важных вопросов (типа «Может ли фермер увеличить снимаемый урожай, увеличив количество вносимых азотистых удобрений?»), решение которых сводится к проверке гипотезы о том, что в определенных условиях вероятность p некоторого события принимает значение, превосходящее некоторое заданное число (или, в других задачах, равное некоторому заданному числу). Дальнейшая часть главы посвящена методам решения задач о приемочном контроле и о проверке гипотезы о значении неизвестной нам вероятности p. При рассмотрении задачи приемочного контроля прежде всего дается представление о простейшем плане контроля, заключающемся в том, что из большой партии изделий проверяется небольшее число n и вся партия

отвергается или принимается в зависимости от того, будет ли среди проверенных изделий число бракованных превосходить заданное число a или нет. Далее на конкретных числовых примерах показывается законность использования в этой задаче биномиального распределения (сводящегося к предположению о том, что каждое проверяемое изделие после проверки снова возвращается в партию, так что вероятность извлечь бракованное изделие для всей выборки остается одной и той же). После этого вводится понятие «оперативной характеристики данного метода контроля» — диаграммы, характеризующей вероятность того, что будет принята партия изделий с долей брака p; при этом приводятся конкретные образцы таких оперативных характеристик, указываются общие методы их расчета и обсуждается вопрос о принципах выбора того или иного плана контроля в зависимости от практических запросов задачи. В заключение главы рассматривается гипотеза о том, что вероятность p принимает заданное значение p_0, и подробно разбирается вопрос о ее статистической проверке; в качестве конкретного примера здесь выбрана не слишком серьезная задача о разоблачении мальчика, утверждавшего, что он может не видя отгадать цвет масти выбранной из колоды карты. В этой связи авторы вспоминают, что подобного же рода задача о разоблачении «дамы — дегустатора чая с молоком», утверждавшей, что она может по вкусу определить, было ли молоко налито в чашку до чая или после чая, обсуждалось в первой серьезной книге по планированию статистических экспериментов, написанной крупнейшим английским статистиком XX в. Р. Фишером.

Последняя VIII глава книги посвящена вопросу об оценке среднего значения по данным конечной выборки. Здесь подсчитываются среднее значение и стандартное уклонение выборочного среднего, оценивается различие этих уклонений для случая повторной и бесповторной выборок и приводятся формулы для среднего значения и дисперсии биномиального распределения. В заключение на 1,5 страницах излагается закон больших чисел (со ссылкой на неравенство Чебышева, доказанное в гл. III) и кратко объясняется его смысл.

Книга заканчивается четырьмя дополнениями, в первом из которых приводятся простейшие понятия (и обозначения) теории множеств, во втором излагаются элементы комбинаторики (включая сюда и теорему о биноме Ньютона, причем даже и для нецелых показателей и с конкретными примерами использования этой общей биномиальной теоремы для приближенных вычислений), в третьем, на 7 страницах со многими примерами и упражнениями, разъясняется метод математической индукции и, наконец, в четвертом даются общие указания об использовании индексов и знака суммирования $\sum$ (и индексы и знак $\sum$ широко применяются в книге). В самом конце приводится также список дополнительных упражнений, краткий перечень ответов на все упражнения и указатель использованных в книге символов

и специальных терминов. Кроме того, к книге приложено еще специальное дополнение для учителя («Teachers' Notes and Answer Guide» объемом в 123 стр.), содержащее ряд методических и литературных указаний и дополнений, а также полное решение всех имеющихся в основной книге упражнений.

В целом книга свидетельствует о большом значении, придаваемом в США конкретным практическим задачам вообще и математической статистике в частности. Как элементарное введение в математическую статистику, книга обладает большими методическими достоинствами: она живо и доходчиво написана, содержит много полезных практических советов и ярких примеров и в целом дает достаточно отчетливое представление о содержании математической статистики как науки и о ее приложениях. Вообще, удачное освещение вопроса о приложениях математики безусловно является сильнейшей стороной книги — значительная часть содержащегося здесь материала (например, советы об использовании в работе специальных таблиц, о целесообразном расположении вычислений и о применении приближенных методов счета) будет очень полезной для школьника даже и вне связи с теорией вероятностей и математической статистикой. Сугубому «практицизму» авторов несколько противоречит их стремление широко (но иногда, пожалуй, слишком формально) использовать общие теоретико-множественные концепции; здесь, по-видимому, большую роль сыграли определенные общие тенденции широкого внедрения в преподавание теоретико-множественных представлений, характерные для современной американской педагогики. С другой стороны, эта книга — конечно, не учебник теории вероятностей; принципиальным вопросам теории здесь уделено значительно меньше места, чем в гораздо меньшей по объему книжке Б. В. Гнеденко и А. Я. Хинчина «Элементарное введение в теорию вероятностей». В первую очередь здесь следует отметить отсутствие четкого разъяснения самого смысла понятия вероятности и полное игнорирование вопроса о статистических закономерностях в физике и естествознании; явно недостаточное место уделено обсуждению понятий зависимости и независимости и закона больших чисел; отсутствует даже и самый термин «случайная величина» (при наличии большого числа специальных статистических терминов).

Вряд ли можно сомневаться в том, что перед советской школой также очень скоро встанет вопрос о включении в курс математики элементов теории вероятностей; однако решать этот вопрос у нас, по-видимому, будут совсем иначе. При всём том широкое знакомство нашей педагогической и научной общественности со всем зарубежным опытом в этом отношении безусловно было бы весьма желательным; в частности, нам кажется, что перевод рецензируемой здесь книги на русский язык можно было бы только приветствовать.

«ТИПИЧНАЯ ЗАДАЧА» ПО ТЕОРИИ ВЕРОЯТНОСТЕЙ

На дне глубокого сосуда
Лежат спокойно эн шаров.
Попеременно их оттуда
Таскают двое дураков.

Занятье это им приятно,
Они таскают тэ минут,
И каждый шар они обратно,
Его исследовав, кладут.

Ввиду занятия такого
Как вероятность велика,
Что был один глупей другого
И что шаров он вынул ка?

Из фольклора
механико-математического факультета
Московского университета

«МАТЕМАТИК СДЕЛАЕТ ЭТО ЛУЧШЕ»

Известный польский математик Гуго Штейнгауз шутливо утверждает, что существует закон, который формулируется так: ***математик сделает это лучше.*** *А именно, если поручить двум людям, один из которых — математик, выполнение любой незнакомой им работы, то результат всегда будет следующим: математик сделает ее лучше.*

Журнал «Польша», 1960, № 10

ИНТЕРЕСНАЯ КНИГА ПО МЕТОДИКЕ МАТЕМАТИКИ

Ю. М. Гайдук

(Харьков)

André Fouché, La pédagogie des mathématiques, Paris, 1952, 134 стр.

Небольшая книга А. Фуше не является систематическим руководством по методике преподавания математики в школе. Скорее ее можно считать введением, знакомящим с общими методико-математическими принципами, и прежде всего изложением (выдержанным в чисто позитивной, но тем не менее живой манере) педагогического кредо автора. Книге предпослано, однако, предисловие Ж. Дефоржа — генерального инспектора народного образования Франции, и то одобрение, которое она находит у этого официального лица, придает ей значение документа, отражающего в известной степени характер постановки математического преподавания во французской средней школе. Отмеченное обстоятельство повышает интерес иностранного, в том числе советского, читателя к книге французского автора.

Содержание книги состоит из «Вступления» и двух частей: часть I — «Алгебра», часть II — «Геометрия».

Во «Вступлении» дана общая характеристика эволюции педагогики математики. Здесь, впрочем, нет никаких исторических подробностей (в частности, никаких имен); это — сжатый логический итог исторического развития, неизбежно схематичный, но в общем верный и облеченный в образную форму.

Автор обращает внимание читателей прежде всего на борьбу двух противоположных по духу методов преподавания — догматического и эвристического, поочередно сменявших друг друга на авансцене педагогических исканий.

Суровым пастырем, идущим во главе своего стада и безразличным к судьбе отставших (оставляемых в жертву «злому волку — невежеству»), изображает автор догматизм. Девиз этого метода — «сначала выучить, понять потом». От учащихся требуются неограниченная вера, строгое выполнение правил, запоминание наизусть теорем, автоматическая быстрота действий. Ошибка учащегося — его

непростительное преступление. Учитель — сверхчеловек, окруженный ореолом непогрешимости и всеведения.

В ином свете представлен педагог-эвристик. Это — благодушный неторопливый пастырь, замыкающий шествие пасомых и спокойный за их целость. «Сначала понять, а выучить потом» — таков девиз эвристиков. Всё обучение принимает характер отыскания истин самими учащимися. Ошибка — лишь легко исправимая случайность, приносящая даже пользу, когда она анализируется. Учитель — лишь человек, почти старший товарищ учащихся, способный иной раз и ошибиться, но всегда помнящий о том, что́ может затруднять и мучить ребенка.

В условиях массовой школы обнаружилась неэффективность каждого из этих двух «чистых» методов: слишком много жертв требует чистый догматизм, пренебрегающий природой детского восприятия, в слишком большом времени — далеко превосходящем нормы учебных планов — нуждается чистая эвристика. Поэтому в практике школьного преподавания всё сильнее укрепляется стремление к гармоническому сочетанию обоих традиционных методов: пастырь — учитель должен занять место где-то посреди стада, чтобы успешнее руководить движением вперед каждой его части. Но это предполагает активную психологическую установку во всей деятельности учителя, и французский педагог призывает с особой серьезностью отнестись к психологическому аспекту методики математики. Дальнейшие замечания автора показывают, что, пропагандируя приведенный сейчас тезис, он не впадает, однако, в тот вульгарный психологизм, который, гипертрофируя один из аспектов методики математики, игнорирует в то же время другой ее аспект — научно-математический. Напротив, Фуше подчеркивает необходимость полного учета обоих аспектов, учета существующих между ними глубоких и многосторонних связей:

> «Новейшие исследования, проведенные в некоторых классах, связанные с предоставлением человеческой природе большей свободы проявления, выяснили, что не только преподавание нуждается в переменах. Следовало бы кое в чем пересмотреть и логическую структуру самой математики. Повышение эффективности преподавания — это вопрос не только психологического порядка, как думали прежде, но и проблема из области чистой логики» [1]).

Как уже отмечалось, главной проблемой математической дидактики, по Фуше, должно стать отыскание правильной комбинации эвристического и догматического методов. Намеченное самим Фуше

[1]) К сожалению, Фуше не приводит здесь примеров. Со своей стороны, мы могли бы указать на исследования по эмпирической (или «натуральной») геометрии известного датского математика Ельмслева как например интересной математической разработки «психологического заказа», связанного с проблемой начального преподавания геометрии.

решение этой проблемы предусматривает преимущественное применение эвристического метода в ряде наиболее важных моментов курса (в начале изучения каждой новой темы геометрии и в некоторых узловых вопросах алгебры), использование чистого догматизма при изучении ряда правил вычислений в арифметике и обращение к промежуточному методу (когда устанавливаемое положение сначала формулируется, а затем обосновывается) в других частях курса. Автор отмечает при этом, что комбинирование различных методов должно быть гибким, и многое остается предоставленным педагогическому искусству учителя. Возможно полнее должны учитываться индивидуальные особенности учащихся: детям с созерцательным складом ума, пассивным, индифферентным нужно рекомендовать усиленные дозы эвристики, догматизм же следует использовать как лекарство для детей с гипертрофированной, болезненной активностью.

В конце «Вступления» автор формулирует главную цель своего сочинения — проанализировать мотивы и причины господствующей среди учащихся боязни математики и указать педагогические средства, найденные для преодоления этой боязни. Автор решительно отвергает тезис о прирожденном характере математических способностей:

> «...Никакой математической одаренности не существует; эта загадочная и мифическая способность — лишь предрассудок, приводящий в уныние детей, лишающий их веры в свои силы».

Возникновение у учащихся неприязни к математике и последующую неспособность их к ней Фуше объясняет рядом психологических и физиологических причин; по его мнению, эта математическая идиосинкразия вполне доступна лечению и профилактике.

Оставив «Вступление», обратимся теперь к обзору основного материала книги. Вследствие скромного объема сочинения автор ограничился почти исключительно начальной алгеброй и элементами планиметрии; правда, затронуты и некоторые другие вопросы, но лишь постольку, поскольку в отношении этих вопросов автор нашел нужным сделать специальные замечания.

Первая, алгебраическая часть книги охватывает в пяти относящихся сюда главах вопросы арифметического и алгебраического счета, развитие понятия числа (дробные и относительные [1]) числа), элементы учения о функциональной зависимости (функции, производные, идея бесконечности).

Основную трудность при усвоении учащимися алгебры автор видит в том, что им обыкновенно остается неясной «логика целого» алгебраической системы [2]). Эта неясность подвергает тяжелому

[1]) По французской школьной терминологии — «алгебраические».

[2]) Как перекликаются с этим замечанием А. Фуше воспоминания Ч. Дарвина об изучении им в юношеские годы математики! Дарвин с увлечением штудировал «Начала» Евклида, но испытывал, уже будучи студентом Кэм-

испытанию здравый смысл учащихся и препятствует приобретению ими прочных навыков. Для искоренения зла автор полагает необходимым прежде всего внести единство в понимание смысла алгебраических знаков, устранив традиционное противопоставление друг другу роли знаков, обозначающих числа, и роли знаков, обозначающих действия, так как такое противопоставление «искажает идею числа в сознании учащегося, делая из чисел инертный материал, отданный во власть знаков действий, тогда как число... является самым первым из всех операторов». Высказавшись таким образом в пользу операторной трактовки понятия числа в школе, автор с большой последовательностью проводит ее в своем дальнейшем изложении.

Из других особенностей методики Фуше отметим ту большую роль, которую он отводит в начале курса алгебры упражнениям в переводе с алгебраического языка на словесный (и наоборот).

Остановимся на методических рекомендациях автора, относящихся к изучению функциональной зависимости. Он подчеркивает значение этой части курса, равно как и тесно связанного с ней учения об уравнениях, для воспитания математической активности учащихся. С общим понятием функции учащихся очень полезно ознакомить (на примерах из физики) сейчас же после ознакомления с тождественными преобразованиями.

> «Здесь учащиеся встречаются с новым мыслительным процессом сильно затрудняющим их, так как им приходится отваживаться на свободный выбор значений переменной. Их пугает это внезапное отсутствие ограничений».

Большое внимание уделено методике изучения линейных функций. Автор справедливо настаивает на том, чтобы характерное свойство этих функций — пропорциональность приращения [1]) функции приращению независимой переменной — было возможно быстрее усвоено учащимися. В качестве полезного применения линейных функций указывается общее уравнение равномерного движения.

Представляют большой интерес замечания (к сожалению, слишком беглые) об изучении понятия производной. По мнению А. Фуше, этому понятию должно предшествовать понятие дифференциала как главной части приращения (производная вводится затем как «дифференциальный коэффициент»). Автор рекомендует записать приращение Δy как функцию приращения Δx и полученное выражение

бриджского университета, «отвращение» к алгебре, в первых начатках которой он, благодаря «тупому учителю», не мог «найти никакого смысла». (См. Автобиографию Ч. Дарвина в 1-м томе иллюстрированного собрания его сочинений под ред. К. А. Тимирязева, Москва, 1907).

[1]) Во французской терминологии Δy называется «изменением» (variation) переменной y. Было бы целесообразно перейти на этот термин и у нас, отказавшись от термина «приращение», внушающего учащимся мысль об обязательном возрастании переменной, получающей «приращение».

преобразовать таким образом, чтобы выделить из него линейную часть.

> «Тот факт, что затем приращение заставляют стремиться к нулю, совершенно второстепенен, и для средних учащихся лучше предпочесть пользоваться очень малыми приращениями, как это делают физики».

Намеченная французским автором методика изучения производных отличается, как видим, подчеркнуто алгебраическим характером и имеет очевидной целью избежать, насколько это возможно, привлечения понятия предела. Разумеется, строгое проведение такого подхода возможно, лишь если ограничиться рассмотрением простейших функций (рациональных), изучаемых в начальной алгебре.

Не останавливаясь на других методических рекомендациях, носящих частный характер, упомянем еще лишь об одном принципиальном замечании. Речь идет о различении автором двух типов доказательств в учебном курсе — «логических» и «технических». К первым отнесены доказательства, использующие только те предложения, которые приняты в данном курсе за аксиомы. Ко вторым — доказательства, в которых используются предложения, выведенные из аксиом. Фуше подчеркивает бóльшую убедительность для учащихся доказательств «логического» типа и предлагает по возможности шире пользоваться ими (конечно, при условии, что они не окажутся слишком громоздкими).

Перейдем ко второй, геометрической части рецензируемой книги. В семи главах этой части получил отражение следующий круг вопросов: эволюция и структура геометрии, доказательства свойств геометрических фигур, задачи на геометрические места, конструктивные задачи, общие замечания о курсе геометрии, геометрия и алгебра (учение о подобии), геометрия в пространстве.

Автор начинает с выяснения специфики преподавания геометрии. В отличие от алгебры, в геометрии абстрактное всегда тесно связано с конкретным. Поэтому, как бы ни был сложен ход геометрической дедукции, «адаптация» к ее результату никогда не составляет трудности. В геометрии затруднения связаны не с приобретением новых навыков, а, наоборот, с необходимостью очистить воображение от всех предубеждений, научиться хорошо видеть.

Одна из главных задач учителя, по мнению Фуше, состоит в том, чтобы отучить учащихся смотреть на геометрию как на собрание «готовых», сухих доказательств, подлежащих заучиванию: нужно побудить учащихся самих создавать «свою» геометрию. Для этого, пишет автор, следует «убедить учащихся с самого начала, что, обладая терпением, воображением и искренним желанием, каждый из них сможет добиться успеха в этом деле». Важно разуверить учащихся в том, будто для успешных занятий по геометрии нужно заранее обладать «геометрической интуицией» — этой загадочной способностью, присущей математикам.

«Интуиция придет к ним позже, после их первых удач, ибо она — их почти непосредственный результат, но не их причина. Каждый может развить в себе интуицию, в то время как только боязнь ошибиться парализует нормальное развитие».

Главной формой работы учащихся по курсу геометрии Фуше делает «нахождение и решение задач». Теоремы, необходимые для решения задач, должны быть также «открыты» учащимися. При этом целесообразно отказаться сначала от доказательства теорем, являющихся для учащихся очевидными,— к таким теоремам лучше вернуться в конце курса, когда учащиеся, с уже развитым до некоторой степени логико-дедуктивным мышлением, сумеют лучше оценить смысл их доказательства.

Для выяснения педагогической роли геометрических задач и для развития у учащихся навыков их решения Фуше предлагает различать четыре типа задач. К первому типу он относит задачи, в которых требуется «открыть все свойства данной фигуры». Автор отмечает, что этот тип задач, где речь идет о «чистом открытии», мало любим учащимися, так как, не зная определенной цели, они не видят в процессе работы конца своим усилиям. Второй тип образуют задачи на доказательство: доказать, что некоторая фигура, обладающая данными свойствами, обладает также другим определенным свойством.

«Здесь имеется граница для усилий, но ее существование создает нервное напряжение, делающее эти усилия неестественными. Спеша достичь цели, многие учащиеся ведут себя подобно животному, попавшему в ловушку: стремятся скорее угадать, чем найти решение».

К третьему типу принадлежат задачи, в которых требуется построить фигуру, имеющую некоторое свойство. Это — задачи на «чистое исследование»; здесь цель вполне определена, но средства ее достижения не указаны. Наконец, к четвертому типу Фуше относит задачи о геометрических местах (выраженные в форме: «какой будет фигура, все элементы которой обладают данным свойством»); он отмечает, что эти задачи с совершенно неизвестным искомым объектом учащиеся чаще, чем любые другие, пытаются решать простым угадыванием.

Анализ указанных четырех типов задач приводит автора к формулировке «первого рабочего принципа» его методики геометрии «*в геометрии нужно не угадывать, а открывать*».

Для каждого из типов задач Фуше предлагает особую схему действий для нахождения решения задачи. Приведены также характерные примеры решения задач второго, третьего и четвертого типов.

Не вдаваясь здесь в подробный разбор этих схем, отметим лишь, что отправным пунктом решения задачи второго типа (после, разумеется, тщательного ознакомления с ее условием) ставится парадоксальное на первый взгляд требование — «забыть то, что нужно доказать». Цель этого требования — расковать мысль (и энергию) решающего, позволить ей прощупать все возможные пути, исходящие из данных задачи,— одним словом, направить процесс решения не столько по пути *поиска* (recherche), сколько по пути *открытия* (decouverte).

Много внимания уделено у Фуше такому моменту, как составление учащимися окончательной редакции решения задачи, причем подчеркнуто значение этой работы как с точки зрения формирования логического мышления, так и с точки зрения воспитания характера учащегося. Рекомендуется приступать к этому заключительному этапу не сразу после отыскания решения, а спустя известное (быть может, значительное) время, когда уже уляжется аффективное возбуждение, вызванное перипетиями процесса решения. Для развития критического (и, в частности, самокритического) чувства полезно практиковать также коллективное обсуждение учащимися сравнительного достоинства индивидуально найденных ими способов решения и коллективную же разработку окончательного варианта наиболее целесообразного решения.

После рассмотрения вопросов, относящихся к задачам, автор останавливается на методике дедукции (*в конце* курса!) очевидных теорем. Это сделано на примерах теоремы о вертикальных углах и теоремы о единственности перпендикуляра, опущенного на прямую из данной точки. Такого рода рассмотрения автор рекомендует в качестве «прекрасного введения в аксиоматику». Характерно, однако, что, оставаясь здесь на традиционной позиции французских методистов (восходящей к Бертрану и Дюамелю), он считает преждевременным знакомить учащихся [«по крайней мере второго [1]) класса лицеев и коллежей»] с «логической акробатикой, связанной с постулатом Евклида».

Заслуживают серьезного внимания рекомендации Фуше по вопросу о закреплении материала, пройденного в курсе геометрии. В основу этой работы французский педагог кладет систематизацию учащимися приобретенных ими сведений под общим, но отнюдь не жестким руководством учителя. Материальное воплощение такая работа должна получить в изготовлении учащимися таблиц, сводящих воедино свойства тех или других геометрических фигур, в составлении ими схем логической взаимозависимости теорем и т. п.

Из других замечаний в этой части книги остановимся еще только на тех, которые высказаны по поводу некоторых вопросов курса геометрии в пространстве [2]). Основную задачу при изучении этой

[1]) То есть предпоследнего.

[2]) Термины «планиметрия», «стереометрия» во французской терминологии всё более уходят в прошлое. Можно было бы лишь приветствовать аналогичный пересмотр русской школьной математической терминологии.

части курса геометрии автор усматривает в развитии у учащегося «воображения плотника» — умения видеть наряду с «видимыми» также «невидимые» элементы пространственной фигуры. Развивать это умение лучше всего на решении большого количества задач с пересекающимися плоскостями и прямыми. Не следует торопиться с переходом к изучению параллельности, дабы не совмещать с самого начала трудности пространственного представления с трудностями логического порядка. Хорошо упражнять учащихся в решении несложных стереометрических задач «в уме» — без чертежа [1]).

В заключение, говоря о книге в целом, мы считаем нужным отметить педагогический демократизм и оптимизм Фуше, его вдумчивую психологическую установку, его стремление к взаимному сближению между педагогикой математики и математической наукой наших дней. Нельзя не приветствовать основного принципа автора — его требования учить в школе *не только знанию, но и пониманию — активному творческому пониманию — математики* [2]).

Однако не всё в книге Фуше одинаково привлекательно. Так, здоровые в своей основе психологические установки французского педагога выступают порой в облачении терминологии, заимствованной из некоторых модных на Западе буржуазных философских и психологических теорий. Следует также сказать, что далеко не все методические рекомендации автора можно считать бесспорными. Наконец, ряд специальных указаний Фуше применим лишь в условиях специфической организации французской средней школы.

Как бы то ни было, «Педагогика математики» — произведение стимулирующей педагогической мысли; знакомство с ним принесло бы пользу и удовольствие также широкому кругу советских преподавателей математики.

[1]) Подобный прием в свое время пропагандировал М. В. Остроградский.
[2]) Этот принцип тесно сближает установки Фуше с передовыми педагогическими идеями выдающегося французского физика, борца за демократизацию французского народного просвещения Поля Ланжевена.

К СВЕДЕНИЮ АВТОРОВ

Государственное издательство физико-математической литературы доводит до сведения авторов, что шестым выпуском заканчивается издание первой серии сборников «Математическое просвещение».

Приступая к выпуску новой серии, Издательство обращает внимание авторов на то, что сборники «Математическое просвещение» не являются по своему характеру специальным периодическим изданием и не ставят перед собой задачи первичной публикации оригинальных научных результатов, как это делают математические журналы и сборники научных трудов высших учебных заведений и научно-исследовательских институтов. Непринятые редакцией сообщения и заметки, посвященные частным вопросам, будут возвращаться авторам без мотивировки. Во многих случаях возвращение присланных сообщений не будет означать непризнания их научных или методических достоинств, а будет лишь выражать право редакции отклонять материал, публикация которого под силу только регулярно выходящим математическим журналам или «Трудам», «Ученым запискам» и другим периодическим научным изданиям.

Будут приложены все усилия к тому, чтобы сделать и дальнейшие выпуски «Математического просвещения» как можно более интересными, полезными и доступными для читателей, изучающих, преподающих и применяющих математику, а также для широких кругов любителей математики.

СОДЕРЖАНИЕ [1])

ПАМЯТИ А. Я. ХИНЧИНА

Б. В. Гнеденко. Александр Яковлевич Хинчин 3
А. Я. Хинчин. О воспитательном эффекте уроков математики 7
А. Я. Хинчин. О так называемых «задачах на соображение» в курсе арифметики 29

I. ОБЗОРЫ И СТАТЬИ

Р. Л. Добрушин. Математические методы в лингвистике 37
И. М. Яглом. Комплексные числа и их применение в геометрии . . . 61
В. Г. Болтянский и В. А. Ефремович. Очерк основных идей топологии. (*Окончание.*) 107
Ю. М. Безбородов и В. Б. Орлов. Машина играет в шахматы . . 139

II. НАУЧНЫЕ СООБЩЕНИЯ

Ю. Д. Бураго. Задача о круге Юнга для сферы 165
А. Е. Гельман. Рекуррентный метод вычисления корней алгебраического уравнения 171
Я. С. Дубнов. Модель евклидовой геометрии на гиперболической плоскости 181
В. А. Залгаллер. Как выйти из леса? (Об одной задаче Беллмана) . 191
Д. Д. Ивлев. О двойных числах и их функциях 197
Краткие сообщения:
1. *Р. Г. Бинг и Н. Д. Казаринов* (США). О конечности числа отражений, переводящих плоский невыпуклый многоугольник в выпуклый 205
2. *А. Г. Дорфман.* Ромбокубооктаэдры и взаимные им многогранники 207
3. *И. А. Каждан и С. С. Масько.* Вариант понятия степени точки относительно окружности 208
4. *А. В. Кужель.* Исследование корней уравнения $x^n+px^k+q=0$ 210
5. *З. А. Скопец.* Проективное обобщение теоремы Симсона 211
6. *Ю. З. Телесин.* О представлении натуральных чисел в виде суммы значений квадратного двучлена 214

[1]) Материалы, помещенные на стр. 170, 196—204, 236—266, 296, 310—328, 354, 362, в содержание не включены.

III. НАУЧНО-МЕТОДИЧЕСКИЕ СООБЩЕНИЯ

(Опыт преподавания и педагогический эксперимент)

И. И. Жогин. О средних 217
И. П. Натансон. Кривые Пеано 227
Л. З. Румшиский. Вычислительный лабораторный практикум по курсу высшей математики во втузах 237
М. Я. Цыркин. К вопросу об определении тригонометрических функций с помощью функциональных уравнений 245
Краткие сообщения:
1. *В. А. Ефремович.* Наглядный вывод формулы Лобачевского для угла параллельности 255
2. *В. И. Левин.* О числе e 256
3. *Г. Мархасев.* Вычисление почти треугольных определителей . . . 257
4. *Е. Онофраш.* (Румыния). Одно обобщение теоремы об углах равнобедренного треугольника 260
5. К вопросу о решении уравнений методом итераций:
   а) *Л. М. Рыбаков.* Метод последовательного вычисления всех действительных корней уравнения 262
   б) *В. Л. Загускин.* Вычислительная схема для ускоренного вычисления всех корней уравнения 263
От редакции 265

IV. НАУЧНАЯ И ПЕДАГОГИЧЕСКАЯ ХРОНИКА

Я. И. Ривкинд. Кружок по математическому анализу в Гродненском педагогическом институте 267
Е. П. Розенбаум (США). Преподавание элементарной математики в США (*Перевод с английского М. П. Данилова*) 281
И. Я. Депман. Подготовка учителей математики для средней школы в Норвегии 297
В. Г. Болтянский и Э. Р. Розендорн. XXI школьная математическая олимпиада в Москве 301

Сообщения о школьных математических олимпиадах:
1. Молдавская ССР (*К. Спатар*) 311
2. Белгородская область (*Ю. К. Василенко*) 313
3. Гор. Львов (*А. С. Кованько*) 314

Новости математической науки:
1. П. Эрдёш, П. Туран и др. Некоторые новые результаты в области аддитивной теории чисел (*П. Эрдёш*, Венгрия) 315
2. Е. В. Бет, А. Тарский. Равносторонний треугольник — основное понятие геометрии в пространстве (*И. М. Яглом*) 319
3. Применение электронных счетных машин для отыскания совершенных чисел (*И. Я. Депман*) 324

V. ЗАДАЧИ

Под редакцией И. М. Яглома

Задачи 329
Решения задач 337

VI. МАТЕМАТИЧЕСКАЯ ЛИТЕРАТУРА

А. М. Яглом. Экспериментальный учебник теории вероятностей и математической статистики для американских средних школ 355
Ю. М. Гайдук. Интересная книга по методике математики 363

К сведению авторов 371

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
«ФИЗМАТГИЗ»

Москва, В-71, Ленинский проспект, 15

ВЫШЛИ ИЗ ПЕЧАТИ

Э. Борель, Вероятность и достоверность.
Н. Я. Виленкин, Метод последовательных приближений (серия «Популярные лекции по математике»).
М. Г. Гаазе-Рапопорт, Автоматы и живые организмы.
А. П. Доморяд, Математические игры и развлечения.
Дж. Мак Кинси, Введение в теорию игр.
А. И. Маркушевич, Ряды.
Б. А. Трахтенброт, Алгоритмы и машинное решение задач.
А. Тьюринг, Может ли машина мыслить.
А. Я. Хинчин, Цепные дроби.

Книги продаются в книжных магазинах, а также высылаются наложенным платежом республиканскими, краевыми и областными отделениями «Книга — почтой».

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
«ФИЗМАТГИЗ»

Москва, В-71, Ленинский проспект, 15

ГОТОВЯТСЯ К ВЫПУСКУ

Я. С. Дубнов, Измерение отрезков (серия «Библиотека „Математического просвещения“»).

А. И. Жуков, Введение в теорию относительности.

Дж. Литлвуд, Математическая смесь.

В. Литцман, Где ошибка?

О. Оре, Замечательный математик Нильс Хенрик Абель.

Л. Хенкин, О математической индукции (серия «Библиотека „Математического просвещения“»).

И. М. Яглом, Комплексные числа (серия «Библиотека „Математического просвещения“»).

Предварительные заказы на эти книги принимают магазины Книготорга. Оформив заказ на почтовой открытке лично в магазине, Вы получите извещение о поступлении книги в магазин. В случае отказа в приеме предварительного заказа просим сообщить Всесоюзному объединению книжной торговли по адресу: Москва, Ленинский проспект, 15.

В СЛЕДУЮЩЕМ (СЕДЬМОМ) ВЫПУСКЕ «МАТЕМАТИЧЕСКОГО ПРОСВЕЩЕНИЯ» СРЕДИ ДРУГИХ МАТЕРИАЛОВ БУДУТ НАПЕЧАТАНЫ:

Доклад Д. Гильберта на Международном математическом конгрессе 1900 года в Париже (с предисловием акад. П. С. Александрова) и текст четырех проблем Гильберта с комментариями.

Статья польского академика В. Серпинского «О некоторых нерешенных задачах арифметики».