Математический анализ: конспект лекций

В.И. Богачев

мат. фак. ВШЭ, 1 курс, осень 2021

Оглавление

Предисловие

Курс математического анализа является самым объемным на математических факультетах и в целом представляет собой фундамент для изучения всех последующих дисциплин «непрерывной математики». В отличие от остальных курсов этот предмет едва ли может быть отнесен к какой-либо особой области математики. Скорее, подобно армейской службе, он учит жизни. На начальной стадии освоения математический анализ имеет ряд очевидных пересечений (порой похожих на дублирования) с теорией множеств и алгеброй, но на последующих стадиях связи с другими областями (например с геометрией) выглядят уже как приложения. Неудивительно, что за полтора века существования математического анализа в университетских программах (под самыми разными названиями) было издано огромное число учебных пособий по нему, различающихся объемом, методическими принципами и предполагаемой аудиторией. Чрезвычайно велико число таких пособий и на русском языке. В каждом российском университете, включающем факультет математики, есть и свой курс математического анализа (в крупных университетах обычно не один), многие из этих курсов были изданы и нашли широкий круг читателей за пределами своего конкретного университета, поэтому было бы совершенно лишним приводить здесь какой-либо отражающий ситуацию список учебников. Однако представляется уместным отметить, что самым известным пособием по математическому анализу в нашей стране является превосходный трехтомник Г. М. Фихтенгольца [**19**], написанный более 70 лет назад одним из крупнейших специалистов своего времени по теории интеграла и остающийся ценнейшим источником информации по предмету. Трактат Фихтенгольца необходимо иметь преподавателям, в него полезно заглядывать студентам, но вряд ли кто-нибудь сейчас выберет его в качестве основного учебника, а не дополнительного пособия, причем дело здесь не только в его значительном объеме, но скорее в заметно изменившихся принципах изложения многих основных вопросов. Второе обстоятельство стало играть роль уже довольно давно, что привело к появлению также весьма объемных, но более современных курсов С. М. Никольского [**15**], Л. Д. Кудрявцева [**9**], В. А. Зорича [**6**] и других авторов, неоднократно издававшихся большими тиражами. Есть при этом (также вышедшие несколькими огромными тиражами) и гораздо более компактные учебники (естественно, охватывающие меньший материал), из которых можно отметить курс У. Рудина [**18**], впервые вышедший на английском языке в США в 50-х годах, но отнюдь не устаревший (в интересной автобиографии Рудина [**23**] рассказывается, что подвигло его на написание курса анализа вскоре после военной службы в британском флоте во время Второй мировой войны). Весьма велико и число задачников; классикой являются Гюнтер и Кузьмин [**4**] (современным студентам полезно хотя бы полистать задачник, впервые вышедший в 1912 году для студентов Института путей сообщения и многократно переиздававшийся) и Демидович [**5**] (на которого есть и АнтиДемидович [**12**]). Желающим углубить понимание основ анализа стоит познакомиться с книгами [**3**], [**13**], [**17**]; основательная проработка задач из последнего собрания — уже шаг к профессиональному знанию предмета.

Зачем читают лекции по математическому анализу и зачем пишут конспекты? Почему бы не предложить студентам самим почитать какой-либо учебник и сдать экзамен? Хотя последнее в принципе возможно и для отдельных студентов даже может быть более полезным, многолетняя практика убедительно показывает, что для большинства учащихся самым рациональным является обсуждение на лекциях и семинарах всяких разных мелочей, мимо которых легко пройти при самостоятельном чтении. Поэтому полезные лекции должны существенно отличаться и от учебника, и от даже лекторского конспекта самих этих лекций. По этой же причине чужой конспект или конспект лектора не могут заменить ни своего собственного, даже коряво написанного, ни, главное, активного участия в обсуждении предмета на лекциях и семинарах.

Глава 1

Числа и последовательности

1.1. Элементы теории множеств

В данном курсе анализа предполагается, что основные необходимые сведения по теории множеств известны из отдельного вводного курса логики. К этим сведениям относятся базовые неопределяемые понятия множества и отображения множеств. Считаются известными множества натуральных чисел $\mathbb{N}$, целых чисел $\mathbb{Z}$ и рациональных чисел $\mathbb{Q}$.

Оказывается удобным считать множеством пустое множество (не имеющее элементов); его обозначают символом $\varnothing$.

Принадлежность элемента a множеству A обозначают через $a \in A$, включение множества A в множество B через $A \subset B$. Если некая формула $P(x)$ с переменным x имеет место для всех элементов множества A, то пишут кратко $P(x)\,\forall\, x \in A$.

Объединение множеств A и B (совокупность их элементов) обозначается символом $A \cup B$. Пересечение множеств A и B (множество их общих элементов) обозначается через $A \cap B$. Разность множеств A и B, т. е. множество элементов из A, не лежащих в B, обозначается через $A \backslash B$.

Произведение непустых множеств X и Y обозначается через $X \times Y$. Оно состоит из всевозможных пар вида (x, y), где $x \in X$, $y \in Y$. Аналогично определяется произведение $n \geqslant 2$ множеств $X_1, \ldots, X_n$.

Последовательностью элементов множества A называют отображение из $\mathbb{N}$ в A; при этом образ числа n обозначают символом a_n или $a(n)$ и называют элементом с номером n. Рассматривают последовательности чисел, отображений, множеств и т. д.

Отображение $f\colon\ A \to B$ множества A в множество B называется *инъективным*, если $f(x)$ не совпадает с $f(y)$ для всяких несовпадающих x и y. Отображение f называется *сюръективным*, если для каждого b из B есть $a \in A$ с $b = f(a)$. Отображение называется *биективным*, если оно инъективно и сюръективно. Например, отображение $f\colon n \mapsto n+1$ множества $\mathbb{N}$ в себя инъективно, но не сюръективно, а вот на $\mathbb{Z}$ оно оказывается биекцией. В то же время отображение $f\colon n \mapsto n^2$ на $\mathbb{N}$ инъективно, а на $\mathbb{Z}$ нет. Самые простые примеры: $A = \{1, 2\}$, $B = \{1\}$, $f(1) = f(2) = 1$ — сюръекция, но не инъекция; $A = \{1\}$, $B = \{1, 2\}$, $f(1) = 1$ — инъекция, но не сюръекция.

Во многих вопросах анализа используется (часто неявно) так называемая *аксиома выбора*. Она заключается в следующем: если есть семейство непустых попарно непересекающихся множеств, то существует множество, содержащее ровно по одному элементу из каждого из данных множеств.

На первый взгляд кажется, что это нечто само собой разумеющееся, даже кажется странным выделение этого бесспорного утверждения в качестве аксиомы, но при более внимательном рассмотрении (которое началось в начале XX века) оказывается, что утверждение это вовсе не бесспорное из-за того, что вообще непонятно, что такое существование множества (поскольку понятие это неопределяемое). Хуже даже не это (что можно было бы отнести к области интереса философов), а то, что с помощью аксиомы выбора удается доказать такие утверждения, которые кажутся невозможными. Например, с помощью этой аксиомы доказывается, что шар в $\mathbb{R}^3$ можно разделить на конечное число частей, из которых движениями (без каких-либо растяжений) можно собрать два таких же шара. По такого рода причинам еще столетие назад многие видные математики стали считать недопустимыми рассуждения, основанные на аксиоме выбора. Беда, однако, в том, что если ее запретить, то окажется невозможным установить многие привычные утверждения. Например, нельзя будет доказать, что счетное объединение счетных множеств тоже счетно, нельзя будет переставлять суммирования в двойных рядах и многое другое. Правда, некоторым выходом из таких ситуаций служит разрешение использовать аксиому выбора для счетных совокупностей, но не для каких угодно. Такой выход кажется не менее искусственным, чем использование самой аксиомы выбора, поэтому ниже неявное ее использование не будет оговариваться специально.

Во всяком случае, привлечение аксиомы выбора весьма полезно при обсуждении мощностей бесконечных множеств. Само понятие мощности не определяется, говорят лишь о том, что множества A и B имеют равную мощность (равномощны), если между ними можно установить биекцию. Конечно, можно тем самым считать, что вводятся и мощности как классы равномощных множеств (иначе говоря, всякая мощность отождествляется с неким множеством, ее представляющим, причем равномощные с ним множества представляют ту же мощность).

Замечательным образом оказывается, что всякие два множества A и B сравнимы по мощности: справедлива следующая теорема.

1.1.1. Теорема. *Пусть A и B — множества. Тогда они либо равномощны, либо нет, причем в последнем случае либо A равномощно части B, либо B равномощно части A.*

Полезно еще знать следующую теорему Кантора – Бернштейна.

1.1.2. Теорема. *Если A равномощно части B, а B равномощно части A, то A и B равномощны.*

Разумеется, это не исключает того, что A и B могут быть равномощными при том, что A равномощно собственной части B. Скажем, интервал $(0,1)$ и отрезок $[0,1]$ равномощны, как мы увидим ниже.

Обе теоремы довольно неочевидны.

С помощью аксиомы выбора можно показать, что объединение двух бесконечных равномощных множеств равномощно этим множествам (см. [**2**]).

Вот еще несколько простых полезных фактов.

Счетным называют множество, равномощное $\mathbb{N}$.

Континуальным называют множество, равномощное множеству всех бесконечных последовательностей из 0 и 1.

1.1.3. Теорема. *Множество всех бесконечных последовательностей из 0 и 1 несчетно.*

Доказательство. В противном случае мы можем занумеровать его числами $n = 1, 2, \ldots$ и расположить в виде бесконечной таблицы, в которой первая строка — первая последовательность, вторая — вторая и т. д., причем всякая последовательность из 0 и 1 записана некоторой строкой. Чтобы получить противоречие, мы предъявим бы последовательность, заведомо не охваченную нумерацией: можно на первом месте поставить не то, что стоит у первой последовательности, на втором — не то, что у второй, и т. д. Полученная последовательность не может стоять никакой строкой: от написанного в строке с номером n она отличается числом на месте n. $\square$

Произведение двух континуальных множеств тоже континуально, ибо пары последовательностей $(x_n), (y_n)$ можно кодировать одинарными последовательностями $(x_1, y_1, x_2, y_2, \ldots)$, что дает взаимно однозначную кодировку (из всякой одинарной последовательности расщеплением получается единственная пара).

Произведение двух счетных множеств счетно. В самом деле, достаточно проверить, что счетно $\mathbb{N}\times\mathbb{N}$. Запишем это множество в виде таблицы, где первая строка состоит из пар $(1,1)$, $(1,2)$, $(1,3)$ и т. д., вторая — из пар $(2,1)$, $(2,2)$, $(2,3)$ и т. д., а следующие строки образуются аналогично. Это двухиндексное множество можно занумеровать натуральными числами так: берем расширяющиеся квадраты $n\times n$ в этой таблице и постепенно нумеруем их элементы, сначала квадрат 1×1, затем оставшиеся элементы квадрата 2×2 и т. д.

Всякое бесконечное множество A равномощно его объединению со счетным (или конечным) множеством B. В самом деле, в A есть счетная часть S (это и означает бесконечность), поэтому соответствие между $A = (A\backslash S) \cup S$ и $A \cup B$ устанавливается так: на $A\backslash S$ берем тождественное отображение и замечаем (что совсем просто), что S равномощно $S \cup B$.

Множества $\mathbb{Q}$ и $\mathbb{Z}$ счетны. Действительно, $\mathbb{Z}$ можно занумеровать натуральными числами так: использовать нечетные номера больше 1 для нумерации положительных целых чисел, использовать четные номера для нумерации отрицательных целых чисел, а нулю присвоить номер 1. Множество $\mathbb{Q}$ содержит счетное множество $\mathbb{N}$ и инъективно отображается в счетное (по уже доказанному) множество $\mathbb{Z}\times\mathbb{Z}$: несократимой дроби p/q сопоставляем пару (p, q). По теореме Кантора – Бернштейна получаем, что $\mathbb{Q}$ счетно. Конечно, можно и непосредственно занумеровать рациональные числа натуральными, модифицируя рассуждение для $\mathbb{N}\times\mathbb{N}$.

Г. Кантором было замечено, что множество всех подмножеств непустого множества X (оно обозначается через 2^X) всегда больше мощности самого X. В самом деле, мощность 2^X не меньше мощности X. Если бы они были равномощны, то каждой точке x мы могли бы сопоставить множество $M(x) \subset X$ так, что всякое множество $A \subset X$ было бы сопоставлено какой-то точке. Рассмотрим множество Z таких элементов x, что x не лежит в $M(x)$. Множество Z может быть и пустым, но в любом случае оно имеет вид $Z = M(a)$ для некоторого $a \in X$. Если $a \in Z$, то по определению Z мы имеем $a \notin M(a) = Z$. Значит, $a \notin Z = M(a)$. Но тогда a входит в Z! Таким образом, невозможен ни один из вариантов, что означает невозможность указанной нумерации.

Таким образом, можно строить множества все возрастающей мощности. Из этого ясно, что нельзя говорить о «множестве всех множеств», т. е. такой объект был бы внутренне противоречивым.

1.2. Вещественные числа

Вещественные числа можно ввести аксиоматически как множество $\mathbb{R}$, на котором заданы операции сложения, умножения и отношения порядка, подчиненные ряду условий (довольно длинному). Эти условия с целью запоминания (если их можно запомнить) полезно разбить на группы: первая относится к сложению, вторая к умножению, третья их связывает, четвертая говорит об отношении порядка и его связи с операциями.

Аксиомы сложения. Каждой паре чисел $x, y \in \mathbb{R}$ сопоставлено число $x + y$, причем

1. операция коммутативна, т. е. $x + y = y + x$,

2. существует нейтральный элемент 0, для которого выполнено равенство $x + 0 = 0 + x = x$ при всех x,

3. для всякого x имеется противоположный элемент $-x$, для которого $x + (-x) = 0$,

4. операция сложения ассоциативна: $x + (y + z) = (x + y) + z$.

Эти условия означают, что по сложению $\mathbb{R}$ является коммутативной группой.

Аксиомы умножения. Каждой паре чисел $x, y \in \mathbb{R}$ сопоставлено число $x \cdot y$, причем

1. операция коммутативна, т. е. $x \cdot y = y \cdot x$,

2. существует единица 1, отличная от 0, для которой

$$x \cdot 1 = 1 \cdot x = x \quad \text{при всех } x,$$

3. для всякого $x \neq 0$ имеется обратный элемент $x^{-1} = 1/x$, для которого $x \cdot x^{-1} = x^{-1} \cdot x = 1$,

4. операция умножения ассоциативна: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

Эти условия означают, что по умножению $\mathbb{R}\backslash\{0\}$ является коммутативной группой. Нулевой элемент фактически не участвует в аксиомах умножения, за исключением требования, что он не совпадает с единицей. Для сокращения записи вместо $x \cdot y$ пишут xy (кроме случаев, когда это может привести к недоразумениям, скажем, вместо $1 \cdot 2$ не стоит писать 12).

Связь сложения и умножения. Умножение дистрибутивно по отношению к сложению:

$$(x+y)z = xz + yz.$$

Порядок. Для упорядоченных пар чисел введено отношение порядка $\leqslant$ так, что для всякой пары $x, y \in \mathbb{R}$ установлено, что либо $x \leqslant y$, либо $y \leqslant x$, причем (i) $x \leqslant x$, (ii) одновременно $x \leqslant y$ и $y \leqslant x$ лишь при $x = y$, (iii) если $x \leqslant y$ и $y \leqslant z$, то $x \leqslant z$.

Связь порядка со сложением и умножением. Если $x \leqslant y$, то $x + z \leqslant y + z$ при всех z и если $0 \leqslant x$ и $0 \leqslant y$, то $0 \leqslant xy$.

Аксиома полноты в терминах порядка. Если даны непустые множества A и B, причем $a \leqslant b$ для всех $a \in A$ и $b \in B$ (это обозначается через $A \leqslant B$), то найдется $c \in \mathbb{R}$ такое, что $a \leqslant c \leqslant b$ для всех $a \in A$ и $b \in B$, т. е. c порядково разделяет A и B.

По определению полагают $x \geqslant y$, если $y \leqslant x$. Далее, пишут $x < y$, если $x \leqslant y$ и $x \neq y$. Наконец, $x > y$, если $y < x$. Если $x > 0$, то x называют положительным числом, а если $x \geqslant 0$, то x называют неотрицательным числом. Если $x < 0$, то x называют отрицательным числом, а если $x \leqslant 0$, то x называют неположительным числом.

Положим $|x| = x$ при $x \geqslant 0$ и $|x| = -x$ при $x < 0$. Заметим, что $|x| \geqslant 0$, ибо если $x < 0$, то $x + (-x) < 0 + (-x)$, т. е. $0 < -x$. Число $|x|$ называется модулем числа x.

Приведем некоторые следствия аксиом. Натуральные числа (множество $\mathbb{N}$) появляются последовательным добавлением 1 к 1. Целые числа (множество $\mathbb{Z}$) состоит из натуральных, противоположных им и нуля. Множество $\mathbb{Q}$ состоит из произведений целых чисел на обратные к натуральным.

Для множества $\mathbb{N}$ натуральных чисел справедлив принцип индукции: если множество S натуральных чисел содержит 1 и вместе с каждым своим элементом n содержит и $n+1$, то $S = \mathbb{N}$.

Верно равенство $x \cdot 0 = 0$ для всех x, хотя в списке аксиом его нет. Для доказательства заметим, что $0 \cdot x = (0+0)x = 0 \cdot x + 0 \cdot x$. Сложив обе части с $-(0 \cdot x)$, получаем $0 = 0 \cdot x$. Кроме того, в списке нет верного тождества $-x = (-1) \cdot x$. Для проверки заметим, что $(-1) \cdot x + x = (-1) \cdot x + 1 \cdot x = (-1+1) \cdot x = 0$. Прибавив к левой и правой частям полученного равенства $-x$, приходим к нужному тождеству.

Далее, заметим, что 0 и 1 единственны (докажите это!), причем верно неравенство $1 > 0$ (в аксиомах сказано лишь, что $1 \neq 0$). Действительно, пусть $1 < 0$. Тогда $1 + (-1) < -1$, т. е. $-1 > 0$, откуда $(-1)(-1) > 0$. Остается заметить, что $(-1)(-1) = 1$, ибо после прибавления к обеим частям -1 и использования равенства $-1 = (-1)1$ вместе с распределительным законом мы получим нуль. Аналогично получаем $x^2 \geqslant 0$ для всех x: при $0 \geqslant x$ это входит в аксиомы, а при $x < 0$ мы имеем $0 < -x$ (как пояснено выше), значит, $0 \leqslant (-x)(-x) = (-1)^2x^2 = x^2$.

Заметим еще, что $x \leqslant y$ в точности тогда, когда $x - y \leqslant 0$. Кроме того, при $x, y > 0$ соотношение $x < y$ равносильно соотношению $x^{-1} > y^{-1}$ (докажите это!). Далее, если $a \geqslant b$ и $c \geqslant 0$, то $ac \geqslant bc$.

Из аксиом вытекает привычное неравенство

$$|x + y| \leqslant |x| + |y|.$$

При $x \geqslant 0$, $y \geqslant 0$ или $x \leqslant 0, y \leqslant 0$ оно превращается в равенство. Если $x \geqslant 0$, $y \leqslant 0$, то $|x + y| = x + y$ или $-x - y$, что не больше $x - y$, ибо $-x \leqslant 0$, $y \leqslant -y$. Случай $x \leqslant 0$, $y \geqslant 0$ аналогичен.

Разумеется, возникает вопрос, почему существует $\mathbb{R}$, удовлетворяющее этому длинному списку требований. Некоторые соображения на этот счет приведены ниже в § 1.7.

Следствием аксиом является следующий *принцип Архимеда*:

если $0 < a < b$, *то найдется такое* $n \in \mathbb{N}$, *что* $na > b$.

В частности,

для всякого числа $x > 0$ *найдется такое* $n \in \mathbb{N}$, *что* $x < n$. Это равносильно тому, что для всякого числа $a > 0$ найдется такое $n \in \mathbb{N}$, что $n^{-1} < a$.

Докажем этот частный случай. В самом деле, иначе получаем, что для всех натуральных чисел верна оценка $n \leqslant x$, т. е. непусто множество B таких b, что $\mathbb{N} \leqslant b$. Ввиду аксиомы полноты найдется число c, разделяющее $\mathbb{N}$ и B. Это означает, что $n \leqslant c$ при всех $n \in \mathbb{N}$, но меньшего c числа с таким свойством уже нет. Это неверно: оно есть, годится также $c-1$, ибо $n \leqslant c-1$ при всех $n \in \mathbb{N}$, поскольку $n+1 \in \mathbb{N}$. Общий случай вытекает из этого частного применением к $x = ba^{-1}$.

Отметим, что возможно такое построение вещественных чисел, что принцип Архимеда станет аксиомой, а не следствием других аксиом, как было выше.

1.2.1. Следствие. *Для всяких чисел* $a < b$ *найдется такое рациональное число* r, *что* $a < r < b$.

Доказательство. Можно считать, что $b - a \leqslant 1$, иначе вместо b можно взять $b - 1$. Возьмем натуральное число n, для которого $n^{-1} < b - a$ и $n^{-1} < a$. Затем возьмем наименьшее натуральное k, для которого $kn^{-1} > a$. Поскольку $(k-1)n^{-1} \leqslant a$ и $n^{-1} < b - a$, то $kn^{-1} = (k-1)n^{-1} + n^{-1} < b$. Итак, подходит $r = kn^{-1}$. □

Это следствие открывает путь к построению модели вещественных чисел на основе метрического пополнения $\mathbb{Q}$.

Множество A *ограничено сверху*, если есть такое M, что $a \leqslant M$ при всех $a \in A$. При этом пишут $A \leqslant M$.

Множество A *ограничено снизу*, если есть такое m, что $a \geqslant m$ при всех $a \in A$.

Множество называют *ограниченным*, если оно ограничено сверху и снизу.

Для всякого ограниченного сверху множества A существует единственное число $\sup A$, называемое *точной верхней гранью* множества A и являющееся наименьшим из чисел M, для которых $A \leqslant M$.

Действительно, по условию непусто множество B таких чисел b, что $A \leqslant b$. При этом $A \leqslant B$. По аксиоме полноты найдется c, для которого $A \leqslant c \leqslant B$. Это и есть $\sup A$, ибо $A \leqslant c$, но меньшего числа уже нет, так как иначе оно попадет в B и будет неверно, что $c \leqslant B$.

Аналогично для всякого ограниченного снизу множества A существует единственное число $\inf A$, называемое *точной нижней гранью* множества A и являющееся наибольшим из чисел m, для которых имеем $A \geqslant m$.

Существование точной верхней грани для каждого ограниченного сверху множества равносильно аксиоме полноты в терминах порядка. В самом деле, если $A \leqslant B$ и A имеет точную верхнюю грань c, то $c \leqslant b$ при всех $b \in B$, ибо если $b < c$, то неверно, что c — точная верхняя грань.

Всем указанным аксиомам, кроме полноты, удовлетворяет множество рациональных чисел $\mathbb{Q}$. Последняя аксиома для него нарушена: для непустых множеств

$$A = \{r \in \mathbb{Q}\colon\ r > 0, r^2 < 2\} \quad \text{и} \quad B = \{s \in \mathbb{Q}\colon\ s > 0, s^2 > 2\}$$

мы имеем $A \leqslant B$, но нет разделяющего рационального c. В самом деле, если $a \in A$ и $b \in B$, то $b^2 - a^2 > 0$, но $b^2 - a^2 = (a+b)(b-a)$, где $a + b > 0$, поэтому $b - a > 0$, т. е. $b > a$. Предположим, что найдется разделяющее $c \in \mathbb{Q}$. Читатель должен уметь доказывать (от противного), что $c^2 \neq 2$. Если $c^2 < 2$, то при достаточно большом n получаем

$(c+1/n)^2 = c^2+2c/n+1/(n^2) < 2$, ибо при достаточно большом n число $1/n$ становится меньше положительного числа $(2-c^2)/4$. Значит, $c+1/n \in A$, что невозможно, ибо $a \leqslant c$ при всех $a \in A$. Аналогично к противоречию приводит предположение, что $c^2 > 2$. Кстати, попутно доказано, что есть число $c \geqslant 0$, для которого $c^2 = 2$.

Особо отметим, что с порядковой неполнотой оказался связан тот факт, что в $\mathbb{Q}$ не нашлось решения уравнения $x^2 = 2$. Вещественное число, удовлетворяющее уравнению вида

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

с целыми коэффициентами a_i, где $a_n \neq 0$, называется *алгебраическим*. Прочие вещественные числа называются *трансцендентными*. Число $\sqrt{2}$ иррационально, но алгебраично, а вот π и e трансцендентны (что доказывается довольно сложно и было установлено лишь в конце XIX века). Про некоторые числа довольно простого вида до сих пор неизвестно, являются ли они трансцендентными (про $e+\pi$ неизвестна даже иррациональность).

Мы увидим ниже, что все уравнения $x^2 = c > 0$ имеют решения в $\mathbb{R}$, но вот алгебраическое уравнение $x^2 = -1$ решений не имеет. Тем самым поле вещественных чисел оказывается алгебраически неполным, несмотря на выполнение аксиомы полноты для порядка и метрическую полноту. Чтобы исправить положение, вводится поле $\mathbb{C}$ комплексных чисел, в котором разрешимы все нетривиальные алгебраические уравнения с комплексными коэффициентами. Однако в $\mathbb{C}$ нет естественного порядка, связанного с топологией аналогично случаю $\mathbb{R}$ (докажите это; конечно, как-то упорядочить $\mathbb{C}$ можно, например, используя равномощность с $\mathbb{R}$ или введя так называемый лексикографический порядок, при котором $z_1 < z_2$, если либо $\operatorname{Re} z_1 < \operatorname{Re} z_2$, либо $\operatorname{Re} z_1 = \operatorname{Re} z_2$, но $\operatorname{Im} z_1 < \operatorname{Im} z_2$).

Далее используется следующая терминология.

Интервалом с концами a и b, где $a < b$, называется множество

$$(a, b) = \{x\colon\ a < x < b\}.$$

Отрезком с концами a и b называется множество

$$[a, b] = \{x\colon\ a \leqslant x \leqslant b\},\ \text{ где } a \leqslant b.$$

Отрезок без одного из концов называют *полуинтервалом*, а интервалы, полуинтервалы и отрезки называют *промежутками*.

Длиной промежутка с концами в a и b называется $b - a$.

Открытыми лучами называют множества $(-\infty, a)$ и $(b, +\infty)$. Замкнутые лучи — множества $(-\infty, a]$ и $[b, +\infty)$. Значения терминов «открытый» и «замкнутый» прояснится ниже при обсуждении основ топологии прямой.

Такая же терминология (кроме длины) употребляется для любых упорядоченных множеств.

В заключение этого параграфа докажем две очень важные теоремы про отрезки.

1.2.2. Теорема. (Принцип вложенных отрезков) *Пусть отрезки $I_n = [a_n, b_n]$ таковы, что $I_{n+1} \subset I_n$. Тогда их пересечение непусто. Если для всякого $k \in \mathbb{N}$ найдется номер n с $b_n - a_n \leqslant k^{-1}$, то пересечение состоит из одной точки.*

Доказательство. Для множеств $A = \{a_n\}$ и $B = \{b_n\}$ имеем $A \leqslant B$, ибо при $k \geqslant n$ верна оценка $a_n \leqslant a_k \leqslant b_k$. По аксиоме полноты найдется c, для которого $a_n \leqslant c \leqslant b_n$ при всех n, т. е. c лежит в пересечении. Если выполнено второе условие, то не может быть двух точек c_1 и c_2 в пересечении, ибо тогда при $c_2 > c_1$ найдется $k \in \mathbb{N}$ с $k^{-1} < c_2 - c_1$, значит, взяв n, для которого $b_n - a_n < k^{-1}$, получим противоречие. □

1.2.3. Теорема. (Принцип конечных подпокрытий) *Если отрезок покрыт бесконечным набором интервалов, то из них можно выбрать конечное число, объединение которых покрывает этот отрезок.*

Доказательство. Предположим противное. Тогда хотя бы одна из двух половин данного отрезка I (обозначим ее через I_1) не может быть покрыта конечным поднабором данных интервалов. По индукции получаем половины I_n предыдущих половин I_{n-1}, не допускающие конечных подпокрытий. По предыдущей теореме отрезки I_n имеют общую точку. Она лежит внутри одного из интервалов J данного покрытия. Так как длина I_n равна произведению длины I и 2^{-n}, то при достаточно большом n получим, что $I_n \subset J$, что невозможно: ведь I_n не допускал конечного подпокрытия. □

Заметим, что для интервала теорема неверна: $(0, 1)$ покрывается всеми интервалами $(n^{-1}, 1)$, но какой-либо конечной их части уже не хватит. Кроме того, неверно будет утверждение, если покрывать не интервалами, а отрезками: достаточно взять покрытие $[0, 1]$ одноточечными отрезками $[x, x]$ или такое покрытие нетривиальными отрезками: $[0, 1/2]$ и $[x, 1]$, где $x > 1/2$. Таким образом, определенная асимметрия в условии теоремы существенна.

1.3. Сходимость последовательностей

Последовательность чисел x_n обозначается через $\{x_n\}$.

1.3.1. Определение. *Последовательность чисел x_n называют сходящейся к числу x, называемому пределом последовательности $\{x_n\}$, если для каждого $\varepsilon > 0$ найдется такой номер N_ε, что*

$$|x - x_n| < \varepsilon \quad \textit{при всех } n \geqslant N_\varepsilon.$$

Обозначения: $x = \lim\limits_{n\to\infty} x_n$ *или* $x_n \to x$.

Отметим, что $x_n \to x$ в точности тогда, когда $x_n - x \to 0$. Поэтому последовательность $\{x_n\}$ сходится к x в точности тогда, когда

$$x_n = x + \varepsilon_n, \quad \text{где } \varepsilon_n \to 0.$$

Ясно, что последовательность может иметь только один предел и что сходящаяся последовательность ограничена.

Еще одна полезная банальность: если $\{x_n\}$ сходится, то туда же сходится и $\{x_{n+1}\}$.

Надо привыкнуть к тому, что в определении не требуется явно указывать зависимость N_ε от ε; номер N_ε не обязан быть наименьшим, с которого начинается нужная оценка: важно лишь, чтоб когда-то последовательность уже не выходила из ε-зазора вокруг x. Никакой монотонности и регулярности стремления к пределу не требуется; также нет ограничений на скорость сходимости, она может быть сколь угодно медленной.

Частичный предел последовательности есть предел какой-либо ее сходящейся подпоследовательности. Сходящаяся последовательность не может иметь двух различных частичных пределов (ибо достаточно малые окрестности двух точек не пересекаются и нельзя в них одновременно находиться).

Вот простейшие факты, связанные с пределами.

1.3.2. Предложение. (i) *Если $x_n \to 0$ и последовательность $\{y_n\}$ ограничена, то $x_n y_n \to 0$.*

(ii) *Если $\{x_n\}$ сходится к x, а $\{y_n\}$ сходится к y, то*

$$x_n + y_n \to x + y \quad \textit{и} \quad x_n y_n \to xy.$$

Если при этом $y_n, y \neq 0$, то

$$\frac{x_n}{y_n} \to \frac{x}{y}.$$

(iii) *Если $x_n \to x$ и $y_n \to y$, причем $x_n \leqslant y_n$, то $x \leqslant y$.*

Доказательство. (i) Если $|y_n| \leqslant M$, где $M > 0$, то имеем оценку $|x_n y_n| \leqslant M|x_n| < \varepsilon$, как только N_ε взято так, что $|x_n| < \varepsilon/M$ (это возможно, ибо $x_n \to 0$).

(ii) Берем L_ε так, что $|x_n - x| < \varepsilon/2$ при $n \geqslant L_\varepsilon$, затем $N_\varepsilon \geqslant L_\varepsilon$ так, что $|y_n - y| < \varepsilon/2$ при $n \geqslant N_\varepsilon$, после чего замечаем, что

$$|x_n + y_n - (x + y)| \leqslant |x_n - x| + |y_n - y| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

при $n \geqslant N_\varepsilon$.

Далее, $x_n y_n - xy = (x_n - x)y_n + x(y_n - y)$, где каждое из слагаемых стремится к нулю в силу (i), поэтому по доказанному стремится к нулю и сумма.

Пусть теперь $y_n, y \neq 0$. Достаточно проверить, что $1/y_n \to 1/y$. Замечаем, что

$$\frac{1}{y_n} - \frac{1}{y} = \frac{y - y_n}{y_n y}.$$

Так как $y - y_n \to 0$, то ввиду (i) достаточно проверить, что последовательность чисел $1/y_n$ ограничена. Иначе говоря, надо установить, что $|y_n| \geqslant c$ при некотором $c > 0$. Это ясно из того, что $|y_n - y| \geqslant |y|/2$ при всех n с некоторого номера N_1, т. е. $|y_n| \geqslant |y|/2$ для таких n, а конечное число оставшихся членов отличны от нуля.

(iii) Если $x > y$, то для $\varepsilon = (x - y)/4$ получаем, что $x_n > x - \varepsilon$ и $y_n < y + \varepsilon$ при всех достаточно больших n, но тогда $x_n > y_n$, ибо $x - \varepsilon > y + \varepsilon$. □

Следующее простое утверждение называют «леммой о двух милиционерах».

1.3.3. Лемма. *Если две последовательности $\{x_n\}$ и $\{y_n\}$ сходятся к числу a и $\{z_n\}$ такова, что $x_n \leqslant z_n \leqslant y_n$, то $z_n \to a$.*

Доказательство. Замечаем, что

$$z_n - a = z_n - y_n + y_n - a \leqslant y_n - a, \quad z_n - a = z_n - x_n + x_n - a \geqslant x_n - a,$$

поэтому достаточно взять N_ε так, что $y_n - a < \varepsilon$ и $x_n - a > -\varepsilon$ при всех $n \geqslant N_\varepsilon$. □

Для ограниченной числовой последовательности $\{x_n\}$ символом $\overline{\lim\limits_{n\to\infty}}\, x_n$ (или символом $\limsup\limits_{n\to\infty} x_n$) обозначают ее наибольший частичный предел, т. е. точную верхнюю грань частичных пределов (легко видеть, что она тоже является частичным пределом). Аналогично символом $\varliminf_{n\to\infty} x_n$ (или символом $\liminf_{n\to\infty} x_n$) обозначается наименьший частичный предел. Эти величины называют *верхним пределом* и *нижним пределом* соответственно.

1.3.4. Предложение. *Всякая ограниченная возрастающая последовательность сходится к своей точной верхней грани, а всякая ограниченная убывающая последовательность сходится к своей точной нижней грани.*

Доказательство. Пусть $a_n \leqslant a_{n+1}$ и a — точная верхняя грань последовательности $\{a_n\}$. Пусть $\varepsilon > 0$. Найдется такое число N, что $a_N > a - \varepsilon$ (иначе a не есть точная верхняя грань). Тогда $a_n > a - \varepsilon$ при всех $n \geqslant N$. Кроме того, $a_n \leqslant a$ при всех n. □

Конечно, не всякая ограниченная последовательность сходится, например, $x_n = (-1)^n$ не имеет предела. Однако следующая теорема Больцано – Вейерштрасса дает некоторую компенсацию.

1.3.5. Теорема. *Всякая ограниченная последовательность содержит сходящуюся подпоследовательность.*

Доказательство. Пусть $\{x_n\}$ лежит в отрезке I. Хотя бы в одну половину I_1 отрезка I попадает бесконечно много членов последовательности. Пусть x_{n_1} — один из них. Хотя бы в одну половину I_2 отрезка I_1 попадает бесконечно много членов последовательности. Пусть x_{n_2} — один из них, причем $n_2 > n_1$. По индукции находим возрастающие номера n_k, для которых элемент x_{n_k} попадает в половину I_k отрезка I_{k-1}, в которую попали бесконечно многие элементы последовательности. Полученные вложенные отрезки со стремящимися к нулю длинами имеет единственную общую точку. Из построения очевидно, что к ней и сходится $\{x_{n_k}\}$. □

Введем еще одно важное понятие.

1.3.6. Определение. *Последовательность чисел x_n называют фундаментальной* (*или последовательностью Коши*)*, если для каждого $\varepsilon > 0$ найдется такой номер N_ε, что*

$$|x_n - x_k| < \varepsilon \quad \textit{при всех } n, k \geqslant N_\varepsilon.$$

Всякая фундаментальная последовательность $\{x_n\}$ ограничена, т. е. лежит в некотором отрезке. Действительно, найдется такой номер N_1, что $|x_n - x_{N_1}| < 1$ при всех $n \geqslant N_1$, поэтому все члены этой последовательности с номерами не менее N_1 лежат в отрезке $[x_{N_1} - 1, x_{N_1} + 1]$, расширив который, легко охватить и первые N_1 членов.

1.3.7. Лемма. *Пусть последовательность $\{x_n\}$ фундаментальна. Если ее подпоследовательность $\{x_{n_k}\}$ сходится к x, то вся последовательность сходится к x.*

Доказательство. Пусть $x = \lim\limits_{k\to\infty} x_{n_k}$ и $\varepsilon > 0$. Найдем номер N с $|x_n - x_k| < \varepsilon/2$ при $n \geqslant N$. Увеличив N, можно считать, что верна оценка $|x_{n_k} - x| < \varepsilon/2$ при $k \geqslant N$. Теперь при $n \geqslant n_N$ получаем

$$|x_n - x| \leqslant |x_n - x_{n_N}| + |x_{n_N} - x| < \varepsilon,$$

что доказывает равенство $x = \lim\limits_{n\to\infty} x_n$. □

Легко видеть также, что сходящаяся последовательность $\{x_n\}$ фундаментальна: если N_ε взято так, что $|x - x_n| < \varepsilon/2$ при всех $n \geqslant N_\varepsilon$, то

$$|x_n - x_k| \leqslant |x_n - x| + |x - x_k| < \varepsilon$$

при всех $n \geqslant N_\varepsilon$. Для всего множества $\mathbb{R}$ верно и обратное (это утверждение называют *критерием сходимости Коши*).

1.3.8. Теорема. *Всякая фундаментальная последовательность в $\mathbb{R}$ сходится.*

Доказательство. Данная фундаментальная последовательность $\{x_n\}$ лежит в некотором отрезке $[-M, M]$. Значит, она содержит сходящуюся подпоследовательность. По предыдущей лемме вся исходная последовательность сходится. □

Важность этого несложного утверждения в том, что оно дает существование предела без знания самого предела (в определении фундаментальности никакого предела не фигурирует).

Вот еще одно очень простое полезное соображение, связанное с фундаментальностью.

1.3.9. Предложение. *Во всякой фундаментальной последовательности $\{a_n\}$ есть такая подпоследовательность $\{a_{n_k}\}$, что справедлива оценка $|a_{n_k} - a_{n_j}| \leqslant 2^{-k}$ при всех k и $j \geqslant k$.*

Доказательство. Находим номер n_1 так, что $|a_n - a_k| < 2^{-1}$ при $n, k \geqslant n_1$, затем $n_2 > n_1$ так, что $|a_n - a_k| < 2^{-2}$ при $n, k \geqslant n_2$, и т. д. □

Отметим следующую особенность последовательностей с оценкой типа

$$|a_{n+1} - a_n| \leqslant c2^{-n}$$

лишь между соседними членами. В этом случае имеем при всех $k > n$ оценку

$$|a_k - a_n| \leqslant |a_k - a_{k-1}| + \cdots + |a_{n+1} - a_n| \leqslant c2^{-k+1} + \cdots + c2^{-n} \leqslant 2c2^{-n},$$

что дает фундаментальность последовательности $\{a_n\}$.

Для фундаментальности НЕДОСТАТОЧНО, чтобы $a_n - a_{n+1} \to 0$.

1.3.10. Пример. (Гармонический ряд) Пусть

$$a_n = 1 + 2^{-1} + \cdots + n^{-1}.$$

Тогда $a_{n+1} - a_n = (n+1)^{-1} \to 0$, но конечного предела $\{a_n\}$ не имеет и возрастает к бесконечности. В самом деле,

$$a_{2n} - a_n = (n+1)^{-1} + \cdots + (2n)^{-1},$$

где стоит n слагаемых, каждое не меньше последнего, поэтому имеем $a_{2n} - a_n \geqslant 1/2$, значит, $\{a_n\}$ не фундаментальна.

Напомним, что определение сходимости не требует как-то конструктивно задавать N_ε через ε, не требуется также искать наименьший подходящий номер, надо лишь уметь для каждого ε обеспечить каким угодно способом какой-то. Очень редко можно явно указать N_ε. Например, если $a_n = n^{-1}$, то для $n \geqslant N_\varepsilon = [\varepsilon^{-1}] + 1$, где $[a]$ — целая часть числа a, получаем $a_n < \varepsilon$. Вот примеры, в которых N_ε явно не предъявляется.

1.3.11. Пример. (i) При $A > 1$ имеем

$$\lim_{n\to\infty} \frac{n}{A^n} = 0.$$

Сначала заметим, что, взяв $q \in (0, A^{-1})$, для всех достаточно больших номеров n получим

$$\frac{a_{n+1}}{a_n} = \frac{n+1}{n} A^{-1} \leqslant q < 1,$$

ибо $(n+1)/n = 1 + n^{-1} \to 1$. Значит, начиная с некоторого номера, $\{a_n\}$ убывает и имеет предел $a \geqslant 0$. Случай $a > 0$ невозможен, ибо тогда по свойствам пределов $a_{n+1}/a_n \to a/a = 1$, хотя мы видели, что $a_{n+1}/a_n \leqslant q < 1$ при больших n.

(ii) При $A > 1$ и всех k имеем

$$\lim_{n\to\infty} \frac{n^k}{A^n} = 0.$$

Это достаточно проверить для $k \in \mathbb{N}$, но тогда по (i) имеем соотношение $n/(A^{1/k})^n \to 0$, а теперь можно перемножить k таких соотношений.

(iii) Для всякого A имеем

$$\lim_{n\to\infty} \frac{A^n}{n!} = 0.$$

Достаточно проверить это для $A \in \mathbb{N}$. Опять берем отношение

$$\frac{a_{n+1}}{a_n} = \frac{A}{n+1},$$

замечаем, что при $n > A$ последовательность убывает, опять имеет предел $a \geqslant 0$, опять невозможен случай $a = 0$.

(iv) Число e (латинское е) задается равенством

$$e = \lim_{n\to\infty} \Bigl(1 + \frac{1}{n}\Bigr)^n.$$

Почему существует предел и почему он конечен? В задаче 1.9.15 предлагается проверить, что новая последовательность

$$y_n = (1 + 1/n)^{n+1}$$

убывает. Она уж точно имеет некоторый предел e, но

$$x_n = y_n(1 + 1/n)^{-1},$$

где второй сомножитель стремится к единице, поэтому $\{x_n\}$ имеет тот же предел. В следующем параграфе будет указано хорошее приближение e, позволяющее оценивать его с любой разумной точностью без калькуляторов. Окажется, что $e = 2,71\ldots$, но попробуйте непосредственно установить, что $2 < e < 3$.

(v) Верно равенство

$$\lim_{n\to\infty} \Bigl(1 - \frac{1}{n}\Bigr)^n = e^{-1}.$$

Действительно, из предыдущего равенства, задающего e как предел последовательности $\{y_n\}$, и свойств пределов следует, что e^{-1} является пределом последовательности чисел

$$\Bigl(\frac{n}{n+1}\Bigr)^{n+1} = \Bigl(1 - \frac{1}{n+1}\Bigr)^{n+1}.$$

(vi) Предыдущие примеры очень характерны и бывают полезны для анализа многих других сходимостей, но надо помнить, что никаких «контрольных» последовательностей, определяющих сходимость, нет. В частности, какова бы ни была сходящаяся к нулю последовательность чисел $\varepsilon_n > 0$, найдется медленнее сходящаяся к нулю последовательность, т. е. найдется неограниченно возрастающая последовательность чисел $C_n > 0$, для которых $C_n\varepsilon_n \to 0$. В самом деле, найдем возрастающие номера N_k с тем свойством, что $|\varepsilon_n| < k^{-2}$ при $n \geqslant N_k$. Теперь возьмем C_n так: $C_1 = 1$ при $n < N_2$, $C_2 = 2$ при $N_2 \leqslant n < N_3$,..., $C_n = k$ при $N_k \leqslant n < N_{k+1}$ и т. д.

Наконец, полезно допустить в качестве пределов бесконечности. Говорят, что $\{x_n\}$ стремится к плюс бесконечности (или расходится к плюс бесконечности) и пишут $\lim\limits_{n\to\infty} x_n = +\infty$, если для всякого $R > 0$ найдется такой номер n_R, что $x_n > R$ при всех $n \geqslant n_R$. Аналогично определяется стремление к минус бесконечности. Чтобы хоть как-то отличить это от случая конечного предела, такие последовательности не называют сходящимися. Разумеется, фундаментальности здесь нет. Понятно, что для $x_n > 0$ соотношение $\lim\limits_{n\to\infty} x_n = +\infty$ равносильно соотношению $\lim\limits_{n\to\infty} x_n^{-1} = 0$.

1.3.12. Пример. (i) Рассмотрим последовательность

$$a_n = \frac{an+b}{cn+d}, \quad c \neq 0.$$

Здесь при ненулевом a числитель и знаменатель имеют бесконечные пределы, поэтому неприменимо указанное выше правило обращения с дробями. Однако после деления на n мы получаем отношение последовательностей $a + b/n$ и $c + d/n$, сходящихся к a и c соответственно, т. е. теперь правило уже применимо.

(ii) Рассмотрим последовательность

$$a_n = \frac{P(n)}{Q(n)},$$

где $P(x) = c_k x^k + \cdots + c_0$, $Q(x) = b_k x^k + \cdots + b_0$, причем $b_k \neq 0$. Здесь такая же первоначальная трудность, как и в (i), но деление на n ее не снимает (если $k > 1$), легко догадаться, что делить надо на n^k. Тогда оказывается, что $P(n)/n^k = c_k + c_{k-1}n^{-1} + \cdots + c_0 n^{-k} \to c_k$ при $n \to \infty$, аналогично $Q(n)/n^k \to b_k$, так что искомый предел есть отношение c_k/b_k.

(iii) Рассмотрим последовательность

$$a_n = \frac{3^n(2n+1) - 2^n}{3^n(n-1) + n2^n}.$$

Сразу приходит в голову поделить на 3^n, но это не помогает, ибо снова имеем выражения с бесконечными пределами. Затем догадываемся еще и на n поделить. Тогда получаем $2 + n^{-1} - (2/3)^n n^{-1}$ в числителе и $1 - n^{-1} + (2/3)^n$ в знаменателе. Это уже подпадает под действие правила и дает ответ 2. Почему же не стали делить на 6^n? Это сразу дало бы конечные пределы, но, к сожалению, в знаменателе был бы нулевой предел, что запрещено нашим правилом.

Таким образом, неочевидные ситуации возникают, когда надо найти предел дроби, у которой числитель и знаменатель имеют оба бесконечный предел или нулевой предел. В большинстве случаев бесконечные пределы числителя и знаменателя сводятся к нулевым, а для эффективного исследования последних нужны асимптотические разложения, основанные на формуле Тейлора из главы 3.

1.4. Ряды

Сходимость рядов представляет собой частный случай сходимости последовательностей, но специфика этого случая столь существенна, а роль рядов в приложениях традиционно столь велика, что этому случаю уделяют даже большее внимание, чем случаю последовательностей.

1.4.1. Определение. *Пусть дана последовательность вещественных чисел a_n. Если последовательность сумм*

$$S_n := a_1 + \cdots + a_n = \sum_{i=1}^{n} a_i$$

имеет конечный предел S, то говорят, что ряд с общим членом a_n сходится к S и пишут $S = \sum_{n=1}^{\infty} a_n$.

Указанные суммы S_n называют частичными суммами ряда с общим членом a_n.

В случае $a_n \geqslant 0$ пишут $\sum_{n=1}^{\infty} a_n = \infty$, если S_n возрастают к бесконечности.

Случай $a_n \geqslant 0$ удобен тем, что здесь сходимость ряда сводится к ограниченности частичных сумм S_n: ведь получается возрастающая последовательность. Конечно, в общем случае не так.

Из критерия Коши сходимости последовательности вытекает, что сходимость ряда из a_n равносильна тому, что для всякого $\varepsilon > 0$ найдется такой номер N_ε, что $|S_n - S_m| < \varepsilon$ при $n, m \geqslant N_\varepsilon$, что записывается

как
$$\Big|\sum_{i=n}^{m} a_i\Big| < \varepsilon, \quad n, m \geqslant N_\varepsilon.$$
Из этого видно, что сходимость ряда влечет сходимость a_n к нулю, но расходящийся гармонический ряд из n^{-1} показывает (пример 1.3.10), что последнее недостаточно для сходимости ряда.

Поскольку $\Big|\sum_{i=n}^{m} a_i\Big| \leqslant \sum_{i=n}^{m} |a_i|$, то сходимость ряда $\sum_{n=1}^{\infty} |a_n|$ (так называемая *абсолютная сходимость* ряда) влечет сходимость исходного ряда $\sum_{n=1}^{\infty} a_n$.

Обратное неверно: рассмотрим знакопеременный ряд Лейбница
$$1 - \frac{1}{2} + \frac{1}{3} + \cdots, \quad a_n = (-1)^{n+1} n^{-1}.$$
Мы знаем, что ряд из модулей (гармонический ряд) расходится. Однако без модулей ряд сходится: для этого достаточно заметить, что сумма
$$\frac{1}{N} - \frac{1}{N+1} + \frac{1}{N+2} + \cdots + \frac{(-1)^{M+1}}{M}$$
всегда лежит между 0 и N^{-1}, что обеспечивает фундаментальность последовательности частичных сумм. Это же рассуждение доказывает более общий факт.

1.4.2. Предложение. *Пусть последовательность чисел $a_n > 0$ монотонно убывает к нулю. Тогда знакопеременный ряд с общим членом $(-1)^n a_n$ сходится.*

Из сказанного выше ясно также, что сходимость ряда из a_n (даже абсолютная) обеспечивается сравнением $|a_n| \leqslant b_n$ с членами b_n неотрицательного сходящегося ряда. Этот, на первый взгляд очень банальный, способ чаще всего и используется. Более того, чаще всего используются сравнения с некоторыми специальными ряда, рассмотренными ниже, а также с геометрической прогрессией.

Напомним, что для всякого $q \neq 1$ верно равенство
$$1 + q + \cdots + q^n = \frac{q^{n+1} - 1}{q - 1}.$$
Оно доказывается умножением обеих частей на $q - 1$. Поэтому при $|q| < 1$ ряд сходится:
$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

При $0 < q < 1$ получается полезная оценка при всех m, l:

$$q^m + q^{m+1} + \cdots + q^{m+l} \leqslant \frac{q^m}{1-q}.$$

Сравнением с подходящей геометрической прогрессией проверяется сходимость ряда в признаках Коши и Даламбера.

1.4.3. Теорема. (Признак Коши) *Пусть*

$$q = \overline{\lim_{n\to\infty}} |a_n|^{1/n}.$$

Если $q < 1$, то ряд $\sum_{n=1}^{\infty} a_n$ сходится, причем абсолютно, если $q > 1$, то этот ряд не сходится. Для $q = 1$ бывают сходящиеся и не сходящиеся ряды.

Доказательство. При $q < 1$ возьмем $q_1 \in (q, 1)$ и найдем N такое, что $|a_n|^{1/n} \leqslant q_1$ при $n \geqslant N$. Тогда $|a_n| \leqslant q_1^n$, что обеспечивает абсолютную сходимость. Если же $q > 1$, то для бесконечно многих n имеем $|a_n| \geqslant 1$, так что нет даже сходимости к нулю общего члена. Ряды с общими членами n^{-1} и n^{-2} дают примеры расходящихся и сходящихся рядов с $a_n^{1/n} \to 1$. □

Как видно из доказательства, условие Коши равносильно оценке

$$|a_n| \leqslant Cq^n$$

с некоторым $C \geqslant 0$ и $q \in [0, 1)$. Поэтому этот признак не может быть очень эффективным, он не работает уже для хорошо сходящегося ряда из n^{-2} и не отличает его от гармонического ряда. Следующий признак Даламбера еще слабее, но бывает удобен из-за того, что привлекает отношение соседних членов ряда.

1.4.4. Теорема. (Признак Даламбера) *Пусть*

$$q = \overline{\lim_{n\to\infty}} \frac{|a_{n+1}|}{|a_n|}.$$

Если $q < 1$, то ряд $\sum_{n=1}^{\infty} a_n$ сходится, причем абсолютно.

Доказательство. Если $q < 1$, то при некотором номере N для всех $n \geqslant N$ имеем $|a_{N+n}| \leqslant q|a_{N+n-1}| \leqslant \cdots \leqslant q^n|a_N|$, что дает сходимость из сравнения с прогрессией. □

Ясно, что при выполнении условия Даламбера выполнено и условие Коши. Однако если взять даже очень хорошо сходящийся ряд с положительными членами, скажем, прогрессию $\{q^n\}$, и писать его члены по два раза, скажем, $1, 1, q, q, q^2, q^2, \ldots$, то для условия Даламбера это даст бесконечно много единиц, так что распознать сходимость не удастся, условие же Коши такой фокус с прогрессией (или иным рядом, сходимость которого оно определяло) раскусит. Отметим еще, что в условии Даламбера оценка $q > 1$ не дает отрицательных выводов: можно в примере выше брать вместо второго одинакового члена удвоенный, т. е. $1, 2, q, 2q, q^2, 2q^2$ и т. д.

Позже мы рассмотрим несколько более тонких признаков сходимости рядов, но сейчас приведем лишь один полезный результат о сходимости рядов с положительными членами (его называют признаком Гаусса или признаком Коши).

1.4.5. Предложение. *Пусть последовательность чисел $a_n > 0$ монотонно убывает к нулю. Тогда сходимость ряда с общим членом a_n равносильна сходимости ряда с общим членом $2^k a_{2^k}$. Иначе говоря, $\sum_{n=1}^\infty a_n < \infty \iff \sum_{k=1}^\infty 2^k a_{2^k} < \infty$.*

Доказательство. Заметим, что члены ряда с номерами между 2^k и $2^{k+1} - 1$ не меньше $a_{2^{k+1}}$ и не больше a_{2^k}. Поэтому их сумма лежит между $2^k a_{2^{k+1}}$ и $2^k a_{2^k}$. Тем самым сходимость последнего ряда обеспечивает ограниченность частичных сумм исходного, а сходимость исходного ряда влечет сходимость ряда из $2^k a_{2^{k+1}}$, равносильную сходимости ряда из $2^k a_{2^k}$. □

Кажется удивительным, что о сходимости ряда можно судить по очень редкой подпоследовательности индексов 2^k, но не следует забывать требование монотонности (без него, конечно, критерий неверен). Приведем примеры применения этого критерия.

1.4.6. Пример. (i) Гармонический ряд с $a_n = 1/n$ просто берется этим критерием, ибо $2^k a_{2^k} = 1$.

(ii) Ряд $a_n = n^{-p}$ при $p > 1$ тоже охватывается:

$$2^k a_{2^k} = 2^k 2^{-pk} = (2^{1-p})^k = q^k,$$

где $q = 2^{1-p} < 1$. Случай $p < 1$ дает расходящийся ряд из-за сравнения с гармоническим, но тоже решается критерием.

(iii) Для ряда $a_n = n^{-1}(\ln(n+1))^{-1}$ получаем

$$2^k a_{2^k} = 2^k 2^{-k} (\ln(2^k + 1))^{-1} = (\ln(2^k + 1))^{-1}.$$

Так как $\ln(2^k+1) < \ln 2^{k+1} = (k+1)\ln 2$, а ряд из $(k+1)^{-1}$ расходится (в силу расходимости гармонического ряда), то расходится и наш ряд.

(iv) Для ряда $a_n = n^{-1}(\ln(n+1))^{-2}$ получаем

$$2^k a_{2^k} = 2^k 2^{-k}(\ln(2^k+1))^{-2} = (\ln(2^k+1))^{-2}.$$

Здесь $\ln(2^k+1) > \ln 2^k = k\ln 2$, что дает сходимость нашего ряда с учетом сходимости ряда из k^{-2} (см. (ii)). Таким образом, гармонический ряд расходился, уменьшение членов делением на $\ln(n+1)$ не помогло, но деление на $(\ln(n+1))^2$ привело к сходимости.

1.5. Азы топологии прямой

1.5.1. Определение. *Открытым называется множество на прямой, вместе со всякой своей точкой содержащее некоторый включающий ее интервал. Пустое множество тоже считается открытым.*

Множество замкнуто, если его дополнение открыто.

Иначе говоря, непустое открытое множество есть некоторое объединение интервалов. Обратно, всякое объединение интервалов открыто. Из определения следует, что объединение всякого семейства открытых множеств открыто. Легко видеть, что пересечение двух открытых множеств открыто (ибо это верно для интервалов). Значит, пересечение любого конечного числа открытых множеств открыто. В задаче 1.9.6 предлагается доказать, что всякий набор интервалов содержит не более чем счетный поднабор с тем же объединением.

С замкнутыми множествами наоборот: любые их пересечения замкнуты и замкнуты конечные объединения.

Полуинтервал $[0,1)$ не открыт и не замкнут, т. е. быть открытым или замкнутым — не альтернатива.

Прямая и пустое множество одновременно открыты и замкнуты. Других одновременно открытых и замкнутых множеств на прямой нет. В самом деле, пусть A открыто и замкнуто, причем непусто и отлично от прямой. Таково же и его дополнение. Значит, можно взять точки $a \in A$ и $b \notin A$. Можно считать, что $a < b$. Рассмотрим множество C таких точек $x \in A$, что $x < b$. Оно непусто и ограничено сверху, значит, имеет точную верхнюю грань c. Если $c \in A$, то приходим к противоречию из-за того, что ввиду открытости A справа от c найдутся точки из A, причем $c < b$ (ибо $b \notin A$). Если же $c \notin A$, то тоже нехорошо: теперь уже из-за открытости дополнения найдется такое $r > 0$, что $(c-r, c]$ не пересекается с A, т. е. c вовсе не точная верхняя грань.

1.5.2. Лемма. *Замкнутое множество в отрезке содержит свои точную нижнюю и точную верхнюю грани.*

Доказательство. Если точная верхняя грань v данного замкнутого множества A в него не входит, то она лежит в открытом дополнении. Значит, при некотором $r > 0$ в дополнении лежит $(v - r, v + r)$, но тогда v не является точной верхней гранью. Случай нижней грани аналогичен. □

1.5.3. Теорема. *Если замкнутое подмножество отрезка покрыто бесконечным набором интервалов (т. е. объединение этих интервалов содержит данное множество), то из них можно выбрать конечное число, объединение которых покрывает это подмножество.*

Доказательство. Пусть замкнутое множество A в $[a, b]$ покрыто интервалами J_α. По определению оно получено удалением из прямой некоторого набора интервалов U_β. Вместе интервалы J_α и U_β покрывают $[a, b]$ (они вообще $\mathbb{R}$ покрывают), поэтому можно найти конечный поднабор $J_{\alpha_1}, \ldots, J_{\alpha_n}$, $U_{\beta_1}, \ldots, U_{\beta_m}$, покрывающий $[a, b]$. При этом отдельно интервалы $J_{\alpha_1}, \ldots, J_{\alpha_n}$ покрывают A, поскольку U_β лежат в дополнении A и не могут помочь в покрытии A. □

1.5.4. Следствие. *Если дан такой набор замкнутых ограниченных множеств, что всякий его конечный поднабор имеет непустое пересечение, то все эти множества обладают общей точкой.*

Доказательство. Предположим, что пересечение всех данных замкнутых множеств F_α пусто. Тогда прямая является объединением их открытых дополнений $U_\alpha = \mathbb{R} \backslash F_\alpha$. Возьмем какое-либо F_{α_0} из наших замкнутых множеств. В частности, и оно покрыто данными множествами U_α. По предыдущему следствию найдется конечный набор $U_{\alpha_1}, \ldots, U_{\alpha_n}$, покрывающий F_{α_0}. По условию имеется точка, входящая во все $F_{\alpha_0}, \ldots, F_{\alpha_n}$. Ясно, что эта точка из F_{α_0} не может быть покрыта дополнениями $F_{\alpha_1}, \ldots, F_{\alpha_n}$. □

Установленное в теореме выше свойство ограниченных замкнутых множеств называют *компактностью*: множество K компактно, если из всякого его покрытия открытыми множествами можно выбрать конечное подпокрытие.

Например, интервал $(0, 1)$ некомпактен: из покрытия интервалами $(1/n, 1)$ нельзя выбрать конечное.

В этих терминах ограниченные замкнутые множества оказываются компактными. С другой стороны, всякое компактное множество на прямой ограничено и замкнуто. В самом деле, каждую точку можно покрыть интервалом и выбрать конечное подпокрытие, что даст ограниченное множество. Далее, если данное компактное множество

K незамкнуто, то имеется точка $a\ /K$, являющаяся пределом последовательности точек $a_n \in K$. Возьмем интервалы $(a + 1/n, n)$ и $(-n, a - 1/n)$. Заметим, что они покрывают K (ибо $K \subset \mathbb{R}\backslash\{a\}$). Никакой конечный их поднабор не покрывает K, так как вне него всегда есть интервал вокруг a, а в таком интервале найдутся точки из $\{a_n\}$ ввиду сходимости $a_n \to a$.

1.5.5. Предложение. *Всякое непустое открытое множество на прямой есть конечное или счетное объединение попарно непересекающихся открытых лучей или интервалов.*

Доказательство. Пусть U открыто и непусто. По определению это есть объединение некоторого набора интервалов J_α. Такие интервалы назовем эквивалентными, если они пересекаются. Получено отношение эквивалентности (проверьте!). Поэтому исходные интервалы разбиты на классы эквивалентности. Объединение всех интервалов, попавших в один класс, оказывается либо лучом, либо интервалом. В самом деле, пусть a и b попали в такое объединение W. Значит, a пришло с одним интервалом J исходного набора, b с каким-то интервалом I. При этом I и J пересекаются. Значит, $I \cup J$ есть интервал, лежащий в U, ибо $I, J \subset U$. Тем самым $(a, b) \subset W$. Два разных объединения, соответствующие разным классам эквивалентности, не могут пересекаться. Таким образом, U разбито на дизъюнктные лучи или интервалы. Их только конечное или счетное число, ибо в каждом есть рациональная точка, а одна точка не может попасть в разные интервалы из-за их дизъюнктности. □

1.5.6. Теорема. (Теорема Бэра) *Пусть* $\mathbb{R} = \bigcup_{n=1}^\infty M_n$, *где* M_n — *замкнутые множества. Тогда найдется* M_n, *содержащее интервал.*

Доказательство. Пусть ни одно M_n не содержит интервалов. Тогда открытое дополнение M_1 непусто, поэтому в нем есть отрезок I_1 ненулевой длины. Так как он не может полностью входить в M_2, то ввиду открытости дополнения M_2 найдется отрезок $I_2 \subset I_1$ ненулевой длины с $I_2 \cap M_2 = \varnothing$. По индукции строим вложенные отрезки I_n с $I_n \cap M_n = \varnothing$. Их пересечение содержит какую-то точку. Ясно, что она не входит ни в одно M_n — противоречие. □

Нетрудно заметить, что такое же рассуждение применимо к любому замкнутому множеству $F \subset \mathbb{R}$ вместо всей прямой и приводит к такому выводу: если $F = \bigcup_{n=1}^\infty F_n$, где все F_n замкнуты, то найдется такой интервал (a, b), что $F \cap (a, b) \subset F_n$. Разумеется, здесь уже нельзя утверждать, что F_n будет содержать интервал, если в самом F не было интервалов; последнее вполне может случиться: достаточно взять $F = \{1/n\} \cup \{0\}$.

1.5.7. Теорема. *Множество действительных чисел континуально, т. е. его мощность — континуум.*

Доказательство. Нам надо доказать равномощность $\mathbb{R}$ и множества последовательностей из 0 и 1. Проще всего сделать это для отрезка $[0,1]$, откуда будет вытекать и континуальность полуинтервала и всей прямой. Числа x из $[0,1]$ будем записывать в двоичной системе следующим образом. Точка попадает в одну из половин отрезка. Если $x \in [0,1/2)$, то положим $x_1 = 0$, иначе $x_1 = 1$. Затем x_2 полагаем равным 0 или 1 в зависимости от того, в левую или правую половину предыдущей половины попала точка x. Середина относится к правой половине. Таким образом, точке x сопоставлена последовательность $\{x_n\}$ из 0 и 1. Обратно, последовательности $\{y_n\}$ из 0 и 1 мы можем сопоставить точку y, полученную как пересечение вложенных отрезков I_n, заданных так: $I_1 = [0,1/2]$ при $y_1 = 0$, $I_1 = [1/2,1]$ при $y_1 = 1$, I_{n+1} есть левая половина I_n при $y_{n+1} = 0$ и правая половина I_n при $y_{n+1} = 1$. Для точек, не являющихся концами двоичных отрезков, построенной точке y будет соответствовать как раз $\{y_n\}$, но, скажем, точке $y = 1/2$ сопоставлена последовательность $(1,0,\ldots)$, хотя она получится и в качестве пересечения для последовательности $(0,1,1,\ldots)$, это связано с тем, что последовательность $(x_1,\ldots,x_n,1,1,1,\ldots)$, где $x_n = 0$, соответствует той же точке, что и $(x_1,\ldots,x_n+1,0,0,0,\ldots)$. Тем не менее за вычетом счетного множества соответствие взаимно однозначно, что доказывает равномощность. $\square$

Как видно, всякое непустое открытое множество тоже континуально. Для замкнутых множеств это, конечно, неверно, но для них верно следующее: всякое непустое замкнутое множество в $\mathbb{R}$ либо конечно, либо счетно, либо континуально (задача 2.5.30).

Одним из интереснейших замкнутых множеств является троичное множество Кантора, названное в честь выдающегося немецкого математика Георга Кантора (G. Cantor).

1.5.8. Пример. (Канторовское множество) Из отрезка $[0,1]$ удалим среднюю треть $J_1 = (1/3,2/3)$. Из двух оставшихся отрезков удалим средние трети $J_{2,1}$, $J_{2,2}$ и продолжим процесс индуктивно.

В итоге будут удалены на шаге n средние трети $J_{n,1}, \ldots, J_{n,2^n}$ оставшихся к тому моменту отрезков. Положим $C = [0,1] \setminus \bigcup_{n,k} J_{n,k}$. Осталось замкнутое множество. Оно непусто, ибо содержит концы $J_{n,k}$, но не только их.

В самом деле, будем записывать числа $x \in [0,1]$ в троичной системе $x = (x_n)$, где $x_n = 0, 1, 2$: если $x \in [0, 1/3]$, то $x_1 = 0$, если $x \in (1/3, 2/3)$, то $x_1 = 1$, если $x \in [2/3, 1]$, то $x_1 = 2$. Последующие компоненты $x_n \in \{0, 1, 2\}$ выбираются последовательно по тому же правилу в зависимости от того, в какой очередной трети оказывается x. За исключением концевых точек интервалов $J_{n,k}$ (т. е. за исключением счетного множества) возникает взаимно однозначное соответствие с последовательностями из $0, 1, 2$. Неоднозначность в указанных точках видна из того, что, скажем, точке $1/3$ при нашей системе соответствует $(0, 2, 2, \ldots)$, но при обратном сопоставлении она получится и из последовательности $(1, 0, 0, \ldots)$. Можно убедиться, что точки C оказываются во взаимно однозначном соответствии с последовательностями только из 0 и 2 (интересно, что для таких последовательностей неприятностей с неоднозначностью уже нет: одна точка не может кодироваться двумя разными последовательностями только из 0 и 2). Таким образом, множество Кантора континуально, хотя мы «видим» только счетную часть его типа точек $1/3$, $2/3$; сначала даже кажется, что больше вообще ничего нет.

Интересно еще, что в C нет изолированных точек. Действительно, пусть $x \in C$ соответствует последовательности $\{x_n\}$ из 0 и 2 и $r > 0$. Если среди x_n бесконечно много нулей, то заменив 0 на позиции с номером n таким, что $2^{-n} < r$, получим число $y \in C$ с $|x - y| < r$. Если $x_n = 2$ при $n \geqslant N$, то на далекой позиции заменим 2 на 0 с таким же эффектом.

1.6. Понятие метрического пространства

1.6.1. Определение. *Метрикой на множестве X называется функция $d\colon\ X \times X \to [0, +\infty)$, удовлетворяющая следующим условиям*:

1) $d(x, y) = 0$ *в точности тогда, когда* $x = y$,

2) $d(x, y) = d(y, x)$ *для всех* $x, y \in X$,

3) $d(x, z) \leqslant d(x, y) + d(y, z)$ *для всех* $x, y, z \in X$.

Множество с заданной на нем метрикой называют метрическим пространством.

Условие 3) называется *неравенством треугольника*. Метрическое пространство X с метрикой d обозначается через (X, d). На одном и том же множестве могут быть заданы разные метрики. Выражение «метрическое пространство X» без указания метрики подразумевает, что некоторая метрика выбрана. Если выполнены лишь условия 2) и 3), то d называют *предметрикой*.

Разные метрики на множестве могут иметь различный практический смысл. Скажем, в группе поселков, соединенных дорогами и тропинками, расстояние можно измерять как по прямой по карте, так и как кратчайшее расстояние по дорогам или же как кратчайшее расстояние и по дорогам, и по тропинкам. В общем случае такие метрики могут существенно отличаться, но во всех случаях расстояние от A до C не больше суммы расстояний от A до B и от B до C. Стандартная евклидова метрика на $\mathbb{R}^n$ имеет вид

$$d(x,y) = \bigl((x_1-y_1)^2 + \cdots + (x_n-y_n)^2\bigr)^{1/2}.$$

Эта метрика порождается нормой

$$\|x\| = \bigl(x_1^2 + \cdots + x_n^2\bigr)^{1/2}$$

по формуле

$$d(x,y) = \|x-y\|.$$

В свою очередь, норма порождается по формуле

$$\|x\| = \sqrt{(x,x)}$$

скалярным произведением

$$(x,y) = x_1y_1 + \cdots + x_ny_n.$$

Позднее будут обсуждаться общие нормы и скалярные произведения на линейных пространствах. Тот факт, что указанная выше метрика на $\mathbb{R}^n$ удовлетворяет неравенству треугольника, равносилен тому, что для соответствующей нормы выполнено неравенство треугольника

$$\|x+y\| \leqslant \|x\| + \|y\|.$$

Для проверки возводим обе части в квадрат и пользуемся равенством $(x+y,x+y) = (x,x) + (y,y) + 2(x,y)$. После упрощения приходим к равносильному неравенству Коши

$$(x,y)^2 \leqslant (x,x)(y,y).$$

Его можно проверить разными способами, один из которых таков: неравенство

$$(x+ty,x+ty) \geqslant 0$$

говорит, что дискриминант трехчлена $t^2(y,y) + 2t(x,t) + (x,x)$ неположителен.

Всякое подмножество метрического пространства само является метрическим пространством относительно метрики объемлющего пространства.

1.6.2. Определение. *Последовательность точек x_n в X сходится к точке x по метрике d, если для всякого $\varepsilon > 0$ найдется такой номер n_ε, что $d(x_n, x) < \varepsilon$ для всех $n \geqslant n_\varepsilon$. Тогда x называется пределом $\{x_n\}$ и обозначается через $\lim\limits_{n\to\infty} x_n$.*

Последовательность $\{x_n\} \subset X$ называется фундаментальной или последовательностью Коши по метрике d, если для всякого $\varepsilon > 0$ найдется такой номер n_ε, что $d(x_n, x_k) < \varepsilon$ для всех $n, k \geqslant n_\varepsilon$.

Сходящаяся последовательность фундаментальна, так как $d(x_n, x_k) \leqslant d(x_n, x) + d(x, x_k)$. Ясно, что предел может быть только один (ввиду неравенства треугольника и положительности расстояния между различными точками). Фундаментальная последовательность ограничена, ибо объединение шара и конечного множества ограничено.

1.6.3. Пример. (i) Любое множество X становится метрическим пространством с метрикой $d(x, y) = 1$ при $x \neq y$, $d(x, x) = 0$.

(ii) Множество всех ограниченных функций $B(\Omega)$ на непустом множестве Ω является метрическим пространством с метрикой

$$d(f, g) = \sup_{x\in\Omega} |f(x) - g(x)|. \tag{1.6.1}$$

Частным случаем является пространство l^∞ всех ограниченных последовательностей на $\mathbb{N}$ с метрикой

$$d(x, y) = \sup_n |x_n - y_n|, \quad x = (x_n),\ y = (y_n).$$

Сходимость функций в $B(\Omega)$ — так называемая равномерная сходимость, более подробно обсуждаемая в следующем семестре, где она возникает по делу.

(iii) Множество $\mathbb{R}^\infty$ всех бесконечных вещественных последовательностей $x = (x_n)$ можно наделить метрикой

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \min(|x_n - y_n|, 1). \tag{1.6.2}$$

Сходимость по этой метрике есть просто покоординатная сходимость.

1.6.4. Определение. *Открытым шаром радиуса $r > 0$ с центром в точке x_0 называется множество*

$$B(x_0, r) := \{x \in X\colon\ d(x, x_0) < r\}.$$

Замкнутый шар радиуса $r > 0$ с центром в x_0 есть множество

$$\{x \in X\colon\ d(x, x_0) \leqslant r\}.$$

Множество A называется *ограниченным*, если оно содержится в каком-либо шаре. Его *диаметр* есть $\operatorname{diam} A := \sup\{d(x, y)\colon\ x, y \in A\}$.

Как и для множеств на прямой, множество называют открытым, если вместе со всякой своей точкой оно содержит некоторый открытый шар с центром в этой точке. Пустое множество тоже считается открытым. Замкнутые множества — дополнения открытых. Непосредственно проверяется, что любое объединение открытых множеств открыто. Поэтому любое пересечение замкнутых множеств замкнуто.

Замыканием множества называется пересечение всех содержащих его замкнутых множеств (т.,е. наименьшее содержащее его замкнутое множество).

Нетрудно проверить, что множество замкнуто в точности тогда, когда оно содержит пределы всех сходящихся последовательностей его элементов.

1.6.5. Определение. *Изометрией двух метрических пространств (M_1, d_1) и (M_2, d_2) называется такое взаимно однозначное отображение J пространства M_1 на M_2, что*

$$d_2\bigl(J(x), J(y)\bigr) = d_1(x, y) \quad \text{для всех } x, y \in M_1.$$

1.6.6. Определение. *Метрическое пространство называется полным, если всякая фундаментальная последовательность в нем сходится.*

1.6.7. Пример. (i) Полуинтервал $[0, 1)$ с обычной метрикой не является полным пространством, так как последовательность $\{1 - 1/n\}$ фундаментальна, но не сходится к точке из $[0, 1)$. Однако этот же интервал оказывается полным метрическим пространством с метрикой $d(x, y) = |\operatorname{tg}(\pi x/2) - \operatorname{tg}(\pi y/2)|$.

(ii) Пространство $B(\Omega)$ с метрикой (1.6.1) полно. В самом деле, пусть $\{f_n\}$ — фундаментальная последовательность. Ясно, что для каждого фиксированного x числовая последовательность $\{f_n(x)\}$ фундаментальна и потому сходится. Ее предел мы обозначим через $f(x)$. Ограниченность фундаментальной последовательности $\{f_n\}$ показывает, что функция f ограничена. Остается заметить, что $\{f_n\}$ сходится к f по метрике $B(\Omega)$, а не только поточечно. Для этого при фиксированном $\varepsilon > 0$ находим такое число n_ε, что $|f_n(x) - f_k(x)| < \varepsilon$ для всех $n, k \geqslant n_\varepsilon$ и всех $x \in \Omega$. Если теперь зафиксировать номер $n \geqslant n_\varepsilon$ и устремить k к бесконечности, то получим оценку $|f_n(x) - f(x)| \leqslant \varepsilon$ для всех $x \in \Omega$, т. е. справедливо неравенство $d(f_n, f) \leqslant \varepsilon$.

(iii) Замкнутое множество в полном пространстве является полным пространством. Напомним, что все пространство всегда замкнуто, даже если оно неполно.

Многие свойства метрических пространств инвариантны относительно изометрий. К таким свойствам относятся сепарабельность, ограниченность и вводимая ниже полнота. Замечательной особенностью пространств ограниченных функций $B(\Omega)$, на первый взгляд весьма специальных, является то, что с точностью до изометрий они содержат вообще все метрические пространства.

1.6.8. Теорема. *Всякое метрическое пространство* (M, d) *изометрично части пространства* $B(M)$.

Доказательство. Пусть M непусто. Зафиксируем $x_0 \in M$ и для каждой точки $x \in M$ зададим функцию f_x на M с помощью равенства $f_x(y) = d(y, x) - d(y, x_0)$. Заметим, что в силу неравенства треугольника $|f_x(y)| \leqslant d(x, x_0)$, т. е. функция f_x ограничена. Для фиксированных $x_1, x_2 \in M$ в силу неравенства треугольника имеем

$$|f_{x_1}(y) - f_{x_2}(y)| = |d(y, x_1) - d(y, x_2)| \leqslant d(x_1, x_2),$$

что дает $d(f_{x_1}, f_{x_2}) \leqslant d(x_1, x_2)$. Поэтому $d(f_{x_1}, f_{x_2}) = d(x_1, x_2)$ ввиду равенства $|f_{x_1}(x_1) - f_{x_2}(x_1)| = d(x_1, x_2)$. Итак, построенное отображение $J\colon\ x \mapsto f_x$ из M в $B(M)$ является изометрией M и $J(M)$. $\square$

Множество A называется всюду плотным в метрическом пространстве M, если в каждом шаре положительного радиуса есть точка из A.

1.6.9. Определение. *Пополнением метрического пространства* M *называется такое полное метрическое пространство* $\widetilde{M}$, *что* M *изометрично всюду плотному множеству в* $\widetilde{M}$.

1.6.10. Предложение. *Всякое метрическое пространство имеет единственное с точностью до изометрии пополнение.*

Доказательство. Дадим даже два доказательства. В силу теоремы 1.6.8 можно считать, что M лежит в пространстве $B(\Omega)$ с $\Omega = M$. Так как $B(M)$ полно, то искомым пополнением является замыкание множества M в $B(\Omega)$. Остается проверить, что если M' — некоторое пополнение M, то $\overline{M}$ и M' изометричны. Это мы оставим в качестве задачи. $\square$

1.7. Вещественные числа как пополнение рациональных

Предыдущая конструкция пополнения на самом деле не вполне подходит к построению модели вещественных чисел, ибо уже их использует при определении метрики (правда, это можно обойти, определяя и метрику через фундаментальные последовательности). Важнее то, что в приведенных аксиомах никакой метрики нет, а полнота вводится через порядок. Однако можно так модифицировать построение пополнения, что ссылок на $\mathbb{R}$ и метрику уже не будет. Для этого просто не будем вводить никаких метрик на $\mathbb{Q}$, а фундаментальность последовательности рациональных чисел r_n определим так: для всякого $k \in \mathbb{N}$ найдется такой номер N_k, что

$$-k^{-1} < r_n - r_m < k^{-1} \quad \text{при всех } n, m \geqslant N_k.$$

Как видим, используется только порядок. Две последовательности рациональных чисел $\{r_n\}$ и $\{s_n\}$ назовем эквивалентными, если последовательность $r1, s_1, r_2, s_2, \ldots$ фундаментальна. Непосредственно проверяется, что получено отношение эквивалентности. В качестве пополнения возьмем множество классов эквивалентности фундаментальных последовательностей, но и на нем пока не будем вводить метрику (ее нет в аксиомах), а введем операции и порядок. С операциями все тривиально: сумма и произведение представителей $\{r_n\}$ и $\{p_n\}$ двух классов эквивалентности задаются представителями $\{r_n+p_n\}$ и $\{r_np_n\}$ соответственно, проверяется независимость от выбора представителей, легко проверяются все нужные соотношения.

Как ввести порядок? Это чуть менее банально: нельзя сказать, что $x = \{x_n\} \leqslant 0$, если $x_n \leqslant 0$, ведь $x = \{n^{-1}\} = 0$. Выход такой: $x = \{p_n\} \leqslant y = \{q_n\}$, если для каждого $k \in \mathbb{N}$, найдется такой номер N_k, что $p_n \leqslant q_n + k^{-1}$ при всех $n \geqslant N_k$. Это же можно еще и так сформулировать: $x = \{p_n\} < y = \{q_n\}$, если найдется такой номер N, что $p_n < q_n$ при всех $n \geqslant N$, причем $\{p_n\}$ и $\{q_n\}$ не являются эквивалентными (но последнее условие нельзя убрать!). Как обычно, следует проверить еще независимость от представителей, что несложно, а также, что получен порядок, в частности, что все элементы сравнимы! Проверим лишь последнее. Пусть даны две неэквивалентные фундаментальные последовательности $x = \{p_n\}$ и $y = \{q_n\}$. Их неэквивалентность означает, что при некотором $k \in \mathbb{N}$ найдется бесконечно много номеров n с $|x_n - y_n| > k^{-1}$; можно считать, что есть бесконечно много номеров с $x_n < y_n - k^{-1}$. В силу фундаментальности с некоторого номера N_k верны неравенства $|x_n - x_m| < (4k)^{-1}$, $|y_n - y_m| < (4k)^{-1}$ при всех $n, m \geqslant N_k$, но тогда для таких номеров мы получаем $x_n < y_n - (4k)^{-1}$, поскольку можно взять далекий номер $m > n$ с $x_m < y_m - k^{-1}$ и оценить x_n так:

$$\begin{aligned} x_n = x_n - x_m + x_m < x_n - x_m + y_m - k^{-1} = \\ = x_n - x_m + y_n + y_m - y_n - k^{-1} \leqslant \\ \leqslant y_n + (4k)^{-1} + (4k)^{-1} - k^{-1}. \end{aligned}$$

Столь же несложно проверить нужное согласование с операциями. Разумеется, в качестве нуля надо взять класс с представителем $(0, 0, \ldots)$, а в качестве единицы — класс с представителем $(1, 1, \ldots)$/

Остается проверка аксиомы полноты. Пусть $A \leqslant B$ для двух множеств нашего пространства (т. е. наборов представителей классов эквивалентности). Ищем разделяющую их точку, т. е. фундаментальную последовательность рациональных чисел r_n, для которой $A \leqslant \{r_n\} \leqslant B$. Здесь становится очень важно, что наш порядок не требует покоординатного сравнения (это было бы катастрофой, ибо не всегда удалось бы находить рациональные разделители). На самом деле мы найдем «вещественного разделителя» (пока не существующего), но заменим его приближающей рациональной последовательностью. Нам дано следующее условие: $\{a_n\} \leqslant \{b_n\}$ для всех $\{a_n\} \in A, \{b_n\} \in B$ в смысле несколько странного на первый взгляд порядка, введенного выше.

Докажем вспомогательное утверждение: для данного ограниченного сверху (рациональным числом Q) множества S рациональных чисел и фиксированного n найдется такое рациональное число r (приближенная точная верхняя грань), что $s \leqslant r$ при всех $s \in S$, но имеется также и точка $s_n \in S$, для которой $s_n \geqslant r - 2^{-n}$. Это делается так. Если найдется число $s_1 \in S$ с $s_1 > Q - 2^{-n}$, то годится $r = Q$. Если такого s_1 нет, то берем любое $s_1 \in A$, делим отрезок рациональных чисел с концами в s_1 и Q на конечное число подотрезков с рациональными концами $t_1, \ldots, t_m$ длины менее 2^{-n} и берем самый правый, содержащий точки из S. Его правый конец годится в качестве r.

Будем строить разделитель A и B. Воспользуемся тем, что представителей можно менять на эквивалентных, а также тем, что во всякой фундаментальной последовательности $\{a_n\}$ есть подпоследовательность с $|a_{n_k} - a_{n_{k+1}}| \leqslant 4^{-k}$. Поскольку подпоследовательность фундаментальной последовательности эквивалентна ей, мы можем произвести замену всех представителей в A и B так, что будет выполнено указанное выше неравенство. Тем самым в качестве представителей теперь используем такие последовательности, что при $m \geqslant n$ верны оценки $|a_n - a_m| \leqslant 2^{-n}$, $|b_n - b_m| \leqslant 2^{-n}$.

Поскольку в B есть хоть один элемент $b^0 = \{b_n^0\}$ и $A \leqslant b_0$, то наш трюк приводит к тому, что все компоненты всех элементов A оказываются ограниченными числом $|b_1^0| + 2$. При произвольном выборе представителей ничего такого не будет, ведь можно как угодно менять начальные куски последовательностей. Напоминаем, что мы все еще находимся среди рациональных чисел, никаких других пока нет. Теперь в качестве разделителя хотелось бы взять последовательность $\{s_n\}$, где s_n — точная верхняя грань множества A_n n-х компонент представителей элементов из A, выбранных указанным выше образом. Но здесь сразу две проблемы: никаких точных верхних граней пока нет, а если бы и были, то могли бы оказаться иррациональными. Эти проблемы обходятся взятием приближенных рациональных точных верхних граней A_n (как объяснено выше), т. е. чисел $r_n \in \mathbb{Q}$ с $A_n \leqslant r_n$, для которых в A_n есть числа α_n с $\alpha_n > r_n - 2^{-n}$.

Покажем, что $\{r_n\}$ разделяет A и B. Пусть $\{a_n\} \in A, \{b_n\} \in B$. По построению $a_n \leqslant r_n$. Если бы оказалось, что $\{b_n\} < \{r_n\}$, то при некотором $k \in \mathbb{N}$ мы бы имели $b_n < r_n - k^{-1}$ при всех $n \geqslant N_1$ для некоторого N_1. Можно считать, что $5 \cdot 2^{-N_1} < (4k)^{-1}$.

В A_{N_1} есть число, большее $r_{N_1} - 2^{-N_1}$. Это означает, что в A есть элемент $\{a_n\}$ с $a_{N_1} > r_{N_1} - 2^{-N_1}$. Для номеров $n > N_1$ имеем

$$\begin{aligned} a_n \geqslant a_{N_1} - 2^{1-N_1} > r_{N_1} - 3 \cdot 2^{-N_1} > b_{N_1} + k^{-1} - 3 \cdot 2^{-N_1} \geqslant \\ \geqslant b_n - 2^{1-N_1} + k^{-1} - 3 \cdot 2^{-N_1} > b_n + (4k)^{-1}, \end{aligned}$$

т. е. $\{a_n\} > \{b_n\}$, что невозможно.

1.8. Комплексные и p-адические числа

Выше были упомянуты три вида полноты поля: порядковая, метрическая, алгебраическая. Первая связана с порядком и состоит в существовании точных верхних граней ограниченных сверху множеств, вторая связана с метрикой и состоит в существовании пределов фундаментальных последовательностей, третья состоит в разрешимости полиномиальных уравнений. Тем самым каждая обеспечивает отсутствие каких-то «дыр»: порядковых, метрических, алгебраических. Поле $\mathbb{R}$ обладает первыми двумя видами полноты, но не третьим. Здесь кратко упомянуты еще два вида полей, связанных с рациональными числами и имеющих какие-то виды полноты.

Множество комплексных чисел $\mathbb{C}$ состоит из пар (x, y) вещественных чисел, символически записываемых в виде $z = x + iy$, где x называют вещественной частью числа и обозначают через $\operatorname{Re} z$, а y называют мнимой частью числа и обозначают через $\operatorname{Im} z$. Сложение определяется просто:

$$x + iy + u + iv = (x + u) + i(y + v),$$

а умножение определяется посредством соглашения $i \cdot i = -1$. Таким образом,

$$(x + iy) \cdot (u + iv) = xu - yv + i(xv + yu).$$

Эта немного странное на первый взгляд определение оказывается весьма разумным и дает гораздо более содержательную систему чисел, чем более «симметричное» умножение, при котором надо было бы отдельно перемножать вещественные и мнимые части. Модуль (норма) комплексного числа z задается формулой

$$|z| = \sqrt{x^2 + y^2}.$$

Если сопряженное число $\bar{z}$ задать формулой $\bar{z} = x - iy$, то получим

$$|z|^2 = z\bar{z}.$$

Несложно проверяется, что $\mathbb{C}$ полно с метрикой, порожденной указанной нормой по формуле $d(z_1, z_2) = |z_1 - z_2|$. В терминах вещественных и мнимых частей чисел $z_1 = x_1 + iy_1$ и $z_2 = x_2 + iy_2$ это просто обычное расстояние на плоскости: $d(z_1, z_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$.

Множество $\mathbb{C}$ оказывается полем, а также двумерным линейным пространством над $\mathbb{R}$ и одномерным над самим собой. Можно сказать, что это поле получено расширением поля $\mathbb{R}$ добавлением одного числа i (мнимой единицы). Это число делает разрешимым уравнение $z^2 = -1$, не имевшее корней в $\mathbb{R}$, но чудесным образом оказывается, что поле $\mathbb{C}$ алгебраически полно, т. е. в нем разрешимы все полиномиальные уравнения $z^n + c_{n-1}z^{n-1} + \dots + c_1 z + c_0 = 0$ любых степеней. Однако в $\mathbb{C}$ нет порядка, в котором порядковые интервалы порождают имеющуюся топологию (систему открытых множеств). Конечно, порядок ввести можно, скажем, лексикографический, в котором $z_1 < z_2$, если либо $\operatorname{Re} z_1 < \operatorname{Re} z_2$, либо $\operatorname{Re} z_1 = \operatorname{Re} z_2$, но $\operatorname{Im} z_1 < \operatorname{Im} z_2$. Однако этот порядок не имеет той связи с топологией, что была на $\mathbb{R}$.

Отметим, что введенная метрика на $\mathbb{C}$ позволяет говорить о сходимости рядов с комплексными членами: ряд $\sum_{n=1}^{\infty} z_n$ сходится к числу S, если его частичные суммы S_n стремятся к S, т. е. $|S_n - S| \to 0$.

Нетрудно проверить, что это равносильно тому, что ряд из вещественных частей z_n сходится к числу $\operatorname{Re} S$, а ряд из мнимых частей сходится к числу $\operatorname{Im} S$.

Поле комплексных чисел наделяется еще одним полезным объектом: скалярным произведением, которое упорядоченной паре комплексных чисел z_1, z_2 сопоставляет число $z_1\overline{z_2}$. По своим свойствам этот объект близок школьному скалярному произведению векторов на плоскости (если $\mathbb{C}$ представлять как $\mathbb{R}^2$), но не совпадает с ним. Скажем, скалярное произведение комплексных чисел 1 и i равно $-i$, хотя соответствующие этим числа векторы плоскости ортогональны.

Всякое комплексное число z с $|z| = 1$ можно записать в виде $z = \cos\theta + i\sin\theta$, где $\theta \in [0, 2\pi)$ называют аргументом z. Пишут также

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Пока эта запись (формула Эйлера) выглядит как символическое обозначение (даже и синусов с косинусами пока нет), но позже, при обсуждении разложений в ряды, оно станет верным тождеством. Из определения умножения комплексных чисел следует, что $z^n = e^{in\theta}$. Более общим образом, если $z_1 = |z_1|e^{i\theta_1}$, $z_2 = |z_2|e^{i\theta_2}$, то

$$z_1 z_2 = |z_1||z_2|e^{i(\theta_1+\theta_2)},$$

т. е. модули перемножаются, аргументы складываются.

Поле вещественных чисел получает еще серию конкурентов, если обратиться к описанной метрической процедуре пополнения $\mathbb{Q}$. Еще раз посмотрим на использованную метрику: она получена по формуле $d(x, y) = |x - y|$ из модуля. Нормированием поля $\mathbb{K}$ называют отображение $\|\cdot\|\colon \mathbb{K} \to [0, +\infty)$ с такими свойствами: $\|x\| = 0$ лишь при $x = 0$, $\|x + y\| \leqslant \|x\| + \|y\|$, $\|xy\| = \|x\|\,\|y\|$. Нормирование порождает метрику по формуле $d(x, y) = \|x - y\|$. Имеется тривиальное нормирование, при котором $\|x\| = 1$ при всех $x \neq 0$. При обсуждении нормирований считается, что $\mathbb{R}$ уже есть.

На множестве $\mathbb{Q}$ можно ввести необозримое число метрик, пополнения по которым дадут с точностью до изометрий все сепарабельные метрические пространства, но оказывается, что лишь очень немногие из них связаны с нормированиями.

Приведем нетривиальный пример нормирования $\mathbb{Q}$, приводящий к p-адическим числам. Пусть p — простое число. Для каждого ненулевого целого числа m можно найти наибольшую возможную степень

$p(m)$ числа p, на которую m делится нацело. Для несократимой рациональной дроби $r = m/n$ положим

$$\|r\|_p = p^{-d(r)}, \quad d(r) = p(m) - p(n).$$

Положим также $\|0\|_p = 0$. Можно проверить, что задано нормирование поля $\mathbb{Q}$ (конечно, оно использует вещественные числа!), соответствующая метрика обозначается через d_p. Пополнение $\mathbb{Q}$ по этой метрике и есть множество p-адических чисел, обозначаемое через $\mathbb{Q}_p$. Оно тоже поле, но не алгебраически полное. Его расширяют до алгебраически полного поля, но процедурой несравнимо более сложной, чем добавление мнимой единицы. Замечательная теорема Островского (А. Островский — швейцарский математик российского происхождения, не путать с писателем А. Н. Островским, если кто о таком слыхал) утверждает, что если метрика на $\mathbb{Q}$ получена из нетривиального нормирования, то множество соответствующих фундаментальных последовательностей совпадает с множеством таких последовательностей либо по обычной норме, либо по метрике d_p с некоторым простым числом p. Тем самым множество пополнений по таким метрикам — $\mathbb{R}$ и все $\mathbb{Q}_p$, см. книгу [**8**].

1.9. Задачи

1.9.1. Докажите, что простых чисел бесконечно много, рассуждая от противного. Заметим, что неизвестно, бесконечно ли множество простых чисел среди чисел вида $n^2 + 1$, где n натурально. Неизвестно также, бесконечно ли число пар простых чисел p, q с $p - q = 2$.

1.9.2. При каких $a \in \mathbb{R}$ ограничено множество $\{a^n/n! | \, n \in \mathbb{N}\}$?

1.9.3. (i) Доказать, что для замкнутых подмножеств прямой выполнена аксиома полноты в терминах порядка. (ii) Бывают ли незамкнутые множества, для которых выполнена аксиома полноты в терминах порядка?

1.9.4. Докажите, что уравнение $x^n = a$ при $n \in \mathbb{N}$ и $a > 0$ имеет положительный корень (его обозначают символом $a^{1/n}$).

1.9.5. Доказать, что объединение континуума континуальных множеств континуально.

1.9.6. Доказать, что из всякого набора интервалов на прямой можно выбрать конечный или счетный поднабор с тем же объединением.

1.9.7. Исследовать сходимость следующих последовательностей:
(i) $a_n = (-1)^n n^{-1}$,
(ii) $a_n = 1 + (1 + (-1)^n)/n$,
(iii) $a_n = (1 + n)^{1/n}$.

1.9.8. Пусть последовательность $\{a_n\}$ имеет предел a. Доказать, что такой же предел имеет последовательность средних арифметических $b_n = (a_1 + \cdots + a_n)/n$.

Верно ли, что из сходимости $\{b_n\}$ следует сходимость $\{a_n\}$?

1.9.9. Доказать, что всякая ограниченная последовательность содержит возрастающую или убывающую подпоследовательность.

1.9.10. Сходится ли последовательность $a_n = \sin n$?

1.9.11. Найти предел последовательности, заданной формулой

$$a_1 = 1, \quad a_{n+1} = \sqrt{1 + a_n}.$$

1.9.12. Сходится ли последовательность, заданная следующей формулой: $a_1 = 1, \quad a_{n+1} = \sin a_n$?

1.9.13. Всякое ли замкнутое подмножество прямой совпадает с множеством частичных пределов какой-либо последовательности?

1.9.14. Докажите, что при $a > -1$ при всех $n \in \mathbb{N}$ верно неравенство Бернулли:

$$(1 + a)^n \geqslant 1 + na.$$

1.9.15. (i) Докажите, что последовательность $\left(1+\dfrac{1}{n}\right)^{n+1}$ монотонно убывает, а последовательность $\left(1 + \dfrac{1}{n}\right)^n$ возрастает.

(ii) Докажите, что последовательность $\left(1 + \dfrac{1}{n}\right)^n$ сходится. Предел обозначается латинской буквой e.

1.9.16. Докажите, что последовательность чисел

$$a_n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}$$

сходится к числу e.

1.9.17. Докажите, что число e иррационально (рассуждая от противного, оцените величину $|a_n - e|$ и покажите, что она не может быть рациональна). На самом деле число e (как и π) даже трансцендентно, т. е. не является корнем полиномиального уравнения с целыми коэффициентами, но доказывается это заметно сложнее (см. [**14**]). Отметим, что про число $e+\pi$ до сих пор неизвестно, является ли оно иррациональным.

1.9.18. Доказать, что
$$\lim_{n\to\infty} n(e^{1/n}-1) = \lim_{n\to\infty} n(1-e^{-1/n}) = 1.$$
Для этого использовать, что $(1+1/n)^n < e < (1+1/n)^{n+1}$ и заметить, что $n(e^{1/n}-1) > 1$, а также $n(e^{1/n}-1) < 1+1/k$ при $n > k$.

1.9.19. (i) Доказать, что
$$\Bigl(\frac{n}{e}\Bigr)^n < n! < \Bigl(\frac{n+2}{e}\Bigr)^{n+3}, \quad n \in \mathbb{N}.$$
Позже будет выведена более точная формула Стирлинга.

(ii) Доказать, что
$$\frac{1}{n+1} < \ln\Bigl(1+\frac{1}{n}\Bigr) < \frac{1}{n}, \quad n \in \mathbb{N}.$$

1.9.20. Найти пределы
$$\lim_{n\to\infty} \sin\Bigl(\pi\sqrt{n^2+1}\Bigr), \quad \lim_{n\to\infty} \Bigl|\sin\Bigl(\pi\sqrt{n^2+n}\Bigr)\Bigr|^2.$$

1.9.21. Доказать, что если последовательность $\{a_n\}$ ограничена и $\lim\limits_{n\to\infty}(a_{n+1}-a_n) = 0$, то ее частичные пределы заполняют отрезок от нижнего до верхнего пределов.

1.9.22. Пусть последовательность $\{a_n\}$ такова, что при всех n, k верно неравенство $0 \leqslant a_{n+k} \leqslant a_n + a_k$. Доказать, что последовательность $\{a_n/n\}$ сходится.

1.9.23. Пусть $a_n > 0$ и последовательность $\{a_{n+1}/a_n\}$ сходится. Доказать, что такой же предел имеет последовательность $\{a_n^{1/n}\}$.

1.9.24. Доказать, что $\lim\limits_{n\to\infty} \dfrac{n}{(n!)^{1/n}} = e$.

1.9.25. Доказать теорему Штольца: если последовательности $\{x_n\}$ и $\{y_n\}$ таковы, что $y_{n+1} > y_n$, $y_n \to \infty$, $\dfrac{x_{n+1}-x_n}{y_{n+1}-y_n} \to A$, то $\dfrac{x_n}{y_n} \to A$.

1.9.26. Исследовать сходимость рядов со следующим общим членом:

(i) $a_n = (-1)^{[\sqrt{n}]}/n$,

(ii) $a_n = (-1)^{n^2}/n$,

(iii) $a_n = 1 - e^{-1/n}$,
(iv) $a_n = 1 - e^{-1/n^2}$,
(v) $a_{2n-1} = b_n$, $a_{2n} = c_n$, где ряды из b_n и c_n сходятся,
(vi) $a_n = (-1)^n \sin(1/n)$.

1.9.27. Дан знакочередующийся ряд $\sum_{n=1}^{\infty}(-1)^n a_n$ с монотонно убывающими к нулю числами $a_n > 0$, который сходится по признаку Лейбница. Всегда ли сходится ряд $\sum_{n=1}^{\infty}(-1)^n b_n$, в котором $\lim\limits_{n\to\infty} b_n/a_n = 1$?

1.9.28. Числа Фибоначчи задаются формулой $f_n = f_{n-1} + f_{n-2}$ при $n > 2$, $f_1 = f_2 = 1$. Докажите, что существует предел $\lim\limits_{n\to\infty} \dfrac{f_{n+1}}{f_n}$ и найдите его.

1.9.29. Докажите, что для всякого вещественного числа x существует бесконечно много пар целых чисел p, q, таких что

$$\left|x - \frac{p}{q}\right| \leqslant \frac{1}{q^2}.$$

1.9.30. Применить предыдущую задачу для исследования сходимости последовательности $a_n = (1 + \ln n)^{-1/2} \ln|\sin n|$.

1.9.31. Пусть s_0 — целое число и $s_1, s_2, \ldots$ — некоторая последовательность натуральных чисел. Соответствующая этой последовательности *цепная дробь* записывается символически как

$$s_0 + \cfrac{1}{s_1 + \cfrac{1}{s_2 + \cdots}},$$

что определяется как предел последовательности $\{a_n\}$, получаемой, если предыдущую формулу оборвать на s_n. Например,

$$a_3 = s_0 + \cfrac{1}{s_1 + \cfrac{1}{s_2 + \cfrac{1}{s_3}}}.$$

Пусть $a_n = p_n/q_n$ — несократимая дробь.

(i) Докажите, что $q_n p_{n-1} - p_n q_{n-1} = (-1)^n$.

(ii) Докажите, что дроби p_{2n}/q_{2n} возрастают, дроби p_{2n+1}/q_{2n+1} убывают, $p_{2n}/q_{2n} \leqslant p_{2k+1}/q_{2k+1}$ при всех n, k.

(iii) Докажите, что $\{a_n\}$ имеет конечный предел, причем для этого предела $\alpha = [s_0; s_1, s_2, \ldots]$ верно неравенство $|\alpha - p_n/q_n| < (q_n q_{n+1})^{-1}$.

(iv) Докажите, что всякое вещественное число α представимо в виде цепной дроби, причем для иррационального числа α числа s_n задаются формулами $s_n = [\alpha_n]$, $\alpha_0 = \alpha$, $\alpha_{n+1} = (\alpha_n - s_n)^{-1}$.

(v) Найдите число, заданное цепной дробью $[1; 1, 1, \ldots]$.

(vi) Докажите, что $(1+\sqrt{17})/2 = [2; 1, 1, 3, 1, 1, 3, 1, 1, 3, \ldots]$.

Указание: см. [**1**], [**14**], [**21**], где есть интересная дополнительная информация о цепных дробях и их применениях.

1.9.32. Доказать, что для всякой последовательности чисел $a_n > 0$ верно неравенство

$$\varlimsup_{n\to\infty} \left(\frac{a_1 + a_{n+1}}{a_n}\right)^n \geqslant e.$$

1.9.33. Можно ли прямую представить в виде объединения попарно непересекающихся отрезков с положительными длинами?

1.9.34. Доказать, что если из замкнутого множества F в $[0,1]$ удалить все его изолированные точки, то останется замкнутое множество. Оно называется производным множеством F и обозначается через F'.

Индуктивно определяются множества $F^{(n)} = (F^{(n-1)})'$. Может ли случиться, что они все различны?

Глава 2

Непрерывность

В этой главе изложены базовые сведения о непрерывности функций и пределах функций. Подход основан на использовании сходимости последовательностей.

2.1. Непрерывные функции

Непрерывность функции в точке означает, что при стремлении аргумента к этой точке значения функции стремятся к значению в ней. Точный смысл этому можно придать в терминах уже знакомой сходимости последовательностей, а также на более топологическом языке окрестностей.

2.1.1. Определение. *Функция f, заданная на множестве $E \subset \mathbb{R}$, называется непрерывной в точке $x_0 \in E$, если для каждой последовательности точек $x_n \in E$, сходящейся к x_0, мы имеем*

$$f(x_0) = \lim_{n\to\infty} f(x_n).$$

Непрерывной называется функция, непрерывная в каждой точке области определения.

Точно так же определяется и непрерывность функций на метрических пространствах. Такое определение непрерывности называют непрерывностью по Гейне, а ниже приведено равносильное ему определение непрерывности по Коши ((ε, δ)-определение).

В случае общих топологических пространств именно его подходящее обобщение обычно берется в качестве первичного определения (хотя для определения Гейне тоже есть подходящее равносильное обобщение через сходимость направленностей).

Простейшие примеры непрерывных функций: $f(x) = c$ — постоянная и $f(x) = x$.

2.1.2. Теорема. *Непрерывность функции f на множестве E в точке $x_0 \in E$ равносильна так называемому (ε, δ)-определению: для каждого $\varepsilon > 0$ найдется такое $\delta = \delta_\varepsilon > 0$, что $|f(x) - f(x_0)| < \varepsilon$ для всех таких $x \in E$, что $|x - x_0| < \delta$.*

Доказательство. Пусть есть непрерывность в x_0 и $\varepsilon > 0$. Если нужного δ нет, но не годится никакое $\delta = 1/n$. Это означает следующее: найдется точка $x_n \in E$, для которой $|x - x_0| < 1/n$, но $|f(x_n) - f(x_0)| \geqslant 1/n$. Тогда $x_n \to x_0$, но $f(x_n) \not\to f(x_0)$, что противоречит непрерывности. Значит, на самом деле в качестве δ можно взять некоторое $1/n$.

Пусть выполнено (ε, δ)-условие, $x_n \in E$ и $x_n \to x_0$. Надо проверить, что $f(x_n) \to f(x_0)$. Пусть $\varepsilon > 0$. Возьмем то $\delta > 0$, для которого при $x \in E \cap (x_0 - \delta, x_0 + \delta)$ мы имеем $|f(x) - f(x_0)| < \varepsilon$. Остается вспомнить, что по определению найдется такое N, что $x_n \in (x_0 - \delta, x_0 + \delta)$ при всех $n \geqslant N$. □

Как и в определении сходимости последовательности, от нас не требуется явно указывать зависимость δ_ε от ε или брать какое-либо «оптимальное» подходящее. Важно лишь, чтоб какое-то было. Как мы увидим ниже, почти никогда проверка непрерывности не производится путем явного указания δ_ε. Однако есть один важный случай, когда его можно указать.

2.1.3. Пример. Функция f на множестве E удовлетворяет условию Липшица с постоянной L (и называется липшицевой), если

$$|f(x) - f(y)| \leqslant L|x - y|, \quad x, y \in E.$$

Такая функция непрерывна: для данного $\varepsilon > 0$ в качестве δ в (ε, δ)-определении можно взять $\delta_\varepsilon = \varepsilon/(L + 1)$ (если $L > 0$, то достаточно на L делить).

Важно иметь в виду, что непрерывность в точке никоим образом не обеспечивает непрерывность в близких точках. Например, функция Дирихле из задачи 2.5.1, разрывная в каждой точке, после умножения на аргумент x превращается в функцию, непрерывную в нуле, но во всех прочих точках такое произведение по-прежнему разрывно. Заметим, кстати, что любая ограниченная функция после умножения на x (или на иную функцию, непрерывную и равную нулю в нуле) становится непрерывной в нуле.

Заметим, что для изолированных точек вообще ничего не требуется, так что в изолированных точках множества E функция всегда непрерывна.

Из свойств пределов последовательностей сразу следуют такие свойства непрерывных функций.

2.1.4. Предложение. *Если функции f и g на E непрерывны в точке x_0, то таковы же $f+g\colon x\mapsto f(x)+g(x)$ и $fg\colon x\mapsto f(x)g(x)$.*

Если $g(x)\neq 0$, то непрерывна и функция $x\mapsto f(x)/g(x)$.

Например, для суммы имеем: если $x_n\to x_0$, то $f(x_n)\to f(x_0)$ и $g(x_n)\to g(x_0)$, значит, $f(x_n)+g(x_n)\to f(x_0)+g(x_0)$.

Таким, непрерывны многочлены и отношения многочленов вне нулей знаменателя.

2.1.5. Предложение. *Если функция f на E непрерывна в точке x_0, то такова же и композиция $\varphi(f)$ для всякой непрерывной функции φ на $\mathbb{R}$ (или на множестве значений f).*

Доказательство. Если $x_n\in E$ и $x_n\to x_0$, то $f(x_n)\to f(x_0)$, откуда $\varphi(f(x_n))\to\varphi(f(x_0))$. □

2.1.6. Пример. Функция $|x|$ очевидным образом непрерывна, поэтому модуль непрерывной функции непрерывен. Конечно, это легко видно и непосредственно из определения.

2.1.7. Теорема. *Образом отрезка при непрерывной функции является отрезок (который может быть точкой).*

Образом замкнутого ограниченного множества при непрерывной функции является ограниченное замкнутое множество. При этом функция принимает наибольшее и наименьшее значения.

Доказательство. Докажем сначала второе утверждение. Если бы нашлась последовательность таких точек x_n из данного ограниченного замкнутого множества E, что значения данной функции f в них стремились бы к $+\infty$ или $-\infty$, то можно было бы перейти к сходящейся подпоследовательности $x_{n_k}\to x_0$ и получить противоречие с тем, что по непрерывности $f(x_{n_k})\to f(x_0)$. Итак, множество значений функции f ограничено.

Покажем, что оно замкнуто. Если это неверно, то найдется число y, не входящее в множество значений, для которого $y=\lim\limits_{n\to\infty} f(x_n)$, $x_n\in E$. Функция $g(x)=(f(x)-y)^{-1}$ непрерывна по доказанному выше, но не является ограниченной — противоречие. Ограниченное замкнутое множество имеет наименьшую и наибольшую точки (лемма 1.5.2), которые и будут наименьшим и наибольшим значениями функции.

Есть и другое обоснование первого утверждения, основанное на понятии компактности (для этого надо вспомнить, что замкнутые ограниченные множества на прямой и есть компакты на прямой). Оно сразу дает компактность образа компакта при непрерывной функции: если образ покрыт открытыми множествами U_α, то исходное множество покрыто их прообразами $f^{-1}(U_\alpha)$, которые открыты из-за непрерывности f. Значит, из них можно выделить конечное подпокрытие, что даст и конечное подпокрытие для $\{U_\alpha\}$.

Обратимся к первому утверждению. Мы уже знаем, что образ данного отрезка $[a, b]$ — замкнутое множество в отрезке с концами в наименьшем и наибольшем значениях. Без ущерба для общности можно считать, что минимум достигается в a, максимум в b. Пусть какая-то точка y из $[f(a), f(b)]$ не входит в множество значений. Рассмотрим множество $A = \{x \in [a, b]\colon\ f(x) < y\}$. Тогда $a \in A$, $b \notin A$. Пусть $\alpha = \sup A$. Из непрерывности f следует, что $f(\alpha) \leqslant y$. Так как при этом $f(\alpha) \neq y$, то $f(\alpha) < y$. Значит, $\alpha < b$. В силу непрерывности точка α имеет правую окрестность, в которой значения тоже меньше y, что противоречит выбору α. $\square$

2.1.8. Пример. Из предыдущего утверждения очевидным образом вытекает, что непрерывная функция на отрезке принимает все промежуточные значения между всякими двумя своими значениями. Из этого можно извлекать разные полезные заключения. Например, если $f\colon\ [0, 1] \to [0, 1]$ непрерывна, то найдется такая точка x_0 («неподвижная точка»), что $f(x_0) = x_0$. В самом деле, это следует из доказанного, если функция $f(x) - x$ имеет минимум и максимум разных знаков, но одного знака они могут быть только при $f(0) = 0$ или $f(1) = 1$, ибо $f(0) \geqslant 0$ и $f(1) \leqslant 1$, так что невозможны неравенства $f(x) - x > 0$ или $f(x) - x < 0$ при всех x.

Близким к понятию непрерывности в точке является понятие предела функции в точке.

2.1.9. Определение. *Пусть дана функция f на множестве $E \subset \mathbb{R}$. Говорят, что в точке x_0 (необязательно лежащей в E) она имеет конечный предел L по множеству E, если для всякого $\varepsilon > 0$ найдется такое $\delta > 0$, что $|f(x) - f(x_0)| < \varepsilon$ при всех $x \in E \backslash \{x_0\}$ с $|x - x_0| < \delta$. Это равносильно тому, что после определения f в точке x_0 значением L она становится непрерывной в точке x_0 на множестве $E \cup \{x_0\}$.*

Обозначение предела в точке: $\lim\limits_{x \to x_0} f(x)$, *а если надо подчеркнуть участие множества E, то* $\lim\limits_{x \to x_0, x \in E \backslash \{x_0\}} f(x)$.

Конечно, это определение можно сформулировать и без ссылки на непрерывность или (ε, δ) с помощью последовательностей, но здесь возникает такой же нюанс, как и с (ε, δ): в последовательность нельзя включать точку x_0 (хотя бы потому, что в ней f не обязана быть изначально определенной): надо требовать, чтобы $f(x_n) \to L$ как только $x_n \in E$, $x_n \neq x_0$ и $x_n \to x_0$.

Совершенно аналогично вводится и бесконечный предел в точке. Например, будем говорить, что $\lim\limits_{x \to x_0} f(x) = +\infty$, если для всякого числа R найдется такое $\delta > 0$, что $f(x) > R$ при всех $x \in E \cap (x_0 - \delta, x_0 + \delta)$.

Наконец, бывает полезно рассмотреть и пределы в бесконечности. Для функции f на луче вида $(a, +\infty)$ Будем говорить, что

$$\lim_{x \to +\infty} f(x) = L,$$

где L — число, если для всякого $\varepsilon > 0$ найдется такое $R > a$, что $|f(x) - L| < \varepsilon$ при всех $x > R$. Аналогично определяется $\lim\limits_{x \to -\infty} f(x)$. Ясно, что здесь можно ввести похожим образом и пределы $+\infty$, $-\infty$.

Далее такие определения понадобятся не из-за обилия множеств, а просто для единообразия рассмотрения функций на промежутках с возможным включением или исключением концов.

Предел справа в точке x_0 определяется как предел в x_0 по множеству $E \cap (x_0, \infty)$, *предел слева* в точке x_0 определяется как предел в x_0 по множеству $E \cap (-\infty, x_0)$.

Если E содержит окрестность x_0, то такие пределы обозначаются через $\lim\limits_{x \to x_0+} f(x)$ и $\lim\limits_{x \to x_0-} f(x)$ соответственно.

Например, если $f(x) = 0$ при $x \leqslant 0$ и $f(x) = 1$ при $x > 0$, то предел слева равен нулю и совпадает со значением в нуле, предел справа равен единице. В самом определении пределов слева и справа значение в самой точке никак не используется (оно может быть и не задано). Переопределение значения в точке никак не влияет на какой-либо предел в этой точке.

2.1.10. Пример. Пусть функция f на (a, b) возрастает и ограничена. Тогда существуют пределы $\lim\limits_{x \to a+} f(x)$ и $\lim\limits_{x \to b-} f(x)$ (докажите!).

Пределы функций в точке полезны в следующей ситуации: функция не была определена в концевых точках интервала, но было бы хорошо продолжить ее до непрерывной функции на всем отрезке. Это не всегда бывает очевидно. Например, мы увидим, что $\lim\limits_{x \to 0} \sin x / x = 1$, но $\lim\limits_{x\to 0}(\cos x - 1)/x = 0$. Кроме того, понятие предела в точке задействовано в определении производной — важнейшего объекта анализа.

2.1.11. Замечание. К утверждению о непрерывности композиции непрерывных функций близко (и вытекает из него) такое утверждение: пусть существуют пределы $b = \lim\limits_{x\to a} g(x)$, $c = \lim\limits_{y\to b} f(y)$, причем $f(b) = c$. Тогда $\lim\limits_{x\to a} f(g(x)) = c$. От условия $f(b) = c$ здесь нельзя отказаться (можно лишь заменить его требованием, что $g(x) \neq b$ при $x \neq a$): если взять $f(y) = 0$ при $y \neq 0$ и $f(0) = 1$, то для $g(x) = x$ при $x \neq 1/n$ и $g(1/n) = 0$ мы получим, что $f(g(1/n)) = 1$ и $f(g(x)) = 0$ при $x \neq 1/n$.

Простейшим разрывом функции в точке x_0 может быть скачок: так называется ситуация, когда имеются конечные пределы слева и справа, но хотя бы один из них не совпадает со значением в точке. Конечно, если оба предела есть и равны значению в точке, то очевидным образом функция непрерывна в этой точке. Возможна ситуация, когда нет конечного предела слева или справа. Например, так будет с функцией, равной 0 в нуле и всех точках, кроме точек вида $1/n$, в которых она принимает значения n (или $1 + 1/n$).

Количественной характеристикой разрыва в точке x_0 служит колебание функции f в точке x_0 по множеству E:

$$\omega(f, E, x_0) := \lim_{\varepsilon\to 0+} \sup_{x,y\in E, |x-x_0|<\varepsilon, |y-x_0|<\varepsilon} |f(x) - f(y)|.$$

Из (ε, δ)-определения ясно, что непрерывность в x_0 равносильна условию $\omega(f, E, x_0) = 0$.

2.1.12. Определение. *Функция f называется равномерно непрерывной на множестве E, если для каждого $\varepsilon > 0$ найдется такое число $\delta = \delta_\varepsilon > 0$, что*

$$|f(x) - f(y)| < \varepsilon \quad \textit{при всяких } x, y \in E \textit{ с } |x - y| < \delta.$$

Например, липшицева функция равномерно непрерывна, а функция $f(x) = 1/x$ на $(0, 1)$ не является равномерно непрерывной: мы не сможем найти нужное δ даже для $\varepsilon = 1$, ибо при всяком $\delta > 0$ вблизи нуля найдутся точки x, y с $f(x) - f(y) > 1$, скажем, $x = \delta$, $y = \delta/2$.

Тем не менее верен следующий факт.

2.1.13. Теорема. *Всякая непрерывная функция на отрезке равномерно непрерывна.*

Доказательство. Пусть $\varepsilon > 0$. Для каждой точки x отрезка I, на котором функция f непрерывна, найдется такое $\delta_x > 0$, что

$$|f(u) - f(v)| < \varepsilon/4 \quad \text{при всех } u, v \in I \cap (x - \delta_x, x + \delta_x).$$

Интервалы $(x-\delta_x/4, x+\delta_x/4)$ покрывают I, и мы знаем, что можно выбрать конечное подпокрытие этими интервалами с некоторыми центрами $x_1, \ldots, x_n$. Возьмем в качестве δ минимальное из $\delta_{x_1}/4, \ldots, \delta_{x_n}/4$ и проверим, что оно подходит. Пусть $x, y \in I$ и $|x - y| < \delta$. Точка x попала в какой-то интервал $(x_i - \delta_{x_i}/4, x_i + \delta_{x_i}/4)$ из выбранного конечного семейства. Так как $|y - x| < \delta \leqslant \delta_{x_i}/4$, то точка y попадает в интервал $(x_i - \delta_{x_i}, x_i + \delta_{x_i})$, что влечет оценку $|f(y) - f(x)| < \varepsilon$.

Есть простое доказательство и без покрытий. А именно, если предположить противное, то найдется такое $\varepsilon > 0$, что для всякого n найдется пара таких точек x_n, y_n, что $|x_n - y_n| < 1/n$, но $|f(x_n) - f(y_n)| \geqslant \varepsilon$. Найдется подпоследовательность $\{x_{n_k}\}$, сходящаяся к некоторой точке $x \in I$. Тогда $y_{n_k} \to x$, значит, $f(x_{n_k}) - f(y_{n_k}) \to 0$, что дает противоречие. □

Так как мы использовали лишь наличие конечного подпокрытия (или возможность выбора сходящейся подпоследовательности), то доказанная теорема верна для непрерывных функций на произвольных замкнутых ограниченных множествах прямой или комплексной плоскости.

Отметим еще такой простой факт, с помощью которого часто проверяется непрерывность.

2.1.14. Предложение. *Пусть функция f на множестве E такова, что для каждого $\varepsilon > 0$ найдется такая непрерывная на E функция f_ε, что $|f(x) - f_\varepsilon(x)| \leqslant \varepsilon$ сразу для всех $x \in E$, т. е. функция f равномерно приближается непрерывными. Тогда f тоже непрерывна.*

Доказательство. Пусть $x_0 \in E$ и $\varepsilon > 0$. По условию найдется непрерывная на E функция g, для которой $|f(x) - g(x)| \leqslant \varepsilon/4$ сразу для всех $x \in E$. Теперь по непрерывности g можно найти $\delta > 0$ такое, что $|g(x_0) - g(x)| < \varepsilon/4$ при всех $x \in E$ с $|x - x_0| < \delta$. Для таких x получаем оценку

$$|f(x) - f(x_0)| = |f(x) - g(x) + g(x) - g(x_0) + g(x_0) - f(x_0)| \leqslant \\ \leqslant |f(x) - g(x)| + |g(x) - g(x_0)| + |g(x_0) - f(x_0)| < \varepsilon,$$

ибо каждое из трех слагаемых в правой части меньше $\varepsilon/4$ (по разным причинам: первое и третье из-за оценки приближения, второе в силу выбора δ для g). □

Доказанное выше свойство непрерывных функций достигать максимумы и минимумы на отрезках в $\mathbb{R}$ переносится дословно с тем же доказательством на вещественные функции на замкнутых кругах в $\mathbb{C}$ (в случае комплексной непрерывной функции f на круге это можно применить к вещественной функции $|f|$). Это позволяет просто доказать так называемую *основную теорему алгебры*.

2.1.15. Теорема. *Всякий отличный от константы многочлен с комплексными коэффициентами* $f(z) = c_n z^n + c_{n-1}z^{n-1} + \cdots + c_0$ *имеет корень, т. е. найдется* $z_0 \in \mathbb{C}$ *с* $f(z_0) = 0$.

Доказательство. Можно считать, что $c_n \neq 0$. Пусть

$$\mu = \inf\{|f(z)|\colon z \in \mathbb{C}\}, \quad |c_k| \leqslant M,\ M > 1.$$

Заметим, что при $|z| \geqslant R = 2Mn/|c_n| + \mu + 1$ мы имеем

$$|f(z)| \geqslant |c_n||z|^n - nM|z|^{n-1} = |z|^{n-1}(|c_n||z| - nM) \geqslant |z|^{n-1}nM \geqslant \mu+1.$$

Значит, непрерывная функция $|f|$ (ее непрерывность ясна из определения, как и для вещественных функций) на замкнутом круге радиуса R принимает минимум μ во внутренней точке z_0 этого круга. Вместо оценки выше можно было бы использовать такое соображение:

$$|c_{n-1}|\,|z|^{-1} + \cdots + |c_1|\,|z|^{1-n} + |c_0|\,|z|^{-n} \to 0 \quad \text{при } z| \to \infty.$$

Поэтому $|f(z)| = |z|^n|c_n + c_{n-1}z^{-1} + \cdots + c_1 z^{1-n} + c_0 z^{-n}| \to +\infty$.

Пусть $\mu > 0$. Заменой $z - z_0$ приходим к случаю, когда $|f|$ имеет положительный минимум $|c_0|$ в нуле. Покажем, что это невозможно.

Можно считать, что $\mu = c_0 = 1$, поделив на c_0. Пусть k — наименьший номер, больший нуля, для которого $c_k \neq 0$, т. е.

$$f(z) = c_n z^n + \cdots + c_k z^k + 1, \quad c_k \neq 0.$$

Тогда $f(z) = 1 + z^k(z^{n-k} + \cdots + c_k)$. При достаточно малом $r > 0$ и $|z| \leqslant r$ имеем $|z^{n-k} + \cdots + c_{k+1}z| \leqslant |c_k|/2$.

Подберем z с $|z| = r$ так, что $|f(z)| < 1$; это даст противоречие. Возьмем z так, что $|z| = r$ и $c_k z^k = -|c_k|r^k$. Для этого запишем c_k в виде $c_k = |c_k|e^{i\theta}$ и возьмем $z = re^{-i\theta/k + i\pi/k}$, т. е. $z^k = -r^k e^{-i\theta}$, ибо $e^{i\pi} = -1$. Тогда при малых r (таких, что $|c_k|r^k < 1$)

$$|1 + c_k z^k| = 1 + c_k z^k = 1 - |c_k|r^k,$$

т. е. имеем $|f(z)| \leqslant |1 + z^k c_k| + |c_n z^n + \cdots + c_{k+1}z^{k+1})|$, что не больше $1 - |c_k|r^k + r^{k+1}Mn \leqslant 1 - |c_k|r^k/2$ при достаточно малом $r > 0$ (подходит $r \leqslant |c_k|/(2Mn)$). $\square$

Из этой важной теоремы вытекает, что

$$f(z) = c(z - z_1) \cdots (z - z_n),$$

где z_k — корни f (возможно, кратные). В самом деле, можно применить индукцию. При $n = 1$ утверждение верно. При $n > 1$ один корень z_1 есть. Сдвигом можно считать, что $z_1 = 0$, т. е. $c_0 = 0$. Тогда $f(z) = zg(z)$, где многочлен g имеет меньшую степень, так что к нему применимо предположение индукции.

Введем еще одно понятие. Пусть даны две функции φ и $\psi > 0$ на множестве $(x_0 - r, x_0) \cup (x_0, x_0 + r)$. Говорят, что φ представляет собой o-малое от ψ при $x \to x_0$ и пишут $\varphi = o(\psi)$, если $\lim\limits_{x \to x_0} \varphi(x)/\psi(x) = 0$. Например, $x^2 = o(x)$ при $x \to 0$.

2.2. Разрывы монотонных функций и обратные функции

Особо просто устроены точки разрыва монотонных функций.

2.2.1. Теорема. *Пусть функция f монотонна на отрезке I. Тогда она имеет не более счетного множества точек разрыва, причем все ее точки разрыва — скачки.*

Доказательство. Можно считать, что $f(y) \geqslant f(x)$ при $y \geqslant x$. Заметим, что во всякой точке x имеются конечные односторонние пределы $\lim\limits_{y \to x-} f(y) \leqslant \lim\limits_{y \to x+} f(y)$. Действительно, положим

$$s = \sup\{f(y)\colon\ y < x\}, \quad S = \inf\{f(y)\colon\ y > x\}.$$

Эти величины конечны, ибо значения f заключены между значениями в концах I. Ввиду возрастания f имеем $s \leqslant S$. Покажем, что верны равенства $s = \lim\limits_{y \to x-} f(y)$, $S = \lim\limits_{y \to x+} f(y)$. Пусть $\varepsilon > 0$. По определению s найдется число $y_0 < x$ с $f(y_0) \geqslant s - \varepsilon$. Тогда при всех $y \in (y_0, x)$ имеем $f(y_0) \leqslant f(y) \leqslant s$, т. е. $f(y) \in (s - \varepsilon, s]$. Аналогично доказывается второе равенство. Если $s < S$, то независимо от значения в x получаем скачок. Если же $s = S$, то (в отличие от общего случая, когда $f(x)$ может отличаться от s) мы получаем $s \leqslant f(x) \leqslant S$, т. е. $f(x) = s = S$, значит, функция f непрерывна в x.

Много ли скачков может произойти? Заметим, что в силу возрастания f при наличии скачка в x в интервале (s, S) нет значений функции f, кроме, возможно, числа $f(x)$ (всякое значение либо не больше s, либо не меньше S). Если $s < f(x) < S$, то не принимаются значения из двух интервалов, а при $f(x) = s$ или $f(x) = S$ не принимаются значения из полуинтервала. Поэтому таких промежутков — лакун в области значений — может быть самое большее счетное число (в каждом есть рациональная точка, а одна точка не может попасть сразу в два интервала из-за монотонности). □

Эта полезная теорема может провоцировать попытку извлечь из нее больше, чем она дает, но надо быть осторожным: все же общая монотонная функция не слишком похожа на ступенчатую с конечным числом ступенек (задача 2.5.9).

2.2.2. Теорема. *Пусть функция f строго возрастает на $[a,b]$. Тогда на отрезке $[f(a),f(b)]$ существует возрастающая непрерывная обратная функция.*

Аналогичное верно для строго убывающих функций.

Доказательство. Мы уже знаем, что образ $[a,b]$ — невырожденный отрезок $[f(a),f(b)]$. Из строгого возрастания следует, что определена обратная функция $f^{-1}\colon [f(a),f(b)] \to [a,b]$, причем она тоже строго возрастает. Эта функция не имеет скачков (иначе ее образ не был бы равен $[a,b]$), поэтому она непрерывна. □

2.3. Некоторые элементарные функции

Наиболее просто вводятся многочлены

$$P(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$$

с коэффициентами c_k. Их непрерывность уже была отмечена. Функция $x^{1/n}$ при $x \geqslant 0$ задается как обратная к строго возрастающей функции x^n на $[0,+\infty)$. Множество значений $x^{1/n}$ тоже $[0,+\infty)$.

Для задания степенной функции x^a и показательной функции a^x достаточно задать e^x и $\ln x$; после этого полагают

$$a^x = e^{x\ln a}, \quad x^a = e^{a\ln x}.$$

Важнейшая в анализе и приложениях функция e^x может быть введена разными способами. Один вкратце состоит в следующем. Сначала определим e^x для рациональных $x = m/n$ как $(e^m)^{1/n}$. Заметим, что $(e^m)^{1/n} = (e^{1/n})^m$, ибо в степени n правая часть равна $(e^{1/n})^{m\cdot n} = (e^{1/n})^{n\cdot m} = ((e^{1/n})^n)^m = e^m$. На рациональных числах легко проверяется (с помощью равенства $(a^{1/n}b^{1/n})^n = ab$) тождество

$$e^{x+y} = e^x e^y.$$

Затем при рациональных $y \in [0, 1/2]$ несложно проверяется неравенство $|e^y - 1| \leqslant 6y$ (задача 2.5.5), из которого следует, что на положительных рациональных числах функция e^x липшицева на каждом отрезке: при $x \leqslant R$ имеем $e^{x+h} - e^x = e^x(e^h - 1) \leqslant 6e^R h$. Это позволяет доопределить ее по непрерывности (задача 2.5.6), причем с сохранением указанного тождества. Наконец, формулой $e^{-x} = 1/e^x$ экспонента доопределяется на всей оси. Далее будет доказано, что экспонента задается рядом

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots ,$$

который абсолютно сходится при всех $x \in \mathbb{R}$ (и даже при всех комплексных x). Такой способ задания может быть альтернативным указанному выше, он имеет много достоинств. При этом способе равенство $e^1 = e$ будет не определением, а следствием доказанного ранее про число e. Наконец, можно задать e^x и как предел $\lim\limits_{n\to\infty}(1 + x/n)^n$. Множество значений функции e^x на $[0, +\infty)$ есть $[1, +\infty)$, что видно из ее возрастания, равенства $e^0 = 1$ и неограниченности. Поэтому множество значений на $(-\infty, 0]$ равно $(0, 1]$.

Функция $\ln x$ задается на $(0, +\infty)$ как обратная к e^x, что дает ее непрерывность и равенство

$$\ln(xy) = \ln + \ln y.$$

Аккуратное определение элементарных функций $\cos x$ и $\sin x$ отнюдь не элементарно. Наглядное школьное определение трактует эти функции как координаты точки на единичной окружности, полученной поворотом $(1, 0)$ на x радиан против часовой стрелки, но строгое определение должно опираться на понятие длины криволинейной дуги. Поэтому здесь мы не будем этим заниматься. Если считать известными формулы сложения, то из них легко вывести непрерывность $\cos x$ и $\sin x$, заметив, что в нуле синус непрерывен (что тоже апеллирует к длинам, как объяснено ниже).

Имеется аналитический способ задания синуса и косинуса без апелляции к дугам, а именно можно положить

$$e^{ix} = \cos x + i \sin x, \ x \in \mathbb{R},$$

где левая часть задана рядом из $(ix)^n/n!$. Выделение вещественной и мнимой частей дает равенства

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots , \quad \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots ,$$

где ряды сходятся абсолютно. При таком определении непосредственно проверяется, что $\overline{e^{ix}} = e^{-ix}$ и $|e^{ix}| = 1$ при всех $x \in \mathbb{R}$, что дает тождество $(\sin x)^2 + (\cos x)^2 = 1$. При задании синуса рядом обсуждаемый ниже замечательный предел $\lim\limits_{x\to 0} \sin x/x = 1$ обосновывается очень просто и без всяких дуг и хорд. Основные тригонометрические формулы становятся следствием равенства $e^{ix+iy} = e^{ix}e^{iy}$. Правда, это равенство надо проверить (оно пока лишь для вещественных чисел известно), но это можно сделать, перемножая ряды для e^{ix} и e^{iy} и сравнивая результат с рядом для $e^{ix+iy} = e^{ix}e^{iy}$. Однако становится совершенно неочевидным, что $\sin\pi = 0$. Да и что такое π при этом подходе (при школьном подходе это половина длины окружности единичного радиуса)? Число π здесь можно определить как наименьший строго положительный корень уравнения $\sin x = 0$ (можно доказать, что такой есть, хотя при задании рядом это отнюдь не является очевидным; например, можно показать с помощью ряда, что $\sin 2 > 0.8$, откуда $(\cos 2)^2 < 0.5$, значит, $\cos 4 < 0$, а тогда $\sin 8 < 0$). В качестве следствия получим равенства $\sin(2\pi) = 0$, $\cos(2\pi) = 1$, а также 2π-периодичность e^{ix}, а потому и $\sin x$ и $\cos x$. Все это можно проделать совершенно строго, но отнюдь не быстро, поэтому мы не будем заниматься деталями. Отметим лишь, что иррациональность числа π при таком подходе доказывается сравнительно несложно (это предмет задачи 3.8.28).

2.4. Замечательные пределы

Ряд особо важных не вполне очевидных пределов функций называют замечательными пределами. Приведем некоторые из них (существование используемых в них функций обсуждалось выше).

С помощью соотношения $e = \lim\limits_{n\to\infty} (1 + 1/n)^n$ доказываются (задача 2.5.3) соотношения

$$\lim_{x\to+\infty}\Bigl(1 + \frac{1}{x}\Bigr)^x = e, \quad \lim_{x\to+\infty}\Bigl(1 - \frac{1}{x}\Bigr)^x = \frac{1}{e}. \qquad (2.4.3)$$

Первый из этих пределов можно записать как

$$\lim_{x\to 0+} (1 + x)^{1/x} = e,$$

что логарифмированием дает равенство

$$\lim_{x\to 0} \frac{\ln(1 + x)}{x} = 1. \qquad (2.4.4)$$

К этому сводится следующий полезный предел:

$$\lim_{x\to 0}\frac{e^x-1}{x}=1. \tag{2.4.5}$$

Для обоснования положим $y=e^x-1$. Тогда имеем $x=\ln(1+y)$ и $y\to 0$ при $x\to 0$, что позволяет применить (2.4.4) в форме $y/\ln(1+y)\to 1$ при $y\to 0$.

Наконец, еще один важный предел связан с $\sin$:

$$\lim_{x\to 0}\frac{\sin x}{x}=1. \tag{2.4.6}$$

Это соотношение объясняется так (здесь речь идет именно об объяснении, а не доказательстве, позже появится формальное определение синуса, из которого будет вытекать сказанное). При $x\in(0,\pi/2)$ значение $\sin x$ считаем ординатой точки B, полученной на единичной окружности из точки $A=(1,0)$ поворотом на x радиан (конечно, предполагая, что откуда-то известно, что это можно сделать). Абсцисса этой точки есть $\cos x$. Продолжая предполагать известными весьма неочевидные вещи, связанные с площадями, замечаем, что площадь равнобедренного треугольника с вершиной в нуле и двумя другими вершинами в A и B, равная $\sin x/2$, заключена между площадями сектора, порожденного дугой длины x между A и B, и гомотетичного сектора, проходящего через $(0,\cos x)$. Площадь первого сектора равна половине длины дуги, т.е. $x/2$, площадь меньшего равна $x(\cos x)^2/2$. Сейчас не будем обсуждать, откуда такие сведения о площадях. Итак,

$$x(\cos x)^2\leqslant \sin x\leqslant x.$$

Значит,

$$1-(\sin x)^2\leqslant\frac{\sin x}{x}\leqslant 1.$$

Из правой оценки следует, что $\sin x\to 0$ при $x\to 0$. Это доказывает равенство (2.4.6). Далее мы увидим, что с помощью (2.4.6) и формул сложения можно найти все производные $\sin x$ и разложить $\sin x$ в ряд по степеням x. Этот ряд очень хорошо сходится, и в принципе можно было бы взять его за определение $\sin x$ (как было отмечено в предыдущем параграфе). Тогда (2.4.6) мгновенно вытекало бы из определения. К сожалению, основные свойства синуса при таком определении совершенно неочевидны (и трудно доказываются), скажем, периодичность при таком подходе не видна. Поэтому для аккуратного проведения описанной программы надо было бы изложить основные факты теории площадей и длин кривых, что в программе анализа делается позже. С другой стороны, вряд ли можно рекомендовать вводить синус через ряд.

Замечательные пределы помогают находить другие неочевидные пределы.

2.4.1. Пример. (i) Равенство

$$\lim_{x\to 0}\frac{\sqrt{1+x}-1}{x}=\frac{1}{2}$$

доказывается так: умножаем и делим числитель и знаменатель на функцию $\sqrt{1+x}+1$, что имеет предел 2 при $x\to 0$; в числителе получаем ровно x, сокращающийся с x в знаменателе, после чего никаких неопределенностей уже нет.

(ii) При вычислении предела

$$\lim_{x\to 0}\frac{\sin(\sin x)}{\ln(1+x)}$$

записываем отношение в нем как

$$\frac{\sin(\sin x)}{x}\frac{x}{\ln(1+x)},$$

затем первое отношение записываем как

$$\frac{\sin(\sin x)}{\sin x}\frac{\sin x}{x}$$

и замечаем, что оба сомножителя имеют предел 1.

(iii) Для нахождения предела

$$\lim_{x\to 0}\frac{(1+x)^\alpha-1}{x}$$

можно использовать замену $y=\alpha\ln(1+x)$ и тот факт, что $y\to 0$ при $x\to 0$. Тогда

$$\frac{(1+x)^\alpha-1}{x}=\frac{e^y-1}{e^{y/\alpha}-1}=\alpha\frac{e^y-1}{y}\frac{y/\alpha}{e^{y/\alpha}-1},$$

что стремится к α при $y\to 0$.

Вообще, имея какой-то предел $\lim\limits_{x\to x_0} f(x)/g(x)=L$, можно вместо x подставить функцию $\varphi(x)$, которая при $x\to x_1$ имеет предел x_0, причем отлична от x_0 при $x\neq x_1$; тогда очевидным образом получим $\lim\limits_{x\to x_1} f(\varphi(x))/g(\varphi(x))=L$. Если при этом каким-то преобразованием еще скрыть указанную структуру отношения (так что не будет сразу видно участие $\varphi(x)$), то можно изготовить не вполне банальные задачи для контрольной студентам. Для тренировки очень полезно самостоятельно порешать задачи из Демидовича [**5**], а также посмотреть, что пишут в АнтиДемидовиче [**12**].

2.5. Задачи

2.5.1. Функция Дирихле D равна 1 в рациональных точках и 0 в иррациональных. Доказать, что она разрывна в каждой точке прямой, но отдельно на множествах $\mathbb{Q}$ и $\mathbb{R}\backslash\mathbb{Q}$ непрерывна.

2.5.2. Функция Римана R равна 0 в иррациональных точках, в рациональной точке вида $x = m/n$, где дробь несократима, полагаем $R(x) = 1/n$, причем $0 = 0/1$. Доказать, что эта функция непрерывна в иррациональных точках и разрывна в рациональных.

2.5.3. Доказать соотношения (2.4.3).

2.5.4. Доказать, что ограниченная функция f на отрезке $[a, b]$ непрерывна в точности тогда, когда она имеет замкнутый график, т. е. замкнуто множество $\Gamma_f := \{(x, f(x))\colon\ x \in [a, b]\}$.

Привести пример, показывающий, что это может быть неверно для неограниченных функций.

2.5.5. Доказать, что при всех рациональных $h \in [0, 1/2]$ верно неравенство $e^h - 1 \leqslant 6h$. Для этого использовать, что $e < 3$, и доказать неравенство $3^{p/q} \leqslant 1+6p/q$, т. е. $3^p q^q < (q+6p)^q$, для натуральных чисел p, q с $2p \leqslant q$, рассмотрев слагаемое $q^{q-p}(6p)^p q!/(p!(q-p)!)$ в разложении правой части по степеням q и заметив, что $p! \leqslant p^p$, $q(q-1)\cdots(q-p+1) > (q/2)^p$. Можно также заметить, что при $1/(n+1) \leqslant h \leqslant 1/n$ имеем $e^h \leqslant e^{1/n} \leqslant 1 + 3/n$, ибо $(1 + 3/n)^n \geqslant 1 + n3/n = 4$. При этом $1/n \leqslant 2h$.

2.5.6. Пусть функция f на множестве E удовлетворяет условию Липшица: $|f(x)-f(y)| \leqslant L|x-y|$, $x, y \in E$. Доказать, что f единственным образом продолжается на замыкание E до липшицевой функции.

2.5.7. (i) Привести пример ограниченной непрерывной функции на множестве рациональных чисел, не имеющей непрерывного продолжения на всю прямую.

(ii) Пусть функция непрерывна на множестве рациональных чисел и в каждой точке прямой имеет конечный предел. Будет ли непрерывна функция на прямой, доопределенная этими пределами в иррациональных точках?

2.5.8. Пусть E — непустое замкнутое множество в $\mathbb{R}$ и функция f непрерывна на E. Рассмотрим дополнение к E, представляющее собой конечное или счетное объединение дизъюнктных интервалов или открытых лучей, и линейно доопределим на каждом из таких промежутков f так: в случае интервала с концами a, b положим $f(at + (1-t)b) = tf(a) + (1-t)f(b)$ при $t \in (0,1)$, для луча $(a, +\infty)$ положим $f(t) = f(a)$ при $t > a$, аналогично для $(-\infty, a)$. Доказать, что получена непрерывная на прямой функция.

2.5.9. Привести пример возрастающей функции на $[0,1]$ со всюду плотным множеством точек разрыва.

2.5.10. Докажите, что непрерывная функция на отрезке обладает обратной в точности тогда, когда она строго монотонна.

2.5.11. Доказать, что функция f непрерывна на прямой в точности тогда, когда множество $f^{-1}(U)$ открыто для всякого открытого множества U. Это равносильно также тому, что прообраз всякого замкнутого множества замкнут.

2.5.12. (i) Верно ли, что функция $f\colon \mathbb{R} \to \mathbb{R}$ непрерывна в точности тогда, когда она переводит отрезки в отрезки и прообраз $f^{-1}(y)$ каждой точки $y \in \mathbb{R}$ замкнут?

(ii) Пусть функция $f\colon \mathbb{R} \to \mathbb{R}$ переводит каждый отрезок в отрезок. Верно ли, что она непрерывна?

2.5.13. (i) Дана функция f на $[0,1]$, $\omega(x)$ — ее колебание в точке x. Доказать, что множества $\{x\colon\ \omega(x) \geqslant r\}$ замкнуты.

(ii) Вывести из этого, что множество точек разрыва произвольной функции f на $[0,1]$ представляет собой счетное объединение замкнутых множеств.

(iii) Показать, что для всякого множества F в $[0,1]$, представляющего собой счетное объединение замкнутых множеств, найдется функция f на $[0,1]$, множество точек разрыва которой есть F.

2.5.14. Найти пределы

(i) $\lim\limits_{x\to 0}((1+x)^{1/n} - 1)/x$, (ii) $\lim\limits_{x\to 1}\Big((1-x)^{-1} - 3(1-x^3)^{-1}\Big)$,

(iii) $\lim\limits_{x\to 0} \dfrac{x^4 + 3x^2}{x^5 + x^3 + 2x^2}$, (iv) $\lim\limits_{x\to 1} \dfrac{3x^4 - 4x^3 + 1}{(x-1)^2}$,

(v) $\lim\limits_{x\to +\infty} \Big(\dfrac{2x+3}{2x+1}\Big)^{x+1}$, (vi) $\lim\limits_{x\to +\infty} (\sin\sqrt{x+1} - \sin\sqrt{x})$.

2.5.15. Из неравенств $\left(1+\dfrac{1}{n}\right)^n < e < \left(1+\dfrac{1}{n}\right)^{n+1}$ вывести, что
$$1+\frac{1}{2}+\cdots+\frac{1}{n} = \ln n + C + \varepsilon_n,$$
где C — постоянная (число Эйлера), $\varepsilon_n \to 0$.

2.5.16. (i) Доказать, что
$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x.$$
(ii) Пусть $\varepsilon_n \to 0$ при $n \to \infty$. Доказать, что
$$\lim_{n\to\infty}\left(1+\frac{x+\varepsilon_n}{n}\right)^n = e^x.$$

2.5.17. Для $a, b > 0$ найти предел
$$\lim_{n\to\infty}\left(\frac{a^{1/n}+b^{1/n}}{2}\right)^n.$$

2.5.18. Доказать, что при $n \to \infty$
$$\cos\frac{1}{n\sqrt{n}}\cdot\cos\frac{2}{n\sqrt{n}}\cdots\cos\frac{n}{n\sqrt{n}} \to e^{-1/6}.$$

2.5.19. Пусть f_n — непрерывные функции на $[0,1]$, причем для каждого $x \in [0,1]$ существует конечный предел $f(x) = \lim\limits_{n\to\infty} f_n(x)$. Доказать, что во всяком отрезке ненулевой длины в $[0,1]$ найдется точка непрерывности f.

Привести пример, показывающий, что может не существовать отрезка ненулевой длины, на котором f непрерывна.

2.5.20. Пусть функция f на $[0,+\infty)$ ограничена на каждом отрезке.

(i) Предположим, что $\lim\limits_{x\to\infty}(f(x+1)-f(x)) = A$. Доказать, что $\lim\limits_{x\to\infty} f(x)/x = A$.

(ii) Предположим, что $f(x) \geqslant c > 0$ и $\lim\limits_{x\to\infty} f(x+1)/f(x) = A$. Доказать, что $\lim\limits_{x\to\infty}(f(x))^{1/x} = A$.

2.5.21. Доказать, что множество непрерывных функция на прямой континуально.

2.5.22. Используя алгебраический базис пространства $\mathbb{R}$ над $\mathbb{Q}$ доказать существование разрывной в каждой точке прямой функции f с тем свойством, что $f(x+y) = f(x) + f(y)$ при всех x, y.

2.5.23. Пусть функция $f\colon\ [0,1]\to[0,1]$ монотонно возрастает (ее непрерывность не предполагается!). Докажите, что существует точка x, для которой $f(x)=x$.

2.5.24. (i) Даны две непрерывных функции f и g из $[0,1]$ в $[0,1]$. Известно, что они коммутируют (т. е. $f(g(x))=g(f(x))$ для всех x), причем хотя бы одна из них монотонна. Докажите, что у них есть общая неподвижная точка.

(ii)* Докажите, что без предположения о монотонности заключение может быть неверно.

2.5.25. Имеются две дороги, ведущие из пункта A в пункт B, причем два человека, связанные веревкой длины 20 метров, идущие по разным дорогам, смогли перейти из пункта A в пункт B. Всегда ли смогут по этим дорогам разъехаться два воза с сеном, едущие между этими же пунктами в противоположных направлениях, если радиусы перевозимых стогов больше 10 метров?

2.5.26. Пусть функция f непрерывна и ограничена на $[0,+\infty)$. Доказать, что при каждом $T>0$ найдется такая последовательность точек $x_n\to\infty$, что $\lim\limits_{n\to\infty}(f(x_n+T)-f(x_n))=0$.

2.5.27. *Канторовой лестницей* называют функцию H: $[0,1]\to[0,1]$, заданная следующим образом: $H(0)=0$, $H(1)=1$, на средней трети $(1/3,2/3)$ она принимает значение $1/2$, далее, на каждом шаге с номером n индуктивного построения множества Кантора, когда удаляются интервалы $U_{n,k}$, $k=1,\ldots,2^n-1$, длины 3^{-n}, на интервале $U_{n,k}$ полагаем H равной $(2k-1)2^{-n}$. В оставшихся точках (т. е. в точках множества Кантора C) полагаем $H(x)=\inf_{y>x,y\notin C}H(y)$. Докажите, что построенная функция H монотонна и непрерывна.

2.5.28. Пусть F — непустое замкнутое множество в $[0,1]$ без изолированных точек (примером служит множество Кантора). Доказать, что существует непрерывная возрастающая функция f на $[0,1]$ с $f(0)=0$ и $f(1)=1$, которая постоянна на каждом интервале открытого дополнения к F. Вывести из этого, что F континуально.

2.5.29. Доказать, что для всякого непустого замкнутого множества Z в $[0,1]$ найдется непрерывная функция на $[0,1]$, отображающая множество Кантора на Z.

2.5.30. Доказать, что всякое непустое замкнутое множество в $\mathbb{R}$ либо конечно, либо счетно, либо континуально.

Глава 3

Производная

В этой главе обсуждается важнейшее понятие непрерывного анализа — производная.

3.1. Дифференцируемые функции

3.1.1. Определение. *Функция f, определенная в окрестности точки $x_0 \in \mathbb{R}$, называется дифференцируемой в этой точке, если существует конечный предел*

$$f'(x_0) := \lim_{h\to 0, h\neq 0} \frac{f(x_0+h)-f(x_0)}{h}, \qquad (3.1.7)$$

называемый производной f в точке x_0.

В терминах последовательностей это означает, что для всякой последовательности ненулевых чисел $h_n \to 0$ мы имеем

$$f'(x_0) = \lim_{n\to\infty} \frac{f(x_0+h_n)-f(x_0)}{h_n}. \qquad (3.1.8)$$

Существование производной равносильно также тому, что есть такое число L, что приращение $f(x_0+h)-f(x_0)$ имеет вид

$$f(x_0+h)-f(x_0) = Lh + r(h), \quad \text{где } \lim_{h\to 0, h\neq 0} \frac{r(h)}{h} = 0. \qquad (3.1.9)$$

Напомним, что в этом случае пишут $r(h) = o(h)$ и называют $r(h)$ o-маленьким от h. Ясно, что тогда $f'(x_0) = L$.

Смысл такого переписывания определения довольно глубок: при фиксированном x_0 приращение как функция h записывается как линейная функция с точностью до «малых более высокого порядка». Именно в таком виде удобнее давать определение производной в многомерных или бесконечномерных пространствах. В одномерном же случае линейная функция определяется своим коэффициентом, который с ней и отождествляется, так что здесь говорят о производной в точке как о числе, а не о линейной функции.

Линейную функцию $h \mapsto f'(x_0)h$ называют также *дифференциалом* функции f в точке x_0 (но с этим термином надо проявлять осторожность, поскольку нередко дифференциалом называют и само значение $f'(x_0)h$, скажем, пишут $df(x) = f'(x)dx$).

Из определения очевидно, что постоянная имеет нулевую производную, линейная функция $x \mapsto Lx$ имеет одну и ту же производную L в каждой точке x_0. Для более общих функций, конечно, производные в разных точках могут быть весьма различны.

Указанное свойство линейной функции говорит, что поиск производной f сводится к поиску линейной функции $h \mapsto Lh$, для которой функция $f(x) - Lx$ имеет нулевую производную в x_0.

Очевидно, что производная постоянной равна нулю: $\mathrm{const}' = 0$. Ниже мы увидим, что верно и обратное.

Заметим еще, что из дифференцируемости в точке следует непрерывность в этой точке: это очевидно из (3.1.9) или же из того, что при достаточно малых h выполнено неравенство $|f(x_0+h) - f(x_0)| \leqslant (|f'(x_0)| + 1)|h|$.

Приведем простейшие свойства производных.

3.1.2. Теорема. *Пусть функции f и g дифференцируемы в точке x_0. Тогда в этой точке дифференцируемы их сумма и произведение, причем*

$$(f+g)'(x_0) = f'(x_0) + g'(x_0),$$

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Если $g \neq 0$ в окрестности x_0, то

$$\Big(\frac{f}{g}\Big)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

В частности,

$$\Big(\frac{1}{g}\Big)'(x_0) = \frac{-g'(x_0)}{g(x_0)^2}.$$

Доказательство. Мы имеем

$$f(x_0+h) - f(x_0) = f'(x_0)h + r_1(h),$$

$$g(x_0+h) - g(x_0) = g'(x_0)h + r_2(h),$$

где отношения $r_1(h)$ и $r_2(h)$ к h стремятся к нулю при $h \to 0$. Ясно, что так будет и для $r_1(h) + r_2(h)$.

В случае произведения из этих же равенств находим, что

$$f(x_0+h)g(x_0+h)-f(x_0)g(x_0)=f(x_0)g'(x_0)h+g(x_0)f'(x_0)h+ \\ +f(x_0)r_2(h)+g(x_0)r_1(h)+r_1(h)r_2(h).$$

Ясно, что отношения трех последних слагаемых к h стремятся к нулю при $h\to 0$.

Случай частного теперь сводится к $1/g$. Здесь мы имеем

$$\frac{1}{g(x_0+h)}-\frac{1}{g(x_0)}=\frac{g(x_0)-g(x_0+h)}{g(x_0)g(x_0+h)},$$

причем после деления на $h\neq 0$ числитель при $h\to 0$ стремится к числу $-g'(x_0)$, а знаменатель стремится к $g(x_0)^2$.

Конечно, все эти соотношения легко получить и непосредственно из свойств пределов: если $h_n\to 0$, $h_n\neq 0$, то по свойствам пределов

$$\frac{(f+g)(x+h_n)-(f+g)(x)}{h_n}= \\ =\frac{f(x+h_n)-f(x)}{h_n}+\frac{g(x+h_n)-g(x)}{h_n}\to f'(x)+g'(x),$$

$$\frac{(fg)(x+h_n)-(fg)(x)}{h_n}=\frac{[f(x+h_n)-f(x)]g(x+h_n)}{h_n}+ \\ +\frac{[f(x)[g(x+h_n)-g(x)]}{h_n}\to f'(x)g(x)+f(x)g'(x),$$

ибо $g(x+h_n)\to g(x)$ из-за непрерывности в точке дифференцируемости. Аналогично рассматривается частное. □

Формула дифференцирования произведения $(fg)'=f'g+fg'$ называется *формулой Лейбница*.

Из формулы Лейбница и равенства $x'=1$ получаем по индукции формулу $(x^n)'=nx^{n-1}$, ибо $(x^n)'=x'x^{n-1}+x(x^{n-1})'$.

Имеет место также следующее правило дифференцирования композиции.

3.1.3. Теорема. *Пусть функция f дифференцируема в точке x_0, а функция g дифференцируема в точке $f(x_0)$. Тогда композиция $g\circ f$ дифференцируема в точке x_0 и*

$$(g\circ f)'(x_0)=g'(f(x_0))f'(x_0). \tag{3.1.10}$$

Доказательство. Заметим, что в силу непрерывности в x_0 функции f при малых h значения $f(x_0+h)$ попадают в окрестность $f(x_0)$, где g определена. Пусть $h_n \to 0$, $h_n \neq 0$. Рассмотрим отношение

$$[g(f(x_0+h_n)) - g(f(x_0))]/h_n.$$

Если $\delta_n = f(x_0+h_n) - f(x_0) = 0$, то это отношение равно нулю. Для оставшихся номеров запишем

$$\frac{g(f(x_0+h_n)) - g(f(x_0))}{h_n} = \\ = \frac{g(f(x_0+h_n)) - g(f(x_0))}{f(x_0+h_n) - f(x_0)} \frac{f(x_0+h_n) - f(x_0)}{h_n}.$$

Так как $\delta_n \to 0$ и $\delta_n \neq 0$ для рассматриваемых номеров, то предел первого сомножителя в правой части равен $g'(f(x_0))$, поскольку $f(x_0+h_n) = f(x_0) + \delta_n$, а предел второго равен $f'(x_0)$, что дает нужный ответ. Действительно, если есть бесконечное множество номеров, для которых $f(x_0+h_n) - f(x_0) = 0$, то $f'(x_0) = 0$, поэтому $(g \circ f)'(x_0) = 0$ независимо от того, сколько остальных номеров. Если же таких номеров конечное число, то при нахождении предела их можно не учитывать.

Возможно объяснение и через главную линейную часть приращения: так как $f(x_0+h) = f(x_0) + f'(x_0)h + r(h)$, где $r(h) = o(h)$, $g(f(x_0)+s) - g(f(x_0)) = g'(f(x_0))s + r_2(s)$, где $r_2(s) = o(s)$, то

$$g(f(x_0+h)) - g(f(x_0)) = g(f(x_0) + f'(x_0)h + r(h)) - g(f(x_0)) = \\ = g'(f(x_0))[f'(x_0)h + r(h)] + r_2(f'(x_0)h + r(h)).$$

Ясно, что $g'(f(x_0))r(h) = o(h)$ и $r_2(f'(x_0)h + r(h)) = o(h)$. В самом деле, при достаточно малых h имеем $|f'(x_0)h+r(h)| \leqslant (|f'(x_0)|+1)|h|$. Так как $r_2(s) = s\alpha(s)$, где $\lim\limits_{s\to 0} \alpha(s) = 0$, то

$$|r_2(f'(x_0)h + r(h))| \leqslant (|f'(x_0)| + 1)|h|\,|\alpha(f'(x_0)h + r(h))|,$$

$\lim\limits_{h\to 0} \alpha(f'(x_0)h + r(h)) = 0$, ибо $f'(x_0)h + r(h) \to 0$ при $h \to 0$. $\square$

Из этой теоремы следует, что если дифференцируемая функция f имеет дифференцируемую обратную функцию $g = f^{-1}$, то выполнено равенство $g'(y) = 1/f'(f^{-1}(y))$, ибо $g(f(x)) = x$, что после дифференцирования дает равенство $g'(f(x))f'(x) = 1$. Однако обратная функция, даже если и существует, не обязана быть дифференцируемой. Скажем, для $f(x) = x^3$ обратная функция $g(y) = y^{1/3}$ не имеет производную в нуле.

Для дифференцируемости обратной функции нужно отличие от нуля производной исходной функции.

3.1.4. Теорема. *Пусть функция f взаимно однозначно отображает интервал I на интервал J и имеет в точке $x_0 \in I$ отличную от нуля производную $f'(x_0)$, причем обратная функция f^{-1} непрерывна в точке $y_0 = f(x_0)$. Тогда функция $g = f^{-1}$ дифференцируема в y_0 и*

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'\bigl(f^{-1}(y_0)\bigr)}.$$

Доказательство. Пусть $h_n \to 0$, $h_n \neq 0$. При всех достаточно больших n имеем $y_0 + h_n \in J$. Положим $x_n = g(y_0 + h_n)$. Тогда $f(x_n) = y_0 + h_n$ и $x_n \to g(y_0) = x_0$. Значит,

$$\frac{g(y_0 + h_n) - g(y_0)}{h_0} = \frac{x_n - x_0}{f(x_n) - f(x_0)} \to \frac{1}{f'(x_0)},$$

ибо $f(x_n) \neq f(x_0)$. $\square$

Стоит напомнить, что если f непрерывна и строго монотонна на I (это и есть наиболее типичное условие обратимости), то функция f^{-1} непрерывна на J. Поэтому в такой ситуации нужно проверять лишь условие $f' \neq 0$. Понятно, что непрерывность g в y_0 необходима для дифференцируемости, поэтому в общем случае без нее не обойтись (она не вытекает из взаимной однозначности).

Хотя из дифференцируемости следует непрерывность самой функции, производная не обязана быть непрерывной (правда, она не может быть всюду разрывной).

3.1.5. Пример. Положим $f(x) = x^2 \sin(x^{-3})$ при $x \neq 0$ и $f(0) = 0$. Из формул дифференцирования композиции и частного и приводимой ниже формулы $(\sin x)' = \cos x$ при $x \neq 0$ получаем

$$f'(x) = 2x \sin(x^{-3}) - 3x^{-2} \cos(x^{-3}).$$

В точке 0 непосредственно по определению находим $f'(0) = 0$ в силу ограниченности синуса. Однако в нуле функция f' разрывна, так как имеется стремящаяся к нулю последовательность точек x_n, в которых $\cos(x_n^{-3}) = 0$, скажем, $x_n = (2\pi n + \pi/2)^{-1/3}$. Таким образом, ни в какой окрестности нуля f' не ограничена.

Отметим, что с помощью производной можно ввести и интеграл (интеграл Ньютона) так: дифференцируемую функцию F на прямой назовем первообразной функции f, если $F'(x) = f(x)$. Тогда интегралом от f по отрезку $[a, b]$ можно назвать $F(b) - F(a)$:

$$\int_a^b f(x)\,dx := F(b) - F(a).$$

Позже будет доказано, что непрерывные функции имеют первообразные. Из такого определения сразу следует, что многочлены имеют первообразные, а также функции вида $x^n e^{kx}$, $x^n \sin x$, $x^n \cos x$. Из формулы дифференцирования $(FG)' = FG' + F'G$ вытекает, что FG служит первообразной для $FG' + F'G$. Тем самым оказывается верной формула интегрирования по частям

$$\int_a^b F'(x)G(x)\,dx = F(b)G(b) - F(a)G(a) - \int_a^b F(x)G'(x)\,dx.$$

При таком способе действия не нужны никакие интегральные суммы. Простейшие свойства интеграла здесь становятся следствием свойств производной. Например, интеграл от суммы равен сумме интегралов, а если $f = F' \geqslant 0$, то интеграл от f неотрицателен просто в силу монотонности функции с неотрицательной производной (см. ниже). Однако такой способ интегрирования имеет и недостатки, поэтому позже интегралы Римана и Лебега будут введены иным образом.

3.2. Теорема о среднем

Важнейшая из теорем, связанных с производными, — теорема о среднем, связывающая приращение функции с производной. Однако сначала полезно получить специальный случай — теорему Ролля, который, в свою очередь, вытекает из теоремы Ферма.

Точка c называется точкой *локального максимума*, если мы имеем $f(c + h) \leqslant f(c)$ при всех достаточно малых по модулю h. Аналогично определяются точки *локального минимума*.

3.2.1. Теорема. (Теорема Ферма) *Если функция f дифференцируема в интервале (a, b), то $f'(c) = 0$ для всякой точки $c \in (a, b)$ являющейся точкой локального минимума или локального максимума.*

Доказательство. Пусть c — точка локального максимума. Тогда при малых $h > 0$ имеем $f(c+h) - f(c) \leqslant 0$, т. е. $[f(c+h) - f(c)]/h \leqslant 0$, откуда $f'(c) \leqslant 0$. При $h < 0$ мы также имеем $f(c - h) - f(c) \leqslant 0$, т. е.

$[f(c-h)-f(c)]/(-h) \geqslant 0$, откуда $f'(c) \geqslant 0$. Итак, $f'(c)=0$. Случай локального минимума аналогичен. $\square$

3.2.2. Теорема. (Теорема Ролля) *Пусть функция f непрерывна на $[a,b]$ и дифференцируема в (a,b), причем $f(a)=f(b)$. Тогда найдется такая точка $c \in (a,b)$, что $f'(c)=0$.*

Доказательство. Мы знаем, что функция f имеет точку минимума x_1 и точку максимума x_2 в отрезке $[a,b]$. Если $x_1=x_2$, то функция постоянна. Если $x_1 \neq x_2$, то хотя бы одна из этих точек является внутренней. Остается применить теорему Ферма. $\square$

Следующая теорема Лагранжа, другого французского классика, — одна из важнейших в дифференциальном исчислении, несмотря на тривиальность доказательства (геометрическое пояснение см. на с. 79).

3.2.3. Теорема. (Теорема Лагранжа) *Пусть функция f непрерывна на отрезке $[a,b]$ и дифференцируема в (a,b). Тогда найдется такая точка $c \in (a,b)$, что*

$$f(b)-f(a)=f'(c)(b-a).$$

Доказательство. Применим теорему Ролля к функции

$$g(x)=f(x)-x\frac{f(b)-f(a)}{b-a},$$

имеющей равные значения в концах. Тогда $g'(c)=0$ в некоторой внутренней точке c, что дает нужное равенство. $\square$

3.2.4. Следствие. *Если функция f дифференцируема в (a,b), то для всяких двух значений y_1 и y_2 ее производной, принимаемых в каких-то точках (a,b), функция f' принимает и всякое значение y между y_1 и y_2.*

Доказательство. Общий случай сразу сводится к случаю нулевого промежуточного значения переходом к функции $g(x)=f(x)-yx$, для которой

$$g'(x_1)=f'(x_1)-y=y_1-y<0,\ g'(x_2)=f'(x_2)-y=y_2-y>0,$$

где $y_1=f(x_1)<y_2=f(x_2)$. Итак, считаем, что есть точки x_1 и x_2 с $f'(x_1)>0$, $f'(x_2)<0$, надо найти точку x с $f'(x)=0$. Пусть $x_1<x_2$. Тогда $f(x)>f(x_1)$ при $x>x_1$, близких к x_1. Значит, точка максимума x_0 на $[x_1,x_2]$ отлична от x_1. Она отлична и от x_2, ибо $f(x)>f(x_2)$ при $x<x_2$, близких к x_2. По теореме Ферма $f'(x_0)=0$. Случай $x_2<x_1$ аналогичен, надо лишь рассмотреть точку минимума. $\square$

3.2.5. Следствие. *Дифференцируемая на (a, b) функция f является возрастающей, т. е. $f(y) \geqslant f(x)$ при $y \geqslant x$ в точности тогда, когда $f' \geqslant 0$ на (a, b).*

Дифференцируемая на (a, b) функция f является убывающей, т. е. $f(y) \leqslant f(x)$ при $y \geqslant x$ в точности тогда, когда $f' \leqslant 0$ на (a, b).

Значит, f постоянна на (a, b) в точности тогда, когда $f' = 0$ на (a, b).

Доказательство. Если f возрастает, то $[f(x+h) - f(x)]/h \geqslant 0$ при $h > 0$, откуда $f'(x) \geqslant 0$. Обратное ясно из теоремы о среднем. Для убывающей функции аналогично. □

Это следствие имеет весьма нетривиальное частичное усиление (см. [**13**]): возрастающая функция имеет неотрицательную производную во всех точках отрезка, кроме множества меры нуль, т. е. множества, которое можно покрыть счетным набором интервалов со сколь угодно малой суммой. Однако последнее свойство еще не характеризует возрастание. Если же рассматривать так называемые обобщенные производные, то неотрицательность производной оказывается равносильной возрастанию функции уже без предположений о дифференцируемости.

3.2.6. Следствие. *Дифференцируемая на (a, b) функция удовлетворяет условию Липшица с постоянной L в точности тогда, когда $|f'(x)| \leqslant L$ на (a, b).*

Конечно, липшицева функция не обязана быть дифференцируемой (пример: $|x|$ в нуле), однако оказывается, что липшицева функция имеет много точек дифференцируемости, а именно всюду плотное множество (а множество точек недифференцируемости имеет нулевую длину). Этот факт, открытый Лебегом, весьма нетривиален: липшицевость дает лишь ограниченность отношений $(f(x+h) - f(x))/h$, но почему есть много точек, где они сходятся?

Совершенно аналогично теореме Лагранжа доказывается следующая теорема Коши.

3.2.7. Теорема. *Пусть функции f и g непрерывны на $[a, b]$ и дифференцируемы в (a, b). Тогда найдется такая точка $c \in (a, b)$, что*

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c).$$

Если $g'(t) \neq 0$ при всех $t \in (a, b)$, то найдется такая точка $c \in (a, b)$, что

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Доказательство. Функция
$$h(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x)$$
имеет равные значения в a и b, значит, ее производная имеет нуль внутри. $\square$

Из теоремы Коши вытекает так называемое *правило Лопиталя*. Оно позволяет находить предел отношения $f(x)/g(x)$ в точке в том случае, когда обе функции имеют нулевой предел в этой точке, так что неприменимо стандартное правило обращения с пределом частного (как всегда, мешает нуль знаменателя).

3.2.8. Теорема. *Пусть функции f и g дифференцируемы на интервале (a, b), g' не имеет нулей, $\lim\limits_{x\to a+} f(x) = \lim\limits_{x\to a+} g(x) = 0$ и существует конечный или бесконечный предел $\lim\limits_{x\to a+} f'(x)/g'(x)$. Тогда существует и совпадает с ним $\lim\limits_{x\to a+} f(x)/g(x)$. Аналогичное утверждение верно в случаях $x \to a-$, $x \to +\infty$ и $x \to -\infty$.*

Доказательство. Из условия и теоремы Ролля следует, что в (a, b) функция g имеет не более одного нуля, поэтому можно считать, что нулей нет вообще. Пусть предел L отношения производных конечен и $\varepsilon > 0$. Найдем такое $\delta > 0$, что при $x \in (a, a+\delta]$ предел отношения производных лежит в $(L-\varepsilon, L+\varepsilon)$. Для каждого такого x ввиду равенства нулю пределов самих функций найдется такая достаточно близкая к a точка $u = u(x) \in (a, x)$, что $f(x)/g(x)$ отличается от $[f(x)-f(u)]/[g(x)-g(u)]$ меньше, чем на ε. Найдется точка $c \in (u, x)$, для которой
$$\frac{f(x) - f(u)}{g(x) - g(u)} = \frac{f'(c)}{g'(c)}.$$
Так как правая часть лежит в $(L-\varepsilon, L+\varepsilon)$, то $f(x)/g(x)$ лежит в $(L-2\varepsilon, L+2\varepsilon)$. В случае бесконечного предела, скажем, $+\infty$, рассуждения аналогичны, но $\delta > 0$ берется так, что отношение производных в $(a, a+\delta)$ будет больше заданного большого числа. Конечно, аналогичное утверждение верно и для пределов слева. Случаи $a = +\infty$ и $a = -\infty$ разбираются так же, но их можно вывести из доказанного переходом к $f(1/y)$, $y \to 0$. $\square$

3.2.9. Замечание. От условия равенства нулю пределов самих функций отказаться нельзя (скажем, для $f(x) = x+1$ и $g(x) = x$ в нуле), но это условие можно заменить таким условием: $\lim\limits_{x\to a+} g(x) = +\infty$. Это видно из того, что в аналогично полученных неравенствах

$$L - \varepsilon < \frac{f(x) - f(a+\delta)}{g(x) - g(a+\delta)} < L + \varepsilon$$

можно числитель и знаменатель разделить на $g(x)$ и воспользоваться тем, что обе величины $f(a+\delta)/g(x)$ и $g(a+\delta)/g(x)$ стремятся к нулю при $x \to a+$.

Отметим еще, что правило Лопиталя необратимо: верно равенство

$$\lim_{x\to 0} \frac{x^2 \sin \frac{1}{x}}{x} = 0,$$

но отношение производных предела не имеет. Аналогичное явление имеет место для предела $\lim\limits_{x\to 0}(x - \sin x)/(x + \sin x)$.

3.3. Вторая производная и исследование функций с помощью первых двух производных

Из предыдущего обсуждения ясно, что на участках положительности производной функция возрастает, на участках отрицательности убывает, в точках локального минимума и максимума производная равна нулю (но не наоборот: функция x^3 имеет нулевую производную в нуле, однако не меняет возрастание на убывание). Обычно еще полезно знать нули функции, а значения производной в нулях дает угол наклона касательной, что в совокупности уже позволяет набросать очертания графика функции. Для более точной картины нужна вторая производная, причем во многих приложениях этого бывает достаточно.

3.3.1. Предложение. *Пусть функция f дифференцируема в некоторой окрестности точки x_0 и $f'(x_0) = 0$. Если в x_0 есть вторая производная и $f''(x_0) > 0$, то точка x_0 является точкой локального минимума, а если $f''(x_0) < 0$, то это точка локального максимума.*

Доказательство. Достаточно рассмотреть первый случай. Пусть $h > 0$. По теореме о среднем имеем $f(x_0 + h) - f(x_0) = f'(\xi)h$, где $\xi \in (x_0, x_0 + h)$. В силу условия $f''(x_0) > 0$ имеем $f'(\xi) > 0$ при достаточно малых h, ибо иначе $[f'(\xi) - f'(x_0)]/(\xi - x_0) \leqslant 0$, откуда при $h \to 0$ мы бы получили $\xi \to x_0$ и тогда $f''(x_0) \leqslant 0$. Итак, $f(x_0 + h) > f(x_0)$ при малых $h > 0$. Аналогично $f(x_0 + h) < f(x_0)$ при малых $h < 0$. $\square$

По-прежнему, случай $f''(x_0) = 0$ не позволяет сделать никакого заключения.

С помощью второй производной можно продолжить разложение функции с точностью до более малых более высокого порядка. А именно верна следующая *формула Тейлора со второй производной.*

3.3.2. Теорема. *Если функция f дифференцируема в интервале (a,b), $x_0 \in (a,b)$ и существует $f''(x_0)$, то при достаточно малых h имеем*

$$f(x_0+h)-f(x_0)=f'(x_0)h+\frac{1}{2}f''(x_0)h^2+r_2(h), \quad r_2(h)=o(h^2).$$

Доказательство. Функция

$$r_2(h)=f(x_0+h)-f(x_0)=f'(x_0)h+\frac{1}{2}f''(x_0)h^2$$

определена в окрестности нуля и дважды дифференцируема в нуле, причем $r_2(0)=r_2'(0)=r_2''(0)=0$. Покажем, что $r_2(h)=o(h^2)$. По теореме о среднем $r_2(h)=r_2'(\xi_h)h$, где $\xi_h\in(0,h)$ при $h>0$ и аналогично при $h<0$. Так как $r_2''(0)=0$, то $r_2'(h)/h\to 0$ при $h\to 0$, т. е. $r_2'(h)=hr(h)$, где $\lim\limits_{h\to 0} r(h)=0$. Поэтому $r_2'(\xi_h)=r(\xi_h)\xi_h=\alpha r(\xi_h)h$, где $|\alpha|\leqslant 1$. Итак, $|r_2(h)|\leqslant h^2|r(\xi_h)|$, причем $r(\xi_h)\to 0$ при $h\to 0$. □

Из этой формулы также видна роль знака $f''(x_0)$ в вопросе о локальном экстремуме, когда $f'(x_0)=0$.

В следующем параграфе формула Тейлора будет получена для производных всех порядков.

Знак второй производной на интервале дает также выпуклость (вверх или вниз) функции на этом интервале.

3.3.3. Определение. *Функция f на промежутке J с концами a,b называется выпуклой* (*или выпуклой вниз*)*, если*

$$f(tx+(1-t)y)\leqslant tf(x)+(1-t)f(y) \quad \forall\, x,y\in J,\ \forall\, t\in[0,1].$$

Функция f называется вогнутой (*или выпуклой вверх*) *на J, если*

$$f(tx+(1-t)y)\geqslant tf(x)+(1-t)f(y) \quad \forall\, x,y\in J,\ \forall\, t\in[0,1].$$

Вогнутость f равносильна выпуклости $-f$. Выпуклость означает, что график функции f на $[x,y]$ лежит под графиком аффинной функции, совпадающей с f в концах x,y, причем так происходит для всякого отрезка в J.

Очевидно, что аффинная функция выпукла и вогнута одновременно, причем других таких нет. Функция x^2 выпукла, что легко проверить непосредственно (но сейчас это будет выведено из более общего признака). Этот пример показывает, что выпуклость не совсем связана с монотонностью (но связь есть, как мы увидим). Далее, функция $|x|$ выпукла, но не везде имеет производную (тем более не везде имеет вторую производную).

3.3.4. Предложение. *Пусть функция f на интервале (a,b) дважды дифференцируема. Тогда ее выпуклость на (a,b) равносильна условию $f''(x) \geqslant 0$ при всех $x \in (a,b)$.*

Доказательство. Пусть $x, y \in (a,b)$, $x < y$ и $t \in (0,1)$. Введем еще точку $z = tx + (1-t)y$. Легко проверить, что неравенство выпуклости равносильно неравенству

$$\frac{f(z)-f(x)}{z-x} \leqslant \frac{f(y)-f(z)}{y-z}. \tag{3.3.11}$$

Пусть $f'' \geqslant 0$. Тогда $f'(\xi) \leqslant f'(\eta)$ при $\xi \leqslant \eta$. Поэтому предыдущее неравенство выполнено в силу теоремы о среднем, которая представляет указанные в нем отношения как значения производных в некоторых точках из (x,z) и (z,y).

Предположим теперь, что (3.3.11) верно для всех троек x, z, y из (a,b) с $x < z < y$ и покажем, что $f'' \geqslant 0$. При $z \to x$ получаем

$$f'(x) \leqslant \frac{f(y)-f(x)}{y-x}.$$

Аналогично при $y \to x$ получаем

$$\frac{f(y)-f(x)}{y-x} \leqslant f'(y).$$

Значит, $f'(x) \leqslant f'(y)$ при $x < y$, откуда

$$f''(x) = \lim_{y\to x+} (f(y)-f(x))/(y-x) \geqslant 0,$$

что и требовалось. □

3.3.5. Замечание. Из доказательства видно, что выпуклость однократно дифференцируемой функции f равносильна тому, что f' возрастает (необязательно строго). Можно показать, что без каких-либо дополнительных предположений выпуклая функция непрерывна и имеет производную всюду, кроме не более чем счетного множества точек.

На самом деле выпуклость f равносильна неравенству $f'' \geqslant 0$ для любых функций, но производные тогда надо рассматривать в смысле обобщенных функций.

3.3.6. Пример. Функции $|x|^p$ при $p \geqslant 1$ и e^x выпуклы на всей прямой, функции $\ln x$ и $x^{1/p}$ вогнуты на $(0, +\infty$.

С помощью первых двух производных можно довольно неплохо нарисовать график не слишком сложной функции.

Во-первых, заметим, что геометрический смысл первой производной функции f в точке x_0 таков: это тангенс угла наклона касательной к графику в точке $(x_0, f(x_0))$. Касательная задается уравнением

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Не так просто формально и точно определить, что такое касательная. График есть множество пар $\{(x, f(x)\}$. Если функция выпукла, то множество над графиком тоже выпукло, тогда касательная оказывается под графиком (см. рис. 1).

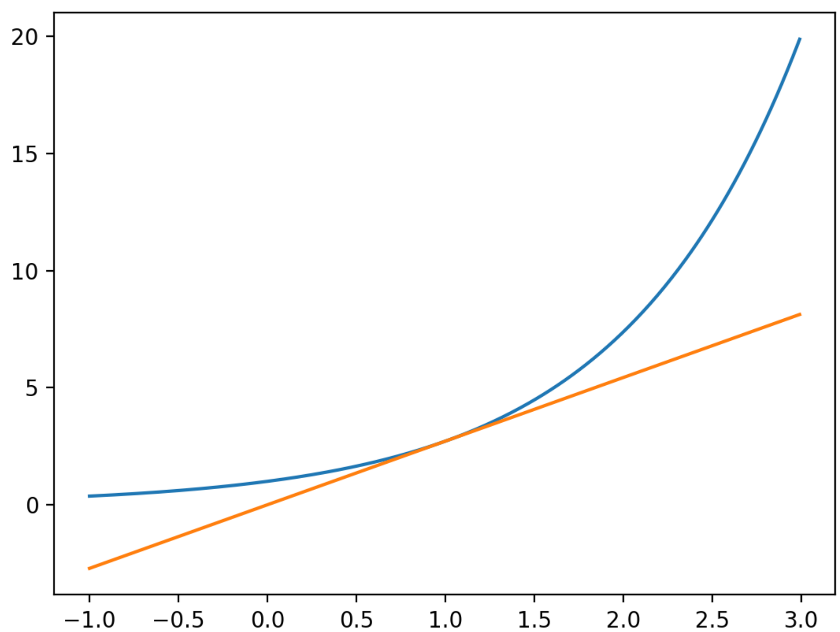

Рис. 1. Касательная к выпуклому графику

Для вогнутой функции касательная лежит над графиком (см. рис. 2).

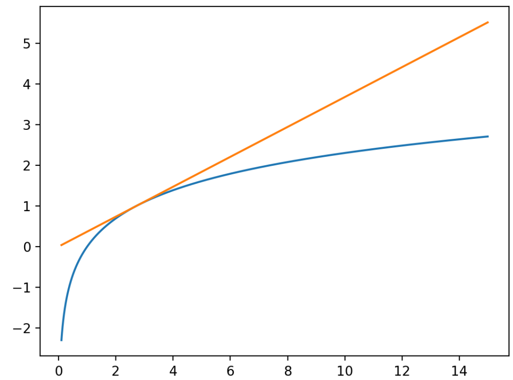

Рис. 2. Касательная к вогнутому графику

Может быть выпуклость с одной стороны и вогнутость с другой — точка перегиба (см. рис. 3).

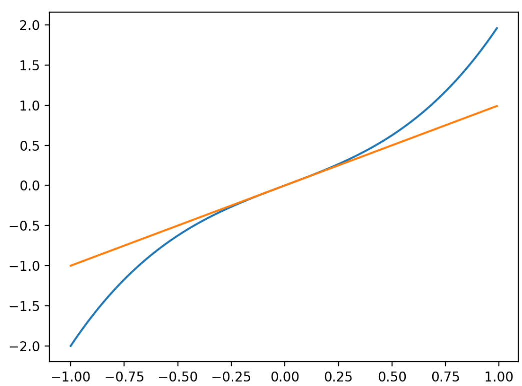

Рис. 3. Точка перегиба

В общем случае никакой геометрической наглядности нет. Скажем, функция $f(x) = x^2 \sin(1/x)$, $f(0) = 0$, имеет нулевую производную в нуле, но картинка, нарисованная компьютером (см. рис. 4), не очень хорошо передает сильную осцилляцию около нуля (ну, тяжело ему находить $1/x$ при очень малых x).

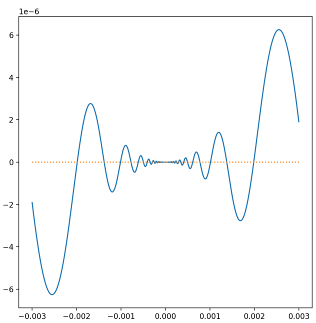

Рис. 4. Осцилляция

Картинка (см. рис. 5) может также объяснить (но не доказать) теорему о среднем: ищем точку, в которой касательная к графику параллельна хорде (тангенс угла наклона хорды равен $(f(b)-f(a))/(b-a)$).

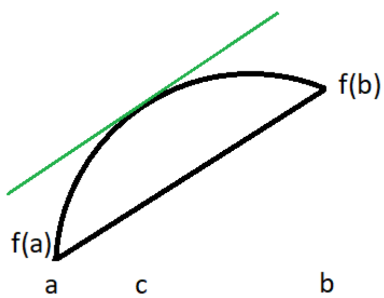

Рис. 5. Теорема о среднем

Однако для относительно простых функций можно предложить такую методику исследования графика: 1) находим нули производной, между ними функция монотонна; 2) однако как она монотонна, может подсказать вторая производная, поэтому полезно знать и ее нули. Если на промежутке $f' > 0$, $f'' > 0$, то картинка похожа на x^2 или e^x, если $f' > 0$, $f'' < 0$, то похожа $\ln x$ или $x^{1/2}$. Если $f''(x_0) = 0$, а слева и справа f'' имеет разные знаки, то x_0 называют точкой перегиба. Скажем, $x_0 = 0$ — точка перегиба для x^3.

3.4. Формула Тейлора

Теорема Тейлора (или формула Тейлора) дает аналог теоремы о среднем для старших производных и служит эффективным средством приближения гладких функций многочленами (а также дает разложения в степенные ряды для некоторых функций).

3.4.1. Теорема. (Формула Тейлора с остаточным членом в форме Лагранжа) *Пусть функция f на интервале (a, b) дифференцируема $n+1$ раз. Тогда для всяких точек x и $x+h$ из (a, b) найдется такая точка $\xi \in (x, x+h)$, что*

$$f(x+h) - f(x) = \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(x) h^k + \frac{1}{(n+1)!} f^{(n+1)}(\xi) h^{n+1}. \quad (3.4.12)$$

Доказательство. Можно считать, что $x = 0$. Рассмотрим на $[0, h]$ (если $h < 0$, то на $[h, 0]$) вспомогательную функцию

$$F(t) = f(h) - \Big[f(t) + f'(t)(h-t) + \cdots + \frac{1}{n!} f^{(n)}(t)(h-t)^n\Big].$$

Непосредственное вычисление по формуле Лейбница для произведений дает равенство

$$F'(t) = -\frac{1}{n!} f^{(n+1)}(t)(h-t)^n.$$

Применим теорему о среднем в форме Коши к двум функциям $F(t)$ и $G(t) = (h-t)^{n+1}$. Это даст точку $\xi \in (0, h)$, для которой

$$\frac{F(h) - F(0)}{G(h) - G(0)} = \frac{F'(\xi)}{G'(\xi)}.$$

При этом $G(h) - G(0) = -h^{n+1}$, $G'(\xi) = -(n+1)(h-\xi)^n$,

$$F'(\xi) = -f^{(n+1)}(\xi)(h-\xi)^n/n!,$$

так что $F'(\xi)/G'(\xi) = f^{(n+1)}(\xi)/(n+1)!$. Значит,

$$F(h) - F(0) = -h^{n+1} f^{(n+1)}(\xi)/(n+1)!.$$

Поскольку $F(h) - F(0) = -F(0)$ совпадает с

$$-f(h) + f(0) + f'(0)h + \cdots + f^{(n)}(0)h^n/n!$$

получаем доказываемое равенство. □

Заменив x на x_0, а $x+h$ на x, Полученную формулу можно записать в следующем виде с некоторым $\xi \in (x_0, x)$:

$$f(x) - f(x_0) = \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k + \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x-x_0)^{n+1}.$$

Отметим еще, что h может быть отрицательным в доказанной теореме, а x может быть меньшим x_0 в последней формуле.

Имеется несколько более общая (но и более техническая) формула Тейлора, из которой предыдущая следует при $\varphi(t) = (x+h-t)^{n+1}$.

3.4.2. Теорема. *Пусть функция f на интервале (a, b) удовлетворяет условиям предыдущей теоремы, а функция φ дифференцируема в (a, b), причем $\varphi' \neq 0$. Тогда для всяких точек x и $x+h$ из (a, b) найдется такая точка $\xi \in (x, x+h)$, что*

$$\begin{aligned} f(x+h) - f(x) &= \\ &= \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(x)h^k + \frac{\varphi(x+h) - \varphi(x)}{\varphi'(\xi)n!} f^{(n+1)}(\xi)(x+h-\xi)^n. \end{aligned}$$

Доказательство. То же рассуждение надо применить к функции $G(t) = \varphi(t)$. □

Если во второй теореме взять $\varphi(t) = x+h-t$, то приходим к такой формуле

3.4.3. Следствие. (Формула Тейлора с остаточным членом в форме Коши) *Пусть функция f на интервале (a, b) дифференцируема $n+1$ раз. Тогда для всяких точек x и $x+h$ из (a, b) найдется такая точка $\xi \in (x, x+h)$, что*

$$f(x+h) - f(x) = \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(x)h^k + r_n(x, h),$$

где остаточный член имеет вид

$$r_n(x, h) = \frac{1}{n!} f^{(n+1)}(\xi)(x+h-\xi)^n h, \quad \xi \in (x, x+h). \tag{3.4.13}$$

Эта менее изящная формула оказывается полезной, как мы увидим ниже при разложении логарифма.

Возникает вопрос, что можно получить в качестве остаточного члена без производной порядка $n+1$ в духе известного представления $f(x+h)-f(x)=f'(x)h+o(h)$ в случае однократной дифференцируемости. Ответ оказывается аналогичным и дается формулой Тейлора с остаточным членом в форме Пеано.

3.4.4. Теорема. *Пусть функция f на интервале (a,b) дифференцируема n раз. Тогда для всякой точки x из (a,b) при достаточно малых h имеем*

$$f(x+h)-f(x)=\sum_{k=1}^{n}\frac{1}{k!}f^{(k)}(x)h^k+r_n(x,h), \quad r_n(x,h)=o(h^n).$$

Доказательство — предмет задачи 3.8.3.

При фиксированном x разложение Тейлора без остаточного члена представляет собой многочлен по h. Поэтому для $x=0$ естественно вместо h использовать переменную x, что дает выражение

$$f(x)=f(0)+f'(0)x+\cdots+\frac{f^{(n)}(0)x^n}{n!}+r_n(x),$$

называемое формулой Маклорена.

Даже если функция f бесконечно дифференцируема на всей прямой, ряд Тейлора может не сходиться вне нуля, а если даже и сходится, то сумма может быть отлична от f. Например, пусть $f(x)=e^{-1/x^2}$ при $x\neq 0$ и $f(0)=0$. Тогда $f^{(n)}(0)=0$ при всех n, поэтому все члены ряда Тейлора при разложении в нуле равны нулю, так что ряд сходится превосходно, но вовсе не к f.

Как узнать, когда ряд сходится к f? Простым достаточным условием сходимости ряда Тейлора к функции f на (x_0-r,x_0+r) служит оценка $|f^{(n)}(\xi)|\leqslant Cn!r^{-n}$ при всех n. При выполнении этой оценки и $|h|\leqslant\delta<r$ остаточный член не превосходит по модулю $C(\delta/r)^n$, где $\delta/r<1$, что обеспечивает сходимость ряда к $f(x_0+h)$ при всех $|h|<r$. Указанное условие несложно проверяется для основных элементарных функций. Ниже будет доказано и в определенном смысле обратное утверждение: представимые степенными рядами функции бесконечно дифференцируемы и их ряды Тейлора дают исходные представления.

Разложения Тейлора полезны при нахождении пределов отношений. Предположим, что даны гладкие функции f и g в окрестности точки x_0, для которых $f(x_0)=g(x_0)=0$. Тогда без дополнительных условий ничего нельзя сказать про предел отношения $f(x)/g(x)$ при

$x \to x_0$, даже если $g(x) \neq 0$ при $x \neq x_0$: предела может не быть вовсе, а если и есть, то может оказаться любым. Однако если нам известны разложения $f(x) = a_k(x-x_0)^k + o(x^k)$, $g(x) = b_k(x-x_0)^k + o(x^k)$, где $b_k \neq 0$, то после деления на x^k мы приходим к отношению функций с ненулевым пределом b_k знаменателя, так что $\lim\limits_{x\to x_0} f(x)/g(x) = a_k/b_k$. Нужные разложения как раз и получаются с помощью формулы Тейлора, но надо иметь в виду, что нередко бывает проще свести задачу к известным пределам с помощью замены переменной, чем вычислять тейлоровские коэффициенты.

3.5. Производные элементарных функций

Для степеней x^n из определения и вычисления $(x+h)^n - x^n$ сразу получаем $(x^n)' = nx^{n-1}$, что выше было выведено из формулы Лейбница. По формуле для обратной функции находим $(x^{1/n})' = n^{-1}x^{1-1/n}$.

Для синуса и косинуса имеем

$$(\sin x)' = \cos x, \quad (\cos x)' = -\sin x.$$

Эти формулы доказываются так. Замечательный предел $\lim\limits_{x\to 0} \sin x/x = 1$ дает первую формулу в нуле. В других точках она получается из известной формулы для $\sin(x+h) - \sin x = 2\sin(h/2)\cos(x+h/2)$. Случай косинуса аналогичен, но может быть выведен и из формула для синуса с помощью равенства $\cos x = \sin(x+\pi/2)$.

Для экспоненты имеем

$$(e^x)' = e^x,$$

что вытекает из формулы $e^{x+h} - e^x = e^x(e^h - 1)$ и соотношения $\lim\limits_{h\to 0}(e^h - 1)/h = 1$, см. (2.4.5).

Следовательно, для обратной функции $\ln x$ получаем

$$(\ln x)' = \frac{1}{x}.$$

Далее,

$$(x^\alpha)' = \alpha x^{\alpha-1}.$$

$$(a^x)' = (\ln a)a^x,$$

что вытекает из равенств $a^x = e^{x\ln a}$, $x^\alpha = e^{\alpha\ln x}$ и правила дифференцирования композиции.

Рассмотрим несколько подробнее разложения Тейлора ряда элементарных функций.

Для e^x мы имеем $(e^x)^{(n)} = e^x$, поэтому разложение Тейлора дает $e^x = 1 + x + \frac{x^n}{n!} + \frac{e^\xi x^{n+1}}{(n+1)!}$, где $\xi \in (0, x)$. Таким образом, остаточный член оценивается через $x^{n+1}e^x/(n+1)!$, что для всех x из фиксированного отрезка $[-R, R]$ дает универсальную оценку через величину $R^{n+1}e^R/(n+1)!$, которая, как мы знаем, стремится к нулю при $n \to \infty$. Таким образом, получаем разложение в сходящийся ряд

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots .$$

Аналогично для функций $\sin x$ и $\cos x$ имеем

$$(\sin x)^{(4n)} = \sin x, \quad (\sin x)^{(4n+1)} = \cos x,$$

$$(\sin x)^{(4n+2)} = -\sin x, \quad (\sin x)^{(4n+3)} = -\cos x.$$

Поэтому получаем разложения

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + \cdots ,$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + \cdots .$$

Далее, при $|x| < 1$ имеем

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + \frac{(-1)^n}{n}x^n + \cdots ,$$

при $|x| < 1$ и всех α имеем

$$(1+x)^\alpha = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + \cdots .$$

Из сказанного выше очевидно, что разложения экспоненты, синуса и косинуса сходятся во всех точках к исходным функциям, ибо остаточные члены в форме Лагранжа стремятся к нулю. Для двух последних функций это очевидно лишь при $x \geqslant 0$. Скажем, для $\ln(1+x)$ мы получаем

$$\frac{(-1)^{n+1}}{n}\frac{x^{n+1}}{(1+\xi)^n}.$$

При x около -1 расположение ξ неочевидно, что не позволяет заключить, что второй сомножитель оценивается единицей. Однако остаточный член в форме Коши позволяет это сделать. Здесь при $x \in (-1, 0)$

мы имеем (конечно, с иным ξ) величину

$$(-1)^n x\Big(\frac{x-\xi}{1+\xi}\Big)^n, \quad \xi(x,0),$$

так что модуль величины в скобках равен $(|x|-|\xi|)/(1-|\xi|) < 1$, поэтому остаток стремится к нулю. Обратим внимание, что в этом примере ряд сходится при $x \in (-1,1]$, хотя представленная этим рядом функция бесконечно дифференцируема и при $x > 1$. Причина этого явления обсуждается в комплексном анализе.

3.6. Степенные ряды

Степенным рядом называют ряд вида

$$\sum_{n=1}^{\infty} c_n x^n$$

или в комплексном случае

$$\sum_{n=1}^{\infty} c_n z^n.$$

Числа c_n называют коэффициентами степенного ряда. Аналогично рассматриваются также степенные ряды вида

$$\sum_{n=1}^{\infty} c_n (x-x_0)^n, \quad \sum_{n=1}^{\infty} c_n (z-z_0)^n,$$

сводящиеся к указанным заменами $x-x_0$ и $z-z_0$.

Причина полезности привлечения комплексных рядов даже в случае исходного вещественного ряда раскрывается в комплексном анализе. Степенные ряды могут быть полезны для нахождения сумм числовых рядов. Например, если нам известно разложение функции f в ряд $f(x)=\sum_{n=1}^{\infty} c_n x^n$, сходящийся при $x=1$, то $f(1)$ будет служить суммой ряда из c_n. Скажем, используя тейлоровские разложения различных функций, можно находить довольно неочевидные суммы. На самом деле для всякой не слишком банальной функции со сходящимся рядом Тейлора можно подставить какое-нибудь число из области сходимости и получить такой ряд, что не знающему исходную разложенную функцию будет вряд ли возможно найти сумму.

Сразу ясно, что при $x=0$ ряд сходится, но других точек сходимости может и не быть. Например, так происходит для ряда из $n^n x^n$, общий член которого не стремится к нулю при $x \neq 0$.

Основной результат о сходимости степенных рядов состоит в следующем.

3.6.1. Теорема. *Если ряд $\sum_{n=1}^{\infty} c_n z^n$ сходится в некоторой точке z_0 комплексной плоскости, то он сходится абсолютно при всех z в открытом круге $|z| < |z_0|$. Радиус R максимального круга, внутри которого сходится этот ряд, задается формулой*

$$R^{-1} = \overline{\lim_{n\to\infty}} |a_n|^{1/n}.$$

Доказательство. Пусть сходится ряд из $c_n z_0^n$. Тогда при всех достаточно больших n мы имеем $|c_n||z_0|^n \leqslant 1$. Можно считать, что это верно для всех n. Тогда $|c_n||z|^n \leqslant (|z|/|z_0|)^n$, так что ряд из $c_n z^n$ сходится абсолютно.

Пусть $|z| < R - \delta$, где $\delta > 0$ и R задано указанной выше формулой. Выберем $\varepsilon > 0$ столь малым, что $q = (R-\delta)(R^{-1}+\varepsilon) < 1$. Для всех достаточно больших n имеем $|c_n|^{1/n} \leqslant R^{-1} + \varepsilon$. Тогда

$$|c_n|^{1/n}|z| \leqslant (R^{-1}+\varepsilon)(R-\delta) = q,$$

поэтому сходимость ряда из $|c_n||z|^n$ следует из признака Коши.

Наконец, если наш ряд сходится в точке z, то $|c_n||z|^n \leqslant 1$ при достаточно больших n, откуда $|z||c_n|^{1/n} \leqslant 1$, значит, $|z|R^{-1} \leqslant 1$, т. е. верно неравенство $|z| \leqslant R$. □

На границе максимального круга сходимости поведение ряда может быть различным. Может иметь место сходимость во всех граничных точках. Например, ряд $\sum_{n=1}^{\infty} n^{-2} z^n$ имеет круг сходимости радиуса $R = 1$ и сходится на границе. Ряд $\sum_{n=1}^{\infty} z^n$ имеет такой же круг сходимости, но не сходится в граничных точках, ибо общий член там не стремится к нулю. Ряд $\sum_{n=1}^{\infty} (-1)^n n^{-1} z^n$ с тем же радиусом сходимости сходится в точке $z = 1$ и расходится в точке $z = -1$. Множество точек сходимости на граничной окружности может быть сложно устроено (но это не может быть произвольное множество). В курсе комплексного анализа выясняется, что на границе круга сходимости в $\mathbb{C}$ обязательно есть точка, в окрестность которой нельзя аналитически продолжить функция аргумента z, заданную суммой ряда внутри круга сходимости. Однако таким свойством обладает не всякая точка границы круга сходимости.

Доказанная теорема показывает, что известные нам разложения элементарных функций позволяют продолжить их в комплексную область. Для некоторых из них (типа $f(x) = (1 + x^2)^{1/2}$) разложения сходятся не во всей комплексной плоскости, что указывает на наличие особенностей, невидимых с вещественной прямой, на которой такие функции вполне могут быть бесконечно дифференцируемыми. Скажем, для указанной функции проблемными оказываются точки i и $-i$. В некоторых случаях с особенностями можно бороться путем существенного усложнения области определения, что приводит к важному понятию римановой поверхности, но это уже не сюжет действительного анализа.

Выше мы нашли разложения Тейлора функций e^x, $\sin x$ и $\cos x$ на вещественной прямой и заметили, что они сходятся при всех x. Из этого вытекает, что эти ряды сходятся и при всех комплексных z, так что мы получаем продолжения наших функций до функций e^z, $\sin z$ и $\cos z$ комплексного переменного. Эти новые функции, совпадающие со старыми, теряют многие привычные свойства. Скажем, уже неверно, что экспонента положительна, а синус по модулю не больше 1. Например, можно показать, что $e^{i\pi} = -1$. Для этого следует заметить, что из определения посредством рядов дает тождество

$$e^{iz} = \cos z + i \sin z.$$

Оно встречалось нам ранее как тождество Эйлера для вещественных значений z лишь в символическом виде, так как не были определены комплексные степени. Здесь может возникнуть вопрос, почему бы сразу не задать синус и косинус рядами (вместо нестрогих апелляций к длинам дуг)? В принципе, так сделать можно, но попробуйте при таком определении доказать, что синус периодичен и равен нулю в π.

В заключение мы покажем, что представленная степенным рядом функция бесконечно дифференцируема в интервале сходимости ряда, а ее разложение Тейлора в нуле и даст исходный ряд, так что получается некоторая завершенность описания, хотя полная картина открывается лишь при выходе в комплексную плоскость (а для функций типа корней и логарифма при выходе на римановы поверхности).

3.6.2. Предложение. *Пусть* $f(x) = \sum_{n=1}^{\infty} c_n x^n$ *при* $|x| < R$ *и* c_n *вещественны. Тогда функция* f *бесконечно дифференцируема в* $(-R, R)$ *и* $f^{(n)}(0) = n! c_n$, *т. е. исходный ряд есть ее ряд Тейлора.*

Доказательство. Достаточно проверить, что функция f один раз дифференцируема в $(-R, R)$ и

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1},$$

ибо радиус сходимости нового ряда тоже равен R в силу соотношения $n^{1/n} \to 1$. Тогда последующее дифференцирование даст ряды с общими членами вида $n(n-1)\cdots(n-k+1)c_n x^{n-k}$. Пусть $|x_0| < R$. Можно считать, что $x_0 > 0$. При достаточном малом h имеем

$$f(x_0+h) = \sum_{n=1}^{\infty} c_n (x_0+h)^n.$$

Таким образом,

$$f(x_0+h) - f(x_0) = h\sum_{n=1}^{\infty} n c_n x_0^{n-1} + r(h),$$

$$r(h) = \sum_{n=1}^{\infty} c_n\bigl((x_0+h)^n - x_0^n - n x_0^{n-1} h\bigr).$$

Нам надо проверить, что $r(h)/h \to 0$ при $h \to 0$. Для упрощения выкладок считаем, что $R > 1$ (к этому случаю легко перейти заменой $y = x(R+1)$). Тогда можно считать, что $|c_n| \leqslant 1$ при всех n, ибо $|c_n| \leqslant 1$ при достаточно больших n, а для конечных сумм утверждение верно. Кроме того, $|x_0+h| \leqslant q < 1$ при всех достаточно малых h. По формуле Тейлора

$$(x_0+h)^n - x_0^n - n x_0^{n-1} h = \frac{h^2}{2} n(n-1)\xi^{n-2}, \quad \xi \in (x_0, x_0+h).$$

Значит, $|r(h)| \leqslant |h|^2 M$, $M = \sum_{n\geqslant 1} n^2 q^{n-2}$, где ряд сходится. Поэтому $|r(h)/h| \leqslant M|h| \to 0$ при $h \to 0$. $\square$

Из формулы видно, что c_n определяются однозначно по сумме ряда. Кроме того, $f^{(k)}(x) = \sum_{n\geqslant k} n(n-1)\cdots(n-k+1)c_n x^{n-k}$.

Формула для производной позволяет просто оценить $f^{(k)}(x)$. Для упрощения формул опять будем считать, что $R > 1$, тогда $|c_k| \leqslant C$ при некотором C. Сумма $\sum_{n\geqslant k} n(n-1)\cdots(n-k+1)t^{n-k}$ равна производной порядка k от $\sum_n t^n = (1-t)^{-1}$, т. е. $k!(1-t)^{-k-1}$. Поэтому

$$|f^{(k)}(x)| \leqslant \sum_{n\geqslant k} n(n-1)\cdots(n-k+1)|c_n|\,|x|^{n-k} \leqslant Mk!(1-|x|)^{-k-1}.$$

Следовательно, при $|x_0| < R$ в достаточно малой окрестности точки x_0 имеет место разложение

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k.$$

Иначе говоря, если функция задана степенным рядом по x в круге с центром в нуле, то в окрестности всякой точки x_0 внутри этого круга ее можно переразложить по степеням $x - x_0$, используя производные в x_0.

Легко заметить, что если две функции, заданные степенными рядами в окрестности нуля, совпадают в точках $x_n \to 0$, то они равны в этой окрестности (все производные в нуле равны). Однако из того, что бесконечно дифференцируемая функция задается степенным рядом в какой-то окрестности, еще не следует, что она всюду задается степенным рядом. Например, функция $f(x) = \exp(-x^2)$, доопределенная нулем при $x \leqslant 0$, задается нулевым рядом около -1, задается ненулевым рядом около 1, но не задается рядом около нуля.

Резюмируем сказанное про разложения в степенные ряды (для упрощения в точке 0):

• пусть функция f бесконечно дифференцируема в окрестности нуля; тогда возникает формальный степенной ряд

$$\sum_{n \geqslant 0} c_n x^n, \quad c_n = f^{(n)}(0)/n!;$$

такой ряд сходится при $|x| < R$, где $1/R = \limsup_{n\to\infty} |c_n|^{1/n}$ (это верно и для произвольных чисел c_n, необязательно полученных из какой-то функции f); в частности, может сходиться лишь при $x = 0$; при $|x| > R$ ряд не сходится;

• заданная при $|x| < R$ функция $F(x) = \sum_{n\geqslant 0} c_n x^n$ бесконечно дифференцируема и $F^{(n)}(0)/n! = c_n$, т. е. она служит своим рядом Тейлора; кроме того, для каждой точки x_0 с $|x_0| < R$ верно равенство $F(x) = \sum_{n\geqslant 0} F^{(n)}(x_0)(x - x_0)^n/n!$ при достаточно малых $|x - x_0|$;

• исходная функция f может совпадать с F лишь в нуле, а для совпадения в интервале $(-\delta, \delta)$ достаточно иметь оценку

$$|f^{(n)}(x)| \leqslant Cn!\delta^{-n}.$$

Важно, что такая выполнена не в одной точке, а на целом интервале.

Позже степенные ряды будут обсуждаться при изучении функций комплексного переменного и в связи с функциональными рядами.

3.7. Дифференцируемые функции многих переменных

Здесь мы начнем обсуждение функций многих переменных, которое продолжится в следующем семестре. Стандартной метрикой на пространстве $\mathbb{R}^n$ векторов $x = (x_1, \ldots, x_n)$ с вещественными компонентами x_i является евклидова метрика

$$d(x, y) = |x - y| = \Big(\sum_{i=1}^{n} |x_i - y_i|^2\Big)^{1/2}.$$

Она порождена евклидовой нормой

$$|x| = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2},$$

которая в свою очередь порождается скалярным произведением

$$(x, y) = \sum_{i=1}^{n} x_i y_i$$

по формуле $|x| = \sqrt{(x, x)}$. Это частный вид нормы $\|\cdot\|$ на $\mathbb{R}^n$, т. е. такой функции $x \mapsto \|x\|$ со значениями в $[0, +\infty)$, что $\|x\| = 0$ только при $x = 0$, $\|\lambda x\| = |\lambda|\, \|x\|$ для всех чисел λ и всех векторов x, а также $\|x + y\| \leqslant \|x\| + \|y\|$ для всех $x, y \in \mathbb{R}^n$.

Выполнение неравенства треугольника для нормы $|x| = \sqrt{(x, x)}$ проверяется так. После возведения в квадрат приходим к неравенству $(x, y)^2 \leqslant (x, x)(y, y)$, называемому неравенством Коши. В координатах оно выглядит так:

$$\Big(\sum_{i=1}^{n} x_i y_i\Big)^2 \leqslant \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2.$$

Неравенство Коши можно доказывать разными способами. Например, можно заметить, что оно выражает неположительность дискриминанта трехчлена $(x + ty, x + ty) \geqslant 0$.

Помимо стандартной нормы на $\mathbb{R}^n$ есть континуум иных существенно иных норм, скажем, $\|x\|_\infty = \max_i |x_i|$. Однако стандартная имеет много преимуществ. Не все нормы порождаются скалярными произведениями. Кроме того, можно показать, что для всякой иной нормы $\|\cdot\|$ найдутся такие числа $c_1, c_2 > 0$, что выполнены неравенства $c_1\|x\| \leqslant |x| \leqslant c_2\|x\|$. Поэтому сходимость последовательностей по всем нормам на $\mathbb{R}^n$ сводится к сходимости по стандартной, которая очевидным образом равносильна покоординатной.

Открытым шаром с центром в x_0 радиусом $r > 0$ называется множество

$$\{x\colon |x - x_0| < r\}.$$

Замкнутый шар есть множество $\{x\colon |x - x_0| < r\}$.

Как и в общем метрическом пространстве, возникают понятия открытого и замкнутого множества. А именно: множество U открыто, если всякая его точка u входит в U с некоторым открытым шаром с центром в u.

Пустое множество тоже объявляется открытым. Замкнутые множества — дополнения открытых.

Замкнутость множества A равносильна тому, что A одержит пределы всех сходящихся последовательностей точек из A.

Определение непрерывности отображений $f\colon \mathbb{R}^n \to \mathbb{R}^k$ вводится точно так же, как на прямой или в общем метрическом пространстве.

3.7.1. Определение. *Отображение f непрерывно в точке x, если при $x_i \to x$ имеем $f(x_i) \to f(x)$.*

Как и выше, это равносильно тому, что для всякого $\varepsilon > 0$ найдется такое $\delta > 0$, что при $|x - y| < \delta$ имеем $|f(x) - f(y)| < \varepsilon$.

В силу сказанного выше свойство непрерывности не зависит от выбора метрики на $\mathbb{R}^n$, порожденной нормой. Конечно, бывают и метрики, не порождаемые нормой, которые могут привести к иным свойствам непрерывности, как, скажем, дискретная метрика, но здесь обсуждается исключительно стандартная евклидова метрика.

Непрерывная функция на $\mathbb{R}^n$ непрерывна по каждому переменному отдельно, но обратное неверно. Например, функция на плоскости, равная 0 всюду, кроме точек вида (t, t^2), в которых она равна 1, разрывна в нуле, но сужение ее на каждую прямую непрерывно. Как и для общих метрических пространств, эффективно проверяемым достаточным условием непрерывности является условие Липшица.

Как и на прямой, появляется понятие компактного множества или компакта, т. е. такого множеств, что всякое его покрытие открытыми множествами содержит конечное подпокрытие. Это свойство равносильно тому, что из всякой бесконечной последовательности в данном множестве можно выбрать подпоследовательность, сходящуюся к точке из этого множества. Для наших целей удобнее принять последнее в качестве определения.

3.7.2. Определение. *Множество $K \subset \mathbb{R}^n$ назовем компактным, если всякая бесконечная последовательность в K одержит подпоследовательность, сходящуюся к точке из K. Пустое множество тоже считаем компактным.*

По этому определению конечные множества компактны.

Как и на прямой, компакты в $\mathbb{R}^n$ — в точности ограниченные замкнутые множества.

3.7.3. Предложение. *Множество $K \subset \mathbb{R}^n$ компактно в точности тогда, когда оно ограничено и замкнуто.*

Доказательство. Обоснование совершенно аналогично случаю прямой. Если K компактно и $x_j \to x$, где $x_j \in K$, то имеется подпоследовательность, сходящаяся к точке $x' \in K$, а тогда $x = x' \in K$. Поэтому K замкнуто. Если $x_j \in K$ и $|x_j| \to +\infty$, то в $\{x_j\}$ нет сходящейся подпоследовательности. Значит, K ограничено.

Обратно, если K ограничено и замкнуто и $\{x_j\} \subset K$, то компоненты векторов x_j равномерно ограничены. Поэтому можно выбрать подпоследовательность векторов x_{j_k}, сходящуюся покоординатно к некоторому вектору x. В силу замкнутости имеем $x \in K$.

Если пользоваться определением через открытые покрытия, то также работает аналогичное случаю отрезка рассуждение. В самом деле, если множество K замкнуто и лежит в кубе, скажем, $[0,1]^n$, а открытые множества U_α покрывают K, то при отсутствии конечного подпокрытия есть меньший куб с ребром $1/2$, не покрытый никаким конечным поднабором множеств U_α. Далее строим вложенную последовательность кубов с ребрами 2^{-k}, также не покрытых никакими конечными поднаборами множеств U_α. Они стягиваются к некоторой точке a из $[0,1]^n$, что следует из теоремы о вложенных отрезках. Эта точка входит в какое-то U_α, но тогда в U_α лежат и построенные кубы при достаточно большом k. Обратно, компакт K ограничен (можно покрыть конечным набором шаров) и замкнут, ибо если K имеем предельную точку $a \notin K$, то $\mathbb{R}^n$ без a можно покрыть открытыми шарами, не содержащими a. Тогда всякий конечный набор таких шаров не пересекается с достаточно малым шаром вокруг a, поэтому не покрывает всего K. Обратно, замкнутость и ограниченность легко вывести из свойства конечного подпокрытия (оставим это в качестве упражнения). $\square$

3.7.4. Теорема. *Непрерывная функция f на компакте $K \subset \mathbb{R}^n$ ограничена и достигает максимума и минимума.*

Доказательство. Пусть $x_j \in K$ и $f(x_j) \to \sup_{x \in K} f(x)$. Перейдя к подпоследовательности, можно считать, что $x_j \to x$, где $x \in K$. Тогда $f(x_j) \to f(x)$. Значит, x — точка максимума. Аналогично с минимумом. $\square$

Основная идея производной отображений многомерных (или даже бесконечномерных) пространств — приближение линейными отображениями с точностью до «малых высших порядков». Напомним, что линейное отображение $A\colon \mathbb{R}^n \to \mathbb{R}^k$ удовлетворяет условиям

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay$$

для всех векторов x и y и всех скаляров α и β. Линейные отображения называют также линейными операторами. При $k = 1$ линейное отображение называется линейной функцией или линейным функционалом. Линейный оператор A задается матрицей $(a_{ij})_{i\leqslant k, j\leqslant n}$, столбцы которой высоты k представляют собой векторы $Ae_1, \ldots, Ae_n$, где e_i — векторы стандартного базиса в $\mathbb{R}^n$. Для $x = x_1e_1 + \cdots + x_ne_n$ имеем $Ax = x_1Ae_1 + \cdots + x_nAe_n$, поэтому вектор-столбец Ax получается умножением матрицы оператора A на вектор-столбец x (k-я компонента есть скалярное произведение k-й строки матрицы и x).

Для обычной нормы $|x|$ в $\mathbb{R}^n$ и $\mathbb{R}^k$ норма линейного оператора $A\colon \mathbb{R}^n \to \mathbb{R}^k$ задается формулой

$$\|A\| = \sup_{|x|\leqslant 1} |Ax|.$$

Это число есть минимум таких C, что $|Ax| \leqslant C|x|$ для всех x. Из определения следует, что линейный оператор A удовлетворяет условию Липшица с постоянной $\|A\|$, ибо $|Ax - Ay| = |A(x-y)| \leqslant \|A\|\,|x-y|$.

Напомним еще, что для линейного оператора $A\colon \mathbb{R}^n \to \mathbb{R}^k$ имеется сопряженный оператор $A^*\colon \mathbb{R}^k \to \mathbb{R}^n$, для которого

$$(Ax, y) = (x, A^*y), \quad x \in \mathbb{R}^n,\ y \in \mathbb{R}^k.$$

Матрица A^* получается из матрицы A транспонированием.

3.7.5. Определение. *Пусть $U \subset \mathbb{R}^n$ — открытое множество. Отображение $F\colon U \to \mathbb{R}^k$ называется дифференцируемым в точке $x \in U$, если есть такое линейное отображение $A\colon \mathbb{R}^n \to \mathbb{R}^k$, что*

$$F(x+h) - F(x) = Ah + r(x,h), \quad \text{где } r(x,h) = o(h),$$

т. е. $\lim\limits_{|h|\to 0} \dfrac{|r(x,h)|}{|h|} = 0$. *Оператор A называется производной отображения F в точке x и обозначается символами $F'(x)$, $DF(x)$.*

Обратим внимание, что $F(x+h)$ определено при всех достаточно малых h, ибо x входит в U вместе с некоторым шаром. Для прямой получаем прежнее определение. При этом называемое производной число $F'(x)$ надо рассматривать как линейную функцию $h \mapsto F'(x)h$, что дает полное соответствие с общим определением. Именно эта линейная функция приближает приращение $F(x+h) - F(x)$ с точностью до $o(h)$. С многомерной точки зрения, $F'(x) = F'(x)1$.

Из дифференцируемости в точке следует непрерывность в этой точке: при $h \to 0$ имеем $F(x+h) - F(x) \to 0$, так как $Ah \to 0$ и $r(x,h) \to 0$.

Из определения следует существование производной вдоль всякого фиксированного вектора:

$$\partial_h F(x) = \lim_{t\to 0} \frac{F(x+th) - F(x)}{t},$$

равных $F'(x)(h)$. В самом деле,

$$F(x+th) - F(x) = tAh + r(th),$$

где $|r(th)|$ есть $o(t)$, ибо $r(th)/t = |h|r(th)/|th|$.

Для базисных векторов e_i, где $e_1 = (1, 0, \ldots, 0$ и т. д., получаем существование *частных производных*

$$\partial_{x_i} F(x) = \lim_{t\to 0} \frac{F(x+te_i) - F(x)}{t}.$$

Отображение F имеет компоненты $F_1, \ldots, F_k$. Поэтому дифференцируемость F равносильна дифференцируемости в данной точке всех компонент. При этом

$$\partial_{x_i} F(x) = (\partial_{x_i} F_1(x), \ldots, \partial_{x_i} F_k(x)).$$

Из сказанного ясно, что производная $F'(x)$ определяется однозначно. Матрица производной в стандартном базисе состоит из n столбцов $F'(x)(e_i)$, имеющих k координат, являющихся частными производными $\partial_{x_i} F_j(x)$. Таким образом, $F'(x)$ задается матрицей

$$\begin{pmatrix} \partial_{x_1} F_1 & \partial_{x_2} F_1 & \ldots & \partial_{x_n} F_1 \\ \partial_{x_1} F_2 & \partial_{x_2} F_2 & \ldots & \partial_{x_n} F_2 \\ \ldots & \ldots & \ldots & \ldots \\ \partial_{x_1} F_k & \partial_{x_2} F_k & \ldots & \partial_{x_n} F_k \end{pmatrix}$$

В этой матрице n столбцов высоты k. Есть два особых случая, когда вместо матрицы удобнее говорить о векторе: $k = 1$ и $n = 1$.

Первый случай возникает, когда F — числовая функция. Тогда линейная функция $F'(x)$ задается вектором в виде $F'(x)(h) = (h, v)$. Этот вектор v называется *градиентом* функции F и обозначается символами

$$\nabla F(x) \quad \text{или} \quad \operatorname{grad} F(x).$$

Ясно, что

$$\nabla F(x) = \operatorname{grad} F(x) = (\partial_{x_1} F(x), \ldots, \partial_{x_n} F(x)), \ \partial_h F(x) = (\nabla F(x), h).$$

Итак, производная k-мерного отображения $F = (F_1, \ldots, F_k)$ задается матрицей со строками $\nabla F_1, \ldots, \nabla F_k$. Из определения следует, что дифференцируемость F в точке равносильна дифференцируемости в этой точке всех компонент F_i, поэтому проверка дифференцируемости в принципе сводится к указанному первому случаю.

Второй особый случай: $n = 1$, т. е. отображение из прямой в многомерное пространство. Здесь тоже обычно удобнее считать, что

$$F'(t) = \bigl(F_1'(t), \ldots, F_k'(t)\bigr)$$

есть вектор в $\mathbb{R}^k$, хотя формально надо говорить о задаваемом им линейном отображении $h \mapsto hF'(t)$.

Следует иметь в виду, что существование частных производных или даже производных по всем направлениям еще не означает дифференцируемость.

3.7.6. Пример. Зададим функцию f на $\mathbb{R}^2$ так: $f(x,y) = 1$, если $y = x^2$ и $x > 0$, $f(x,y) = 0$ в остальных точках. Тогда $\partial_h f(0,0) = 0$ для всех векторов h, ибо при фиксированном h имеем $f(th_1, th_2) = 0$ для всех достаточно малых t. Однако $f(t,t^2) = 1$ при всех $t > 0$, что даже не стремится к нулю при $t \to 0$.

Простейшим дифференцируемым отображением является постоянная $F(x) = c$. Ясно, что $F'(x) = 0$ для всех x.

Другой простой пример — отображение вида

$$F(x) = Ax + c,$$

где A — линейный оператор и c — постоянный вектор. Для него $F'(x) = A$ для всех x, ибо $A(x+h) - Ax = Ah$, так что $r(x,h) = 0$. Таким образом, производная линейного отображение есть само это отображение. Здесь не возникает противоречия с привычным свойством линейных функций иметь в качестве производной константу: в многомерном случае это означает, что производная одна и та же во всех точках.

Рассмотрим еще такой пример (квадратичная форма):

$$F(x) = (Qx, x),$$

где Q — симметричный линейный оператор, т. е. $Q^* = Q$. Тогда

$$(Qx + Qh, x + h) = (Qx, x) + 2(Qx, h) + (Qh, h).$$

Поскольку $(Qh,h)=o(h)$ в силу оценки $|(Qh,h)|\leqslant \|Q\|\,|h|^2$, получаем

$$F'(x)(h)=2(Qx,h),\quad \nabla F(x)=2Qx.$$

Из определения очевидно, что если отображения F и G дифференцируемы в точке x, то их сумма тоже дифференцируема и

$$D(F+G)(x)=DF(x)+DG(x).$$

3.7.7. Теорема. *Если отображение $F\colon U\to \mathbb{R}^k$ дифференцируемо в точке $x\in U$ и принимает значения в открытом множестве $V\subset \mathbb{R}^k$, а отображение $G\colon V\to\mathbb{R}^m$ дифференцируемо в точке $F(x)$, то композиция $G\circ F$ дифференцируема в x, причем*

$$(G\circ F)'(x)=G'(F(x))F'(x).$$

Иначе говоря, $(G\circ F)'(x)(h)=G'(F(x))(F'(x)(h))$.

Доказательство. Положим $y=F(x)$. Мы имеем

$$F(x+h)=F(x)+F'(x)(h)=r_1(x,h),$$

$$G(y+u)=G(y)+G'(y)(u)+r_2(y,u),$$

где $r_1(x,h)=o(h)$, $r_2(y,u)=o(u)$. Поэтому

$$\begin{aligned} &G(F(x+h))-G(F(x))=\\ &=G'(y)(F(x+h)-F(x))+r(y,F(x+h)-F(x))=\\ &=G'(y)(F'(x)(h))+G'(y)r(x,h)+r(y,F(x+h)-F(x)).\end{aligned}$$

При этом $G'(y)r(x,h)=o(h)$, ибо

$$|G'(y)r(x,h)|\leqslant \|G'(y)\|\,|r(x,h)|.$$

Наконец, $r(y,F(x+h)-F(x))=o(h)$, ибо

$$\frac{|r(y,F(x+h)-F(x))|}{|h|}=\frac{|r(y,F(x+h)-F(x))|}{|F(x+h)-F(x)|}\frac{F(x+h)-F(x)}{|h|}$$

при $F(x+h)\neq F(x)$. В этом отношении

$$|F(x+h)-F(x)|/|h|\leqslant \|F'(x)\|+|r_1(x,h)|/|h|,$$

что остается ограниченным при $h\to 0$. Отношение

$$|r(y,F(x+h)-F(x))|/|F(x+h)-F(x)|$$

стремится к нулю при $h\to 0$, так как $F(x+h)-F(x)\to 0$. $\square$

3.7.8. Пример. Если G линейно, то $(G \circ F)'(x) = GF'(x)$.
Если F линейно, то $(G \circ F)'(x) = G'(F(x))F$.
Если A и B — линейные операторы, то

$$(AFB)'(x) = AF'(Bx)B.$$

3.7.9. Пример. (i) Для функции $f(x) = |x|$ вне нуля имеем

$$\nabla |x| = \frac{x}{|x|}.$$

В самом деле: $\partial_{x_i}(x_1^2 + \cdots + x_n^2)^{1/2} = x_i(x_1^2 + \cdots + x_n^2)^{-1/2}$.

(ii) Для $f(x) = \varphi(|x|)$, где функция φ дифференцируема на прямой, имеем

$$\nabla f(x) = \varphi'(|x|)\frac{x}{|x|}.$$

(iii) Если функция f на $\mathbb{R}^n$ дифференцируема и A — линейный оператор в $\mathbb{R}^n$, то

$$\nabla (f \circ A)(x) = A^* \nabla f(Ax).$$

Это видно из формулы композиции, так как $(f \circ A)'(x) = f'(Ax)A$, поэтому

$$(\nabla (f \circ A)(x), h) = (f \circ A)'(x)(h) = f'(Ax)(Ah) = (\nabla f(Ax), Ah) = \\ = (A^* \nabla f(Ax), h).$$

Аналогом теоремы о среднем является такое утверждение. Отрезком $[a, b]$ в $\mathbb{R}^n$ называется множество точек вида $ta + (1-t)b$, где $t \in [0, 1]$.

3.7.10. Предложение. *Пусть отображение F дифференцируемо в открытом множестве $U \subset \mathbb{R}^n$, $[a, b] \subset U$, причем при некотором M*

$$\|F'(c)\| \leqslant M \quad \text{\textit{для всех} } c \in [a, b].$$

Тогда

$$|F(b) - F(a)| \leqslant M|b - a|.$$

Доказательство. Возьмем единичный вектор u, для которого $|F(b) - F(a)| = (F(b) - F(a), u)$. Рассмотрим функцию

$$f(t) = (F(ta + (1-t)b), u), \quad t \in [0, 1].$$

По предыдущей теореме эта функция дифференцируема на $[0, 1]$, причем

$$f'(t) = (F'(ta + (1-t)b)(a - b), u),$$

так как линейное отображение $t \mapsto ta + (1-t)b$ из $\mathbb{R}$ в $\mathbb{R}^n$ имеет производную $a-b$, а линейная функция $y \mapsto (y, u)$ имеет производную $h \mapsto (h, u)$. Значит, $|f'(t)| \leqslant M|b-a|$. По теореме о среднем

$$|F(b) - F(a)| = |f(0) - f(1)| = |f'(\xi)|$$

при некотором $\xi \in [0, 1]$. □

Это предложение слабее одномерной теоремы, в которой

$$f(b) - f(a) = f'(\xi)(b-a)$$

для некоторой точки $\xi \in (a, b)$. В такой формулировке многомерный вариант неверен. Например, для

$$f(t) = (\cos t, \sin t)$$

имеем $f(0) = f(2\pi)$, но $f'(t) = (-\sin t, \cos t)$ не обращается в нуль. Однако есть более тонкий многомерный вариант:

$F(b) - F(a)$ лежит в замкнутой выпуклой оболочке векторов вида $F'(\xi)(b-a)$, где $\xi \in [a, b]$, т. е. в наименьшем замкнутом выпуклом множестве, содержащем эти векторы.

3.7.11. Следствие. *Если $F'(x) = 0$ для всех x из шара U, то F постоянно в U.*

Если же $F'(x) = A$ в U, где A — линейный оператор, не зависящий от x, то $F(x) = Ax + c$ в U, для некоторого постоянного вектора c.

Доказательство. Первое утверждение очевидно из теоремы. Второе следует из первого, ибо $F(x) - Ax$ имеет нулевую производную. □

3.7.12. Следствие. *Если $\|F'(x)\| \leqslant M$ в шаре U, то F на U удовлетворяет условию Липшица с постоянной M, т. е. выполнено неравенство $|F(x) - F(y)| \leqslant M|x-y|$.*

Отметим еще, что из доказательства теоремы о среднем видно, что если имеется лишь оценка

$$|\partial_{x_i} F(x)| \leqslant M, \quad x \in [a, a + e_i]$$

для одной частной производной, то выполнено более слабое неравенство

$$|F(a + te_i) - F(a)| \leqslant M|t|, \ t \in [0, 1].$$

Если на шаре ограничены все частные производные $\partial_{x_i} F(x)$, то отображение F липшицево на этом шаре, хотя и не обязано быть всюду дифференцируемым (впрочем, можно показать, что оно обязано иметь много точек дифференцируемости).

Большое практическое значение имеет следующий факт, отчасти реабилитирующий частные производные.

3.7.13. Теорема. *Если отображение* $F\colon \mathbb{R}^n \to \mathbb{R}^k$ *в шаре* U *имеет непрерывные частные производные* $\partial_{x_i} F$*, то* F *дифференцируемо в* U.

Доказательство. Как отмечено выше, достаточно рассмотреть случай $k = 1$. Покажем, что F дифференцируемо в $x_0 \in U$. Пусть $c_i = \partial_{x_i} F(x_0)$. Линейная функция $c_1x_1 + \cdots + c_nx_n$ дифференцируема и имеет в x_0 такие же частные производные. Вычитая ее из F, можно считать, что $c_i = 0$. Кроме того, вычитая $F(x_0)$, можно считать, что $F(x_0) = 0$. Наконец, можно считать, что $x_0 = 0$. Проверим, что $F'(0) = 0$. Пусть $\varepsilon > 0$. В силу непрерывности частных производных найдется такое $r > 0$, что $|\partial_{x_i} F(x)| \leqslant \varepsilon n^{-1}$ при $|x| < r$. Согласно сказанному выше о теореме о среднем для частных производных

$$|F(x + te_i) - F(x)| \leqslant \varepsilon n^{-1} \quad \text{при } |x| < r/2, |t| < r/2.$$

Поэтому при $|h| < r/2$ имеем

$$\begin{aligned}
&|F(h_1e_1 + \cdots + h_ne_n)| = \\
&\qquad = |F(h_1e_1 + \cdots + h_ne_n) - F(h_1e_1 + \cdots + h_{n-1}e_{n-1}) + \\
&+ F(h_1e_1 + \cdots + h_{n-1}e_{n-1}) - F(h_1e_1 + \cdots + h_{n-2}e_{n-2}) + \cdots + F(h_1e_1)| \leqslant \\
&\qquad\qquad \leqslant \varepsilon n^{-1}(|h_n| + \cdots + |h_1|) \leqslant \varepsilon |h|.
\end{aligned}$$

Это означает, что $F(h) = o(h)$ при $h \to 0$, т. е. $F'(0) = 0$. □

Важную роль играют обратимые дифференцируемые отображения. Если отображение $F\colon \mathbb{R}^n \to \mathbb{R}^n$ отображает открытое множество U взаимно однозначно на открытое множество V, причем F и F^{-1} имеют непрерывные производные, то F называют диффеоморфизмом. В следующем семестре будет доказано, что необходимое и достаточное условие того, что взаимно однозначное отображение F с непрерывной производной — диффеоморфизм, состоит в том, что его производная в каждой точке — обратимый оператор. Более того, если отображение F имеет непрерывную производную $F'(x)$, причем в некоторой точке x_0 оператор $F'(x_0)$ обратим, то F диффеоморфно отображает некоторый шар с центром в x_0 на окрестность точки $F(x_0)$.

Можно показать, что не бывает диффеоморфизмов между пространствами разных размерностей.

3.8. Задачи

3.8.1. Найти производные функций

(i) $\operatorname{tg}^2 x + \ln\cos x$, (ii) $\sqrt{x+\sqrt{x}}$, (iii) $\ln\dfrac{e^x+2+\sqrt{e^{2x}+e^{4x}+1}}{e^x+2-\sqrt{e^{2x}+e^{4x}+1}}$, (iv) $\operatorname{arctg} e^{\sin x}$, (v) $\ln\dfrac{\sin x+1}{x^2+1}$.

3.8.2. Найти пределы (i) $\lim\limits_{x\to 0}\dfrac{\ln\cos\alpha x}{\ln\cos\beta x}$, (ii) $\lim\limits_{x\to 0}\dfrac{e^x-e^{-x}-2x}{x-\sin x}$, (iii) $\lim\limits_{x\to\pi/4}\dfrac{\operatorname{tg} x-1}{\sin(4x)}$, (iv) $\lim\limits_{x\to 0}\dfrac{\ln^2(1+x)-\sin^2 x}{1-e^{-x^2}}$, (v) $\lim\limits_{x\to\pi/2}(\sin x)^{\operatorname{tg} x}$, (vi) $\lim\limits_{x\to 1}\Bigl(\dfrac{\pi}{2}\operatorname{tg}(\pi x/2)-(1-x)^{-1}\Bigr)$.

3.8.3. Доказать теорему 3.4.4 по индукции, используя теорему о среднем при индуктивном переходе.

3.8.4. Может ли функция на прямой быть дифференцируемой только в точке 0 и точках $1/n$, где $n\in\mathbb{N}$?

3.8.5. Построить графики функций $x\ln x$, xe^{-x}, x^2e^{-x}.

3.8.6. Во всех ли точках отрезка $[0,1]$ дифференцируема канторовская лестница?

3.8.7. Пусть функция f непрерывна на отрезке $[0,1]$ и дифференцируема в $(0,1)$, причем $|f'(c)|\leqslant|f(1)-f(0)|$ для всех $c\in(0,1)$. Доказать, что f имеет вид $f(x)=kx+C$.

3.8.8. В эллипс $(x/a)^2+(y/b)^2=1$ вписать прямоугольник наибольшей площади, стороны которого параллельны осям эллипса.

3.8.9. Найти минимум функции $A/\sin x+B/\cos x$ на $[0,\pi/2]$, где $A>B>0$. Использовать эту задачу для ответа на вопрос: какова наибольшая длина соломинки, которую муравей может горизонтально пронести через длинный подземный ход с шириной 10 см, поворачивающий перпендикулярно и после поворота переходящий в длинный переход с шириной 5 см?

3.8.10. Исследовать на выпуклость и вогнутость функцию $x\ln x$ на луче $(0,\infty)$.

3.8.11. Доказать, что для выпуклой на промежутке J функции f при всех $x_1,\ldots,x_n\in J$ и $t_1,\ldots,t_n\in[0,1]$ с $\sum_{i=1}^n t_i=1$ верно неравенство

$$f(t_1x_1+\cdots+t_nx_n)\leqslant t_1f(x_1)+\cdots+t_nf(x_n).$$

3.8.12. (i) Доказать, что непрерывная на (a,b) функция f выпукла в точности тогда, когда для всех $x,y \in (a,b)$ верно неравенство

$$f\Bigl(\frac{x+y}{2}\Bigr) \leqslant \frac{f(x)+f(y)}{2}.$$

Используя задачу 2.5.22, показать, что от непрерывности отказаться нельзя.

(ii) Вывести из (i), что непрерывная на (a,b) функция f выпукла в точности тогда, когда ее дискретная вторая производная неотрицательна, т. е.

$$\frac{f(x+h)+f(x-h)-2f(x)}{h^2} \geqslant 0$$

при всех $x \in (a,b)$ и всех $h \neq 0$, для которых $x+h \in (a,b)$.

3.8.13. Пусть функция f непрерывна на прямой, выпукла на полупрямых $(-\infty,0]$ и на $[0,+\infty)$. Верно ли, что она выпукла на всей оси?

3.8.14. (i) Доказать, что всякая выпуклая функция непрерывна.

(ii) Доказать, что выпуклая на интервале функция на всяком внутреннем отрезке удовлетворяет условию Липшица.

(iii) Пусть функция f выпукла на $[a,b]$. Доказать, что она либо монотонна на всем отрезке, либо найдется такая точка $c \in [a,b]$, что f монотонна на $[a,c]$ и $c,b]$.

3.8.15. (i) Найти радиусы сходимости степенных рядов с коэффициентами $n!$, 2^n, $n\ln n$, $3^n/n!$, $(n!)^2/(2n)!$.

(ii) Найти радиусы сходимости степенных рядов $\sum_{n=1}^\infty a^{n^2}x^n$ и $\sum_{n=1}^\infty a^n x^{n^2}$.

3.8.16. Пусть f — выпуклая функция на (a,b), g — возрастающая выпуклая функция на промежутке, содержащем множество значений f. Доказать, что композиция $g\circ f$ также выпукла.

3.8.17. Пусть функция f всюду дифференцируема в $(0,1)$. Доказать, что f' имеет точки непрерывности.

3.8.18. (i) Построить пример такой всюду дифференцируемой функции на прямой, что ее производная разрывна в точках всюду плотного множества.

(ii) Построить пример такой всюду дифференцируемой функции на прямой, что ее производная разрывна в точках континуального множества.

3.8.19. Разложить по степеням x функции (i) $\sqrt{1+x}$, (ii) $\ln\dfrac{1+x}{1-x}$, (iii) $\ln(2-3x+x^2)$. (iv) $\operatorname{arctg} x$, (v) $\arcsin x$, (vi) $\ln\bigl((1+x)/(1-x)\bigr)$.

3.8.20. Найти предел

$$\lim_{x\to 0}\frac{\sin(tgx)-\operatorname{tg}(\sin x)}{\operatorname{arctg}(\arcsin x)-\arcsin(\operatorname{arctg} x)}.$$

3.8.21. Разложив e^{e^x} двумя способами, доказать равенство

$$\frac{1^3}{1!}+\frac{2^3}{2!}+\frac{3^3}{3!}+\cdots+\frac{n^3}{n!}+\cdots=5e.$$

3.8.22. Доказать тождество

$$\frac{\sin x}{1}+\frac{\sin 3x}{3}+\cdots+\frac{\sin(2n-1)x}{2n-1}+\cdots=\frac{\pi}{4},\quad x\in(0,\pi).$$

3.8.23. Пусть f — дифференцируемая функция на прямой. Доказать, что множество таких точек x, что $f(x)=0$, но $f'(x)\neq 0$, не более чем счетно.

3.8.24. Пусть функция f дифференцируема $n+1$ раз в окрестности точки x_0 и $f^{(n+1)}(x_0)\neq 0$. Тогда для всякого достаточно малого $h>0$ по теореме Тейлора найдется число $\theta(h)\in(0,1)$ (единственное из-за монотонности $f^{(n)}$ в окрестности x_0), для которого

$$f(x_0+h)-f(x_0)=f'(x_0)h+\cdots+\frac{f^{(n-1)}(x_0)h^{n-1}}{(n-1)!}+\frac{f^{(n)}(x_0+\theta(h)h)h^n}{n!}.$$

Доказать, что $\lim\limits_{h\to 0}\theta(h)=\dfrac{1}{n+1}$.

3.8.25. Пусть $f(x)=x\sin\ln x$ при $x\neq 0$, $f(0)=0$. Доказать, что в теореме о среднем $f(x)-f(0)=xf'(\xi(x))$ нельзя выбрать точку $\xi(x)$ непрерывно зависящей от x на $(0,1)$.

3.8.26. Пусть функция f на прямой бесконечно дифференцируема и для каждой точки x найдется номер производной n_x, для которого $f^{(n_x)}(x)=0$. Доказать, что f является многочленом.

3.8.27. (Задача Теренса Тао) Существует ли такая бесконечно дифференцируемая функция f на прямой, что ее ряд Тейлора

$$\sum_{n=0}^{\infty} f^{(n)}(x_0)(x-x_0)^n/n!$$

сходится при всех x для каждой точки x_0, но ни на каком интервале заданная рядом функция не совпадает с f?

3.8.28. Разобрать детали следующего обоснования иррациональности числа π, определяемого как наименьший корень $x > 0$ уравнения $\sin x = 0$.

Пусть $\pi = a/b$, где a и b натуральны. При $n \in \mathbb{N}$ положим

$$A_n(b) = b^n \int_0^\pi \frac{x^n(\pi - x)^n}{n!} \sin x \, dx.$$

По определению функция под интегралом положительна на интервале, поэтому $A_n(b) > 0$. Кроме того, функция под интегралом оценивается через $C^n/n!$, поэтому $A_n(b) \to 0$ при $n \to \infty$.

С другой стороны, введя многочлен степени $2n$ по формуле

$$f(x) = \frac{x^n(a - bx)^n}{n!},$$

после $2n + 1$ интегрирований по частям получим равенство

$$\begin{aligned} A_n(b) &= \int_0^\pi f(x) \sin x \, dx \\ &= \Big[-f(x) \cos x \Big]_{x=0}^{x=\pi} - \Big[-f'(x) \sin x \Big]_{x=0}^{x=\pi} + \cdots \\ &\pm \Big[f^{(2n)}(x) \cos x \Big]_{x=0}^{x=\pi} \pm \int_0^\pi f^{(2n+1)}(x) \cos x \, dx. \end{aligned}$$

Последний интеграл равен нулю, ибо $f^{(2n+1)} = 0$, так как f — многочлен степени $2n$. Синус и косинус в точках 0 и π принимают целые значения. В предположении $\pi = a/b$ это же верно и для $f^{(k)}$ при $0 \leqslant k \leqslant 2n$, так как при наличии обоих множителей x и $a - bx$ в производной получаем нули, а отсутствие одного из множителей означает, что $k \geqslant n$, так что факториал в знаменателе сократился. Итак, $A_n(b)$ является положительным целым числом. Это противоречит стремлению к нулю.

3.8.29. Пусть функция f на $\mathbb{R}^2$ такова, что функции $x \mapsto f(x, y)$ непрерывны и $|\partial_y f(x, y)| \leqslant 1$. Доказать, что она непрерывна.

3.8.30. Пусть функция f на $\mathbb{R}^n$ является многочленом по каждому переменному при фиксированных остальных. Доказать, что она многочлен.

3.8.31. Для $z \in \mathbb{C}$ пусть $f(z) = \exp(-z^{-4})$ при $z \neq 0$ и $f(0) = 0$. Доказать, что f как функция на $\mathbb{R}^2$ имеет все частные производные

$\partial_x^n \partial_y^k f$ на всей плоскости, но в нуле разрывна, поэтому не дифференцируема.

3.8.32. Пусть функция f на $\mathbb{R}^2$ имеет всюду частные производные $\partial f/\partial x$, $\partial f/\partial y$. Доказать, что в некоторой точке она дифференцируема.

3.8.33. (i) Доказать, что прямую можно непрерывно отобразить на плоскость. (ii) Доказать, что прямую нельзя отобразить на плоскость посредством липшицева отображения.

Литература

[1] Арнольд В.И. Цепные дроби. Изд-во МЦНМО, М., 2009; 40 с.

[2] Верещагин Н.К., Шень А. Лекции по математической логике и теории алгоритмов. Часть 1. Начала теории множеств. МЦНМО, М., 1999; 128 с.

[3] Гельбаум Б., Олмстед Дж. Контрпримеры в анализе. Мир, М., 1967; 252 с.

[4] Гюнтер Н.М., Кузьмин Р.О. Сборник задач по высшей математике. Т. 1 – 2. 13-е изд. Т. 3. 4-е изд. ГИФМЛ, М., 1958, 1951; 284 с., 286 с., 268 с. (переиздан в 2003 г., изд. Лань, СПб.).

[5] Демидович Б.П. Сборник задач и упражнений по математическому анализу. 13-е изд., испр. Изд-во Моск. ун-та ЧеРо, М., 1997; 624 с.

[6] Зорич В.А. Математический анализ: В 2 т. Наука, М., 1981, 1984; 544 с., 640 с.

[7] Кановей В.Г. Аксиома выбора и аксиома детерминированности. Наука, М., 1984; 65 с.

[8] Коблиц Н. p-адические числа, p-адический анализ и дзета функции. Мир, М., 1982; 192 с.

[9] Кудрявцев Л.Д. Курс математического анализа. Т. 1–3. Дрофа, М., 2003, 2004, 2006; 704 с., 720 с., 351 с.

[10] Кудрявцев Л.Д. и др. Сборник задач по математическому анализу. Т. 1–3. 2-е изд. Физматлит, М., 2003; 496 с., 505 с., 473 с.

[11] Львовский С.М. Лекции по математическому анализу. МЦНМО, М., 2008; 296 с.

[12] Ляшко И.И., Боярчук А.К., Гай Я.Г., Головач Г.П. Справочное пособие по высшей математике (АнтиДемидович). Т. 1. Ч. 1, 2. Изд-во ЛКИ, М., 2007; 240 с., 224 с.

[13] Натансон И.П. Теория функций вещественной переменной. 3-е изд. Наука, М., 1974; 480 с.

[14] Нестеренко Ю.В. Теория чисел. Изд. центр «Академия», М., 2008; 272 с.

[15] Никольский С.М. Курс математического анализа. Т. 1, 2. 3-е изд. Наука, М., 1983; 468 с., 451 с.

[16] Никольский С.М. Курс математического анализа. 6-е изд., ФИЗМАТЛИТ, М., 2001; 592 с.
[17] Полиа Г., Сеге Г. Задачи и теоремы из анализа. Т. 1, 2. 3-е изд. Наука, М., 1978; 392 с., 432 с.
[18] Рудин У. Основы математического анализа. Мир, М., 1976; 320 с.
[19] Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. Т. 1, 2, 3. 7-е изд. Наука, М., 1970.
[20] Хавин В.П. Основы математического анализа. Дифференциальное и интегральное исчисление функций одной вещественной переменной. Лань, СПб., 1998; 448 с.
[21] Хинчин А.Я. Цепные дроби. Наука, М., 1978; 112 с.
[22] Шилов Г.Е. Математический анализ. Функции нескольких вещественных переменных. Ч. 1–2. М., Наука, 1972; 624 с.
[23] Rudin W. The way I remember it. Amer. Math. Soc., Rhode Island, Providence, 1996; 191 p.

Программа коллоквиума

1. Множество действительных чисел. Существование точной верхней грани и точной нижней грани ограниченного множества.
2. Принцип вложенных отрезков. Равномощность вещественной прямой множеству последовательностей из нулей и единиц и ее несчетность. Принцип конечного подпокрытия отрезка интервалами.
3. Предел последовательности. Операции над сходящимися последовательностями.
4. Сходимость ограниченной монотонной последовательности. Число e.
5. Частичные пределы. Теорема Больцано – Вейерштрасса об ограниченных последовательностях.
6. Фундаментальные последовательности. Ограниченность фундаментальной последовательности и фундаментальность сходящейся. Теорема Коши о сходимости фундаментальных последовательностей.
7. Сходимость числового ряда, абсолютная сходимость, сходимость абсолютно сходящегося ряда. Расходимость гармонического ряда и сходимость ряда Лейбница.
8. Равносильность сходимости ряда с убывающими к нулю членами a_n сходимости ряда из $2^k a_{2^k}$. Примеры применения.
9. Открытые и замкнутые множества. Строение открытых множеств на прямой.
10. Непрерывность функции в точке: равносильность определений Гейне и Коши. Предел функции в точке. Точки разрыва монотонных функций.
11. Основные свойства непрерывных функций (сложение, умножение, частное, композиция).
12. Существование максимума и минимума непрерывной функции на замкнутом ограниченном множестве. Теорема о промежуточных значениях для непрерывной функции на отрезке. Непрерывность обратной к строго монотонной непрерывной функции.
13. Производная функции. Основные свойства (дифференцируемость сумм, произведений, частных, композиций).
14. Теорема Ферма о точках экстремума. Теорема Ролля. Теоремы о среднем Лагранжа и Коши.
15. Правило Лопиталя.

Программа семестра

1. Множество действительных чисел. Существование точной верхней грани и точной нижней грани ограниченного множества.
2. Принцип вложенных отрезков. Равномощность вещественной прямой множеству последовательностей из нулей и единиц и ее несчетность. Принцип конечного подпокрытия отрезка интервалами.
3. Предел последовательности. Операции над сходящимися последовательностями.
4. Сходимость ограниченной монотонной последовательности. Число e.
5. Частичные пределы. Теорема Больцано – Вейерштрасса об ограниченных последовательностях.
6. Фундаментальные последовательности. Ограниченность фундаментальной последовательности и фундаментальность сходящейся. Теорема Коши о сходимости фундаментальных последовательностей.
7. Сходимость числового ряда, абсолютная сходимость, сходимость абсолютно сходящегося ряда. Расходимость гармонического ряда и сходимость ряда Лейбница.
8. Равносильность сходимости ряда с убывающими к нулю членами a_n сходимости ряда из $2^k a_{2^k}$. Примеры применения.
9. Открытые и замкнутые множества. Строение открытых множеств на прямой.
10. Непрерывность функции в точке: равносильность определений Гейне и Коши. Предел функции в точке. Точки разрыва монотонных функций.
11. Основные свойства непрерывных функций (сложение, умножение, частное, композиция).
12. Существование максимума и минимума непрерывной функции на замкнутом ограниченном множестве. Теорема о промежуточных значениях для непрерывной функции на отрезке. Непрерывность обратной к строго монотонной непрерывной функции.
13. Производная функции. Основные свойства (дифференцируемость сумм, произведений, частных, композиций).
14. Теорема Ферма о точках экстремума. Теорема Ролля. Теоремы о среднем Лагранжа и Коши.
15. Правило Лопиталя.
16. Производная композиции и обратной функции.

17. Формула Тейлора для дважды дифференцируемой функции. Формула Тейлора для n-кратно дифференцируемой функции. Формула Тейлора для $n+1$-кратно дифференцируемой функции с остаточным членом Лагранжа.
18. Исследование локальных экстремумов в терминах двух первых производных.
19. Выпуклые и вогнутые функции. Условия выпуклости в терминах второй производной. Примеры.
20. Разложения Тейлора для e^x, $\sin x$, $\cos x$, $\ln(1+x)$, $(1+x)^\alpha$. Степенные ряды.
21. Метрические пространства. Пространство $\mathbb{R}^n$. Непрерывные отображения, равносильные описания непрерывности. Условие Липшица.
22. Компактные множества в $\mathbb{R}^n$. Непрерывные отображения компактов.
23. Дифференцируемые функции многих переменных. Производная, градиент. Теорема о среднем.
24. Дифференцируемость функции с непрерывными частными производными.

Вопросы экзамена

1. Предел последовательности. Компактные множества на прямой.
2. Производная функции на прямой.
3. Теорема об обратной функции для дифференцируемой функции. Производная обратной функции.
4. Выпуклые и вогнутые функции. Характеризация через знак второй производной для дважды дифференцируемых функций. Примеры.
5. Формула Тейлора для дважды дифференцируемой функции.
6. Общая формула Тейлора с остаточным членом в форме Лагранжа.
7. Степенные ряды. Разложения Тейлора для e^x, $\sin x$, $\cos x$, $\ln(1+x)$.
8. Пространство $\mathbb{R}^n$, его скалярное произведение, норма и метрика. Сходимость в $\mathbb{R}^n$. Непрерывные отображения. Условие Липшица. Линейные операторы в $\mathbb{R}^n$ и их матрицы. Липшицевость линейных операторов в $\mathbb{R}^n$.
9. Компактные множества в $\mathbb{R}^n$. Существование максимума и минимума непрерывной функции на компакте.
10. Дифференцируемость функций и отображений многих переменных. Производная, градиент. Производная композиции дифференцируемых отображений.

11. Теорема о среднем для вещественной функции на $\mathbb{R}^n$ и для отображения из $\mathbb{R}^n$ в $\mathbb{R}^k$.
12. Дифференцируемость функции с непрерывными частными производными.