Version released: July 18, 2013

Латвийский университет
Институт математики и информатики

К. М. Подниекс

ВОКРУГ ТЕОРЕМЫ ГЕДЕЛЯ

Рига "Зинатне" 1992

УДК 164 519.9

П о д н и е к с К. М. Вокруг теоремы Геделя. – Рига: Зинатне, 1992. – 191 с. - ISBN 5-7966-0928-9.

Проведен методологический анализ природы математики. Показано, что сущность математического метода состоит в исследовании застывших моделей. Обоснована несостоятельность утверждений об ограниченности аксиоматического метода. Предлагается следующая методологическая оценка теоремы Геделя о неполноте:

Всякая формальная теория с методологической точки зрения является моделью некоторой **застывшей системы мышления**. С учетом этого основной вывод из теоремы о неполноте можно переформулировать так: всякая достаточно всеобъемлющая, но застывшая система мышления неизбежно оказывается несовершенной – в ней содержатся либо противоречия, либо проблемы, для решения которых данной (застывшей!) системы недостаточно. **Именно в строгом доказательстве принципиального несовершенства всякой застывшей системы мышления состоит подлинный диалектический смысл достижений Геделя**.

Изложены важнейшие результаты математической логики XX в., знание которых необходимо для понимания предлагаемой методологической концепции.

Библиогр. 33 назв.

Научный редактор канд.физ.-мат.наук В. К. Детловс

Р е ц е н з е н т ы:

проф., д-р физ.-мат.наук Р. В. Фрейвалд

проф., д-р физ.-мат.наук В. А. Успенский

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Является ли Ваша философия математики платонистской или нет, это можно определить с помощью следующего теста. Рассмотрим последовательность простых чисел-близнецов:

(3,5), (5,7), (11,13), (17,19), (29,31), (41,43), ...

Гипотеза: существует бесконечно много пар близнецов. Это предположение не доказано (и не опровергнуто) до сих пор. Верите ли Вы, что несмотря ни на что гипотеза должна быть "объективно" истинной или ложной? Для обоснования своей веры Вы можете воспользоваться следующим рассуждением. Представим себе, что мы продвигаемся вперед вдоль последовательности натуральных чисел

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,...

и время от времени встречаем пары близнецов:

(3,5), (5,7), (11,13), (17,19), (29,31), (41,43),...

Существует ведь только две возможности: а) мы доходим до последней пары близнецов и больше их не встречаем (в этом случае гипотеза оказывается ложной), б) пары близнецов появляются все время (тогда гипотеза истинна).

Рассуждая таким образом, Вы демонстрируете свой платонизм. Вы привыкли оперировать натуральными числами так, как будто они составляют некий специфический "мир", который очень похож на мир повседневных вещей. Вы привыкли думать, что на практике любое достаточно определенное утверждение должно быть либо истинным, либо ложным. Поэтому Вы не в состоянии представить третью возможность: количество пар близнецов **не является ни конечным, ни бесконечным**. Однако такая возможность не будет нас удивлять, если мы осознаем, что система натуральных чисел содержит не только некоторую информацию о действительном мире, но и множество элементов **фантазии**. Почему Вы полагаете, что этот фантастический мир людям удалось "сфантазировать" так идеально правильно, что на вопрос о количестве близнецов обязательно будет существовать ответ?

Данная монография предназначена для математиков, физиков, философов (в том числе для студентов старших курсов этих специаль-

ностей) и всех интересующихся методологическими проблемами науки.

Философы и физики для знакомства с основными методологическими выводами могут ограничиться чтением разделов 1, 2.1, 2.4, 5.1, 5.3, 5.4, 6.1–6.3, пропуская непонятные математические подробности. Математики должны, вообще говоря, изучить весь материал, за исключением разделов 2.4 и 4.3–4.7, которые могут изучаться факультативно.

Издание может использоваться в качестве учебного пособия по курсу математической логики (как вторая его часть – после изучения исчисления высказываний и исчисления предикатов).

В разделе 4 изложено решение десятой проблемы Гильберта – одно из самых красивых рассуждений в истории математики.

Автор выражает признательность научному редактору В. К. Детловсу за множество предложений, способствовавших улучшению текста.

Октябрь 1991 г. К. П о д н и е к с

1. ПРИРОДА МАТЕМАТИКИ

1.1. Платонизм − философия работающих математиков

Французский математик Шарль Эрмит сказал как-то: он убежден в том, что числа и функции − это не изобретения математиков, они существуют независимо от нас, как существуют вещи реального мира. Было время, когда это высказывание цитировалось как свидетельство "стихийного материализма выдающихся ученых".

Однако такие высказывания математиков свидетельствуют совсем о другом − об их стихийном **платонизме**. Платонистское отношение математиков к объектам своих исследований обусловлено самой природой математического метода. Но, разумеется, при решении методологических вопросов такой философии придерживаться уже нельзя. Как трудно, однако, изменить привычки, обретенные в ходе повседневной работы, когда переходишь в сферу методологии...

Но сначала − о платонизме самого Платона (427−347 гг. до н.э.), который жил на закате "золотого века" Древней Греции. В 431−404 гг. до н.э. велась обескровившая Грецию Пелопоннесская война, а в 337 г. − через 10 лет после смерти Платона − Грецию покорила Македония. В своей конкретной форме философия Платона сформировалась под влиянием греческой математики.

Развитие греческой математики в VI−V вв. до н.э. привело к образованию математических объектов в современном смысле этого слова: представления о числах, точках, прямых и т.д. **стабилизировались** и тем самым **оторвались** от своего первоисточника − свойств и отношений объектов реального мира. "Математическая прямая не имеет ширины, а точка вообще не имеет размеров". Ничего в точности такого в реальном мире нет: вместо прямых встречаются более или менее гладкие полосы, а вместо точек − пятна различной формы и размеров. Однако без этого перехода к идеализированному (но зато стабильному, застывшему) миру точек, прямых и т.д. математические знания остались бы на уровне ремесла, так и не достигнув уровня науки. Только идеализация (упрощение, исключение второстепенных деталей) сделала возможным такой эффективный инструмент, как евклидова геометрия.

В свою очередь, понятие натурального числа (1, 2, 3, 4, ...) возникло в ходе оперирования совокупностями несливающихся предметов. Процесс становления этого понятия завершился по-существу уже в VI в. до н.э., когда во времена Пифагора были доказаны первые теоремы о системе натуральных чисел **в целом**, например, теорема о том, что простых чисел существует "больше любого наперед заданного количества". Ясно, что об эмпирической проверке таких утверждений речи быть не может. Но в то время понятие натурального числа уже оторвалось от своего реального источника − "количественных закономерностей совокупностей несливающихся предметов", и стало функционировать самостоятельно − как **модель**. Натуральный ряд чисел − это **идеализация** упомянутых количественных закономерностей. Человек абстрагировал его на основе практического опыта с небольшими совокупностями (1, 2, 3, 10, 100, 1000 и т.д. предметов). Для совокупностей гораздо больших (многие миллионы предметов) он предположил аналогичные закономерности и тем самым идеализировал (а может быть, как заметил П. К. Рашевский [1973], даже **исказил**) реальную ситуацию.

В самом деле, количество атомов в данном листе бумаги − четное или нечетное? С точки зрения традиционной арифметики оно "обязано" (в каждый момент времени) быть либо четным, либо нечетным. В действительности же лист бумаги никакого точного числа атомов не имеет (хотя бы из-за сотен тысяч ядерных реакций, происходящих каждую секунду под воздействием космических лучей). Кроме того, согласно новейшим космологическим теориям, полное число элементарных частиц во Вселенной значительно меньше 10^{200}. Как мы должны тогда относиться к утверждениям вроде "$10^{200}+1$ − нечетное число"? Очевидно, таким образом, что арифметика занимается не только практически полезными алгоритмами вычисления, но и вещами совершенно фантастическими, лишенными непосредственного реального смысла.

Разумеется, древние греки не могли видеть все это столь ясно. Рассуждая о количестве простых чисел, они думали, что обсуждают вещи столь же реальные, как те совокупности предметов, от которых понятие натурального ряда было абстрагировано.

Итак, первый в истории математики процесс идеализации закончился стабилизацией понятий о числах, точках, прямых и т.д. Эти понятия определились и надолго стали общепринятыми в обществе математиков. Этот момент наступил еще в V в. до н.э. Стабилизация понятий свидетельствует об их **отделении** от реальных объектов, обращение с которыми привело людей к выработке этих понятий. Ведь

застывшим может стать только понятие, уже **оторванное** от своих реальных прообразов, продолжающих самостоятельную жизнь и содержащих огромное разнообразие второстепенных и изменяющихся нюансов. Работая в области геометрии, математик исследует не непосредственно отношения реальных объектов, а свое сложившееся (**застывшее**) представление о них − идеализированный "мир" точек, прямых и т.д. Если бы он во время своих размышлений постоянно вспоминал об особенностях реальных вещей (о степени их гладкости и т.п.), то вместо науки (общих, эффективных и далеко идущих геометрических методов) мы имели бы только простейшие, специфические алгоритмы, найденные путем проб и ошибок или с помощью элементарной интуиции. Именно на таком уровне остановилось развитие математики Древнего Востока.

Для греческих философов появление математического метода было новостью: исследовать не непосредственно природу, а какое-то застывшее представление о ней, субъективно воспринимаемое в процессе исследования как "последняя" реальность, дальше которой ничего нет. Рождение математического метода разные философы отметили по-разному (но, разумеется, никто из них не сумел тогда дать правильную оценку такого сложного явления). Платон, изучая математику, пришел к весьма оригинальному мировоззрению, согласно которому существует два мира: мир идей (идеально строгий и точный, упорядоченный и гармоничный − как мир геометрических образов) и мир вещей (несовершенный, "размытый", хаотический). Каждая реальная вещь представлялась Платону несовершенной, приблизительной реализацией своей "идеи" (которая существует независимо от самой вещи в мире идей). Характерно также остроумное, но совершенно фантастическое представление Платона о природе математического исследования: перед рождением человека его душа обитает в мире идей, а во время своей земной жизни, занимаясь математикой, она постепенно вспоминает опыт, обретенный в мире идей. Разумеется, это перевернутое вверх ногами представление о действительной природе математического метода. Конечный результат развития математических понятий − застывшая система идеализированных объектов − принимается Платоном за исходную позицию, вокруг которой "танцуют" вещи реального мира. Платон старался по-своему объяснить стороны процесса познания, которые были недоступны философии его времени из-за недостаточной конкретно-научной базы. В данном случае речь шла об объяснении природы идеализированных математических объектов. Для правильного ее объяснения греческая наука не имела достаточной базы в таких областях, как физика, биология, физиология и психология.

Сегодня мы называем платонизмом любую философскую

позицию, которая какую-либо систему идеальных объектов человеческой мысли трактует как особый, независимо существующий мир. Именно такой оказывается и философия "работающих" математиков, которые, как правило, не задумываются о "природе" своей деятельности.

Платонистское отношение к объектам своего исследования неизбежно для математика: в своей повседневной работе он оперирует числами, функциями, точками, прямыми и т.д. как объектами, составляющими некоторое подобие мира, как "последней реальностью", за которой нет никакой другой, "более подлинной" реальности. Да и как иначе можно глубоко исследовать систему понятий, которая стабилизировалась и оторвалась от своего первоисточника? Именно в платонизме работающих математиков − одна из "тайн" творческой силы математики!

Это объясняет и неизбежность элементов платонизма в мировоззрении математиков, как правило, не очень искушенных в философии. Привычки, обретенные в повседневной работе, обладают огромной силой. Когда математик, не имеющий достаточной философской подготовки, берется за решение методологических вопросов, за объяснение природы своих результатов, он невольно привносит в рассуждения элементы платонизма. Этим страдали и страдают в равной мере как рядовые, так и великие математики. Заявления математиков об объективном характере своих результатов − как правило, не материализм, а платонизм! Самые выдающиеся среди немногих исключений из этого правила – А. Н. Колмогоров и В. М. Глушков.

Правда, платоник в некотором смысле "лучше" субъективного идеалиста, утверждающего, что математические объекты − произвольные творения человеческого ума. Следует, однако, различать людей, которые просто объявляют свои построения объективными, существующими независимо от нас, людей, и материалистов, которые пытаются объяснить происхождение математических понятий и определить закономерности их развития.

Является ли Ваша собственная философия математики платонистской или нет, это легко определить с помощью следующего теста. Рассмотрим последовательность простых чисел-близнецов:

(3,5), (5,7), (11,13), (17,19), (29,31), (41,43), ...

(простые числа принято называть близнецами, если их разность равна 2). Гипотеза: существует бесконечно много пар близнецов. Это предположение не доказано (и не опровергнуто) до сих пор. Верите ли Вы, что несмотря ни на что, гипотеза должна быть "объективно"

истинной или ложной? Для обоснования своей веры Вы можете воспользоваться следующим рассуждением. Представим себе, что мы продвигаемся вперед вдоль последовательности натуральных чисел

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, ...

и время от времени встречаем пары близнецов:

(3,5), (5,7), (11,13), (17,19), (29,31), (41,43), ...

Существует ведь только две возможности: а) мы доходим до последней пары близнецов и больше их не встречаем (в этом случае гипотеза оказывается ложной), б) пары близнецов появляются все время (тогда гипотеза истинна).

Рассуждая таким образом, Вы демонстрируете свой платонизм. Вы привыкли оперировать натуральными числами так, как будто они составляют некий специфический мир, который очень похож на мир повседневных вещей. Вы привыкли думать, что на практике любое достаточно определенное утверждение должно быть либо истинным, либо ложным. Поэтому Вы и не в состоянии представить третью возможность: количество пар близнецов не является ни конечным, ни бесконечным. Однако такая возможность не будет нас удивлять, если мы вспомним, следуя П. К. Рашевскому, что система натуральных чисел содержит не только некоторую информацию о действительном мире, но и множество элементов фантазии. Почему Вы полагаете, что этот фантастический мир людям удалось "сфантазировать" так идеально правильно, что на вопрос о количестве близнецов обязательно будет существовать ответ?

(Другая иллюстрация платонистского подхода к методологическим вопросам математики − высказывание Н. Н. Лузина о континуум-проблеме в разделе 2.4.)

И Ваш платонизм, и платонизм Н. Н. Лузина − это нормальный платонизм работающего математика, стимулирующий занятие проблемами любой сложности − ведь заранее никогда неизвестно, разрешима проблема или нет.

Однако, переходя к решению **методологических** вопросов, уже нельзя давать волю платонистским привычкам (полагая, что несмотря на неразрешимость проблемы близнецов "для нас, людей", их количество "объективно" является либо конечным, либо бесконечным). Это означает допускать существование мира идей (мира чисел), не зависящего от аксиом, используемых в рассуждениях математиков. Тогда платонизм математический превращается в платонизм философский. Такие люди утверждают, что традиционные аксиомы не передают адекватно все богатство содержательной математики, что надо искать более адекватные

аксиомы, и даже − что никакая фиксированная система аксиом не в состоянии представить богатство математики полностью. Это погоня за миражами − никакого подлинного мира математики, не зависящего от аксиом, с помощью которых он исследуется, разумеется, не существует. Правильная же оценка ситуации состоит в следующем.

Если обнаружено, что традиционные аксиомы математики не позволяют решить какую-либо проблему, то это свидетельствует **о внутреннем несовершенстве** данных аксиом (а не об их неадекватности какому-то "миру"). Возможно, следует заняться совершенствованием аксиом. И всегда оказывается, что вариантов развития, как правило, несколько. Например, можно принять так называемую аксиому конструктивности или противоречащую ей аксиому детерминированности (см. раздел 2.4). Так как эти варианты противоречат друг другу, то о приближении к **единственному** "подлинному миру математики" здесь не может быть и речи.

Наш главный вывод состоит в следующем: хотя повседневная работа математиков постоянно толкает их в платонизм (и как творческий метод этот платонизм весьма эффективен), при решении методологических вопросов от него следует сознательно отказываться. Игнорирование этой проблемы − основной недостаток многих философских сочинений, посвященных математике.

1.2. Исследование застывших моделей − сущность математического метода

Термин "модель" используется ниже в смысле, принятом в прикладной математике, а не в логике (т.е. мы будем обсуждать модели природных процессов и технических устройств, а не модели множеств формул).

Что характерно для математического подхода к решению какой-либо (физической, технической и т.п.) проблемы? Характерно прежде всего стремление как можно скорее "покончить с реальностью", перейти к исследованию определенной (фиксированной) математической модели. Поэтому в процессе формулирования задачи часто задаются вопросы: можно ли предположить, что данная зависимость линейна, можно ли пренебречь такими-то возмущениями, можно ли считать данное распределение вероятностей равномерным (нормальным или пуассоновским) и т.д., и т.п. Во всем этом видно стремление скорее и с использованием по возможности меньшего числа исходных принципов сформулировать математическую задачу, решение которой несмотря на

сделанные упрощения дало бы какое-то решение исходной проблемы.

Математики приучают себя к жизни (именно к жизни!) в мире математических понятий, а в отдельные периоды (пока идет разработка конкретной проблемы) − даже в узко-специальном мире определенной модели. После того как модель создана, для математиков ее исследование становится самоцелью. В процессе работы они отвлекаются от отражающего аспекта модели, совершенно игнорируют его. Именно в этом − причина платонистского отношения математиков к объектам своих исследований. Именно в этом − источник творческой силы математики, источник "непостижимой эффективности математики в естествознании и технике" (Е. Вигнер). Благодаря такому подходу математики умеют **извлекать максимум следствий из минимума посылок**. Именно разработанность математических моделей (наличие готовых алгоритмов, общих методов) делает их применение таким эффективным. Ведь модель, если ее единственное достоинство − адекватность оригиналу, сама по себе бесполезна, если нет методов и алгоритмов, позволяющих в реальное время вывести заключения, дающие новые знания об оригинале. И ключ к этой разработанности − умение математиков (буквально) жить в мире разрабатываемой модели, забывая обо всем другом. Это даже создает для некоторых из них репутацию сухарей, отшельников и чудаков.

Таким образом, платонизм является фактически **психологией** работающих математиков, и философией он оказывается только с их собственной субъективной точки зрения.

С появлением математики все научные теории следовало бы делить на два класса:

а) теории с развивающейся системой принципов,

б) теории с застывшей системой принципов.

Теории класса а) в ходе своего развития обогащаются новыми принципами, которые нельзя обосновать ранее принятыми. Появлению таких принципов мы обязаны фантазии специалистов, которые опираются на все более совершенную экспериментальную базу. Прогресс теории состоит здесь прежде всего в этом процессе обогащения.

С другой стороны, в математике, физике, отчасти в химии и совсем редко в других науках встречаются теории, принципиальная основа которых со временем не меняется, а если и меняется − это изменение квалифицируется как **переход к новой теории**. Так, на теорию относительности А. Эйнштейна можно смотреть как на уточнение классической механики И. Ньютона, как на дальнейшее развитие той же ньютоновской теории. Но поскольку обе теории очень точно определены,

то на переход "от Ньютона к Эйнштейну" можно смотреть и как на переход к другой теории. Развитие этих теорий продолжается по сей день: доказываются новые теоремы, изобретаются новые методы расчетов и т.д. Однако принципиальная основа (исходные постулаты) каждой из них остается неизменной (такой, какой она была при жизни их создателей). Только те положения признаются относящимися к данной теории, которые можно вывести из (давно известных) ее основных принципов. Все, что выходит за рамки этих принципов, относится уже к другой теории.

Застывшая система основных принципов − отличительная особенность всякой математической теории. Математическая модель какого-либо явления природы или технического устройства − это непременно **застывшая модель**, сближению которой с оригиналом положен предел. Только такую модель может исследовать математик. Всякая попытка уточнить модель (видоизменить ее определение с целью еще больше приблизить к оригиналу) приводит к новой модели, которая опять должна "застыть", чтобы ею мог заниматься математик. Формирование математических моделей, их уточнение − это не собственно математическая деятельность, она относится к той отрасли науки или техники, которая заинтересована в конечном результате исследования.

Сформулированная выше концепция сущности математики не является общепринятой и, как правило, воспринимается с трудом. Что же препятствует восприятию математических теорий как застывших? "Во-первых, математические теории почти никогда не рассматриваются изолированно одна от другой.... теория множеств одновременно с ее возникновением начала применяться к изучению геометрических объектов (собственно, для этого она и была создана). Чем продуктивнее, чем ближе к практике математическая теория, тем сильнее проявляется эта тенденция. Во-вторых, внутри любой теории ее теоремы состоят, как правило, из двух частей: условия и заключения. Заключение теоремы является, таким образом, следствием не только застывшей совокупности аксиом, но и конкретного, специфического для данной теоремы условия. А что такое условие, как не расширение застывшей системы принципов? В-третьих, любая математическая теория открыта для пополнения новыми **понятиями**. Так, в анализе вслед за понятием непрерывности функции вводятся: понятие точки разрыва, классификация таких точек, понятие функции, непрерывной на отрезке,..., равномерной непрерывности, условие Лифшица,... и т.д. Исследуются свойства каждого нового понятия, и эти свойства постепенно оттесняют далеко на задний план исходную совокупность аксиом.... Все это нисколько не противоречит тезису о неизменности исходной системы принципов

(аксиом и правил вывода), но препятствует восприятию математических теорий как "застывших" работающими математиками." (из письма С. С. Лаврова, 1988 г.)

Итак, математический метод − исследование застывших моделей. Очень важно, что математическая модель (именно потому, что она застывшая) уже не привязана жестко к оригиналу. Может оказаться, что модель была выбрана неудачно (плохо отвечает оригиналу), однако это не препятствует ее исследованию − ведь застывшая модель точно определена. Можно сказать поэтому, что математическая модель "не нуждается" в оригинале, она не обязательно является моделью чего-то, она − модель "сама по себе". Такую модель можно видоизменить (получив новую), руководствуясь уже не интересами соответствия оригиналу, а просто ради эксперимента или исходя из эстетических соображений. Так легко получаются "модели сами по себе", не имеющие реальных оригиналов. Застывший характер математических моделей делает это явление возможным и даже неизбежным.

Если математический метод − исследование застывших моделей, то чем является в таком случае сама математика? Модели могут быть более или менее общими (сравним, например, школьную арифметику, теорию относительности и какую-либо конкретную модель Солнечной системы). Частные модели лучше всего исследовать под руководством специалистов, которые эти модели строят и используют. Сочетание специальной подготовки с достаточной математической подготовкой (в одном человеке или в коллективе) будет здесь наиболее эффективным. Исследование же моделей, которые представляют более общий (или даже всеобщий) интерес и имеют широкую область применения (применимы в исследовании целого ряда более специальных моделей), составляет содержание особой науки, которую принято называть **математикой**. Так, своей широкой применимостью в различных областях науки примечателен математический анализ (дифференциальное и интегральное исчисление). Это типичный пример модели (теории), которая относится к математике. С другой стороны, конкретная модель Солнечной системы (используемая, в частности для точного предсказания солнечных затмений) является слишком специальной, чтобы относить ее к математике (хотя это и математическая модель).

Застывший характер математических моделей и теорий составляет как силу, так и слабость математики. Извлечь максимум информации из минимума посылок − это умение математиков многократно доказало свою эффективность в науке и технике. Однако обратной стороной такой силы является слабость: никакая конкретная застывшая модель (теория) не в состоянии решить все проблемы, возникающие в науке (или даже только в математике). Этот диалектический тезис блестяще подтвердился

в знаменитой теореме Геделя о неполноте.

И еще одна слабость: математика, оторвавшись от действительных проблем, управляемая только своими "внутренними потребностями", на наших глазах расплывается и разбухает... Создаются теории и целые отрасли математики, которые еще долго не будут (а может быть, принципиально не могут) применяться к исследованию реальных проблем. Как пошутил польский писатель С. Лем в своей книге "Сумма технологий", математик − сумасшедший портной, который шьет "всевозможные одежды", надеясь сшить и кое-что пригодное для одевания. Как мы видели, эти отрицательные явления − неизбежное следствие самой природы математического метода.

1.3. Интуиция и аксиоматизация

Застывший характер математических моделей и теорий не всегда бросается в глаза − мешает платонистская привычка смотреть на объекты математики как на особый "не зависящий от нас мир", который "нами только изучается".

Мало кто будет оспаривать застывший характер теории, которая полностью аксиоматизирована. В аксиомах такой теории выражены все принципы рассуждения, которые в ней допускаются. Тем самым принципиальная основа теории зафиксирована и всякое ее изменение выражается в явных изменениях аксиом.

Каким образом, однако, можно считать застывшими теории, которые не аксиоматизированы? Так, все математики сходятся во мнении относительно того, какие рассуждения о свойствах натуральных чисел следует признать доказательными, а какие приводят только к гипотезам или ошибкам. И это − несмотря на то, что большинство математиков не знают ничего о каких-либо аксиомах арифметики! И даже в случаях, когда теория вроде бы построена на аксиомах (как, например, геометрия в "Началах" Евклида), в ее рассуждениях могут быть обнаружены моменты, не вызывающие разногласий относительно их справедливости, но в аксиомах тем не менее не отраженные. Например, различные свойства отношения "точка А лежит на прямой между точками В и С" используются у Евклида без всякого обоснования. Только в XIX в. М. Паш ввел "аксиомы порядка", характеризующие это отношение. Тем не менее и раньше все математики рассуждали о нем одинаково, не сознавая, как это у них получается.

Пытаясь объяснить это явление, приходим к понятию **интуиции**. Обычно оно почему-то связывается с творческим мышлением, с

"непосредственным постижением истины" и т.п. Но здесь нас интересует гораздо более прозаический аспект интуиции.

Человеческий мозг является настолько сложной системой связей и процессов, что нет никакой надежды, что "слабый свет сознания" будет держать под контролем все детали этого электрохимического фейерверка. Это заставляет признать, что кроме мыслительных процессов, которые осознаются (полностью или частично), в мозгу постоянно происходит масса процессов **бессознательного** мышления. Причем, как показывает опыт, если бессознательный процесс приводит к результатам, имеющим большое значение для данной личности, то результат иногда распознается сознанием. Однако сам процесс, приведший к результату, при этом может остаться скрытым от сознания (отсюда и впечатление "откровения", см. Ж. Адамар [1945], А. Пуанкаре [1908]).

Если существуют бессознательные процессы мышления, то должны существовать и неосознанные "разумные принципы", регулирующие это мышление (ведь оно приводит не только к беспорядочным сновидениям, но и к разумному решению реальных проблем). Такие принципы могут управлять и теми процессами мышления, которые частично осознаются.

В случае математических теорий мы как раз имеем дело с определенным комплексом бессознательных принципов, которые наряду с аксиомами (или даже совсем без них) регулируют наши рассуждения. Такие бессознательные регулирующие факторы, вырабатываемые в ходе интенсивных умственных занятий в определенной области, и следует **называть интуицией**. Можно говорить поэтому, что помимо явно сформулированных аксиом и правил вывода теория может быть зафиксирована и в особой интуиции. Можно говорить об интуиции натурального ряда, которая (без каких-либо аксиом) однозначно регулирует наши рассуждения о натуральных числах, или о "евклидовой интуиции", которая делает геометрию вполне определенной, хотя в аксиомах Евклида содержатся далеко не все предпосылки геометрических рассуждений.

Но как объяснить возникновение интуиций, одинаково управляющих рассуждениями стольких людей? По-видимому, решающим здесь является то, что люди − существа примерно одинаковые, что все они имеют дело с примерно одинаковым внешним миром и в процессе обучения, воспитания, практической и научной деятельности они стремятся к согласию между собой.

Со временем, когда исследования достигают определенного уровня сложности, постоянство (определенность) интуитивных моделей становится недостаточным. Тогда среди специалистов начинают

возникать разногласия: какие способы рассуждений допустимы, а какие − нет. Но даже если постоянства и достаточно, оно может оказаться "не того рода". Может оказаться, например, что допустимые (по общему мнению) способы рассуждения приводят к нелепым выводам. В истории математики подобные ситуации бывали: крах дискретной геометрической интуиции в результате открытия несоизмеримых отрезков (конец VI в. до н.э.), подозрительность к отрицательным и комплексным числам (до конца XVIII в.), спор Л. Эйлера и Ж. д'Аламбера относительно понятия функции (XVIII в.), плохо обоснованное обращение с расходящимися рядами (XVIII–XIX вв.), трудности восприятия теории множеств Кантора, парадоксы в этой теории (XIX в.), скандал с аксиомой выбора (начало XX в.). Все это − следствие неизбежной неконтролируемости бессознательных процессов. По-видимому, "принципы", регулирующие эти процессы, отбираются и укрепляются посредством своеобразного "естественного отбора на полезность". Однако мы знаем, что естественный отбор страдает "близорукостью", что он не способен на безошибочную далеко идущую координацию. Поэтому появление парадоксов (как действительных, так и мнимых) неудивительно.

Определяющая интуиция теории не всегда остается постоянной − особенно много изменений происходит на начальном этапе становления теории (когда интуиция, как и сама теория, еще не сложилась окончательно). На этом, самом деликатном, этапе между специалистами возникают особенно резкие разногласия (над новаторами издеваются и т.п.).

Надежный выход из положения может состоять только в переводе (хотя бы части) бессознательных принципов в сознательные (с последующим исследованием их согласованности). В буквальном смысле такой перевод невозможен: мы не можем знать внутреннюю дифференциацию факторов, составляющих интуицию. Поэтому речь может идти только о **реконструкции** этого "черного ящика" явными средствами.

Существует два метода такой реконструкции: генетический и аксиоматический. С помощью **генетического метода** пытаются моделировать интуицию средствами другой теории (которая сама может также быть интуитивной). Таким образом, "подозрительная" интуиция моделируется на базе более надежной интуиции. Этим способом удалось преодолеть подозрительность к комплексным числам, вводя их геометрическую интерпретацию (каждое комплексное число представляется точкой на плоскости, т.е. средствами евклидовой геометрии). В результате даже такие необычные свойства этих чисел, как бесконечное множество значений *log(x)* для отрицательного *x*, превратились в простые теоремы геометрического или топологического

характера. И споры по поводу всевозможных кажущихся парадоксов, связанных с комплексными числами, утратили почву. Там, где раньше могли ориентироваться только выдающиеся математики, теперь легко ориентируется любой школьник.

Аналогичное прояснение ситуации наступило с определением основных понятий математического анализа (предел последовательности, непрерывность и т.д.) в терминах "эпсилон-дельта". Однако оказалось, что некоторые из понятий, реконструированных в терминах "эпсилон-дельта", обладают **неожиданными свойствами**, которых у их интуитивного прообраза не было. Так, раньше полагали, что всякая непрерывная функция дифференцируема почти всюду, кроме отдельных исключительных точек "разлома". Однако если руководствоваться точным определением непрерывности, то оказывается, что можно построить непрерывные функции, не имеющие производной ни в одной точке (пример нигде не дифференцируемой функции К. Вейерштрасса). У А. Пуанкаре эти функции вызывали отвращение и он называл их язвой...

Появление у реконструированных понятий неожиданных свойств, во-первых, свидетельствует о том, что здесь происходит именно реконструкция (а не простое копирование интуитивных понятий), а во-вторых, это заставляет серьезно рассматривать вопрос об **адекватности реконструкций**.

Генетический метод "проясняет" одну интуицию средствами другой, т.е. действует относительно. **Аксиоматический метод**, напротив, действует "абсолютно" и состоит в следующем. Среди общепризнанных утверждений об объектах теории выделяются некоторые объявляемые аксиомами, т.е. "истинами", не требующими доказательства. После этого остальные утверждения теории уже требуется доказывать на основе аксиом. Доказательства могут содержать и интуитивные элементы, которые должны иметь более элементарный (более очевидный) характер по сравнению с тем, что выражено в аксиомах. Часто эти элементы сводятся к интуитивному использованию чисто логических средств рассуждения, арифметики целых чисел, математического анализа или теории множеств. Наиболее известные случаи, когда применялся аксиоматический метод: аксиомы Дж. Пеано для арифметики целых чисел, аксиомы Евклида, аксиомы Д. Гильберта для той же евклидовой геометрии, аксиомы Э. Цермело и А. Френкеля для теории множеств.

Аксиоматизация (так же как генетический метод) дает всего лишь **реконструкцию** интуитивных понятий. Проблема адекватности реконструкции здесь обычно сводится к вопросу: все ли существенные характеристики интуитивных понятий отражены в аксиомах или какая-то часть забыта? Более сложными являются случаи, когда аксиоматизация

применяется не просто для реконструкции существующей интуитивной теории "один к одному", а для спасения последней, когда та запуталась в парадоксах. Система аксиом Цермело-Френкеля для теории множеств была создана именно в такой ситуации: в интуитивной теории множеств Г. Кантора были обнаружены парадоксы и аксиоматизация явилась единственным выходом из положения. Проблема адекватности реконструкции здесь особенно сложна: сохранено ли все **положительное** содержание интуитивной теории?

В чем может состоять **критерий** адекватности реконструкции? Рассмотрим в качестве примера определение понятия действительного числа через рациональные числа. Такие определения были предложены в 1870-х гг. одновременно несколькими математиками (Р. Дедекиндом, Г. Кантором и др.). В результате многие подразумеваемые ранее свойства действительных чисел превратились в теоремы. Но почему мы считаем эти реконструкции удовлетворительными? Достаточно ли точно и полно передают они исходное интуитивное понятие действительного числа? Как обосновать точность и полноту реконструкции, если исходное понятие существует только в интуиции и всякое его выделение оттуда становится новой реконструкцией, адекватность которой опять нуждается в обосновании? Другого пути нет: мы должны руководствоваться только тем, как интуитивное понятие **проявляет себя в практике** математических рассуждений. Если все свойства действительных чисел, которые ранее считались очевидными и которые хотя бы раз фиксировались на бумаге, доказаны как теоремы (на основе нового, реконструированного понятия), если все теоремы математического анализа, доказанные ранее с использованием интуитивного понятия, передоказаны на основе реконструированного понятия, то те стороны интуитивного понятия действительного числа, которые успели проявить себя в математической практике, в реконструкции отражены.

Но, быть может, некоторые стороны интуитивного понятия еще не проявили себя, но могут проявить в будущем? Оспаривать такое предположение, казалось бы, очень трудно. В самом деле, допустим, что так оно и случится: явится через 100 лет математик X и докажет новую теорему математического анализа, используя свойство действительных чисел, которое ранее никто не использовал (т.е. оно себя в математической практике никак не проявляло). И тогда **все** сразу согласятся, что это неотъемлемое свойство действительных чисел? И что оно подразумевалось и 100 лет назад? Последнее во всяком случае уже нельзя будет проверить − никто из ныне живущих математиков до открытия X не доживет! Другими словами, предполагать, что в интуитивных математических понятиях скрыты какие-то аспекты,

которые очень долго не проявляют себя на практике ("на бумаге"), − это все тот же математический платонизм, считающий мир математических объектов существующим независимо от рассуждений математиков.

В ряде случаев дополнительным аргументом в пользу совпадения интуитивных понятий и их реконструкций оказывается построение нескольких принципиально различных, но **эквивалентных реконструкций**. Так, при уточнении понятия действительного числа в 1870-х гг. Г. Кантор определял действительные числа как сходящиеся последовательности рациональных чисел, Р. Дедекинд − как "сечения" во множестве рациональных чисел. Можно строго доказать эквивалентность этих реконструкций.

Другим впечатляющим примером является уточнение (казалось бы, весьма неопределенного) интуитивного понятия **вычислимости** (или понятия алгоритма). Начиная с 1930-х гг. было предложено множество различающихся по форме уточнений понятия алгоритма: рекурсивные функции, машины Тьюринга, лямбда-исчисление А. Черча, канонические системы Э. Поста, нормальные алгорифмы А. А. Маркова и др. И во всех случаях была строго доказана эквивалентность этих уточнений.

Эквивалентность различных реконструкций одного интуитивного понятия свидетельствует, что объем реконструированных понятий не является случайным. Это очень важный аргумент в пользу замены интуитивного понятия реконструкцией.

Тенденция перехода от интуитивных понятий к более или менее явным их реконструкциям в истории математики проявляется достаточно четко. Интуитивные теории не могут развиваться без этих реконструкций: усложнение понятий и методов приводит к необходимости их явной реконструкции просто для обеспечения **нормального развития** теории. В большинстве случаев реконструкция выполняется генетическим методом, а когда дело касается фундаментальных математических понятий (например, понятия множества) − аксиоматическим методом (фундаментальные понятия потому и фундаментальны, что их нельзя "генетически" свести к другим понятиям).

Теорема К. Геделя о неполноте (см. раздел 5.3) породила множество рассуждений о том, что аксиоматический метод недостаточен для реконструкции "живого, содержательного" математического мышления. Аксиоматику сравнивали с прокрустовым ложем, которое не в состоянии вместить все богатство содержательной математики. Это рецидив платонизма. Разве могут в математике какие-либо доказательные рассуждения происходить иначе, как по схеме "посылки − заключение"? Если так и всякое математическое рассуждение сводится к цепи

заключений, то можно спросить: эти заключения происходят по **определенным** правилам (т.е таким, которые не меняются от одного случая к другому и от одного математика к другому)? И если правила являются определенными, то, будучи функцией человеческого мозга, могут ли они быть такими, что их нельзя никак явно сформулировать? Если какие-либо "правила" нельзя явно сформулировать, то, следовательно, нельзя **доказать** их определенность! Ну, а полагать, что в математике кроме рассуждений (по определенным правилам) имеются "объекты", существующие независимо от этих рассуждений, означает впасть в обыкновенный платонизм работающего математика.

Таким образом, преждевременно говорить об ограниченности аксиоматизации − границы ее применимости, по-видимому, совпадают с границами применимости самой математики (см. раздел 6.1).

В процессе развития математических теорий аксиоматизация и интуиция взаимодействуют. Аксиоматизация "проясняет" интуицию, когда та "запуталась в себе". Но аксиоматизация влечет за собой и неприятные последствия: многие рассуждения, которые в интуитивной теории опытный специалист проводит очень быстро и представляет компактно, в аксиоматической теории оказываются очень громоздкими. Поэтому после замены интуитивной теории аксиоматической (особенно если эта замена неэквивалентна по причине недостатков интуитивной теории) специалисты развивают новую интуицию, которая восстанавливает способность теории к творческому развитию. Пример тому − история аксиоматизации теории множеств. Когда в интуитивной теории множеств Кантора в 1890-х гг. были обнаружены противоречия, от них удалось избавиться путем аксиоматизации. Естественно, что созданная аксиоматическая теория множеств Цермело-Френкеля отличалась от интуитивной теории Кантора не только формой, но и отдельными аспектами содержания. Для работы в новой теории специалисты развили модифицированную интуицию (в том числе особую интуицию множеств и классов). Ныне вполне нормальной считается работа в теории Цермело-Френкеля на интуитивном уровне. Именно на таком уровне доказываются серьезные новые теоремы этой теории.

Какую пользу дает аксиоматизация?

Во-первых, аксиоматизация позволяет "подправить" интуицию: устранить неточности, двусмысленности и парадоксы, которые иногда возникают из-за неполной контролируемости бессознательных процессов. Самый впечатляющий пример − историю аксиоматизации теории множеств, мы только что отметили.

Во-вторых, аксиоматизация позволяет подвергнуть подробному исследованию отношения между принципами теории (прежде всего

установить их зависимость или независимость), а также между этими принципами и теоремами теории. Для доказательства конкретной теоремы иногда требуются не все аксиомы теории, а только их часть. Исследования такого рода могут привести к созданию более общих теорий, которые применимы в различных конкретных теориях. Характерными примерами являются теория групп и многочисленные ее алгебраические ответвления.

В-третьих, нередко после аксиоматизации удается установить недостаточность данной теории для решения отдельных проблем, естественно возникающих в ней. Именно так произошло с континуум-проблемой в теории множеств. В таких случаях можно ставить вопрос о необходимости совершенствования системы аксиом теории, о развитии альтернативных вариантов теории и т.д.

1.4. Формальные теории

Насколько далеко может зайти процесс аксиоматизации теории? Возможно ли **полное** изгнание интуиции из рассуждений теории, т.е. исчерпывающее сведение теории к системе аксиом и правил вывода?

В трудах Г. Фреге, Б. Рассела и Д. Гильберта, относящихся к концу XIX − началу XX вв., процесс аксиоматизации ряда серьезных математических теорий действительно удалось довести до конца. Они были представлены в виде исчерпывающей системы аксиом и правил вывода, без всякой примеси интуиции. Логическая техника, разработанная этими корифеями, позволяет полностью аксиоматизировать любую теорию, которая основана на застывшей системе принципов (т.е. любую математическую теорию).

Как же выглядят такие полностью аксиоматизированные теории? Чаще всего их называют **формальными теориями**, подчеркивая, что в них ни один шаг рассуждения нельзя сделать, не сославшись на "документ" − точно сформулированный список аксиом и правил вывода. Даже "самоочевидные" логические принципы вроде "если A влечет B и B влечет C, то A влечет C" должны выводиться из явно сформулированных аксиом и правил вывода.

Более точно понятие формальности можно определить в терминах теории алгоритмов: теорию T можно считать формальной, если построен алгоритм (механически применяемая процедура вычисления) для проверки правильности рассуждений с точки зрения принципов теории T. Это значит, что если некто предлалагает математический текст, являющийся, по его мнению, доказательством некоторой теоремы в

теории Т, то механически применяя алгоритм, мы можем проверить, действительно ли предложенный текст соответствует стандартам правильности, принятым в Т. Таким образом, стандарт правильности рассуждений для теории Т определен настолько точно, что проверку его соблюдения можно передать вычислительной машине (следует помнить, что речь идет о **проверке правильности** готовых доказательств, а не об их поиске!). Если проверку правильности доказательств в какой-либо теории нельзя передать вычислительной машине и она доступна в полной мере только человеку, значит, еще не все принципы теории аксиоматизированы (то, что мы не умеем передать машине, остается в нашей интуиции и "оттуда" регулирует наши рассуждения).

В качестве несерьезного примера формальной теории можно рассматривать игру в шахматы − назовем это теорией Ш. Утверждениями в Ш будем считать **позиции** (всевозможные расположения фигур на доске вместе с указанием "ход белых" или "ход черных"). Тогда аксиомой теории Ш естественно считать **начальную позицию**, а правилами вывода − **правила игры**, которые определяют, какие ходы допустимы в каждой позиции. Правила позволяют получать из одних утверждений другие. В частности, отправляясь от нашей единственной аксиомы, мы можем получать теоремы Ш. Общая характеристика теорем Ш состоит, очевидно, в том, что это − всевозможные позиции, которые могут получиться, если передвигать фигуры, соблюдая правила.

Упражнение 1.1. Найдите пример недоказуемого утверждения теории Ш.

В чем выражается формальность теории Ш? Если некто предлагает нам "математический текст" и утверждает, что это − доказательство теоремы А в теории Ш, то ясно, что речь идет о непроверенной записи шахматной партии, законченной (или отложенной) в позиции А. Проверка не является, однако, проблемой: правила игры сформулированы настолько точно, что можно составить программу для вычислительной машины, которая будет осуществлять такие проверки. (Еще раз напомним, что речь идет о проверке правильности записи шахматной партии, а не о проверке того, можно ли заданную позицию получить, играя по правилам, − эта задача намного сложнее!)

Упражнение 1.2. Оцените объем текста этой программы на одном из языков высокого уровня.

Несколько серьезнее другой пример формальной теории (мы заимствуем его у П. Лоренцена). Утверждениями в теории L являются всевозможные цепочки, составленные из букв a, b например a, aa, aba, abaab. Единственной аксиомой L является цепочка a, наконец, в L имеется два правила вывода:

$$\frac{X}{Xb} \; ; \quad \frac{X}{aXa} \; .$$

Такая запись означает, что в теории L из цепочки X непосредственно выводятся Xb и aXa. Примером теоремы L является цепочка aababb:

$$a \vdash ab \vdash aaba \vdash aabab \vdash aababb.$$

Этот факт обычно записывается так: $L \vdash aababb$ (читается: в теории L доказуемо утверждение aababb).

Упражнение 1.3. а) Напишите все теоремы L, содержащие не более 7 букв.

б) Опишите алгоритм, отличающий теоремы L от других ее утверждений.

Очень важное общее свойство формальных теорий дает следующее

Упражнение 1.4. Покажите, что множество всех теорем формальной теории является **эффективно перечислимым** (по другой терминологии − рекурсивно перечислимым).

Таким образом, теоретически для каждой формальной теории существует вычислительная машина, которая печатает на бумаге подряд все ее теоремы (и ничего кроме теорем). К сожалению, в общем случае такая машина мало подходит для решения проблемы, которая обычно интересует математиков: доказуемо ли в данной теории данное утверждение? Если, сидя возле машины, мы дождались момента, когда интересующее нас утверждение напечатано, то проблема решена (утверждение оказалось доказуемым). Но пока этот момент не наступил, мы не можем знать, будет ли утверждение напечатано через некоторое время или не будет напечатано вообще.

Теорию T принято называть **разрешимой** (или эффективно разрешимой), если существует алгоритм, распознающий теоремы T среди всех ее рассуждений. В упражнении 1.3 Вы доказали, что теория L является разрешимой.

Языки серьезных формальных теорий содержат символ отрицания "¬". В таких теориях решение проблемы, выраженной в некотором утверждении A, означает либо доказательство A, либо его опровержение (т.е. доказательство ¬A). Если для решения проблемы попытаться воспользоваться перечисляющей машиной из упражнения 1.4, то мы, сидя возле машины, должны дожидаться печати утверждения A или ¬A. Если будут напечатаны оба утверждения, это будет означать, что теория T **противоречива** (в ней можно доказать некоторое утверждение вместе с его отрицанием). Однако всего здесь четыре возможности: а) будет

напечатано A, но не ¬A, б) будет напечатано ¬A, но не A, в) будет напечатано и A, и ¬A (тогда теория T противоречива), г) не будет напечатано ни A, ни ¬A. В случае г) мы можем сидеть у перечисляющей машины сколь угодно долго, однако ни напечатания A, ни ¬A не дождемся. В этом случае теорию T принято называть **неполной** (а **полной** называется теория, в которой любое утверждение, которое можно сформулировать средствами языка теории, можно либо доказать, либо опровергнуть).

Упражнение 1.5. Докажите, что всякая полная формальная теория разрешима.

1.5. Логика

Логикой принято называть набор средств рассуждения, применяемых во многих теориях. Соответственно и каждая серьезная формальная теория должна иметь среди своих аксиом **логические аксиомы**, а среди своих правил вывода − **логические правила**.

Внешней оболочкой каждой теории является ее **язык**, на котором записываются утверждения теории (аксиомы, теоремы, гипотезы и т.д.). Первичными неделимыми единицами языка серьезных формальных теорий считаются:

а) **переменные** (в своем интуитивном понимании теории мы всегда "неофициально" приписываем переменным какую-либо − одинаковую для всех − область значений: "все натуральные числа", "все множества" и т.п.),

б) **константы** (например, "0" в арифметике интуитивно мы приписываем каждой константе "неофициальное" конкретное значение из области значений переменных),

в) **функциональные символы** (например, "+" в арифметике "неофициально" это функция x+y),

г) **предикатные символы** (язык любой серьезной теории содержит как минимум символ "=", интуитивно понимаемый как равенство "объектов" теории),

д) **логические связки и кванторы** (отрицание "¬", дизъюнкция "∨", конъюнкция "∧", импликация "→", квантор существования "∃", квантор всеобщности "∀"),

е) **скобки и запятые**.

Из переменных, констант и функциональных символов (а также

скобок и запятых) по особым для каждого языка правилам составляются **термы**. Например, в арифметике возможен терм (x+y)+1, где x, y − переменные, 1 − константа, + − функциональный символ. Интуитивно, терм − либо составное обозначение для "объекта" из области значений переменных (например, (1+1)+1 − обозначение числа 3), либо обозначение функции.

Далее, из термов и предикатных символов составляются **атомарные формулы**. Например, $(t_1=t_2)$, где t_1, t_2 − любые термы теории. Атомарная формула $P(t_1, ..., t_n)$, где t_i − термы, P − предикатный символ, представляет собой утверждение, что "объекты", обозначенные термами $t_1, ... ,t_n$, находятся в отношении, обозначенном через P.

Из атомарных формул, логических связок и кванторов по обычным правилам составляются **формулы** теории, например

$$(\forall x)((x=0)\vee\neg(x=0)).$$

Это пример **замкнутой формулы**, все переменные которой связаны кванторами. Замкнутая формула − "определенное утверждение об объектах теории", тогда как "истинность" формулы, имеющей свободные переменные, может зависеть от того, какие конкретные значения эти переменные принимают. Например, формула $(x=0)\vee(x=1)$ оказывается "истинной" при x=1 и "ложной" при x=2.

Покажем теперь, как можно сформулировать список логических аксиом и правил вывода, достаточный для воспроизведения общепринятых логических средств рассуждения. Большинство аксиом будут представлены **схемами аксиом** (каждая схема включает в себя бесконечное, но легко распознаваемое семейство аксиом). Эквивалентность не считается самостоятельной логической связкой, $A\leftrightarrow B$ определяется как $(A\rightarrow B)\wedge(B\rightarrow A)$.

Сначала список аксиом:

L1) $A\rightarrow(B\rightarrow A)$,

L2) $(A\rightarrow(B\rightarrow C))\rightarrow((A\rightarrow B)\rightarrow(A\rightarrow C))$,

L3) $A\rightarrow(B\rightarrow A\wedge B)$,

L4) $A\wedge B\rightarrow A$,

L5) $A\wedge B\rightarrow B$,

L6) $A\rightarrow A\vee B$,

L7) $B\rightarrow A\vee B$,

L8) $(A\rightarrow C)\rightarrow((B\rightarrow C)\rightarrow(A\vee B\rightarrow C))$,

L9) (A→B)→((A→¬B)→¬A),

L10) ¬A→(A→B) (из противоречия следует все),

L11) A∨¬A (закон исключенного третьего),

L12) (∀x)D(x)→D(t),

L13) D(t)→(∃x)D(x).

Здесь A, B, C − произвольные формулы, D − формула и t − терм такие, что кванторы D не связывают переменных, входящих в t.

Это были схемы аксиом **исчисления предикатов**. Далее следуют аксиомы, описывающие **свойства равенства**:

L14) x=x,

L15) x=y → y=x,

L16) x=y → ((y=z)→(x=z)),

L17) x=t → (D(x, x)→D(x, t)).

Здесь D − формула, содержащая x и переменные терма t только свободно. Обозначение D(x, x) предполагает, что все вхождения переменной x в формулу D разбиты на две группы. Если все вхождения второй группы заменить на t, получится формула D(x, t).

Кроме аксиом нужны еще три **правила вывода**:

$$\frac{A,\ A\to B}{B}\ ,\quad \frac{C\to D(x)}{C\to(\forall x)D(x)}\ ,\quad \frac{D(t)\to C}{(\exists x)D(x)\to C}\ .$$

Первое правило принято называть MODUS PONENS, в нем A, B − произвольные формулы теории. В остальных двух правилах формула C не должна содержать x и переменные терма t.

Приведенный здесь список аксиом и правил вывода достаточен для построения обычных логических средств рассуждения, используемых в математических теориях. Определение всякой серьезной формальной теории должно включать либо этот список, либо один из многих возможных его эквивалентов. Особенности же каждой отдельной теории проявляются: а) в количестве констант, функциональных и предикатных символов, б) в правилах образования термов и атомарных формул, в) в **собственных аксиомах** теории. Собственных правил вывода формальные теории обычно не имеют.

Полный набор перечисленных выше аксиом задает так называемую **классическую логику**. Она используется в подавляющем большинстве математических теорий. Исключив схему аксиом L11: A∨¬A (закон исключенного третьего), получаем так называемую

конструктивную (или **интуиционистскую**) **логику**. В теориях, использующих конструктивную логику, невозможны доказательства существования объектов методом "от противного". Такие неконструктивные доказательства существования основаны на схеме ¬¬A→A, которая равносильна L11 (с целью доказать утверждение A мы принимаем ¬A в качестве гипотезы, выводим противоречие, т.е. доказываем ¬¬A, и делаем вывод об истинности A). Подробнее о конструктивной математике см. А. Г. Драгалин [1979], Б. А. Кушнер [1973].

Интересную роль играет схема L10: ¬A→(A→B). Согласно ей если в теории обнаружено **противоречие** (выведены одновременно некоторая формула A и ее отрицание ¬A), то тем самым становится выводимой **любая формула** B. Таким образом, для доказательства непротиворечивости какой-либо теории достаточно установить невыводимость в этой теории хотя бы одной формулы (например, 0=1)).

Следствием логических аксиом является также формула (∃x)x=x (проверьте). Таким образом, принятие одних только логических аксиом уже гарантирует "существование" по крайней мере одного "объекта" теории.

Важным логическим принципом является **теорема дедукции**: если в теории T, приняв в качестве гипотезы формулу A, можно доказать формулу B (сокращенная запись T, A ├ B), причем в этом доказательстве к свободным переменным из A не применяются второе и третье правила вывода, то T ├ A→B, т.е. уже без всяких гипотез в теории T доказуема импликация A→B.

Упражнение 1.6. Покажите, используя теорему дедукции, что аксиомы L15, L16 следуют из остальных аксиом.

1.6. Программа Гильберта

К началу XX в. честь математики была поставлена под серьезное сомнение: в теории множеств были обнаружены парадоксы − самые настоящие противоречия. К этому времени теория множеств уже успела показать себя как естественная основа и плодотворнейшее орудие математики. Для спасения "основы и орудия" немецкий математик Давид Гильберт предложил в 1904 г. свою программу перестройки оснований математики, которая состояла из двух частей:

а) Представить существующую математику (включая очищенный от парадоксов вариант теории множеств) в виде формальной теории.

б) Доказать **непротиворечивость** полученной теории (т.е. доказать, что в этой теории никакое утверждение не может быть доказано вместе со своим отрицанием).

Решить задачу п. а) означало довести до конца процесс аксиоматизации математики, который в XIX в. и так уже продвинулся решительным образом (уточнение понятий функции, непрерывности, действительного числа, аксиоматизация арифметики натуральных чисел, геометрии и т.д.). Задача же п. б) была радикальным нововведением − попытаться **доказать** непротиворечивость полученной в п. а) всеобъемлющей теории. Д. Гильберт первым понял, что решение до конца задачи а) делает возможной постановку задачи б).

Дело в том, что, не решив до конца а), т.е. оставаясь отчасти в области интуитивной математики, нельзя говорить об абсолютном доказательстве непротиворечивости математики. В интуитивной теории можно надеяться обнаружить противоречие, т.е. можно надеяться доказать, соблюдая общепринятые (интуитивные) правила рассуждения, некоторое утверждение вместе с его отрицанием. Но никак нельзя даже пытаться доказывать непротиворечивость интуитивной теории, поскольку утверждение о непротиворечивости относится к множеству всех теорем, доказуемых в теории, т.е. к совокупности, четкого определения которой мы как раз не имеем.

Однако, если вместо интуитивной теории взять формальную, положение изменяется. Множество теорем формальной теории является уже точно определенным объектом. Несмотря на это оно все же бесконечно. Каким же образом Д. Гильберт рассчитывал получить доказательство утверждения о непротиворечивости, относящегося ко всему этому бесконечному множеству?

Вернемся к нашим примерам формальных теорий. В теории Ш множество всех теорем оказывается, правда, конечным (хотя конечность эта с практической точки зрения ближе к бесконечности). Легко доказать следующее утверждение, относящееся ко **всем** теоремам Ш: ни в одной теореме белые не имеют 10 ферзей. В самом деле, достаточно заметить, что в аксиоме Ш белые имеют 1 ферзь и 8 пешек и что по правилам игры белым ферзем может стать только белая пешка. Остальное решает арифметика: 1+8<10. Таким образом, мы подметили в системе аксиом и правил вывода теории Ш особенности, которые делают справедливым наше общее утверждение о теоремах Ш.

Аналогичные возможности имеем в случае теории L. Можно доказать, например, следующее утверждение, относящееся ко всем теоремам L: если X − теорема, то aaX − тоже теорема. В самом деле, если X=a (X − аксиома), то L ├ aaa по второму правилу вывода. Это базис

индукции. Шаг индукции: если для теоремы X выводимость aaX уже установлена, то для теорем aXa и Xb имеем aaX ├ aaaXa, aaX ├ aaXb. Таким образом, индукцией по структуре вывода справедливость нашего утверждения установлена.

Итак, если множество всех теорем теории точно определено, можно доказывать утверждения, относящиеся ко всем теоремам одновременно. Д. Гильберт полагал, что утверждение о непротиворечивости теории не будет исключением. Грубо говоря, он надеялся подметить особенности системы аксиом всей математики, которые делают вывод противоречия невозможным.

Заметим, однако, что утверждение, относящееся к бесконечному множеству объектов, не может быть доказано простой эмпирической проверкой (перебором объектов). Всякое его доказательство неизбежно должно быть **теоретическим**. Например, при доказательстве утверждения L ├ X → L ├ aaX мы воспользовались математической индукцией. Так **в какой же теории** следует доказывать непротиворечивость формальной теории, охватывающей всю существующую математику? Ясно, что средства рассуждения, используемые для обоснования непротиворечивости некоторой теории T, должны быть более надежными по сравнению со средствами, допускаемыми в самой теории T. В самом деле, можно ли доверять доказательству непротиворечивости, если в нем используются сомнительные средства, сами нуждающиеся в обосновании? Но если теория T охватывает всю математику, то никаких средств рассуждения, выходящих за рамки T, математик знать не может. Поэтому необходимые для доказательства непротиворечивости средства рассуждения мы вынуждены черпать из самой (универсальной) теории T − из той ее части, которая представляется нам наиболее надежной.

В математике четко выделяется два уровня "надежности" рассуждений:

1) **арифметические** ("дискретные") рассуждения используют только понятие целого числа и аналогичные дискретные объекты,

2) **теоретико-множественные** рассуждения используют канторовское понятие о произвольном множестве.

Первый уровень считается надежным (его мало кто подвергает сомнению), второй − опасным (его только недавно "очистили" от явных противоречий). Д. Гильберт рассчитывал, разумеется, на доказательство непротиворечивости всей математики средствами первого уровня.

Сразу, как только Д. Гильберт объявил о своем проекте (1904 г.), французский математик А. Пуанкаре высказал сомнения в его

реальности. По мнению А. Пуанкаре, Д. Гильберт, доказывая непротиворечивость математики с помощью математической индукции (самое важное средство первого уровня), допускает в своих рассуждениях **порочный круг**. Непротиворечивость математики означает и непротиворечивость математической индукции,... доказанную с ее же помощью! Мало кто тогда (включая самого Д. Гильберта) мог осознать серьезность этого намека... Но через 25 лет К. Гедель доказал, что А. Пуанкаре был прав: абсолютное доказательство непротиворечивости математики невозможно (см. раздел 5.4).

2. АКСИОМАТИЧЕСКАЯ ТЕОРИЯ МНОЖЕСТВ

2.1. Возникновение интуитивной теории множеств

Возникновение интуиции произвольного бесконечного множества – закономерный результат развития математики XIX в. Приемы образования новых математических понятий, характерные для первой половины XIX в., в теории множеств оказались доведенными до логического конца. Последний шаг был сделан немецким математиком Георгом Кантором под влиянием конкретной проблемы из "старой" математики. К интуитивному понятию о произвольном бесконечном множестве Г. Кантора привела, как ни странно, проблема сходимости рядов Фурье. В 1872 г. он доказал теорему о единственности разложения функции в ряд Фурье для рядов, которые сходятся в интервале (a, b) всюду, исключая, возможно, конечное число точек, и задался вопросом: насколько можно эту теорему обобщить?

В то время математики почти не думали и не говорили об отрезке прямой как о множестве точек. Они представляли себе отрезок как континуум – непрерывную, сплошную среду, в которой можно отмечать отдельные точки, но не задумывались над тем, "состоит" ли среда целиком из точек, исчерпывается ли ими. Сегодня трудно вообразить такое, но в середине XIX в. эта точка зрения господствовала.

Представление об отрезках прямых как совокупностях точек появилось в самом начале развития "чистой" математики – в VI в. до н.э. Естественно, тогда речь могла идти только о точках положительных размеров и о конечном их числе в каждом отрезке. Если считать при этом, что все точки – одинаковые, то легко сделать вывод, что отношение любых двух отрезков должно выражаться рациональным числом. Такая интуиция успела уже достаточно сильно укорениться, когда было обнаружено существование несоизмеримых отрезков (прежде всего несоизмеримость диагонали квадрата и его стороны). С этим открытием связан первый кризис основ математики: дискретная геометрическая интуиция потерпела крах. Выход из кризиса был найден в V в. до н.э. – от представления об отрезке прямой как о совокупности точек пришлось отказаться. Вместо него отрезок представлялся как сплошная среда, в которой можно отмечать отдельные точки, можно определить отношение двух отрезков и т.д. Эта новая интуиция сохранялась практически без

изменений до 70-х гг. XIX в., и ее не смогло поколебать даже изобретение бесконечно малых величин.

Итак, Г. Кантор не мог сразу задаться вопросом: останется ли в силе теорема о единственности разложения, если множество исключительных точек будет бесконечным? У него не могло быть тогда еще **общего** понятия о бесконечном множестве точек. Поэтому он начал с простейшего вида бесконечных множеств – с множеств, имеющих одну предельную точку, например

$$\{ \ \frac{1}{n} \ | \, n \geq 1 \ \}.$$

Предельной точкой этого множества является нуль. И Г.Кантору удалось доказать теорему единственности для случая, когда множество исключительных точек имеет одну предельную точку. Дальнейшее обобщение на случай множества с конечным числом предельных точек уже тривиально.

Следующий шаг – к простейшему множеству, имеющему бесконечно много предельных точек. Такое множество имеет единственную предельную точку второго порядка, вокруг которой сгущаются "обычные" предельные точки, например

$$\{ \ \frac{1}{m}+\frac{1}{n} \ | \, m, n \geq 1\}.$$

И для таких множеств Г.Кантору удалось доказать свою теорему единственности. Дальше можно идти к предельным точкам третьего, четвертого и других порядков, вовлекая в работу все более сложные бесконечные множества точек.

Так постепенно, в процессе работы со все более сложными множествами точек, у Г. Кантора начало формироваться интуитивное понятие о произвольном бесконечном множестве. Уже попытка систематизации множеств с предельными точками определенных порядков провоцирует введение этого понятия. Вводится понятие так называемого производного множества: если P – множество точек, то P' – множество предельных точек P. По индукции можно определять дальше P", P"', P"",...

Упражнение 2.1. Покажите, что при фиксированном k для множества

$$Q_k = \{ \ \frac{1}{n_1}+...+\frac{1}{n_k} \ | \, n_1 ,..., n_k \geq 1 \ \}$$

k-е производное множество $Q_k^{(K)}$ состоит из единственной точки 0.

Свою теорему о единственности разложения в ряд Фурье Г. Кантору удалось обобщить на любые множества исключительных точек конечного порядка, т.е. на любые множества P, для которых одна из производных $P^{(k)}$ конечна.

Но и отсюда можно двинуться дальше. Можно перейти к предельным точкам "бесконечного порядка" – так естественно называть точку, которая является предельной точкой k-го порядка для любого натурального числа k. Это значит, что производное множество бесконечного порядка $P^{(\omega)}$ определяется как пересечение $P' \cap P'' \cap P''' \cap ...$ Разумеется, Г. Кантор доказал теорему о единственности и для случая, когда для множества исключительных точек P производное множество $P^{(\omega)}$ является конечным. Но и отсюда можно двинуться дальше – к производным множествам

$P^{(\omega+1)} = (P^{(\omega)})'$, $P^{(\omega+2)}$, $P^{(\omega+3)}$,...,

$P^{(\omega\cdot 2)}$, $P^{(\omega\cdot 2+1)}$,..., $P^{(\omega\cdot 3)}$, $P^{(\omega\cdot 3=1)}$,...,

$P^{(\omega\cdot\omega)}$, $P^{(\omega\cdot\omega+1)}$, ...

Таким образом, те же ряды Фурье привели Г. Кантора не только к интуиции произвольного бесконечного множества, но и к расширению понятия натурального числа – к так называемым бесконечным порядковым числам (или ординалам, см. дальше).

Упражнение 2.2. Попробуйте построить множество P такое, что $P^{(\omega)} = \{0\}$, или $P^{(\omega\cdot 2)} = \{0\}$, или $P^{(\omega\cdot\omega)} = \{0\}$. Указание: возьмите объединение всех множеств $R_k = Q_k \cap [\,0, \frac{1}{k}\,]$.

К осени 1873 г. наступил решающий перелом: 29 сентября Г. Кантор пишет в письме Р.Дедекинду, что можно установить взаимно однозначное соответствие между рациональными и натуральными числами. Это была известная теперь всем конструкция:

$$\begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \ldots \\ \frac{1}{1} & \frac{1}{2} & \frac{2}{1} & \frac{1}{3} & \frac{3}{1} & \frac{1}{4} & \frac{2}{3} & \frac{3}{2} & \frac{4}{1} & \ldots \\ 2 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & \ldots \end{array} .$$

Сначала идут несократимые дроби с суммой числителя и знаменателя, равной 2, затем – с суммой, равной 3, и т.д. И еще: в этом же письме Г. Кантор спрашивает: а не удастся ли и все действительные числа перенумеровать с помощью натуральных чисел?

Это была целая революция в представлениях о математическом

континууме! Г. Кантор уже считает числовую прямую множеством точек – не просто средой, в которой можно отмечать отдельные точки, а средой, которая **состоит** из точек, исчерпывается ими! Это закономерный результат занятий Г. Кантора все более сложными множествами точек – в его глазах континуум "расчленился" на отдельные точки. Сейчас представление о континууме как о множестве точек опять (через 2000 лет!) прочно вошло в математику, поэтому трудно вообразить, что когда-то здесь могли быть какие-то проблемы.

В ответном письме Р. Дедекинд показал, как можно перенумеровать натуральными числами все алгебраические числа. Но перенумеровать все действительные числа ему не удалось...

Разумеется, это не случайно, поскольку в своем следующем письме Р. Дедекинду (7 декабря 1873 г.) Г. Кантор показывает, что взаимно однозначное соответствие между натуральными и действительными числами невозможно. В своем доказательстве Г. Кантор применил конструкцию, названную впоследствии диагональным методом. Он исходил из произвольной последовательности действительных чисел a_1 , a_2 ,..., a_n ,... и произвольного интервала (b, c), делил интервал на три части, брал ту из частей, которая не содержит a_1 , затем делил на три эту часть и брал ту треть, которая не содержит a_2 , и т.д. В результате получалась последовательность стягивающихся интервалов (b_i, c_i):

$$b_1 \le b_2 \le b_3 \le ... \le c_3 \le c_2 \le c_1.$$

Общая точка (предел) этих интервалов и представляет собой действительное число, не входящее в последовательность a_1 , a_2 ,..., a_n ,... Таким образом, никакая последовательность, пронумерованная натуральными числами, не может исчерпать все действительные числа.

Это была еще одна революция – в представлениях о математической бесконечности. Оказывается, наряду с бесконечным множеством натуральных чисел существует "еще более бесконечное" множество действительных чисел, т.е. существуют бесконечности по крайней мере **двух** типов. Теорема Кантора дает также поразительно простое доказательство существования трансцендентных чисел (и одновременно доказательство того, что трансцендентных чисел "гораздо больше", чем алгебраических, которые можно перенумеровать с помощью натуральных чисел). Правда, конкретные трансцендентные числа построил еще в 1844 г. Ж. Лиувилль, а в 1873 г. Ш. Эрмит доказал, что трансцендентным является число e.

Обнаружив существование двух типов бесконечности, Г. Кантор

пошел дальше: в письме Р. Дедекинду от 5 января 1874 г. он пишет о своих попытках сравнить континуумы различной размерности. Например, где больше точек: внутри квадрата или в отрезке прямой? Казалось бы, чем больше размерность, тем больше должно быть точек. Г. Кантор также поверил в это и более 3 лет пытался доказать. Только потом он наконец решил попытаться доказать противное (невероятное!) – что в квадрате столько же точек, сколько в отрезке. И это ему сразу же удалось, о чем он сообщил в письме Р. Дедекинду от 20 июня 1877 г. Конструкцию Г. Кантора легко объяснить любому школьнику. Отображение квадрата [0, 1)×[0, 1) в отрезок [0,1] задается с помощью десятичных разложений координат:

(x, y) № z,

x=0,abcd...,

y=0,ABCD...,

z=0,aAbBcCdD...

Г. Кантору показалось, что его доказательство "уничтожило" понятие размерности. В ответном письме Р. Дедекинд указал, что отображение Г. Кантора, будучи взаимно однозначным, не является непрерывным (при непрерывном отображении размерность вроде бы должна сохраняться?). Однако позднее – Дж. Пеано в 1890 г. и Д. Гильберт в 1891 г. сумели построить непрерывное отображение отрезка на квадрат (но это отображение взаимно однозначным уже не оказалось). И только в 1911 г. голландский математик Л. Брауэр (позднее – основоположник интуиционизма) доказал, что размерность сохраняется при отображениях, которые одновременно являются и непрерывными, и взаимно однозначными.

Тогда же, в 1877 г. Г. Кантор пришел к **континуум-проблеме**: поработав с самыми различными множествами точек (на прямой, на плоскости, в пространстве), он обнаружил только два типа бесконечных множеств:

– **счетные** множества (их элементы можно перенумеровать с помощью натуральных чисел),

– множества, эквивалентные всему континууму (например, отрезку прямой).

Никаких множеств "промежуточной мощности" (содержащих элементов больше, чем натуральных чисел, но меньше, чем континуум) обнаружено не было (подробнее см раздел 2.4). Поэтому Г. Кантор предположил, что таких множеств вообще не существует. Это предположение принято называть **континуум-гипотезой**: всякое бесконечное множество точек на

прямой либо является счетным, либо эквивалентно всему континууму.

Много лет потратил Г. Кантор, пытаясь доказать эту гипотезу. Континуум-проблема – одна из самых красивых проблем во всей математике – ее суть легко объяснить любому школьнику. Решить ее не удалось ни Г. Кантору, ни многочисленным его последователям.

В 1895 г. Г. Кантора, изнуренного безуспешными попытками доказать континуум-гипотезу, настигает еще один удар – он обнаруживает в своей теории множеств... противоречие! Об этом он сообщает в письме Д. Гильберту. В 1897 г. еще одно противоречие обнаруживает (и немедленно публикует) итальянский математик Ч. Бурали-Форти...

Более основательное изучение ранней истории теории множеств по книгам Ф. А. Медведева [1965, 1982] только усилит впечатление, что эта теория является закономерным результатом развития математики XIX в.: в канторовском понятии о произвольном бесконечном множестве доведены до логического конца принципы математического мышления, характерные для всего предыдущего периода. И что обнаруженные противоречия столь же закономерны.

2.2. Формализация противоречивой теории множеств

Сейчас мы будем заниматься формализацией теории множеств в том виде, в каком она была создана Г. Кантором. В основу этой теории положено интуитивное представление о "мире множеств", в котором все множества (и конечные, и бесконечные) существуют одновременно и в "готовом виде". В своих аксиомах теории множеств мы хотим отразить законы, характеризующие этот застывший "мир множеств".

Уже в самом начале возникает, однако, такой вопрос: возможен ли мир, состоящий **только** из множеств? Множество может состоять из элементов-множеств. Но что лежит тогда в основании этой "башни"? Т.е. если x_2 – элемент множества x_1 , а x_3 – элемент x_2 , x_4 – элемент x_3 и т.д., то неужели в этой цепи никогда не появится что-то более "осязаемое" по сравнению с множествами, состоящими из множеств? Как показал в 1925 г. Дж. фон Нейман, теория множеств вполне может обойтись без введения "осязаемых" объектов. В самом деле, если бы "в мире множеств ничего не было", то этот мир представлял бы собой пустое множество, а это уже кое-что. Обозначим это "кое-что" через 0. Тогда мы можем образовать еще одно множество, состоящее из одного элемента 0, т.е. множество $\{0\}$. Следующим шагом может быть образование множества

из двух элементов: {0,{0}}, и т. д.:

$$x_0 = 0,\ x_1 = \{0\},\ x_2 = \{0,\{0\}\},\ldots,\ x_{n+1} = x_n \cup \{x_n\},\ldots$$

Таким образом, даже предположив, что "ничего нет", мы получаем бесконечно много множеств, т.е. если в основу "башни множеств" положить пустое множество, это нас по существу ничем не ограничивает (ср. К. Дэвлин [1977]).

Теперь мы можем определить **язык теории множеств**. Переменными будут служить, как обычно, x, y, z, ... – с индексами или без них. Значениями переменных (интуитивно) будем считать произвольные множества (поскольку в "мире множеств" существуют только множества).

Константы (например, 0 для обозначения пустого множества) мы вводить не станем. Позднее увидим, что без них можно обойтись. Совершенно необходимо, однако, ввести особый предикатный символ "$\in$" ("принадлежит") – в дополнение к общему для всех теорий символу "=". В результате появятся **атомарные формулы** двух видов: $x \in y$ ("x принадлежит y" или "x является элементом множества y") и $x=y$ ("x, y – одно и то же множество"). Атомарные формулы будем соединять с помощью логических связок и кванторов, создавая **формулы теории множеств**. Например, формула $(\forall y)\neg(y \in x)$ утверждает, что x есть пустое множество, а формула

$$(\exists y)(\forall z)(z \in x \leftrightarrow z=y)$$

– что x содержит единственный элемент.

Как и всякая серьезная формальная теория, теория множеств принимает логические аксиомы и правила вывода (классическую логику). Переходя к собственным аксиомам теории множеств, мы должны прежде всего определить специфическое для этой теории понятие равенства: "два множества равны, если они состоят из одних и тех же элементов". Несмотря на кажущуюся тривиальность это определение требует специальной аксиомы. Из логических аксиом оно не выводится. Из логических аксиом можно вывести (проверьте), что

$$x=y \rightarrow (\forall z)(z \in x \leftrightarrow z \in y).$$

Это естественно: если x равно y, то все, что можно сказать об x, можно сказать и об y. Однако если множества x,y определены по-разному и лишь ценой некоторых усилий мы сумели установить, что $(\forall z)(z \in x \leftrightarrow z \in y)$, то отсюда вытекает $x=y$ лишь при специфическом для теории множеств понятии равенства. Логика безразлична к символу "$\in$", она не приписывает ему никаких индивидуальных свойств. Именно поэтому мы и должны ввести специальную аксиому:

$$(\forall x \forall y)((\forall z)(z \in x \leftrightarrow z \in y) \rightarrow x=y). \qquad (Z1)$$

Это так называемая **аксиома экстенсиональности**. Она и определяет специфическое понимание равенства в теории множеств (термин "экстенсиональность" означает независимость от конкретного вида определения).

Уже само принятие логических аксиом и правил вывода (то, что мы относим их к "миру множеств") гарантирует существование по крайней мере одного множества: следствием логических аксиом является формула $(\exists x)x=x$. Здесь просто утверждается существование некоторого множества x, без каких-либо индивидуальных свойств. Чтобы получить множества, обладающие конкретными свойствами, нужны специальные аксиомы.

Г. Кантор писал в свое время: "Множество – это многое, мыслимое нами как единое". Но каким образом многое можно мыслить как единое? Только задав отличительный признак того, что входит в многое. Это и будет тем единым, что может связать многое вместе. Обозначения, отражающие такой подход, прочно вошли в математику, например $K=\{x \mid K(x)\}$, где K – свойство "быть крокодилом". По этому поводу принято говорить, что K – "множество всех крокодилов", и обращаться с ним как с единым объектом.

Немецкий математик Готлоб Фреге (первый ученый, достигший серъезных успехов в формализации математики) в своей формальной системе реализовал эту идею объединения многого в единое. На современном нам языке принцип Кантора-Фреге формулируется в виде так называемой **схемы аксиом свертывания**. Пусть F(y) – формула на языке теории множеств (кроме переменной y она может содержать и другие свободные переменные, тогда мы будем считать их параметрами). На F(y) можно смотреть как на утверждение: "множество y обладает свойством F". Следуя подходу Кантора-Фреге, все множества y, обладающие свойством F, можно объединить в единое множество $x=\{y \mid F(y)\}$. Таким образом, мы принимаем схему аксиом

$$(\exists x)(\forall y)(y \in x \leftrightarrow F(y)) \qquad (Z2\{F\})$$

(предполагается, что формула F не содержит x). Для конкретной формулы F получается конкретная аксиома свертывания.

В частности, аксиома $Z2\{\neg(y=y)\}$ обеспечивает существование пустого множества:

$$(\exists x)(\forall y)(y \in x \leftrightarrow \neg(y=y)).$$

Ясно, что x здесь – пустое множество: $(\forall y)\neg(y \in x)$.

Упражнение 2.3. Выведите из схемы свертывания существование

других множеств: {0}, {0, {0}}.

Для полноценного воспроизведения интуитивной теории множеств нам не хватает еще только знаменитой аксиомы выбора. Однако мы не будем пока ее обсуждать, так как уже аксиомы, которые приняты нами до сих пор..., приводят к противоречию!

В 1902 г., когда Г. Фреге уже правил корректуру второго, завершающего, тома книги, излагающей его формальную систему математики, он получил письмо от английского философа Б. Рассела, который, ознакомившись с первым томом, обнаружил, что принципы, принятые в системе Фреге, очень быстро приводят к противоречию. В наших терминах рассуждения Б.Рассела выглядели бы так. Некоторые множества не являются элементами самих себя, таково, например, пустое множество: $\neg(0 \in 0)$. Рассмотрим множество **всех** таких множеств, существующее согласно аксиоме свертывания $Z2\{\neg(y \in y)\}$:

$$(\exists x)(\forall y)(y \in x \leftrightarrow \neg(y \in y)).$$

В частности, полагая y=x:

$$(\exists x)(x \in x \leftrightarrow \neg(x \in x)),$$

т.е. существует множество x, для которого приводит к противоречию и предположение $x \in x$, и предположение $\neg(x \in x)$. Так появился на свет **парадокс Рассела**. Как ни странно, оказалось, что некоторые из аксиом свертывания приводят к противоречиям. Это значит, что в самом общем виде схема свертывания не может служить основой теории множеств.

Замечание. Текст письма Б. Рассела к Г. Фреге см. в книге Б. В. Бирюкова [1985]. Еще до Б. Рассела в 1895 г. противоречие в своей теории множеств обнаружил Г.Кантор. В 1897 г. еще один парадокс теории множеств обнаружил и опубликовал Ч. Бурали-Форти. Их рассуждения были значительно сложнее рассуждений Б. Рассела (см. раздел 2.4). (Подробнее об истории открытия парадоксов в теории множеств см. книгу Ф. А. Медведева [1965].)

Сегодня, почти 100 лет спустя, открытие парадоксов уже не кажется чем-то катастрофическим. Но какое впечатление оно должно было произвести (и произвело) на Г. Фреге, Г. Кантора и Р. Дедекинда, которые поверили в неограниченную применимость принципа свертывания еще в 70-е гг. XIX в. и прожили с этой верой более 20 лет! Г.Фреге свою формальную систему математики, а Г. Кантор – свою теорию множеств считали главным делом всей жизни. Удар оказался очень тяжелым: после 1902 г. они не опубликовали ни одной работы, сколько-нибудь сравнимой с их выдающимися работами предыдущего периода. Найти выход из положения они уже не могли. Выход нашли математики следующего поколения.

2.3. Аксиомы Цермело-Френкеля

Решающая идея принадлежит Эрнсту Цермело. Он предложил принять в качестве аксиом теории множеств только те варианты аксиом свертывания, которые **реально используются** в математике. К 1908 г. Э. Цермело уже завершил "переучет" средств, которые используются в математике для образования множеств, и опубликовал свой вариант аксиом теории множеств. Оставим пока в стороне вопрос о том, какие конкретные аксиомы содержала эта система. Понятно, что аксиомы Z2{$\neg(y \in y)$}, приводящей к парадоксу Рассела, в ней не было, поскольку рассуждения такого рода в математической практике не встречаются.

Но как же быть в случаях, когда с помощью некоторой формулы F(y) из всех множеств y выделяются те, которые "обладают свойством F", однако, предположив, что все выделенные y, собранные вместе, образуют множество, приходится констатировать противоречие? Так, {$y|\neg(y \in y)$} является, по-видимому, некоторым "собранием" множеств (в частности, пустое множество 0 является элементом этого собрания: $\neg(0 \in 0)$). Однако если объявить такое собрание, множеством и обозначить его через x, то получится парадокс Рассела: $x \in x$ приводит к $\neg(x \in x)$ и наоборот. Как же быть? Решение, достойное царя Соломона: выведенное нами противоречие будем считать **доказательством** (от противного), что собрание Рассела... **не является множеством**!

Действуя в таком духе, мы должны ввести понятие "собраний" множеств, более широкое, чем понятие множества. Этим более широким понятием является понятие **класса**, которое определяется следующим образом. Пусть $F(y, z_1, ..., z_n)$ – формула на языке теории множеств ($z_1, ..., z_n$ могут отсутствовать). Тогда будем говорить, что при фиксированных $z_1, ..., z_n$ формула F определяет nкласс

$$\mathbf{A} = \{y \mid F(y, z_1, ..., z_n)\}.$$

Ясно, что, изменяя параметры $z_1, ..., z_n$, получим, вообще говоря, другой класс. Например, класс $\{y \mid \neg(y=z_1)\}$ состоит из всех множеств, исключая z_1, и поэтому зависит от того, как зафиксировано z_1. С другой стороны, класс

$$\mathbf{V} = \{y \mid y=y\}$$

состоит из всех вообще множеств и ни от каких параметров не зависит.

Всякое множество является классом: множество x совпадает с

классом $\{y \mid y \in x\}$. Однако не всякий класс можно отнести к множествам. Например, класс Рассела

$$\mathbf{R} = \{y \mid \neg(y \in y)\}$$

не совпадает ни с одним множеством (если допустить, что **R** совпадает с x, то $y \in x$ равносильно $\neg(y \in y)$ и при y=x получается противоречие). Такие классы принято называть **собственными классами**. То, что раньше считалось парадоксами теории множеств, теперь будет считаться доказательствами "собственности" соответствующих классов.

Классы будем обозначать прописными буквами: **A, B, C, ...** (в отличие от множеств, обозначаемых строчными буквами: a, b, c, ...). Такие обозначения должны лишний раз напомнить, что записи вроде

$$y \in \mathbf{A},\ \mathbf{A}=\mathbf{B},\ \mathbf{A}\subseteq\mathbf{B},\ \mathbf{A}\cap\mathbf{B},\ \mathbf{A}\cup\mathbf{B},\ \mathbf{A}-\mathbf{B}$$

не относятся непосредственно к языку теории множеств (в котором нет "прописных переменных"), а являются всего-навсего более наглядной для нашего восприятия записью следующих формул (формула F определяет класс **A**, формула G – класс **B**):

$$F(y),\ (\forall y)(F(y)\leftrightarrow G(y)),\ (\forall y)(F(y)\rightarrow G(y)),$$
$$F(y)\wedge G(y),\ F(y)\vee G(y),\ F(y)\wedge\neg G(y).$$

А теперь мы должны, следуя Э. Цермело, сформулировать те варианты схемы свертывания, которые реально используются в математике. Наше изложение отличается от изложения Э. Цермело некоторыми усовершенствованиями, предложенными позднее.

Самый, по-видимому, распространенный способ образования новых множеств – выделение с помощью явно сформулированного условия части элементов уже известного множества. Так определяются, например, различные подмножества натуральных чисел – четные числа, простые числа и т.д. В общем же случае имеем следующую картину: условие F(y), представленное формулой из языка теории множеств (эта формула может содержать кроме y еще и параметры $z_1, ..., z_n$), выделяет среди элементов множества x те, которые удовлетворяют F. Эти выделенные элементы образуют подмножество множества x, обозначенное через z.

----------	---------- $y \in x$ ----------	----------
----------	---------- $F(y)$ ----------	----------
----------	---------- $y \in z$ ----------	----------

Чтобы узаконить такой способ рассуждения, вводится следующая **схема аксиом выделения**: если $F(y, z_1, ..., z_n)$ – любая формула со

свободными переменными y, z_1, ..., z_n, не содержащая переменных x, z, то следующая формула объявляется аксиомой:

$$(\forall z_1 \ldots \forall z_n)(\exists z)(\forall y)(y \in z \leftrightarrow y \in x \wedge F(y, z_1, \ldots, z_n)). \qquad (Z21\{F\})$$

Ясно, что это частный случай схемы свертывания, а именно $Z2\{y \in x \wedge F(y, z_1, \ldots, z_n)\}$. Особенно наглядно аксиомы выделения формулируются в терминах классов: если формула F определяет класс **A**, то аксиома Z21{F} утверждает, что пересечение $\mathbf{A} \cap x$ (класса **A** и множества x) является множеством: $\mathbf{A} \cap x = z$.

Теперь можно доказать существование пустого множества, т.е. формулу $(\exists x)(\forall y)\neg(y \in x)$. В самом деле, мы уже знаем, что какие-то множества существуют (это следствие аксиом логики). Пусть x_0 – одно из таких множеств. Тогда с помощью невозможного условия $\neg(y=y)$ и аксиомы $Z21\{\neg(y=y)\}$ получаем множество x со свойством

$$(\forall y)(y \in x \leftrightarrow y \in x_0 \wedge \neg(y=y)),$$

т.е. $(\forall y)\neg(y \in x)$, что и требовалось.

С другой стороны, аксиома Z1 гарантирует, что существует только **одно** пустое множество: если множества x_1, x_2 оба пусты, т.е. $(\forall y)\neg(y \in x_1)$ и $(\forall y)\neg(y \in x_2)$, то $(\forall y)(y \in x_1 \leftrightarrow y \in x_2)$, а отсюда согласно Z1 вытекает, что $x_1 = x_2$.

Для удобства записи хотелось бы ввести для пустого множества особое обозначение, например 0. Но, к сожалению, в языке теории множеств нет констант. Если мы введем одну такую константу, это не решит проблемы, поскольку затем захочется ввести еще одну константу и т.д. без конца. Оказывается, однако, что можно обойтись вообще без констант, не ограничивая возможности рассуждений.

Если мы хотим использовать в своих рассуждениях константу 0, не входящую в язык теории множеств, мы должны указать общий метод, позволяющий переводить наши утверждения, содержащие 0, в "чистую форму" языка теории множеств. И это, оказывается, совсем просто: если мы утверждаем, например, что множество 0 обладает свойством, выраженным в формуле F(x), т.е. утверждаем, что имеет место F(0), это можно выразить и с помощью формулы, не содержащей символ 0:

$$(\forall x)((\forall y)\neg(y \in x) \rightarrow F(x)),$$

т.е. если x – пустое множество, то F(x). Так как существование и единственность пустого множества мы уже доказали, никаких неточностей здесь не возникает. Допустимо, таким образом, использовать

в рассуждениях символ 0 – его всегда можно исключить, не изменяя смысл рассуждений (но усложняя их запись).

Аналогично можно поступать и в других случаях, когда удается доказать существование и единственность множества, удовлетворяющего некоторой формуле $\tau(x)$, т.е. удается доказать формулы

$$(\exists x)\tau(x),\ (\forall x_1\,\forall x_2)(\tau(x_1)\wedge\tau(x_2)\rightarrow x_1 = x_2).$$

Мы можем в таком случае ввести особое обозначение t для этого единственного объекта, и пользоваться им в своих рассуждениях. При этом всякое утверждение F(t) ("t обладает свойством F") можно сформулировать и без t: $(\forall x)(\tau(x)\rightarrow F(x))$. Вполне допустимо, что формула $\tau(x)$ содержит не только свободную переменную x, но и параметры, скажем, z_1, z_2, т.е. формула имеет вид $\tau(x, z_1, z_2)$. В этом случае, если доказано, что каждой паре z_1, z_2 соответствует единственное x, формула τ определяет по существу некоторую **операцию** над множествами. Если обозначить эту операцию через %, то $x=z_1\%z_2$ будет более наглядной записью формулы $\tau(x, z_1, z_2)$. Этим обозначением можно свободно пользоваться, поскольку утверждение, что имеет место, например, $F(z_1\%z_2)$, можно переписать без символа "%":

$$(\forall x)(\tau(x, z_1, z_2)\rightarrow F(x)).$$

Возможностью введения и исключения новых констант и операций широко пользуются в полуформальных рассуждениях теории множеств. Полная формализация означала бы исключение всех вспомогательных символов.

Упражнение 2.4. а) Докажите, что для любых двух множеств x,y их пересечение $x\cap y$ и разность $x–y$ также являются множествами. Обоснуйте (описанным выше способом) допустимость использования символов операций "$\cap$" и "–".

б) Докажите, что если $\mathbf{A}\subseteq\mathbf{B}$ и **A** – собственный класс, то и **B** – собственный класс.

Утверждение б) проливает свет на причины, почему некоторые классы являются собственными. Это их "величина" – если класс охватывает слишком большое разнообразие элементов-множеств, то его самого уже нельзя считать множеством. В частности, нельзя считать множеством **класс всех множеств**:

$$\mathbf{V}=\{y \mid y=y\},$$

поскольку $\mathbf{R}\subseteq\mathbf{V}$, где $\mathbf{R}=\{y \mid \neg(y\in y)\}$ – класс Рассела.

К сожалению, аксиомы выделения позволяют получать только "меньшие" множества из "больших". В итоге оказывается, что сами по себе они гарантируют существование только наименьшего множества – пустого. Поэтому придется ввести также увеличивающие аксиомы.

Простейшей из них является **аксиома пары**. Если x_1, x_2 – множества, то можно образовать новое множество, обозначаемое обычно через $\{x_1, x_2\}$. Чтобы узаконить такой способ рассуждения, аксиомой объявляется следующая формула:

$$(\forall x_1 \forall x_2)(\exists z)(\forall y)(y \in z \leftrightarrow y=x_1 \vee y=x_2). \qquad (Z22)$$

Ясно, что это – частный случай схемы свертывания, а именно $Z2\{y=x_1 \vee y=x_2\}$.

Очень важным элементом языка математики является понятие **упорядоченной пары**. Такую пару, состоящую из x_1 и x_2, будем обозначать через (x_1, x_2), чтобы подчеркнуть отличие от **неупорядоченной** пары $\{x_1, x_2\}$. Но каким образом ввести (x_1, x_2) в нашу теорию множеств? Ведь у нас "все есть множество"! Как добиться нужного порядка x_1 и x_2? В 1921 г. К.Куратовский предложил следующее остроумное определение упорядоченной пары через неупорядоченную:

$$(x_1, x_2) = \{\{x_1\},\{x_1, x_2\}\}.$$

Интуитивно ясно, что x_1 играет здесь отличную роль по сравнению с x_2, т.е. между ними установлен некоторый порядок.

Упражнение 2.5. Докажите корректность определения К.Куратовского:

$$\neg(x_1=y_1 \wedge x_2=y_2) \rightarrow (x_1, x_2)\neq(y_1, y_2).$$

Отсюда, в частности, вытекает, что $x_1\neq x_2 \rightarrow (x_1, x_2)\neq(x_2, x_1)$.

Теперь мы можем ввести обычное для математики понятие **декартова произведения**:

$$\mathbf{A}\times\mathbf{B} = \{(u, v) \mid u \in A \wedge v \in B\},$$

или, более точно:

$$\mathbf{A}\times\mathbf{B} = \{z \mid (\exists u \exists v)(u \in A \wedge v \in B \wedge z=(u,v)\}.$$

Хотелось бы полагать, что декартово произведение $x\times y$ двух множеств x, y окажется множеством. Этот вопрос будет решен положительно после принятия аксиомы множества подмножеств (см.

дальше).

Теперь мы можем ввести также понятие **отношения**. В самом общем случае отношением следует называть произвольный класс, состоящий только из упорядоченных пар. Т.е. класс Q={y | F(y)} будет называться отношением, если

$$(\forall y)(y \in Q \rightarrow (\exists u \exists v) y=(u, v)).$$

Если через F(u,v) обозначить формулу $(\exists y)(y=(u, v) \wedge F(y))$, то можно писать также: Q={(u, v) | F (u,v)}.

Несколько позднее мы докажем, что некоторые отношения являются собственными классами, например:

$$\mathbf{E} = \{(u, v) \mid u=v\}, \mathbf{C} = \{(u, v) \mid u \in v)\}.$$

Наша следующая аксиома должна обеспечить возможность произвольных **объединений** множеств. Кроме широко используемых конечных объединений типа $x \cup y$ в математике нужны объединения бесконечного числа множеств. Самый же общий случай – объединение произвольного "множества множеств":

$$\cup x=\{y \mid (\exists z)(y \in z \wedge z \in x)\}.$$

Таким образом, класс $\cup x$ получается, если объединить все элементы множества x, т.е. собрать вместе все "элементы элементов" x. В частности, $\cup\{x, y\}$ – это обычное $x \cup y$. Например, $\cup\{0, \{0\}\}=\{0\}$.

Аксиома объединения (или аксиома суммы) объявляет класс $\cup x$ множеством:

$$(\forall x)(\exists u)(\forall y)(y \in u \leftrightarrow (\exists z)(z \in x \wedge y \in z)) \quad (Z23)$$

Это опять частный случай схемы свертывания, а именно $Z2\{(\exists z)(z \in x \wedge y \in z)\}$.

Упражнение 2.6. а) Докажите, что если **A –** собственный класс, а x – множество, то **A**–x является собственным классом.

б) Покажите, что с помощью аксиом выделения, пары и объединения можно доказать существование любого конкретного конечного множества, построенного на базе 0, например {0, {0, {0}}, {0, {0, {0}}}}.

С помощью аксиомы объединения можно доказать, что если **A, B –** непустые классы и декартово произведение **A**×**B –** множество, то **A, B** также являются множествами. В самом деле, возьмем какой-нибудь элемент $v_0 \in \mathbf{B}$, тогда для каждого $u \in \mathbf{A}$:

$$(u,v_0) \in \mathbf{A} \times \mathbf{B},$$

$$\{\{u\},\{u,v_0\}\} \in \mathbf{A}\times\mathbf{B},$$

$$\{u\} \in \cup(\mathbf{A}\times\mathbf{B}),$$

$$u \in \cup\cup(\mathbf{A}\times\mathbf{B}).$$

Таким образом, $\mathbf{A} \subseteq \cup\cup(\mathbf{A}\times\mathbf{B})$. Поскольку $\mathbf{A}\times\mathbf{B}$ **–** множество, то по аксиоме объединения, $\cup(\mathbf{A}\times\mathbf{B})$ – множество, т.е. класс **A** содержится в множестве, поэтому сам является множеством. Аналогично доказывается, что множеством является **B**.

Упражнение 2.7. Покажите, что отношения **E**, **C**, определенные выше, являются собственными классами.

Следующий способ свертывания вошел в математику в 70-х гг. прошлого века – вместе с точными определениями системы действительных чисел на основе системы рациональных чисел. Будь то определение посредством "сходящихся в себе" последовательностей рациональных чисел (Г. Кантор), "сечений" множества рациональных чисел (Р. Дедекинд) или посредством бесконечных двоичных дробей, говоря о "произвольном" действительном числе, приходится привлекать понятие "произвольного" множества натуральных чисел. Лучше всего это видно на примере двоичных дробей: всякую бесконечную дробь, например

$$0{,}1010110011000010111...,$$

можно истолковать как характеристическую функцию некоторого множества натуральных чисел, в данном случае

$$\{1, 3, 5, 6, 9, 10, 15, 17, 18, 19, ... \}.$$

Обратное также верно. Таким образом, исследуя "произвольные" действительные числа, мы тем самым вводим в работу систему **всех** множеств натуральных чисел. И нам кажется, мы настолько хорошо знакомы с "устройством" этой системы, что можем считать его множеством.

Чтобы сделать эту новую операцию законной, нужна специальная аксиома свертывания. Через $x\subseteq y$ будем обозначать формулу

$$(\forall z)(z \in x \rightarrow z \in y),$$

т.е. формулу, утверждающую, что x есть подмножество y.

Аксиома множества подмножеств (ее называют также аксиомой степени):

$$(\forall x)(\exists z)(\forall y)(y \in z \leftrightarrow y\subseteq x). \qquad (Z24)$$

Это частный случай схемы свертывания, а именно $Z2\{y\subseteq x\}$. В терминах

классов смысл аксиомы Z24 сводится к следующему. Для всякого множества x определяется класс

$$P(x)=\{y \mid y\subseteq x\}$$

всех его подмножеств. Аксиома Z24 утверждает, что если x – множество, то P(x) – также множество.

Из аксиомы множества подмножеств вытекает, что множеством является **декартово произведение** любых двух множеств x и y:

$$x\times y=\{(u, v) \mid u \in x \wedge v \in y\}.$$

В самом деле, если $u \in x$ и $v \in y$, то $\{u\} \in P(x)$, $\{u, v\} \in P(x\cup y)$ и, следовательно, $(u, v) = \{u, \{u, v\}\} \in PP(x\cup y)$. Класс $x\times y$ содержится, таким образом, в множестве $PP(x\cup y)$. Из соответствующей аксиомы выделения тогда вытекает, что сам этот класс также является множеством.

Отношение $F=\{(u,v) \mid F(u,v)\}$ называется **функцией**, если каждому u соответствует не более одного v:

$$(u, v_1) \in F \wedge (u, v_2) \in F \rightarrow v_1 = v_2,$$

другими словами,

$$(\forall u \forall v_1 \forall v_2)(F(u, v_1) \wedge F(u, v_2) \rightarrow v_1 = v_2).$$

В этом случае будем записывать $(u, v) \in F$ в виде $F(u)=v$.

Каждой функции соответствуют **область определения** и **область значений**. Формально эти области можно определить и для произвольного отношения Q:

$$dom(Q)=\{u \mid (\exists v)(u, v) \in Q\},$$

$$rng(Q)=\{v \mid (\exists u)(u, v) \in Q\}$$

(domain, range). Если Q – класс, то, вообще говоря, dom(Q), rng(Q) также будут классами и необязательно – множествами. Например, для отношения $\mathbf{E} = \{(u, v) \mid u=v\}$ (оно является по существу тождественной функцией: $\mathbf{E}(u)=u$) имеем $dom(\mathbf{E})=rng(\mathbf{E})=\mathbf{V}$.

Упражнение 2.9. Докажите, что отношение Q является множеством, если и только если dom(Q) и rng(Q) – множества.

Функция, тождественно равная 0, т.е. $\mathbf{O}=\{(u, v) \mid v=0\}$, дает пример функции – собственного класса, область значений которой является множеством: $rng(\mathbf{O})=\{0\}$. Ясно, что область определения такой функции должна быть собственным классом (и действительно: $dom(\mathbf{O})=\mathbf{V}$). Но если известно, что F – функция, а dom(F) – множество, то может ли оказаться, что rng(F) является собственным классом? Кажется, этого не

должно быть: если область определения **функции** является множеством, то и область значений должна быть множеством. Однако, чтобы сделать такой способ рассуждения законным, нужны специальные аксиомы, так называемая схема аксиом подстановки. Э. Цермело упустил эту схему в своей первоначальной системе аксиом 1908 г. То, что она необходима при доказательстве некоторых теорем "высшей" теории множеств, заметил лишь в 1922 г. А. Френкель.

Если F – функция и **A –** класс, то через F"**A** принято обозначать **образ A** при отображении посредством F:

$$F"\mathbf{A}=\{v \mid (\exists u)(u \in A \wedge F(u)=v)\}.$$

Упражнение 2.10. Покажите, что если f – функция-множество, то для любого класса **A** образ f"**A** является множеством.

В частности, если f – функция-множество, то для любого множества a образ f"a также является множеством. Но если вместо f взять функцию-класс F? Будет ли тогда образ F"a множеством, если множеством является a? Если F является собственным классом, то в нашей теории множеств такую функцию нельзя рассматривать как вещь. В этом случае F – соответствие, заданное некоторой формулой

$$F=\{(u, v) \mid F(u, v)\}.$$

Говоря, что F – функция, мы имеем в виду только то, что формула F(u, v) сопоставляет u не более одного v. Если мы возьмем элементы u множества a и **подставим** вместо каждого такого u соответствующее ему v, то во что превратится множество a? Разумеется, в другое **множество**. Такой способ рассуждений кажется естественным, однако он редко используется в математике (возможно, поэтому Э. Цермело и не обратил на него внимания). Обычно математические функции строятся сразу как некоторые множества пар (натуральных, действительных или комплексных чисел), а не как заданные формулами соответствия, которые нельзя считать множествами. В теории множеств, однако, такие соответствия встречаются, поэтому чтобы сделать указанный выше способ рассуждения законным, следуя А. Френкелю, принимаем **схему аксиом подстановки**:

$$F \text{ – функция} \rightarrow (\forall a)(\exists b)\ F"a=b,$$

или, более точно, если формула $F(u, v, z_1, ..., z_n)$ не содержит переменных v_1, v_2, x, y, a, b, то аксиомой Z25{F} объявляется следующая формула:

$$(\forall u \forall v_1 \forall v_2)(F(u, v_1, z_1, ..., z_n) \wedge F(u, v_2, z_1, ..., z_n) \rightarrow v_1 = v_2) \rightarrow$$

$$\rightarrow(\forall a)(\exists b)(y \in b \leftrightarrow (\exists x)(x \in a \wedge F(x, y, z_1, ..., z_n))). \qquad Z25\{F\}$$

Здесь мы опять имеем дело с частным случаем схемы свертывания, а именно с аксиомами $Z2\{(\exists x)(x \in a \wedge F(x, y))\}$, однако применять эти аксиомы разрешается, только если доказано, что каждому x соответствует не более одного y такого, что F(x, y).

Упражнение 2.11. а) Покажите, что если **A –** собственный класс и a – множество, то невозможна одно-однозначная (инъективная) функция, отображающая **A** в a.

б) Покажите, что схема подстановки влечет схему выделения Z21.

в) Покажите, следуя Я. Мыцельскому, что аксиому пары можно вывести из остальных аксиом (примените дважды Z24, а затем – Z25).

Сформулированные до сих пор аксиомы сами по себе способны доказать только существование конечных множеств, построенных на базе пустого множества 0, например {{0}, {0, {0}}, {{0}}, 0} (см. упражнение 2.6б). В самом деле, все эти аксиомы выполняются в области конечных множеств (проверьте). Следуя предложению Дж. фон Неймана, некоторые из этих множеств можно использовать в качестве **натуральных чисел**: 0 можно считать нулем, {0} – единицей, {0, {0}}={0, 1} – числом 2, {0, 1, 2} – числом 3 и т.д. Если множество c_n объявлено числом n, то $c_{n+1} = c_n \cup \{c_n\}$ объявляется числом n+1. Казалось бы, мы имеем теперь и множество всех натуральных чисел:

$$\{0, 1, 2, 3, ..., n, n+1, ... \}.$$

К сожалению, так как принятые до сих пор аксиомы свертывания выполняются в области конечных множеств, то с их помощью невозможно доказать существование **множества**, состоящего из всех натуральных чисел. Непросто определить (формулой) даже **класс** всех натуральных чисел. Ведь бесконечные дизъюнкции вроде

$$x=c_0 \vee x=c_1 \vee x=c_2 \vee ... \vee x=c_n \vee...$$

недопустимы в нашем языке теории множеств. По Дж. фон Нейману, понятие натурального числа в теории множеств определяется путем сочетания двух следующих свойств:

а) **Транзитивности**. Множество x называется транзитивным, если вместе с каждым своим элементом оно содержит также все элементы этого элемента:

$$(\forall u \forall v)(u \in v \wedge v \in x \rightarrow u \in x).$$

Легко проверить, что определенные выше натуральные числа являются транзитивными множествами.

б) **Идеального порядка**. Отношение "<" называется отношением

порядка на множестве x, если

$$(\forall a)(a \in x \rightarrow \neg(a<a)),$$

$$(\forall a \forall b \forall c)(a,b,c \in x \rightarrow (a<b \wedge b<c \rightarrow a<c)).$$

Будем говорить, что "<" идеально упорядочивает x, если всякое непустое подмножество x содержит как наименьший, так и наибольший элементы. Интуитивно ясно, что в последнем случае множество x должно быть конечным. Возможность идеального порядка можно принять даже за определение конечности множества.

Дж. фон Нейман предложил определить класс **N** всех натуральных чисел следующим образом:

$$\mathbf{N} = \{y \mid y \text{ транзитивно} \wedge y \text{ идеально упорядочено отношением принадлежности } "\in"\}.$$

Упражнение 2.12. а) Покажите, что "стандартные" натуральные числа все являются элементами класса **N**, т.е. что $c_n \in \mathbf{N}$ для всех n. Заметьте, что это утверждение является **схемой теорем**, которую мы не умеем "свернуть" в одну общую теорему.

б) Докажите, что если $x \in \mathbf{N}$, $x \neq 0$ и y – наибольший элемент x, то

$$x - \{y\} \in \mathbf{N}, \; x - \{y\} \in x, \; x - \{y\} = y.$$

Выведите отсюда, что **N** содержит только одно множество, содержащее ровно n элементов – c_n.

в) (курсовая работа) Покажите, что в классе **N** выполняются аксиомы элементарной арифметики (т.е. теории EA из раздела 3.1). Ср. упражнение 2.13.

Если мы намерены считать, что натуральные числа образуют множество, то следует принять соответствующую аксиому свертывания Z2$\{y \in \mathbf{N}\}$ или

$$(\exists x)(\forall y)(y \in x \leftrightarrow y \in \mathbf{N}). \qquad (Z26)$$

Так как аксиома Z26 обеспечивает существование бесконечных множеств, ее принято называть **аксиомой бесконечности**.

Замечание. Обычно аксиома бесконечности формулируется более просто, например

$$(\exists x)(0 \in x \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x)).$$

Здесь речь идет о множестве, содержащем все натуральные числа. Из этой аксиомы можно вывести Z26 (и наоборот).

Аксиома Z26 гарантирует существование только **счетных**

множеств. Существование **несчетных** множеств вытекает (теперь, после принятия Z26) из аксиомы множества подмножеств Z24: если через ω обозначить множество всех натуральных чисел, то множество P(ω) будет несчетным.

Упражнение 2.13. Покажите, что в множестве ω выполняются аксиомы арифметики Пеано (см. раздел 3.1).

Этим завершается наш "переучет" случаев схемы свертывания, которые применяются в математике (и допустимость которых почти не подвергается сомнению).

В своем изложении философии множеств мы отметили, что для построения нашего мира множеств не требуется никакого исходного материала – все можно построить, отправляясь от пустого множества 0. Однако принятые нами до сих аксиомы не могут гарантировать, что всякое множество, о котором идет речь в нашей теории, действительно построено "из ничего". Ведь в этих аксиомах идет речь о том, какие множества **существуют**, и не сказано ничего о множествах, которые "не должны существовать". Наши аксиомы не исключают, например, существования множества x такого, что $x \in x$ (т.е. x – элемент самого себя), или существования пары множеств y,z таких, что $y \in z \wedge z \in y$.

Каким образом, однако, записать утверждение, что "всякое множество построено из ничего"? Один из возможных подходов состоит в следующем: если мы переходим от множества x_0 к его элементу x_1, затем к x_2 – элементу x_1 и т.д.:

$$... \in x_3 \in x_2 \in x_1 \in x_0 ,$$

то этот "спуск" должен где-то оборваться. Окажись он безграничным, мы получили бы множество $x = \{x_0, x_1, x_2, x_3, ...\}$, обладающее свойством: для любого $y \in x$ найдется $z \in x$ такое, что $z \in y$. В запрещении таких множеств и состоит смысл **аксиомы регулярности** (или аксиомы фундирования):

$$\neg(\exists x)(\neg(x=0) \wedge (\forall y)(y \in x \rightarrow (\exists z)(z \in x \wedge z \in y))). \quad (Z3)$$

Аксиому регулярности ввел в 1925 г. Дж. фон Нейман. Почему так поздно? Дело в том, что работая в "положительной" теории множеств, т.е. занимаясь построением различных множеств с помощью аксиом свертывания (натуральных чисел, действительных чисел, функций и т.д.) и изучением их свойств, нет надобности в постулировании несуществования прочих множеств. Все построенные множества автоматически обладают свойством, утверждаемым в аксиоме регулярности. И только переходя к исследованию **принципиальных**

возможностей аксиом теории множеств, приходится учитывать свойства, которыми должны обладать **все** множества вообще.

Упражнение 2.14. Выведите из аксиомы регулярности, что

а) **R=V** (здесь **R –** класс Рассела, **V –** класс всех множеств), т.е. покажите, что $\neg(x \in x)$ для всех x,

б) невозможны множества y, z такие, что $y \in z \wedge z \in y$.

Формальную теорию, собственными аксиомами которой являются Z1 (аксиома экстенсиональности), Z21 (схема выделения), Z22 (аксиома пары), Z23 (аксиома объединения), Z24 (аксиома множества подмножеств), Z25 (схема подстановки), Z26 (аксиома бесконечности) и Z3 (аксиома регулярности), принято называть **теорией множеств Цермело-Френкеля** и обозначать через ZF (Zermelo, Fraenkel).

Правда, сам Э. Цермело включал в свою систему аксиом еще **аксиому выбора**.

Если элементами множества x являются только непустые множества, то можно попытаться определить функцию f, которая каждому $y \in x$ сопоставляет f(y), равное некоторому $z \in y$ (поскольку y непусто). Такую функцию естественно назвать **функцией выбора** для (семейства множеств) x: в каждом из непустых членов семейства она выбирает по одному элементу. Для всякого ли семейства непустых множеств существует функция выбора? Желая всегда иметь положительный ответ на этот вопрос, мы можем принять специальную аксиому – это и есть аксиома выбора:

$$(\forall x)((\forall y)(y \in x \rightarrow y \neq 0) \rightarrow$$

$$\rightarrow(\exists f)(f \text{ – функция} \wedge \mathrm{dom}(f)=x \wedge (\forall y)(y \in x \rightarrow f(y) \in y))).$$

В 1904 г. Э.Цермело использовал этот "принцип произвольного выбора" для доказательства теоремы, согласно которой любое множество можно вполне упорядочить. (Пара $(x, <)$ называется вполне упорядоченным множеством, если всякое непустое подмножество x содержит наименьший элемент в смысле отношения "<".)

Упражнение 2.15. а) Докажите обратное утверждение: если множество $\cup x$ можно вполне упорядочить, то для x существует функция выбора.

б) Докажите, используя аксиому выбора, что каждое бесконечное множество содержит счетное подмножество.

Следует особо отметить, что аксиома выбора **не является аксиомой свертывания**: функция f, существование которой постулируется, не определяется формулой, выражающей отношение

f(u)=v. Именно этот неконструктивный характер аксиомы выбора подчеркивается в названии "принцип **произвольного** выбора" (выбор считается осуществимым несмотря на то, что никакого определенного формулой **правила** выбора мы не имеем). И именно этим ее характером были вызваны протесты французских математиков после опубликования работы Э. Цермело в 1904 г. Главными оппонентами были Э. Борель, А. Лебег и Р. Бэр. По их мнению, бесконечное число актов произвольного выбора – это способ рассуждения, выходящий за пределы математики и относящийся скорее к области черной магии (более подробно об истории аксиомы выбора см. книгу Ф. А. Медведева [1982]).

И сейчас среди математиков нет единства по вопросу, следует принять аксиому выбора или отвергнуть ее. Поэтому для формальной теории множеств, в которой наряду с аксиомами ZF принимается и аксиома выбора, используется особое обозначение ZFC (от *axiom of choice*). В теории ZFC можно воспроизвести все результаты интуитивной теории множеств Г. Кантора и его последователей (см. книгу Т. Йеха [1973]). И поскольку остальные математические теории могут быть воспроизведены средствами теории множеств (это было осознано еще в конце XIX в.), то ZFC – формальная теория, которая охватывает всю математику. **Так был выполнен первый этап программы Гильберта.**

2.4. Вокруг проблемы континуума

Мы уже видели, каким образом Г. Кантор пришел в 1877 г. к проблеме континуума. Изучая самые различные, в том числе очень сложные бесконечные множества точек, он обнаружил только два типа таких множеств:

– счетные множества (их элементы можно перенумеровать с помощью натуральных чисел),

– множества, эквивалентные отрезку прямой (их можно взаимно однозначно отобразить на отрезок прямой, в этом смысле они имеют мощность континуума).

Поэтому Г. Кантор решил, что всякое бесконечное множество точек либо должно быть счетным, либо имеет мощность континуума, и попытался свое предположение доказать. Однако это ему не удалось ни тогда, ни в последующие 10 и 20 лет. Не удалось это и его многочисленным последователям. В том смысле, как понимал ее Г. Кантор, континуум-проблема не решена до сих пор.

Работая над доказательством континуум-гипотезы Г. Кантор

создал **теорию порядковых чисел** (или – как сегодня принято говорить – **ординалов**). Порядковые числа – обобщение натуральных чисел. Если натуральные числа используются для подсчета элементов конечных множеств, то порядковые числа – для подсчета элементов любых множеств (в том числе бесконечных).

К идее таких обобщенных чисел Г. Кантора привели еще исследования сходимости рядов Фурье (см.раздел 2.1): определив производные множества конечных порядков

$$P', P'', P''', ..., P^{(n)}, ...,$$

он сообразил, что пересечение этих множеств является "производным множеством бесконечного порядка", которое можно обозначить через $P^{(\omega)}$, где ω – первое бесконечное порядковое число. Но если от $P^{(\omega)}$ взять производную первого порядка, то результат естественно обозначить уже через $P^{(\omega+1)}$, дальше следует $P^{(\omega+2)}$, $P^{(\omega+3)}$ и т.д. За всеми этими "числами" следует, разумеется, $\omega+\omega$ (или $\omega\cdot 2$). А еще дальше:

$\omega\cdot 2+1, \omega\cdot 2+2, ..., \omega\cdot 3, \omega\cdot 3+1, ...,$

$\omega\cdot\omega, \omega\cdot\omega+1,...$

Однако определить общее понятие о бесконечных порядковых числах непросто. Вот как это делается в теории ZF (следуя идеям Дж.фон Неймана).

Вспомним (см.раздел 2.3), что натуральными числами мы назвали транзитивные множества, идеально упорядоченные отношением принадлежности "$\in$", т.е. x – натуральное число, если

а) $(\forall z \forall y)(z \in y \in x \rightarrow z \in x)$ (транзитивность),

б) отношение "$\in$" является отношением порядка на x,

в) всякое непустое подмножество x имеет как наибольший, так и наименьший элементы в смысле отношения "$\in$".

Если отбросить в п. в) требование о существовании наибольшего элемента (оставив только наименьший), то получится определение порядкового числа (или ординала) по Дж. фон Нейману: **ординал –** это транзитивное множество, вполне упорядоченное отношением принадлежности "$\in$".

Упражнение 2.16. Напишите условие "x – ординал" в виде формулы языка теории множеств.

Таким образом, ординалы образуют класс, который принято обозначать через **On** ("Ordinal numbers"). Сами ординалы будем обозначать через α, β, γ и т.д.

Естественным отношением порядка среди ординалов является “$\in$”:

$$\alpha<\beta \leftrightarrow \alpha \in \beta.$$

Покажем, что

$$\alpha \in \beta \wedge \text{"}\beta \text{ – ординал"} \rightarrow \text{"}\alpha \text{ – ординал"},$$

т.е. элементами ординала (как множества) могут быть только ординалы. Для этого мы должны проверить, что α – транзитивное множество, вполне упорядоченное отношением “$\in$”. Транзитивность α следует из транзитивности β и того, что “$\in$” является отношением порядка на β:

$$x \in y \in \alpha \in \beta \rightarrow y \in \beta \wedge x \in \beta,$$

поэтому $x \in y \wedge y \in \alpha \rightarrow x \in \alpha$. Вполне упорядоченность α вытекает из транзитивности и вполне упорядоченности β: из $\alpha \in \beta$ следует, что α – подмножество β (транзитивность β), поэтому если x – непустое подмножество α, то x – непустое подмножество β, т.е. x имеет наименьший элемент в смысле отношения “$\in$”.

Отсюда для любого ординала α: $\alpha=\{\beta \mid \beta<\alpha\}$. Это обобщение равенства $n = \{0, 1, 2, ..., n–1\}$ (n – натуральное число).

Упражнение 2.17. Проверьте, что

а) если α – ординал, то $\alpha\cup\{\alpha\}$ – также ординал (причем – наименьший, превосходящий α; вместо $\alpha\cup\{\alpha\}$ принято писать $\alpha+1$),

б) каждый непустой класс ординалов содержит наименьший элемент,

в) если x – множество ординалов, то $\cup x$ – также ординал (причем – точная верхняя грань x).

Отсюда, в частности, следует, что **On** – собственный класс (в этом состоял смысл парадокса Бурали-Форти: если предположить, что On – множество, то получается противоречие).

Ординал α называется **последующим ординалом**, если $\alpha=\beta+1$ для некоторого β. В противном случае α называют **предельным ординалом**. Таким образом, наименьший предельный ординал – 0, второй предельный ординал принято обозначать через ω.

Упражнение 2.18. а) Проверьте, что ординалы, меньшие ω, – натуральные числа (т.е. ω – множество всех натуральных чисел).

б) Проверьте, что если α – предельный ординал, то $\alpha=\cup\alpha$.

Теперь легко доказать **принцип трансфинитной индукции**: если **A –** класс ординалов такой, что

а) $0 \in \mathbf{A}$,

б) $\alpha \in \mathbf{A} \rightarrow \alpha+1 \in \mathbf{A}$,

в) $x \subseteq \mathbf{A} \rightarrow \cup x \in \mathbf{A}$,

то $\mathbf{A}=\mathbf{On}$. Это обобщение принципа математической индукции (дополнительно появился п. в)). В самом деле, если $\mathbf{A} \neq \mathbf{On}$, то пусть β – наименьший ординал, не входящий в **A**. Ясно, что $\beta \neq 0$. Если $\beta=\alpha+1$, то $\alpha \in \mathbf{A}$ и поэтому $\beta \in \mathbf{A}$. Если же β – предельный ординал, то все $\alpha<\beta$ принадлежат **A**, т.е. $\beta = \{\alpha \mid \alpha<\beta\} \subseteq \mathbf{A}$ и $\cup\beta \in \mathbf{A}$. Но $\beta=\cup\beta$. Теорема доказана.

Упражнение 2.19. Докажите **теорему о трансфинитной рекурсии**: если **G** – функция-класс, то существует единственная функция-класс **F**, такая, что $\mathrm{dom}(\mathbf{F})=\mathbf{On}$ и для всех $\alpha \in \mathbf{On}$ $\mathbf{F}(\alpha)=\mathbf{G}(\mathbf{F}|\alpha)$. Через $\mathbf{F}|\alpha$ принято обозначать ограничение функции **F** на "область" α, т.е.

$$\mathbf{F}|\alpha = \{(u, v) \mid u \in \alpha \wedge (u, v) \in \mathbf{F}\}.$$

Функция **G** играет здесь роль шага рекурсии: если известны значения **F** для всех $x<\alpha$ (т.е. все значения функции $\mathbf{F}|\alpha$), то **G** позволяет "вычислить" $\mathbf{F}(\alpha)$.

Если натуральные числа позволяют "пересчитать" любое идеально упорядоченное (конечное) множество, то ординалы позволяют "пересчитать" любое вполне упорядоченное множество, т.е. множество, на котором введено отношение порядка такое, что любое непустое подмножество имеет наименьший элемент. Оказывается, что каждое такое множество изоморфно единственному ординалу.

Упражнение 2.20. Докажите это с помощью трансфинитной рекурсии.

Каждое натуральное число может использоваться для "пересчета" (упорядочения) конечных множеств, но оно может использоваться и для измерения количества элементов множества. С бесконечными порядковыми числами ситуация сложнее: например, ординалы

$$\omega,\omega+1,\omega+2,...,\omega\cdot 2,...,\omega\cdot\omega,...$$

соответствуют различным упорядочениям счетных множеств. В самом деле, $\omega+2$ соответствует порядку

$$2, 3, 4, 5,..., 0, 1,$$

а $\omega\cdot 2$ – порядку

$$1, 3, 5, 7, 9,..., 2, 4, 6, 8, 10,...$$

Наконец, $\omega\cdot\omega$ соответствует матрице из бесконечного числа бесконечных строк. Все это – только счетные множества. Для измерения количества элементов в них естественно использовать один и тот же наименьший бесконечный ординал ω.

Таким образом, целесообразно следующее определение: ординал α называется **кардиналом**, если α нельзя отобразить одно-однозначно ни на какой ординал $\beta<\alpha$. Именно кардиналы естественно использовать для измерения количества элементов бесконечных множеств. Легко видеть, что все натуральные числа – кардиналы, а ω – наименьший бесконечный кардинал. А дальше – за ω – существуют еще кардиналы?

Используя аксиому множества подмножеств Z24 и схему подстановки Z25, можно доказать, что за каждым кардиналом k существует кардинал $l>k$, т.е. кардинал, который невозможно одно-однозначно отобразить на k (это теорема Ф. Хартогса, доказанная в 1915 г.). В самом деле, покажем, что все ординалы, которые можно одно-однозначно отобразить на k, образуют множество. Отсюда вытекает (поскольку **On** – собственный класс), что должны существовать кардиналы, большие k.

Итак, рассмотрим сначала все отношения на множестве k. Каждое такое отношение – подмножество $k\cdot k$, т.е. элемент множества $P(k\cdot k)$. Нетрудно написать формулу, выделяющую среди всех отношений те, которые вполне упорядочивают k. Из схемы выделения Z21 вытекает, что все эти отношения образуют множество $z\subseteq P(k\cdot k)$. Мы уже знаем, что каждому элементу $r \in z$ соответствует единственный ординал α, такой, что вполне упорядоченное множество (k, r) изоморфно $(\alpha, \in)$. Определив это соответствие в виде формулы $F(r, \alpha)$, можно применить схему подстановки Z25, сделав вывод, что F"z – множество. Но F"z – это как раз все ординалы, которые можно взаимно однозначно отобразить на k (ведь каждый такой ординал определяет некоторое вполне упорядочение множества k). Тем самым теорема Ф. Хартогса доказана.

Упражнение 2.21. Напишите обе формулы, о которых шла речь в доказательстве.

Таким образом, последовательность кардиналов оказывается неограниченной. Для обозначения бесконечных кардиналов принято использовать первую букву еврейского алфавита – $\aleph$ (алеф) с индексами. Так, $\aleph_0$ обозначает ω, т.е. первый бесконечный кардинал (счетная мощность). Далее следует $\aleph_1$ – первая несчетная мощность, $\aleph_2$ – вторая несчетная мощность и т.д. А за всеми $\aleph_n$ (n – натуральные числа) следует новый кардинал – $\aleph_\omega$.

Упражнение 2.22. а) Проверьте, что $\aleph_\omega = \cup\{ \aleph_n \mid n \in \omega\}$.

б) Докажите более общий результат: если x – множество, состоящее только из кардиналов, то $\cup x$ – кардинал.

Имея эти результаты, можно определить $\aleph_\alpha$ для каждого

ординала α:

$\aleph_0 = \omega$,

$\aleph_{\alpha+1}$ = наименьший кардинал, превосходящий $\aleph_\alpha$,

$\aleph_\alpha = \cup\{ \aleph_\beta \mid \beta<\alpha\}$, если α – предельный ординал.

Упражнение 2.23. Проверьте, что для каждого кардинала k существует ординал α такой, что $k = \aleph_\alpha$ (т.е. что алефы исчерпывают все кардиналы).

Таков аппарат, созданный Г. Кантором для измерения "мощности" (количества элементов) бесконечных множеств. Правда, алефы подходят для измерения мощности только тех бесконечных множеств, которые можно вполне упорядочить. Каждое ли бесконечное множество можно вполне упорядочить? Г. Кантор считал, что можно... и что это не требует доказательства. Сегодня мы знаем, что утверждение о возможности вполне упорядочить любое бесконечное множество равносильно аксиоме выбора (см. выше).

Как выглядит в свете созданного аппарата континуум-проблема? Если мы принимаем аксиому выбора (работаем в теории ZFC), то множество всех действительных чисел можно вполне упорядочить и поэтому его мощность (обозначим ее через c) можно измерить с помощью какого-либо $\aleph_\alpha$:

$$(\exists\alpha)\ c = \aleph_\alpha$$

(причем мы уже знаем, что α>0). Континуум-гипотеза утверждает, что всякое бесконечное множество действительных чисел либо является счетным (имеет мощность $\aleph_0$), либо имеет мощность континуума (мощность c). В таком случае на шкале алефов между $\aleph_0$ и c никакие мощности находиться не могут и поэтому $c = \aleph_1$. Так просто формулируется континуум-гипотеза в терминах созданного Г. Кантором аппарата алефов. Разумеется, это могло только усилить веру Г. Кантора в близость окончательного решения континуум-проблемы...

Однако только в 1905 г. Й. Кениг сумел доказать, что c не равно $\aleph_\omega$ (и далее: $c \neq \aleph_\alpha$, если $\alpha=\omega\cdot2, \omega\cdot3,..., \omega\cdot\omega,...$ и вообще, если α – счетный предельный ординал). По существу это все, что известно до сих пор. Никто не сумел доказать ни $c \neq \aleph_2$, ни $c \neq \aleph_3$ и т.д.

Теперь мы знаем, что эти трудности не случайны. Начало решения загадки принадлежит К. Геделю, который доказал в 1939 г., что континуум-гипотеза, если ее принять без доказательства, не создает новых противоречий. Более точно, если через Con(T) обозначить утверждение "теория T непротиворечива" (Con – consistent), то результат

К. Геделя выглядит так:

$$Con(ZF) \rightarrow Con(ZFC+"c = \aleph_1 ").$$

Таким образом, если бы принятие аксиомы выбора и континуум-гипотезы привело к противоречиям, то противоречие можно было бы найти уже в теории ZF. К. Гедель доказал, таким образом, и "безопасность" весьма сомнительной аксиомы выбора! Идея, использованная К. Геделем, в общей форме была предложена Д. Гильбертом. Исходя из того факта, что несмотря на всевозможные ухищрения никому не удается **построить** множество точек с мощностью между $\aleph_0$ и $\aleph_1$, Д. Гильберт предложил попытаться доказать, что такие множества и нельзя построить (может быть, они "существуют", но их нельзя построить). Прошли годы, и только в 1939 г. К. Гедель сумел реализовать эту идею.

К. Гедель определяет последовательность множеств $\{L_\alpha \mid \alpha \in \mathbf{On}\}$ (наше изложение следует книге К. Дэвлин [1977]):

$L_0 = 0$,

$L_{\alpha+1} = Def(L_\alpha)$ (т.е. все множества, которые можно определить, используя L_α, см. дальше),

$L_\alpha = \cup\{L_\beta \mid \beta<\alpha\}$, если α – предельный ординал.

Класс $\mathbf{L} = \cup\{L_\alpha \mid \alpha \in \mathbf{On}\}$ называется классом **конструктивных множеств** (можно показать, что $\mathbf{On} \subseteq \mathbf{L}$, поэтому **L** оказывается собственным классом). Почему элементы класса **L** следует считать конструктивными множествами? "Секрет" – во второй строчке определения (третья строчка ничего нового не создает – она только собирает вместе все построенное на предыдущих этапах). Что такое $Def(L_\alpha)$, или в общем случае – Def(m) для множества m? Это все **определимые** подмножества m (т.е. $Def(m) \subseteq P(m)$). Определимые с помощью формул языка теории множеств исходя из m. Кванторы в этих формулах должны быть ограничены множеством m, т.е. они могут использоваться только в двух следующих контекстах:

$$... (\forall x)(x \in m \rightarrow ...)... \text{ или короче: } ...(\forall x \in m)...,$$

$$... (\exists x)(x \in m \wedge ...)... \text{ или короче: } ...(\exists x \in m)...$$

Если формула содержит свободные переменные (параметры), то они могут принимать в качестве значений только элементы из m. Например, формула F(y, z):

$$y=z \vee (\exists x \in m)(\forall u \in m)(u \in y \leftrightarrow u=x)$$

для каждого значения параметра z, принадлежащего m, определяет подмножество m:

$$\{y \mid y \in m \wedge F(y,z)\}.$$

Элементы m, которые совпадают с z или являются одноэлементными множествами. Разумеется, все это вполне корректно только в случае, когда m – транзитивное множество (т.е. множество, содержащее также элементы своих элементов). Поскольку формул в языке теории множеств существует только счетное число, то, функцию Def(m) можно определить абсолютно корректно (один из вариантов см. в книге Т. Йеха [1973]).

К. Гедель показал, что если к теории ZF присоединить в качестве аксиомы утверждение V=L ("все множества конструктивны"), то можно доказать (как теоремы) и аксиому выбора, и (как ожидал Д. Гильберт) – континуум-гипотезу ($c = \aleph_1$). Далее, К. Гедель показал, что

$$Con(ZF) \rightarrow Con(ZF+"V=L").$$

Отсюда и вытекает, что и "скандальную" аксиому выбора, и континуум-гипотезу можно принять в качестве аксиом, не боясь впасть в противоречие. Правда, континуум-гипотеза не может претендовать на роль настоящей аксиомы – настолько бедны следствия, которые можно получить из нее. А утверждение V=L ("все множества конструктивны"), напротив, оказывается, настолько богато следствиями, что его стали называть **аксиомой конструктивности**.

Результат К. Геделя, полученный в 1939 г., не противоречил надеждам Г. Кантора (умершего в 1918 г.) на решение континуум-проблемы. Однако прошло еще 25 лет и американский математик Поль Коэн доказал (в 1963 г.), что эти надежды все же беспочвенны:

$$Con(ZF) \rightarrow Con(ZFC+"c = \aleph_2 "),$$

$$Con(ZF) \rightarrow Con(ZFC+"c = \aleph_3 "),$$

...

и вообще

$$Con(ZF) \rightarrow Con(ZFC+"c = \aleph_{\alpha+1} ")$$

для любого конечного или счетного ординала α. Таким образом, мы можем, не впадая в противоречие, принять в качестве аксиомы любое из следующих утверждений:

$$c = \aleph_1 , c = \aleph_2 , c = \aleph_3 ,...$$

и даже (как пошутил однажды Н. Н. Лузин) $c = \aleph_{17}$. Соединяя вместе доказанное К. Геделем и П. Коэном, приходится констатировать, что

аксиомы ZF, даже если присоединить к ним аксиому выбора, для решения континуум-проблемы недостаточны. Как оценивать этот вывод?

Примечание. Метод Коэна позволяет доказать также независимость аксиомы выбора от аксиом ZF, причем в очень сильной форме (см. Т. Йех [1973]):

$$\mathrm{Con(ZF)} \rightarrow \mathrm{Con(ZF+Q)},$$

где Q – следующее утверждение: существует счетное множество x, состоящее из неупорядоченных пар (элементами пар являются множества действительных чисел), такое, что для x не существует функции выбора. Таким образом, из аксиом ZF невозможно вывести существование функции выбора даже для счетного множества пар!

Итак, как действовать в ситуации, когда аксиомы теории множеств оказались недостаточными для решения очень важной проблемы (а также ряда других проблем, см. дальше)? Типичную реакцию работающего математика на такую ситуацию показывает следующее высказывание Н. Н. Лузина, сделанное в 1927 г. (цит. по статье Л. В. Келдыш [1974]):

"Мощность континуума, если только мыслить его как множество точек, есть единая некая реальность, и она должна находиться на алефической шкале **там, где она на ней есть** (подчеркнуто мною. – К.П.), нужды нет, если определение этого места затруднительно или, как прибавил бы Ж. Адамар, даже невозможно для нас, людей".

$$\aleph_0 \quad \aleph_1 \quad \aleph_2 \quad \ldots \quad \aleph_n \quad \aleph_{n+1} \quad \ldots \quad \aleph_\omega \quad \ldots$$

|_____|_____|____... _|_____|____... _|_____...

Аксиомы теории множеств не позволяют решить, **где именно** на алефической шкале находится мощность континуума, хотя они и позволяют доказать, что она на этой шкале находится. Математик-платоник (см. раздел 1.1), глядя на геометрический образ алефической шкалы, ищет это место **глазами**! Он не может представить себе ситуацию, когда точка находится на прямой (доказано, что находится!), однако определить точное ее местоположение невозможно. Это нормальный платонизм работающего математика, стимулирующий занятие проблемами любой сложности – ведь заранее никогда неизвестно, разрешима проблема или нет. Однако, переходя к решению **методологических** вопросов, например, о значении результатов П. Коэна, уже нельзя давать волю платонистским привычкам (полагая, что несмотря на неразрешимость проблемы континуума "для нас, людей", определенное место на алефической шкале для мощности континуума все же "объективно" существует). Это означает допускать существование мира идей ("мира множеств"), не зависящего от аксиом теории множеств,

используемых в рассуждениях математиков. Тогда платонизм математический превращается в платонизм философский. Такие люди утверждают, что аксиомы теории множеств не отражают адекватно "подлинный мир множеств", что надо искать "более адекватные" аксиомы, и даже – что никакая фиксированная система аксиом не в состоянии представить мир множеств полностью. Это погоня за миражами – никакого "подлинного мира множеств", не зависящего от аксиом, с помощью которых он исследуется, разумееется, не существует. Корректная же оценка результатов П. Коэна состоит в следующем.

Итак, доказана теорема $(\exists\alpha)\ c = \aleph_\alpha$, но невозможно определить конкретное значение α. Этот эффект свидетельствует о **внутреннем несовершенстве** традиционной теории множеств (а не о ее неадекватности "миру множеств"). Возможно, следует заняться совершенствованием аксиом теории. И оказывается, что вариантов развития здесь может быть несколько.

Значительный интерес представляет **аксиома конструктивности** (V=L). Она позволяет решить не только континуум-проблему (доказать, что $c = \aleph_1$), но и другие проблемы, недоступные аксиомам ZF (даже вместе с аксиомой выбора). В качестве второго примера рассмотрим проблему, сформулированную русским математиком М. Суслиным в 1920 г.

Проблема Суслина. Пусть (p,<) – упорядоченное множество, причем

а) p не имеет ни наименьшего, ни наибольшего элементов,

б) порядок "<" – плотный (т.е. если x<y, то существует z такое, что x<z<y),

в) множество p – полное (т.е. всякое непустое ограниченное подмножество p имеет наименьшую верхнюю грань и наибольшую нижнюю грань),

г) каждое семейство попарно непересекающихся интервалов p не более чем счетно.

Этими свойствами обладает множество всех действительных чисел. М. Суслин предположил, что всякое упорядоченное множество, обладающее свойствами а) – г), должно быть изоморфно множеству всех действительных чисел (гипотеза Суслина).

Проблема Суслина, казалось бы, касается самих основ "природы" действительных чисел (так же как континуум-проблема). Однако и эта проблема неразрешима в системе аксиом ZFC (см. Т. Йех [1973]).

Упражнение 2.24. Покажите, что проблема Суслина сводится к

вопросу: всякое ли упорядоченное множество, обладающее свойствами а) – г), содержит счетное плотное подмножество (среди действительных чисел эту роль играют рациональные числа)?

Из аксиомы конструктивности вытекает отрицательное решение проблемы Суслина: существует упорядоченное множество, обладающее свойствами а) – г), которое не содержит счетных плотных подмножеств (и поэтому не изоморфно множеству всех действительных чисел). Это доказал в 1968 г. Р. Енсен (см. К. Дэвлин [1977]).

Аналогично, аксиома конструктивности приводит к решению и других проблем, которые недоступны ZFC. Ряд таких проблем рассматривается в книге К. Дэвлин [1977]. Теория множеств ZF+"V=L" оказывается, таким образом, весьма богатой. Однако многие математики не согласны считать V=L "полноценной" аксиомой. Во-первых, потому, что развернутая ее запись на языке теории множеств содержит несколько тысяч символов, во-вторых, потому, что "неочевидно", почему все множества "должны быть" конструктивными.

Проблема. Мы уже видели, что аксиому бесконечности Z26: (∃x) x=**N** (ее развернутая запись также оказывается достаточно длинной) можно заменить очень короткой аксиомой (см. раздел 2.3). Насколько коротким можно сделать аналог аксиомы конструктивности?

Аксиома конструктивности кажется подозрительной и противникам аксиомы выбора: обе эти аксиомы позволяют доказать, что множество всех действительных чисел можно вполне упорядочить. Им представляется, что такой результат противоречит "самой природе" действительной прямой. Причем из V=L выводится гораздо более сильный результат: можно написать формулу F(α, x), которая задает функцию, отображающую **On** на **V**, т.е. функцию, которая "пересчитывает" класс всех множеств.

Другой путь развития теории множеств, позволяющий преодолеть слабости теории ZFC, был предложен в 1962 г. польскими математиками Я. Мыцельским и Г. Штейнгаузом – так называемая **аксиома детерминированности** (наше изложение следует книге В. Г. Кановея [1984]).

Каждое множество A последовательностей натуральных чисел (функций типа ω→ω) задает некоторую **игру**. Игрок 1 пишет натуральное число a_0. Затем игрок 2 пишет a_1, а игрок 1 – число a_2 и т.д. Цель игрока 1 – добиться, чтобы полученная в результате бесконечная последовательность

$$a_0, a_1, a_2, ..., a_n,...$$

оказалась элементом множества A (тогда игрок 1 считается победителем). Игрок 2 преследует прямо противоположную цель – чтобы последовательность не стала элементом A.

Стратегией принято называть функцию S, которая каждой конечной последовательности a_0, a_1, ..., a_n сопоставляет натуральnное число $b = S(a_0, a_1, ..., a_n)$. В игре A стратегия S называется **выигрышной** для игрока 1, если при любой игре игрока 2:

$$a_1, a_3, a_5, a_7, a_9,...$$

последовательность

$a_0 = S(\lambda)$ (λ – пустая последовательность),

a_1,

$a_2 = S(a_0, a_1)$,

a_3,

$a_4 = S(a_0, a_1, a_2, a_3)$,

...

...

принадлежит множеству A. Аналогично можно определить понятие выигрышной стратегии для игрока 2.

Множество A называется **детерминированным**, если один из игроков имеет выигрышную стратегию в игре A.

Упражнение 2.25. Покажите, что если A – конечное или счетное множество, то игрок 2 имеет выигрышную стратегию. Таким образом, все счетные множества оказываются детерминированными.

Можно доказать детерминированность и других типов множеств. Однако, используя аксиому выбора, можно доказать существование **недетерминированного** множества. Но поскольку это никому еще не удалось сделать без аксиомы выбора, то многие специалисты полагают, что утверждение

"каждое множество детерминировано"

(это и есть **аксиома детерминированности** – AD) не противоречит аксиомам ZF (хотя она и противоречит аксиоме выбора).

Важным аргументом в пользу "естественности" AD может служить ее запись в следующей бесконечнокванторной форме (см. В. Г. Кановей [1984]):

$$(\exists a_0)(\forall a_1)(\exists a_2\ ) \ldots (a_0, a_1, a_2, \ldots) \in A) \vee$$
$$\vee\ (\forall a_0)(\exists a_1)(\forall a_2) \ldots \neg(a_0, a_1, a_2, \ldots) \in A).$$

Первая часть "формулы" выражает существование выигрывающей стратегии для игрока 1, вторая часть – для игрока 2. Если вторую часть преобразовать к виду

$$\neg(\exists a_0)(\forall a_1)(\exists a_2) \ldots (a_0, a_1, a_2, \ldots) \in A,$$

то "формула" принимает вид закона исключенного третьего, т.е. становится "очевидной".

Из AD можно вывести **счетную аксиому выбора**, утверждающую, что функция выбора существует для любого счетного семейства непустых множеств действительных чисел (но AD противоречит **полной** аксиоме выбора!). Это очень важно, поскольку без счетной аксиомы выбора не может обойтись даже математический анализ. Без нее невозможно было бы доказать, например, что каждое бесконечное множество действительных чисел содержит счетное подмножество и что объединение счетного множества счетных множеств является счетным.

Из аксиомы детерминированности можно вывести, что **каждое множество действительных чисел измеримо по Лебегу**. Существование неизмеримого множества (известный пример Витали) можно доказать, используя аксиому выбора.

Из AD вытекает также **континуум-гипотеза** в следующей форме: каждое бесконечное множество действительных чисел либо является счетным, либо имеет мощность континуума. Однако, в теории ZF+AD невозможно доказать вполне упорядочиваемость множества всех действительных чисел (из AD вытекает, что всякое вполне упорядоченное множество действительных чисел является не более чем счетным). Поэтому мощность континуума оказывается здесь **несравнимой** с алефами (кроме $\aleph_0$).

Таким образом, оба направления (аксиома конструктивности и аксиома детерминированности) уже дали богатый набор красивых и интересных результатов. Эти две теории множеств ничуть не хуже традиционной теории, но они противоречат друг другу. Таким образом, о приближении к **единственному** "подлинному миру множеств" здесь не может быть и речи.

3. ЭЛЕМЕНТАРНАЯ АРИФМЕТИКА

3.1. От аксиом Пеано до аксиом элементарной арифметики

Попытаемся получить независимую формализацию средств математического рассуждения, использующих только понятие натурального числа (не упоминая ни действительных чисел, ни – тем более – произвольных множеств Кантора). Это самая надежная часть арсенала математики, не скомпрометировавшая себя парадоксами. Естественно назвать ее **элементарной арифметикой** (по аналогии с термином "элементарная теория чисел" – в отличие от аналитической теории чисел). Именно эти средства Д. Гильберт хотел использовать для доказательства непротиворечивости всей математики.

Итак, мы будем строить формальную теорию EA, областью изменения переменных которой будет множество {0,1,2,3,...}.

В 1880-х гг. Р. Дедекинд и Дж. Пеано сформулировали аксиомы арифметики, используя единственную константу 0 (нуль) и функциональный символ S (функция следования: S(x)=x+1):

P1) $(\forall x)\neg(0=Sx)$ (нуль не следует ни за каким числом),

P2) $(\forall x\forall y)(x\neq y \rightarrow Sx\neq Sy)$ (однозначность функции следования),

P3) если F – любое "свойство" натуральных чисел, то

$$F(0) \wedge (\forall x)(F(x)\rightarrow F(Sx)) \rightarrow (\forall x)F(x)$$

(принцип математической индукции).

Если бы мы занимались построением арифметики натуральных чисел в теории множеств ZF (см. раздел 2.3), то могли бы определить 0 как пустое множество, Sx – как $x\cup\{x\}$ и, взяв за основу множество ω, могли бы доказать аксиомы Пеано как теоремы ZF. Вместо свойства F фигурировало бы тогда произвольное подмножество ω:

$$(\forall z)(z\leq\omega \wedge 0 \in z \wedge (\forall x)(x \in z \rightarrow x\cup\{x\} \in z) \rightarrow z=\omega).$$

Но как быть, если мы не хотим вводить в рассуждения понятие произвольного множества натуральных чисел (произвольного подмножества ω), что связано с выходом в теорию множеств Кантора? По-видимому, мы должны при формулировке принципа математической

индукции ограничиться свойствами натуральных чисел, которые можно записать формулами нашей теории ЕА. Но оказывается, что в таком случае "способности" теории зависят от выбора языка (точнее – от выбранного набора функциональных символов).

Если взять по минимуму (как Дж.Пеано) только константу 0 и функциональный символ S, то понятие терма теории ЕА (для большей точности будем писать EA_0) определяется очень просто:

а) константа 0 и любая переменная – терм EA_0,

б) если t – терм, то St – также терм EA_0.

Таким образом, имеется только два типа термов: SS... S0 (обозначения конкретных натуральных чисел) и SS... Sx (обозначения функций типа x+n, где n – конкретное натуральное число). Элементарными формулами EA_0 будем считать только равенства ($t_1 = t_2$), где t_1, t_2 – термы ЕА (таким образом, EA_0 не имеет собственных предикатных символов). Из элементарных формул с помощью логических связок $\neg$, $\wedge$, $\vee$, $\rightarrow$ и кванторов $\forall$, $\exists$ строятся более сложные формулы EA_0. В качестве аксиом EA_0 принимаются аксиомы P1, P2, а вместо P3 – **схема аксиом**:

$$B(0) \wedge (\forall x)(B(x) \rightarrow B(Sx)) \rightarrow (\forall x)B(x),$$

где B(x) – формула на языке EA_0 с одной выделенной свободной переменной x. Кроме x формула B может содержать и другие свободные переменные (параметры).

Теория EA_0 оказывается очень слабой.

Упражнение 3.1. (курсовая работа). Покажите, что в теории EA_0 невозможно определить операцию сложения натуральных чисел – невозможна формула, означающая x+y=z, т.е.формула PLUS(x,y,z) такая, что

$EA_0 \vdash PLUS(x,0,x)$, (x+0=x)

$EA_0 \vdash (\forall u)(PLUS(x,y,u) \rightarrow PLUS(x,Sy,Su))$. (x+Sy=S(x+y))

Найдите также алгоритм, позволяющий по замкнутой формуле языка ЕА определить, доказуема ли она на основе аксиом EA_0 (и убедитесь, что всякая такая формула либо доказуема, либо опровержима) (см. Д. Гильберт, П. Бернайс [1934]). Возможна ли аналогичная формула, выражающая отношение x<y?

Таким образом, теория EA_0 не может представлять серьезного

интереса – в ней невозможно обсуждать даже операцию сложения натуральных чисел. Простейший шаг вперед – добавить к языку EA_0 "недостающий" функциональный символ "+" для операции сложения. В этом случае определение терма (новой теории, которую будем обозначать через EA_1) выглядит так:

а) константы 0,1 и любая переменная – термы EA_1 (вместо символа S вводим константу 1, поскольку операция сложения позволяет получить Sx в виде x+1),

б) если t_1 , t_2 – термы EA , то (t_1+t_2) – также терм EA_1. Соответственно мы должны изменить и дополнить аксиомы:

P1) $\neg(0=x+1)$,

P2) $\neg(x=y) \rightarrow \neg(x+1=y+1)$,

P3) $x+0=x$,

P4) $x+(y+1)=(x+y)+1$,

P5) $B(0) \wedge (\forall x)(B(x)\rightarrow B(x+1)) \rightarrow (\forall x)B(x)$.

Здесь B(z) – формула уже на языке EA_1, т.е. схема аксиом индукции в теории EA_1 должна быть "сильнее" схемы индукции в EA_1.

Это действительно так. Теория EA_1 уже достаточно сложна. Так, например, в ней легко определить формулой отношение x<y:

$$x<y \leftrightarrow (\exists z)x+z+1=y.$$

К 1929 г. теория EA была подробно исследована М. Пресбургером, поэтому ее принято называть **арифметикой Пресбургера**. Прежде всего М. Пресбургер построил алгоритм, который позволяет для каждой замкнутой формулы на языке EA_1 определить, "истинна" ли она при нашем естественном понимании формул EA_1. Имея этот алгоритм, можно показать, что всякая замкнутая формула на языке EA_1 либо доказуема, либо опровержима на основе аксиом EA_1. Одновременно была доказана непротиворечивость EA_1, причем в этом доказательстве были использованы только средства, разрешенные в программе Гильберта. Это были уже весьма нетривиальные результаты. Полученные в 1929 г., они еще больше укрепили веру Д. Гильберта в реализуемость его программы (веру в возможность доказательства непротиворечивости всей математики с помощью надежных средств рассуждения). Казалось, остался еще один чисто математический, хотя и технически сложный шаг, – и цель будет достигнута. Однако уже в 1930 г. последовала совсем

неожиданная развязка...

Упражнение 3.2. (курсовая работа). Изучите алгоритм Пресбургера (см. Д. Гильберт, П. Бернайс [1934]) и покажите, что невозможна формула, означающая x*y=z, т.е. формула MULT(x,y,z) такая, что

$EA_1 \vdash MULT(x, 0, 0)$, (x*0=0)

$EA_1 \vdash (\forall u \forall v)(MULT(x, y+1, u) \wedge MULT(x, y, v) \rightarrow u=v+x)$.

(x*(y+1)=(x*y)+x).

В 1974 г. М. Фишер и М. Рабин показали, что проблема разрешения для формул EA_1 очень сложна: решение вопроса об истинности формулы длины n требует $2^{2^{cn}}$ единиц времени (см. М. Рабин [1977]).

Итак, теория EA_1 также не может нас удовлетворить: в ней невозможно обсуждать операцию умножения натуральных чисел и тем самым – понятие простого числа и другие важные понятия теории чисел. По-видимому, мы должны сделать еще один шаг: введем функциональный символ "*" для операции умножения, дополнив определение терма (в новой теории, которую будем обозначать просто через EA):

а) константы 0,1 и любая переменная – терм EA,

б) если t_1, t_2 – термы, то (t_1+t_2) и (t_1*t_2) – также термы EA.

Очевидно, термы EA – это обозначения полиномов (от нескольких, вообще говоря, переменных) с натуральными коэффициентами. Например, терм (((x*x)+(((1+1)*x)*y))+(y*y)) представляет полином $x^2+2xy+y^2$.

Теория EA использует только предикатный символ равенства "=" (понимаемый как равенство натуральных чисел). Если t_1, t_2 _ термы EA, то $(t_1=t_2)$ – элементарная формула EA. Из элементарных формул с помощью логических связок и кванторов строятся более сложные формулы EA, причем $\exists x$ мы понимаем как "существует натуральное число", а $\forall x$ – как "для всех натуральных чисел".

Средствами языка EA легко записываются простейшие утверждения о свойствах натуральных чисел, например

"x – четное число" $\leftrightarrow (\exists y)x=y+y$,

"x – простое число" $\leftrightarrow 1<x \wedge \neg(\exists y \exists z)(y<x \wedge z<x \wedge x=y*z)$,

"существует бесконечно много простых чисел" ↔

↔(∀x)(∃y)(x<y ∧ "y – простое число").

Упражнение 3.3. Перепишите средствами языка EA следующие утверждения: "существует бесконечно много простых чисел-близнецов", "при делении x на y получается остаток z", "x и y не имеют общих делителей".

Все, что мы до сих пор говорили о нашем "понимании" языка элементарной арифметики, является нашим личным делом и к определению теории EA не относится. В этой теории о свойствах введенных нами символов известно только то, что сформулировано в аксиомах. Часть этих свойств уже содержится в логических аксиомах и правилах вывода, которые безоговорочно принимаются в EA. Для определения наиболее существенных свойств мы вводим собственные аксиомы EA:

E1) ¬(0=x+1),

E2) ¬(x=y) → ¬(x+1=y+1),

E3) x+0=x,

E4) x+(y+1)=(x+y)+1,

E5) x*0=0,

E6) x*(y+1)=(x*y)+x,

E7) B(0) ∧ (∀x)(B(x)→B(x+1)) → (∀x)B(x).

Здесь дополнительно к аксиомам EA введено рекурсивное определение умножения (аксиомы E5, E6), а в схеме аксиом индукции (E7) в качестве формулы B может выступать любая формула B(x) языка EA с выделенной свободной переменной x (при этом кроме x формула может содержать и другие свободные переменные – параметры).

Ясно, что теория EA еще "сильнее" EA_1. В свете результата Фишера-Рабина о сложности разрешающего алгоритма для EA_1 (см. выше), у нас мало шансов построить такой алгоритм для EA (а в разделе 6.4 будет доказано, что он вообще невозможен). Но стоит ли заниматься всем этим, если может оказаться, что в EA невозможно определить операцию возведения в степень:

$x^0 = 1$,

$x^{y+1} = (x^y)*x$?

В разделе 3.3 будет доказана **теорема о представимости**, согласно которой в EA можно определить (с помощью подходящей формулы)

любую вычислимую функцию, т.е. дополнение языка ЕА новыми функциональными символами ничего к "мощности" теории ЕА прибавить не может. Это свидетельствует о **неслучайном характере** того объема средств рассуждения, который заключен в аксиомах ЕА.

С точки зрения этого результата теория ЕА представляет собой одну из возможных формализаций **дискретной математики –** как раз тех "надежных средств", которые Д. Гильберт предполагал использовать для доказательства непротиворечивости всей математики. Поэтому теория ЕА заслуживает серьезного изучения.

С интуитивным пониманием формул языка ЕА связан укоренившийся платонистский предрассудок, согласно которому всякая замкнутая формула на языке ЕА либо "истинна", либо "ложна". Разберемся в этом. Интуиция руководила нами при отборе аксиом. Мы руководствуемся ею также в случаях, когда требуется установить формальный аналог какой-либо теоремы из обычной (интуитивной) теории чисел (вспомним упражнение 3.3). Всякая попытка идти еще дальше и придать нашей интуиции натурального ряда какой-то более глубокий смысл приводит к платонизму. Однако многие идут этим путем и серьезно рассматривают следующее определение "истинности" формул языка ЕА:

а) Элементарные формулы ЕА имеют вид $(t_1=t_2)$, т.е. это утверждения о равенстве значений полиномов. При любых конкретных значениях переменных, входящих в термы t_1, t_2, их значения могут быть вычислены и тем самым может быть решен вопрос об истинности равенства $t_1=t_2$.

б) В теории ЕА имеется четыре логические связки: $\neg$, $\wedge$, $\vee$, $\rightarrow$. Если решен вопрос об истинности формул A, B, то по обычным правилам алгебры логики решается вопрос об истинности $\neg B$, $A\wedge B$, $A\vee B$, $A\rightarrow B$.

в) В теории ЕА имеется два квантора: $\exists$, $\forall$. Формула $(\exists x)C(x)$ считается истинной, если формула C(x) истинна при хотя бы одном значении переменной x. Формула $(\forall x)C(x)$ считается истинной, если C(x) истинна при любом значении переменной x.

Итак, вроде получается, что всякая замкнутая формула языка ЕА (формула, не имеющая свободных переменных) должна оказаться либо "истинной", либо "ложной". Это хорошо согласуется с нашей привычной интуицией: замкнутая формула языка ЕА выражает вполне определенное свойство системы натуральных чисел (вроде бесконечности множества простых чисел), которым эта система либо обладает, либо не обладает.

Заметим, однако, что в своем определении "истинности" формул

мы воспользовались оборотами типа "C(x) истинна при любом значении x". Согласно нашему интуитивному пониманию природы натуральных чисел, всевозможных значений x существует бесконечно много, т.е. мы не можем установить "истинность" (∀x)C(x) простой проверкой "истинности" C(x) для конкретных значений x. Истинность (∀x)C(x) мы в состоянии установить (если это вообще удается) только **теоретическим** путем – доказывая это утверждение на основе каких-либо аксиом, т.е. в определенной теории (например, EA или ZF). Можем ли мы утверждать, что всякую замкнутую формулу EA можно либо "теоретически доказать", либо "теоретически опровергнуть"? Нам хотелось бы так утверждать..., но какую **конкретную теорию** мы имеем в виду при этом – EA, ZF или другую? Чтобы придать нашему определению "истинности" точный смысл, мы должны сделать выбор, иначе определение будет висеть в воздухе.

По-видимому, одним из источников упрямой веры в осмысленность приведенного определения "истинности" является закон исключенного третьего, принятый в классической логике. Согласно этому закону, в частности,

$$EA \vdash A \vee \neg A$$

для любой формулы A. Однако следует ли отсюда, что если A – замкнутая формула, то либо EA ├ A, либо EA ├ ¬A? Из теоремы Геделя о неполноте вытекает, что ни EA, ни ZF, ни какая-либо другая серьезная математическая теория этим идеальным свойством обладать не могут – несмотря на постулирование закона исключенного третьего в их аксиомах! (Подробнее см. раздел 5.)

Основным средством вывода теорем в теории EA является, как и следовало ожидать, схема индукции. Рассмотрим в качестве примера вывод формулы 0+x=x (она отличается от аксиомы x+0=x!). Обозначим 0+x=x через A(x). Сначала мы должны доказать A(0), т.е. 0+0=0, но это частный случай упомянутой только что аксиомы. Теперь можем доказать A(x)→A(x+1). Предполагая воспользоваться теоремой дедукции, возьмем A(x) в качестве гипотезы:

A(x) или 0+x=x (гипотеза),

0+(x+1)=(0+x)+1 (частный случай аксиомы),

0+x=x → (0+x)+1=x+1 (свойство равенства),

(0+x)+1=x+1 (MODUS PONENS),

0+(x+1)=x+1 или A(x+1) (транзитивность равенства).

По теореме дедукции отсюда следует EA ├ A(x)→A(x+1), а затем EA ├ (∀x)(A(x)→A(x+1)). Так как A(0) уже доказано, то по схеме индукции

получаем EA $\vdash (\forall x)A(x)$ или EA $\vdash 0+x=x$.

Аналогично доказываются другие простые теоремы EA. Следует помнить, однако, что перед тем как доказывать какую-либо теорему (например, коммутативность умножения: x*y=y*x), полезно уже знать некоторые теоремы. Таким образом, даже доказательство простых теорем EA содержит в себе творческий момент – он состоит в наиболее рациональном выборе порядка, в котором эти теоремы следует доказывать.

Упражнение 3.4. Докажите коммутативность умножения, пройдя следующую цепь теорем (заимствованную из книги Э.Мендельсона [1976]):

$$x=y \rightarrow x+z=y+z,\ x+z=y+z \rightarrow x=y,$$

$$0+x=x,\ (x+1)+y=(x+y)+1,\ x+y=y+x,\ (x+y)+z=x+(y+z),$$

$$x=y \rightarrow x*z=y*z,\ 0*x=0,\ (x+1)*y=(x*y)+y,\ x*y=y*x.$$

В теории EA можно доказать все обычные свойства отношения $x<y$, определенного формулой $(\exists z)(x+z+1=y)$. В частности, можно доказать схему теорем, равносильную схеме индукции, – так называемый **принцип наименьшего числа**:

$$EA \vdash \neg(\forall x)C(x) \rightarrow (\exists y)(\neg C(y) \wedge (\forall z<y)C(z)).$$

В теории EA можно свободно трактовать **делимость** чисел. Если через R(x,y,z) обозначить формулу $(\exists u)(x=y*u+z \wedge z<y)$, т.е. "при делении x на y получается остаток z", то можно доказать, что

$$EA \vdash 0<y \rightarrow (\exists z)R(x,y,z),$$

$$EA \vdash R(x,y,z_1) \wedge R(x,y,z_2) \rightarrow z_1 = z_2,$$

т.е. теоремы о существовании и единственности остатка.

Все недостающие здесь доказательства можно найти в книге Э. Мендельсона [1976], в которой изложены формальные выводы из аксиом EA основных свойств натуральных чисел. После того, как средствами EA воспроизведены **основы** теории, можно приступить к формальному доказательству уже более серьезных теорем, например о бесконечности количества простых чисел. Потратив определенные усилия, мы могли бы "передоказать" средствами EA все теоремы, содержащиеся в элементарных учебниках теории чисел. Эту техническую работу мы здесь проводить не будем. Поверим на слово тем, кто уже проделал ее – для себя и для нас. Итак, следующее утверждение является эмпирически установленным фактом: все рассуждения обычной (интуитивной) теории чисел, которые не апеллируют к произвольным действительным числам и функциям, могут быть формально воспроизведены в EA.

Введем особые обозначения для некоторых специальных термов ЕА: **2** – для (1+1), **3** – для ((1+1)+1) и т.д. Термы такого рода принято называть **нумералами –** стандартными обозначениями конкретных натуральных чисел. Нумералы, используемые в схемах формул ЕА, будем обозначать через **k**, **l**, **m**, **n**, **p**, **q**, **r**.

Отметим теперь ряд простых свойств отношения "<", которые понадобятся нам в дальнейшем. Если k>0, то

$$EA \vdash x<\mathbf{k} \leftrightarrow (x=\mathbf{0})\vee(x=\mathbf{1})\vee\ldots\vee(x=\mathbf{k-1}),$$

$$EA \vdash (\exists x<\mathbf{k})C(x) \leftrightarrow C(\mathbf{0})\vee C(\mathbf{1})\vee\ldots\vee C(\mathbf{k-1}),$$

$$EA \vdash (\forall x<\mathbf{k})C(x) \leftrightarrow C(\mathbf{0})\wedge C(\mathbf{1})\wedge\ldots\wedge C(\mathbf{k-1}).$$

Разумеется, это **схемы** теорем.

Упражнение 3.5. Если в элементарной формуле $t_1=t_2$, содержащей переменные ("неизвестные"), все слагаемые правой части перенести в левую, то получим **диофантово уравнение** (см. раздел 4.1). Нас интересует решение таких уравнений в натуральных числах. Покажите, что если найдено конкретное решение $(b_1, b_2,..., b_n)$ уравнения $t_1 = t_2$, то

$$EA \vdash (\exists x_1 \exists x_2 \ldots \exists x_n)t_1 = t_2.$$

(Указание: сначала индукцией по структуре термов покажите, что в ЕА доказуема любая истинная формула $t_1 = t_2$, не содержащая переменных.)

3.2. Натуральные числа в других теориях

Формальная арифметика ЕА представляет простейший уровень математических рассуждений – в которых участвуют только целые числа (и не участвуют произвольные действительные числа, не говоря уже о произвольных множествах Кантора). Более сложные рассуждения формализуются и более сложными (по сравнению с ЕА) формальными теориями. "Силу" этих более сложных теорий составляет прежде всего их способность обсуждать более сложные объекты (действительные числа, функции действительных и комплексных переменных и т.д.), которые недоступны в ЕА. Однако не может ли оказаться, что более сложная теория в состоянии доказать некоторые утверждения, трактующие исключительно о свойствах натуральных чисел, которые не в состоянии доказать ЕА? Не может ли оказаться, например, что ЕА не в состоянии доказать бесконечность количества простых чисел-близнецов, однако с привлечением действительных чисел это доказательство удается? На первый взгляд такое невозможно – ведь натуральные числа определяются

независимо от действительных чисел (и "до них"). На самом деле ситуация сложнее...

Каким образом выделить в некоторой формальной теории T ту ее часть, которая относится к компетенции EA? Этот вопрос решается очень естественно с помощью так называемых **относительных интерпретаций**. Чтобы воспроизвести в теории T арифметику, прежде всего какие-то объекты из области значений переменных T должны быть объявлены натуральными числами. Это связано с выделением в языке T некоторой формулы N(x) (с единственной свободной переменной x), которая "утверждает", что x является натуральным числом. Далее, необходимо отобразить в теории T элементарные формулы EA, т.е. формулы вида $t_1 = t_2$, трактующие о значениях полиномов t_1, t_2 с натуральными коэффициентами. В самом общем виде это будет некоторое вычислимое отображение π, переводящее всякую элементарную формулу F из языка EA в формулу π(F) из языка T. При этом естественно требовать, чтобы формула π(F) всегда имела в точности те свободные переменные, которые имеет F. Такое отображение можно доопределить для произвольных формул EA:

$\pi(\neg A)=\neg\pi(A)$, $\pi(A\wedge B)=\pi(A)\wedge\pi(B)$,

$\pi(A\vee B)=\pi(A)\vee\pi(B)$, $\pi(A\to B)=\pi(A)\to\pi(B)$,

$\pi((\exists x)C(x))=(\exists x)N(x)\wedge\pi(C(x))$,

$\pi((\forall x)C(x))=(\forall x)(N(x)\to\pi(C(x)))$.

Пару (N, π) будем называть **относительной интерпретацией** EA в теории T, если для каждой аксиомы A теории EA

$$T \vdash N(x_1)\wedge\ldots\wedge N(x_k) \to \pi(A),$$

где $x_1,\ldots, x_k$ – свободные переменные аксиомы A. Например, для аксиомы $x+1=y+1 \to x=y$ требуется

$$T \vdash N(x)\wedge N(y) \to (\pi(x+1=y+1)\to\pi(x=y)).$$

Так как теория T содержит полный набор логических средств рассуждения, то отсюда следует, что для любой замкнутой формулы C из языка EA: если $EA \vdash C$, то $T \vdash \pi(C)$. Это означает, что, имея относительную интерпретацию EA в T, в теории T можно доказать любое свойство натуральных чисел, которое доказуемо в EA. И если нас интересует только "арифметическое содержание" теории T, мы можем, допуская вольность, даже писать $T \vdash C$ вместо $T \vdash \pi(C)$. Учитывая роль системы натуральных чисел в математике, формальную теорию, в которой относительно интерпретируема теория EA (и которая содержит в этом смысле полноценное понятие натурального числа), будем называть

фундаментальной теорией. Простейшей из фундаментальных теорий является, конечно, сама теория EA.

Теория множеств ZF, разумеется, также оказывается фундаментальной. Формула N(x), определяющая относительную интерпретацию EA в ZF, означает здесь просто "$x \in \omega$", где ω – множество всех натуральных чисел. Однако если написать "$x \in \omega$" в развернутом виде средствами языка ZF, то получится довольно сложная формула (см.раздел 2.3). Таким же сложным оказывается и отображение π. Ограничимся простейшими примерами. Во-первых, $\pi(x=y)=(x=y)$. Поскольку нуль определяется как пустое множество 0, единица – как {0}, 2 – как {0, {0}} и т.д., то $\pi(x=0)=(\forall y)\neg(y \in x)$, $\pi(x=1)=(\forall y)(y \in x \leftrightarrow \pi(y=0))$ и т.д. Формула $\pi(x=2)$ получится, если развернуть формулу

$$(\forall y)(y \in x \leftrightarrow \pi(y=0) \vee \pi(y=1)).$$

Вопрос, с которого мы начали этот раздел, можно теперь переформулировать следующим образом: возможно ли, что для некоторой фундаментальной теории T (которая "сильнее" теории EA) найдется в языке EA формула C такая, что $T \vdash C$, но не $EA \vdash C$? Ответ на этот вопрос будет получен в разделе 6.5 (см. также приложение 2).

Упражнение 3.6. Обобщая результат упражнения 1.4, покажите, что множество $\pi^{-1}(T)$ "арифметических теорем" всякой фундаментальной теории T является эффективно перечислимым. (Воспользуйтесь вычислимостью отображения π.)

3.3. Теорема о представимости

Как можно трактовать в теории EA операцию возведения в степень, если соответствующего символа в языке теории нет? Другой вопрос: мы уже несколько раз писали формулы, заменяющие в EA интуитивно понимаемые предикаты вроде "x – простое число":

$$1<x \wedge \neg(\exists y \exists z)(y<x \wedge z<x \wedge x=y*z).$$

Какие требования мы должны предъявлять к такого рода формулам? Если кто-то предлагает формулу и утверждает, что она "выражает" то-то и то-то, как это проверить? Для таких ситуаций вводятся понятия выразимости предикатов и представимости функций.

Будем говорить, что формула A(x,y) (x, y – единственные ее свободные переменные) **выражает** в теории EA **предикат** a(x,y) (x,y пробегают по натуральным числам), если для любых натуральных чисел m, n

а) если имеет место a(m, n), то ЕА ├ A(**m**, **n**),

б) если a(m, n) не имеет места, то ЕА ├ ¬A(**m**, **n**).

Напоминаем, что **m**, **n** – термы ЕА, обозначающие числа m, n: **2** – это 1+1 и т.д. Можно спорить, являются ли эти требования достаточными, чтобы признать формулу A действительно выражающей предикат a, но, вне всякого сомнения, эти требования необходимы. (Мы сформулировали определение для случая двухместных предикатов, аналогично определяется выразимость одноместных, трехместных и т.д. предикатов.)

Легко проверить, что формула x=y выражает в ЕА обычное интуитивно понимаемое равенство натуральных чисел. В самом деле, для любых m, n: а) если m=n, то ЕА ├ **m**=**n** (термы **m**, **n** в этом случае просто совпадают), б) если m≠n, то ЕА ├ ¬(**m**=**n**) (проверьте).

Интересно было бы получить какую-то характеристику класса тех предикатов, которые выразимы в ЕА. Если теория ЕА противоречива, то в ней доказуема любая формула и поэтому – в смысле принятого нами определения – в ЕА можно выразить любой предикат (например, формула x=y выражает любой двухместный предикат). Если же теория ЕА непротиворечива, то каждая формула может выражать не более одного предиката, и поскольку формул существует только счетное число, то и выразимых предикатов будет не более чем счетное число. Но даже всевозможных одно-местных предикатов "существует" несчетное число, т.е. должны "существовать" невыразимые предикаты.

Какие же предикаты выразимы в ЕА, а какие – нет (если предположить, что эта теория непротиворечива)? Легко заметить, что всякий выразимый в ЕА предикат должен быть эффективно разрешимым (по другой терминологии – рекурсивным) предикатом. В самом деле, если формула A(x,y) выражает в ЕА предикат a(x,y), то для решения вопроса об истинности a(m, n) мы должны установить, которая из формул A(**m**, **n**), ¬A(**m**, **n**) доказуема в ЕА (точно известно, что одна из них доказуема). Множество всех теорем ЕА эффективно перечислимо (см.упражнение 1.4). Перечисляя всевозможные теоремы ЕА, мы рано или поздно встретим либо формулу A(**m**, **n**), либо ¬A(**m**, **n**) (но только одну из них, поскольку ЕА – по предположению – непротиворечива). Это и решит вопрос об истинности a(m, n). Таким образом, a(x,y) – эффективно разрешимый предикат.

Решить вопрос, противоречива теория ЕА или нет, мы не в состоянии. Все, что можно сделать, – это попытаться доказать **независимо** от предположения о непротиворечивости ЕА, что **все** эффективно разрешимые предикаты выразимы в ЕА. Несколько позднее мы это сделаем.

Будем говорить, что формула F(x,y,z) **представляет** в ЕА **функцию** f(x,y) (отображающую натуральные числа в натуральные), если для всех k, m, n таких, что f(k, m)=n,

а) ЕА ├ F(**k**, **m**, **n**),

б) ЕА ├ (∀z)(¬(z=**n**)→¬F(**k**, **m**, z)).

Можно спросить по поводу этого определения: если отождествлять функцию f с предикатом f(x,y)=z, то почему бы не считать ее представлением формулу, выражающую в ЕА этот предикат? Это означало бы замену условия б) условием

б1) если f(k, m)≠n, то ЕА ├ ¬F(**k**, **m**, **n**).

Легко видеть, что из б) вытекает б1). В самом деле, если f(k, m)≠n, но f(k,m)=q, то ЕА ├ ¬(**n**=**q**), а согласно условию б) ЕА ├ ¬(**n**=**q**)→¬F(**k**, **m**, **n**), т.е. ЕА ├ ¬F(**k**, **m**, **n**), что и требовалось. Однако вывести б) из б1) мы уже не в состоянии. Поэтому можно сказать, что формула F, обладающая свойствами а+б), представляет функцию f "лучше", чем формула, которая обладает только свойствами а+б1). Если в случае а+б1) для каждого n, отличного от f(k, m), достаточно иметь отдельное доказательство формулы ¬F(**k**, **m**, **n**), то в случае а+б) для всех таких n должно быть дано **единое** доказательство: ЕА ├ ¬(z=**q**)→¬F(**k**, **m**, z), где q=f(k, m).

Если теория ЕА противоречива, то в ней представима любая функция.

Упражнение 3.7. Убедитесь, что если ЕА непротиворечива, то всякая всюду определенная функция, представимая в ЕА, должна быть вычислимой.

Мы хотим доказать, что в ЕА (**независимо** от предположения непротиворечивости) представима любая вычислимая функция.

Между выразимостью предикатов и представимостью функций существует простая и естественная связь.

Упражнение 3.8. Докажите, что некоторый предикат выразим в ЕА тогда и только тогда, когда его характеристическая функция представима в ЕА.

Достаточно, таким образом, установить представимость в ЕА всякой вычислимой функции, чтобы одновременно можно было утверждать, что всякий разрешимый предикат выразим в ЕА.

ТЕОРЕМА О ПРЕДСТАВИМОСТИ. Всякая вычислимая функция представима в теории ЕА.

Доказательству теоремы о представимости посвящена оставшаяся часть этого раздела.

Сначала рассмотрим одно из точных определений понятия **вычислимой функции**: функция, отображающая натуральные числа в натуральные числа, называется вычислимой, если существует **машина Тьюринга**, вычисляющая ее значения. Всякая машина Тьюринга M определяется:

а) алфавитом внутренних состояний (программист может считать их состояниями оперативной памяти, тогда $m=2^k$, где k – количество бит в памяти):

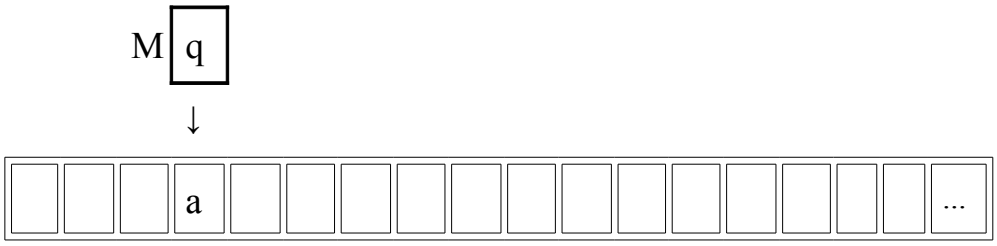
$$Q_M = \{q_{start}, q_{stop}, q_1, ..., q_m\},$$

б) алфавитом ленты:

$$A_M = \{0, 1, a_1, ..., a_n\},$$

в) программой P_M, состоящей из конечного числа команд вида qa → q'a'e, где q, q' $\in Q_M$ (q – "старое", q' – "новое" состояние машины), a,a' $\in A_M$ (a – "старый", a' – "новый" символ в клетке, обозреваемой головкой машины), e = –1, 0, +1 (указание, куда переместить головку). В программе не должно быть двух команд с одинаковыми левыми частями.

Под этим формальным определением скрыта следующая наглядная модель:



В каждый момент (выражаемый, например, целым числом секунд)

а) машина M находится в одном из состояний Q_M,

б) каждая клетка ленты (она бесконечна только вправо) находится в одном из состояний A_M,

в) головка машины обозревает некоторую клетку ленты. Если в момент t машина M находилась в состоянии q, головка обозревала в клетке символ a и в программе P_M содержится команда qa → q'a'e, то к моменту t+1 машина перейдет в состояние q', символ a она заменит символом a' и в зависимости от e головка перейдет: если e=–1, то на одну клетку влево, если e=0 – останется на месте, если e=+1 – на одну клетку

вправо (если головка обозревает самую левую клетку ленты, то e=–1 равносильно e=0).

Мы будем говорить, что машина M вычисляет (всюду определенную) функцию f(x,y), если для всех значений x, y, отправляясь от ситуации, в которой

а) M находится в состоянии q_{start},

б) головка обозревает левую клетку ленты,

в) в начале ленты находится последовательность символов

$$0\underbrace{11...1}_{x}0\underbrace{11...1}_{y}0 \quad ,$$

через конечное число шагов, следующих по программе P_M, машина приходит к ситуации, в которой:

а) она находится в состоянии q_{stop},

б) в начале ленты находится последовательность символов

$$0\underbrace{11...1}_{f(x,y)}0 \quad .$$

Если для функции f(x,y) удалось построить такую машину M, мы будем называть f вычислимой. Это определение оказывается эквивалентным всем другим точным определениям понятия вычислимой функции.

В качестве примера напишем программу для машины Тьюринга, которая вычисляет функцию f(x)=x+1. Задача этой программы – пройти через массив единиц до нуля, заменить нуль единицей, а рядом записать новый нуль:

$Q_M = \{q_{start}, q_{stop}, q, q'\}$, $A_M = \{0,1\}$,

$P_M = \{q_{start}, 0 \to q, 0, +1;$

$q, 1 \to q, 1, +1;$

$q, 0 \to q', 1, +1;$

$q', 0 \to q_{stop}, 0, 0;$

$q', 1 \to q_{stop}, 0, 0\}$

Упражнение 3.9. а) Напишите программы для машин Тьюринга, вычисляющих функции x+y, x·2, x·y, 2^x, $[\log_2 x]$.

б) Используя какой-либо язык программирования, напишите

интерпретатор машин Тьюринга. Это программа, которая получает программу машины и начальное состояние ленты, а затем "симулирует" работу введенной программы и после ее остановки печатает заключительное состояние ленты. Такой интерпретатор позволяет отлаживать программы для машин Тьюринга, физически этих машин не имея.

Для доказательства теоремы о представимости мы должны записать отношение f(x, y)=z в виде некоторой формулы F(x, y, z) на языке EA. При этом будем использовать то, что между парой аргументов x, y и значением z должна существовать связь в виде вычислительного процесса, следующего по программе P_M подходящей машины Тьюринга.

Своей цели мы достигли бы значительно быстрее, если бы в языке EA имелись специальные переменные для так называемых ситуаций. **Ситуацией** (**полным состоянием** машины) в момент времени t будем называть последовательность

$$(q, l, d_0, d_1, ..., d_{s-1}),$$

где q – внутреннее состояние машины в данный момент, l – номер клетки, обозреваемой головкой машины (нумерация начинается с нуля), d_0, d_1, ..., d_{s-1} – символы в клетках с номерами 0, 1, ..., s–1. Если C_1, C_2,... – переменные, значениями которых являются такие ситуации, было бы удобно иметь в языке EA также специальные символы для функций:

q(C)=q в ситуации C,

l(C)=l в ситуации C,

s(C)= число клеток, учитываемых в C,

$d_i(C)=d_i$ в ситуации C.

Имея все это в языке EA, мы могли бы поступить следующим образом. Сначала строим формулы START и STOP.

Формула START(C, x, y) должна утверждать, что C – начальная ситуация, в которой на ленте задана пара аргументов x, y:

$$q(C)=q_{start} \wedge l(C)=0 \wedge x+y+\mathbf{3}=s(C) \wedge d_0(C)=0 \wedge$$

$$\wedge \forall i\,(0<i<x+y+\mathbf{3} \rightarrow d_i(C)=1 \vee d_i(C)=0 \wedge (i=x \vee i=x+y+1)).$$

Формула STOP(C,z) должна утверждать, что C – заключительная ситуация, в которой вычислено значение z:

$$q(C)=q_{stop} \wedge z<s(C) \wedge d_0(C)=0 \wedge \forall i\,(0<i<z+1 \rightarrow d_i(C)=1 \vee (d_i(C)=0 \wedge i=z)).$$

Далее, каждой команде $\eta \in P_M$ (где M – машина, вычисляющая функцию f), имеющей вид qa → q'a'e, мы сопоставим формулу $STEP_\eta$ (C_1, C_2), которая утверждает, что ситуация C_2 получается из ситуации C_1 в результате выполнения команды η. Рассмотрим случай e=+1:

$$(\exists u \exists v)[q(C_1)=q \wedge l(C_1)=u \wedge s(C_1)=v \wedge d_u(C_1)=a \wedge$$

$$\wedge q(C_2)=q' \wedge d_u(C_2)=a' \wedge l(C_2)=u+1 \wedge (\forall i \le v)(i=u \vee (\exists j)(d_i(C_1)=j \wedge d_i(C_2)=j))$$

$$\wedge ((u<v \wedge s(C_2)=v) \vee (u=v \wedge s(C_2)=v+1))].$$

Упражнение 3.10. Напишите аналоги этой формулы для случаев e=–1, 0.

Следующий этап – формула $CHAINE_M(C_1, C_2)$, утверждающая, что ситуация C_2 получается из C_1 с помощью конечного числа шагов по программе P_M. Здесь удобно ввести уже специальные переменные L_1, L_2, принимающие в качестве значений **конечные последовательности** ситуаций. Соответственно нам потребуются также символы функций:

δ(L)= число ситуаций в последовательности L,

$\gamma_i(L)$= i-я ситуация в L.

В языке EA таких средств нет, но мы будем пока игнорировать это обстоятельство. Формула $CHAINE_M(C_1, C_2)$ имеет вид

$$(\exists L \exists w)[\delta(L)=w+1 \wedge \gamma_0(L)=C_1 \wedge \gamma_w(L)=C_2 \wedge$$

$$\wedge (\forall i<w)(\exists C_3 \exists C_4)\ (\gamma_i(L)=C_3 \wedge \gamma_{i+1}(L)=C_4 \wedge STEP_M(C_3, C_4)],$$

где $STEP_M(C_3, C_4)$ – формула

$$STEP_{k1}(C_3, C_4) \vee STEP_{k2}(C_3, C_4) \vee ... \vee STEP_{kt}(C_3, C_4),$$

а команды k1, k2, ..., kt составляют программу P_M.

Если машина M вычисляет функцию f(x, y), то формулу F(x, y, z), утверждающую, что f(x,y)=z, можно написать теперь следующим образом:

$$(\exists C_1 \exists C_2\ )(START(C_1, x, y) \wedge CHAINE_M\ (C_1, C_2) \wedge STOP(C_2, z)).$$

Одержав эту победу, мы вынуждены, однако, вернуться к языку EA. В нем нет переменных для ситуаций, нет функциональных символов, которые мы успели ввести, не говоря уже о переменных для

последовательностей ситуаций. Поэтому, если мы хотим, чтобы формула F(x,y,z) действительно "вошла" в язык ЕА, мы должны воспроизвести все "излишества" средствами ЕА.

Константы вроде q_{start}, q_{stop}, 0, 1, q, q', a, a' и т.д. "вложить" в ЕА нетрудно – каждую из них можно заменить некоторым натуральным числом – ее кодом (при этом будет использован только конечный набор чисел – алфавиты Q_M , A_M конечны). Но что делать с ситуациями? Чтобы они "вошли" в ЕА, их также нужно закодировать натуральными числами. Если внутренние состояния машины и символы на ленте уже закодированы натуральными числами, то ситуации превратились в конечные последовательности таких чисел. Мы должны научиться, следовательно, представлять любые конечные последовательности натуральных чисел **одним** натуральным числом (в худшем случае – двумя или тремя). Причем сделать это нужно так, чтобы функции q(C), d_i(C) и т.д. можно было представить формулами ЕА.

К. Геделю принадлежит идея использовать для этой цели так называемую **китайскую теорему об остатках**. Рассмотрим следующую задачу: найти число, которое при делении на 3 дает остаток 2, при делении на 5 – остаток 3, при делении на 7 – остаток 4. Если повезет, решение этой задачи можно угадать:

$$53=3\cdot17+2=5\cdot10+3=7\cdot7+4.$$

Рассмотрим, однако, общий случай: найти число, которое при делении на заданные числа дает заданные остатки. Заданные числа обозначим через u_1, u_2, ..., u_n, это натуральные числа ≥ 2. Заданные остатки обозначим через v_1, v_2, ..., v_n; таким образом, $0 \leq v_i < u_i$. Если числа u_i, u_j имеют общий делитель, то остатки при делении на эти числа уже нельзя выбрать независимо. Например, если u_i , u_j – оба четные, то соответствующие остатки должны иметь одинаковую четность, иначе задача не будет иметь решения:

$$x=u_i y_i+v_i = u_j y_j +v_j,\ u_i = 2u_i',\ u_j = 2u_j',\ v_i - v_j = 2(u_i'y_i - u_j'y_j).$$

Поэтому ситуация упростится, если мы потребуем, чтобы числа u_1, u_2, ..., u_n были **попарно взаимно простыми** (т.е. чтобы никакие два из них не имели общих делителей). В таком случае задача всегда имеет решение, как утверждает

КИТАЙСКАЯ ТЕОРЕМА ОБ ОСТАТКАХ. Если даны попарно взаимно простые числа u_1, u_2, ..., u_n ($u_i \geq 2$ для всех i) и числа v_1, ..., v_n ($0 \leq v_i < u_i$ для всех i), то найдется число p, меньшее произведения

$u_1u_2...u_n$, которое при делении на u_i дает остаток v_i (и так для всех i).

Д о к а з а т е л ь с т в о. Каждому числу p, где

$$0 \le p < P = u_1u_2...u_n,$$

сопоставим "вектор остатков" при делении на u_1, ..., u_n. Покажите, что два числа могут иметь один и тот же вектор остатков, только если их разность делится на P.

С помощью китайской теоремы об остатках мы могли бы кодировать последовательность натуральных чисел v_0, v_1, ..., v_{n-1}, представляя каждое v_i как остаток от деления некоторого числа p на число u_i из набора u_0, u_1, ..., u_{n-1} , который мы должны научиться генерировать каким-то достаточно простым способом. Иначе кодирование не получится. Самый простой способ – представить u_i как линейную функцию от i: u = yi+z. Но как подобрать коэффициенты y, z? Числа u_i должны быть взаимно простыми. Если d общий делитель u_i и u_j , то d делит также разность $u_i - u_j = y(i-j)$. Если возьмем z=y+1, т.е. $u_i = 1+y(1+i)$, то делители y не смогут быть общими делителями u_i, u_j. Ими могут быть тогда только делители разности |i–j|, т.е. числа ≤n–1. Но если к тому же еще y делится на (n–1)!, то и эти числа не смогут быть общими делителями u_i , u_j, т.е. числа u_0, ..., u_{n-1} будут взаимно простыми. И если, кроме того, y настолько велико, что $u_0 > v_0$, $u_1 > v_1$, ..., $u_{n-1} > v_{n-1}$, то согласно китайской теореме об остатках найдется число p, которое при делении на u_0 дает остаток v_0, при делении на u_1 – остаток v_1, ... , при делении на u_{n-1} – остаток v_{n-1}.

Пару чисел (x, y) мы могли бы считать кодом последовательности v_0, v_1, ..., v_{n-1}, но в этих числах никак не отражается число n. Если же вместо v_0, v_1, ..., v_{n-1} закодировать указанным способом последовательность n, v_0, v_1, ..., v_{n-1}, то соответствующая пара (x, y) стала бы действительно полноценным кодом v_0, v_1, ..., v_{n-1}.

Функцию

$$\beta(x, y, i) = \text{остаток от деления x на } 1+y(1+i)$$

принято называть **бета-функцией Геделя**. Мы установили, что для любого набора натуральных чисел v_0, ..., v_{n-1} можно подобрать числа x, y такие, что

$$\beta(x, y, 0)=n, \beta(x, y, 1)=v_0, ..., \beta(x, y, n)=v_{n-1}.$$

Отметим также, что бета-функцию можно представить в теории EA следующей формулой B(x, y, i, j) (утверждающей, что β(x, y, i)=j):

$$(\exists z)(x=(1+y*(1+i))*z+j \wedge j<1+y*(1+i)).$$

Теперь можно переписать формулы START, STOP и т.д. средствами языка EA. О константах мы уже условились – все состояния машины M, все символы на ленте будем кодировать натуральными числами. Далее, ситуация – последовательность $(q, l, d_0, d_1, ..., d_{s-1})$ становится теперь последовательностью натуральных чисел, которую можно закодировать двумя натуральными числами a, b, такими, что

$$\beta(a, b, 0)=s, \beta(a, b, 1)=q, \beta(a, b, 2)=l, \beta(a, b, i+3)=d_i$$

(для всех i). Кванторы вида ∃C, где C – переменная для ситуаций, можно заменить теперь кванторами ∃a∃b, где a, b – уже переменные для натуральных чисел. Равенства, содержащие "незаконные" функциональные символы:

$$s(C)=s, q(C)=q, l(C)=l, d_i(C)=d_i,$$

можно заменить равенствами

$$\beta(a, b, 0)=s, \beta(a, b, 1)=q, \beta(a, b, 2)=l , \beta(a, b, i+3)=d_i.$$

Устранение "незаконных" символов из неравенств вида $s_1 < s(C)$ также не составляет проблемы:

$$(\exists s_2)(\beta(a, b, 0)=s_2 \wedge s_1 < s_2).$$

Упражнение 3.11. Перепишите средствами EA формулы START, STOP, $STEP_\eta$, $STEP_M$. Определите длину каждой формулы.

В формуле $CHAINE_M$ появились новые, еще более сложные переменные – для конечных последовательностей ситуаций, а также функциональные символы δ(L), $\gamma_i(L)$. Отдельную ситуацию мы научились кодировать двумя натуральными числами. Если ситуация C_i закодирована парой чисел (a_i , b_i), то последовательность

$$L=\{C_0, C_1, ..., C_{n-1}\}$$

мы можем закодировать как последовательность чисел

$$n, a_0, b_0, a_1, b_1, ..., a_{n-1}, b_{n-1}.$$

Если (a, b) – код этой последовательности, то формулу δ(L)=w можно

заменить на $\beta(a, b, 0)=w$, а формулу $\gamma_i(L)=C$ – на

$$\beta(a, b, 2i+1)=a' \wedge \beta(a, b, 2i+2)=b',$$

где (a', b') – код ситуации C. Квантор $\exists L$, где L – переменная для последовательностей ситуаций, следует заменить кванторами $\exists a \exists b$.

Упражнение 3.12. Сделайте все эти замены в формуле CHAINE_M и определите ее длину. То же сделайте и для формулы F(x, y, z), утверждающей, что f(x, y)=z.

Итак, для каждой вычислимой функции f(x, y) исходя из машины Тьюринга, вычисляющей ее, мы умеем строить формулу F(x, y, z) в языке EA, которая (в нашем интуитивном ее понимании) утверждает, что f(x, y)=z. Чтобы завершить доказательство теоремы о представимости, мы должны показать, что в EA доказуемы формулы

$$F(\mathbf{k}, \mathbf{m}, \mathbf{n}), \neg(z=\mathbf{n}) \rightarrow \neg F(\mathbf{k}, \mathbf{m}, z),$$

(при условии, что f(k, m)=n). Мы не будем этим заниматься (как не занимались выводом из аксиом EA сколько-нибудь полной системы основных свойств натуральных чисел). Желающие проследить за всеми деталями могут обратиться к книге Э. Мендельсона [1976].

Будем считать теорему о представимости доказанной.

4. ДЕСЯТАЯ ПРОБЛЕМА ГИЛЬБЕРТА

4.1. История проблемы и ее решения

Задачами, приводящими к решению уравнений в целых числах, математики интересовались со времен Пифагора (VI в. до н.э.). Давно известно, что некоторые уравнения вообще не имеют решений в целых числах (например, 2x+2y=1, поскольку при любых целых x,y левая часть является четным числом). Другие имеют конечное число решений (например, $x^2+y^2=2$ сводится к $x^2=1$, $y^2=1$, т.е. получаются четыре решения). Наконец, бывают уравнения, имеющие бесконечно много решений в целых числах. В качестве примера рассмотрим уравнение 3x–7y=1. Решая относительно x, получаем

$$x=\frac{7y+1}{3}=2y+\frac{y+1}{3} \quad .$$

Число $\frac{y+1}{3}$ должно быть целым, обозначим его через t, тогда y=3t–1, x=2y+t=7t–2. Какое бы t мы ни взяли, получается целое решение уравнения.

По каким признакам определить, имеет ли данное уравнение решения в целых числах, и если имеет – то сколько их? Прежде чем искать ответ на этот вопрос, следует уточнить класс уравнений, о котором идет речь. Имеются в виду уравнения типа P=Q, где P, Q – выражения, составленные из символов неизвестных (их может быть один, два, три и больше), из целых чисел и операций сложения, вычитания и умножения. Уравнения такого рода (при условии, что нас интересуют только целые решения) принято называть **диофантовыми уравнениями** (в честь Диофанта, который в III в. н.э. занимался задачами, приводящими к таким уравнениям). Говоря современным языком, речь идет о решении в целых числах уравнений вида P=0, где P – полином (от одной или нескольких переменных) с целыми коэффициентами.

Следующий способ решения любого уравнения **первой степени** с двумя неизвестными был известен еще в средние века. Пусть дано уравнение ax+by=c. Если наибольший общий делитель чисел a, b не является делителем числа c, то уравнение не имеет целых решений. Если

является – делим обе стороны уравнения на этот общий делитель и начинаем применять метод редукции коэффициентов (как выше, при решении уравнения 3x–7y=1). Через конечное число шагов приходим к формулам вида x=dt+e, y=gt+h, которые при любом t дают решение исходного уравнения ax+by=c (решений в этом случае получается бесконечно много, поскольку всегда d, g≠0).

Далее, общий метод решения диофантовых уравнений **второй степени** с двумя неизвестными был найден Ж. Лагранжем в XVIII в.

К сожалению, это почти все сколько-нибудь общие результаты исчерпывающего характера. Остальные исследования давали множество тонких, но **частных** методов, применимых к уравнениям третьей, четвертой степени и т.д. весьма специальных видов (а то и только к отдельному уравнению). Чем объясняется этот контраст между первоначальными успехами и отсутствием дальнейшего продвижения (несмотря на исключительное развитие математики в целом)?

В августе 1900 г. в Париже состоялся II Международный конгресс математиков. 8 августа Д. Гильберт прочитал на нем доклад "Математические проблемы". Среди 23 проблем, решение которых (по мнению Д. Гильберта) совершенно необходимо было получить в наступающем XX в., десятую проблему он определил следующим образом (см. Д. Гильберт [1900]):

"Пусть задано диофантово уравнение с произвольным числом неизвестных и целыми рациональными числовыми коэффициентами. Указать способ, при помощи которого возможно после конечного числа операций установить, разрешимо ли это уравнение в целых рациональных числах".

Сама формулировка показывает, что тогда, в 1900 г., с определенностью можно было говорить только о **положительном** решении **десятой проблемы Гильберта –** об "указании способа". Что для всего огромного разнообразия уравнений **единый** "способ решения" может отсутствовать – мысль о такой возможности в 1900 г. никому не приходила в голову. Лишь в 30-х гг. оформилось математически точное понятие **алгоритма** ("способа, который после конечного числа операций дает ответ", см. раздел 3.3). Пока класс всех возможных "способов" не был определен с математической точностью, нельзя было серьезно говорить о доказательстве невозможности общего "способа" для решения какого-либо вида задач.

К концу 40-х гг. вера в адекватность математически точного понятия алгоритма укоренилась уже настолько, что можно было серьезно ставить вопрос об **отрицательном** решении десятой проблемы Гильберта

(и других классических проблем, касающихся существования алгоритмов). Т.е. можно было говорить о строгом доказательстве невозможности алгоритма, который по данному диофантову уравнению определял бы, имеет ли оно решение в целых числах или нет. Гипотезу, что десятая проблема Гильберта **алгоритмически неразрешима**, первым выдвинул (с достаточным на то основанием) американский математик М. Дэвис в 1949 г. Доказательство этой гипотезы растянулось на 20 лет – последний шаг был сделан только в 1970 г.

М. Дэвис наметил следующий путь доказательства своей гипотезы. Сначала он перешел от решения уравнений в целых числах к решению в натуральных числах (область, более привычная для теории алгоритмов).

Упражнение 4.1. Покажите, что алгоритм, определяющий разрешимость диофантовых уравнений в целых числах, существует, если и только если существует алгоритм, определяющий их разрешимость в натуральных числах. (В своих рассуждениях можете использовать тот факт, что полином $x^2+y^2+z^2+t^2$ при всевозможных целых значениях x, y, z, t принимает в качестве значений все натуральные числа. Это известная теорема Лагранжа о том, что каждое натуральное число представимо в виде суммы четырех квадратов, см. например, А. А. Бухштаб [1966].)

Итак, М. Дэвис предлагал доказывать невозможность алгоритма, определяющего разрешимость диофантовых уравнений в натуральных числах. Каким же образом?

Упражнение 4.2. Пусть $P(a, x_1, ..., x_n)=0$ – диофантово уравнение, содержащее параметр a. Проверьте, что множество S тех значений a, для которых уравнение разрешимо в натуральных числах, является эффективно перечислимым.

Если бы удалось построить уравнение (с параметром a) такое, что множество S оказалось бы неразрешимым, то цель была бы достигнута. Простейший способ доказать, что S может быть неразрешимым (будучи эффективно перечислимым), – показать, что в качестве S, при подходящем выборе уравнения, может выступать **произвольное** эффективно перечислимое множество (как известно, среди эффективно перечислимых множеств существуют неразрешимые).

Исходя из таких соображений, М. Дэвис ввел понятие диофантовых представлений. Пусть $R(a_1, ..., a_m)$ – предикат для натуральных чисел. Формулу

$$(\exists x_1)...(\exists x_n)\ P(a_1, ..., a_m, x_1, ..., x_n)=0,$$

где P – полином с целыми коэффициентами, будем называть

диофантовым представлением предиката R, если эта формула истинна для тех и только тех наборов (a_1, ... a_m), для которых истинным является предикат R. Например, предикат "a – четное число" имеет следующее диофантово представление:

$$(\exists x)\ a–2x=0.$$

Таким образом, диофантово представление предиката R – это диофантово уравнение с параметрами (a_1, .., a_m), которое разрешимо (в натуральных числах), если и только если $R(a_1, ..., a_m)$ истинно.

Ясно, что всякий предикат, обладающий диофантовым представлением, эффективно перечислим (упражнение 4.2). М. Дэвис предположил, что для всякого эффективно перечислимого предиката существует диофантово представление. Если это так, то отсюда вытекает алгоритмическая неразрешимость десятой проблемы Гильберта. Возьмем эффективно перечислимое неразрешимое множество S, построим диофантово представление $a \in S$:

$$(\exists x_1)...(\exists x_n)P(a, x_1, ..., x_n) = 0.$$

Тогда невозможен алгоритм, определяющий по числу a, разрешимо ли уравнение $P(a, x_1, ..., x_n)=0$ в натуральных числах (что и требовалось).

В 1949 г. М. Дэвису удалось сделать только первый шаг в намеченном направлении: он показал, что всякий эффективно перечислимый предикат $R(a_1, ..., a_m)$ можно представить в виде

$$(\exists y)(\forall z\leq y)(\exists x_1)...(\exists x_n)\ P(a_1, ..., a_m, y, z, x_1, ..., x_n) = 0.$$

Избавиться от квантора $\forall z\leq y$ М. Дэвису тогда не удалось.

Следующий успех был достигнут в 1960 г.: М. Дэвис совместно с Х. Патнэмом доказали, что для любого эффективно перечислимого предиката $R(a_1, ..., a_m)$ можно получить представление вида

$$(\exists x_1)...(\exists x_n)\ X(a_1, ..., a_m, x_1, ..., x_n) = 0,$$

где выражение X получено из символов $a_1, ..., a_m, x_1, ..., x_n$ и натуральных чисел применением операций сложения, вычитания, умножения и **возведения в степень** (например, $x^{2y+z}–yz+3=0$). Правда, доказательство Дэвиса-Патнэма содержало пробел (они воспользовались недоказанной до сих пор гипотезой о существовании произвольно длинных арифметических прогрессий, состоящих только из простых чисел). Восполнить этот пробел удалось Дж. Робинсон в 1961 г. Уравнения вида $X(a_1, ..., a_m, x_1, ..., x_n) = 0$ были названы **показательно-диофантовыми**

уравнениями. Из теоремы Дэвиса-Патнэма-Робинсон вытекала невозможность алгоритма, определяющего, разрешимо ли данное показательно-диофантово уравнение в натуральных числах. Это был крупный успех, однако до полной победы оставалось еще почти 10 лет.

Более того, даже после 1961 г. не все еще поверили, что дело удастся намеченным путем довести до конца (показав, что всякий эффективно перечислимый предикат обладает диофантовым представлением). Если взять предикаты "a – простое число" и "a – степень числа 2" и вообразить, что построены их диофантовы представления:

$$\text{"a – простое число"} \leftrightarrow (\exists x_1)...(\exists x_n)\, P(a, x_1, ..., x_n) = 0,$$

$$\text{"a – степень двойки"} \leftrightarrow (\exists y_1)...(\exists y_k)\, Q(a, y_1, ..., y_k) = 0,$$

то получится, что уравнение P=0 разрешимо, если и только если a – простое число, а уравнение Q=0 – если и только если a имеет вид 2^m. Такая возможность сильно противоречила сложившейся теоретико-числовой интуиции (согласно этой интуиции, простые числа и экспонента настолько далеко отстоят от полиномов, что между ними никогда нельзя будет установить простую связь). И тем не менее в 1970 г. советскому математику Юрию Владимировичу Матиясевичу (см. Ю. В. Матиясевич [1970]) удалось сделать последний и решающий шаг – получить диофантово представление **экспоненты**:

$$a=b^c \leftrightarrow (\exists x_1)...(\exists x_n)\, P(a, b, c, x_1, ..., x_n) = 0.$$

С помощью этого представления легко исключить операцию возведения в степень из представлений Дэвиса-Патнэма-Робинсон и таким образом получить диофантово представление для любого эффективно перечислимого предиката. Тем самым в 1970 г. алгоритмическая неразрешимость десятой проблемы Гильберта была **доказана полностью**.

Алгоритмическая неразрешимость объясняет отмеченные выше трудности с решением диофантовых уравнений высоких степеней. **Общего** метода, определяющего, разрешимо ли любое данное диофантово уравнение, не существует. Поэтому всякий метод определения разрешимости неизбежно оказывается **частным** – применимым только к уравнениям специального вида. Одновременно такая принципиальная ограниченность всякого метода обеспечивает неограниченный простор для творчества в данной области математики.

В последующих разделах излагается доказательство существования диофантовых представлений для любого эффективно

перечислимого предиката.

Упражнение 4.3. Покажите, что проблема разрешимости произвольного диофантова уравнения может быть сведена к проблеме разрешимости: а) системы диофантовых уравнений 2-й степени или б) диофантова уравнения 4-й степени. Таким образом, проблема разрешимости систем уравнений 2-й степени и уравнений 4-й степени оказывается уже алгоритмически неразрешимой. Не случайно поэтому, что до сих пор не найдены сколько-нибудь общие методы решения диофантовых уравнений 4-й степени.

4.2. Начало и план доказательства

Диофантово представление предиката является по существу формулой специального вида в языке ЕА (нужно только в уравнении P=0 перенести в правую сторону члены с отрицательными коэффициентами). Вспомним, что, доказывая представимость в ЕА любой вычислимой функции, мы уже получили некоторые специальные формулы ЕА, представляющие предикаты вида $f(a_1, ..., a_m) = b$, где f – произвольная вычислимая функция.

Упражнение 4.4. Проверьте по тексту раздела 3.3, что эти формулы были образованы с помощью только следующих средств:

а) элементарных формул вида s=t, s<t, где s, t – термы ЕА (полиномы с целыми положительными коэффициентами),

б) конъюнкции и дизъюнкции,

в) квантора существования,

г) ограниченного квантора всеобщности, а именно квантора $\forall x \le U$, где U – линейная функция вида $b_1y + b_2y_2 + ... + b_ky_k + c$ с натуральными коэффициентами $b_1, b_2, ..., b_k, c$.

В своей работе 1949 г. М. Дэвис также исходил из подобного результата. Важнейшей (с точки зрения наших целей) особенностью представляющих формул является отсутствие в них символов отрицания и неограниченных кванторов всеобщности. "Недостатком" отрицаний и кванторов всеобщности является то, что они выводят за пределы класса эффективно перечислимых предикатов (если R – эффективно перечислимый предикат, то $\neg R$ уже не обязательно таков, аналогично – в случае перехода от R(x,y) к $(\forall y)R(x,y)$). Формулы, содержащие такие средства, в общем случае невозможно перевести в диофантовы

представления.

Итак, мы начинаем с эффективно перечислимого предиката $R(a_1, ..., a_m)$. Построим машину Тьюринга M(R), которая печатает на своей ленте все возможные наборы чисел $a_1, ..., a_m$, удовлетворяющие предикату $R(a_1, ..., a_m)$. Тогда следующая функция вычислима:

$f(a_1, ..., a_m, t) = 1$, если после t шагов работы машины M(R) набор чисел $a_1, ..., a_m$ на ленте уже напечатан,

$f(a_1, ..., a_m, t) = 0$, иначе.

По теореме о представимости можно построить формулу ЕА, которая представляет функцию f и содержит только средства, перечисленные в пп. а)–г) упражнения 4.4. Если $F(a_1, ..., a_m, t, w)$ – эта формула (w соответствует значению функции), то

$$R(a_1, ..., a_m) \leftrightarrow (\exists t)\, F(a_1, ..., a_m, t, 1).$$

Тем самым доказано, что предикат R равносилен некоторой формуле, построенной только с помощью средств, перечисленных в а)–г). Теперь мы должны научиться преобразовывать такие формулы в диофантовы представления. Начнем "изнутри" – с элементарных формул. Формула вида s=t уже является диофантовым представлением (без кванторов). Формулу вида s<t можно заменить формулой $(\exists x)(s+x+1=t)$, которая уже является диофантовым представлением. Теперь можно переходить к средствам б)–г), с помощью которых строятся более сложные формулы.

Упражнение 4.5. Покажите, что конъюнкцию и дизъюнкцию двух диофантовых представлений $(\exists)P=0$, $(\exists)Q=0$ можно преобразовать в диофантово представление.

Таким образом, если в нашем процессе преобразования, начинающемся с элементарных формул, встречаются символы $\wedge$, $\vee$, мы знаем, как от них освободиться. Если встретился символ $\exists$, он может остаться на своем месте (если $(\exists)P=0$ – диофантово представление, то $(\exists\exists)P=0$ – также). Но что делать, если встретился квантор $\forall z \le t$? Пусть то, что находилось за ним, мы уже привели к виду диофантова представления

$$(\forall z \le t)(\exists x_1)...(\exists x_n)\, P(b_1, ..., b_k, y, x_1, ..., x_n) = 0 \qquad (1)$$

(напомним, что t – линейная функция с натуральными коэффициентами от $b_1, ..., b_k$). Мы хотим заменить (1) диофантовым представлением вида

$$(\exists y_1)...(\exists y_s)Q(b_1, ..., b_k, y, ..., y_s) = 0, \qquad (2)$$

но как это сделать?

Оказывается, что это очень сложная задача: преобразовать формулу вида (1) в формулу вида (2). Будем называть ее задачей **устранения ограниченного квантора всеобщности** (короче: устранения $\forall\leq$). Решив эту задачу, мы тем самым решим до конца задачу преобразования формул, образованных с помощью средств а)–г), в диофантовы представления, т.е. одновременно докажем существование представлений для любых эффективно перечислимых предикатов.

Итак, решаем задачу устранения $\forall\leq$. План наших действий:

1) Подробно исследовать свойства решений уравнения $x^2-(a^2-1)y^2=1$ $(a>1)$. Оказывается, это уравнение имеет в натуральных числах бесконечно много решений. Если n-е решение обозначить через $(x_n(a), y_n(a))$, то $x_n(a)$ и $y_n(a)$ растут экспоненциально по n. Этим и обусловлен наш интерес к данному уравнению.

2) Опираясь на полученную информацию, построить диофантово представление предиката

$$R(a, x, y, n) \leftrightarrow x=x_n(a) \wedge y=y_n(a),$$

т.е. предиката, который при фиксированном a требует от x и y экспоненциального роста по n.

3) Используя диофантово представление предиката R, получить диофантово представление экспоненты, т.е. предиката $x=y^z \wedge v\geq 3$.

4) Получить диофантовы представления для числа сочетаний и факториала (т.е. для предикатов $x=C_y^z$ и $x=y!$).

5) Используя все полученные диофантовы представления, научиться устранять $\forall\leq$.

Задачи 1), 2) были решены Ю. В. Матиясевичем, задачи 3), 4) – Дж. Робинсон, задача 5) – совместно М. Дэвисом, Х. Патнэмом и Дж. Робинсон.

Чтобы сделать наши рассуждения по возможности наглядными, воспользуемся языком **сравнений**. Сравнение – это нечто вроде равенства, только не точное равенство, а равенство с точностью до слагаемого, кратного **модулю** (по которому рассматривается сравнение). Например, число 18 сравнимо с 78 по модулю 10:

$$18 \equiv 78 \bmod 10,$$

поскольку $78=18+6\cdot 10$. Число сравнимо с нулем по модулю m, если и только если оно делится на m: $x \equiv 0 \bmod m$ означает, что $x=0+k\cdot m$ для

некоторого k.

Упражнение 4.6. Докажите следующие свойства сравнений (позволяющие обращаться с ними как с обычными равенствами):

$a \equiv a \bmod m$,

$a \equiv b \bmod m \rightarrow b \equiv a \bmod m$,

$a \equiv b \bmod m \wedge b \equiv c \bmod m \rightarrow a \equiv c \bmod m$,

$a \equiv b \bmod m \wedge a_1 \equiv b_1 \bmod m \rightarrow a+a_1 \equiv b+b_1 \bmod m$,

$a \equiv b \bmod m \wedge a_1 \equiv b_1 \bmod m \rightarrow a \cdot a_1 \equiv b \cdot b_1 \bmod m$,

$a \cdot c \equiv b \cdot c \bmod m \rightarrow a \equiv b \bmod m$, если c взаимно просто с m,

$a \cdot c \equiv b \cdot c \bmod (c \cdot m) \rightarrow a \equiv b \bmod m$, если c взаимно просто с m.

4.3. Исследование уравнения Ферма

Уравнение $x^2 - Dy^2 = 1$ (где $D>0$) играет центральную роль в теории диофантовых уравнений второй степени с двумя неизвестными. Если коэффициент D является полным квадратом ($D=k^2$), то решение уравнения сводится к системам линейных уравнений:

$$x^2 - k^2y^2 = (x - ky)(x + ky) = 1,$$

$$x - ky = 1 \wedge x + ky = 1 \rightarrow x = 1 \wedge y = 0,$$

$$x - ky = -1 \wedge x + ky = -1 \rightarrow x = -1 \wedge y = 0.$$

Таким образом, получаются только два решения.

Значительно более интересен случай, когда D не является полным квадратом. Совершенно неожиданно в этом случае уравнение $x^2 - Dy^2 = 1$ всегда имеет бесконечно много решений в натуральных числах! Это знал еще П. Ферма в XVII в., но первое строгое доказательство было дано только Ж. Лагранжем (XVIII в.). Простейшим для анализа уравнение Ферма оказывается при $D=a^2-1$:

$$x^2 - (a^2 - 1)y^2 = 1.$$

Два его решения можно угадать (проверьте): $x=1 \wedge y=0$, $x=a \wedge y=1$. Остальные решения (в натуральных числах) получаются с помощью следующего остроумного рассуждения. Возьмем выражение $\left(a+\sqrt{a^2-1}\right)^n$ и разложим его по формуле бинома Ньютона. В результате часть членов будут целыми числами, другая же часть будет

содержать множитель $\sqrt{a^2-1}$. Например, для n=2:

$$\left(a+\sqrt{a^2-1}\right)^2=a^2+2a\sqrt{a^2-1}+\left(a^2-1\right) \quad .$$

Сводя вместе члены в каждой части, получаем

$$\left(a+\sqrt{a^2-1}\right)^n=x_n(a)+y_n(a)\sqrt{a^2-1} \quad ,$$

где $x_n(a)$, $y_n(a)$ – натуральные числа (например, $x_2(a)=2a^2-1 \wedge y_2(a)=2a$). Аналогично

$$\left(a-\sqrt{a^2-1}\right)^n=x_n(a)-y_n(a)\sqrt{a^2-1}$$

(с теми же x_n, y_n – проверьте, что это действительно так!). Перемножая оба последних равенства, получаем

$$(a^2-a^2+1)^n = x_n^{\,2} - (a^2-1)y_n^{\,2} ,$$

$$x_n^{\,2} - (a^2-1)y_n^{\,2} = 1.$$

Таким образом, при любом $n\geq 0$ пара чисел $x=x_n(a)$, $y=y_n(a)$ является решением уравнения $x^2-(a^2-1)y^2=1$. При n=0, 1 получаются уже известные нам тривиальные решения (1, 0), (a, 1), при n=2 – новое решение $(2a^2-1, 2a)$.

Из нашего определения чисел $x_n(a)$, $y_n(a)$ легко получить рекуррентные соотношения, позволяющие вычислить x_{m+n}, y_{m+n}, если уже известны x_m, y_m, x_n, y_n $(m, n \geq 0)$:

$$x_{m+n}(a) = x_m(a)\, x_n(a) + y_m(a)\, y_n(a)(a^2-1),$$

$$y_{m+n}(a) = x_m(a)\, y_n(a) + y_m(a)\, x_n(a).$$

В частности, при m=1:

$$x_{n+1}(a) = a\, x_n(a) + (a^2-1)\, y_n(a),$$

$$y_{n+1}(a) = x_n(a) + a\, y_n(a).$$

Упражнение 4.7. Докажите эти соотношения. Докажите также, что $x_n(a)$, $y_n(a)$ возрастают по n (т.е. что действительно получается бесконечно много решений уравнения $x^2-(a^2-1)y^2=1$).

Оказывается, что последовательность $\{(x_n, y_n) \mid n\geq 0\}$ исчерпывает **все** решения нашего уравнения.

ЛЕММА 1. При a>1

$$x^2-(a^2-1)y=1 \leftrightarrow (\exists n)(x=x_n(a) \wedge y=y_n(a)).$$

Д о к а з а т е л ь с т в о. 1) Влево. Это мы уже знаем.

2) Вправо. Пусть числа x,y удовлетворяют уравнению. Если x≤1, то x=1 и y=0, т.е. $x=x_0(a)$ и $y=y_0(a)$.

Пусть теперь x>1. Тогда y>0. Если мы рассчитываем показать, что $x=x_n(a) \wedge y=y_n(a)$ для некоторого n>0, то x,y должны выражаться через x_{n-1}, y_{n-1} в соответствии с известными нам рекуррентными соотношениями, т.е. должно существовать решение (u,v) нашего уравнения такое, что

$$x = au + (a^2-1)v,$$

$$y = u+av.$$

Решая эту систему относительно u, v, получаем

$$u = ax - (a^2-1)y,$$

$$v = -x + ay. \qquad (3)$$

Таким образом, u, v – целые числа.

Упражнение 4.8. Проверьте, что $u^2-(a^2-1)v^2=1$ (т.е. что (u, v) является решением уравнения), а также, что 0<u<x и v≥0.

Итак, если пара (x, y) является решением уравнения $x^2-(a^2-1)y^2=1$, то числа x, y выражаются по формулам (3) через другое решение (u, v) этого уравнения, такое, что u<x. Если оказывается, что также u>1, то пара (u, v) выражается аналогично через решение (u', v') такое, что u'<u. "Спуск" может удаваться только конечное число (скажем, n) раз, и после этого будет достигнута ситуация, когда u≤1, т.е. $u=x_0(a)$ и $v=y_0(a)$ и, таким образом, $x=x_n(a) \wedge y=y_n(a)$.

Лемма 1 доказана.

Все это очень красиво, но почему уравнением Ферма заинтересовались, решая десятую проблему Гильберта? Заинтересовались им при поиске диофантова представления экспоненты. Найти такое представление, скажем, для предиката

$$Q(b) \leftrightarrow (\exists n)b=2^n$$

– это значит найти диофантово уравнение $P(b, z_1, ..., z_k)=0$ с параметром b, такое, что решение $(z_1, ..., z_k)$ существует, если и только если b

является степенью числа 2. Таким образом, диофантово условие P=0 должно "заставить" параметр b расти со скоростью экспоненты. Уравнение Ферма дает как раз нечто подобное.

ЛЕММА 2. При a>1 и n≥0

$$a^n \le x_n(a) \le \left(a+\sqrt{a^2-1}\right)^n \quad .$$

Д о к а з а т е л ь с т в о.

$$x_n(a)+y_n(a)\sqrt{a^2-1}=\left(a+\sqrt{a^2-1}\right)^n=a^n+C_n^1 a^{n-1}\sqrt{a^2-1}+\ldots \quad ,$$

что и требовалось доказать.

Таким образом, $x_n(a)$ растет по n со скоростью экспоненты (хотя и не следует точно ни одной из экспонент вида $(a+a_1)^n$), и это достигнуто через диофантово условие на x

$$(\exists y)(x^2-(a^2-1)y^2=1).$$

Именно поэтому уравнение Ферма интересно как отправная точка в поисках диофантова представления экспоненты. (Эти соображения принадлежат Дж. Робинсон и относятся еще к 1952 г.)

По идее Ю. В. Матиясевича, теперь мы должны провести исследование остатков, получающихся при взаимном делении чисел $x_n(a)$, $y_n(a)$.

Сначала, пусть n фиксировано, n>0 (при n=0 мы имели бы $x_0(a)=1$, и ничего интересного не получается). Будем изучать остатки от деления $x_N(a)$ на $x_n(a)$, где N=0, 1, 2, ... Для этого рассмотрим **по модулю** $x_n(a)$ известные рекуррентные соотношения для x_{m+n}, y_{m+n} (по модулю x_n – это значит, что мы будем пренебрегать слагаемыми, кратными x_n):

$$x_{m+n}(a) = x_m(a)\, x_n(a) + y_m(a)\, y_n(a)(a-1) \equiv (a^2-1)y_m y_n,$$

$$y_{m+n}(a) = x_m(a)\, y_n(a) + y_m(a)\, x_n(a) \equiv x_m y_n.$$

Подставляя m+n вместо m, получаем

$$x_{m+2n} \equiv (a^2-1)y_{m+n}\, y_n \equiv (a^2-1)x_m y_n^{\,2},$$

$$y_{m+2n} \equiv x_{m+n} y_n \equiv (a^2-1)y_m y_n^{\,2} .$$

Заметим теперь, что $x_n^{\,2}-(a^2-1)y_n^{\,2}=1$, т.е. по модулю x_n

$$(a^2-1)y_n^{\,2} \equiv x_n^{\,2} - 1 \equiv -1.$$

Подставляя вместо $(a^2–1)y_n$ число –1, получаем

$$x_{m+2n} \equiv – x_m, \qquad (4)$$

$$y_{m+2n} \equiv – y_m.$$

Подставляя здесь m+2n вместо m, имеем

$$x_{m+4n} \equiv – x_{m+2n} \equiv x_m,$$

$$y_{m+4n} \equiv – y_{m+2n} \equiv y_m.$$

Таким образом, остатки от деления $x_N(a)$ на $x_n(a)$ меняются по N с периодом 4n. Поэтому достаточно изучить поведение остатков при N=0, 1, 2, ..., 4n–1.

Согласно (4)

$$x_0 \equiv x_0,\ x_1 \equiv x_1,\ ...,\ x_{2n–1} \equiv x_{2n–1},$$

$$x_{2n} \equiv – x_0,\ x_{2n+1} \equiv – x_1,\ ...,\ x_{4n–1} \equiv – x_{2n–1}.$$

Аналогично для $y_N(a)$. Дело, однако, еще не доведено до конца, поскольку числа x_{n+1} , ..., $x_{2n–1}$, участвующие в характеристике остатков, все еще больше делителя x_n. Чтобы довести дело до конца, рассмотрим соотношения, выражающие x_{2n}, y_{2n} через $x_{2n–m}$, $y_{2n–m}$, x_m, y_m :

$$x_{2n} = x_{2n–m}\, x_m + (a^2–1)y_{2n–m}\, y_m,$$

$$y_{2n} = x_{2n–m}\, y_m + y_{2n–m}\, x_m.$$

Решая эту систему относительно $x_{2n–m}$, $y_{2n–m}$, получаем

$$x_{2n–m} = x_{2n}\, x_m – (a^2–1)y_{2n}\, y_m,$$

$$y_{2n–m} = y_{2n}\, x_{–m} – x_{2n}\, y_m.$$

Учитывая, что $x_{2n} \equiv – x_0 \equiv –1$, $y_{2n} \equiv – y_0 \equiv 0$ (по модулю x_n), имеем

$$x_{2n–m} \equiv – x_m,$$

$$y_{2n–m} \equiv y_{2m}.$$

Теперь нашу характеристику остатков от деления $x_N(a)$ на $x_n(a)$ (внутри периода 4n) можно довести до конца:

$$x_0 \equiv x_0,\ x_1 \equiv x_1,\ ...,\ x_{n–1} \equiv x_{n–1},$$

$$x_n \equiv -x_n,\ x_{n+1} \equiv -x_{n-1},\ \ldots,\ x_{2n-1} \equiv -x_1,$$

$$x_{2n} \equiv -x_0,\ x_{2n+1} \equiv -x_1,\ \ldots,\ x_{3n-1} \equiv -x_{n-1},$$

$$x_{3n} \equiv x_n,\ x_{3n+1} \equiv x_{n-1},\ \ldots,\ x_{4n-1} \equiv x_1.$$

Ясно, что вместо x_n и $-x_n$ здесь можно было написать просто нуль, но это нарушило бы симметрию.)

Имея такую характеристику, можно доказать следующую лемму Ю. В. Матиясевича.

ЛЕММА 3. Пусть $a\geq 3$, $n\geq 1$, $0<m<n$. Тогда для всех N

$$x_N(a) \equiv x_m(a) \bmod x_n(a) \leftrightarrow (N \equiv +m \bmod 4n) \vee (N \equiv -m \bmod 4n).$$

Д о к а з а т е л ь с т в о. 1) Влево. Если $N=4kn+m$ или $N=4kn-m$, то $x_N=x_m \bmod x_n$ вытекает непосредственно из полученной выше характеристики.

2) Вправо. Пусть известно, что $x_N \equiv x_m \bmod x_n$, где $0<m<n$. Разделим число N на 4n и найдем остаток: $N=4kn+m'$, где $0\leq m'<4n$. Если $0<m'<n$, то из нашей характеристики следует, что $m'=m$ и $N=4kn+m$, что и требовалось. Если $3n<m'$, то из характеристики следует, что $m'=4n-m$ и $N=4(k+1)n-m$.

Упражнение 4.9. Покажите, что третий случай $m'=0$ или $n\leq m'\leq 3n$ невозможен (учтите, что при $a>2$: $i<n \rightarrow x_i(a) < x_n(a)/2$).

Лемма 3 доказана.

Теперь мы должны провести аналогичное исследование остатков от деления $y_N(a)$ на $y_n(a)$ ($n\geq 1$ фиксировано, N=0, 1, 2, ...).

Упражнение 4.10. Проведите это исследование самостоятельно. Период получится длиной в 2n, а внутри периода остатки будут вести себя таким образом:

$$y_0 = y_0,\ y_1 = y_1,\ \ldots,\ y_{n-1} = y_{n-1},$$

$$y_n = -y_n,\ y_{n+1} = -y_{n-1},\ \ldots,\ y_{2n-1} = -y_1.$$

Из всего этого нас интересует только условие, при котором $y_N(a)$ делится на $y_n(a)$ (еще одна лемма Ю. В. Матиясевича):

ЛЕММА 4. Пусть $a\geq 2$, $n\geq 1$. Тогда $y_N(a)$ делится на $y_n(a)$, если и только если N делится на n.

Д о к а з а т е л ь с т в о вытекает непосредственно из полученной

характеристики.

Еще одна важная лемма Ю. В. Матиясевича дает условие, при котором $y_N(a)$ делится не только на $y_n(a)$, но и на $y_n^2(a)$.

ЛЕММА 5. Пусть $a\geq 2$. Тогда $y_N(a)$ делится на $y_n(a)$, если и только если N делится на $ny_n(a)$.

Д о к а з а т е л ь с т в о. Легко проверить (индукцией по k), что по модулю y_n^2

$$x_{kn} = x_n^k,$$

$$y_{kn} = kx_n^{k-1}y_n.$$

1) Импликация вправо. Если y_N делится на y_n^2, то по лемме 4: $N = kn$. Если y_{kn} делится на y_n^2, то число $kx_n^{k-1}y_n$ также должно делиться на y_n^2, т.е. kx_n^{k-1} должно делиться на y_N. Поскольку $x_n^2-(a^2-1)y_n^2=1$, то x_n не может иметь общих делителей с y_n, поэтому на y_n должно делиться само число k. Так как $N=kn$, то теперь ny_n делит N, что и требовалось.

2) Влево. Если ny_n делит N, то $N=kn$, где y_n делит k. Поэтому y_n^2 делит $kx_n^{k-1}y_n$, т.е. y_n^2 делит и $y_{kn} = y_N$, что и требовалось.

Лемма 5 доказана.

В дальнейшем нам потребуются еще три леммы. Первая из них принадлежит Дж. Робинсон, остальные две тривиальны.

ЛЕММА 6. При $a\geq 2$ и $n\geq 0$

$$x_n(a) \equiv 1 \bmod(a-1),$$

$$y_n(a) \equiv n \bmod(a-1).$$

ЛЕММА 7. При $a, a'\geq 2$ и $b\geq 1$, если $a \equiv a' \bmod b$, для всех n:

$$x_n(a) \equiv x_n(a') \bmod b,$$

$$y_n(a) \equiv y_n(a') \bmod b.$$

ЛЕММА 8. При $a\geq 2$ и $k\geq 0$ по модулю 2

$$x_{2k} \equiv 1,\ x_{2k+1} \equiv a,\ y_{2k} \equiv 0,\ y_{2k} \equiv 1.$$

Упражнение 4.11. Докажите эти леммы с помощью индукции.

4.4. Диофантово представление последовательности решений уравнения Ферма

Сейчас мы должны построить диофантово представление для предиката

$$Q(a, x, y, n) \leftrightarrow a\geq 3 \wedge x=x_n(a) \wedge y=y_n(a).$$

Какие "диофантовы условия" следует наложить на числа x, y, чтобы "заставить" их равняться $x_n(a)$ и $y_n(a)$? Прежде всего, разумеется, условие

$$\mathbf{E_1}: x^2 - (a^2-1)\, y^2 = 1.$$

Отсюда следует, что существует ν такое, что $x=x_\nu(a)$ и $y=y_\nu(a)$. Само значение ν мы пока не знаем. Но какие условия следует наложить на x, y, чтобы оказалось, что ν=n? Из леммы 6 мы знаем, что $y \equiv n \bmod(a-1)$, поэтому можно было бы потребовать $y = n \bmod(a-1)$, тогда отсюда следовало бы $\nu \equiv n \bmod(a-1)$. К сожалению, если $n\geq a-1$, то отсюда еще нельзя будет вывести, что ν=n.

Чтобы обойти эту трудность, приходится идти в обход в самом прямом смысле. Введем новое уравнение Ферма со свободным параметром a' и обозначим некоторое его решение через (x', y'):

$$\mathbf{E_2}: x'^2 - (a'^2-1)\, y'^2 = 1.$$

И теперь потребуем не $y \equiv n \bmod(a-1)$, а

$$\mathbf{E_3}: y' \equiv n \bmod(a'-1)$$

(в расчете, что a'–1 можно будет сделать больше n). Поскольку для некоторого ν' имеет место $x' = x_{\nu'}(a') \wedge y' = y_{\nu'}(a')$, то по лемме 6 $y' \equiv \nu' \bmod(a'-1)$ и поэтому

$$\nu' \equiv n \bmod(a'-1). \qquad (1)$$

Но так как мы "ушли в сторону" от исходного уравнения, от успеха пользы никакой не будет, если мы не сумеем найти "обратный путь" – нужно подходящим образом связать решение (x', y') с интересующим нас решением (x, y). Введем для этого новый модуль сравнения, обозначим его через X и потребуем, чтобы выполнялось условие

$$\mathbf{E_4}: a' \equiv a \bmod X \wedge x' \equiv x \bmod X.$$

Тогда новые числа a', x' не будут "слишком сильно отличаться" от старых a, x, причем, изменяя X, мы можем надеяться добиться максимально тесной связи. Теперь по лемме 7 $a' \equiv a \bmod X$ дает

$$x = x_\nu(a) \equiv x_\nu(a') \bmod X,$$

$$x' = x_{\nu'}(a') \equiv x_{\nu'}(a) \bmod X.$$

По условию $x' \equiv x \bmod X$, отсюда получается

$$x_{\nu'}(a) \equiv x_\nu(a) \bmod X. \qquad (2)$$

Чтобы создать условия для применения леммы 3 (она здесь сама напрашивается), модуль X следует сделать решением уравнения Ферма с параметром a. Введем поэтому еще одно число Y и условие

$$\mathbf{E_4}: X^2 - (a^2-1)\,Y^2 = 1.$$

Отсюда $X=x_N(a)$ и $Y=y_N(a)$ для некоторого N и (2) принимает вид

$$x_{\nu'}(a) \equiv x_\nu(a) \bmod x_N(a).$$

При $a\geq 3$ здесь можно было бы применить лемму 3, однако надо обеспечить еще $0<\nu<N$, поэтому введем условие

$$\mathbf{E_6}: 0<x<X$$

(поскольку $0<x_\nu(a)=x<X=x_N(a)$ и $x_i(a)$ возрастает по i).

Наконец, по лемме 3

$$\nu' \equiv \pm\nu \bmod 4N. \qquad (3)$$

Сравним это с (1):

$$\nu' \equiv n \bmod(a'-1).$$

Наша конечная цель – обеспечить $\nu=n$, т.е. оба последних сравнения нужно "свести вместе" к одному модулю. Для этого нужно найти достаточно большой общий делитель чисел 4N и $a'-1$. Числом $a'-1$ мы можем распоряжаться относительно свободно, но как получить делитель числа N, которое само нам неизвестно? Здесь помощь оказывает лемма 5: $y_\nu^2(a)$ делит $y_N(a)$, если и только если $\nu\cdot y_\nu(a)$ делит N. Короче: y^2 делит Y, если и только если $\nu\cdot y$ делит N. Поэтому, если мы потребуем

$$\mathbf{E_7}: y \text{ делит } Y,$$

то 4y будет делителем 4N (мы опускаем неизвестное число ν, которое не сумели бы сделать делителем модуля $a'-1$). Теперь нужно потребовать еще

$$\mathbf{E_8}: 4y \text{ делит } a'-1$$

(будем надеяться, что это требование не будет противоречить остальным условиям, наложенным на a'). Тогда (1) вместе с (3) дает

$$\nu' \equiv \pm\nu \bmod 4y \wedge \nu' \equiv n \bmod 4y$$

и отсюда

$$n \equiv \pm\nu \bmod 4y.$$

Другими словами, $n+\nu$ или $n-\nu$ делится на $4y$. Поскольку $y=y_\nu(a)$ возрастает по ν, то $y \geq \nu$, поэтому мы можем смело потребовать также

$$\mathbf{E_9}: n \leq y$$

(напомним, что мы добиваемся равенства $\nu=n$, т.е. желательно "наделить" n свойствами, присущими ν). Рассмотрим теперь отдельно две упомянутые возможности:

1) $n+\nu$ делится на $4y$. Поскольку $n+\nu \leq 2y$, то это возможно только при $n=\nu=0$, что и требовалось.

2) $n-\nu$ делится на $4y$. Поскольку $|n-\nu| \leq y$, то это возможно только при $n=\nu$, что и требовалось.

Вспомнив, что $x=x_\nu(a)$ и $y=y_\nu(a)$, мы можем утверждать теперь, что из условия

$$a \geq 3 \wedge \exists a'x'y'XY\ (E_1 \wedge E_2 \wedge E_3 \wedge E_4 \wedge E_5 \wedge E_6 \wedge E_7 \wedge E_8 \wedge E_9) \qquad (4)$$

вытекает, что $x=x_n(a)$ и $y=y_n(a)$, т.е. $Q(a, x, y, n)$.

Упражнение 4.12. Покажите, как (4) можно преобразовать в диофантово представление вида $(\exists)P=0$. Оцените число кванторов $\exists$, степень полинома P и сумму модулей его коэффициентов.

Решение нашей задачи будет завершено лишь, если удастся показать, что из $Q(a, x, y, n)$, т.е. из $a \geq 3 \wedge x=x_n(a) \wedge y=y_n(a)$ также следует условие (4). (В частности, только тогда и будет установлена взаимная непротиворечивость требований E_i.)

Итак, зная, что $a \geq 3 \wedge x=x_n(a) \wedge y=y_n(a)$, мы должны найти числа a', x' ,y' ,X, Y такие, что имеют место E_i для всех $i=1, 2, ..., 9$. Отметим сразу, что выполнение E_1 уже обеспечено леммой 1, а выполнение E_9 – тем обстоятельством, что $y_n(a) \geq n$ для всех n.

Числа X, Y (решение того же уравнения, что x, y) определим следующим образом: пусть N – любое число, делящееся на $ny_n(a)=ny$; возьмем $X=x_N(a)$ и $Y=y_N(a)$. Этим будет обеспечено условие E_5, а по лемме 5 тогда y делит Y, что обеспечивает E_7.

Остается указать число a', определяющее вспомогательное

уравнение, и его решение x', y'. При этом мы должны выполнить следующие условия:

$$\mathbf{E_2}: x'^2 - (a'^2-1)\, y'^2 = 1,$$

$$\mathbf{E_3}: y' \equiv n \bmod (a'-1),$$

$$\mathbf{E_4}: a' \equiv a \bmod X \wedge x' \equiv x \bmod X,$$

$$\mathbf{E_8}: a' \equiv 1 \bmod 4y.$$

Случай n=0 (тогда x=1, y=0) здесь приходится разбирать отдельно. Тогда E_8 требует a'=1, затем E_3 требует y'=0, E_2 – x'=1, наконец, E_4 требует $a \equiv 1 \bmod X$. Только последнее требование грозит "нарушить гармонию", но, к счастью, из-за y=0 мы вынуждены были выбрать N=0, поэтому X=1!

Пусть теперь n>0. Тогда y>0. Руководствуясь E_4 и E_7 , сначала укажем a'. Если бы модули X, 4y оказались взаимно простыми, существование числа a' вытекало бы из китайской теоремы об остатках (см. раздел 3.3). Убедимся, что X и 4y действительно взаимно просты. Во-первых, X должно быть нечетным числом. Этого можно добиться, если взять число N четным (лемма 8). (Напомним, что до сих пор мы требовали от N лишь делимость на ny.) Во-вторых, y и X не имеют общих делителей, поскольку y^2 делит Y, а $X^2-(a^2-1)Y^2=1$. Таким образом, существование числа a', удовлетворяющего условиям E_4 и E_8, обеспечено (причем, очевидно, можно выбрать a'>1).

Остается определить x', y'. Возьмем $x'=x_n(a')$ и $y'=y_n(a')$, тогда автоматически выполняется E_2, а по лемме 6 – и E_3. Наконец, так как $x=x_n(a)$ и $a' \equiv a \bmod X$, то по лемме $x' \equiv x \bmod X$, т.е. обеспечена вторая половина условия E_4.

Это все, что требовалось: из Q(a, x, y, n) мы вывели (4).

4.5. Диофантово представление экспоненты

Воспользуемся теперь полученным в предыдущем разделе диофантовым представлением предиката Q(a, x, y, n) для построения диофантова представления экспоненты, т.е. предиката n

$$E(u, v, n) \leftrightarrow u=v^n \wedge v \geq 3.$$

Будем следовать рассуждениям Дж. Робинсон, относящимся еще к 1952 г. (она установила тогда, что предикат $u=v^n$ "диофантово выразим" через предикат Q, но, разумеется, не имела диофантова представления для Q, полученного Ю. В. Матиясевичем 18 лет спустя).

Возьмем наше основополагающее равенство

$$\left(a+\sqrt{a^2-1}\right)^n = x_n(a)+y_n(a)\sqrt{a^2-1}$$

и обозначим $v=a+\sqrt{a^2-1}$. Тогда слева будем иметь просто v^n , а справа вместо $\sqrt{a^2-1}$ можем подставить v–a. Таким образом, уравнение

$$v^n - x_n(a) - y_n(a)\,(v - a) = 0$$

имеет корень $v_1=a+\sqrt{a^2-1}$. Поскольку коэффициенты уравнения рациональны, то число $v_2=a-\sqrt{a^2-1}$ также должно быть его корнем. С другой стороны, v_1, v_2 – корни квадратного уравнения

$$v^2 - 2av + 1 = 0.$$

Это значит, что трехчлен $v^2–2av+1$ является делителем полинома $v^n–x_n(a)–y_n(a)(v–a)$ в поле рациональных чисел. Более того, частное от этого деления должно быть полиномом с **целыми** коэффициентами (старший коэффициент делителя равен 1). Отсюда следует, что если v – целое число, то $v^n–x_n(a)–y_n(a)(v–a)$ делится на $v^2–2av+1$. В этом и состоит утверждение следующей леммы Дж. Робинсон:

ЛЕММА 9. При $a\ge1$ и $n\ge0$

$$v^n \equiv x_n(a)+y_n(a)(v - a) \bmod(v^2–2av+1).$$

Упражнение 4.13. а) Проверьте, что лемма 9 действительно выполняется и для n=0, 1 (наше рассуждение проходит только при $n\ge2$).

б) Проверьте, что на самом деле

$$v^n - x_n(a) - y_n(a)(v - a) = (v^2–2av+1)(y_1v^{n–2} + y_2v^{n–3} + ... + y_{n–2}v + y_{n–1}).$$

(Это будет по существу новым доказательством леммы 9, уже без использования алгебры полиномов.)

Значение леммы 9 состоит в том, что она связывает произвольную степень v с хорошо изученными числами $x_n(a)$, $y_n(a)$, причем связь эта достигается посредством полиномов **фиксированной** степени (v–a, $v^2–2av+1$). Отсюда легко получается диофантово представление для $u=v^n$.

В самом деле, мы должны исходить из произвольных u, v, n, и, накладывая на них диофантовы условия, добиться выполнения правенства $u=v^n$. Возьмем сначала числа a, x, y и подчиним их условию

$$\mathbf{F_1}: Q(a, x, y, n),$$

т.е. $x=x_n(a) \wedge y=y_n(a)$. Тогда по лемме 9

$$v^n \equiv x+y(v-a) \bmod(v^2-2av+1).$$

Чтобы как-то "привязать" число u к числу v , потребуем

$$\mathbf{F_2}: u \equiv x+y(v-a) \bmod(v^2-2av+1).$$

Тогда

$$u \equiv v^n \bmod(v^2-2av+1).$$

Чтобы отсюда можно было заключить, что $u=v^n$, надо обеспечить достаточную величину модуля – он должен быть больше как u, так и v^n. Этого можно добиться, увеличивая свободный параметр a – тогда модуль будет расти (по абсолютному значению). В частности, условие

$$\mathbf{F_3}: u < 2av-v^2-1$$

обеспечивает половину требуемого. Но как (с помощью диофантовых условий) добиться неравенства $v^n<2av-v^2-1$? Проще ли это по сравнению с обеспечением сразу равенства $u=v^n$? Проще, поскольку с помощью диофантовых условий мы умеем "заставить" величину расти **со скоростью** экспоненты, но не умеем пока "заставить" ее **в точности следовать** экспоненте. В самом деле, из леммы 2 мы знаем, что $x_n(v)\geq v^n$, а обеспечить условие $x_n(v)<2av-v^2-1$ очень легко. Вводим числа X, Y такие, что

$$\mathbf{F_4}: Q(v, X, Y, n),$$

т.е. $X=x_n(v) \wedge Y=y_n(v)$, и требуем, чтобы число a было настолько велико, чтобы имело место

$$\mathbf{F_5}: X < 2av -v^2 - 1.$$

Отсюда следует

$$v^n < x_n(v) = X < 2av-v^2-1,$$

что совместно с F_3 и (1) влечет $u=v^n$.

Таким образом, мы сумели вывести $u=v^n$ из условия

$$(\exists axyXY)\ (F_1 \wedge F_2 \wedge F_3 \wedge F_4 \wedge F_5). \qquad (2)$$

Одновременно мы обеспечили также неравенство $v \geq 3$ (оно содержится в F_4), т.е. из (2) выведена истинность предиката E(u, v, n).

Упражнение 4.14. а) Покажите, как преобразовать (2) в диофантово представление вида $(\exists)P=0$. Оцените число кванторов $\exists$, степень полинома P и сумму модулей его коэффициентов.

б) Чтобы завершить доказательство, покажите, что (2) вытекает из E(u, v, n), т.е. $u=v^n \wedge v \geq 3$.

Итак, следуя работам Ю. В. Матиясевича и Дж. Робинсон, мы получили для предиката $u=v^n \wedge v \geq 3$ диофантово представление вида

$$(\exists z_1 ... \exists z_k)\ P(u, v, n, z_1, ..., z_k) = 0.$$

Полагая v=3 и добавив квантор $\exists n$, мы получаем диофантово представление предиката:

"u является степенью числа 3".

Существует, таким образом, диофантово уравнение

$$P(u, v_1, ..., v_s)=0$$

с параметром u, которое имеет решения в натуральных числах, если и только если u имеет вид 3^n. Этот результат оказался неожиданным для многих специалистов по теории чисел.

4.6. Диофантовы представления числа сочетаний и факториала

Еще в 1952 г. Дж.Робинсон показала, что предикаты $x=C_y^z$, $x=z!$ диофантово выразимы через экспоненту. Теперь, имея диофантово представление экспоненты, можно построить диофантовы представления и этих предикатов.

Метод Дж. Робинсон, относящийся к сочетаниям, был усовершенствован Ю. В. Матиясевичем. Будем исходить из формулы бинома Ньютона

$$(1+p)^y = \sum_{z=0}^{y} C_y^z p^z \qquad (1)$$

При p=1 мы имели бы

$$2^y = \sum_{z=0}^{y} C_y^z \quad ,$$

таким образом, $C_y^z \le 2^y$ для всех z≤y. Поэтому если p – большое натуральное число, например, $p=3^y$, то $C_y^z < p$ для всех z≤y, и мы можем смотреть на числа C_y^z как на цифры в системе счисления с основанием p. Сумма в (1) представляет тогда p-ичную запись числа $(1+p)^y$. Если мы хотим теперь наложить на число x условие, приводящее к равенству $x= C_y^z$, надо потребовать, чтобы x было цифрой при степени p^z в записи числа $(1+p)^y$, т.е.

$$x < p \wedge (1+p)^y = u + xp^z + vp^{z+1},$$

где

$$u = \sum_{i=0}^{z-1} C_y^i p^i \quad , \quad v\,p^{z+1} = \sum_{i=z+1}^{y} C_y^i p^i \quad .$$

Число v однозначно определяется тем, что оно стоит у множителя p^{z+1} (как частное от деления $(1+p)^y$ на p^{z+1}). Для выделения же u (не обращаясь к формуле, содержащей C_y^i) мы должны потребовать, чтобы имело место неравенство $u<p^z$ (тогда u определяется однозначно как остаток от деления $(1+p)^y$ на p^z). В самом деле (учитывая, что $C_y^i < p$),

$$\sum_{i=0}^{z-1} C_y^i p^i \le (p-1)\sum_{i=0}^{z-1} p^i = (p-1)\frac{p^z-1}{p-1} < p^z \quad .$$

Таким образом, если $x=C_y^z \wedge z\le y$, то

$$(\exists puv)\ (p=3^y \wedge (1+p)^y = u + xp^z + vp^{z+1} \wedge x<p \wedge u<p^z). \qquad (2)$$

Покажем теперь, что из z≤y вместе с (2) вытекает $x=C_y^z$. Согласно (2) число u является остатком от деления $(1+p)^y$ на p^z. Согласно же (1) этот остаток выражается также суммой

$$\sum_{i=0}^{z-1} C_y^i p^i \quad ,$$

т.е. u равно этой сумме. Рассмотрим тогда число

$$q = \frac{(1+p)^y - u}{p^z} = x + vp \quad .$$

Так как x<p, то x является остатком от деления q на p. С другой стороны,

$$q = C_y^z + \sum_{i=z+1}^{y} C_y^i p^{i-z} \quad ,$$

где $C_y^z \le 2^y < p$ и большая сумма делится на p. Это значит, что C_y^z также является остатком от деления q на p, т.е. $x = C_y^z$, что и утребовалось.

Упражнение 4.15. Покажите, как преобразовать $z \le y \wedge (2)$ в диофантово представление вида $(\exists)P=0$. Оцените число кванторов $\exists$, степень полинома P и сумму модулей его коэффициентов.

Займемся, наконец, предикатом x=z!. Рассуждения Дж. Робинсон исходят из следующей идеи. Как известно,

$$C_y^z = \frac{y(y-1)\ldots(y-z+1)}{z!} \quad .$$

Если y значительно больше z, то числитель незначительно отличается от y^z, т.е. $z! \approx y^z / C_y^z$. Чтобы оценить необходимую величину y, изучим частное y^z / C_y^z подробнее:

$$\frac{y^z}{C_y^z} = z! \frac{y}{y} \cdot \frac{y}{y-1} \cdot \ldots \cdot \frac{y}{y-z+1} \quad .$$

Отсюда вытекает, что

$$z! \le \frac{y^z}{C_y^z} \le z! \left(\frac{y}{y-z}\right)^z = z! \left(1 + \frac{z}{y-z}\right)^z \quad .$$

Если возьмем y=z+zt, то

$$z! \le \frac{y^z}{C_y^z} \le z! \left(1 + \frac{1}{t}\right)^z = z! \left(1 + \sum_{i=1}^{z} C_z^i t^{-i}\right) \quad .$$

Учитывая, что $C_z^i \le 2^z$, возьмем $t = 2^z u$, тогда

$$z! \le \frac{y^z}{C_y^z} \le z! \left(1 + \frac{z}{u}\right) \quad .$$

Наконец, взяв $u = 2z \cdot z^z$, получаем (поскольку $z! \le z^z$)

$$z! \le \frac{y^z}{C_y^z} \le z! + \frac{1}{2} \quad .$$

Таким образом, если $y = z + 2z^2 2^z z^z$, то

$$z\,! = \left[\frac{y^z}{C_y^z} \right] \quad ,$$

где [ ] – символ целой части (наибольшее целое, не превосходящее число в скобках).

Упражнение 4.16. Покажите, как отсюда получить для предиката x=z! диофантово представление вида (∃)P=0. Оцените число кванторов ∃, степень полинома P и сумму модулей его коэффициентов.

В 1960 г. Х. Патнэм заметил, что всякое множество A натуральных чисел, имеющее диофантово представление, т.е. обладающее свойством

$$x \in A \leftrightarrow (\exists z_1 ... \exists z_n)\ P(x, z_1, ..., z_n) = 0$$

для некоторого полинома P с целыми коэффициентами, представимо как множество всех положительных значений некоторого (другого) полинома Q (также с целыми коэффициентами).

Упражнение 4.17. Покажите, что можно взять $Q = x\,(1 - P^2)$.

Упражнение 4.18. Исходя из теоремы Вильсона (см. А. А. Бухштаб [1966]):

$$\text{"x – простое число"} \leftrightarrow x>1 \wedge x^2 \text{ делит } x!+x,$$

постройте **диофантово представление для множества всех простых чисел**. Оцените, как всегда, число кванторов и степень полинома P под ними. Полином $Q = x\,(1 - P^2)$ представляет множество всех простых чисел как множество всех своих положительных значений. Таким образом, "формула для простых чисел", о невозможности которой все время говорили специалисты по теории чисел, в определенном смысле все-таки существует.

4.7. Устранение ограниченного квантора всеобщности

Теперь мы подошли к своей конечной цели – научиться преобразовывать формулы вида

$$(\forall y \le t)(\exists z_1 ... \exists z_n)\ P(a_1, ..., a_m, y, z_1, ..., z_n) = 0 \qquad (1)$$

(P – полином с целыми коэффициентами, t – полином первой степени с натуральными коэффициентами от a_1, a_2, ..., a_m) в диофантовы представления вида

$$(\exists v_1 ... \exists v_k)\ Q(a_1, ..., a_m, v_1, ..., v_k) = 0.$$

Именно этого не хватило нам в разделе 4.2 для преобразования формул, представляющих вычислимые функции, в диофантовы представления. Наши рассуждения будут следовать в основном по пути, намеченному в 1960–1961 гг. М. Дэвисом, Х. Патнэмом и Дж. Робинсон. Некоторые усовершенствования были внесены (уже после 1970 г.) совместно Ю. В. Матиясевичем и Дж. Робинсон.

При фиксированных a_1, ..., a_m формула (1) действительно представляет собой утверждение о **существовании** (несмотря на присутствие квантора всеобщности) – о существовании (t+1)n чисел (по n на каждое значение y = 0, 1, ..., t). Введем специальные обозначения для этих чисел:

$$y=0;\ z_1^{(0)}, ..., z_n^{(0)},$$

$$y=1;\ z_1^{(1)}, ..., z_n^{(1)},$$

...

$$y=t;\ z_1^{(t)}, ..., z_n^{(t)},$$

Избавиться от квантора $\forall\leq$ можно, закодировав эту таблицу ограниченным (не зависящим от t) числом натуральных чисел. Можно попытаться, например, закодировать одним числом каждый из n столбцов таблицы. Для кодирования можно использовать китайскую теорему об остатках. Если бы имелись попарно взаимно простые числа u_0, u_1, ..., u_z, то по этой теореме можно было бы подобрать числа w_1, ..., w_n такие, чтобы всякое $z_i^{(y)}$ оказалось остатком от деления w_i на u_y, т.е.

$$z_i^{(y)} < u_y \wedge w_i \equiv z_i^{(y)} \bmod u_y \qquad (2)$$

для всех $y\leq t$ и i=1, ..., n (для этого делители u_y должны быть, разумеется, достаточно большими).

Но если все это будет сделано, какие условия надо наложить на w_1, ..., w_n, чтобы остатки $z_i^{(y)}$ удовлетворяли уравнению (1)? Выбор у нас небольшой – попытаемся подставить числа w_1, ..., w_n в уравнение P=0 вместо чисел $z_i^{(y)}$. Вместо y подставим неопределенное пока число x. Что можно сказать о значении $P(a_1, ..., a_m, x, w_1, ..., w_n)$? Если в дополнение к (2) мы потребуем еще, чтобы для всех y = 0, 1, ..., t имело место

$$x \equiv y \bmod u_y, \qquad (3)$$

то можно будет утверждать, что

$$P(a_1, ..., a_m, x, w_1, ..., w_n) \equiv P(a_1, ..., a_m, y, z_1^{(y)}, ..., z_n^{(y)}) \bmod u_y.$$

Так как значение P справа равно нулю, то получаем

$$P(a_1, ..., a_m, x, w_1, ..., w_n) \equiv 0 \bmod u_y$$

для всех $y \le t$, т.е. значение слева делится на все числа u_y. Но эти числа попарно взаимно просты, поэтому значение P должно делиться и на их произведение:

$$P(a_1, ..., a_m, x, w_1, ..., w_n) \equiv 0 \bmod u_0 u_1 ... u_z. \qquad (4)$$

Посмотрим теперь на (4) с другой стороны – не как на следствие каких-то предположений, а как на **условие**, налагаемое на числа $w_1, ... ,w_n$. Тогда если числа $z_i^{(y)}$ – остатки от деления w_i на u_y, то согласно (1) и (2) должно иметь место

$$P(a_1, ..., a_m, y, z_1^{(y)}, ..., z_n^{(y)}) \equiv 0 \bmod u_y.$$

Чтобы получить отсюда не только $\equiv 0 \bmod u_y$, но и $= 0$, значение слева должно было меньше модуля u_y. Грубую оценку значения полинома P можно получить следующим образом. Через z обозначим число, которое больше всех чисел $z_i^{(y)}$, через N – степень полинома P, через M – сумму модулей его коэффициентов. Тогда

$$|P(a_1, ..., a_m, y, z_1^{(y)}, ..., z_n^{(y)})| \le M ((a_1+1)...(a_m+1)(t+1)(z+1))^N.$$

Упражнение 4.19. Убедитесь, что это действительно так.

Выражение справа обозначим через T. Чтобы интересующие нас значения полинома P оказались меньше всех u_y, мы должны употребовать:

а) чтобы остаток от деления чисел w_i на u_y всегда был i меньше z,

б) чтобы все u_y были больше T (зависящего от z). Кроме того, мы должны найти достаточно простой генератор (больших и попарно взаимно простых) чисел u_y. В принципе идею генератора можно было бы позаимствовать у бета-функции Геделя (см. раздел 3.3), определив

$$u_y = 1 + T\, t!\, (1+y),$$

тем более что диофантово представление факториала мы уже имеем. К сожалению, в таком случае придется проводить специальную работу, чтобы получить диофантово представление для произведения $u_0 u_1 ... u_z$, т.е. для модуля в сравнении (4). В свое время М. Дэвис, Х. Патнэм и Дж.

Робинсон приблизительно так и сделали.

Более удобным оказывается, однако, другой генератор чисел u_y, предложенный позднее Ю. В. Матиясевичем и Дж. Робинсон. Представим число C_v^{t+1}(при условии $t+1 \le v$) следующим образом:

$$C_v^{t+1} = \frac{v(v-1)\ldots(v-t)}{(t+1)!} = \left(\frac{v+1}{1}-1\right)\left(\frac{v+1}{2}-1\right)\ldots\left(\frac{v+1}{t+1}-1\right) \quad .$$

Если потребовать, чтобы v+1 делилось на (t+1)!, то все сомножители окажутся целыми числами. Если потребовать еще больше – чтобы v+1 делилось на $((t+1)!)^z$, то эти сомножители окажутся взаимно простыми.

Упражнение 4.20. Проверьте, что это действительно так (если d – общий делитель i-го и j-го сомножителей, рассмотрите их разность).

Итак, потребуем, чтобы v+1 делилось на $((t+1)!)^z$, и возьмем

$$u_y = \frac{v+1}{y+1} - 1 \quad .$$

Тогда с произведением $u_0 u_1 \ldots u_z$ никаких проблем не будет – оно равно C_v^{t+1}, а диофантово представление для числа сочетаний мы строить умеем. Теперь мы можем объединить все свои условия (*res(a, b)* означает "остаток от деления a на b"):

$\mathbf{G_1}$: $P(a_1, \ldots, a_m, x, w_1, \ldots, w_n) \equiv 0 \bmod C_v^{t+1}$,

$\mathbf{G_2}$: $(\forall y \le t)\ x \equiv y \bmod \left(\frac{v+1}{y+1}-1\right)$,

$\mathbf{G_{3i}}$: $(\forall y \le t)\ \mathrm{res}\left(w_i, \frac{v+1}{y+1}-1\right) < z$,

$\mathbf{G_4}$: $\frac{v+1}{y+1}-1 > M((a_1+1)\ldots(a_m+1)(t+1)(z+1))^N$,

$\mathbf{G_5}$: v+1 делится на $((t+1)!)^z$.

Упражнение 4.21. Покажите, что (1) равносильно формуле

$$\exists x\ \exists v\ \exists z\ \exists w_1 \ldots \exists w_n\ G_1 \wedge G_2 \wedge G_{31} \wedge \ldots \wedge G_{3n} \wedge G_4 \wedge G_5.$$

Не забудьте воспользоваться китайской теоремой об остатках при построении чисел $x, w_1, \ldots, w_n$.

К сожалению, последнюю формулу нельзя сразу преобразовать в диофантово представление. Мешают этому кванторы $\forall y \le t$ в условиях G_2,

G_{3i}. Однако в отличие от ∀y≤t в исходной формуле (1), здесь этот квантор стоит не над произвольным диофантовым представлением, а над конкретными и достаточно простыми предикатами. Можно надеяться, что эта простота позволит избавиться от кванторов ∀y≤t полностью.

Сначала заметим, что если взять v=x, то условие G_2 будет выполнено автоматически, т.е. в нашем списке требований его можно опустить. В самом деле,

$$\frac{x+1}{y+1}-1=\frac{x-y}{y+1}$$

является целым числом, которое делит x–y (частное равно y+1), т.е.

$$x \equiv y \bmod \left(\frac{x+1}{y+1}-1\right) .$$

Займемся теперь условиями G_{3i}. Если остаток от деления w_i на $\frac{x+1}{y+1}-1$ меньше z, то одно из чисел w_i, w_i–1, ..., w_i–z+1 делится на $\frac{x+1}{y+1}-1$, т.е. на это число делится и произведение

$$w_i(w_i-1)...(w_i-z+1)=\frac{w_i!}{(w_i-z)!} .$$

Поскольку числа $\frac{x+1}{y+1}-1$ для различных *y* попарно взаимно просты, то условие

$$\forall y\le t\ \left(\ \frac{x+1}{y+1}-1 \ \text{ делит } \ \frac{w_i!}{(w_i-z)!}\ \right)$$

равносильно условию

$$\mathbf{G_{3i}'}\colon C_x^{\,t+1} \text{ делит } \frac{w_i!}{(w_i-z)!}$$

(т.е. исчезает квантор ∀y≤t). Это тем более приятно, что мы умеем строить диофантово представление факториала и для нас не составит труда записать G_{3i}' диофантовыми средствами. Но, к сожалению, G_{3i}' не равносильно G_{3i} ! Из G_{3i} вытекает G_{3i}' (как мы только что убедились), однако из G_{3i}' **не** вытекает G_{3i} – если произведение делится на некоторое число, нет гарантии, что на это число будет делиться один из сомножителей. Что же тогда можно гарантировать?

Если число R делит прозведение $P_1P_2..\ P_k$, то R можно разложить в

прозведение $R_1R_2...R_k$, где каждое R_i делит свое P_i. Если R_i – наибольший из сомножителей R, то $R_i^k \geq R$ и $R_i \geq \sqrt[k]{R}$. Таким образом, если произведение $P_1P_2...P_k$ делится на R, можно утверждать только, что R и один из сомножителей P_i имеют общий делитель $\geq \sqrt[k]{R}$. Большего гарантировать нельзя.

Поэтому если вместо условий G_{3i} возьмем (с одной стороны, более удобные) условия G_{3i}', то можно гарантировать только, что некоторое $w_i - j$ (где $0 \leq j \leq z$) имеет общий делитель с числом

$$u_y = \frac{x+1}{y+1} - 1 \quad ,$$

который по величине $\geq \sqrt[z]{u_y}$. К счастью, этого оказывается достаточно. В самом деле, "пройдем" следующим образом от w_1 до w_n (при фиксированном $y \leq t$). При некотором $z_1^{(y)} < z$ разность $w_1 - z_1^{(y)}$ делится на некоторый делитель $S_1 \geq \sqrt[z]{u_y}$ числа u_y. Далее, при некотором $z_2^{(y)} < z$ разность $w_2 - z_2^{(y)}$ делится на некоторый делитель $S_2 \geq \sqrt[z]{S_1}$ числа S_1, т.е. $w_2 - z_2^{(y)}$ делится на делитель $S_2 \geq \sqrt[z^2]{u_y}$ числа u_y. И так далее, до разности $w_n - z_n^{(y)}$, которая делится на делитель $S_n \geq \sqrt[z^n]{u_y}$ числа u_y (здесь также $z_n^{(y)} < z$).

Подводя итог, заключаем, что для всех i=1, ..., n

$$w_i \equiv z_i^{(y)} \bmod S_n,$$

где S_n делит u_y (а следовательно, и C_i^{t+1}) и $S_n \geq \sqrt[z^n]{u_y}$. По условию G_1 имеем

$$P(a_1, ..., a_m, x, w_1, ..., w_n) \equiv 0 \bmod S_n.$$

Отсюда вытекает, что

$$P(a_1, ..., a_m, y, z_1^{(y)}, ..., z_n^{(y)}) \equiv 0 \bmod S_n,$$

причем все $z_i^{(y)} < z$. Значение слева должно будет равняться нулю, если оно будет (по абсолютному значению) меньше модуля S_n, т.е. если $\sqrt[z^n]{u_y}$ будет больше

$$T = M((a_1+1)...(a_m+1)(t+1)(z+1))^N.$$

Это значит, что условие G_4 можно заменить условием

$$\mathbf{G_4}'\colon\ \frac{x+1}{t+1}-1>T^{z^n}\quad .$$

Упражнение 4.22. Проверьте еще раз, что (1) действительно равносильно формуле

$$(\exists x\ \exists v\ \exists z\ \exists w_1\ ...\ \exists w_n)\ G_1' \wedge G_{31}' \wedge\ ...\ \wedge G_{3n}' \wedge G_4' \wedge G_5'.$$

где G_1', G_5' отличаются от G_1, G_5 тем, что вместо v подставлено x. Покажите, как преобразовать эту последнюю формулу в диофантово представление.

Тем самым задача устранения $\forall\leq$ решена до конца.

5. ТЕОРЕМЫ О НЕПОЛНОТЕ

5.1. Парадокс лжеца

Классической формой парадокса лжеца является высказывание Евбулида (IV в. до н.э.): "Я лгу". Евбулид предложил своим слушателям определить, является ли это высказывание истинным или ложным. И что же получается? Если предположить, что высказывание истинно, это означало бы, что Евбулид лжет, т.е. что его высказывание ложно. Если же предположить, что высказывание Евбулида ложно, это означало бы, что он не лжет, т.е. что он говорит правду и поэтому его высказывание истинно.

Казалось бы, утверждения вроде "Я пишу" **все** должны иметь определенное значение истинности (истинно или ложно). "Я лгу" является таким же утверждением, однако ему нельзя приписать ни истинность, ни ложность, не впадая в противоречие. Парадоксы такого рода вызывают горячие споры вот уже более двух тысяч лет. С одной стороны, ведутся поиски "законов правильной речи", которые высказывание Евбулида "нарушает" (и поэтому его вообще нельзя считать высказыванием). С другой стороны, всякий такой новонайденный закон сразу же подвергается сомнению (либо как запрещающий наряду с парадоксами также и совсем безобидные высказывания, либо как запрещающий одни парадоксы, но допускающий другие). Нас здесь будет интересовать именно эта другая сторона, так как она раскрывает творческий потенциал, заключенный в парадоксах.

Чтобы отмести возражение, что высказывание Евбулида "слишком неясно" (не ясно, относится ли оно к себе или к другим высказываниям Евбулида, сделанным в течение его жизни), французский логик XIV в. Иоанн Буридан предложил следующую форму парадокса: обозначим через p высказывание, содержащееся в рамке:

p: | p ложно |

Здесь уже не может быть никаких сомнений, что к чему относится, но противоречие все равно возникает.

Для тех, кто хотел бы объявить высказывание Евбулида неправильным как относящееся к самому себе ("самоссылающееся"),

немецкий логик XIV в. Альберт Саксонский предложил следующие парадоксы:

p_1:	p_2 ложно	q_1:	q_2 ложно
p_2:	p_1 истинно	q_2:	q_3 ложно
		q_3:	q_1 ложно

Здесь каждое высказывание относится не к себе, а к другому высказыванию. И тем не менее ни одно из них нельзя признать истинным или ложным, не впадая в противоречие.

Упражнение 5.1. Убедитесь, что это действительно так. (Подробнее о дискуссиях по поводу парадоксов в средние века см. книгу Н. И. Стяжкина [1967].)

Мы можем попытаться "принять" парадокс лжеца, вводя новую, более совершенную классификацию высказываний (взамен обычного подразделения на истинные и ложные):

а) истинные высказывания,

б) ложные высказывания,

в) высказывания, не имеющие значения истинности.

Сделав это, возьмем теперь высказывание

q:	q ложно или q не имеет значения истинности

Если q истинно,то q не ложно и q **имеет** значение истинности. Это значит, что высказывание, содержащееся в рамке, является ложным ..., но это и есть q! Аналогично, если q ложно, то q истинно. Наконец, если q не имеет значения истинности, то q истинно! Наша классификация высказываний опять не является исчерпывающей. (Последний парадокс принято называть Усиленным Лжецом.)

Упражнение 5.2. Вводя свою классификацию, мы по существу пытались заменить обычную двузначную логику трехзначной. С этой точки зрения, Лжец – парадокс двузначной, Усиленный Лжец – парадокс трехзначной логики. Сформулируйте аналогичные парадоксы для четырехзначной логики и т.д.

5.2. Лемма об автоссылках

Попытаемся воспроизвести рассмотренные выше парадоксы

средствами теории ЕА. Чтобы воспроизвести классический парадокс лжеца, мы должны построить формулу Q, "утверждение" которой состояло бы в том, что "Q можно опровергнуть в ЕА" (в теории ЕА ложным считается то, что можно опровергнуть исходя из аксиом). Но как добиться, чтобы формула относилась к себе самой, "говорила о себе"?

Формулы ЕА "умеют говорить" только о натуральных числах. Чтобы формула Q могла говорить "о формулах и о себе", все формулы должны быть **закодированы** натуральными числами. С этой целью нумеруем сначала все символы языка ЕА (будем считать, что переменные ЕА образованы из символов x, a следующим образом: x, xa, xaa,...):

x	a	0	1	+	*	=	(	)	¬	∧	∨	→	∃	∀
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Тогда всякая формула ЕА становится последовательностью натуральных чисел, например, x=1 превращается в 0, 6, 3. С помощью бета-функции Геделя любую такую последовательность можно закодировать двумя натуральными числами (впереди каждой последовательности ставится ее длина). Формулу x=1, например, кодируют числа a, b такие, что

β(a, b, 0)=3 (длина формулы),

β(a, b, 1)=0, β(a, b, 2)=6, β(a, b, 3)=3.

Мы знаем, что такие числа a, b существуют. Пару (a, b) легко закодировать одним числом, например

$$c = (a+b)^2 + a.$$

Упражнение 5.3. а) Покажите, как по числу c восстановить числа a, b.

б) Оцените сверху код формулы 0+1=1.

Итак, каждую формулу F из языка ЕА мы умеем кодировать одним натуральным числом. Соответствующий этому числу терм ЕА будем обозначать через **F** и называть **геделевым номером** F (в честь К. Геделя, который первым ввел кодирование формул числами, чтобы получить возможность "говорить" о формулах на языке арифметики). По формуле F мы умеем строить ее номер **F**, а по номеру – восстанавливать саму формулу.

Если теперь для некоторой формулы A(x) и другой формулы B удается установить, что ЕА ⊢ A(**B**), можно сказать: в теории ЕА доказано, что формула B "обладает свойством A". Если же удастся установить

$$\text{ЕА} \vdash B \leftrightarrow A(\mathbf{B}),$$

то получится: формула B "утверждает", что... она обладает свойством A.

ЛЕММА ОБ АВТОССЫЛКАХ. Для любой формулы A(x) из языка EA, имеющей единственную свободную переменную x, можно построить замкнутую формулу B (из языка EA), такую, что

$$EA \vdash B \leftrightarrow A(\mathbf{B}).$$

Д о к а з а т е л ь с т в о. Введем следующую так называемую **функцию подстановки** sub(x, y). Ее значением является геделев номер формулы, полученной из формулы с номером x подстановкой вместо всех свободных переменных терма y. Если x – не номер формулы, мы полагаем sub(x, y) = 0 для всех y.

Вне всякого сомнения, sub(x, y) – вычислимая функция. В самом деле, зная x, можно проверить, является ли x номером формулы. Если нет – полагаем значение функции равным нулю, если является – восстанавливаем по x формулу, находим ее свободные переменные и подставляем вместо них терм y. Наконец, находим номер полученной формулы, который и будет значением функции. Мы могли бы написать программу для вычисления sub(x, y), скажем, на языке *Pascal* (это была бы большая работа, но не очень трудная). Несколько труднее написать для вычисления sub(x, y) программу **машины Тьюринга**. Заниматься в деталях этим делом (осуществимость которого ясна каждому сколько-нибудь опытному программисту) мы здесь не будем. Лучше этого сошлемся на **тезис Черча**, согласно которому всякая функция, вычислимая в интуитивном смысле этого слова, вычислима и на подходящей машине Тьюринга.

Итак, будем считать, что машина Тьюринга, вычисляющая функцию sub, построена. Используя эту машину и следуя конструкциям, изложенным в доказательстве теоремы о представимости, можно построить в языке EA формулу SUB(x, y, z) такую, что для всех k, m, n: если sub(k, m)=n, то

а) $EA \vdash SUB(\mathbf{k}, \mathbf{m}, \mathbf{n})$,

б) $EA \vdash \neg(z=\mathbf{n}) \rightarrow \neg SUB(\mathbf{k}, \mathbf{m}, z)$.

Имея формулы SUB и A(x), введем, следуя К. Геделю, формулу $A_1(x)$:

$$(\forall z)(SUB(x, x, z) \rightarrow A(z)).$$

Обратите внимание на повторение переменной x – это ключевая идея! Что же "утверждается" формулой $A_1(x)$? Если взять формулу с номером x и подставить вместо ее свободных переменных терм x (т.е. номер самой формулы), то получится формула, обладающая свойством A. Мы знаем, что формуле SUB(x, x, z) удовлетворяет единственное z для каждого x, поэтому квантор $\forall z$ является здесь не более чем декорацией, которая, однако, необходима, поскольку в языке EA нет специального символа для

функции sub. Таким образом, в формуле $A_1(x)$ идет речь о некоторой операции подстановки, причем утверждается, что результат подстановки будет обладать свойством A. Попытаемся применить эту операцию к **самой формуле** $A_1(x)$. Обозначим через n номер этой формулы и подставим терм **n** вместо x. Полученную формулу $A_1(\mathbf{n})$ обозначим через B. Что же "утверждается" в B?

"Если взять формулу с номером n (т.е. $A_1(x)$) и подставить вместо x ее номер (т.е. терм **n**), то получится формула ($A_1(\mathbf{n})$, т.е. B), обладающая свойством A".

Итак, B "утверждает", что она обладает свойством A! Проследите еще два-три раза за этим рассуждением.

Чтобы завершить доказательство леммы об автоссылках, мы должны доказать в EA, что действительно $B \leftrightarrow A(\mathbf{B})$.

1. Покажем, что $EA \vdash B \rightarrow A(\mathbf{B})$. Значение sub(n, n) равно номеру формулы B, т.е. по определению представимости

$$EA \vdash SUB(\mathbf{n}, \mathbf{n}, \mathbf{B}), \quad EA \vdash \neg(z=\mathbf{B}) \rightarrow \neg SUB(\mathbf{n}, \mathbf{n}, z) \qquad (1)$$

Чтобы воспользоваться теоремой дедукции, возьмем в качестве гипотезы формулу B, т.е. $(\forall z)(SUB(\mathbf{n}, \mathbf{n}, z) \rightarrow A(z))$. Согласно (1) число z может быть равно только **B**, т.е. (чисто логическими рассуждениями) мы выводим отсюда $A(\mathbf{B})$. По теореме о дедукции это означает, что $EA \vdash B \rightarrow A(\mathbf{B})$.

2. Покажем, что $EA \vdash A(\mathbf{B}) \rightarrow B$. Имея $A(\mathbf{B})$ в качестве гипотезы, получаем формулу $SUB(\mathbf{n}, \mathbf{n}, B) \rightarrow A(B)$. С учетом (1):

$$(\forall z)(SUB(\mathbf{n}, \mathbf{n}, z) \rightarrow A(z)),$$

но это и есть формула B. По теореме дедукции отсюда следует, что $EA \vdash A(\mathbf{B}) \rightarrow B$.

Лемма об автоссылках доказана.

Таким образом, для любого свойства формул, которое мы умеем изобразить средствами EA, можно подобрать замкнутую формулу, "утверждающую", что она этим свойством обладает.

Об авторстве. В своей знаменитой работе, опубликованной в 1931 г., К. Гедель использовал конструкцию, которая составляет доказательство леммы об автоссылках, однако в общем виде он эту лемму не сформулировал. На возможность приведенной выше общей формулировки впервые обратил внимание Р. Карнап (см. М. Дэвис [1965]).

Упражнение 5.4 (А. Мостовский, 1961 г.). Покажите, что если A(x, y), B(x, y) – формулы EA с двумя свободными переменными, то найдутся замкнутые формулы C, D, такие, что

$$EA \vdash C \leftrightarrow A(\mathbf{C}, \mathbf{D}), \quad EA \vdash D \leftrightarrow B(\mathbf{C}, \mathbf{D}).$$

Если на самом деле A зависит только от y, а B – только от x, получается, что C и D "наговаривают друг на друга".

5.3. Теорема Геделя о неполноте

Итак, лемма об автоссылках позволяет, по-видимому, воспроизвести парадокс лжеца средствами EA. Какие это будет иметь последствия? Противоречие?

Формула-аналог утверждения "Я лгу" должна утверждать: "Меня можно опровергнуть в EA" (вместо истинности и ложности в EA фигурируют доказуемость и опровержимость). Если

F: "¬F доказуема в EA",

то

¬F: "¬F недоказуема в EA",

поэтому с тем же успехом мы можем рассматривать формулу

F: "F недоказуема в EA", (1)

она также "равносильна" парадоксу лжеца. Именно такой формулой занимался в свое время К. Гедель, и мы не будем нарушать традицию.

Формулу (1) мы могли бы получить из леммы об автоссылках, если бы сумели изобразить в виде формулы EA свойство "формулу с номером x можно доказать в EA".

Понятие формулы мы уже "вложили" в EA – каждая формула кодируется натуральным числом (ее номером). Доказательство в теории EA – это последовательность формул, удовлетворяющая специальным условиям. Если формулы представлены числами, то доказательства превращаются в последовательности чисел. Действуя так же, как в случае формул, мы можем закодировать одним натуральным числом любую последовательность формул. Это число естественно называть **геделевым номером** данной последовательности.

Разумеется, по номеру можно определить, является ли он номером доказательства, принадлежащего теории EA. В самом деле, по номеру последовательности можно восстановить все входящие в нее формулы и

порядок их расположения. Затем мы можем проверить, является ли каждая из этих формул аксиомой EA (логической или собственной) или же она получена из предыдущих формул последовательности с помощью правил вывода. Если для всех формул это так – анализируемый номер представляет доказательство. Можно построить даже машину Тьюринга, реализующую эту процедуру проверки без всякого вмешательства человека.

Таким же образом можно установить разрешимость предиката "y является номером EA-доказательства формулы с номером x". По теореме о представимости найдется формула, выражающая в EA этот предикат. Обозначим ее через PRF(x, y) (*proof* – доказательство).

Теперь можем построить формулу, которая утверждает "Меня нельзя доказать в EA". Взяв в лемме об автоссылках формулу ¬(∃y)PRF(x, y) (т.е. "формула с номером x не имеет доказательства в EA"), получим замкнутую формулу G, такую, что

$$EA \vdash G \leftrightarrow \neg(\exists y)PRF(\mathbf{G},y). \qquad (2)$$

Действительно, "утверждение" G состоит в том, что "G недоказуема в EA".

Попытаемся теперь выяснить, истинно или ложно то, что G "утверждает". С точки зрения теории EA истинно то, что доказуемо из аксиом EA, а ложно – то, что опровержимо с помощью этих аксиом. –

1. Предположим, что $EA \vdash G$. Пусть тогда n – номер доказательства формулы G. Формула PRF(x, y) выражает в EA предикат "y есть доказательство для x", поэтому $EA \vdash PRF(\mathbf{G}, \mathbf{n})$, а также – $EA \vdash (\exists y)PRF(\mathbf{G}, y)$. Однако согласно (2) эта последняя формула эквивалентна ¬G, т.е. получаем, что $EA \vdash \neg G$.

Итак, если дано доказательство (средствами EA) для формулы G, то найдется также доказательство и для отрицания ¬G, т.е. теория EA окажется в таком случае **противоречивой**. Ну, а "на самом деле" – G доказуема в EA? Этого мы не знаем. Если EA непротиворечива, то G нельзя доказать в EA. А "на самом деле" – EA непротиворечива? Этого мы также не знаем.

2. Предположим теперь, что $EA \vdash \neg G$. Согласно (2)

$$EA \vdash (\exists y)PRF(\mathbf{G}, y).$$

В нашем интуитивном понимании формула (∃y)PRF(G, y) утверждает, что в EA существует доказательство формулы G. Это вроде бы опять должно означать, что теория EA противоречива. Не может ли случиться, однако, что, доказав средствами EA формулу (∃y)PRF(G, y), мы, тем не менее, не в состоянии найти конкретное значение y? Разве, перебирая подряд: 0, 1,

2, ..., мы не должны натолкнуться однажды на номер доказательства G?

К сожалению, у нас нет достаточных оснований для такого заключения. Если мы найдем номер доказательства G, то теория EA окажется противоречивой. Но если не найдем? Тогда никакое n не будет номером EA-доказательства формулы G, т.е. для любого n: EA ├ ¬PRF(**G**, **n**). С другой стороны, мы знаем, что EA ├ (∃y)PRF(**G**, y). Противоречие? Не совсем, поскольку формула (∃y)PRF(**G**, y) противоречит формуле (∀y)¬PRF(**G**, y), но **ее** мы еще не доказали в EA. Все, что мы имеем пока, это **бесконечная серия** отдельных доказательств: для ¬PRF(**G**, **0**), для ¬PRF(**G**, **1**) и т.д. Можем ли мы надеяться свернуть всю серию в **единое конечное** EA-доказательство формулы (∀y)¬PRF(**G**, y)? До сих пор это никому не удалось.

Итак, из предположения EA ├ ¬G мы не сумели вывести противоречие в теории EA. Максимум, что мы можем утверждать: если EA ├ ¬G, то найдется формула C(y) с одной свободной переменной y такая, что

а) EA ├ (∃y)C(y),

б) для каждого n: EA ├ ¬C(**n**).

Формула C(y) не дает "настоящего" противоречия (доказательства формулы вместе с ее отрицанием). Однако, если подобная формула C(y) имеется, это означает все же, что в EA "не все в порядке". "Непорядок" такого рода принято называть **ω-противоречием**. (Это понятие было введено К. Геделем в ходе рассуждений, подобных приведенным выше.)

Упражнение 5.5. Покажите, что "настоящее" противоречие (т.е. формула D такая, что EA ├ D и одновременно EA ├ ¬D) означает также ω-противоречие.

Нами доказана знаменитая

ТЕОРЕМА ГЕДЕЛЯ О НЕПОЛНОТЕ (для теории EA). Можно построить замкнутую формулу G из языка EA, такую, что

а) если G доказуема в EA, то теория EA противоречива,

б) если ¬G доказуема в EA, то теория EA ω-противоречива.

Почему этой теореме придается такое большое значение? Сначала несколько общепринятых терминов. Замкнутую формулу F из языка теории T называют **неразрешимой в** T, если ни F, ни ¬F нельзя доказать средствами T (F предсказывает вполне определенное свойство "объектов" теории T, однако это предсказание нельзя средствами T ни доказать, ни опровергнуть).

Теорию, содержащую неразрешимые формулы, принято называть

неполной. Отсюда и название – "теорема о неполноте".

Не следует, однако, думать, что нами **доказана** неполнота теории EA. Неразрешимость средствами EA формулы G будет доказана только..., если удастся доказать, что теория EA **ω-непротиворечива** (т.е. что в ней не могут возникать ω-противоречия). До тех пор мы вправе утверждать только, что доказали **несовершенство** аксиом EA – эти аксиомы либо ω-противоречивы, либо с их помощью нельзя решить некоторые проблемы, касающиеся натуральных чисел (одна такая проблема выражена в формуле G – несмотря на все наши разговоры о том, что G "занимается" собственной доказуемостью, G – замкнутая формула в языке EA и как таковая выражает вполне определенное свойство натуральных чисел).

Несовершенную систему аксиом следует совершенствовать. Может быть, мы "забыли" какие-то важные аксиомы? Следует найти их, присоединить к аксиомам EA, и в результате мы получим... совершенную систему?

К сожалению, рассуждения К. Геделя проходят и для любого расширения EA. Как бы мы ни расширяли EA, в результате должна получиться все же некоторая формальная теория T (язык которой совпадает с языком EA) – предикат "y является номером T-доказательства формулы с номером x" должен быть разрешимым. (Возможность "механической" проверки доказательств является ли предложенный текст корректным доказательством, – это отличительная черта формальной теории.) Отсюда вытекает, что найдется формула $PRF_T(x, y)$, выражающая в EA этот предикат. Далее, по лемме об автоссылках можно будет получить замкнутую формулу G_T, такую, что

$$EA \vdash G_T \leftrightarrow \neg(\exists y)PRF_T(\mathbf{G_T}, y).$$

Т.е. формула G "утверждает", что она недоказуема в теории T.

Упражнение 5.6. Покажите, что если в теории T доказуемы все теоремы EA, то

а) если формула G_T доказуема в T, то эта теория противоречива,

б) если $\neg G_T$ доказуема в T, то эта теория ω-противоречива.

(Указание: повторите – с изменениями – рассуждения, доказывающие теорему Геделя для EA.)

Таким образом, никакие новые аксиомы не могут привести к "совершенной" системе аксиом арифметики. Метод Геделя позволяет доказать **принципиальное** несовершенство всякой системы аксиом арифметики: каждая такая система неизбежно является либо ω-

противоречивой, либо недостаточной для решения некоторых проблем, касающихся свойств натуральных чисел.

Курт Гедель (Kurt Gödel) родился 28 апреля 1906 г. в г. Брно (в то время Чехия входила в состав Австро-Венгрии и город назывался Брюнн). Высшее образование получил в Венском университете (изучал математику и физику). Там же в 1930 г. ему была присуждена ученая степень доктора философии. Свою знаменитую теорему о неполноте К. Гедель изложил 23 октября 1930 г. на заседании одной из секций Венской академии наук. Статья с развернутым изложением поступила в редакцию 17 ноября и вышла в следующем, 1931 г. (Оригинальные формулировки К. Геделя см. в книге В.А.Успенского [1982]. По поводу утверждений, что результат К. Геделя был в известной степени предвосхищен П. Финслером в 1926 г., см. статью З. А. Кузичевой [1970].) С 1933 по 1938 г. К. Гедель являлся доцентом математики Венского университета. Общался с философами-позитивистами Венского кружка (Р. Карнап и др.). Когда в 1938 г. Гитлер захватил Австрию, К. Гедель эмигрировал в США. С 1940 г. постоянно работал в Институте высших исследований г. Принстона (имел тесные контакты с А. Эйнштейном). Гражданство США получил в 1948 г. В 1949 г. К. Гедель предложил новый тип решения некоторых уравнений общей теории относительности, что было оценено А. Эйнштейном как "важный вклад" в эту теорию. Умер в США 14 января 1978 г.

Упражнение 5.7. Какого рода теоремы о неполноте получаются из парадоксов Альберта Саксонского, приведенных в разделе 5.1? Возьмите в помощь результат упражнения 5.4.

Понятие об ω-противоречивости несколько "портит" теорему К. Геделя о неполноте. Однако сам Гедель не пытался освободиться от него, впервые это удалось только Б. Россеру в 1936 г. – в теореме Россера вместо ω-противоречивости фигурирует "обычная" противоречивость.

ТЕОРЕМА ГЕДЕЛЯ В ФОРМЕ РОССЕРА. В языке всякой фундаментальной теории T найдется замкнутая формула R_T ("выражающая" некоторое свойство натуральных чисел), такая, что если $T \vdash R_T$ или $T \vdash \neg R_T$, то теория T противоречива.

В этой формулировке сделано еще одно усиление теоремы о неполноте, не имеющее отношения к методу Россера. Если теорему Геделя мы обсуждали только для тех формальных теорий, язык которых совпадает с языком ЕА, то сейчас речь идет о теории T с **произвольным** языком. Это освобождает нас от подозрений, что принципиальное несовершенство всякой системы аксиом арифметики кроется в неудачном выборе языка ЕА. Если для теорий с языком ЕА было достаточно

потребовать доказуемость всех теорем ЕА, то теперь – для теорий с произвольным языком – мы требуем относительную интерпретируемость ЕА в этих теориях. В разделе 3.2 мы назвали такие теории фундаментальными.

Д о к а з а т е л ь с т в о. Из упражнения 1.4 мы знаем, что фундаментальная теория в состоянии доказать только эффективно перечислимое множество формул языка ЕА. Построим машину Тьюринга, перечисляющую эти формулы:

$$F_0 , F_1 , F_2 , F_3 ,... \qquad (1)$$

(таким образом, $T \vdash \pi(F_н)$ для всех n, где π – отображение, переводящее формулы ЕА в формулы теории Т). Предикат

"формула с номером x появляется в перечислении (1) как F_t"

является разрешимым. Пусть формула $PRF_T(x, t)$ выражает в ЕА этот предикат (вводя обозначение $PRF_T(x, t)$, мы как бы считаем число t доказательством формулы x). Разрешим также предикат

"**отрицание** формулы с номером x появляется в перечислении как F_t".

Формулу, выражающую в ЕА этот предикат, обозначим через $REF_T(x, t)$ (*refutation* – опровержение).

Основная идея Б. Россера состояла в следующем: вместо формулы К .Геделя G_T, утверждающей: "Меня нельзя доказать в теории Т", взять формулу R_T , утверждающую:

"Меня легче опровергнуть, чем доказать в Т".

Если под "трудностью" доказательства формулы понимать номер места, в котором она появляется в перечислении (1), то формулу Б. Россера можно получить из леммы об автоссылках, взяв в качестве A(x) формулу

$$(\forall t)(PRF (x,t) \rightarrow (\exists z<t)REF (x,z)).$$

Получится замкнутая формула Q_T из языка ЕА, такая, что

$$EA \vdash Q_T \leftrightarrow (\forall t)(PRF_T(\mathbf{Q_T}, t) \rightarrow (\exists z<t)REF_T(\mathbf{Q_T}, z)). \qquad (2)$$

Покажем, что в качестве искомой формулы R_T (в языке теории Т) можно взять $\pi(Q_T)$.

1. Предположим, $T \vdash R_T$. Тогда формула Q_T появляется в перечислении (1), скажем, под номером n. Отсюда

$$EA \vdash PRF(\mathbf{Q_T}, \mathbf{n}). \qquad (3)$$

Поскольку в теории T доказуемы все теоремы EA (точнее, их переводы), то в ней доказуемы также формулы (2), (3) и их следствие

$$(\exists z<\mathbf{n})REF_T(\mathbf{Q_T}, z). \qquad (4)$$

Если $\neg Q_T$ действительно появляется в перечислении (1) под номером <n, то $T \vdash \neg R_T$ и теория T противоречива. Если же $\neg Q_T$ не появляется в первых n местах перечисления (1), то

$$EA \vdash \neg REF_T(\mathbf{Q_T}, \mathbf{0}) \wedge \neg REF_T(\mathbf{Q_T}, \mathbf{1}) \wedge \ldots \wedge \neg REF_T(\mathbf{Q_T}, \mathbf{n–1}).$$

Отсюда

$$EA \vdash \neg(\exists z<\mathbf{n})REF_T(\mathbf{Q_T}, z).$$

Эта формула, так же как (4), доказуема в теории T, вместе они дают противоречие в T, что и требовалось.

2. Предположим теперь, $T \vdash \neg R_T$. Тогда формула $\neg Q_T$ появляется в перечислении (1), скажем, под номером n. Отсюда

$$EA \vdash REF_T(\mathbf{Q_T}, \mathbf{n}). \qquad (5)$$

Если формула Q_T появляется в (1) раньше $\neg Q_T$, то $T \vdash R_T$ и получается противоречие в теории T. Если же Q_T не появляется в первых n местах перечисления (1), то

$$EA \vdash \neg PRF_T(\mathbf{Q_T}, \mathbf{0}) \wedge \neg PRF_T(\mathbf{Q_T}, \mathbf{1}) \wedge \ldots \wedge \neg PRF_T(\mathbf{Q_T}, \mathbf{n–1}),$$

$$EA \vdash \neg(\exists t<\mathbf{n})PRF_T(\mathbf{Q_T}, t). \qquad (6)$$

Далее, из (5) вытекает, что

$$EA \vdash (\forall t>\mathbf{n})(PRF_T(\mathbf{Q_T}, t) \rightarrow (\exists z<t)REF_T(\mathbf{Q_T}, z)),$$

поскольку при t>n в качестве z можно взять n. Совместно с (6) это дает

$$EA \vdash (\forall t)(PRF_T(\mathbf{Q_T}, t) \rightarrow (\exists z<t)REF_T(\mathbf{Q_T}, z)).$$

Вспомнив (2), можно получить $EA \vdash Q_T$ и $T \vdash R_T$, что опять означает противоречие в теории T.

Теорема Россера доказана.

Теперь мы можем сформулировать установленный К. Геделем "принцип несовершенства" в еще более сильной форме: **всякая фундаментальная теория несовершенна – она либо противоречива, либо недостаточна для решения всех возникающих в ней проблем.**

Фундаментальность теории здесь существенна – нефундаментальная теория **может** оказаться достаточной для решения всех возникающи_в ней проблем. В таких теориях "возникают", однако, только проблемы весьма специального вида (это следствие нефундаментальности, т.е. неспособности воспроизвести полноценное понятие натурального числа).

В качестве примера нефундаментальной теории можно назвать упомянутую в разделе 3.1 **арифметику Пресбургера**. В 1929 г. М. Пресбургер доказал полноту и непротиворечивость этой теории (полученной из ЕА удалением символа умножения). В силу теоремы Геделя-Россера тем самым была доказана и ее нефундаментальность.

5.4. Вторая теорема Геделя

Прежде чем перейти к методологической оценке революционного открытия, сделанного К. Геделем, представим в сжатой форме **чисто математическую** сторону вопроса.

Пусть Т – любая фундаментальная теория (т.е. формальная теория, в которой можно воспроизвести полноценное понятие натурального числа). По методу Геделя-Россера в языке Т строится замкнутая формула R_T, трактующая о свойствах натуральных чисел. Если бы R_T удалось доказать в теории Т, то по другому методу Геделя-Россера мы сумели бы вывести в Т противоречие. Если бы, с другой стороны, удалось формулу R_T опровергнуть средствами Т, то, по тому же методу, мы получили бы в теории Т противоречие. Эти два метода – метод построения формул R_T и метод, преобразующий всякое Т-доказательство или Т-опровержение R_T в Т-доказательство противоречия, и составляют чисто математическую сторону достижений К. Геделя и Б. Россера (без какой-либо философской оценки).

Итак, если Т – фундаментальная теория, то Т либо противоречива, либо недостаточна для решения вопроса об истинности формулы R_T (некоторого утверждения о свойствах натуральных чисел). Поскольку теорию, которая неспособна доказать либо опровергнуть некоторую замкнутую формулу из своего языка, принято называть **неполной**, теорему Геделя-Россера можно сформулировать более кратко: **всякая фундаментальная теория либо противоречива, либо неполна.**

Но почему эту теорему называют **теоремой о неполноте**? Ведь упомянутые выше методы Геделя и Россера не дают никаких средств для

решения, какая из двух возможностей имеет место в случае конкретной теории T – противоречивость или неполнота (противоречивая теория неизбежно оказывается "полной" – в ней доказуема любая формула). Поэтому, чтобы доказать "через Геделя" неполноту хотя бы одной формальной теории, нужно **доказать непротиворечивость** этой теории. Но мы уже знаем (см. раздел 1.6), что к доказательствам непротиворечивости следует предъявлять требования более строгие, чем к обычным математическим доказательствам: если доказывается непротиворечивость теории T, то мы вправе использовать только средства рассуждения, более надежные по сравнению со средствами, содержащимися в самой T. Но если речь идет о теории EA – о простейшей (и самой надежной) из фундаментальных теорий математики, то возможны ли еще более надежные средства рассуждения, не содержащиеся уже в EA? Если нет, то мы вынуждены доказывать непротиворечивость EA средствами самой же EA (или какой-либо "особенно надежной" частью этих средств). Возможно ли такое – теория доказывает свою собственную непротиворечивость?

Чтобы подойти к решению этого вопроса, формулировку его следует уточнить, переведя в формулы. В доказательстве теоремы Геделя-Россера мы построили для фундаментальной теории T формулу $PRF_T(x, y)$ (из языка EA), выражающую в EA предикат

"формула с номером x появляется в перечислении множества $\pi^{-1}(T)$ под номером t". (1)

Напомним, что $\pi^{-1}(T)$ – множество тех формул EA, которые (после перевода в язык T) доказуемы в теории T, поэтому, формула $(\exists t)PRF_T(x, t)$ "утверждает", что перевод EA-формулы с номером x доказуем в T.

Мы знаем, что если теория T противоречива, то в ней доказуема любая формула, в частности, перевод формулы 0=1. И наоборот, если **доказано**, что перевод 0=1 недоказуем в T, то доказана и непротиворечивость теории T. Поэтому формула $\neg(\exists t)PRF_T(\mathbf{0=1},t)$ "утверждает", что теория T непротиворечива. Обозначим эту формулу через Con(T) (*consistent* – согласована, непротиворечива).

Формула Con(T) зависит от выбора формулы PRF_T. Последнюю, оказывается, можно выбрать таким образом, что соответствующая Con(T) становится легко доказуемой (здесь мы следуем книге С. К. Клини [1957], с. 473). Пусть PRF_T – любая формула, выражающая в EA предикат (1). Введем другую формулу $PRF'_T(x, t)$, определяемую как конъюнкция:

$$PRF_T(x,t) \wedge \neg PRF_T(\mathbf{0=1},t).$$

Упражнение 5.8. Проверьте, что если теория T непротиворечива (т.е. перевод формулы 0=1 недоказуем в T), то формула PRF'_T выражает (в смысле строгого определения выразимости) в EA предикат (1).

Соответствующее PRF'_T утверждение о непротиворечивости T обозначим через Con'(T). В развернутом виде это формула

$$\neg(\exists t)(PRF_T(\mathbf{0=1},t) \wedge \neg PRF_T(\mathbf{0=1},t)).$$

Но ведь для доказательства **этой** формулы почти достаточно исчисления высказываний! Исчисление высказываний способно доказать противоречивость любой формальной теории? Да, способно, даже если сама теория противоречива – ведь Con'(T) можно доказать, не зная ничего о теории T. Поэтому грош цена таким "доказательствам". Но в чем же причина их кажущегося "успеха"? В том, что, доказывая Con'(T), мы на самом деле не можем утверждать, что эта формула выражает непротиворечивость теории T. В упражнении 5.8 мы доказали, что формула PRF'_T выражает предикат (1), **предполагая**, что T непротиворечива. Предполагая непротиворечивость, нетрудно... доказать ее в виде формулы Con'(T).

Из этого эксперимента можно извлечь все же один урок: нельзя, оказывается, обсуждать доказуемость (в какой-бы то ни было теории) формулы Con(T), предположив только, что эта формула получена из произвольной формулы PRF_T, выражающей предикат (1). Все рассуждения такого рода проходят и для формулы Con'(T), т.е. серъезных результатов таким путем получить не удастся. Если мы собираемся анализировать, какими средствами можно или нельзя доказать Con(T), надо проследить, какие средства использовались для установления того, что соответствующая формула PRF_T выражает предикат (1).

Оказывается, что если формула PRF_T получена "естественным путем", т.е. путем моделирования машины Тьюринга, перечисляющей $\pi^{-1}(T)$ (как в доказательстве теоремы о представимости), то для установления того, что эта формула выражает предикат (1), достаточно средств теории EA. Убедиться в этом с полной строгостью нелегко, хотя удивительного здесь ничего нет: ведь наши рассуждения о формулах и машинах Тьюринга идут в рамках средств, которые формализованы в EA.

Но какие средства нужны, чтобы доказать соответствующую ("естественную") формулу Con(T) (если это вообще возможно, т.е. если теория T "действительно" непротиворечива)? Предположим, что нам удалось каким-то образом доказать Con(T). Какие следствия можно извлечь отсюда? Единственными известными нам теоремами,

позволяющими извлечь хотя бы какие-то выводы из непротиворечивости теории, являются теоремы о неполноте. "Если теория T непротиворечива, то формула Геделя G_T недоказуема в T". Но ведь G_T и "утверждает", что она недоказуема в T! Т.е. если Con(T), то G_T. Формально, $Con(T) \rightarrow G_T$.

Эта импликация может быть не только формально записана, ее удается формально и доказать. Если формула PRF_T (исходя из которой строятся как Con(T), так и G_T) получена "естественным путем", то всегда оказывается, что

$$EA \vdash Con(T) \rightarrow G_T. \qquad (2)$$

Это неудивительно: ведь рассуждения К.Геделя используют вполне элементарные средства (но в остроумном сочетании), которые все формализованы в EA.

Если после всего этого удалось бы доказать средствами EA формулу Con(T) (т.е. непротиворечивость теории T), то в силу (2) отсюда следовало бы, что $EA \vdash G_T$. Поскольку T – фундаментальная теория, то мы получили бы $T \vdash G_T$. Однако из теоремы Геделя о неполноте мы знаем, что доказуемость G_T влечет противоречие в теории T. Таким образом, если EA в состоянии доказать непротиворечивость T, то "на самом деле" ... T противоречива. В частности, если удастся доказать средствами EA непротиворечивость самой теории EA, то тем самым в EA будет найдено противоречие!

Этот вывод К. Геделя (его вторая теорема в исторической статье, опубликованной в 1931 г.) показывает, что программа Гильберта (см. раздел 1.6) не может быть реализована до конца. Вспомним, что программа включала два этапа:

а) построение формальный теории, охватывающей всю существующую математику,

б) доказательство непротиворечивости полученной теории (в этом доказательстве Д. Гильберт предполагал использовать только надежные средства рассуждения, не выходящие за пределы EA).

Осуществление этапа а) удалось – созданная к началу 30-х гг. формальная теория множеств Цермело-Френкеля действительно охватывает всю математику. Однако трудности, встретившиеся на этапе б), оказались принципиальными – средствами, формализованными в EA, нельзя доказать даже непротиворечивость самой EA, не говоря уже о доказательстве непротиворечивости всей математики.

Может быть, неудача Д. Гильберта объясняется узостью средств

рассуждения, которые он допускал в доказательствах непротиворечивости? Обобщение результатов К. Геделя показывает, что причина лежит гораздо глубже. И этим обобщениям мы обязаны в значительной мере самому Д. Гильберту (идея приведенной ниже формулировки второй теоремы Геделя заимствована из книги "Основания математики", написанной совместно Д. Гильбертом и П. Бернайсом [1934, 1939]).

Вместо формулы $PRF_T(x, t)$, выражающей предикат (1), введем в рассмотрение формулу $(\exists t)PRF_T(x, t)$, которую будем обозначать через $PR_T(x)$. Очевидно, что $PR_T(x)$ означает утверждение, что перевод (π-образ) формулы с номером x доказуем в теории T. Основное свойство формулы Геделя G_T можно теперь представить так: G_T – любая формула, обладающая свойством

$$T \vdash G_T \leftrightarrow \neg PR(\mathbf{G_T}).$$

Формула Con(T) выражается через $PR_T(x)$ как $\neg PR_T(\mathbf{0=1})$, где **0=1** – номер формулы 0=1.

Забудем теперь о происхождении формулы $PR_T(x)$ и будем говорить, что теория T "понимает", что формула Con(T) выражает ее непротиворечивость, если соответствующая формула $PR_T(x)$ удовлетворяет следующим **условиям Гильберта:** для любых формул A, B из языка EA

1. если $T \vdash A$, то $T \vdash PR_T(\mathbf{A})$,
2. $T \vdash PR_T(\mathbf{A}) \rightarrow PR_T(\mathbf{PR_T(A)})$,
3. $T \vdash PR_T(\mathbf{A}) \wedge PR_T(\mathbf{A \rightarrow B}) \rightarrow PR_T(\mathbf{B})$.

По условию 2) теория T "сознает", что формула PR_T выражает понятие T-доказуемости: ведь смысл импликации 2) состоит в том, что "если $T \vdash A$, то $T \vdash PR_T(\mathbf{A})$". По условию 3) теория T "сознает", что множество ее (арифметических) теорем замкнуто относительно правила MODUS PONENS: "если $T \vdash A$ и $T \vdash A \rightarrow B$, то $T \vdash B$".

ВТОРАЯ ТЕОРЕМА ГЕДЕЛЯ. Пусть фундаментальная теория T "понимает", что формула Con(T) выражает ее непротиворечивость. Тогда либо T противоречива, либо Con(T) недоказуема в T.

Д о к а з а т е л ь с т в о. Проведем формальное (в теории T) доказательство первой части теоремы Геделя о неполноте: если $T \vdash G_T$, то теория T противоречива. Другими словами, покажем, что

$$T \vdash PR_T(\mathbf{G_T}) \rightarrow PR_T(\mathbf{0=1}). \qquad (3)$$

Так как $T \vdash \neg G_T \leftrightarrow PR_T(\mathbf{G_T})$, то отсюда будет следовать, что $T \vdash Con(T) \rightarrow G_T$. Поэтому если $T \vdash Con(T)$, то $T \vdash G_T$, что сразу влечет противоречивость теории T.

Итак, остается доказать (3). Воспользуемся теоремой дедукции: возьмем в качестве гипотезы формулу $PR_T(\mathbf{G_T})$ и попытаемся вывести $PR_T(\mathbf{0=1})$. По условию 2)

$$T \vdash PR_T(\mathbf{G_T}) \rightarrow PR_T(\mathbf{PR_T(G_T)}).$$

При нашей гипотезе тогда получается $PR_T(\mathbf{PR_T(G_T)})$. Поскольку формула Геделя обладает свойством $PR_T(\mathbf{G_T}) \rightarrow \neg G$, то по условию 1)

$$T \vdash PR_T(\mathbf{PR_T(G_T)} \rightarrow \neg \mathbf{G_T}).$$

По условию 3) имеем

$$T \vdash PR_T(\mathbf{PR_T(G_T)}) \wedge PR_T(\mathbf{PR_T(G_T)} \rightarrow \neg \mathbf{G_T}) \rightarrow PR_T(\neg \mathbf{G_T}).$$

Таким образом, из нашей гипотезы выводится формула $PR_T(\neg \mathbf{G_T})$. Теперь, имея в виду саму гипотезу $PR_T(\mathbf{G_T})$, заметим, что

$$T \vdash G_T \rightarrow (\neg G_T \rightarrow 0=1)$$

("из противоречия следует все, что угодно"). Повторяя дважды проделанную только что манипуляцию с условием 3), получаем формулу $PR_T(\mathbf{0=1})$, что и требовалось. По теореме дедукции отсюда следует (3).

Вторая теорема Геделя доказана.,

Вернемся теперь к нашей "патологической" формуле Con(T), которую можно было доказывать почти в исчислении высказываний. Если бы формула $PR'_T(x)$, определяемая как $(\exists t)PRF'_T(x, t)$, удовлетворяла условиям Гильберта, то по второй теореме Геделя отсюда следовало бы, что теория T противоречива. С точки зрения нашего определения это означает, что если T непротиворечива, то она "не понимает", что Con(T) выражает ее непротиворечивость. Может доказать Con(T), но не понимает ее!

Если же формула Con(T) получена "естественным путем", то всегда оказывается, что условия Гильберта выполнены (и даже более того – для доказательств, существование которых требуется в этих условиях, достаточно средств EA). Соответственно для таких формул Con(T) справедливо и заключение второй теоремы Геделя: либо теория T

противоречива, либо Con(T) нельзя доказать в T. Таким образом, если некоторая фундаментальная теория может доказать свою собственную непротиворечивость, то... эта теория противоречива. Если необходимым элементом обоснования теории считать доказательство ее непротиворечивости, этот вывод можно сформулировать еще более эффектно: **фундаментальная теория не может сама себя обосновать**.

Ну, а **нефундаментальные** теории? Они не в состоянии даже **поставить** "своими силами" проблему своего обоснования. Либо их язык не позволяет написать что-либо подобное формуле Con(T), либо (в случае достаточно богатого языка) их аксиомы не позволяют доказать требуемое в условиях Гильберта (по нашей терминологии, это означает, что такие теории "не понимают" свою формулу Con(T)).

В некоторых случаях одна фундаментальная теория способна доказать непротиворечивость другой ("более слабой"). Так, в теории множеств ZF можно доказать непротиворечивость EA (множество ω оказывается моделью, в которой выполняются все аксиомы EA, см. приложение I). Формула Con(EA) недоказуема в EA (если эта теория непротиворечива), однако перевод ее в язык теории множеств можно доказать с помощью аксиом ZF. Формула Con(EA) – замкнутая формула языка EA, т.е. вполне определенное утверждение о свойствах натуральных чисел. Это утверждение в EA недоказуемо, но его можно доказать в теории множеств. Это ответ на вопрос, сформулированный в разделе 3.2: некоторые утверждения о свойствах натуральных чисел, требующие для своей **формулировки** только понятие натурального числа (и, казалось бы, касающиеся только этих чисел), требуют для своего **доказательства** сложные понятия, не укладывающиеся в рамки EA (более впечатляющий пример см. в приложении 2).

6. ВОКРУГ ТЕОРЕМЫ ГЕДЕЛЯ

6.1. Методологическое значение теорем о неполноте

Нередко из теорем о неполноте пытаются сделать вывод о принципиальном превосходстве "живого, творческого, содержательного, человеческого мышления" над любой формальной теорией, о невозможности исчерпать фиксированной формальной теорией "все богатство содержательной математики". С этим можно было бы согласиться, если бы указанное "превосходство" не понималось как способность "творческого мышления" **безошибочно** находить истины, которые нельзя доказать в заданной формальной теории. Сторонники подобной точки зрения обычно рисуют следующую картину.

Возьмем любую формальную теорию T, которая содержит полноценное понятие натурального числа (т.е. фундаментальную теорию, по нашей терминологии). Формула Геделя G_T, построенная для теории T, утверждает: "Я недоказуема в T". И К. Гедель показал, что она действительно недоказуема в T. Таким образом, К. Гедель умел доказывать **истинность** формулы G_T (которая – как формула из языка EA – является также утверждением о свойствах натуральных чисел). Это значит, что какую бы формальную теорию T мы ни взяли, К. Гедель с помощью своего "творческого мышления" докажет нам истинность некоторого утверждения G_T о свойствах натуральных чисел, которое нельзя доказать в теории T. Таким образом, никакая формальная теория не в состоянии отразить в себе "живое, человеческое" понятие о ряде натуральных чисел (не говоря уже об остальных "богатствах содержательной математики"). Как только формальная теория T зафиксирована, "творческий разум" **безошибочно** находит истину G_T, выводящую за пределы этой теории.

То, что мы узнали о теореме Геделя в предыдущих разделах, должно заставить нас кое-что в этом рассуждении пересмотреть. К. Гедель смог доказать, что формула G_T недоказуема в теории T, только **предположив**, что эта теория непротиворечива. Т.е. он выводил истинность G из **предположения** непротиворечивости T. Иначе и быть не могло: если доказана истинность G_T, то тем самым доказана

непротиворечивость теории T (истинность G_T означает, что эта формула недоказуема в T, недоказуемость же в T хотя бы одной формулы означает непротиворечивость T). Таким образом, не зная ничего о противоречивости или непротиворечивости теории T, мы не можем ничего сказать и об истинности или ложности формулы G . Как же могут относиться к вопросу о непротиворечивости те, кто считает G_T безошибочно найденной истиной, недоказуемой в T?

Во-первых, они не могут полагать, что **всякая** формальная теория непротиворечива. Искусственный пример противоречивой теории получить легко: добавим к теории EA аксиому 0=1. Так как в EA доказуема формула ¬(0=1), то полученная теория противоречива. Однако с такого рода искусственными примерами мало кто станет считаться. Более серъезно необходимо относиться к следующему факту: не существует алгоритма, который по системе аксиом и правил вывода определял бы, противоречива ли соответствующая формальная теория.

Упражнение 6.1. Предполагая, что теория EA непротиворечива, покажите (ср. раздел 6.3), что невозможен алгоритм, определяющий по замкнутой формуле F, противоречива теория EA+F или нет. (Теория EA+F получена присоединением к EA формулы F в качестве аксиомы.)

Таким образом, проблема непротиворечивости теории не может решаться механически, одним и тем же методом во всех случаях. Каждая теория требует в этом вопросе **конкретного** подхода.

Наконец, следует упомянуть о теориях, которые в свое время их создатели считали непротиворечивыми, но которые позднее оказались все же противоречивыми. Такая судьба постигла первую в истории формальную систему математики, разработанную еще в прошлом веке Г. Фреге. В 1902 г., когда второй, завершающий том книги с изложением системы Г. Фреге находился уже в печати, Б.Рассел обнаружил, что эта система содержит противоречие. Г. Фреге прожил до 1925 г., но серьезных работ больше не опубликовал – такой страшный удар нанес парадокс Рассела по абсолютной уверенности автора в безупречности исходных посылок его системы. (Следует, однако, заметить, что Г. Фреге зря принял этот удар как лично его касающийся. Знакомство с принципами его системы производит на всякого непредубежденного человека, не знающего о парадоксе Рассела, такое же впечатление абсолютной надежности, как на самого Г. Фреге. В этом смысле Г. Фреге принял на себя удар, "причитающийся" всему предшествующему математическому образу мышления.)

Несколькими годами раньше такая же судьба постигла другое великолепное творение математики XIX в. – теорию множеств Г.

Кантора. Так же как система Г. Фреге, теория Г. Кантора производит при знакомстве с ее основами впечатление абсолютной надежности и фундаментальности. Кажется, тут не может быть места сомнениям. Принципы математического мышления доведены у Г. Кантора до логического конца – понятия произвольного бесконечного множества. И тем не менее в 1895 г. Г. Кантор сам обнаруживает, что из его принципов можно вывести противоречие...

Со времен Г. Фреге и Г. Кантора формальные теории математики значительно усовершенствованы. Противоречия в них пока не найдены. Тем не менее горький опыт этих великих мыслителей должен был научить нас по крайней мере одному: никакая наша "внутренняя уверенность" в надежности исходных посылок теории (сколько бы людей ни разделяли эту уверенность) не может быть абсолютной гарантией от противоречий.

Поскольку решение вопроса об истинности утверждений, выдвигаемых в качестве "истин", недоказуемых в какой-либо формальной теории, сопряжено с решением вопроса о непротиворечивости этой теории, то ни о каком общем методе, безошибочно находящем "истины", выводящие за пределы любой заданной формальной теории, речи быть не может. Вдруг теория, которую мы собираемся "превзойти", окажется противоречивой? Тогда формула Геделя, построенная для этой теории, окажется ложной. Какая же это "превосходящая истина"?

Пожалуй, наиболее ярко нелепость критикуемой точки зрения проявляется в следующем "силлогизме":

а) Чтобы "превзойти через Геделя" формальную теорию T, нужно доказать истинность формулы G_T.

б) Чтобы доказать истинность G_T, нужно доказать непротиворечивость теории T.

в) Чтобы доказать непротиворечивость T, нужно применить средства, выходящие за пределы этой теории (вторая теорема Геделя).

В итоге имеем: чтобы "превзойти через Геделя" формальную теорию, ... нужно применить средства, выходящие за пределы этой теории. Какая великолепная... тавтология!

Будучи не в состоянии доказать непротиворечивость какой-либо теории, но будучи уверенными, что "на самом деле" так оно и есть, мы можем попытаться принять утверждение о непротиворечивости в качестве **новой аксиомы**. И изучить следствия, вытекающие отсюда. Такой подход не является чем-то необычным для математики (в теории чисел изучаются следствия гипотезы Римана, в теории множеств –

следствия континуум-гипотезы Кантора и т.п.). Заметим, однако, что здесь все время говорится о **гипотезах**. Такой же гипотезой следует считать и утверждение о непротиворечивости теории, в надежности которой мы уверены. Нужно всегда помнить об этом: принятие гипотезы в качестве аксиомы – это **постулирование**, т.е. принятие без достаточных на то оснований. В ходе изучения следствий из гипотезы могут быть обнаружены противоречия, и тогда гипотезу придется отвергнуть.

Из всего сказанного можно сделать вывод, что теоремы о неполноте не дают никакого общего метода, позволяющего безошибочно (с первой попытки) находить истины, недоказуемые в заданной формальной теории. И, по-видимому, такого общего метода вообще не существует: серьезные теории настолько сложны, что "превосходящие" их средства приходится подбирать для каждой из них **конкретно.**

Методологическое значение теорем о неполноте заключается совсем в другом. **Всякая фундаментальная теория либо противоречива, либо недостаточна для решения некоторых возникающих в ней проблем**. Мы уже знаем, что методы К. Геделя и его последователей не дают средств, позволяющих решить, является ли конкретная теория противоречивой или неполной. Поэтому точнее было бы сказать не "теоремы о неполноте", а "**теоремы о несовершенстве**". Всякая фундаментальная теория несовершенна – она либо противоречива, либо недостаточна для решения некоторых возникающих в ней проблем.

Несовершенную теорию необходимо совершенствовать. Возникшие противоречия должны устраняться через совершенствование аксиом теории. Проблемы, неразрешимость которых в данной теории доказана (или подозревается), следует решать, испытывая различные **гипотезы** в качестве дополнительных аксиом (т.е. на шаткой основе, а вовсе не на основе способности неформального мышления сразу, с первой попытки находить новые "надежные" истины, выводящие за рамки старых аксиом).

Всякая формальная теория с методологической точки зрения является моделью некоторой **застывшей системы мышления**. С учетом этого основной вывод из теорем о неполноте можно переформулировать так: всякая достаточно всеобъемлющая (фундаментальная?), но застывшая система мышления неизбежно оказывается несовершенной – в ней содержатся либо противоречия, либо проблемы, для решения которых данной (застывшей!) системы недостаточно. **Именно в строгом доказательстве принципиального несовершенства всякой застывшей системы мышления состоит подлинный диалектический смысл достижений Геделя**. (Однако отсюда не следует, что незастывшая

система мышления может существовать в качестве **математической** теории.)

Замечание. Иногда теории определяются как произвольные множества формул, называемых теоремами. Такая точка зрения очень абстрактна. Сущность теории не исчерпывается множеством ее теорем. Теория – это достаточно определенная **система мышления** (а формальная теория – модель абсолютно определенной, т.е. застывшей, системы мышления), а не только совокупность результатов (теорем), которые можно получить с помощью средств данной системы, она включает и сами эти средства. Волновая механика Э. Шредингера и матричная механика В. Гейзенберга – различные теории, хотя доказано, что по своим результатам они совпадают. С упомянутой абстрактной точки зрения (сводящей теорию к множеству ее теорем), противоречивая теория кажется "пустой" (в ней доказуемы все формулы без разбора, т.е. такая теория как будто не делает различия между истиной и ложью). Однако, если, совершенствуя аксиомы теории, противоречия (обнаруженного типа) удается исключить, то неужели при этом совершенствуется "пустота"? Если смотреть на теорию **конкретно –** как на систему мышления, противоречия представляются уже не как "пустота", а как **несовершенство** аксиом теории. Несовершенные аксиомы должны совершенствоваться.

6.2. Теорема о двойной неполноте

Неразрешимость формулы Б. Россера R_T, построенной для теории T, можно было доказать, предположив только непротиворечивость T. В остальном рассуждения Б. Россера проходят в рамках теории EA. Это значит, что доказательство неразрешимости формулы R_T можно формализовать в теории EA+Con(T) (т.е. в теории, полученной присоединением к EA гипотезы о непротиворечивости T). Теорию, с помощью которой изучаются свойства другой теории, принято называть **метатеорией**. Итак, неразрешимость R_T можно доказать в метатеории EA+Con(T). Вероятно, в этой метатеории можно доказать неразрешимость (средствами теории T) и многих других формул. Не может ли оказаться, однако, что неразрешимость некоторых формул недоказуема в EA+Con(T) (т.е. для установления их неразрешимости недостаточно предположить только непротиворечивость теории T)?

Ответ на этот вопрос можно получить путем моделирования парадокса **Усиленный Лжец** (см. раздел 5.1):

q: | q ложно или q не имеет значения истинности |

Все три возможные альтернативы (q истинно, q ложно, q неразрешимо) приводят к противоречию. Если теория T обсуждается в метатеории M, мы можем попытаться получить формулу H, которая утверждала бы: "H опровержима в T или в M доказуема T-неразрешимость H". Это действительно можно сделать, в результате получается первый ("геделевский") вариант теоремы о двойной неполноте: если теории T, M ω-непротиворечивы, то формула H неразрешима в теории T, однако этот факт нельзя доказать в метатеории M (см. К. М. Подниекс [1975]). Отсюда и название – "теорема о двойной неполноте".

Ниже мы докажем сразу усиленный ("россеровский") вариант этой теоремы (К. М. Подниекс [1976]). Сначала уточним смысл отношения "M – метатеория для T". Пусть T, M – фундаментальные теории. Через (N_T, π_T), (N_M, π_M) будем обозначать относительные интерпретации теории EA в теориях T, M соответственно. Будем говорить, что M – **метатеория для** T, если зафиксированы формулы $PR_T(x)$, $RF_T(x)$ из языка EA такие, что для любой формулы F (из языка EA):

а) если $T \vdash \pi_T(F)$, то $M \vdash \pi_M(PR_T(\mathbf{F}))$,

б) если $T \vdash \pi_T(\neg F)$, то $M \vdash \pi_M(RF_T(\mathbf{F}))$.

Таким образом, теория M "знает кое-что" об арифметических утверждениях, которые доказуемы или опровержимы в теории T. Чтобы не загромождать запись символами π_T, π_M, мы будем далее писать просто

$$T \vdash F, \quad T \vdash \neg F, \quad M \vdash PR_T(F), \quad M \vdash RF_T(F)$$

вместо $T \vdash \pi_T(F)$, $M \vdash \pi_M(PR_T(\mathbf{F}))$ и т.п. Такая вольность записи допустима, если мы интересуемся только **арифметическими** утверждениями, которые доказуемы в теориях T, M.

ТЕОРЕМА О ДВОЙНОЙ НЕПОЛНОТЕ. Пусть T, M – фундаментальные теории, причем M – метатеория для T. Тогда найдется замкнутая формула H из языка EA, такая, что если теории T, M непротиворечивы, то H неразрешима в T, однако в M нельзя доказать ни $\neg PR_T(\mathbf{H})$, ни $\neg RF_T(\mathbf{H})$ (т.е. в метатеории M нельзя доказать ни T-недоказуемость, ни T-неопровержимость формулы H).

Д о к а з а т е л ь с т в о. Возьмем машину, которая перечисляет все арифметические теоремы теорий T, M:

$$(A_0, T)\ (A_1, M)\ (A_2, M)\ (A_3, T)...$$

Появление пары (A_i, T) означает $T \vdash A_i$, появление пары (A_i, M) – $M \vdash A_i$. Такая перечисляющая машина существует, поскольку множество всех теорем всякой формальной теории эффективно перечислимо. Наша цель – получить формулу H такую, что ни одна из следующих четырех выводимостей невозможна:

$$T \vdash H, \quad T \vdash \neg H, \quad M \vdash \neg PR_T(\mathbf{H}), \quad M \vdash \neg RF_T(\mathbf{H}). \qquad (1)$$

Поэтому будем называть формулу Q из языка EA **позитивной**, если в указанном перечислении раньше появляется пара (Q, T) или $(\neg RF_T(\mathbf{Q}), M)$, и **негативной –** если раньше появляется $(\neg Q, T)$ или $(\neg PR_T(\mathbf{Q}), M)$. (Формула H, которую мы ищем, не должна быть ни позитивной, ни негативной.) Номер (в перечислении) первой появившейся пары будем называть (позитивным или негативным) номером формулы Q. Очевидно, следующие предикаты разрешимы:

a(x, y) = "y есть **позитивный** номер формулы с номером x",

b(x, y) = "y есть **негативный** номер формулы с номером x".

Пусть формулы A(x, y), B(x, y) выражают эти предикаты в теории EA.

Следуя методу Россера, возьмем теперь формулу

$$(\forall y)(A(x,y) \rightarrow (\exists z<y)B(x,z))$$

и применим лемму об автоссылках. В результате получится замкнутая формула H, обладающая свойствами

$$EA \vdash H \leftrightarrow (\forall y)(A(\mathbf{H},y) \rightarrow (\exists z<y)B(\mathbf{H},z))$$

("Если я позитивна, то я негативна, причем с меньшим номером".)

Упражнение 6.2. Следуя доказательству теоремы Россера (см. раздел 5.3), покажите, что выводимости (1) несовместимы с непротиворечивостью теорий T, M.

Теорема о двойной неполноте доказана.

Если взять M=EA+Con(T), т.е. если обсуждать теорию T средствами EA, предполагая непротиворечивость T, то оказывается, что существуют формулы, неразрешимость которых нельзя доказать исходя только из этого предположения. Для доказательства неразрешимости таких формул (если они получены из теоремы о двойной неполноте) нужно предположить Con(EA+Con(T)), т.е. непротиворечивость самой метатеории. Это ответ на вопрос, сформулированный в начале раздела.

Открытие явления неполноты привело к новому методу развития математических теорий. Если в теории T не удается утверждение F ни доказать, ни опровергнуть, мы можем попытаться принять F (или ¬F) в

качестве новой аксиомы. Однако такой прием связан с определенной опасностью: не исключено, что **в будущем** утверждение F будет доказано (тогда теория T+¬F окажется противоречивой) или опровергнуто (тогда противоречивой окажется теория T+F). Но до того как это произошло, трудно сделать выбор с гарантией.

Хотелось бы, принимая F в качестве новой аксиомы, "обезопасить" себя от возможных противоречий. Как это сделать? Нужно **доказать непротиворечивость** новой теории T+F. Однако из второй теоремы Геделя мы знаем, что **абсолютное** доказательство непротиворечивости фундаментальной теории невозможно (для такого доказательства нужны средства, выходящие за рамки самой теории, т.е. средства, менее надежные, чем сама теория). Таким образом, **абсолютные** гарантии от противоречий невозможны. Но мы можем попытаться получить гарантию **относительную –** доказать, что принятие новой аксиомы не порождает новых противоречий (кроме тех, которые, возможно, уже содержатся в "старой" теории T). Чтобы такому доказательству можно было доверять, его следует проводить в рамках теории T, предполагая только ее непротиворечивость. Т.е. речь идет о доказательстве непротиворечивости теории T+F средствами метатеории T+Con(T). Если оно удается, то, обнаружив противоречие, вытекающее из принятой новой аксиомы F, мы сумеем преобразовать вывод этого противоречия в вывод противоречия, не зависящий от F, в вывод противоречия средствами теории T (которой **сейчас** неограниченно доверяем). Таким образом, если непротиворечивость теории T+F доказана в метатеории T+Con(T), то тем самым доказано, что принятие F в качестве аксиомы столь же "безопасно", как принятие аксиом теории T.

Это новый подход к развитию математических теорий, возможность которого была осознана в начале XIX в. – с изобретением неевклидовых геометрий. Будучи не в состоянии решить вопрос об истинности утверждения F на основе общепризнанных аксиом, мы пытаемся доказать, что принятие F (или ¬F) столь же "безопасно", как принятие общепризнанных аксиом. И затем присоединяем F (или ¬F) к этим аксиомам. (Заметим, что здесь появляется возможность **ветвления** в развитии теории – можно изучать как теорию T+F, так и T+¬F, ср. раздел 2.4.)

Раньше казалось, что единственный способ развития математической теории – вывод все новых следствий при **неизменном** списке аксиом. Теоремы о неполноте показали принципиальную недостаточность этого подхода – некоторые проблемы неизбежно не смогут быть решены, оставаясь в рамках существующих аксиом. При упомянутом новом подходе мы уже не считаем список аксиом неприкосновенным – к нему разрешается добавлять новые аксиомы, даже противоречащие друг

другу (разумеется, не вместе, а порознь – получая из одной теории несколько различных). Неизменной здесь должна оставаться только "безопасность" от противоречий – аксиому F разрешается принять, если средствами метатеории T+Con(T) доказана непротиворечивость теории T+F. Таким образом, отвергая принцип "неизменных аксиом", новый подход сохраняет принцип "неизменной безопасности".

Теорема о двойной неполноте показывает недостаточность и этого принципа (он совершеннее принципа неизменных аксиом, но все же не является абсолютно совершенным). В самом деле, положив M=T+Con(T), мы получаем формулу H, которая неразрешима в теории T, однако средствами теории M нельзя доказать ни $\neg PR_T(\mathbf{H})$ (непротиворечивость теории T+¬H), ни $\neg RF_T(\mathbf{H})$ (непротиворечивость теории T+H). Аксиома детерминированности (AD, см. раздел 2.4) является, возможно, примером аксиомы, "безопасность" принятия которой, т.е.

$$Con(ZF) \rightarrow Con(ZF+AD),$$

доказать никогда не удастся.

Интересно отметить, что явление "двойной неполноты" предвидел еще в 1926 г. П. Леви (см. П. Леви [1926]).

6.3. Проблема творчества в математике

Все математические теоремы выводятся из фиксированного набора аксиом. Иногда говорят по этому поводу, что в математике не может появиться ничего такого, чего уже ранее не было в аксиомах. Итак, в процессе развития математики не может появиться "ничего нового"?

Как относиться к такого рода рассуждениям? Если новым считать новое положение, которое привносится в исходные принципы теории (и **не вытекает** из этих принципов), то действительно – отличительной чертой **математических** теорий является именно застывшая система исходных принципов. Новое исходное положение приводит к новой теории. Однако если новое понимать так узко, то "откровение" предыдущего абзаца становится тавтологией.

Но, может быть, здесь хотят утверждать, что в математике не может быть ничего нового, поскольку математика – "механическая" наука, в которой аксиомы и правила вывода, а не "живой человек" определяют, что и как делать?

Мы уже знаем, что множество всех теорем формальной теории является эффективно перечислимым – что в принципе осуществимо

механическое устройство (машина Тьюринга), печатающее на бесконечной ленте все теоремы:

$$F_0 , F_1 , F_2 , F_3 ,...$$

Если просидеть достаточно долго возле этой ленты, можно дождаться любой наперед заданной теоремы (если мы заранее уверены, что теорема действительно доказуема в данной теории). Вся теория, казалось бы, заключена в механическом устройстве...

Но если решение задачи нам неизвестно, мы высказали некоторое предположение F и хотим узнать, истинно ли оно (рассуждая в рамках теории T)? Т.е. нас интересует, будет ли $T \vdash F$ или $T \vdash \neg F$. Здесь указанное перечисляющее устройство мало чем может помочь: из теорем о неполноте мы знаем, что возможна ситуация, когда ни $T \vdash F$, ни $T \vdash \neg F$ (формула F неразрешима в теории T). В этом случае мы можем просидеть у ленты сколько угодно долго, но ни формула F, ни ¬F напечатаны не будут – решения задачи от машины мы не получим.

Вообще говоря, творчество можно было бы считать устраненным из теории T, если бы удалось "механизировать" различение формул трех родов:

а) доказуемые в T,

б) опровержимые в T,

в) неразрешимые в T.

Разумеется, если теория T противоречива, то в ней доказуема любая формула, и классы а), б) в этом случае совпадают, а класс в) оказывается пустым. Если противоречие в T действительно найдено (найдена формула A такая, что $T \vdash A$ и $T \vdash \neg A$ одновременно), то по аксиоме исчисления высказываний $\neg A \rightarrow (A \rightarrow B)$ можно уже без всяких усилий доказать произвольную формулу B. Таким образом, после того как противоречие в теории T найдено, в рассуждениях этой теории пропадает всякий творческий момент, все удается без труда "механизировать". (Правда, само нахождение противоречия **в серьезной теории** обычно бывает настоящим творческим актом – см., например, раздел 2.2.)

Отсюда следует, что проблему "механизации" различения классов а)–в) можно обсуждать только в такой метатеории, которая способна доказать непротиворечивость теории T (или эту непротиворечивость **постулирует**). Так мы и будем поступать.

Итак, пусть T – фундаментальная формальная теория. Предполагая непротиворечивость T, покажем, что класс а) **эффективно-неотделим** (см. дальше) от класса б). Отсюда будет следовать, что "механизация"

различения классов а)–в) невозможна (и, в частности, что классы а),б), являясь эффективно перечислимыми, не являются, тем не менее, разрешимыми, а класс в) не является даже эффективно перечислимым).

Допустим от противного, что классы а), б) эффективно отделимы. Тогда существует разрешимый предикат s(x) (x пробегает натуральные числа), такой, что для любого n, являющегося номером формулы в языке EA,

1) если s(n), то перевод формулы с номером n нельзя опровергнуть в теории T (т.е. этот перевод не принадлежит классу б)).

2) если не s(n), то перевод формулы с номером n нельзя доказать в T (т.е. этот перевод не принадлежит классу а)).

Таким образом, предикат s(x) отделяет множество π^{-1}(а) от π^{-1}(б). Пусть истинность предиката s(x) распознается машиной Тьюринга M. Через C(x, t), D(x, t) обозначим формулы, выражающие в теории EA рекурсивные предикаты:

"машина M, работая с аргументом x, останавливается ровно за t шагов и выдает результат "истинно",

"машина M, работая с аргументом x, останавливается ровно за t шагов и выдает результат "ложно".

Используя идею конструкции Россера, по лемме об автоссылках можно получить формулу E такую, что

$$EA \vdash E \leftrightarrow (\forall t)(C(\mathbf{E}, t) \rightarrow (\exists z<t)D(\mathbf{E}, z)).$$

Упражнение 6.3. Следуя доказательству теоремы Россера (см. раздел 5.3), покажите, что если s(E) истинно, то $EA \vdash \neg E$, а если s(E) ложно, то $EA \vdash E$.

Таким образом, значение s(E) не может быть ни истинным, ни ложным (оба предположения приводят к противоречию). Отсюда вытекает, что предикаты, отделяющие T-доказуемые формулы от T-опровержимых формул, не могут быть разрешимыми. В частности, не может быть речи о "механизации" различения классов а)–в).

Если, работая в фундаментальной теории T, мы задались гипотезой F и хотим решить, будет ли $T \vdash F$ или $T \vdash \neg F$ (или, быть может, F неразрешима средствами T), то мы должны искать те **конкретные** особенности своей гипотезы, которые делают ее доказуемой, опровержимой или неразрешимой в теории T. **Общего метода**, пригодного для решения всевозможных таких вопросов, как мы только что показали, не существует. Ну а конкретный подход к проблеме (при условии невозможности общего метода) требует самого настоящего

творчества. Подметив особенности гипотезы H_1, которые делают ее доказуемой в теории T, мы не можем с уверенностью рассчитывать, что этих наших идей будет достаточно для решения вопроса об истинности другой гипотезы H_2 и т.д.

(Эта часть нашей новой "философии творчества" действительна лишь в предположении непротиворечивости теории T. Но если T "на самом деле" противоречива – и мы об этом не знаем, нам предстоит совершить также творческий акт – **найти это противоречие**. Ведь из упражнения 6.1 мы знаем, что не существует общего метода, который определял бы по описанию любой теории, противоречива она или нет. Итак, без творчества не обойтись и в этом случае.)

Таким образом, на новом, более высоком, методологическом уровне (на уровне формальных теорий, являющихся моделями математических теорий, которые – модели – сами допускают изучение математическими средствами) подтвердилось положение, известное любому человеку, достаточно хорошо знакомому с историей математики. А именно: никакой фиксированный метод, никакая ограниченная "идейная база" не дают возможности решать все проблемы математики (даже если аксиомы остаются неизменными и достаточными). Все время появляются новые проблемы, решение которых требует новых идей. В этом смысле математика – неисчерпаемый источник идей.

6.4. Теорема о сокращении доказательств

Специалисты по теории чисел уже давно заметили, что некоторые известные теоремы могут быть доказаны более просто, если исходить из **гипотезы Б. Римана**. Эта гипотеза (касающаяся расположения комплексных нулей дзета-функции Римана) не доказана до сих пор (хотя и хорошо проверена эмпирически с помощью ЭВМ). Однако ее принятие значительно упрощает доказательства некоторых теорем, которые могут быть доказаны и без нее (но более сложным путем).

Если мы присоединяем к теории T в качестве новой аксиомы гипотезу H, которая неразрешима в T, то получаем новую теорию T+H, которая "сильнее" T. И нет ничего удивительного в том, что в T+H не только можно доказывать новые теоремы (недоказуемые ранее в T), но к тому же еще многое, что в T делалось с трудом, в T+H становится проще.

То, что сказано в двух предыдущих абзацах, нельзя все же считать **строгим** обоснованием возможности этого явления. Если теорема как-то доказана в теории T и (более просто) – в теории T+H, то еще нельзя

считать доказанным, что эту теорему нельзя доказать проще и в **самой теории** Т. Недостающее строгое обоснование было дано К. Геделем в 1936 г.

ТЕОРЕМА О СОКРАЩЕНИИ ДОКАЗАТЕЛЬСТВ. Пусть Т – фундаментальная теория, в которой недоказуема замкнутая формула L. Тогда для любой вычислимой функции f(x), которая монотонно стремится к бесконечности, найдется теорема К теории Т такая, что $l_{T+L}(K) < f(l_T(K))$.

Формулировка нуждается, видимо, в комментариях. Во-первых, символы $l_{T+L}(K)$, $l_T(K)$ обозначают "длину" кратчайшего доказательства теоремы К соответственно в теориях Т+L и Т (более подробно о понятии длины доказательства см. дальше). Во-вторых, о роли функции f(x). Если взять конкретно f(x)=[x/100], то по теореме о сокращении доказательств найдется теорема К теории Т такая, что $l_{T+L}(K) < \frac{l_T(K)}{100}$ (при условии, что дополнительная аксиома L недоказуема в Т). Т.е. кратчайшее доказательство К в теории Т+L более чем в сто раз короче кратчайшего доказательства этой теоремы в теории Т. Можно взять в качестве f(x) еще более медленно растущие функции, например [x/1000]], $[\sqrt{x}]$, $[\log_{10} x]$ и т.д.

Теперь более подробно о понятии длины доказательства, которое используется в теореме. Дело в том, что эта теорема остается справедливой при **любом** естественном способе измерения длины доказательств. Самый простой способ измерения – по количеству символов. Доказательство – это некоторая последовательность формул $F_0, F_1, ..., F_n$, оканчивающаяся интересующей нас теоремой. Если через |F| обозначить число символов, образующих формулу F , длину нашего доказательства мы могли бы определить как сумму $|F_0|+|F_1|+ ... +|F_n|$.

Достаточными условиями для способа измерения длины доказательств, при которых оказывается справедливой теорема о сокращении, являются следующие:

а) длину доказательства можно вычислить, зная само доказательство (т.е. существует алгоритм, который по каждому Т-доказательству вычисляет его длину),

б) для любого числа t существует только конечное число доказательств длины $\leq t$, и все эти доказательства можно выписать, зная t (т.е. существует алгоритм, который по числу t перечисляет все Т-доказательства длины $\leq t$ и, сделав это, "ставит точку", показывая, что больше доказательств не будет).

Упражнение 6.4. Учитывая **конечность алфавита** формальной теории, убедитесь, что способ измерения длины доказательств "по числу символов" удовлетворяет условиям а), б).

Не следует думать, что способ "по числу символов", будучи самым простым, является и самым естественным из возможных способов измерения длины доказательств. При оценке "трудности" доказательства его физическая длина (после того, как выписаны все детали) не является решающей. Несколько аналогичных рассуждений, отличающихся легко контролируемыми деталями, мы обычно принимаем за **одно** рассуждение и при оценке "трудности" считаем только один раз. Очевидно, можно придумать много способов измерения длины доказательств, все более приближающихся к человеческому представлению о "трудности". Но все эти способы несмотря на их различия должны обладать свойствами а),б) (иначе их нельзя считать **полноценными** способами измерения длины). Поэтому теорема о сокращении доказательств в равной мере будет применима к любому из них.

Д о к а з а т е л ь с т в о т е о р е м ы. Будем рассуждать от противного. Пусть f(x) – вычислимая функция (монотонно стремящаяся к бесконечности), такая, что для всех формул K, доказуемых в теории T,

$$f(l_T(K)) \le l_{T+L}(K). \qquad (1)$$

Выведем отсюда, что в таком случае множество всех теорем теории T+¬L разрешимо. Это даст нужное нам противоречие, так как T+¬L – фундаментальная теория (вместе с T), условие недоказуемости L средствами T предполагает непротиворечивость T+¬L, а из раздела 6.3 мы знаем, что тогда множество всех теорем T+¬L должно быть неразрешимым.

Если некоторая формула K доказуема в теории T+¬L, то по теореме дедукции T ├ ¬L→K. Применяя (1), получаем T

$$f(l_T(\neg L \to K)) \le l_{T+L}(\neg L \to K).$$

Заметим теперь, что формула ¬L→K доказывается в теории T+L **очень легко**. В самом деле, будем исходить из того, что в исчислении высказываний доказуема формула L→(¬L→K) ("из противоречия следует все, что угодно"):

...

...

L→(¬L→K) в исчислении высказываний

L аксиома теории T+L

¬L→K по правилу MODUS PONENS

Очевидно, длина этого вывода является вычислимой функцией g(K, L) от формул K, L (свойство а), см. выше), тогда

$$l_{T+L}(\neg L \to K) \le g(K, L).$$

Отсюда

$$f(l_T(\neg L \to K)) \le g(K,L). \qquad (2)$$

Таким образом, если формула K доказуема в теории T+¬L, то $T \vdash \neg L \to K$ и имеет место (2). Но, используя методы вычисления функций f, g и зная, что f монотонно стремится к бесконечности, мы можем получить отсюда вычислимую функцию h(K, L), такую, что если K доказуема в T+¬L, то

$$l_T(\neg L \to K) \le h(K,L). \qquad (3)$$

Упражнение 6.5. Покажите, что это действительно так. Как вычисляется функция h?

Имея функцию h, можно предложить следующий метод для решения вопроса о том, доказуема ли формула K в теории T+¬L, Если $T+\neg L \vdash K$, то $T \vdash \neg L \to K$, причем длина кратчайшего доказательства удовлетворяет (3). Вычисляем h(K, L) и выписываем (по свойству б)) все T-доказательства длины ≤h(K, L). Если среди них имеется доказательство формулы ¬L→K, то $T+\neg L \vdash K$, если нет – K недоказуема в T+¬L.

Теорема о сокращении доказательств доказана.

6.5. Теорема Геделя в диофантовой форме

Как бы мы доказывали теорему Геделя о неполноте, используя то, что всякое эффективно перечислимое множество имеет диофантово представление (см. раздел 4.1)?

Пусть T – фундаментальная теория, рассмотрим для нее предикат:

"x есть номер формулы EA, перевод которой доказуем в T". (1)

Это эффективно перечислимый предикат, пусть

$$(\exists z_1 \ldots \exists z_n) P_T(x, z_1, \ldots, z_n) = 0$$

– его диофантово представление (P_T – полином с целыми коэффициентами, число n может зависеть от теории T). По лемме об автоссылках найдется замкнутая формула D_T, такая, что

$$EA \vdash D_T \leftrightarrow \neg(\exists z_1 \ldots \exists z_n) P_T(\mathbf{D_T}, z_1, \ldots, z_n) = 0.$$

Ясно, что D_T – диофантова версия формулы Геделя G_T. Как же она поведет себя?

Если $T \vdash D_T$, то номер $\mathbf{D_T}$ удовлетворяет предикату (1) и уравнение

$$P_T(\mathbf{D_T}, z_1, \ldots, z_n) = 0 \qquad (2)$$

должно иметь решения в натуральных числах. Обозначим одно из этих решений через $(b_1, \ldots, b_n)$, тогда

$$EA \vdash P_T(\mathbf{D_T}, b_1, \ldots, b_n) = 0$$

(поскольку речь идет о **числовом** равенстве, не содержащем переменных, см. упражнение 3.5). Отсюда вытекает, что

$$EA \vdash (\exists z_1 \ldots \exists z_n) P_T(\mathbf{D_T}, z_1, \ldots, z_n) = 0$$

и $T \vdash \neg D$. Таким образом, если $T \vdash D$, то теория T оказывается противоречивой.

Если же теория T непротиворечива, то формула D_T недоказуема в ней и поэтому номер $\mathbf{D_T}$ уже не удовлетворяет предикату (1). Как следствие, уравнение (2) не имеет решений в натуральных числах. Однако формула

$$\neg(\exists z_1 \ldots \exists z_n) P_T(\mathbf{D_T}, z_1, \ldots, z_n) = 0,$$

утверждающая этот факт, будучи равносильной D_T, недоказуема в теории T.

Таким образом, нами доказана

ТЕОРЕМА О НЕПОЛНОТЕ В ДИОФАНТОВОЙ ФОРМЕ. Пусть T – фундаментальная теория. Тогда существует диофантово уравнение $Q_T(z_1, \ldots, z_n) = 0$ такое, что либо теория T противоречива (тогда данное уравнение имеет решение в натуральных числах), либо уравнение не имеет решений в натуральных числах, однако формула

$$\neg(\exists z_1 \ldots \exists z_n) Q_T(z_1, \ldots, z_n) = 0,$$

утверждающая этот факт, недоказуема в T.

В частности, уравнение $Q_{EA} = 0$ таково, что, найдя его решение, мы придем к противоречию в теории EA, однако если оно неразрешимо в натуральных числах, то этот факт нельзя будет доказать средствами только EA. Таким образом, в элементарной арифметике можно решать

отнюдь не все вопросы, касающиеся разрешимости диофантовых уравнений.

Но неразрешимость уравнения $Q_{EA}=0$ удается доказать в теории множеств Цермело-Френкеля. Однако, если взять уравнение $Q_{ZFC}=0$, ситуация повторится: найдя решение этого уравнения, мы придем к противоречию в теории Цермело-Френкеля: если же решений не существует, этот факт нельзя будет доказать средствами данной теории (а так как теория ZFC формализует **все** средства рассуждения, признаваемые в математике, можно сказать даже, что неразрешимость уравнения $Q_{ZFC}=0$ нельзя доказать средствами, общепризнанными сегодня в математике).

Такое положение вещей противоречит распространенному мнению, что понятие натурального ряда в математике является первичным, не зависящим от более сложных понятий действительного числа и произвольного множества (эффектная формулировка этого мнения принадлежит Л. Кронекеру: "Бог создал целые числа, все остальное – дело рук человеческих"). Даже такие "коренные" вопросы, касающиеся натуральных чисел, как решение диофантовых уравнений, в некоторых случаях оказываются разрешимыми только с привлечением более сложных понятий. Приходится заключить, что понятие натурального ряда в истории математики **развивалось** – введение действительных чисел прибавило новые черты и натуральному ряду. Введение Г. Кантором понятия произвольного множества также прибавило новые черты понятию натурального ряда (в частности, стало возможным доказательство неразрешимости уравнения $Q_{EA}=0$). Еще более впечатляющий пример см. в приложении 2.

6.6. Теорема Леба

Формула G_T, с помощью которой К. Гедель доказывал свою теорему о неполноте, обладала свойством

$$EA \vdash G_T \leftrightarrow \neg PR_T(\mathbf{G_T}).$$

Она утверждала: "Я недоказуема в теории T", и оказалось, что если теория T непротиворечива, то G_T действительно недоказуема в T.

По аналогии с этой ситуацией в 1952 г. Л. Хенкин поставил следующий вопрос. Будем рассматривать формулу, которая (в противоположность формуле Геделя) утверждает:

"Я доказуема в теории T",

т.е. обладает свойством

$$EA \vdash A \leftrightarrow PR_T(\mathbf{A}).$$

Вопрос: формула A действительно доказуема в T?

Положительный ответ на этот вопрос был получен в 1955 г. М. Лебом.

ТЕОРЕМА ЛЕБА. Пусть T – фундаментальная теория, а $PR_T(x)$ – формула из языка EA, удовлетворяющая условиям Гильберта. Тогда для любой формулы A, $T \vdash PR_T(A) \rightarrow A$ влечет $T \vdash A$.

Д о к а з а т е л ь с т в о. Имеем:

$$T \vdash \neg A \rightarrow \neg PR_T(\mathbf{A}).$$

Таким образом,

$$T+\neg A \vdash \neg PR_T(\mathbf{A}),$$

т.е. в теории $T+\neg A$ можно доказать, что формула A недоказуема в теории T. Но если A недоказуема в T, то теория $T+\neg A$ непротиворечива (если она противоречива, то в T предположение $\neg A$ влечет противоречие, что является с точки зрения доказательством A). Таким образом, теория $T+\neg A$ доказывает свою собственную непротиворечивость. По второй теореме Геделя, это означает, что $T+\neg A$ – противоречивая теория, т.е. что формула A доказуема в теории T.

Теорема Леба доказана.

Упражнение 6.6. Изложенное доказательство содержит множество пробелов. Сначала попытайтесь самостоятельно определить их (и только потом читайте дальше). Во-первых, если к теории $T+\neg A$ применяется вторая теорема Геделя, то должна быть определена формула $Con(T+\neg A)$, выражающая непротиворечивость теории. К сожалению, стандартный способ построения формул Con в нашем случае непригоден – у нас имеется формула PR_T, но нет формулы $PR_{T+\neg A}$. Мы можем попытаться определить $Con(T+\neg A)$ как $\neg PR_T(\neg\mathbf{A}\rightarrow\mathbf{0=1})$..., но тогда придется передоказывать вторую теорему Геделя (применительно к теории $T+\neg A$). Сделайте это: предположите, что существует формула L такая, что $EA \vdash L \leftrightarrow \neg PR_T(\neg\mathbf{A}\rightarrow\mathbf{L})$, и покажите, что тогда $T \vdash Con(T+\neg A)\rightarrow L$. Отсюда будет следовать, что если $T+\neg A \vdash Con(T+\neg A)$, то $T+\neg A \vdash L$. Покажите, что последнее означает противоречие в теории $T+\neg A$ (повторив первую часть доказательства теоремы Геделя о неполноте). В конечном счете должно получиться: если $T+\neg A \vdash Con(T+\neg A)$, то теория $T+\neg A$ противоречива. Но

как доказать существование формулы L? Если формула F(x, y) представляет в EA следующую функцию f: f(**B**)=¬**A**→**B**, то по лемме об автоссылках найдется формула L такая, что

$$EA \vdash L \leftrightarrow (\exists y)(\neg PR_T(y) \wedge F(\mathbf{L},y)).$$

Поскольку единственное значение y, удовлетворяющее формуле F(**L**, y) – это ¬**A**→**L**, то получаем также

$$T \vdash L \leftrightarrow \neg PR_T(\neg\mathbf{A}\rightarrow\mathbf{L}).$$

Этим завершается доказательство второй теоремы Геделя в нужной нам форме. Другой пробел в нашем доказательстве теоремы Леба – переход от $T+\neg A \vdash \neg PR_T(\mathbf{A})$ к $T+\neg A \vdash Con(T+\neg A)$ был у нас неформальным. Для устранения этого пробела достаточно показать, что

$$T \vdash PR_T(\neg\mathbf{A}\rightarrow\mathbf{0=1})\rightarrow PR_T(\mathbf{A}).$$

Сделайте это. Какой еще (последний) пробел остается?

П р и л о ж е н и е I

ИЗ ТЕОРИИ МОДЕЛЕЙ

Ряд устойчивых платонистских заблуждений связан с другими важными результатами математической логики, которые в основном тексте книги не рассматривались: теорема Геделя о полноте, теорема Левенгейма-Сколема, теорема о категоричности аксиом Пеано. Эти результаты и их методологические последствия (или отсутствие таковых) кратко обсуждаются в данном приложении.

Все они связаны с особым подходом к изучению формальных теорий – с так называемой **теорией моделей**. В этой теории принято использовать в полном объеме средства рассуждения, характерные для теории множеств. Доказательства всех результатов, которые рассматриваются ниже, легко формализуются в теории ZFC. Можно считать, что **теория моделей – это исследование формальных теорий в метатеории** ZFC.

Имея в своем распоряжении произвольные множества, теория моделей исследует формальные теории, используя **интерпретации**. Пусть L – некоторый язык формальной теории (другой термин – язык первого порядка, см. раздел 1.5), имеющий константы c_1, ..., c_k, функциональные символы f_1, ..., f_m, предикатные символы p_1, ..., p_n.

Интерпретацией I языка L принято называть набор следующих объектов:

а) область интерпретации – некоторое непустое множество D_I (оно станет областью изменения переменных языка L),

б) отображение int_I , сопоставляющее:

– каждой константе c_i некоторый элемент множества D_I : $int_I(c_i) = \mathbf{c_i} \in D_I$; это естественно – ведь константы "призваны" обозначать конкретные объекты в области интерпретации,

– каждому функциональному символу f_i некоторую функцию из D_I в D_I , т.е. $int_I(f_1) = \mathbf{f_i}$, где $\mathbf{f_i}$: $D_i \times ... \times D_I \rightarrow D_I$ (естественно, количество аргументов $int_I(f_i)$ совпадает с количеством аргументов f_i),

– каждому предикатному символу p_i некоторое отношение на D_I, т.е. $int_I(p_i) = \mathbf{p_i} \subseteq D_i \times \ldots \times D_I$ (естественно, количество аргументов $int_I(p_i)$ совпадает с количеством аргументов p_i).

В качестве примера рассмотрим так называемую **стандартную интерпретацию** S **элементарной арифметики** EA:

а) область интерпретации $D_S = \{0, 1, 2, \ldots\}$, т.е. $D_S = \omega$ в терминах ZF,

б) отображение int_S сопоставляет: константе 0 – число 0 (пустое множество), константе 1 – число 1 (множество {0}), функциональному символу "+" – функцию x+y (сложение натуральных чисел), функциональному символу "*" – функцию x·y (умножение натуральных чисел), предикатному символу "=" – отношение x=y (равенство натуральных чисел).

При заданной интерпретации I (некоторого языка L) определяется **понятие истинности формул** языка L. Определение начинается с интерпретации термов в виде функций из D_I в D_I. Каждый терм языка является либо константой, либо переменной, либо комбинацией, использующей функциональные символы. В первых двух случаях интерпретацией терма становится либо постоянная функция $c(x) = c_i$, либо тождественная функция e(x) = x, а в последнем случае, если $t=f(t_1, \ldots, t_n)$, то $int_I(t)$ – функция, получаемая путем подстановки функций $int_I(t_1)$, ..., $int_I(t_n)$ в функцию $int_I(f)$. Например, интерпретацией терма (x+y)*(x+y) является функция $(x+y)^2$.

Далее естественным образом определяется истинность элементарных формул (при заданных значениях свободных переменных из области D_I): истинность формулы $p_i(t_1, \ldots, t_n)$ устанавливается путем "вычисления" значений термов $t_1, \ldots, t_n$ и подстановкой этих значений в отношение $int_I(p_i)$. Замечание о свободных переменных здесь существенно – ведь истинность, например, формулы x=1 зависит от конкретного значения x.

Наконец, можно "определить" понятие истинности для произвольных формул языка L при данной интерпретации I (также, естественно, при заданных значениях свободных переменных формулы):

– вопрос об истинности формул ¬A, A∧B, A∨B, A→B тривиальным образом сводится к вопросу об истинности формул A, B,

– формула (∀x)A(x) истинна, если формула A(x) истинна при

любом значении переменной x из области D_I,

– формула $(\exists x)A(x)$ истинна, если формула A(x) истинна при хотя бы одном значении переменной x из области D_I.

Необходимо сознавать, что в случае бесконечной области D_I понятие истинности в такой интерпретации оказывается **неконструктивным**: например, проверка истинности формулы $(\forall x)A(x)$ требует проверки истинности A(x) для бесконечного числа конкретных значений x (ср. рассуждение в разделе 3.1 об истинности формул элементарной арифметики). Еще более неконструктивно понятие истинности формул вида $(\forall x\exists y)B(x, y)$, $(\forall x\exists y\forall z)C(x, y, z)$ и т.д.

Формула языка L называется **тождественно истинной** при данной интерпретации I, если она истинна при любых (взятых из D_I) значениях своих свободных переменных.

Ряд формул тождественно истинны при любых интерпретациях, например

$$A\rightarrow A,$$
$$(A\rightarrow B)\rightarrow((B\rightarrow C)\rightarrow(A\rightarrow C)),$$
$$(\forall x)(C\rightarrow D(x))\rightarrow(C\rightarrow(\forall x)D(x)),$$

где формула C не содержит переменной x. Такие формулы, тождественно истинные при любой интерпретации (т.е. благодаря своей форме), принято называть **логически общезначимыми**. Следует отметить **двойную** неконструктивность этого понятия – на уже неконструктивное понятие истинности здесь накладывается квантор "для всех интерпретаций".

Легко проверить, что логические аксиомы и правила вывода, сформулированные в разделе 1.5, позволяют доказывать только логически общезначимые формулы. Но можно поставить обратный вопрос – всякую ли логически общезначимую формулу можно доказать с помощью этих аксиом? Получить ответ не так просто, он был дан в 1929 г. К. Геделем:

ТЕОРЕМА ГЕДЕЛЯ О ПОЛНОТЕ. В любом языке первого порядка формула является логически общезначимой, если и только если ее можно доказать с помощью логических аксиом и правил вывода из раздела 1.5.

Доказательство этой теоремы довольно сложно. И, разумеется, оно неконструктивно – ведь неконструктивно "в квадрате" само понятие логической общезначимости. В современных учебниках теорема о полноте обычно выводится из следующей теоремы о существовании

модели.

Пусть T – формальная теория. Интерпретация I (языка T) называется **моделью** теории T, если при этой интерпретации тождественно истинны все аксиомы (и все теоремы) теории. Такое понятие модели несколько необычно: в нормальной ситуации сама теория служит моделью явления природы, технического устройства и т.п., а здесь – наоборот! Это, однако, общепринятый жаргон математической логики.

ТЕОРЕМА О СУЩЕСТВОВАНИИ МОДЕЛИ. Если формальная теория непротиворечива, то существует конечная или счетная модель этой теории.

Конечная или счетная модель – интерпретация с конечной или счетной областью D_I. Смысл теоремы можно при желании объяснять так: если в теории нет противоречий, то ее "содержание" непусто – она описывает какую-то "математическую реальность".

Элегантное доказательство теоремы о существовании модели, принадлежащее Л. Генкину и Г. Хазенъегеру, см. в книге Э. Мендельсона [1976].

Если теорема о существовании модели доказана, то теорема Геделя о полноте оказывается простым ее следствием. В самом деле, то, что все формулы некоторого языка первого порядка L, выводимые с помощью логических аксиом и правил вывода, являются логически общезначимыми, можно легко проверить, рассуждая по индукции. Теперь допустим, что какую-либо логически общезначимую формулу F из языка L нельзя доказать с помощью логических аксиом и правил вывода. Это значит, что теория T (на языке L), единственной собственной аксиомой которой является формула ¬F, будет непротиворечивой (будь она противоречивой – это означало бы доказуемость "от противного" формулы F чисто логическими средствами). Таким образом, существует модель теории T, т.е. интерпретация, в которой тождественно истинны все аксиомы T, в том числе формула ¬F. Однако в этой интерпретации должна быть тождественно истинной и (логически общезначимая) формула F! Формулы F и ¬F не могут быть истинными в одной и той же модели, т.е. наше предположение неверно и формулу F можно доказать с помощью логических аксиом и правил вывода, что и требовалось.

Методологическое значение теоремы Геделя о полноте состоит в том, что с ее помощью первоначально неконструктивное понятие логически общезначимой формулы превращается в значительно более конструктивное понятие формулы, доказуемой с помощью логических аксиом и правил вывода.

Примечание 1. Полная конструктивность здесь, к сожалению, недостижима. Как показал в 1936 г. А.Черч, если рассматривать, например, язык EA, то не существует алгоритма, который распознает, является ли данная формула языка логически общезначимой (теорема Черча о неразрешимости исчислений предикатов, см. Э. Мендельсон [1976]).

Примечание 2. Теорема о существовании модели сначала была доказана в более слабой форме: если формальная теория имеет какую-либо модель, то она имеет и конечную или счетную модель (**теорема Левенгейма-Сколема**, доказанная Л. Левенгеймом в 1915 г. и Т. Сколемом в 1920 г.).

С теоремой Левенгейма-Сколема связан **парадокс Сколема**. Если предположить, что теория ZFC имеет какую-либо модель, то согласно этой теореме она имеет и модель, состоящую из **счетного** числа элементов (модель ZFC не может быть конечной из-за аксиомы бесконечности, которая должна выполняться в модели). Но ведь в теории ZFC можно доказать существование **несчетных** множеств, как же она может иметь счетную модель?

Платонистски настроенные мыслители делают отсюда вывод, что теория ZFC не является "полноценной" формализацией теории множеств. Более того, они считают, что "подлинную" теорию множеств вообще невозможно представить в виде формальной теории (ведь для любой такой теории – в случае ее непротиворечивости – найдется счетная модель).

На самом деле парадокс Сколема имеет очень простое объяснение, не оставляющее места для каких-либо спекуляций. Пусть I – интерпретация теории ZFC со счетной областью D_I. В ZFC можно доказать, что, например, множество всех действительных чисел R несчетно, т.е.

$$\neg(\exists f) f : R \overset{1-1}{\to} N \quad , \qquad (1)$$

где N – множество всех натуральных чисел, f – одно-однозначная функция. Разумеется, если взять интерпретации N и R:

$$int_I(N)=n \in D_I, \quad int_I(R)=r \in D_I,$$

то n, r $\subseteq D_I$ и множества n, r оба оказываются счетными, т.е.

$$(\exists f) f : r \overset{1-1}{\to} n \quad .$$

Однако интерпретацией теоремы (1) является утверждение

$$\neg(\exists f \in D_I) f : r \overset{1-1}{\to} n \quad ,$$

поэтому противоречие здесь не возникает: одно-однозначные функции типа f: r→n существуют, но ни одна из них не содержится в модели! "Наблюдателю со стороны" все множества модели кажутся счетными, однако "наблюдатель", находящийся "внутри модели", некоторые функции "не видит", поэтому некоторые счетные множества ему представляются несчетными.

С аксиомами Дж. Пеано (см. начало раздела 3.1) связан еще один платонистский предрассудок. Для этих аксиом можно доказать **теорему о категоричности**, которая утверждает, что "всякие две модели, в которых выполняются аксиомы Пеано, изоморфны". Понятие модели арифметики Пеано легко формулируется в теории ZF: это тройка (v, o, s), где v – множество (его элементы считаются "натуральными числами" модели), o – элемент v ("нуль"), s – функция типа v→v (s(x) играет роль функции x+1). Будем говорить, что в модели (v, o, s) выполняются аксиомы Пеано, если

P1: ¬(s(x)=o) для всех x ∈v,

P2: s(x)=s(y) → x=y для всех x, y ∈v,

P3: если u⊆v, то o ∈u ∧ (∀y)(y ∈u → s(y) ∈u) → u=v.

Естественной моделью арифметики Пеано оказывается множество ω всех натуральных чисел по Дж. фон Нейману (см. раздел 2.3), где нулем является пустое множество 0, а роль s(x) играет x∪{x}. Эту модель принято называть **стандартной моделью**. Наконец, будем говорить, что модель (v, o, s) изоморфна стандартной модели, если существует отображение f: ω→v такое, что

а) f(0)=o,

б) n ∈ω → f(n+1)=s(f(n)) (если x соответствует n, т.е. x=f(n), то s(x) соответствует n+1, т.е. s(x)=f(n+1)),

в) область значений rng(f)=v,

г) f – одно-однозначная функция. Теперь мы можем сформулировать упомянутую теорему.

ТЕОРЕМА О КАТЕГОРИЧНОСТИ. Всякая модель аксиом Пеано изоморфна стандартной модели.

Д о к а з а т е л ь с т в о. Итак, пусть в модели (v, o, s) выполняются аксиомы Пеано. Определим по индукции следующее отображение f типа ω→v:

f(0)=o, f(1)=s(o), f(2)=s(s(o)),..., f(k+1)=s(f(k)),...

Покажем, что f является изоморфизмом:

а) f(0)=o по определению,

б) f(k+1)=s(f(k)) по определению,

в) область значений rng(f) содержится в v, ибо o ∈rng(f), и если y ∈ rng(f), то y=f(k) для некоторого k, и поэтому s(y)=s(f(k))=f(k+1) ∈rng(f). Так как в модели (v,o,s) выполняется аксиома P3, то отсюда вытекает, что rng(f)=v,

г) покажем, что f(m)=f(n) → m=n, т.е. что функция f одно-одно-значна. Итак, пусть f(m)=f(n). Возможны три случая:

г1) m=n=0 (все ясно),

г2) m=0, n>0. Тогда f(m)=o, однако f(n)=s(f(n–1))≠o из-за аксиомы P1, которая выполняется в модели (v,o,s),

г3) m, n>0. Тогда f(m)=s(f(m–1))=f(n)=s(f(n–1)), и по аксиоме P2 получается, что f(m–1)=f(n–1). Продолжая в таком духе, приходим либо к случаю г1), либо к г2).

Теорема доказана.

Итак, получается, что аксиомы Пеано "однозначно определяют" структуру своих моделей. Поэтому теорема о категоричности нередко считается дополнительным аргументом в пользу мнения, что каждое определенное утверждение о свойствах натуральных чисел должно быть либо истинным, либо ложным, т.е. что натуральные числа существуют как "реальность".

Из теоремы Геделя о неполноте (см. раздел 5.3) вытекает, что система аксиом EA (или любая другая система аксиом на языке EA) этим свойством категоричности обладать не может. В самом деле, формула Геделя для теории EA имеет вид ¬(∃x)C(x), причем если теория EA непротиворечива, то для каждого конкретного натурального числа n: EA ├ ¬C(**n**), однако в EA невозможно доказать ¬(∃x)C(x). Это означает, что непротиворечивой является теория EA' = EA+{(∃x)C(x)}. В моделях теории EA' для каждого "стандартного" натурального числа n выполняется ¬C(**n**) (как теорема EA, т.е. и EA'), однако в этих моделях существует объект q, для которого выполняется C(q). Такие объекты принято называть "нестандартными" натуральными числами, а модели, в которых они существуют, – **нестандартными моделями арифметики**.

И формулировку, и доказательство теоремы о категоричности можно осуществить в теории ZF, т.е. ничего "сверхъестественного" в этой теореме быть не может. Ну, а причиной психологического заблуждения, связанного с ней, является закон исключенного третьего (одна из аксиом ZF: A∨¬A для любой формулы A, поэтому каждая формула в данной модели либо истинна, либо ложна).

П р и л о ж е н и е 2

ВОКРУГ ТЕОРЕМЫ РАМСЕЯ

Отношение многих "практикующих" математиков, занятых решением конкретных математических проблем (или даже прикладных вопросов), к теоремам о неполноте можно выразить следующими словами: "Из теорем Геделя **не** вытекает неразрешимость проблемы, которой я сейчас занимаюсь, поэтому оставьте меня в покое!" Некоторый "методологический базис" под такое отношение подводит Р. Парих [1971]:

"Таким образом, операция возведения в степень является не только средством для обозначения "больших чисел", но и средством введения "нематематических" вопросов в теорию чисел.

Почему мы говорим "нематематических"? Рассмотрим формулу Геделя А, которая утверждает: "Я недоказуема". Эта формула выражает свойства N (системы натуральных чисел. – К. П.), поскольку ее можно написать с помощью кванторов и логических связок (т.е. на языке ЕА. – К. П.). Однако чтобы увидеть, что А истинна, но недоказуема, мы используем не свойства N, а свойства интуитивного понятия "доказуемость". Таким образом, разговоры о том, что А является утверждением о числах, похожи на аргументы вроде: поведение человека является проблемой физики, поскольку человеческие существа являются физическими телами. Даже если такое предположение верно, оно оказывается очень теоретическим и малополезным."

Это глубокие соображения и, по-видимому, из них следуют далеко идущие выводы о природе формализации, однако они не могут служить оправданием пренебрежительного отношения "практикующих" математиков к теоремам о неполноте.

Общие теоремы о неполноте доказывают неибежное несовершенство всякой застывшей системы понятий (моделями таких систем являются формальные теории). В ходе развития любой математической теории (это наиболее развитые из застывших систем) неизбежно должны появиться или противоречия, или проблемы, которые в данной теории можно сформулировать, но невозможно решить. Что к этому предсказанию следует относиться вполне серъезно, показало дальнейшее развитие методов математической логики. Сначала, в 1963 г. П. Коэн доказал неразрешимость проблемы континуума в теории

множеств (см. раздел 2.4). С тех пор доказана неразрешимость уже целого ряда классических проблем теории множеств. Поэтому не исключено, что в будущем удастся доказать неразрешимость и некоторых классических проблем теории чисел, например проблемы простых чисел-близнецов.

Первым серьезным шагом в этом направлении является обнаруженная в 1977 г. недоказуемость в теории ЕА так называемой усиленной конечной теоремы Рамсея (если преположить, что ЕА непротиворечива), которую нетрудно доказать в теории множеств (см. дальше). Это первый пример содержательно интересного утверждения о свойствах системы натуральных чисел, для доказательства которого недостаточно "ограниченного" понятия об этих числах, представленного в теории ЕА.

Бесконечная теорема Рамсея

Теорема, доказанная Ф. Рамсеем (1903–1930), относится к области комбинаторной математики.

Для конечного или счетного множества M через |M| будем обозначать количество элементов в M.

Рассмотрим следующую проблему. Пусть M – конечное или счетное множество элементов, которые будем называть "игроками". Если e – натуральное число, то подмножества M, содержащие ровно e элементов, будем называть "e-командами". Пусть задано некоторое разбиение e-команд из M на r непересекающихся классов (например, по "классу игры"). Ф.Рамсей заметил, что если множество M взять достаточно большим, то при любом разбиении e-команд из M на r классов найдется достаточно большое подмножество $H \subseteq M$, такое, что все e-команды из H попадают в один класс разбиения. Такое подмножество H принято называть **однородным** для данного разбиения.

Особенно наглядной проблема становится при e=2. Тогда элементы множества M можно представлять вершинами полного графа, а 2-команды – ребрами графа, окрашивая их в r цветов, в зависимости от класса разбиения, куда каждая команда попадает. Однородным множеством в этом случае является полный подграф, состоящий из ребер одного цвета. Согласно теореме Рамсея в достаточно большом "цветном" полном графе найдется и достаточно большой **одноцветный** полный подграф.

Правда, сам Ф.Рамсей доказал свою теорему для бесконечных множеств.

БЕСКОНЕЧНАЯ ТЕОРЕМА РАМСЕЯ. Пусть e, r>0 – натуральные числа. Если M – счетное множество, то для любого разбиения e-команд из M на r классов найдется бесконечное подмножество H⊆M, все e-команды которого попадают в один класс разбиения.

Д о к а з а т е л ь с т в о (проводится в теории ZFC, см. Р. Грэхем [1984]). Воспользуемся индукцией по e.

1) При e=1 утверждение теоремы очевидно (если M – бесконечное множество, а число классов разбиения конечно, то в один из классов попадает бесконечное число элементов M, это и есть требуемое однородное подмножество H).

2) Идею дальнейшего доказательства проще всего объяснить для e=2. В этом случае (см. выше) ситуацию представляет "цветной" полный граф. Требуется доказать, что в бесконечном полном графе M, ребра которого раскрашены в r цветов, найдется бесконечный полный подграф с ребрами одного цвета.

Определим следующую последовательность вершин графа, цветов и подграфов. Сначала выбираем произвольную вершину $a_0 \in M$. Из нее исходит бесконечное число ребер к остальным вершинам графа. Среди них найдется бесконечно много ребер одного цвета, обозначим этот цвет через c_0, а (бесконечное) множество конечных вершин этих ребер – через M_1 . Таким образом, $M_1 \subseteq M_0 = M$, и вершина a_0 с вершинами из M_1 связана только ребрами цвета c_0. Затем выбираем произвольную вершину $a_1 \in M_1$. Из нее в сторону остальных вершин M_1 исходит бесконечное число ребер, среди них – бесконечное число ребер одного цвета, который обозначим через c_1. Соответствующее (бесконечное) множество конечных вершин этих ребер обозначим через M_2 . Таким образом, $M_2 \subseteq M_1 \subseteq M_0$, и вершина a_1 с вершинами из M_2 связана только ребрами цвета c_1 . Затем выбираем произвольную вершину $a_2 \in M$, и т.д.

В результате этого процесса получаем три бесконечные последовательности:

– вершин: $a_0, a_1, ..., a_n,...$,

– цветов: $c_0, c_1, ..., c_n, ...$,

– бесконечных подграфов: $M = M_0 \supseteq M_1 \supseteq ... \supseteq M_m \supseteq ...$ Здесь всегда $a_i \in M_i - M_{i+1}$, и из a_i в сторону вершин из M_{i+1} ведут только ребра

цвета c_i.

Один из r цветов встречается в последовательности бесконечное число раз, обозначим его через c:

$$c = c_{i0} = c_{i1} = c_{i2} =...$$

Множество соответствующих вершин обозначим через H:

$$H = \{a_{i0}, a_{i1}, a_{i2}, ... \}.$$

Оно бесконечно, и все его вершины связывают ребра цвета c, т.е. мы получили требуемый бесконечный "одноцветный" подграф.

Этим завершается доказательство бесконечной теоремы Рамсея для случая e=2.

3) Докажем теперь общий случай шага индукции: переход от e–1 к e. Мы сможем полностью повторить рассуждения п.2, если будет доказана следующая

ЛЕММА. Пусть даны счетное множество M и разбиение P e-команд из M на r классов. Если бесконечная теорема Рамсея верна для e–1, то для всякого бесконечного подмножества $M' \subseteq M$ и всякого элемента $a' \in M'$ найдется бесконечное подмножество $H' \subseteq M'–\{a'\}$, такое, что все e-команды, состоящие из a' и элементов H', попадают в один класс разбиения P.

Докажем эту лемму. Исходя из разбиения P построим следующее разбиение P' (e–1)-команд из $M' –\{a'\}$ на r классов:

$$\{x_1, ..., x_{e–1}\} \in P_i' \leftrightarrow \{a', x_1, ..., x_{e–1}\} \in P_i.$$

Через P_i, P_i' здесь обозначены i-е классы обоих разбиений, все $x_j \in M'–\{a'\}$. Применим бесконечную теорему Рамсея к (e–1)-командам из $M'–\{a'\}$, получив бесконечное подмножество $H' \subseteq M'–\{a'\}$, такое, что все (e–1)-команды из H' попадают в один класс разбиения P'. Переходя от P' к исходному разбиению P (т.е. добавляя элемент a' к каждой (e–1)-команде из H'), получаем утверждение леммы.

Имея такую лемму, мы можем повторить рассуждение п.2. Лемма нужна для получения бесконечного множества $M_{i+1} \subseteq M_i –\{a_i\}$, такого, что все e-команды из $M_{i+1} \cup\{a_i\}$, содержащие a_i , попадают в один класс разбиения P (этот класс можно обозначить через a).

Тем самым индукцией по e бесконечная теорема Рамсея доказана полностью.

Формулировка и доказательство бесконечной теоремы Рамсея

относятся, разумеется, к теории множеств. Эта теорема оказывается, таким образом, теоремой теории ZFC. Нет никакой возможности даже сформулировать ее средствами теории EA, которые не позволяют обсуждать произвольные бесконечные множества.

Конечная теорема Рамсея

В теории EA можно, однако, сформулировать и доказать аналог бесконечной теоремы Рамсея для конечных множеств:

КОНЕЧНАЯ ТЕОРЕМА РАМСЕЯ. Существует вычислимая функция R(e, r, k), такая, что при любых e, r, k > 0 для любого конечного множества M: если $|M| \geq R(e, r, k)$, то для любого разбиения P e-команд из M на r классов найдется подмножество $H \subseteq M$ такое, что $|H| \geq k$ и все e-команды из H попадают в один класс разбиения P.

И здесь наиболее наглядной оказывается ситуация при e=2: существует функция R(r,k) такая, что любой полный граф, имеющий R(r,k) вершин и ребра, раскрашенные в r цветов, содержат полный подграф из k вершин, все ребра которого раскрашены в один цвет.

Д о к а з а т е л ь с т в о (см. Р. Грэхем [1984], его можно формализовать в теории EA).

1) При r=1 утверждение теоремы очевидно: можно взять R(1, k) = k.

2) Рассмотрим теперь случай r=2, когда все e-команды из M разбиваются на два класса. Оказывается, легче доказать следующее обобщение теоремы: существует вычислимая функция $R'(e, 2, k_1, k_2)$, такая, что при $|M| \geq R'(e, 2, k_1, k_2)$ для дюбого разбиения e-команд из M на два класса найдется либо подмножество H_1 такое, что $|H_1| = k_1$ и все e-команды из H_1 попадают в первый класс разбиения, либо подмножество H_2 такое, что $|H_2| = k_2$ и все e-команды из H_2 попадают во второй класс разбиения.

Доказываем по индукции, переходя от случая e–1 и случаев $(e, k_1–1, k_2)$, $(e, k_1, k_2–1)$ к случаю (e, k_1, k_2).

Б а з и с и н д у к ц и и. При e=1 можно взять

$$R'(1, 2, k_1, k_2) = k_1 + k_2$$

(в этом случае сами элементы M разбиваются на два класса). Далее, рассматривая случай произвольного e и наименьшего k_1, имеем: при k_1=e

можно взять $R'(e, 2, e, k_2) = k_2$ ($k_2 \geq e$), поскольку если существует хотя бы одна e-команда первого класса, то она одна и составляет требуемое H_1, а если все e-команды из M попадают во второй класс, то можно взять $H_2 = M$. Аналогично при $k_2 = e$ можно взять $R'(e, 2, k_1, e) = k_1$ ($k_1 \geq e$).

Ш а г и н д у к ц и и. Пусть $k_1, k_2 > e$. Покажем, что можно взять:

$$R'(e, 2, k_1, k_2) = 1 + R'(e-1, 2, R'(e, 2, k_1-1, k_2), R'(e, 2, k_1, k_2-1)).$$

В самом деле, если $|M| \geq R'(e, 2, k_1, k_2)$, то выбираем один элемент $a \in M$ и рассматриваем все e-команды из M, содержащие a: $\{a, x_1, ..., x_{e-1}\}$. Каждая из них относится к первому или второму классу разбиения P. Определим для (e–1)-команд из $M-\{a\}$ следующее разбиение P' на два класса (i=1, 2):

$$\{x_1, ..., x_{e-1}\} \in P' \leftrightarrow \{a, x_1, ..., x_{e-1}\} \in P .$$

Поскольку

$$|M-\{a\}| \geq R'(e-1, 2, T_1, T_2),$$

где $T_1 = R'(e, 2, k_1-1, k_2)$ и $T_2 = R'(e, 2, k_1, k_2-1)$, то, по теореме Рамсея для случая e–1, найдется подмножество $M' \subseteq M-\{a\}$ такое, что либо $|M'| = T_1$ и все (e–1)-команды из M' попадают в первый класс, либо $|M'| = T_2$ и все (e–1)-команды из M' попадают во второй класс. В первом случае

$$(\forall x_1, ..., x_{e-1} \in M')\ \{a, x_1, ..., x_{e-1}\} \in P_1$$

(P_1 – первый класс разбиения P). Но так как $|M'| = T_1 = R'(e, 2, k_1-1, k_2)$, то по предположению индукции (случай (e, k_1-1, k_2)) найдется подмножество $H \subseteq M'$ такое, что либо $|H| = k_1-1$ и все e-команды из H попадают в первый класс разбиения P (тогда $H \cup \{a\}$ будет искомым подмножеством для случая (e, k_1, k_2)), либо $|H| = k_2$ и все e-команды из H попадают во второй класс разбиения P (тогда это же H подходит и для случая (e, k_1, k_2)).

Второй случай разбирается аналогично.

Для случая r=2 конечная теорема Рамсея доказана.

3) Перейдем теперь к случаю произвольного $r \geq 2$ и проведем индукцию по r.

Б а з и с и н д у к ц и и. При r=2 возьмем

$$R(e, 2, k) = R'(e, 2, k, k).$$

Ш а г и н д у к ц и и. Покажем, что можно взять

$$R(e, r, k) = R(e, 2, R(e, r–1, k)).$$

В самом деле, пусть $|M| \geq R(e, r, k)$ и задано некоторое разбиение P e-команд из M на r классов. Чтобы привести ситуацию к случаю двух классов, зафиксируем один из классов разбиения P (будем называть его первым классом), а остальные объединим (будем называть это вторым классом). Тогда согласно доказанному в случае r=2 найдется подмножество $M' \subseteq M$ такое, что $|M'|=R(e, r–1, k)$ и либо а) все e-команды из M' попадают в первый класс, либо б) все e-команды из M' попадают во второй класс.

В случае а), поскольку $R(e, r–1, k) \geq k$, мы сразу получаем $H \subseteq M'$ такое, что $|H|=k$ и все e-команды H попадают в один класс разбиения P. В случае б) получаем разбиение e-команд из M' на r–1 класс разбиения P. Поскольку $|M'|= R(e, r–1, k)$, то согласно предположению индукции (случай r–1) найдется подмножество $H \subseteq M'$ такое, что $|H|=k$ и все e-команды из H попадают в один класс разбиения P.

Конечная теорема Рамсея доказана полностью.

Хотя в конечной теореме Рамсея идет речь не о натуральных числах, а о произвольных конечных множествах, эту теорему можно сформулировать в теории EA. Достаточно принять какое-либо кодирование конечных множеств натуральными числами, например с помощью бета-функции Геделя (см. раздел 3.3): кодом множества $\{a_1, ..., a_n\}$ будем считать пару чисел a, b такую, что

$$\beta(a, b, 0)=n, \beta eta(a, b, 1)=a_1, ..., \beta(a, b, n)=a_n.$$

Таким образом, конечную теорему Рамсея можно не только сформулировать, но и доказать в теории EA (в этом нет ничего удивительного: ведь все рассуждения в приведенном выше доказательстве элементарны).

Усиленная конечная теорема Рамсея

Мы имеем, таким образом, две теоремы Рамсея:

а) "бесконечную", которую нельзя ни сформулировать, ни (тем более) доказать в элементарной арифметике (теории EA); для этого нужна теория множеств (теория ZFC),

б) "конечную", которую можно сформулировать и доказать в элементарной арифметике.

В 1977 г. Л. Харрингтон, опираясь на новый метод, изобретенный

Л. Кэрби и Дж. Парисом, обнаружил промежуточный вариант – **усиленную конечную теорему Рамсея** (УКТР), которую **можно сформулировать**, **но нельзя доказать** в элементарной арифметике (т.е. в теории ЕА, разумеется, если она непротиворечива). Это первый случай, когда удалось показать, что для доказательства некоторого содержательно интересного утверждения о свойствах натуральных чисел нужны средства, выходящие за рамки теории ЕА.

Чтобы подойти к формулировке усиленной конечной теоремы Рамсея, мы должны рассматривать в конечной теореме Рамсея не произвольные множества М, а множества, которые берутся из определенного счетного "универсума" с отношением порядка, например из множества всех натуральных чисел. Тогда появится возможность сравнивать конечные множества не только по количеству элементов, но и по другим признакам. Например, можно рассматривать "разреженные" множества $H = \{a_1, a_2, \dots, a_n\}$, выделямые условием

$$2^{a1} \le a_1,\ 2^{a2} \le a_2,\ \dots,\ 2^{an} \le a_n,$$

или, наоборот, "плотные" множества:

$$a_1 < a_2 < \dots < a_n \wedge a_1 \le n$$

(или $\min(H) \le |H|$). (Заметим, что в данном случае множество Н может быть "плотным" и "разреженным" одновременно.)

В формулировке УКТР может использоваться широкий класс свойств "плотности". Как мы увидим из доказательства, в качестве свойства плотности в УКТР можно взять любое свойство γ конечных множеств натуральных чисел, удовлетворяющее двум условиям:

а) если $H_1 \subseteq H_2$ и $\gamma(H_1)$, то $\gamma(H_2)$,

б) если H_2 – бесконечное множество натуральных чисел, то найдется конечное подмножество $H_1 \subseteq H_2$, такое, что $\gamma(H_1)$.

Подобное свойство естественно называть плотностью (по отношению к системе всех натуральных чисел).

Свойство $\min(H) \le |H|$ является простейшим из свойств, удовлетворяющих этим условиям:

– если $H_1 \subseteq H_2$ и $\min(H_1) \le |H_1|$, то $\min(H_2) \le |H_2|$,

– если H_2 – бесконечное множество и H_1 – множество наименьших $\min(H_2)$ элементов H_2, то $\min(H_1) \le |H_1|$.

Это пример **разрешимого** свойства "плотности" (по заданному

конечному множеству H можно "вычислить", плотное оно или нет).

УСИЛЕННАЯ КОНЕЧНАЯ ТЕОРЕМА РАМСЕЯ. Для каждого разрешимого свойства плотности γ существует вычислимая функция $R_\gamma(e, r, k)$, такая, что при любых e, r, k>0 для любого конечного множества M: если $|M| \geq R_\gamma(e, r, k)$, то для любого разбиения P e-команд из M на r классов найдется γ-**плотное** подмножество $H \subseteq M$, такое, что $|H| \geq k$ и все e-команды из H попадают в один класс разбиения P.

Эта теорема отличается от обычной конечной теоремы Рамсея по существу "только" одним словом – дополнительно требуется, чтобы множество H было **плотным**.

Д о к а з а т е л ь с т в о (проводится в теории ZFC, см. Дж. Парис, Л. Харрингтон [1977]). Очевидно, можно ограничиться рассмотрением в роли M множеств вида {0,1,...,n}, поэтому мы будем считать, что M={0,1,...,M–1} (как в разделе 2.3). Для данных e, r рассматриваются все $M \geq e$.

Для M=e возможна только одна e-команда {0, 1, ..., e–1} и r ее разбиений на r классов.

Рассмотрим теперь конкретную пару (M,P), где P – разбиение e-команд из M на r классов. Если перейти от M к M+1, т.е.

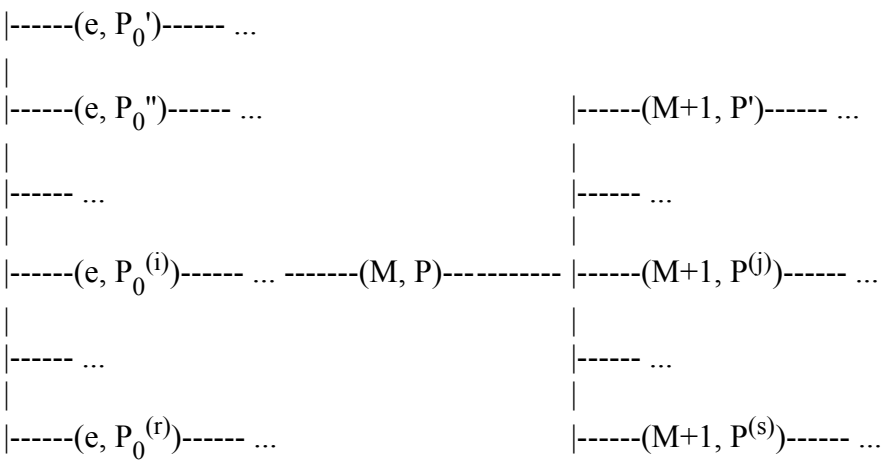
$$M+1 = \{0, 1, ..., M–1, M\} = M \cup \{M\},$$

то из разбиения P можно получить только конечное число различных разбиений e-команд из M+1 на r классов, которые на M согласуются с P. Разбиение P' согласуется с P – это означает, что для любой e-команды $\{x_1, ..., x_e\}$ из M и любого i ($1 \leq i \leq r$): 1 e

$$\{x_1, ..., x_e\} \in P_i' \leftrightarrow \{x_1, ..., x_e\} \in P_i.$$

Разбиение P' отличается от P, таким образом, только тем как оно разбивает те e-команды, которые содержат число M.

Если, начиная с M=e, повторно применять эту процедуру, то в результате мы получим бесконечное дерево с конечными ветвлениями:



Здесь через $P_0^{(i)}$ обозначены все разбиения e–команд из e на r классов, а через $P^{(j)}$ – всевозможные разбиения e–команд из M+1 на r классов, согласованные с разбиением P. Легко проверить, что вершины этого дерева охватывают все возможные пары (M, P) (при фиксированных e, r), одновременно упорядочивая их определенным образом.

Будем называть вершину (M, P) "хорошей", если существует подмножество $H \subseteq M$, обладающее свойством γ, такое что $|H| \geq k$ и все e–команды из H попадают в один класс разбиения P. Легко проверить, что если (M, P) – "хорошая" вершина, то все $(M+1, P^{(j)})$ – также "хорошие".

Из–за конечности ветвлений в вершинах каждый уровень дерева содержит только конечное число вершин, поэтому УКТР будет доказана, если мы сумеем показать, что "нехороших" вершин существует только конечное число.

В самом деле, в таком случае мы могли бы предложить следующий алгоритм вычисления значения функции $R_\gamma(e, r, k)$. Будем проходить все уровни дерева подряд, проверяя каждую вершину (M, P), "хорошая" она или нет (путем полного перебора всех 2^M подмножеств M). Если на каком–то уровне все вершины оказались "хорошими", то номер M этого уровня можно взять в качестве значения $R_\gamma(e, r, k)$. Нужна только гарантия, что такой уровень всегда найдется.

Итак, предположим от противного, что в нашем дереве бесконечно много "нехороших" вершин. Если $(M+1, P^{(j)})$ – "нехорошая" вершина, то (M, P) – также "нехорошая", поэтому в дереве найдется бесконечная ветвь, состоящая только из "нехороших" вершин.

Упражнение 1. Докажите строго, что это действительно так (это

лемма Кенига: если дерево с конечными ветвлениями содержит бесконечно много вершин, то в нем имеется бесконечная ветвь).

Бесконечная ветвь, состоящая из согласованных разбиений (M, P_M), задает одно определенное разбиение P' всех e–команд, состоящих из натуральных чисел, на r классов: класс e–команды $\{i_1, i_2, ..., i_e\}$ определяет разбиение P_M , где

$$M = \max\{i_1, i_2, ..., i_e\},$$

а последующие разбиения P_{M+1}, P_{M+2}, ... не могут это определение изменить, поскольку они согласованы с P_M.

Применив теперь к разбиению P' (всех e–команд из натуральных чисел) бесконечную теорему Рамсея, получим бесконечное множество H', все e–команды которого попадают в один класс разбиения P'. "Плотность" свойства γ позволяет получить конечное подмножество H⊆H', такое, что |H|≥k и γ(H) (сначала находим просто конечное подмножество, обладающее свойством γ, затем, если это необходимо, пополняем его до k элементов из H'). Если взять M=max(H), то все e–команды из H окажутся и в одном классе разбиения P_M. Это противоречит тому, что вершина (M, P_M) – "нехорошая".

Усиленная конечная теорема Рамсея доказана полностью.

УКТР недоказуема в элементарной арифметике

Легко показать, что для разрешимого свойства γ (как min(H)≤|H|) УКТР можно сформулировать в теории EA. В 1977 г. Л. Харрингтон, опираясь на новый метод, изобретенный Л. Кэрби и Дж. Парисом, показал, что для свойства плотности min(H)≤|H| эту теорему **нельзя доказать** в теории EA (разумеется, если она непротиворечива, см. Дж. Парис, Л. Харрингтон [1977]). Это первый случай, когда удалось строго доказать, что для доказательства некоторого содержательно интересного утверждения о свойствах натуральных чисел нужны средства, выходящие за рамки теории EA (мы только что доказали УКТР, используя бесконечную теорему Рамсея, т.е. в теории ZFC). Можно утверждать поэтому, что древние греки, имеющие в своем распоряжении только то изолированное понятие натурального числа, которое формализовано в EA, не могли бы доказать УКТР. Такая возможность появилась только после изобретения теории множеств в XIX в.: введение Г. Кантором понятия произвольного бесконечного множества прибавило, таким образом, новые черты понятию натурального числа.

Если обратиться еще раз к приведенному выше доказательству УКТР, то можно заметить, что сама процедура вычисления функции $R_\gamma(e, r, k)$, разумеется, в средствах теории множеств не нуждается. Эти средства понадобились, чтобы доказать **сходимость алгоритма**: для доказательства, что в дереве разбиений существует уровень, состоящий только из "хороших" вершин (номер первого такого уровня и становится значением $R_\gamma(e, r, k)$).

Однако для проверки истинности предиката $R_\gamma(e, r, k) = q$ средства теории множеств не нужны: проверить, состоит ли q–й уровень в дереве разбиений только из "хороших" вершин (а (q–1)-й – содержит "нехорошую"), можно без каких-либо дополнительных предположений. Поэтому, следуя методу раздела 3.3, мы можем получить формулу $F_\gamma(e, r, k, q)$, выражающую в теории EA предикат $R_\gamma(e, r, k) = q$. Формула

$$(\forall e \forall r \forall k)(\exists q)\, F_\gamma(e, r, k, q),$$

с одной стороны, утверждает, что R_γ является всюду определенной функцией, а с другой стороны, она равносильна УКТР. Таким образом, если теория EA непротиворечива, то в ней невозможно доказать всюду определенность функции R_γ (а в теории ZFC это уже можно сделать).

Еще один удивительный факт связан с чрезвычайной скоростью роста функции R_γ. Уже функция R(e, r, k) из конечной теоремы Рамсея растет очень быстро (проверьте), однако R_γ превосходит в этом смысле всякое воображение. Можно показать, что если в теории EA доказуема всюду определенность функции f(x), т.е.

$$EA \vdash (\forall x)(\exists y)F(x, y),$$

где F(x, y) – формула, представляющая в EA функцию f, то

$$ZFC \vdash (\exists x_0)(\forall x > x_0)(f(x) < R_\gamma(x, x, x+1)).$$

Таким образом, функция $R_\gamma(x, x, x+1)$ растет быстрее любой функции, вычислимость которой можно доказать в элементарной арифметике! (См. Дж. Парис, Л. Харрингтон [1977].)

СПИСОК ЛИТЕРАТУРЫ

Для классических трудов в квадратных скобках дается год публикации оригинала.

А д а м а р Ж. [1945]

Исследование психологии процесса изобретения в области математики. – М.: Сов. радио, 1970. – 121 с.

Б и р ю к о в Б. В. [1985]

Жар холодных чисел и пафос бесстрастной логики. – М.: Знание, 1985. – 192 с.

Б у х ш т а б А. А. [1966]

Теория чисел. – М.: Просвещение, 1966. – 384 с.

Г и л ь б е р т Д. [1900]

Математические проблемы // Проблемы Гильберта. – М.: Наука, 1969. – сс. 11-64.

Г и л ь б е р т Д., Б е р н а й с П. [1934]

Основания математики: Логические исчисления и формализация арифметики. – М.: Наука, 1979. – 557 с.

Г и л ь б е р т Д., Б е р н а й с П. [1939]

Основания математики: Теория доказательств. – М.: Наука, 1982. – 652 с.

Г р э х е м Р. [1984]

Начала теории Рамсея. – М.: Мир, 1984. – 96 с.

Д р а г а л и н А. Г. [1979]

Математический интуиционизм: Введение в теорию доказательств. – М.: Наука, 1979. – 256 с.

Д э в и с М. (Davis M., ed.) [1965]

The Undecidable. Basic papers on undecidable propositions, unsolvable problems and computable functions. – New York: Raven Press, 1965. – 440 p.

Д э в л и н К. (Devlin K. J.) [1977]

The axiom of constructibility. A guide for the mathematician // Lecture Notes in Computer Science. – Berlin, Heidelberg, New-York: Springer, 1977.- Vol.

617. – 96 p.

Й е х Т. [1973]

Теория множеств и метод форсинга. – М.: Мир, 1973. – 150 с.

К а н о в е й В. Г. [1984]

Аксиома выбора и аксиома детерминированности. – М.: Наука, 1984. – 64 с.

К е л д ы ш Л. В. [1974]

Идеи Н.Н.Лузина в дескриптивной теории множеств // Успехи мат. наук. – 1974. – Т.29, вып.5. – сс. 183-196.

К л и н и С. К. [1957]

Введение в метаматематику. – М.: Изд-во иностр.лит., 1957.- 526 с.

К у з и ч е в а З. А. [1970]

К истории теорем о неполноте // История и методол. естеств. наук. Сер. Математика и механика. – 1970. – Вып.9. – сс. 182-189.

К у ш н е р Б. А. [1973]

Лекции по конструктивному математическому анализу. – М.: Наука, 1973. – 425 с.

Л е в и П. (Levy P.) [1926]

Sur le principe du tiers exclu et sur les theoremes non susceptibles de demonstrations // Rev. Metaphys. Morale. – 1926. – Vol.33, N2. – P. 253-258.

М а т и я с е в и ч Ю. В. [1970]

Диофантовость перечислимых множеств // Докл. АН СССР. – 1970. – Т.191, N1. – сс. 279-282.

М е д в е д е в Ф. А. [1965]

Развитие теории множеств в XIX веке. – М.: Наука, 1965. – 350 с.

М е д в е д е в Ф. А. [1982]

Ранняя история аксиомы выбора. – М.: Наука, 1982. – 303 с.

М е н д е л ь с о н Э. [1976] Введение в математическую логику. – М.: Мир, 1976. – 320 с.

П а р и с Дж., Х а р р и н г т о н Л. [1977]

Математическая неполнота в арифметике Пеано // Справочная книга по математической логике: В 4 ч. / Под ред. Дж.Барвайса. – М.: Наука, 1983. – Ч.4.- сс. 319-327.

П а р и х Р. (Parikh R.) [1971]

Existence and feasibility in arithmetic // J. Symbolic Logic. – 1971. – Vol.36, N3. – P. 494-508.

П о д н и е к с К. М. [1975]

Теорема о двойной неполноте // Учен. зап. Латв. ун-та. – 1975. – Т.233. – С. 191-200.

П о д н и е к с К. М. [1976]

Теорема о двойной неполноте // Четвертая Всесоюз. конф. по мат. логике. – Кишинев: Штиинца, 1976. – С. 80.

П о д н и е к с К. М. [1981]

Вокруг теоремы Геделя. – Рига: Латв. ун-т, 1981. – 105 с.

П о д н и е к с К. М. [1988а]

Платонизм, интуиция и природа математики. – Рига: Латв. ун-т, 1988. – 23 с.

П о д н и е к с К. М. [1988б]

Platonism, intuition and the nature of mathematics // Heyting'88. Summer School and Conf. on Math. Logic. September, 1988, Chaika, Bulgaria. Abstracts. – Sofia: BAN, 1988. – P. 50-51.

П у а н к а р е А. [1908]

Математическое творчество // Пуанкаре А. О науке. – М.: Наука, 1983. – С. 309-320.

Р а б и н М. [1977]

Разрешимые теории // Справочная книга по математической логике. В 4 ч. / Под ред. Дж.Барвайса. – М.: Наука, 1982. – Ч.3. – С. 77-111.

Р а ш е в с к и й П. К. [1973]

О догмате натурального ряда. //Успехи мат. наук. – 1973. – Т.28, вып.4. – С. 243-246.

С т я ж к и н Н. И. [1967]

Формирование математической логики. – М.: Наука, 1967. – 400 с.

У с п е н с к и й В. А. [1982]

Теорема Геделя о неполноте. – М.: Наука, 1982. – 112 с.

Карл Мартынович Подниекс
ВОКРУГ ТЕОРЕМЫ ГЕДЕЛЯ
Редактор Н. Д у н д и н а.
Художественный редактор И. Е г е р е.
Технический редактор Л. Д е м и д о в а.
Корректор Н. Л е б е д е в а.

Подписано в печать 05.12.91. Формат 60x84/.16.
Бумага офсетная. Печать офсетная.
11,16 усл.печ.л., 10,52 уч.-изд.л.
Тираж 1000 экз. Заказ N 143.
Издательство "Зинатне" ЛатвАН.
226530 ГСП Рига, ул.Тургенева, 19.
Регистрационное удостоверение N 20250.
Отпечатано в цехе оперативной полиграфии
издательства "Зинатне" ЛатвАН.
226050 Рига, ул. Мейстару, 10