ПРЕДИСЛОВИЕ

Эта книга представляет собой первую часть курса математического анализа, читаемого на математико-механическом факультете Санкт-Петербургского государственного университета. Объем и содержание первой части примерно соответствуют материалу, традиционно входящему в первый семестр пятисеместрового или четырехсеместрового (с отдельным семестровым курсом теории функций комплексной переменной) курса. В нее вошли следующие разделы: введение, теория пределов и непрерывные функции, дифференциальное исчисление функций одной вещественной переменной, неопределенный интеграл.

Несмотря на огромное количество существующих разнообразных курсов математического анализа, мы считаем, что этот учебник будет полезен студентам. Мы надеемся, что наш учебник позволит студентам в основном освоить теоретический курс анализа за сравнительно небольшой срок, так сказать, “в боевой обстановке”.

Это пожелание накладывает некоторые условия на стиль учебника.

Курс не должен быть “неподъемным” для студента, слишком раздутым за счет красивых примеров, приложений, обобщений, исторических замечаний и многочисленных пояснений. Классический трехтомный “Курс дифференциального и интегрального исчисления” Г. М. Фихтенгольца, послуживший основой многих читавшихся курсов, написан очень подробно и содержит необычайно много примеров и приложений. Это достоинство является в то же время и недостатком. Кроме того, по языку он уже довольно архаичен, местами изложение недостаточно строго, и многолетний опыт показал, что многие разделы целесообразнее рассказывать по-другому.

Вместе с тем, курс должен быть по возможности строгим, содержать все необходимые определения и доказательства. Неизбежные “вольности”, на наш взгляд, должны быть сконцентрированы во введении, а после этого логические пробелы в курсе крайне нежелательны. Мы подробно разъясняем детали доказательств и особенно определений, которые часто вызывают трудности у студентов.

Можно сказать, что учебник написан в жанре подробного конспекта лекций.

В первой части уровень общности изложения невысокий; мы ограничиваемся вещественнозначными функциями вещественной переменной. Такой "медленный старт" имеет как достоинства, так и недостатки: с одной стороны, "среднему" студенту легче усвоить теорию в знакомой ситуации, а с другой — с математической точки зрения естественнее все-таки рассказывать теорию пределов в метрических пространствах. И, конечно, необходимым условием такого порядка изложения является наличие достаточного времени (повторим, что эта часть рассчитана на первый семестр пятисеместрового курса).

Каждый лектор строит курс по-своему, так что многие вопросы мы изложили просто по своему вкусу.

Введение в курс (глава 1) достаточно короткое, его основная задача — ввести общеупотребительные термины и дать необходимый минимум предварительных сведений. Подход к теории множеств — "наивный". Из элементов математической логики используется только символика, а кванторы употребляются как стенографические значки. В тех случаях, когда мы приводим сокращенную запись утверждений с кванторами, мы также расшифровываем ее словесно, чтобы студент научился и записывать утверждения, и читать собственные записи. Вещественные числа определяются аксиоматически. Построение множества вещественных чисел в курсе анализа мы считаем излишним, а времени, потраченного в первом семестре, не хватит в четвертом или пятом. Комплексные числа пока что лишь упоминаются. Особняком стоит § 4 введения, где доказываются простейшие свойства счетных множеств.

Теория пределов (глава 2) начинается с более наглядного понятия предела числовой последовательности, и лишь потом (в § 3) изучается предел функции. В первой части курса не упоминается теорема Гейне—Бореля; при таком построении курса ее место в разделе про n-мерное пространство. Мы считаем, что необходимо дать строгое определение степенной и показательной функций и предпочтительно пока что пользоваться "школьным" геометрическим определением синуса и косинуса.

Дифференциальное исчисление (глава 3) строится вокруг основного понятия — многочлена Тейлора, которое предшествует понятию производной. При формулировке глобального варианта формулы Тейлора мы ограничиваемся лагранжевой формой остатка. Отметим удобные технические приемы, связанные с модификацией критерия дифференцируемости и леммы о трех хордах. Последний параграф (§ 8) про выпуклые функции содержит довольно много материала, не всегда входящего в курс: выпуклость и опорные прямые, выпуклость и асимптоты, классические неравенства. Неравенство Иенсена доказывается самым естественным способом — с помощью опорной прямой. Эти вопросы не только интересны сами по себе и иллюстрируют приложения дифференциального исчисления, но без них обсуждение понятия выпуклости было бы неполным.

Из главы 4 про интеграл в первую часть мы посчитали нужным включить лишь первый параграф о первообразной и неопределенном интеграле.

Книга не содержит никакой специально подобранной коллекции задач, но иногда дополнительные сведения сообщаются читателю без доказательства.

Нумерация теорем и лемм отдельная в каждом параграфе; нумерация формул отдельная в каждой главе; нумерация следствий и замечаний отдельная к каждому утверждению или группе утверждений, к которым эти следствия и замечания относятся. Конец доказательства обозначается символом $\square$.

Авторы благодарны А. Н. Подкорытову и А. А. Флоринскому, высказавшим множество ценных замечаний по тексту рукописи, а также всем коллегам по кафедре математического анализа Санкт-Петербургского государственного университета, чьи методические находки использовались в этой книге.

ГЛАВА 1. ВВЕДЕНИЕ

§ 1. Множества

Понятие *множества* будет для нас первичным, неопределяемым. Не определяя понятие множества, мы, тем не менее, обсудим его свойства и правила обращения с ним. Возможен и другой подход: в математической логике множество определяется аксиоматически; аксиомы описывают свойства множеств и правила построения одних множеств из других. Однако, в курсе нам будет удобнее оставаться в рамках "наивной" теории множеств.

Вместо слова "множество" часто употребляются его синонимы: класс, совокупность, набор. Теория множеств была основана Г. Кантором.

Множество состоит из объектов, вещей, называемых его *элементами.* Запись $x \in X$ означает, что объект x принадлежит множеству X, является элементом множества X, а запись $x \notin X$ или $x \overline{\in} X$ — что объект x не принадлежит множеству X, не является элементом множества X.

Множество может задаваться перечислением своих элементов: $\{1, 2, 3\}$, $\{A, B\}$, а также указанием свойства: множество всех элементов, обладающих свойством $\mathcal{P}$, обозначается $\{x : \mathcal{P}(x)\}$ или $\{x \mid \mathcal{P}(x)\}$.

Элементами множеств могут быть множества. Например, в множестве $\big\{1, \{1\}, \{1, 2\}\big\}$ три элемента: число 1, множество, единственным элементом которого является число 1, и множество из двух элементов — чисел 1 и 2. При образовании множеств из элементов следует соблюдать известную осторожность; так, понятие множества всех множеств противоречиво (то есть такого множества нет). Чтобы избежать противоречий, требуют, чтобы элементы были определены раньше множества. В частности, никакое множество не может содержать себя в качестве элемента. Обычно рассматриваются элементы, принадлежащие некоторому основному множеству U (объемлющему множеству, множеству допустимых элементов). Множество U фиксировано и либо ясно из контекста, либо явно указывается. Например, в планиметрии U — это плоскость, в стереометрии — трехмерное пространство и т.д.

Определение. Если каждый элемент множества X принадлежит множеству Y, то говорят, что X *содержится в* Y (а также — что X — *подмножество* Y, X — *часть* Y, Y *содержит* X), и пишут $X \subset Y$ или $Y \supset X$.

Если множества X и Y состоят из одних и тех же элементов, то их называют *равными* и пишут $X = Y$.

Другими словами, $X = Y$ тогда и только тогда, когда $X \subset Y$ и $Y \subset X$.

Удобно ввести в рассмотрение *пустое множество* — множество, в котором нет ни одного элемента. Пустое множество обозначается символом $\varnothing$. Пустое множество содержится в любом множестве Y, так как в нем нет ни одного элемента, не принадлежащего Y.

Несколько числовых множеств имеют общепринятые обозначения:

$\mathbb{N}$ — множество натуральных чисел (натуральный ряд);

$\mathbb{Z}_+$ — множество неотрицательных целых чисел;

$\mathbb{Z}$ — множество целых чисел;

$\mathbb{Q}$ — множество рациональных чисел, то есть дробей вида $\frac{p}{q}$, где $p \in \mathbb{Z}$, $q \in \mathbb{N}$;

$\mathbb{R}$ — множество вещественных (действительных) чисел;

$\mathbb{C}$ — множество комплексных чисел.

Если $\mathcal{P}$ и $\mathcal{Q}$ — два утверждения, то запись $\mathcal{P} \Rightarrow \mathcal{Q}$ называется *импликацией* и означает, что если верно $\mathcal{P}$, то верно и $\mathcal{Q}$. Также говорят, что из $\mathcal{P}$ следует $\mathcal{Q}$, $\mathcal{P}$ влечет $\mathcal{Q}$, $\mathcal{P}$ достаточно для $\mathcal{Q}$, $\mathcal{Q}$ необходимо для $\mathcal{P}$. Запись $\mathcal{Q} \Leftarrow \mathcal{P}$ означает то же самое. Если $\mathcal{P} \Rightarrow \mathcal{Q}$ и $\mathcal{Q} \Rightarrow \mathcal{P}$, то говорят, что утверждения $\mathcal{P}$ и $\mathcal{Q}$ *равносильны* (или *эквивалентны*) или $\mathcal{P}$ необходимо и достаточно для $\mathcal{Q}$, и пишут $\mathcal{P} \Leftrightarrow \mathcal{Q}$. Через $\overline{\mathcal{P}}$ или $\rceil\mathcal{P}$ обозначается отрицание утверждения $\mathcal{P}$.

Для сокращения записи бывает удобно использовать два специальных значка, называемых кванторами:

$\forall$ — *квантор общности* (или *квантор всеобщности*); заменяет слова “любой”, “всякий”, “каждый” или “для любого”, “для всякого”, “для каждого”;

$\exists$ — *квантор существования*; заменяет слова “существует”, “найдется”, “можно подобрать”.

Двоеточие в формулах с кванторами заменяет слова “такой, что”, но будет ставиться не всегда. Например, утверждения

$$\exists x : x > 0,\, x^2 = 2, \qquad \exists x > 0 : x^2 = 2, \qquad \exists x > 0 \quad x^2 = 2$$

означают одно и тоже и читаются: “существует x, такое что x больше нуля и x-квадрат равно двум” или “существует такое положительное x, квадрат которого равен двум”. Фразы

$$\forall x : x > 0,\, x < 2 \quad x^2 < 4 \qquad \text{и} \qquad \forall x > 0 : x < 2 \quad x^2 < 4$$

также означают одно и то же и читаются: “для любого x, такого что x больше нуля и меньше двух, x-квадрат меньше четырех” или “для любого положительного x, меньшего двух, x-квадрат меньше четырех” (а не “для любого x верно, что x больше нуля, x меньше двух и x-квадрат меньше четырех” и не “для любого положительного x верно, что x меньше двух и x-квадрат меньше четырех”!) Некоторые авторы предпочитают записывать последнее утверждение в виде

$$\forall x \ ((x > 0,\, x < 2) \Rightarrow x^2 < 4), \qquad \forall x > 0 \ (x < 2 \Rightarrow x^2 < 4),$$

но мы предпочитаем запись, по возможности близкую к устной речи и менее громоздкую.

Пусть $\mathcal{P}(x)$ — утверждение, зависящее от параметра $x \in X$. Отрицание утверждения “для любого x из X верно $\mathcal{P}(x)$” означает, что “существует x из X, для которого $\mathcal{P}(x)$ неверно”, а отрицание утверждения “для некоторого x из X верно $\mathcal{P}(x)$” означает, что “для любого x из X неверно $\mathcal{P}(x)$”:

$$\overline{\forall x \in X \quad \mathcal{P}(x)} \Leftrightarrow \exists x \in X \quad \overline{\mathcal{P}(x)}, \qquad \overline{\exists x \in X \quad \mathcal{P}(x)} \Leftrightarrow \forall x \in X \quad \overline{\mathcal{P}(x)}.$$

Итак, справедливо такое правило отрицания утверждений, содержащих кванторы: нужно заменить все кванторы на противоположные и самое последнее утверждение на его отрицание.

Мы будем пользоваться логической символикой как вспомогательной, не формализуя ее применение.

Множество определяется своими элементами, поэтому бессмысленно говорить, что какой-то элемент принадлежит множеству в нескольких экземплярах или что множество содержит несколько одинаковых элементов. Чтобы рассматривать несколько экземпляров одного и того же элемента, используют понятие *семейства.* Строго говоря, семейство — это отображение, так что определение семейства будет дано в § 3. Здесь мы опишем понятие семейства неформально.

Пусть X и A — множества, и некоторые элементы множества X снабжены (занумерованы) значками — элементами множества A (индексами). При этом каждый индекс использован ровно один раз, но элементы X могут быть снабжены более чем одним индексом. Последнее и означает, что элементы X могут присутствовать в семействе во многих экземплярах. Для семейства используются обозначения $\{x_\alpha\}_{\alpha\in A}$ или $(x_\alpha)_{\alpha\in A}$. Индекс α в этом обозначении, как говорят, "немой" и может быть заменен другой буквой.

Определение. Пусть $\{X_\alpha\}_{\alpha\in A}$ — семейство множеств. *Объединением* семейства $\{X_\alpha\}_{\alpha\in A}$ называется множество всех элементов, которые принадлежат хотя бы одному из множеств X_α:

$$\bigcup_{\alpha\in A} X_\alpha = \{x : \exists\alpha\in A \ \ x\in X_\alpha\}.$$

Если семейство состоит из двух множеств X и Y, пишут $X\cup Y$ (неважно, как нумеровать множества). Ясно, что

$$(X\cup Y)\cup Z = X\cup(Y\cup Z), \quad X\cup Y = Y\cup X, \quad X\cup X = X\cup\varnothing = X.$$

Определение. Пусть $\{X_\alpha\}_{\alpha\in A}$ — семейство множеств. *Пересечением* семейства $\{X_\alpha\}_{\alpha\in A}$ называется множество всех элементов, которые принадлежат каждому из множеств X_α:

$$\bigcap_{\alpha\in A} X_\alpha = \{x : \forall\alpha\in A \ \ x\in X_\alpha\}.$$

Если семейство состоит из двух множеств X и Y, пишут $X\cap Y$. Ясно, что

$$(X\cap Y)\cap Z = X\cap(Y\cap Z), \quad X\cap Y = Y\cap X, \quad X\cap X = X, \quad X\cap\varnothing = \varnothing.$$

Когда множество индексов A есть $\{1,\ldots n\}$, то пишут $\bigcup\limits_{k=1}^{n} X_k$, $\bigcap\limits_{k=1}^{n} X_k$, а когда $A=\mathbb{N}$, то $\bigcup\limits_{k=1}^{\infty} X_k$, $\bigcap\limits_{k=1}^{\infty} X_k$. Рассматривают также объединение и пересечение не семейства, а множества множеств в том же смысле (впрочем, можно занумеровать каждое множество, используя в качестве индекса само множество, и свести дело к случаю семейства).

Определение. *Разностью* множеств X и Y называется множество всех элементов, которые принадлежат X, но не принадлежат Y:

$$X \setminus Y = \{x : x \in X, x \notin Y\}.$$

В этом определении не предполагается, что $Y \subset X$. Если $Y \subset X$, то разность $X \setminus Y$ называется еще *дополнением* множества Y до множества X. Дополнение X до основного множества U называется короче — дополнением X — и обозначается X^c (иногда также используются обозначения CX, $\overline{X}$, X'). Таким образом, дополнение X есть множество всех элементов основного множества, не принадлежащих X.

Ясно, что равенство множеств равносильно равенству их дополнений. Справедливы также равенства $(X^c)^c = X$, $X \cup X^c = U$, $X \cap X^c = \varnothing$. Соотношения $X \subset Y$, $Y^c \subset X^c$, $X \cap Y^c = \varnothing$ и $Y \cup X^c = U$ равносильны.

Теорема 1. Законы де Моргана. *Пусть* $\{X_\alpha\}_{\alpha\in A}$ *— семейство множеств,* Y *— множество. Тогда*

$$Y \setminus \bigcup_{\alpha\in A} X_\alpha = \bigcap_{\alpha\in A} (Y \setminus X_\alpha), \tag{1}$$

$$Y \setminus \bigcap_{\alpha\in A} X_\alpha = \bigcup_{\alpha\in A} (Y \setminus X_\alpha). \tag{2}$$

Доказательство. Обозначим через Λ и Π соответственно левую и правую части равенства (1). По определению разности соотношение $x \in \Lambda$ означает, что $x \in Y$ и x не принадлежит объединению множеств X_α. По определению объединения это значит, что $x \in Y$ и x не принадлежит ни одному из множеств X_α, то есть

$x \in Y \setminus X_\alpha$ при всех $\alpha \in A$. По определению пересечения последнее значит, что $x \in \Pi$. Равенство $\Lambda = \Pi$ доказано. Соотношение (2) доказывается аналогично. $\square$

Теорема 2. *Пусть* $\{X_\alpha\}_{\alpha \in A}$ *— семейство множеств,* Y *— множество. Тогда*

$$Y \cap \bigcup_{\alpha \in A} X_\alpha = \bigcup_{\alpha \in A} (Y \cap X_\alpha), \tag{3}$$

$$Y \cup \bigcap_{\alpha \in A} X_\alpha = \bigcap_{\alpha \in A} (Y \cup X_\alpha). \tag{4}$$

Доказательство. Обозначим через Λ и Π левую и правую часть равенства (3). По определению пересечения соотношение $x \in \Lambda$ означает, что $x \in Y$ и $x \in \bigcup\limits_{\alpha \in A} X_\alpha$. По определению объединения это значит, что $x \in Y$ и существует такое $\alpha_0 \in A$, что $x \in X_{\alpha_0}$. Другими словами, существует такое $\alpha_0 \in A$, что $x \in Y \cap X_{\alpha_0}$. Последнее означает, что $x \in \Pi$. Равенство $\Lambda = \Pi$ доказано. Соотношение (4) доказывается аналогично. $\square$

Упорядоченная пара — это двухэлементное семейство, где множеством индексов является $\{1, 2\}$. При этом в обозначении упорядоченной пары (a, b) считается, что на первом месте написан элемент, занумерованный индексом 1, а на втором — индексом 2. Порядок элементов удобно указывать с помощью индексации: (x_1, x_2). Подчеркнем, что элементы x_1 и x_2 могут совпадать (в отличие от случая неупорядоченной пары — двухэлементного множества $\{x_1, x_2\}$: если $x_1 = x_2$, то это одноэлементное множество), и что порядок элементов существенен. Равенство пар (a, b) и (c, d) означает, что $a = c$ и $b = d$. Аналогично определяется упорядоченный набор из m элементов $(x_1, \ldots, x_m)$. Элементы упорядоченного набора называются его координатами или компонентами.

Определение. *Декартовым* или *прямым произведением* множеств X и Y называется множество всех упорядоченных пар, таких что первый элемент пары принадлежит X, а второй — Y:

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

Это определение обобщается на несколько сомножителей:

$$X_1 \times \cdots \times X_m = \{(x_1, \ldots, x_m) : x_i \in X_i \text{ при всех } i = 1, \ldots, m\}.$$

Порядок сомножителей существенен; можно сказать, что прямое произведение определено для упорядоченного набора множеств. Если $X_1 = \ldots = X_m = X$, то произведение обозначают также через X^m; в частности, $X^1 = X$. Таким образом, X^m — это множество упорядоченных наборов из m элементов, принадлежащих множеству X. В частности, $\mathbb{R}^m$ (читается: "эр-эм", а не "эр в степени эм") — это множество упорядоченных наборов из m вещественных чисел, а $\mathbb{C}^m$ ("це-эм") — из m комплексных чисел. Точку прямой можно отождествить с вещественным числом, точку плоскости — с парой, а точку пространства — с тройкой чисел, служащих ее координатами. Поэтому множество $\mathbb{R}^1 = \mathbb{R}$ называют прямой, $\mathbb{R}^2$ — плоскостью, а $\mathbb{R}^3$ — трехмерным пространством. По аналогии множество $\mathbb{R}^m$ называют m-мерным (вещественным) пространством, а его элементы — m-мерными точками или векторами. Последний термин объясняется тем, что и вектор можно отождествить с набором чисел — его координат.

§ 2. Вещественные числа

Читателю, конечно, знакомы из школы свойства вещественных (действительных) чисел, хотя четкое определение вещественного числа, скорее всего, неизвестно. В этом параграфе мы определим множество вещественных чисел аксиоматически: множество $\mathbb{R}$ называется множеством вещественных чисел, если выполнен некоторый набор условий (аксиом).

Множество вещественных чисел может быть определено и конструктивно, то есть построено, различными способами: например, с помощью бесконечных десятичных дробей, дедекиндовых сечений или фундаментальных последовательностей. Построение множества $\mathbb{R}$ в этом курсе проводиться не будет.

Для удобства разобьем аксиомы, которых всего шестнадцать, на группы.

I. Аксиомы поля. В множестве $\mathbb{R}$ определены две операции, называемые сложением и умножением, действующие из $\mathbb{R}\times\mathbb{R}$ в $\mathbb{R}$ и удовлетворяющие следующим свойствам.

I.1. Сочетательный закон (ассоциативность) сложения:

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

I.2. Переместительный закон (коммутативность) сложения:

$$x+y = y+x.$$

I.3. Существует вещественное число нуль (0, нейтральный элемент по сложению), такое что $x+0=x$ для всех x.

I.4. Для любого числа x существует такое число $\tilde{x}$, что $x+\tilde{x}=0$ ($\tilde{x}$ называется числом, противоположным x, и обозначается $-x$).

I.5. Сочетательный закон (ассоциативность) умножения:

$$(xy)z = x(yz).$$

I.6. Переместительный закон (коммутативность) умножения:

$$xy = yx.$$

I.7. Существует вещественное число единица (1, нейтральный элемент по умножению), отличное от нуля, такое что $x\cdot 1 = x$ для всех x.

I.8. Для любого числа x, отличного от нуля, существует такое число x', что $xx'=1$ (x' называется числом, обратным к x, и обозначается x^{-1} или $\frac{1}{x}$).

I.9. Распределительный закон (дистрибутивность):

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

Множество, в котором определены две операции, удовлетворяющие свойствам I.1–I.9, называется *полем,* а сами свойства I.1–I.9 — аксиомами поля.

Аксиома I.4 позволяет определить вычитание: $x-y=x+(-y)$, а аксиома I.8 — деление на любое число, отличное от нуля: $\frac{x}{y}=x\cdot\frac{1}{y}$.

II. Аксиомы порядка. Между элементами $\mathbb{R}$ определено отношение $\leqslant$ со следующими свойствами.

II.1. Для любых x, y верно $x \leqslant y$ или $y \leqslant x$.

II.2. Транзитивность: если $x \leqslant y$ и $y \leqslant z$, то $x \leqslant z$.

II.3. Если $x \leqslant y$ и $y \leqslant x$, то $x = y$.

II.4. Если $x \leqslant y$, то $x + z \leqslant y + z$ для любого z.

II.5. Если $0 \leqslant x$ и $0 \leqslant y$, то $0 \leqslant xy$.

Поле, в котором введено отношение порядка, удовлетворяющее свойствам II.1–II.5, называется *упорядоченным.*

Другие знаки неравенств определяются так:

$$a < b \text{ означает, что } a \leqslant b \text{ и } a \neq b,$$
$$a \geqslant b \text{ означает, что } b \leqslant a,$$
$$a > b \text{ означает, что } b \leqslant a \text{ и } a \neq b.$$

Аксиомы поля и порядка — это привычные свойства арифметических действий и неравенств с вещественными числами. Другие знакомые свойства (например, $0 < 1$ или $(-x)(-y) = xy$), при аксиоматическом определении вещественных чисел с помощью этой системы аксиом являются теоремами и могут быть выведены из аксиом. Приведем еще три таких утверждения.

1. Если $x \leqslant y$, то $-x \geqslant -y$.
2. Если $x \leqslant y$ и $a \leqslant b$, то $x + a \leqslant y + b$.
3. Если $0 \leqslant x \leqslant y$ и $0 \leqslant a \leqslant b$, то $ax \leqslant by$.

Мы не будем доказывать все подобные теоремы (это совсем просто, в отличие от осознания, что тот или иной факт требует доказательства).

Множество вещественных чисел удобно изображать в виде числовой прямой, а сами числа — точками на этой прямой. Поэтому числа называют еще и точками.

Наличие порядка позволяет определить промежутки в множестве $\mathbb{R}$. Пусть $a, b \in \mathbb{R}$. Если $a \leqslant b$, то множество

$$[a, b] = \{x \in \mathbb{R} : a \leqslant x \leqslant b\}$$

называется *отрезком* или *сегментом.* Если $a < b$, то множество

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}$$

называется *интервалом,* а множества

$$[a, b) = \{x \in \mathbb{R} : a \leqslant x < b\}, \quad (a, b] = \{x \in \mathbb{R} : a < x \leqslant b\}$$

— *полуинтервалами.* При $a = b$ отрезок состоит из одной этой точки и называется вырожденным. Множества

$$[a, +\infty) = \{x \in \mathbb{R} : a \leqslant x\}, \quad (a, +\infty) = \{x \in \mathbb{R} : a < x\},$$
$$(-\infty, b] = \{x \in \mathbb{R} : x \leqslant b\}, \quad (-\infty, b) = \{x \in \mathbb{R} : x < b\}$$

называются *лучами* (первое и третье — замкнутым лучом, второе и четвертое — открытым). Все множество $\mathbb{R}$ обозначается еще $(-\infty, +\infty)$. Символам $-\infty$ и $+\infty$ в этих обозначениях не приписывается самостоятельного значения. Через $\langle a, b\rangle$ обозначается промежуток любого из четырех типов с концами a и b; через $\langle a, b)$ — любой из двух промежутков (a, b) и $[a, b)$, и т.д. Положим также $[a : b] = [a, b] \cap \mathbb{Z}$. Если $a, b \in \mathbb{R}$, а точка c принадлежит интервалу с концами a и b, то говорят, что *c лежит между a и b.*

Обозначим еще через $\mathbb{R}_+$ множество неотрицательных вещественных чисел, то есть $\mathbb{R}_+ = [\,0, +\infty)$.

Множество $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ называется *расширенной числовой прямой.* Таким образом, в $\overline{\mathbb{R}}$ к вещественным числам добавляются два новых символа (несобственных элемента): $-\infty$ и $+\infty$. Считают, что $-\infty < x < +\infty$ для любого $x \in \mathbb{R}$ и $-\infty < +\infty$. Тогда можно рассматривать промежутки в $\overline{\mathbb{R}}$ вида $\langle a, +\infty]$ или $[-\infty, b\rangle$.

С несобственными элементами можно совершать некоторые операции. Полагают

$$x + (+\infty) = (+\infty) + x = +\infty, \quad x \in \mathbb{R},$$
$$x + (-\infty) = (-\infty) + x = -\infty, \quad x \in \mathbb{R},$$
$$x \cdot (+\infty) = (+\infty) \cdot x = \begin{cases} +\infty, & x > 0, \\ -\infty, & x < 0, \end{cases}$$
$$x \cdot (-\infty) = (-\infty) \cdot x = \begin{cases} -\infty, & x > 0, \\ +\infty, & x < 0, \end{cases}$$
$$(+\infty) + (+\infty) = +\infty, \qquad (-\infty) + (-\infty) = -\infty,$$
$$(+\infty) \cdot (+\infty) = (-\infty) \cdot (-\infty) = +\infty,$$
$$(+\infty) \cdot (-\infty) = (-\infty) \cdot (+\infty) = -\infty.$$

Символам $(+\infty)+(-\infty)$, $(+\infty)-(+\infty)$, $(-\infty)-(-\infty)$, $0\cdot(\pm\infty)$ и $(\pm\infty)\cdot 0$ не приписывается никакого значения.

III. Аксиома Архимеда. *Каковы бы ни были положительные числа $x, y \in \mathbb{R}$, существует такое натуральное число n, что $nx > y$.*

Упорядоченное поле, в котором выполняется аксиома Архимеда, называется *архимедовым.*

Из аксиомы Архимеда следует, что существуют сколь угодно большие натуральные числа.

Сформулированные пятнадцать аксиом еще не определяют множество вещественных чисел полностью: этим аксиомам удовлетворяет, например, и множество рациональных чисел. Поэтому необходимо ввести еще какие-то аксиомы, которые позволили бы различить множества $\mathbb{R}$ и $\mathbb{Q}$. Для этой цели хватает одной аксиомы; ее называют аксиомой полноты или непрерывности. Сформулировать аксиому полноты можно разными способами; мы сделаем это в виде аксиомы о вложенных отрезках. Участвующий в ее формулировке термин "последовательность" строго определяется в § 3.

IV. Аксиома Кантора о вложенных отрезках. *Пусть $\{[a_n, b_n]\}_{n=1}^{\infty}$ — последовательность вложенных отрезков, то есть*

$$a_n \leqslant a_{n+1} \leqslant b_{n+1} \leqslant b_n \text{ для всех } n \in \mathbb{N}.$$

Тогда существует точка, принадлежащая одновременно всем отрезкам $[a_n, b_n]$, то есть

$$\bigcap_{n=1}^{\infty} [a_n, b_n] \neq \varnothing.$$

В множестве $\mathbb{Q}$ это утверждение не выполняется. У читателя появятся все необходимые для доказательства сведения после прочтения главы 2.

Аксиома о вложенных отрезках — способ выразить, что вещественная прямая "сплошная", на ней нет "дырок", в отличие от рациональной прямой, имеющей "дырку" на каждом месте, на котором должно находиться иррациональное число. Известные способы построения множества вещественных чисел как раз и состоят в формализации процедуры "заполнения дырок".

В аксиоме существенно, что речь идет именно об отрезках; пересечение вложенных промежутков другого типа может оказаться пустым. Например,

$$\bigcap_{n=1}^{\infty}\left(0,\frac{1}{n}\right] = \varnothing, \qquad \bigcap_{n=1}^{\infty}[n,+\infty) = \varnothing$$

(это легко следует из аксиомы Архимеда).

При аксиоматическом определении объекта возникают три вопроса: является ли система аксиом непротиворечивой (то есть не следуют ли из нее одновременно некоторое утверждение и его отрицание), независимой (то есть не является ли одна из аксиом следствием остальных — и тогда ее можно удалить) и полной (то есть единственный ли объект описывается системой аксиом). Мы не будем обсуждать эти вопросы и примем на веру, что для приведенной аксиоматики вещественных чисел ответ на них, после некоторых уточнений, положителен.

Определение. Пусть $x \in \mathbb{R}$. Число

$$|x| = \begin{cases} x, & x \geqslant 0, \\ -x, & x < 0 \end{cases}$$

называется *модулем* или *абсолютной величиной* числа x.

Геометрически $|x|$ есть расстояние от точки x на числовой оси до точки 0. График функции модуль изображен на рисунке 1.

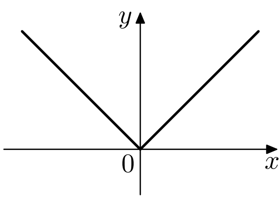

Рис. 1

Напомним известные свойства модуля.

1. $|x| \geqslant 0$.

2. $|-x| = |x|$.
3. $\pm x \leqslant |x|$.
4. $|xy| = |x||y|$.
5. $|\dfrac{x}{y}| = \dfrac{|x|}{|y|}$, $\quad y \neq 0$.
6. $\big||x| - |y|\big| \leqslant |x \pm y| \leqslant |x| + |y|$.

7. *Если $a > 0$, то неравенство $|x| < a$ равносильно двойному неравенству $-a < x < a$.*

Докажем свойство 6; все остальные свойства сразу следуют из определения. Складывая неравенства $\pm x \leqslant |x|$ и $\pm y \leqslant |y|$, находим, что $\pm(x + y) \leqslant |x| + |y|$, откуда вытекает правое неравенство. По доказанному правому неравенству

$$|x| = |x - y + y| \leqslant |x - y| + |y|,$$

то есть $|x| - |y| \leqslant |x - y|$. Меняя числа x и y местами, получаем $|y| - |x| \leqslant |x - y|$. Два последних неравенства и означают, что $\big||x| - |y|\big| \leqslant |x \pm y|$.

Кратко сформулируем определение и простейшие свойства комплексных чисел. Подробнее комплексные числа изучаются в курсе алгебры. *Комплексное число* z — это упорядоченная пара вещественных чисел (x, y), так что как множество $\mathbb{C} = \mathbb{R}^2$. Чтобы получить право называть элементы $\mathbb{C}$ числами, следует определить арифметические действия между ними. Складываются комплексные числа, как и вектора, покоординатно (см. рисунок 2).

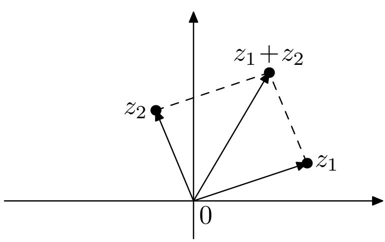

Рис. 2

Если $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$, то

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2).$$

Нулем является точка $(0,0)$, а числом, противоположным z, — число $-z = (-x, -y)$. Умножение комплексных чисел задается формулой

$$z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1). \tag{5}$$

Единицей является точка $(1,0)$. Вещественные числа x отождествляются с комплексными числами вида $(x,0)$; по этому соглашению $\mathbb{R} \subset \mathbb{C}$. Точка $i = (0,1)$ называется *мнимой единицей,* а числа вида $(0,y)$ — *мнимыми* (чисто мнимыми). Всякое комплексное число может быть записано в виде

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

(здесь под iy понимается произведение вектора i на вещественное число y). Равенство $z = x + iy$ называется *алгебраической формой записи* комплексного числа. Легко видеть, что $i^2 = -1$. Умножение комплексных чисел по формуле (5) может быть описано так: нужно раскрыть скобки, положить $i^2 = -1$ и привести подобные члены. Если $z \neq 0$, то обратным к z является число

$$\frac{1}{z} = \frac{x}{x^2 + y^2} + i\frac{-y}{x^2 + y^2}.$$

Легко проверяется, что комплексные числа образуют поле. Однако, это поле неупорядоченное.

Если $z = x + iy$, $x, y \in \mathbb{R}$, то x называют *вещественной частью,* а y — *мнимой частью* z, и пишут $x = \operatorname{Re} z$, $y = \operatorname{Im} z$. Число $\overline{z} = x - iy$ называется *сопряженным* к z. Видно, что

$$\operatorname{Re} z = \frac{z + \overline{z}}{2}, \qquad \operatorname{Im} z = \frac{z - \overline{z}}{2i},$$
$$\overline{\overline{z}} = z, \qquad \overline{z_1 * z_2} = \overline{z_1} * \overline{z_2},$$

где $*$ обозначает одно из четырех арифметических действий. *Модулем* числа z называется длина отвечающего ему вектора:

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

Модуль комплексных чисел обладает следующими свойствами:

$$|z_1 z_2| = |z_1||z_2|, \qquad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|},\ z_2 \neq 0,$$
$$\big||z_1| - |z_2|\big| \leqslant |z_1 \pm z_2| \leqslant |z_1| + |z_2|, \qquad z\overline{z} = |z|^2.$$

Если $z \neq 0$, то угол φ, отсчитанный от вектора 1 до вектора z против часовой стрелки, называется *аргументом* числа z и обозначается $\arg z$. Аргумент определяется не однозначно, а с точностью до слагаемого, кратного 2π. Множество всех значений аргумента z обозначают $\operatorname{Arg} z$, а через $\arg z$ обозначают любой элемент этого множества. Можно дополнительно потребовать, чтобы $\arg z$ принадлежал фиксированному полуинтервалу длины 2π; обычно для этой цели выбирают полуинтервал $(-\pi, \pi]$ или $[0, 2\pi)$, а соответствующее значение аргумента называют *главным*. Числа $r = |z|$ и $\varphi = \arg z$ называются *полярными координатами* точки $z = (x, y)$ на плоскости (см. рисунок 3).

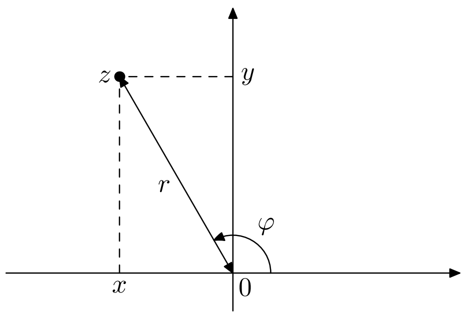

Рис. 3

По известным формулам для прямоугольного треугольника

$$x = r\cos\varphi, \qquad y = r\sin\varphi.$$

Поэтому

$$z = r(\cos\varphi + i\sin\varphi), \qquad r = |z|,\ \varphi = \arg z.$$

Это равенство называется *тригонометрической формой записи* комплексного числа.

При перемножении комплексных чисел их модули перемножаются, а аргументы складываются. Точнее, если $\varphi_1 \in \operatorname{Arg} z_1$, $\varphi_2 \in \operatorname{Arg} z_2$, то $\varphi_1 + \varphi_2 \in \operatorname{Arg}(z_1 z_2)$. Геометрически умножение комплексного числа z_1 на число z_2, равное по модулю 1, означает поворот вектора z_1 на угол φ_2 против часовой стрелки. Умножение комплексных чисел изображено на рисунке 4.

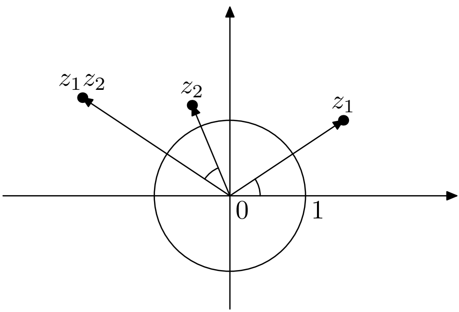

Рис. 4

В частности, если $n \in \mathbb{N}$, $\varphi \in \operatorname{Arg} z$, то $n\varphi \in \operatorname{Arg} z^n$. Отсюда вытекает формула Муавра:

$$z^n = r^n(\cos n\varphi + i \sin n\varphi),$$

верная для всех целых n.

Перейдем к обсуждению некоторых свойств натуральных чисел. Принципом математической индукции называют следующее утверждение.

Принцип математической индукции. *Пусть* $\{\mathcal{P}_n\}_{n=1}^{\infty}$ — *последовательность утверждений. Если*

1) $\mathcal{P}_1$ *верно,*

2) *для любого* $n \in \mathbb{N}$ *из* $\mathcal{P}_n$ *следует* $\mathcal{P}_{n+1}$*,*

то $\mathcal{P}_n$ *верно для всех* $n \in \mathbb{N}$.

Утверждение 1) называется *базой индукции*, а утверждение 2) — *индукционным переходом*; при этом $\mathcal{P}_n$ называют *индукционным предположением*.

Подчеркнем, что при проверке утверждения 2) нужно доказывать не истинность $\mathcal{P}_n$, а тот факт, что из $\mathcal{P}_n$ следует $\mathcal{P}_{n+1}$.

Иногда используется следующая модификация принципа математической индукции. Если последовательность утверждений $\{\mathcal{P}_n\}_{n=m}^{\infty}$ задана для всех целых n, не меньших некоторого целого числа m, то базой индукции служит утверждение $\mathcal{P}_m$, а индукционный переход делается для всех $n \geqslant m$. Для доказательства достаточно применить принцип математической индукции к утверждениям $\mathcal{Q}_n = \mathcal{P}_{n+m-1}$, $n \in \mathbb{N}$.

При аксиоматическом определении множества натуральных чисел принцип математической индукции (или какое-нибудь близкое утверждение) принимается за аксиому. Мы не будем приводить аксиоматику $\mathbb{N}$, а определим $\mathbb{N}$ как подмножество $\mathbb{R}$. Разумеется, чтобы не попасть в порочный круг, множество $\mathbb{N}$ должно быть определено до формулировок аксиом Архимеда и Кантора.

Множество $M \subset \mathbb{R}$ называется *индуктивным,* если $1 \in M$ и вместе с каждым своим элементом x множество M содержит и элемент $x+1$. Индуктивные множества существуют: $\mathbb{R}$ — индуктивное множество.

Определение. Множеством *натуральных чисел* называется минимальное по включению индуктивное подмножество $\mathbb{R}$.

Другими словами,

$$\mathbb{N} = \bigcap_{\substack{M \subset \mathbb{R} \\ M \text{индуктивно}}} M. \tag{6}$$

Действительно, правая часть равенства (6) является индуктивным множеством и содержится в любом индуктивном множестве.

Принцип математической индукции используется в математике гораздо чаще, чем явно упоминается. Если в каком-то определении или рассуждении встретились слова “и так далее”, “продолжим этот процесс неограниченно” или заменяющее их многоточие — это верный признак того, что при формальном изложении должен использоваться принцип математической индукции. Не являются исключением записанные ниже “определения” суммы и произведения нескольких чисел, на самом деле, опирающиеся на индукцию.

Тем не менее, предупредив читателя о принципиальной возможности, а иногда и необходимости расшифровки слов "и так далее", мы не будем заниматься этой расшифровкой.

В качестве одного из применений метода математической индукции докажем формулу бинома Ньютона. Предварительно введем ряд обозначений.

Если $m, n \in \mathbb{Z}$, $m \leqslant n$, то полагаем

$$\sum_{k=m}^{n} a_k = a_m + \ldots + a_n,$$

$$\prod_{k=m}^{n} a_k = a_m \cdot \ldots \cdot a_n.$$

Если $m > n$, то сумма $\sum\limits_{k=m}^{n} a_k$ считается равной 0, а произведение $\prod\limits_{k=m}^{n} a_k$ — равным 1. В более общей ситуации символами $\sum\limits_{\alpha \in A} a_\alpha$ и $\prod\limits_{\alpha \in A} a_\alpha$ обозначаются сумма и произведение конечного числового семейства (то есть такого, что множество индексов конечно). Индексы k и α в этих обозначениях "немые" и могут быть заменены другими буквами, например, $\sum\limits_{k=m}^{n} a_k = \sum\limits_{j=m}^{n} a_j$.

Следующее преобразование называют *сдвигом индекса суммирования:* если $p \in \mathbb{Z}$, то

$$\sum_{k=m}^{n} a_k = \sum_{j=m+p}^{n+p} a_{j-p}.$$

В самом деле, в обеих частях равенства записана сумма чисел $a_m, \ldots, a_n$.

Если $n \in \mathbb{Z}_+$, то произведение $n! = \prod\limits_{k=1}^{n} k$ называется *факториалом* числа n; отметим, что по общему соглашению $0! = 1$. Числа

$$C_n^k = \frac{n!}{k!\,(n-k)!}$$

называются *биномиальными коэффициентами,* в англоязычной литературе они обозначаются $\binom{n}{k}$. Отметим, что

$$C_0^0 = C_n^n = 1, \qquad C_n^1 = C_n^{n-1} = n, \qquad C_{n+1}^k = C_n^{k-1} + C_n^k.$$

Договоримся считать $x^0 = 1$ при всех x, в том числе при $x = 0$.

Теорема 1. Бином Ньютона. *Если* $n \in \mathbb{Z}_+$, $x, y \in \mathbb{R}$ *или* $\mathbb{C}$, *то*

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k}.$$

Доказательство. При $n = 0$ и $n = 1$ равенство очевидно; при $n = 1$ оно служит базой индукции. Сделаем индукционный переход. Пусть формула верна для номера n; докажем, что она верна и для $n+1$. По индукционному предположению

$$(x+y)^{n+1} = (x+y)(x+y)^n = (x+y)\sum_{k=0}^{n} C_n^k x^k y^{n-k} =$$

$$= \sum_{k=0}^{n} C_n^k x^{k+1} y^{n-k} + \sum_{k=0}^{n} C_n^k x^k y^{n+1-k}.$$

В первой сумме сдвинем индекс: $j = k+1$, а затем переименуем j в k и приведем подобные члены:

$$\begin{aligned}
(x+y)^{n+1} &= \sum_{j=1}^{n+1} C_n^{j-1} x^j y^{n+1-j} + \sum_{k=0}^{n} C_n^k x^k y^{n+1-k} = \\
&= \sum_{k=1}^{n+1} C_n^{k-1} x^k y^{n+1-k} + \sum_{k=0}^{n} C_n^k x^k y^{n+1-k} = \\
&= C_n^n x^{n+1} y^0 + \sum_{k=1}^{n} (C_n^{k-1} + C_n^k) x^k y^{n+1-k} + C_n^0 x^0 y^{n+1} = \\
&= C_{n+1}^{n+1} x^{n+1} y^0 + \sum_{k=1}^{n} C_{n+1}^k x^k y^{n+1-k} + C_{n+1}^0 x^0 y^{n+1} = \\
&= \sum_{k=0}^{n+1} C_{n+1}^k x^k y^{n+1-k},
\end{aligned}$$

что и завершает индукционный переход. $\square$

Определение. Множество $E \subset \mathbb{R}$ называется *ограниченным сверху,* если существует такое число $M \in \mathbb{R}$, что $x \leqslant M$ для всех $x \in E$. Число M при этом называется *верхней границей* множества E.

Множество $E \subset \mathbb{R}$ называется *ограниченным снизу,* если существует такое число $m \in \mathbb{R}$, что $x \geqslant m$ для всех $x \in E$. Число m при этом называется *нижней границей* множества E.

Множество $E \subset \mathbb{R}$ называется *ограниченным,* если оно ограничено и сверху, и снизу.

Замечание 1. Если M — верхняя граница множества E, то всякое число, большее M, — тоже верхняя граница множества E. Если m — нижняя граница множества E, то всякое число, меньшее m, — тоже нижняя граница множества E.

Замечание 2. Ограниченность множества E равносильна "ограниченности по модулю", то есть существованию такого числа K, что $|x| \leqslant K$ для всех $x \in E$.

Доказательство. Если $|x| \leqslant K$ для всех $x \in E$, то можно записать, что $-K \leqslant x \leqslant K$ для всех $x \in E$, и положить $m = -K$, $M = K$. Обратно, если $-m \leqslant x \leqslant M$ для всех $x \in E$, то в качестве K можно взять наибольшее из чисел $|m|$ и $|M|$. □

Определение. Число M называется *максимумом* или *наибольшим элементом* множества $E \subset \mathbb{R}$, если $M \in E$ и $x \leqslant M$ для всех $x \in E$.

Число m называется *минимумом* или *наименьшим элементом* множества $E \subset \mathbb{R}$, если $m \in E$ и $x \geqslant m$ для всех $x \in E$.

Максимум и минимум множества E обозначаются соответственно $\max E$ и $\min E$.

Ясно, что максимум множества является его верхней границей, а минимум — нижней границей. В то же время, не всякое, даже ограниченное сверху (снизу), множество имеет максимум (минимум). Например, в интервале $(0, 1)$ нет ни наибольшего, ни наименьшего элемента.

Теорема 2. Существование максимума и минимума конечного множества. *Во всяком конечном подмножестве $\mathbb{R}$ есть наибольший и наименьший элемент.*

Доказательство проведем индукцией по числу n элементов множества. База индукции — случай $n = 1$: если в множестве всего один элемент, то он и наибольший, и наименьший. Для определенности индукционный переход проведем в случае максимума. Пусть всякое n-элементное подмножество $\mathbb{R}$ имеет максимум, E — $(n+1)$-элементное подмножество $\mathbb{R}$:

$$E = \{x_1, \ldots, x_n, x_{n+1}\}.$$

Обозначим

$$c = \max\{x_1, \ldots, x_n\}.$$

Если $c \leqslant x_{n+1}$, то, очевидно, $x_{n+1} = \max E$, а если $c > x_n$, то $c = \max E$. $\square$

Следствие 1. *Во всяком непустом ограниченном сверху (снизу) подмножестве $\mathbb{Z}$ есть наибольший (наименьший) элемент.*

Доказательство. Пусть $E \subset \mathbb{Z}$, $E \neq \varnothing$, E ограничено сверху. Выберем какой-нибудь элемент $n_0 \in E$ и положим

$$E_1 = \{n \in E : n \geqslant n_0\}.$$

Поскольку E ограничено сверху, множество E_1 конечно: если M — натуральное число, служащее верхней границей E, то в множестве E_1 не более $M - n_0 + 1$ элементов. По теореме 2 в множестве E_1 есть наибольший элемент; ясно, что он и будет наибольшим элементом E.

Случай ограниченного снизу множества разбирается аналогично. $\square$

Следствие 2. *Во всяком непустом подмножестве $\mathbb{N}$ есть наименьший элемент.*

Это свойство называют *полной упорядоченностью* множества $\mathbb{N}$.

Определение. Пусть $x \in \mathbb{R}$. Наибольшее целое число, не превосходящее x, называется *целой частью* x и обозначается $[x]$.

Существование целой части обеспечивается следствием 1, поэтому определение корректно.

Замечание 1. Из определения следует, что

$$[x] \leqslant x < [x] + 1, \qquad x - 1 < [x] \leqslant x.$$

Обратно, если $y \in \mathbb{Z}$ и $x - 1 < y \leqslant x$, то $y = [x]$.

Теорема 3. Плотность множества рациональных чисел. *Во всяком интервале есть рациональное число.*

Доказательство. Пусть $a, b \in \mathbb{R}$, $a < b$. Тогда $\frac{1}{b-a} > 0$, и по аксиоме Архимеда найдется такое $n \in \mathbb{N}$, что $n > \frac{1}{b-a}$, то есть $\frac{1}{n} < b - a$. Положим $c = \frac{[na]+1}{n}$. Тогда $c \in \mathbb{Q}$ и

$$c \leqslant \frac{na+1}{n} = a + \frac{1}{n} < a + b - a = b,$$
$$c > \frac{na-1+1}{n} = a,$$

то есть $c \in (a, b)$. $\square$

Свойство, выраженное в теореме 3, называют *плотностью* множества $\mathbb{Q}$ в множестве $\mathbb{R}$.

Следствие 3. *Во всяком интервале бесконечно много рациональных чисел.*

Доказательство. Пусть в некотором интервале (a, b) количество рациональных чисел конечно. Обозначим через x_1 наименьшее из них. Тогда в интервале (a, x_1) нет ни одного рационального числа, что противоречит теореме 3. $\square$

§ 3. Отображения

Пусть X и Y — множества. Если каждому элементу x множества X сопоставлен по определенному правилу f один элемент y множества Y, то говорят, что задано *отображение* множества X в множество Y, и пишут

$$f\colon X \to Y \quad \text{или} \quad X \xrightarrow{f} Y.$$

Это описание нельзя считать строгим определением понятия "отображение", так как оно содержит не получившие до сих пор определения понятия "правило" и "сопоставлен". И хотя можно дать строгое формальное определение отображения на основе понятия множества, нам будет удобнее не проводить эту формализацию и считать понятие отображения неопределяемым.

Отображение — это тройка объектов (X, Y, f). Если из контекста ясно, о каких множествах X и Y идет речь, то упоминание о них опускают, указывают только последний элемент тройки — правило f — и говорят "отображение f".

Множество X называют *областью определения* или *областью задания,* а множество Y — *областью значений* или *областью изменения* отображения.

Тот элемент $y \in Y$, который сопоставляется элементу $x \in X$ по правилу f, обозначается через $f(x)$ и называется *значением* отображения f на элементе x или *образом* элемента x. При этом x называется *аргументом* отображения или *независимой переменной,* а y — *значением* отображения или *зависимой переменной.* Пишут также $x \mapsto f(x)$, $x \in X$.

Разумеется, вместо x, y и f могут использоваться и другие буквы.

Если X и Y — числовые множества, то отображение называют *функцией.* Это соглашение не является общепринятым: иногда термин "функция" употребляют в том же смысле, что и "отображение", а иногда функцией называют отображение с числовой областью значений. В последней ситуации используется также термин *функционал.*

Чаще всего функция (правило) задается аналитически, то есть явной формулой, например: $f(x) = \sin x$ или $g(x) = \sqrt{1 - x^2}$. В этом случае естественной областью определения функции называют множество тех значений аргумента, для которых формула имеет смысл. Здесь для f это будет прямая $\mathbb{R}$, а для g — отрезок $[-1, 1]$. Вместе с тем может оказаться, что функция задана на меньшем множестве. Так, площадь квадрата является функцией S длины его стороны a: $S(a) = a^2$. Правая часть этой формулы имеет смысл для всех чисел a, но областью определения функции S будет лишь множество положительных чисел.

Подчеркнем, что многозначных отображений не бывает: при отображении каждому элементу $x \in X$ сопоставляется один элемент $y \in Y$, так что f в записи $f(x) = \pm\sqrt{x}$ — не функция.

Если элементы множества Y являются множествами, то отображение сопоставляет каждому $x \in X$ множество. Так, каждому $x \in \mathbb{R}$ можно поставить в соответствие множество $F(x)$ вещественных корней уравнения $t^2 = x$ и тем самым определить отображение F из $\mathbb{R}$ в множество всех подмножеств $\mathbb{R}$. Тогда

$$F(x) = \begin{cases} \{-\sqrt{x}, \sqrt{x}\}, & x > 0, \\ \{0\}, & x = 0, \\ \varnothing, & x < 0. \end{cases}$$

Тем не менее, термины “многозначное отображение” и “многозначная функция” используются в математике, но они требуют четких определений, исключающих всякий произвол в выборе значений. Пока нам эти термины не понадобятся, но в теории функций комплексной переменной они играют важную роль.

Определение. Отображение множества натуральных чисел в множество Y называется *последовательностью* в Y. Если Y — числовое множество, то последовательность называется числовой (например, вещественной или комплексной).

Итак, числовая последовательность — это функция натурального аргумента.

Последовательность будет обозначаться символом $\{x_n\}_{n=1}^{\infty}$ или просто $\{x_n\}$. Иногда вместо фигурных скобок используют круглые. Буква x в этой записи означает правило, как раньше f, только вместо $x(n)$ мы пишем x_n, записывая аргумент в виде индекса. Индекс n называется еще номером. Элемент x_n называется n-м членом последовательности.

Запись аргумента в виде индекса типична для последовательности, но употребляется и в других случаях. Семейство $\{f_x\}_{x\in X}$ элементов множества Y есть на самом деле отображение из X в Y, а f_x — другое обозначение для $f(x)$. Индексация и означает сопоставление каждому индексу $x \in X$ элемента $f(x) \in Y$.

Часто термин “последовательность” употребляют в более широком смысле и называют последовательностью отображение, заданное на некотором подмножестве множества $\mathbb{Z}$ целых чисел. Так, говорят о конечной или m-членной последовательности (упорядоченном наборе) $\{x_n\}_{n=1}^{m}$, а также о бесконечной в обе стороны последовательности $\{x_n\}_{n=-\infty}^{\infty}$. Если не оговорено противное, мы будем пользоваться первоначальным определением и считать, что последовательность задана на множестве натуральных чисел.

Над функциями $f, g\colon X \to \mathbb{R}$ $(\mathbb{C})$ и, в частности, числовыми последовательностями, определены арифметические операции: $f+g$, $f-g$, fg, cf (c — число). Например, сумма $f+g$ — это функция, действующая из X в $\mathbb{R}$ или $\mathbb{C}$ по правилу:

$$(f+g)(x) = f(x) + g(x), \qquad x \in X.$$

Если функция g не обращается в нуль на X, то на множестве X определено и частное $\frac{f}{g}$. В общем случае считают, что областью определения частного служит множество $D = \{x \in X : g(x) \neq 0\}$. Аналогичный смысл имеют символы f^2, $|f|$, $\max\{f, g\}$, $\min\{f, g\}$ и другие (в последних двух случаях функции f и g вещественнозначны).

Определение. Пусть $f\colon X \to Y$, A — множество. Множество

$$f(A) = \{y \in Y : \ \exists x \in A \ f(x) = y\}$$

называется *образом* множества A при отображении f. *Множеством значений* отображения f называется множество $f(X)$, то есть образ множества X.

Определение. Пусть $f\colon X \to Y$, B — множество. Множество

$$f^{-1}(B) = \{x \in X : \ f(x) \in B\}$$

называется *прообразом* множества B при отображении f.

Типичны ситуации, когда $A \subset X$ и $B \subset Y$; к ним можно свести и случаи произвольных множеств A и B. В самом деле, так как значения $f(x)$ определены только при $x \in X$ и принадлежат Y, верны равенства $f(A) = f(A \cap X)$, $f^{-1}(B) = f^{-1}(B \cap Y)$.

Если множество B состоит из одной точки: $B = \{y_0\}$, то его прообраз есть множество корней уравнения $f(x) = y_0$:

$$f^{-1}(\{y_0\}) = \{x \in X : f(x) = y_0\}.$$

В обозначении этого прообраза фигурные скобки иногда опускают и пишут $f^{-1}(y_0)$.

Из определения видно, что мы различаем множество значений $f(X)$ и область значений Y отображения f: $f(X)$ всегда содержится в Y, но может не совпадать с Y.

Определение. Пусть $f\colon X \to Y$. Если $f(X) = Y$, то отображение f называется *сюръективным,* или *сюръекцией,* или *отображением “на”* (отображением X на Y).

Другими словами, сюръективность отображения f означает, что при любом $y \in Y$ уравнение $f(x) = y$ имеет хотя бы одно решение в X.

Подчеркнем, что предлог “на” несет дополнительную смысловую нагрузку: выражения “отображение в Y” и “отображение на Y” имеют разный смысл.

Определение. Пусть $f\colon X \to Y$. Если для любых различных элементов X их образы различны, то отображение f называется *инъективным,* или *инъекцией,* или *обратимым* отображением.

Таким образом, f инъективно, если из того, что $x_1, x_2 \in X$, $x_1 \neq x_2$, следует, что $f(x_1) \neq f(x_2)$.

Другими словами, инъективность отображения f означает, что при любом $y \in Y$ уравнение $f(x) = y$ имеет не более одного решения в X.

Определение. Пусть $f\colon X \to Y$. Если отображение f одновременно сюръективно и инъективно, то f называется *биективным,* или *биекцией,* или *взаимно-однозначным* отображением (соответствием).

Другими словами, биективность отображения f означает, что при любом $y \in Y$ уравнение $f(x) = y$ имеет ровно одно решение в X.

Предупредим читателя о том, что некоторые авторы придерживаются другой терминологии и называют биективные отображения обратимыми, а инъективные — взаимно-однозначными (при этом “взаимно-однозначное соответствие” все-таки означает биекцию).

Отметим, что свойство отображения быть сюръективным, инъективным или биективным зависит не только от правила f, но и от множеств X и Y.

Пример. Пусть функция $f\colon X \to Y$ задана формулой $f(x) = x^2$.

1. Если $X = Y = \mathbb{R}$, то f не сюръективна (так как не принимает отрицательных значений) и не инъективна (так как, например, $f(-1) = f(1) = 1$).

2. Если $X = \mathbb{R}$, $Y = \mathbb{R}_+$, то f сюръективна, но не инъективна.
3. Если $X = \mathbb{R}_+$, $Y = \mathbb{R}$, то f инъективна, но не сюръективна.
4. Если $X = Y = \mathbb{R}_+$, то f сюръективна и инъективна, то есть биективна.

Ясно, что если отображение $f\colon X \to Y$ обратимо, то отображение $f\colon X \to f(X)$ биективно (мы сохранили обозначение для правила f).

Определение. Пусть $f\colon X \to Y$. *Графиком* отображения f называется множество

$$\Gamma_f = \{(x, y) : x \in X, y = f(x)\}.$$

Таким образом, $\Gamma_f \subset X \times Y$. В знакомой из школы ситуации, когда f — вещественнозначная функция вещественной переменной, график f есть подмножество плоскости.

График отображения обладает следующим свойством:

$$\text{если } (x, y_1), (x, y_2) \in \Gamma_f, \text{ то } y_1 = y_2.$$

На плоскости это означает, что никакая вертикальная прямая не может иметь двух общих точек с графиком. Обратно, если множество $G \subset X \times Y$ удовлетворяет условию:

$$\text{если } (x, y_1), (x, y_2) \in G, \text{ то } y_1 = y_2, \tag{7}$$

то G есть график некоторого отображения. Его областью определения служит множество

$$E = \{x \in X : \exists y \in Y \ \ (x, y) \in G\},$$

а правило таково: каждому $x \in E$ сопоставляется тот (единственный в силу (7)) элемент $y \in Y$, для которого $(x, y) \in G$.

Пусть $f\colon X \to Y$, f обратимо. Тогда для любого $y \in f(X)$ существует ровно один $x \in X$, такой что $f(x) = y$.

Определение. Пусть $f\colon X \to Y$, f обратимо. Отображение, которое каждому y из множества $f(X)$ сопоставляет то (единственное) значение x из X, для которого $f(x) = y$, называется *обратным* к f и обозначается f^{-1}.

Это определение разъясняет термин "обратимое отображение". Таким образом,

$$f^{-1}: f(X) \to X.$$

Очевидно, что f^{-1} — биекция между $f(X)$ и X. Соотношения $y = f(x)$ и $x = f^{-1}(y)$ для обратимого отображения равносильны. Следовательно, равносильны соотношения $(x, y) \in \Gamma_f$ и $(y, x) \in \Gamma_{f^{-1}}$. Для вещественных функций это значит, что графики обратимой функции и обратной к ней симметричны относительно прямой $y = x$ (рисунок 5). Если f — биекция, то обратное отображение к f^{-1} есть f.

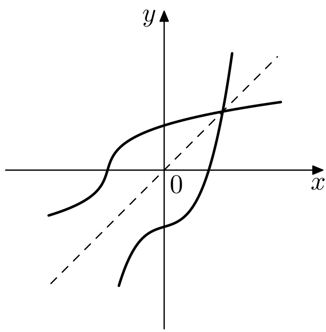

Рис. 5

Обозначение $f^{-1}(B)$ теперь получилось двусмысленным: с одной стороны, так обозначается прообраз множества B при отображении f, а с другой стороны, если f обратимо, — образ множества B при отображении f^{-1}. Это разночтение не приводит к путанице: читателю предлагается доказать, что в случае обратимости f эти два множества совпадают.

Определение. Пусть $f: X \to Y$, $g: Y_0 \to Z$, $f(X) \subset Y_0$. Отображение $h: X \to Z$, действующее по правилу

$$h(x) = g(f(x)), \qquad x \in X,$$

называется *композицией* или *суперпозицией* отображений f и g, а также *сложным отображением*, и обозначается $g \circ f$. При этом g называется *внешним*, а f — *внутренним отображением*.

Итак, по определению $(g \circ f)(x) = g(f(x))$. Для того, чтобы композиция была определена на множестве X, требуется, чтобы множество значений внутреннего отображения содержалось в области определения внешнего. Иногда в определении композиции этого не требуют, и тогда композиция оказывается определенной на множестве $f^{-1}(Y_0)$.

Отметим, что, вообще говоря, $g \circ f \neq f \circ g$, даже если обе композиции определены.

Определение. Отображение $\mathrm{id}_X\colon X \to X$, которое каждому элементу множества X сопоставляет сам этот элемент, называется *тождественным отображением* в множестве X.

Если f обратимо, то из определения обратного отображения вытекают равенства

$$f^{-1} \circ f = \mathrm{id}_X, \qquad f \circ f^{-1} = \mathrm{id}_{f(X)}\,.$$

Определение. Пусть $f\colon X \to Y$, $X_0 \subset X$. Отображение, которое каждому элементу x множества X_0 сопоставляет значение $f(x)$, называется *сужением* отображения f на множество X_0 и обозначается $f|_{X_0}$. Если отображение g есть сужение отображения f, то f называется *продолжением*, *распространением* или *расширением* g.

Таким образом, в определении сужения правило и область значений остаются теми же самыми, а меняется только область определения, то есть $f|_{X_0}\colon X_0 \to Y$. Ясно, что сужение единственно, а продолжение — нет.

Если $D \subset X$, $f\colon D \to Y$, то пишут $f\colon D \subset X \to Y$. Если же область определения отображения f содержит множество D, то говорят, что f задано *по крайней мере* на множестве D.

§ 4. Счетные множества

Пусть даны два конечных множества. Как узнать, одинаково ли число элементов в них? С одной стороны, можно просто пересчитать элементы и сравнить получившиеся в результате числа. С другой стороны, если в множествах одинаковое число элементов, то можно установить между ними взаимно-однозначное соответствие. Верно и обратное: если между множествами можно установить взаимно-однозначное соответствие, то число элементов в них одинаково. Второй способ, в отличие от первого, применим и к бесконечным множествам.

Определение. Множества A и B называют *эквивалентными* или *равномощными* и пишут $A \sim B$, если существует биекция $\varphi\colon A \to B$. Другими словами, два множества называются эквивалентными, если между ними можно установить взаимно-однозначное соответствие.

Примеры.

1. Противоположные стороны прямоугольника, очевидно, равномощны: друг другу сопоставляются противоположные точки.

2. Гипотенуза и катет прямоугольного треугольника также равномощны, хотя и имеют разные длины: взаимно-однозначным соответствием будет проекция гипотенузы на катет (рисунок 6).

3. Любые два невырожденных отрезка $[a, b]$ и $[c, d]$ равномощны: одна из биекций задается формулой

$$y = \frac{x-a}{b-a}\,d + \frac{b-x}{b-a}\,c.$$

Эта же формула задает биекцию интервалов (a, b) и (c, d) (рисунок 7).

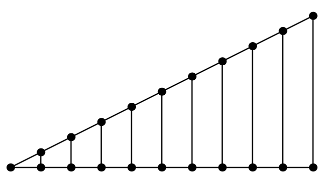

Рис. 6

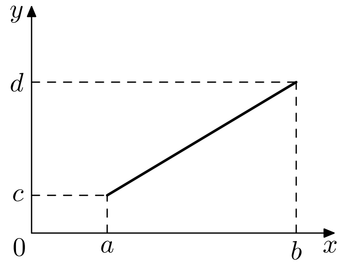

Рис. 7

4. Концентрические окружности равномощны: друг другу сопоставляются точки, лежащие на одном луче, выходящем из центра.

5. Интервал $(-1, 1)$ равномощен $\mathbb{R}$: одна из биекций задается формулой $y = \operatorname{tg} \frac{\pi x}{2}$.

6. Окружность без точки и прямая равномощны. Взаимно-однозначное соответствие, изображенное на рисунке 8, называется *стереографической проекцией.*

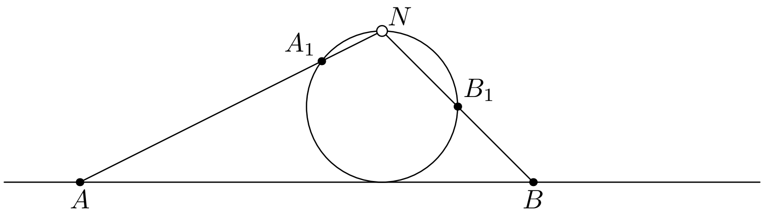

Рис. 8

Эквивалентность множеств является частным случаем общего понятия эквивалентности, которое определяется следующим образом.

Определение. Отношение $\sim$ между элементами некоторого множества называется *отношением эквивалентности,* если оно обладает следующими тремя свойствами.

1. Рефлексивность: $A \sim A$.
2. Симметричность: если $A \sim B$, то $B \sim A$.
3. Транзитивность: если $A \sim B$ и $B \sim C$, то $A \sim C$.

Примерами отношений эквивалентности служат равенство треугольников, подобие треугольников, параллельность прямых (если договориться считать прямую параллельной самой себе).

Замечание 1. Равномощность множеств является отношением эквивалентности.

Доказательство. Так как id_A — биекция множества A на себя, то $A \sim A$. Если $\varphi\colon A \to B$ — биекция, то $\varphi^{-1}\colon B \to A$ — тоже биекция, поэтому из того, что $A \sim B$, следует, что $B \sim A$. Если $\varphi\colon A \to B$ и $\psi\colon B \to C$ — биекции, то $\psi \circ \varphi\colon A \to C$ — тоже биекция, поэтому из того, что $A \sim B$ и $B \sim C$ следует, что $A \sim C$. $\square$

Определение. Множество называется *счетным*, если оно эквивалентно множеству натуральных чисел.

Таким образом, множество A счетно, если его элементы можно расположить в виде последовательности:

$$A = \{a_1, a_2, a_3, a_4, a_5, \dots\},$$

то есть занумеровать натуральными числами так, что каждый элемент будет занумерован ровно один раз и при этом будет израсходован весь натуральный ряд.

Примеры. 1. Множество

$$A = \{2, 4, 6, 8, 10, \dots\}$$

всех четных натуральных чисел счетно. Биекция: $a_n = 2n$.

2. Множество

$$A = \{1, 4, 9, 16, 25, \dots\}$$

всех квадратов натуральных чисел счетно. Биекция: $a_n = n^2$.

3. Множество целых чисел

$$\mathbb{Z} = \{0, -1, 1, -2, 2, \dots\}$$

счетно. Нумерация в указанном порядке задается формулами $a_{2n-1} = n - 1$, $a_{2n} = n$.

Теорема 1. *Всякое бесконечное множество содержит счетное подмножество.*

Доказательство. Пусть множество A бесконечно. Тогда в нем есть элемент a_1. Множество $A \setminus \{a_1\}$ бесконечно, поэтому в нем есть элемент a_2. Множество $A \setminus \{a_1, a_2\}$ также бесконечно, поэтому в нем есть элемент a_3. Ввиду бесконечности множества A этот процесс не оборвется ни на каком шаге; продолжая его и далее, получим множество $B = \{a_1, a_2, a_3, \dots\}$, которое по построению будет счетным подмножеством A. $\square$

Теорема 2. *Всякое бесконечное подмножество счетного множества счетно: если A счетно, $B \subset A$ и B бесконечно, то B счетно.*

Доказательство. Расположим элементы A в виде последовательности:

$$A = \{a_1, a_2, a_3, a_4, a_5, \dots\}.$$

Будем нумеровать элементы B в порядке их появления в этой последовательности. Тем самым каждый элемент B будет занумерован ровно один раз и, так как множество B бесконечно, для нумерации будет использован весь натуральный ряд. □

Теоремы 1 и 2 показывают, что счетные множества — самые "бедные элементами" бесконечные множества. Поэтому часто употребляется следующий термин.

Определение. Пустое, конечное или счетное множество называется *не более чем счетным*.

Замечание 1. Иногда счетными называют множества, которые только что были названы не более чем счетными, то есть множества, равномощные какому-либо подмножеству $\mathbb{N}$. Также нет общепринятого соглашения, считать ли пустое множество конечным или выделять в отдельный класс.

Лемма 1. *Пусть элементы множества A расположены в виде бесконечной в обоих направлениях таблицы (матрицы):*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots \\ a_{21} & a_{22} & a_{23} & \dots \\ a_{31} & a_{32} & a_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

Тогда A счетно.

Замечание 2. Нечеткое выражение "записаны в виде матрицы" на самом деле означает, что элементы множества A занумерованы с помощью упорядоченных пар натуральных чисел (первый элемент пары — номер строки, второй — номер столбца в матрице), причем каждый элемент занумерован ровно один раз и израсходованы все пары. Другими словами, задана биекция множеств $\mathbb{N} \times \mathbb{N}$ и A. Тем самым лемма 1 утверждает счетность множества $\mathbb{N} \times \mathbb{N}$.

Доказательство. Занумеруем элементы множества A “по диагоналям”:

$$A = \{a_{11}, a_{12}, a_{21}, a_{31}, a_{22}, a_{33}, \dots\}$$

или “по квадратам”:

$$A = \{a_{11}, a_{12}, a_{22}, a_{21}, \dots\}. \quad \square$$

Теорема 3. *Не более чем счетное объединение не более чем счетных множеств не более чем счетно.*

Доказательство. Пусть $B = \bigcup\limits_{k=1}^{n} A_k$ или $B = \bigcup\limits_{k=1}^{\infty} A_k$, множества A_k не более чем счетны. Запишем элементы A_1 в первую строку матрицы, элементы $A_2 \setminus A_1$ — во вторую строку и так далее, то есть если задано множество A_k, то элементы $A_k \setminus \bigcup\limits_{j=1}^{k-1} A_j$ запишем в k-ю строку матрицы. Таким образом все элементы множества B окажутся записанными в клетки матрицы (но при этом некоторые клетки могут остаться пустыми). Значит, B равномощно некоторому подмножеству счетного по лемме 1 множества $\mathbb{N} \times \mathbb{N}$. По теореме 2 множество B не более чем счетно. $\square$

Теорема 4. *Множество рациональных чисел счетно.*

Доказательство. Обозначим

$$\mathbb{Q}_+ = \{x \in \mathbb{Q} : x > 0\}, \qquad \mathbb{Q}_- = \{x \in \mathbb{Q} : x < 0\}.$$

При всех $q \in \mathbb{N}$ множество $Q_q = \left\{\frac{1}{q}, \frac{2}{q}, \frac{3}{q}, \dots\right\}$ счетно. По теореме 3 и $\mathbb{Q}_+ = \bigcup\limits_{q=1}^{\infty} Q_q$ счетно. Очевидно, что $\mathbb{Q}_- \sim \mathbb{Q}_+$. Снова по теореме 3 множество

$$\mathbb{Q} = \mathbb{Q}_+ \cup \mathbb{Q}_- \cup \{0\}$$

счетно. $\square$

Следствие 1. *Если $a, b \in \mathbb{R}$, $a < b$, то $\mathbb{Q} \cap (a, b)$ счетно.*

До сих пор мы не привели ни одного примера бесконечного множества, про которое было бы доказано, что оно не является счетным (такие множества называют *несчетными*).

Теорема 5. *Отрезок* $[0,1]$ *несчетен.*

Доказательство. Допустим противное: пусть отрезок $[0,1]$ счетен, то есть все числа отрезка $[0,1]$ можно расположить в виде последовательности:

$$[0,1] = \{x_1, x_2, x_3, x_4, x_5, \dots\}$$

Разобьем отрезок $[0,1]$ на три равных отрезка $\left[0, \frac{1}{3}\right]$, $\left[\frac{1}{3}, \frac{2}{3}\right]$ и $\left[\frac{2}{3}, 1\right]$ и обозначим через $[a_1, b_1]$ тот из них, который не содержит точки x_1 (если таких два, то все равно, какой). Далее разобьем отрезок $[a_1, b_1]$ на три равных отрезка и обозначим через $[a_2, b_2]$ любой из них, который не содержит точки x_2. Этот процесс продолжим неограниченно. В результате мы построим последовательность вложенных отрезков $\{[a_n, b_n]\}_{n=1}^{\infty}$, причем $x_n \notin [a_n, b_n]$ для любого n. По аксиоме о вложенных отрезках существует точка x^*, принадлежащая одновременно всем отрезкам $[a_n, b_n]$. Тем более, $x^* \in [0,1]$. Тогда $x^* = x_m$ при некотором $m \in \mathbb{N}$. Но по построению $x^* \notin [a_m, b_m]$, что противоречит принадлежности x^* всем отрезкам $[a_n, b_n]$. $\square$

Из доказанной теоремы вытекает существование иррациональных чисел.

Следствие 2. *Множества вещественных чисел* $\mathbb{R}$ *и иррациональных чисел* $\mathbb{R} \setminus \mathbb{Q}$ *несчетны.*

Определение. Если множество эквивалентно отрезку $[0,1]$, то говорят, что оно имеет *мощность континуума.*

Отметим без доказательства еще несколько фактов о мощности множеств.

Замечание 1. Ясно, что любой невырожденный отрезок имеет мощность континуума. Можно доказать, что любой невырожденный промежуток и, в частности, вся прямая $\mathbb{R}$, имеет мощность континуума. Более того, при любом $m \in \mathbb{N}$ множество $\mathbb{R}^m$ имеет мощность континуума.

Замечание 2. Множество всех функций $f\colon [0,1] \to \mathbb{R}$ более богато элементами, чем отрезок $[0,1]$, то есть оно не равномощно отрезку, но имеет часть, равномощную отрезку.

Гипотеза континуума. *Всякое бесконечное подмножество* $\mathbb{R}$ *равномощно* $\mathbb{N}$ *или* $\mathbb{R}$.

Эта гипотеза была сформулирована Кантором. В 1934 году австрийский математик К. Гёдель доказал, что гипотеза континуума не противоречит остальным аксиомам теории множеств. В 1964 году американский математик П. Коэн доказал, что ее отрицание также не противоречит остальным аксиомам теории множеств.

Замечание 3. Мы определили равномощность множеств, нигде не определяя, что такое мощность множества, и не пользуясь этим понятием. Чтобы определить понятие мощности множества, можно поступить так. Разобьем все множества на классы эквивалентности: два множества попадают в один класс, если они эквивалентны. После этого каждый класс назовем мощностью. Обычно мощность n-элементного множества (класс n-элементных множеств) так и обозначают числом n, а мощность континуума — буквой c.

Замечание 4. Несложно доказать, что если из бесконечного множества удалить один элемент или даже любое конечное множество элементов, то получится множество, равномощное исходному. Таким образом, всякое бесконечное множество имеет равномощную себе правильную (то есть не совпадающую с ним самим) часть. Конечные множества этим свойством не обладают, поэтому можно даже дать определение бесконечного множества как множества, имеющего равномощную себе правильную часть.

ГЛАВА 2. ТЕОРИЯ ПРЕДЕЛОВ И НЕПРЕРЫВНЫЕ ФУНКЦИИ

§ 1. Предел последовательности

Начнем с определения предела числовой последовательности, вероятно, знакомого читателю из школьного курса.

Мы будем записывать это и многие другие определения кратко с помощью кванторов. Во избежание ошибок читателю рекомендуется вспомнить соглашения о правилах чтения формул с кванторами из § 1 введения. Одновременно мы будем давать словесную формулировку, полностью расшифровывающую краткую запись. Удобно представлять себе, что, в то время как краткая формулировка записывается (лектором на доске или студентом при ответе на экзамене), словесная формулировка произносится.

Определение. Пусть $\{x_n\}_{n=1}^{\infty}$ — последовательность вещественных чисел. Число $a \in \mathbb{R}$ называют *пределом последовательности* $\{x_n\}$ и пишут

$$\lim_{n\to\infty} x_n = a \quad \text{или} \quad x_n \underset{n\to\infty}{\longrightarrow} a,$$

если для любого положительного числа ε существует такой номер N, что для всех номеров n, больших N, выполняется неравенство $|x_n - a| < \varepsilon$:

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \in \mathbb{N} : n > N \;\; |x_n - a| < \varepsilon.$$

Последовательность, имеющая некоторое число своим пределом, называется *сходящейся*; в противном случае (то есть если никакое число не является пределом этой последовательности) — *расходящейся.*

Иногда, если нет опасности недоразумений, запись $n \to \infty$ в обозначении предела опускают и пишут

$$\lim x_n = a \quad \text{или} \quad x_n \to a.$$

Говорят также “x_n сходится к a” и “x_n стремится к a”.

Пример 1. Докажем по определению, что число $a = 0$ является пределом последовательности $x_n = \frac{1}{n}$.

Возьмем $\varepsilon > 0$ (с этой фразы удобно начинать доказательство утверждений на "ε-языке"). Постараемся подобрать такое натуральное N, что для всех номеров $n > N$ будет $\left|\frac{1}{n} - 0\right| < \varepsilon$, то есть $\frac{1}{n} < \varepsilon$. Положим $N = \left[\frac{1}{\varepsilon}\right] + 1$; тогда, если $n > N$, то, тем более, $n > \frac{1}{\varepsilon}$ и $\frac{1}{n} < \varepsilon$. В силу произвольности ε число 0 действительно является пределом последовательности $\frac{1}{n}$.

Пример 2. Пусть все члены последовательности x_n равны одному и тому же значению a. Такая последовательность называется постоянной или *стационарной.* Докажем, что и $\lim x_n = a$.

Действительно, неравенство $|x_n - a| = 0 < \varepsilon$ выполняется для всех $\varepsilon > 0$ и $n \in \mathbb{N}$, поэтому в определении подходит любой номер N, например, $N = 1$.

Пользуясь правилом отрицания утверждений с кванторами, запишем тот факт, что число a не является пределом последовательности $\{x_n\}$:

$$\exists \varepsilon^* > 0 \ \ \forall N \in \mathbb{N} \ \ \exists n \in \mathbb{N} : n > N, \ |x_n - a| \geqslant \varepsilon^*.$$

Пример 3. Докажем, что последовательность $x_n = (-1)^n$ не имеет предела, то есть никакое число a не является ее пределом.

Пусть $a \in \mathbb{R}$. Положим $\varepsilon^* = 1$ и возьмем натуральное N. Положим $n = 2N$, если $a < 0$, и $n = 2N + 1$, если $a \geqslant 0$. Тогда $n > N$ и $|(-1)^n - a| \geqslant 1$, то есть $a \neq \lim\,(-1)^n$.

Сделаем несколько простых, но важных замечаний к определению предела.

Замечание 1. Если для некоторого положительного числа ε_0 нашелся номер N из определения предела, то тот же номер N подходит и для любого числа $\varepsilon > \varepsilon_0$. Действительно, если $|x_n - a| < \varepsilon_0$, а $\varepsilon > \varepsilon_0$, то и подавно $|x_n - a| < \varepsilon$. Поэтому достаточно проверять справедливость определения (существование номера N) лишь для достаточно малых ε (то есть для всех положительных ε, меньших некоторого положительного числа).

Замечание 2. Если для некоторого положительного числа ε_0 нашелся номер N из определения предела, то для того же ε подходит и любой номер N_1, больший N. Действительно, если неравенство $|x_n - a| < \varepsilon$ выполняется для всех номеров $n > N$, а $N_1 > N$, то оно и подавно выполняется для всех номеров $n > N_1$. По той же причине в определении предела не обязательно требовать, чтобы N было натуральным числом. Указание на то, что n — натуральное число, мы также будем обычно опускать для краткости.

Так, в примере 1 можно было положить $N = \left[\frac{1}{\varepsilon}\right]$ (что равно нулю при $\varepsilon > 1$) или даже $N = \frac{1}{\varepsilon}$ (что не обязано быть целым).

Замечание 3. Определение предела останется равносильным исходному, если вместо $n > N$ писать $n \geqslant N$ и (или) вместо $|x_n - a| < \varepsilon$ писать $|x_n - a| \leqslant \varepsilon$.

Поясним это для неравенств с $|x_n - a|$. С одной стороны, из неравенства $|x_n - a| < \varepsilon$ следует неравенство $|x_n - a| \leqslant \varepsilon$. Поэтому если число a удовлетворяет определению со знаком "$<$", то оно и подавно удовлетворяет определению со знаком "$\leqslant$". С другой стороны, пусть число a удовлетворяет определению со знаком "$\leqslant$". Возьмем $\varepsilon > 0$ и подберем такой номер N, что для всех номеров n, больших N, будет $|x_n - a| \leqslant \frac{\varepsilon}{2}$. Тогда тем более $|x_n - a| < \varepsilon$. В силу произвольности ε число a удовлетворяет определению со знаком "$<$".

Замечание 4. Пусть последовательности $\{x_n\}_{n=1}^{\infty}$ и $\{y_n\}_{n=1}^{\infty}$ таковы, что, начиная с некоторого номера n_0, их значения совпадают: $x_n = y_n$ при всех $n \geqslant n_0$. Тогда их пределы одновременно существуют или нет и, если существуют, то равны.

В самом деле, если $a = \lim x_n$ и номер N подходит для числа ε из определения $\lim x_n$, то номер $N_1 = \max\{N, n_0 - 1\}$ подходит для числа ε из определения $\lim y_n$. Аналогичное рассуждение верно, если поменять x_n и y_n ролями.

Таким образом, если изменить у последовательности x_n конечное число членов, то ее предел не изменится. Сказанное позволяет, не меняя определения, говорить о пределе последовательности вида $\{z_n\}_{n=n_0}^{\infty}$, заданной на множестве целых чисел, начиная с некоторого n_0 ($n_0 \in \mathbb{Z}$).

Замечание 5. Неравенство для модуля

$$|x_n - a| < \varepsilon$$

равносильно двойному неравенству

$$a - \varepsilon < x_n < a + \varepsilon.$$

Определение. Интервал $(a - \varepsilon, a + \varepsilon)$ называется *ε-окрестностью* точки a и обозначается $V_a(\varepsilon)$ или V_a, если значение ε несущественно.

Таким образом, определение предела последовательности можно переформулировать, используя понятие окрестности, так: число a называется пределом последовательности $\{x_n\}$, если для любой окрестности точки a все члены последовательности, начиная с некоторого номера, принадлежат этой окрестности.

Отметим еще, что пересечение двух окрестностей точки a есть ее окрестность:

$$V_a(\varepsilon_1) \cap V_a(\varepsilon_2) = V_a\big(\min\{\varepsilon_1, \varepsilon_2\}\big).$$

Теорема 1. Единственность предела последовательности. *Последовательность не может иметь более одного предела: если $a, b \in \mathbb{R}$, $x_n \to a$, $x_n \to b$, то $a = b$.*

Доказательство. Предположим противное: пусть $a \neq b$. Тогда $|a - b| > 0$. Возьмем $\varepsilon = \frac{|a-b|}{2}$. По определению предела найдутся такие номера N_1 и N_2, что $|x_n - a| < \varepsilon$ для всех $n > N_1$, и $|x_n - b| < \varepsilon$ для всех $n > N_2$. Тогда, если $n > \max\{N_1, N_2\}$, то по свойству 5 модуля из § 2 главы 1

$$|a - b| \leqslant |a - x_n| + |x_n - b| < \varepsilon + \varepsilon = |a - b|,$$

что абсурдно. $\square$

Определение. Последовательность $\{x_n\}$ называется *ограниченной,* если множество ее значений ограничено:

$$\exists R > 0 \ \ \forall n \in \mathbb{N} \ \ |x_n| \leqslant R.$$

Теорема 2. *Сходящаяся последовательность ограничена.*

Доказательство. Пусть $x_n \to a$. Взяв $\varepsilon = 1$, подберем такой номер N, что для всех номеров $n > N$ будет $|x_n - a| < 1$. Тогда при всех $n > N$

$$|x_n| \leqslant |a| + |x_n - a| < |a| + 1.$$

Положим

$$R = \max\{|x_1|, \ldots, |x_N|, |a| + 1\};$$

тогда $|x_n| \leqslant R$ при всех $n \in \mathbb{N}$. $\square$

Теорема 3. Предельный переход в неравенстве. *Пусть* $\{x_n\}$, $\{y_n\}$ *— вещественные последовательности,* $x_n \leqslant y_n$ *при всех* $n \in \mathbb{N}$, $a, b \in \mathbb{R}$, $x_n \to a$, $y_n \to b$. *Тогда* $a \leqslant b$. Другими словами: *если* $x_n \leqslant y_n$ *при всех* $n \in \mathbb{N}$ *и существуют пределы* $\{x_n\}$ *и* $\{y_n\}$, *то* $\lim x_n \leqslant \lim y_n$.

Доказательство. Предположим противное: пусть $a > b$. Тогда $\varepsilon = \frac{a-b}{2}$ положительно. По определению предела найдутся такие номера N_1 и N_2, что $a - \varepsilon < x_n$ для всех $n > N_1$, а $y_n < b + \varepsilon$ для всех $n > N_2$. Значит, если $n > \max\{N_1, N_2\}$, то

$$y_n < b + \varepsilon = \frac{a+b}{2} = a - \varepsilon < x_n,$$

что противоречит условию. $\square$

Замечание 1. Как показывает пример последовательностей $x_n = -\frac{1}{n}$ и $y_n = \frac{1}{n}$, стремящихся к нулю, при предельном переходе строгое неравенство может превратиться в нестрогое: из того, что $x_n < y_n$ при всех n, не следует, что $\lim x_n < \lim y_n$.

Следствие 1. 1. *Если* $x_n \leqslant b$ *при всех* $n \in \mathbb{N}$ *и существует* $\lim x_n$, *то* $\lim x_n \leqslant b$.

2. *Если* $x_n \geqslant a$ *при всех* $n \in \mathbb{N}$ *и существует* $\lim x_n$, *то* $\lim x_n \geqslant a$.

3. *Если* $x_n \in [a, b]$ *при всех* $n \in \mathbb{N}$ *и существует* $\lim x_n$, *то* $\lim x_n \in [a, b]$.

Для доказательства первого утверждения нужно взять в качестве $\{y_n\}$ стационарную последовательность: $y_n = b$ при всех n.

Второе утверждение доказывается аналогично, третье следует из первых двух.

Замечание 2. Свойство отрезка из утверждения 3 следствия (неверное для конечных промежутков другого типа) называется *замкнутостью.*

Теорема 4. О сжатой последовательности. *Пусть* $\{x_n\}$, $\{y_n\}$, $\{z_n\}$ — *вещественные последовательности,* $x_n \leqslant y_n \leqslant z_n$ *при всех* $n \in \mathbb{N}$, $a \in \mathbb{R}$, $\lim x_n = \lim z_n = a$. *Тогда предел* $\{y_n\}$ *существует и равен* a.

Доказательство. Возьмем $\varepsilon > 0$. По определению предела найдутся такие номера N_1 и N_2, что $a - \varepsilon < x_n$ для всех $n > N_1$ и $z_n < a + \varepsilon$ для всех $n > N_2$. Положим $N = \max\{N_1, N_2\}$. Тогда при всех $n > N$

$$a - \varepsilon < x_n \leqslant y_n \leqslant z_n < a + \varepsilon.$$

В силу произвольности ε предел $\{y_n\}$ существует и равен a. □

Замечание 1. В одном старом учебнике теорема 4 сопровождалась рисунком, на котором два милиционера ($\{x_n\}$ и $\{z_n\}$) под руки вели нарушителя ($\{y_n\}$) в отделение (a). С тех пор теорему о сжатой последовательности (и ее обобщение — теорему о сжатой функции) называют еще *принципом двух милиционеров.*

Замечание 2. Отметим следующий частный случай теоремы 4: *если* $|y_n| \leqslant z_n$ *при всех* $n \in \mathbb{N}$ *и* $z_n \to 0$, *то* $y_n \to 0$.

Замечание 3. В теоремах 3 и 4 достаточно выполнения неравенств для всех номеров $n \geqslant n_0$, где $n_0 \in \mathbb{N}$.

Определение. Числовая последовательность называется *бесконечно малой,* если она стремится к нулю.

Таким образом, стремление последовательности $\{x_n\}$ к a равносильно тому, что $\{x_n - a\}$ бесконечно мала.

Лемма 1. *Произведение бесконечно малой последовательности на ограниченную есть бесконечно малая: если* $\{x_n\}$, $\{y_n\}$ — *вещественные последовательности,* $\{x_n\}$ — *бесконечно малая,* $\{y_n\}$ *ограничена, то* $\{x_n y_n\}$ — *бесконечно малая.*

Доказательство. В силу ограниченности $\{y_n\}$ найдется такое $K > 0$, что $|y_n| \leqslant K$ при всех n. Возьмем $\varepsilon > 0$. По определению предела последовательности $\{x_n\}$ существует такой номер N, что $|x_n| \leqslant \frac{\varepsilon}{K}$ для всех $n > N$. Но тогда для всех $n > N$

$$|x_n y_n| < \frac{\varepsilon}{K} \cdot K = \varepsilon.$$

В силу произвольности ε это и означает, что $x_n y_n \to 0$. $\square$

Теорема 5. Арифметические действия над сходящимися последовательностями. *Пусть* $\{x_n\}$, $\{y_n\}$ *— вещественные последовательности,* $x_0, y_0 \in \mathbb{R}$, $x_n \to x_0$, $y_n \to y_0$. *Тогда*

1) $x_n + y_n \to x_0 + y_0$;
2) $x_n y_n \to x_0 y_0$;
3) $x_n - y_n \to x_0 - y_0$;
4) $|x_n| \to |x_0|$;
5) *если, кроме того,* $y_n \neq 0$ *при всех* n *и* $y_0 \neq 0$, *то* $\frac{x_n}{y_n} \to \frac{x_0}{y_0}$.

Доказательство. 1. Возьмем $\varepsilon > 0$. По определению предела найдутся такие номера N_1 и N_2, что $|x_n - x_0| < \frac{\varepsilon}{2}$ для всех $n > N_1$ и $|y_n - y_0| < \frac{\varepsilon}{2}$ для всех $n > N_2$. Положим $N = \max\{N_1, N_2\}$. Тогда при всех $n > N$ будет

$$|(x_n + y_n) - (x_0 + y_0)| \leqslant |x_n - x_0| + |y_n - y_0| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

что и доказывает утверждение 1.

2. По свойствам модуля

$$\begin{aligned} |x_n y_n - x_0 y_0| &= |(x_n - x_0) y_n + x_0 (y_n - y_0)| \leqslant \\ &\leqslant |x_n - x_0||y_n| + |x_0||y_n - y_0|. \end{aligned} \tag{1}$$

По условию последовательности $\{|x_n - x_0|\}$ и $\{|y_n - y_0|\}$ бесконечно малые, а по теореме 2 последовательность $\{|y_n|\}$ ограничена, как и постоянная последовательность $\{|x_0|\}$. Следовательно, по лемме 1 оба слагаемых в правой части (1) бесконечно малы, а тогда, по доказанному утверждению 1, и их сумма бесконечно мала. Наконец, $|x_n y_n - x_0 y_0| \to 0$ по замечанию 2 к теореме 4, что и требовалось доказать.

3. Утверждение доказывается аналогично утверждению 1 или применением уже доказанных утверждений 1 и 2:

$$x_n - y_n = x_n + (-1)y_n \to x_0 + (-1)y_0 = x_0 - y_0.$$

4. Утверждение следует из неравенства

$$\big||x_n| - |x_0|\big| \leqslant |x_n - x_0|$$

и замечания 2 к теореме 4.

5. Достаточно доказать, что $\frac{1}{y_n} \to \frac{1}{y_0}$, так как тогда по утверждению 2 получим, что

$$\frac{x_n}{y_n} = x_n \cdot \frac{1}{y_n} \to x_0 \cdot \frac{1}{y_0} = \frac{x_0}{y_0}.$$

Поскольку

$$\frac{1}{y_n} - \frac{1}{y_0} = (y_0 - y_n) \cdot \frac{1}{y_0} \cdot \frac{1}{y_n},$$

последовательность $\{y_0 - y_n\}$ бесконечно малая, а последовательность $\left\{\frac{1}{y_0}\right\}$ ограниченная, по лемме 1 остается доказать ограниченность последовательности $\left\{\frac{1}{y_n}\right\}$.

По определению предела для числа $\varepsilon = \frac{|y_0|}{2}$, которое по условию положительно, существует такой номер N, что $|y_n - y_0| < \varepsilon$ для всех $n > N$. Тогда при всех $n > N$ по свойствам модуля

$$|y_n| = |y_0 + y_n - y_0| \geqslant |y_0| - |y_n - y_0| > |y_0| - \varepsilon = \frac{|y_0|}{2}.$$

Обозначим $k = \min\left\{|y_1|, \ldots, |y_n|, \frac{|y_0|}{2}\right\}$. Тогда $k > 0$ и $|y_n| \geqslant k$ при всех n. Следовательно, $\left|\frac{1}{y_n}\right| \leqslant \frac{1}{k}$ при всех n, что и означает ограниченность последовательности $\left\{\frac{1}{y_n}\right\}$. $\square$

Определение. Говорят, что вещественная последовательность $\{x_n\}$ *стремится к:*

1) *плюс бесконечности*, и пишут

$$\lim_{n\to\infty} x_n = +\infty \quad \text{или} \quad x_n \xrightarrow[n\to\infty]{} +\infty,$$

если для любого положительного числа E существует такой номер N, что для всех номеров n, больших N, выполняется неравенство $x_n > E$:

$$\forall E > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \in \mathbb{N} : n > N \;\; x_n > E;$$

2) *минус бесконечности,* и пишут

$$\lim_{n\to\infty} x_n = -\infty \quad \text{или} \quad x_n \xrightarrow[n\to\infty]{} -\infty,$$

если для любого положительного числа E существует такой номер N, что для всех номеров n, больших N, выполняется неравенство $x_n < -E$:

$$\forall E > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \in \mathbb{N} : n > N \;\; x_n < -E;$$

3) *бесконечности (бесконечности неопределенного знака),* и пишут

$$\lim_{n\to\infty} x_n = \infty \quad \text{или} \quad x_n \xrightarrow[n\to\infty]{} \infty,$$

если для любого положительного числа E существует такой номер N, что для всех номеров n, больших N, выполняется неравенство $|x_n| > E$:

$$\forall E > 0 \;\; \exists N \in \mathbb{N} \;\; \forall n \in \mathbb{N} : n > N \;\; |x_n| > E.$$

Замечание 1. Из определений следует, что если $x_n \to \pm\infty$, то $x_n \to \infty$. Обратное неверно: последовательность $x_n = (-1)^n n$ стремится к бесконечности, но не стремится ни к плюс, ни к минус бесконечности.

Определение. Последовательность, стремящаяся к бесконечности, называется *бесконечно большой.*

Замечание 2. Из определений следует, что если $x_n \to \infty$, то x_n не ограничена. Обратное неверно: последовательность $x_n = (1 + (-1)^n)n$ не ограничена и не стремится к бесконечности.

Замечание 3. Ясно, что последовательность не может одновременно стремиться к числу и к $\pm\infty$ (∞), а также к бесконечностям разных знаков. Другими словами, имеет место единственность предела в $\overline{\mathbb{R}}$: *если* $a, b \in \overline{\mathbb{R}}$, $x_n \to a$, $x_n \to b$, *то* $a = b$.

Замечание 4. Определение достаточно проверять лишь для достаточно больших чисел E; можно также опустить в определении требование $E > 0$.

Замечание 5. Определение сходящейся последовательности не меняется: последовательность называется сходящейся, если она имеет конечный предел. Бесконечно большие последовательности считаются расходящимися.

Замечание 6. Определение предела на языке окрестностей без изменений переносится на бесконечные пределы: $x_n \to a$, если для любой окрестности точки a все члены последовательности, начиная с некоторого номера, принадлежат этой окрестности. Для этого надо определить окрестности бесконечно удаленных точек:

$$V_{+\infty} = (E, +\infty],\ V_{-\infty} = [-\infty, -E),\ V_{\infty} = \{x \in \mathbb{R} : |x| > E\} \cup \{\infty\}.$$

Если же нужно определить именно ε-окрестности, то полагают $V_{+\infty}(\varepsilon) = \left(\frac{1}{\varepsilon}, +\infty\right]$ и т.п.; тогда ε-окрестности сужаются с уменьшением $\varepsilon > 0$.

Замечание 7. 1. *Если* $x_n \geqslant y_n$ *при всех* n *и* $y_n \to +\infty$, *то* $x_n \to +\infty$.

2. *Если* $x_n \leqslant y_n$ *при всех* n *и* $y_n \to -\infty$, *то* $x_n \to -\infty$.

Действительно, в определении предела для числа E и последовательности $\{x_n\}$ подходит тот же номер N, что и для числа E и последовательности $\{y_n\}$. Это замечание дополняет теорему о сжатой последовательности.

Лемма 2. Связь между бесконечно большими и бесконечно малыми. *Пусть* $\{x_n\}$ *— вещественная последовательность,* $x_n \neq 0$ *ни при каком* n. *Тогда последовательность* $\{x_n\}$ *— бесконечно большая в том и только в том случае, когда* $\left\{\dfrac{1}{x_n}\right\}$ *— бесконечно малая.*

Доказательство. Пусть $x_n \to \infty$; докажем, что $\dfrac{1}{x_n} \to 0$. Возьмем $\varepsilon > 0$, и для числа $E = \dfrac{1}{\varepsilon}$ подберем такой номер N, что для всех $n > N$ будет $|x_n| > E$. Последнее равносильно неравенству $\left|\dfrac{1}{x_n}\right| < \varepsilon$, что и означает стремление $\dfrac{1}{x_n}$ к нулю. Доказательство в обратную сторону проводится аналогично. $\square$

Следующая теорема дополняет теорему об арифметических действиях над сходящимися последовательностями. В ней, если не оговорено противное, пределы могут быть и бесконечными.

Теорема 6. Арифметические действия над бесконечно большими. *Пусть $\{x_n\}$, $\{y_n\}$ — вещественные последовательности, $a, b \in \overline{\mathbb{R}} \cup \{\infty\}$.*

1. *Если $x_n \to +\infty$, $\{y_n\}$ ограничена снизу, то $x_n + y_n \to +\infty$.*
2. *Если $x_n \to -\infty$, $\{y_n\}$ ограничена сверху, то $x_n + y_n \to -\infty$.*
3. *Если $x_n \to \infty$, $\{y_n\}$ ограничена, то $x_n + y_n \to \infty$.*
4. *Если $x_n \to \pm\infty$, $y_n \geqslant b > 0$ для всех n (или $y_n \to b > 0$), то $x_n y_n \to \pm\infty$.*
5. *Если $x_n \to \pm\infty$, $y_n \leqslant b < 0$ для всех n (или $y_n \to b < 0$), то $x_n y_n \to \mp\infty$.*
6. *Если $x_n \to \infty$, $|y_n| \geqslant b > 0$ для всех n (или $y_n \to b \neq 0$), то $x_n y_n \to \infty$.*
7. *Если $x_n \to a \neq 0$, $y_n \to 0$, $y_n \neq 0$ при всех n, то $\frac{x_n}{y_n} \to \infty$.*
8. *Если $x_n \to a \in \mathbb{R}$, $y_n \to \infty$, то $\frac{x_n}{y_n} \to 0$.*
9. *Если $x_n \to \infty$, $y_n \to b \in \mathbb{R}$, $y_n \neq 0$ при всех n, то $\frac{x_n}{y_n} \to \infty$.*

Доказательство. Ограничимся доказательством утверждений 1, 6 и 8.

1. Возьмем $E > 0$. По определению ограниченности снизу найдется такое число $m \in \mathbb{R}$, что $y_n \geqslant m$ при всех n. По определению бесконечного предела существует такой номер N, что $x_n > E - m$ для всех $n > N$. Тогда для всех $n > N$

$$x_n + y_n > E - m + m = E.$$

В силу произвольности E это означает, что $x_n + y_n \to +\infty$.

6. Пусть $|y_n| \geqslant b > 0$ для всех n. Возьмем $E > 0$. По определению бесконечного предела существует такой номер N, что $|x_n| > \frac{E}{b}$ для всех $n > N$. Тогда для всех $n > N$

$$|x_n y_n| > \frac{E}{b} \cdot b = E,$$

что и означает стремление $x_n y_n$ к ∞.

Пусть $y_n \to b \neq 0$. Положим $b_1 = \frac{|b|}{2}$, если b — число, и $b_1 = 1$, если b — бесконечность. Тогда, начиная с некоторого номера, $|y_n| \geqslant b_1$, и применимо только что доказанное утверждение.

8. По теореме о пределе произведения и лемме 2

$$\frac{x_n}{y_n} = x_n \cdot \frac{1}{y_n} \to a \cdot 0 = 0.$$

Доказательство остальных утверждений теоремы остается читателю в качестве несложного упражнения. □

Замечание 1. Часть утверждений теорем об арифметических действиях можно объединить следующей формулировкой. *Если $\{x_n\}$, $\{y_n\}$ — вещественные последовательности, $x_n \to a \in \overline{\mathbb{R}}$, $y_n \to b \in \overline{\mathbb{R}}$, знак $*$ означает одно из четырех арифметических действий и $a * b$ определено в $\overline{\mathbb{R}}$, то $x_n * y_n \to a * b$.*

Теоремы об арифметических действиях не позволяют сделать заключение о значении предела в следующих четырех случаях.

1. $x_n \to +\infty$, $y_n \to -\infty$.	$x_n + y_n \to {?}$
2. $x_n \to 0$, $y_n \to \infty$.	$x_n y_n \to {?}$
3. $x_n \to 0$, $y_n \to 0$.	$\dfrac{x_n}{y_n} \to {?}$
4. $x_n \to \infty$, $y_n \to \infty$.	$\dfrac{x_n}{y_n} \to {?}$

В этих ситуациях без дополнительной информации об x_n и y_n ничего нельзя сказать о пределе. Приведем примеры для суммы.

Примеры.

1а. Если $x_n = n + a$ $(a \in \mathbb{R})$, $y_n = -n$, то $x_n + y_n = a \to a$.

1б. Если $x_n = n^2 + n$, $y_n = -n^2$, то $x_n + y_n = n \to +\infty$.

1в. Если $x_n = n^2$, $y_n = -n^2 - n$, то $x_n + y_n = -n \to -\infty$.

1г. Если $x_n = n + (-1)^n$, $y_n = -n$, то $x_n + y_n = (-1)^n$ не имеет ни конечного, ни бесконечного предела.

Таким образом, предел суммы может оказаться равным любому числу, плюс или минус бесконечности или вовсе не существовать.

Читатель сам приведет аналогичные примеры для трех оставшихся случаев.

В ситуациях 1–4 говорят, что имеет место *неопределенность* вида $\infty-\infty$, $0\cdot\infty$, $\frac{0}{0}$ и $\frac{\infty}{\infty}$ соответственно, а нахождение предела называют *раскрытием неопределенности.*

§ 2. Точные границы числовых множеств и монотонные последовательности

Говорят, что $\{[a_n, b_n]\}_{n=1}^{\infty}$ — последовательность *стягивающихся* отрезков, если $a_n \leqslant a_{n+1} \leqslant b_{n+1} \leqslant b_n$ при всех n и $b_n - a_n \to 0$.

Теорема 1. О стягивающихся отрезках. *Пусть $\{[a_n, b_n]\}_{n=1}^{\infty}$ — последовательность стягивающихся отрезков. Тогда пересечение всех отрезков $[a_n, b_n]$ состоит из одной точки:*

$$\exists c \in \mathbb{R} : \quad \bigcap_{n=1}^{\infty} [a_n, b_n] = \{c\},$$

при этом $a_n \to c$ и $b_n \to c$.

Доказательство. То, что пересечение непусто, следует из аксиомы о вложенных отрезках. Пусть $c, d \in \bigcap\limits_{n=1}^{\infty} [a_n, b_n]$. Докажем, что $c = d$. Поскольку $a_n \leqslant c \leqslant b_n$ и $a_n \leqslant d \leqslant b_n$, имеем

$$a_n - b_n \leqslant c - d \leqslant b_n - a_n.$$

По теореме о предельном переходе в неравенстве $0 \leqslant c - d \leqslant 0$, то есть $c = d$. Так как

$$0 \leqslant c - a_n \leqslant b_n - a_n, \qquad 0 \leqslant b_n - c \leqslant b_n - a_n,$$

по теореме о сжатой последовательности $a_n \to c$ и $b_n \to c$. $\square$

Напомним, что множество $E \subset \mathbb{R}$ называется ограниченным сверху, если существует такое число M, что $x \leqslant M$ для всех $x \in E$. Число M при этом называется верхней границей множества E.

Определение. Пусть $E \subset \mathbb{R}$, $E \neq \varnothing$, E ограничено сверху. Наименьшая из верхних границ множества E называется *точной верхней границей,* или *верхней гранью,* или *супремумом* множества E и обозначается $\sup E$.

Аналогично определяется точная нижняя граница.

Определение. Пусть $E \subset \mathbb{R}$, $E \neq \varnothing$, E ограничено снизу. Наибольшая из нижних границ множества E называется *точной нижней границей,* или *нижней гранью,* или *инфимумом* множества E и обозначается $\inf E$.

Замечание 1. Можно записать определение верхней и нижней грани с помощью неравенств:

$$b = \sup E \iff \begin{cases} \forall x \in E \quad x \leqslant b, \\ \forall \varepsilon > 0 \quad \exists x \in E: \quad x > b - \varepsilon; \end{cases}$$
$$a = \inf E \iff \begin{cases} \forall x \in E \quad x \geqslant a, \\ \forall \varepsilon > 0 \quad \exists x \in E: \quad x < a + \varepsilon. \end{cases}$$

Действительно, в определении супремума первая строчка означает, что b — верхняя граница E, а вторая — что никакое число, меньшее b, не является верхней границей E.

Ясно, что если в множестве E есть наибольший элемент (максимум), то он и будет верхней гранью E. Если же в множестве E нет наибольшего элемента, то существование верхней грани требует доказательства. Аналогично положение дел с нижней гранью.

Теорема 2. Существование верхней и нижней грани. *Всякое непустое ограниченное сверху (снизу) подмножество $\mathbb{R}$ имеет верхнюю (нижнюю) грань.*

Доказательство. Для определенности докажем теорему для ограниченного сверху множества; второе утверждение доказывается аналогично. По условию существует верхняя граница E, обозначим ее через M. Возьмем точку $x_0 \in E$; тогда $x_0 \leqslant M$. Обозначим $[a_1, b_1] = [x_0, M]$. Отрезок $[a_1, b_1]$ удовлетворяет двум условиям:

1) $[a_1, b_1] \cap E \neq \varnothing$,
2) $(b_1, +\infty) \cap E = \varnothing$.

Рассмотрим середину отрезка $[a_1, b_1]$ — точку $\frac{a_1+b_1}{2}$. Положим $[a_2, b_2] = \left[a_1, \frac{a_1+b_1}{2}\right]$, если $\left(\frac{a_1+b_1}{2}, b_1\right] \cap E = \varnothing$, и $[a_2, b_2] = \left[\frac{a_1+b_1}{2}, b_1\right]$, если $\left(\frac{a_1+b_1}{2}, b_1\right] \cap E \neq \varnothing$. В обоих случаях

1) $[a_2, b_2] \cap E \neq \varnothing$,
2) $(b_2, +\infty) \cap E = \varnothing$.

Далее рассмотрим середину отрезка $[a_2, b_2]$, и этот процесс продолжим неограниченно. В результате мы построим последовательность вложенных отрезков $[a_n, b_n]$, для которых:

1) $[a_n, b_n] \cap E \neq \varnothing$,
2) $(b_n, +\infty) \cap E = \varnothing$.

При этом отрезки стягивающиеся, так как $b_n - a_n = \frac{b_1 - a_1}{2^{n-1}} \to 0$. По теореме о стягивающихся отрезках существует единственная точка c, принадлежащая одновременно всем отрезкам $[a_n, b_n]$, причем $a_n \to c$ и $b_n \to c$.

Проверим, что $c = \sup E$. Если $x \in E$, $n \in \mathbb{N}$, то $x \leqslant b_n$ по свойству 2). По теореме о предельном переходе в неравенстве $x \leqslant c$, то есть c — верхняя граница E. Возьмем $\varepsilon > 0$ и докажем, что $c - \varepsilon$ не является верхней границей E. Так как $a_n \to c$, найдется номер N, для которого $a_N > c - \varepsilon$ (по определению предела даже все a_n, начиная с некоторого номера, будут удовлетворять этому неравенству). По свойству 1) найдется точка $x \in [a_N, b_N] \cap E$, а тогда $x > c - \varepsilon$. $\square$

Замечание 2. Если множество E не ограничено сверху (снизу), то полагают $\sup E = +\infty$ ($\inf E = -\infty$). При этом определении супремум и инфимум в $\overline{\mathbb{R}}$ существуют у любого непустого множества. Ограниченность E сверху (снизу) равносильна неравенству $\sup E < +\infty$ ($\inf E > -\infty$).

Для пустого множества любое число служит верхней границей и, таким образом, множество верхних границ не ограничено снизу. Поэтому логично положить $\sup \varnothing = -\infty$. По таким же причинам полагают $\inf \varnothing = +\infty$. Это соглашение приводит к несколько странному неравенству $\sup \varnothing < \inf \varnothing$, поэтому иногда, желая его избежать, грани пустого множества не определяют вовсе.

Замечание 3. *Если* $D \subset E \subset \mathbb{R}$, $D \neq \varnothing$, *то* $\sup D \leqslant \sup E$, *а* $\inf D \geqslant \inf E$.

Доказательство. Докажем неравенство для верхних граней; второе утверждение доказывается аналогично. Если $\sup E = +\infty$, то неравенство тривиально. Пусть $\sup E < +\infty$. Если $x \in D$, то $x \in E$ и, следовательно, $x \leqslant \sup E$, то есть $\sup E$ — какая-то верхняя граница множества D. Но $\sup D$ — наименьшая верхняя граница D, поэтому $\sup D \leqslant \sup E$. □

Для $E, F \subset \mathbb{R}$, $t \in \mathbb{R}$ обозначим

$$E + F = \{x + y : x \in E, y \in F\},$$
$$-E = \{-x : x \in E\}, \qquad tE = \{tx : x \in E\}.$$

Замечание 4. *Если* $E, F \subset \mathbb{R}$, $E, F \neq \varnothing$, $t > 0$, *то*

$$\sup(E + F) = \sup E + \sup F, \qquad \inf(E + F) = \inf E + \inf F,$$
$$\sup(tE) = t \sup E, \qquad \inf(tE) = t \inf E,$$
$$\sup(-E) = -\inf E, \qquad \inf(-E) = -\sup E.$$

Эти равенства читатель легко докажет самостоятельно.

Определение. Если функция f задана по крайней мере на множестве D, то под супремумом (инфимумом) функции f на множестве D понимают супремум (инфимум) образа D:

$$\sup_{x \in D} f(x) = \sup\{f(x) : x \in D\} = \sup f(D),$$
$$\inf_{x \in D} f(x) = \inf\{f(x) : x \in D\} = \inf f(D).$$

В частности, речь может идти о гранях последовательности. Аналогично определяются максимум и минимум (наибольшее и наименьшее значение) функции на множестве. При этом немую переменную иногда опускают и, например, наравне с $\sup\limits_{x \in D} f(x)$ пишут $\sup\limits_{D} f$. Функция f называется *ограниченной* (сверху, снизу) на множестве D, если множество $f(D)$ ограничено (сверху, снизу).

Перейдем к вопросу о пределе монотонной последовательности. Начнем с определения монотонности.

Определение. Пусть $D \subset X \subset \mathbb{R}$. Функция $f\colon X \to \mathbb{R}$ называется:

возрастающей на множестве D, если для любых x_1, x_2 из D, таких что $x_1 < x_2$, выполняется неравенство $f(x_1) \leqslant f(x_2)$;

строго возрастающей на множестве D, если для любых x_1, x_2 из D, таких что $x_1 < x_2$, выполняется неравенство $f(x_1) < f(x_2)$;

убывающей на множестве D, если для любых x_1, x_2 из D, таких что $x_1 < x_2$, выполняется неравенство $f(x_1) \geqslant f(x_2)$;

строго убывающей на множестве D, если для любых x_1, x_2 из D, таких что $x_1 < x_2$, выполняется неравенство $f(x_1) > f(x_2)$.
Возрастающие и убывающие функции называются *монотонными*, а строго убывающие и строго возрастающие — *строго монотонными.*

Иногда функции, которые только что были названы возрастающими (убывающими), называют неубывающими (невозрастающими), а те, что были названы строго возрастающими (строго убывающими), называют возрастающими (убывающими). Какой терминологией пользоваться — дело вкуса; мы в курсе будем придерживаться первоначально сформулированных определений.

Сформулируем отдельно частный случай этого определения для последовательностей.

Определение. Вещественная последовательность $\{x_n\}_{n=1}^{\infty}$ называется:

возрастающей, если $x_n \leqslant x_{n+1}$ для всех $n \in \mathbb{N}$;

строго возрастающей, если $x_n < x_{n+1}$ для всех $n \in \mathbb{N}$;

убывающей, если $x_n \geqslant x_{n+1}$ для всех $n \in \mathbb{N}$;

строго убывающей, если $x_n > x_{n+1}$ для всех $n \in \mathbb{N}$.

Легко видеть с помощью индукции, что для последовательностей эти определения равносильны предыдущим. Например, если $x_n \leqslant x_{n+1}$ для всех n, то $x_n \leqslant x_m$ для всех n, m: $n < m$.

Когда говорят, что функция ограничена (возрастает, положительна и т.п.) без указания множества, то имеют в виду, что функция обладает указанным свойством на области определения.

Теорема 3. Предел монотонной последовательности.

1. *Всякая возрастающая ограниченная сверху последовательность сходится.*

2. *Всякая убывающая ограниченная снизу последовательность сходится.*

3. *Всякая монотонная ограниченная последовательность сходится.*

Доказательство. Докажем первое утверждение. Пусть последовательность $\{x_n\}$ ограничена сверху. По теореме 2 существует $\sup\limits_{n\in\mathbb{N}} x_n = c \in \mathbb{R}$. Докажем, что $c = \lim x_n$. Возьмем $\varepsilon > 0$. По определению супремума найдется такой номер N, что $x_N > c - \varepsilon$. В силу возрастания последовательности при любом $n > N$ будет $x_n \geqslant x_N$. Снова по определению супремума $x_n \leqslant c$ при всех n. Итак, для любого $n > N$

$$c - \varepsilon < x_N \leqslant x_n \leqslant c < c + \varepsilon.$$

В силу произвольности ε это значит, что $c = \lim x_n$.

Второе утверждение доказывается аналогично, третье следует из первых двух. $\square$

Замечание 1. *Если последовательность возрастает и не ограничена сверху, то она стремится к плюс бесконечности. Если последовательность убывает и не ограничена снизу, то она стремится к минус бесконечности.*

Доказательство проведем только для возрастающей последовательности $\{x_n\}$. Возьмем $E > 0$. Так как $\{x_n\}$ не ограничена сверху, найдется такой номер N, что $x_N > E$. Тогда для любого номера $n > N$ в силу возрастания последовательности тем более $x_n > E$. $\square$

Замечание 2. Доказано, что *любая монотонная последовательность имеет предел в $\overline{\mathbb{R}}$,* конечный или бесконечный. При этом *для всякой возрастающей последовательности*

$$\lim x_n = \sup x_n,$$

а для всякой убывающей последовательности

$$\lim x_n = \inf x_n.$$

Лемма 1. Неравенство Я. Бернулли. *Если* $n \in \mathbb{Z}_+$, $x > -1$, *то*

$$(1+x)^n \geqslant 1 + nx.$$

Доказательство. При $n = 0$ и $n = 1$ (база индукции) неравенство, очевидно, обращается в верное равенство. Сделаем индукционный переход: пусть неравенство верно для номера n. Тогда

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n(1+x) \geqslant (1+nx)(1+x) = \\ &= 1 + (n+1)x + nx^2 \geqslant 1 + (n+1)x. \qquad \square\end{aligned}$$

Читателю предлагается доказать, что при $n \geqslant 2$ и $x \neq 0$ неравенство Я. Бернулли строгое.

Пример 1. Пусть $q \in \mathbb{R}$, $|q| < 1$. Докажем, что $\lim\limits_{n\to\infty} q^n = 0$. В самом деле, последовательность $\left\{|q|^n\right\}$ убывает и ограничена снизу нулем. Следовательно, существует конечный $\lim\limits_{n\to\infty} |q|^n$; обозначим его через a. Перейдя в равенстве $|q|^{n+1} = |q||q|^n$ к пределу, получим $a = |q|a$ или $(1-|q|)a = 0$. Отсюда $a = 0$, поскольку $|q| < 1$. Таким образом, $|q|^n \to 0$, что равносильно $q^n \to 0$.

Если $|q| > 1$, то $\left|\frac{1}{q}\right| < 1$. По доказанному $\lim\limits_{n\to\infty} \frac{1}{q^n} = 0$, а тогда $\lim\limits_{n\to\infty} q^n = \infty$ по лемме 2 § 1. Если $q = 1$, то очевидно $\lim\limits_{n\to\infty} q^n = 1$. При $q = -1$ было доказано, что последовательность $\{q^n\}$ не имеет предела.

Пример 2. Число e. Докажем сходимость последовательности $x_n = \left(1 + \frac{1}{n}\right)^n$.

Положим $y_n = \left(1 + \frac{1}{n}\right)^{n+1}$. Ясно, что последовательность $\{y_n\}$ ограничена снизу единицей. Кроме того, $\{y_n\}$ убывает:

$$\frac{y_{n-1}}{y_n} = \frac{(1+\frac{1}{n-1})^n}{(1+\frac{1}{n})^{n+1}} = \frac{(\frac{n}{n-1})^n}{(\frac{n+1}{n})^{n+1}} = \left(\frac{n^2}{n^2-1}\right)^{n+1} \frac{n-1}{n} =$$

$$= \left(1 + \frac{1}{n^2-1}\right)^{n+1} \frac{n-1}{n} \geqslant \left(1 + \frac{n+1}{n^2-1}\right) \frac{n-1}{n} = 1$$

(мы воспользовались неравенством Бернулли). Следовательно, $\{y_n\}$ сходится, а тогда по теореме о пределе частного и $x_n = \dfrac{y_n}{1+1/n}$ сходится к тому же пределу.

Определение. Предел последовательности $\left\{\left(1+\frac{1}{n}\right)^n\right\}$ называют *числом Непера* или *основанием натуральных логарифмов* и обозначают буквой e:

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

Из формулы бинома Ньютона следует, что при всех $n \geqslant 2$

$$x_n \geqslant 1 + n\cdot\frac{1}{n} + \frac{n(n-1)}{2}\cdot\frac{1}{n^2} = \frac{5}{2} - \frac{1}{2n} \geqslant \frac{9}{4} > 2.$$

Кроме того, $y_n \leqslant y_5 < 3$ при всех $n \geqslant 5$. Поэтому $2 < e < 3$.

В § 6 главы 3 мы докажем, что число e иррационально (более того, оно трансцендентно, то есть не является корнем никакого многочлена с целыми коэффициентами), и укажем способ его вычисления с любой точностью. Пока отметим без доказательства, что

$$e = 2{,}718281828459045\ldots$$

С помощью теоремы о пределе монотонной последовательности удобно доказывать существование пределов рекуррентно заданных последовательностей вида $x_{n+1} = F(x_n)$. Проиллюстрируем это следующим примером.

Пример 3. Формула Герона. Пусть $a > 0$, $x_0 > 0$,

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \qquad n \in \mathbb{Z}_+. \tag{2}$$

Ясно, что $x_n > 0$ при всех n и, значит, $\{x_n\}$ ограничена снизу. Воспользовавшись очевидным неравенством

$$t + \frac{1}{t} \geqslant 2, \qquad t > 0,$$

мы получим, что при всех $n \in \mathbb{Z}_+$

$$x_{n+1} = \frac{\sqrt{a}}{2}\left(\frac{x_n}{\sqrt{a}} + \frac{\sqrt{a}}{x_n}\right) \geqslant \frac{\sqrt{a}}{2} \cdot 2 = \sqrt{a}.$$

Поэтому при всех $n \in \mathbb{N}$

$$x_{n+1} = \frac{x_n}{2}\left(1 + \frac{a}{x_n^2}\right) \leqslant x_n,$$

то есть последовательность $\{x_n\}$ убывает. Следовательно, она сходится; обозначим $\beta = \lim x_n$. Перейдя к пределу в равенстве (2), получим

$$\beta = \frac{1}{2}\left(\beta + \frac{a}{\beta}\right),$$

откуда $\beta = \sqrt{a}$, так как $\beta \geqslant 0$.

Равенство $\lim x_n = \sqrt{a}$ называется формулой Герона и используется для приближенного вычисления значений корня. Этот способ является частным случаем метода касательных Ньютона приближенного нахождения корней уравнений.

Замечание 3. *Пусть* $x_n > 0$ *при всех* n, $\lim \frac{x_{n+1}}{x_n} < 1$. *Тогда* $x_n \to 0$.

Отсюда следует, что

$$\lim_{n\to\infty} \frac{n^k}{a^n} = 0, \qquad a > 1,\ k \in \mathbb{N},$$
$$\lim_{n\to\infty} \frac{a^n}{n!} = 0, \qquad a \in \mathbb{R},$$
$$\lim_{n\to\infty} \frac{n!}{n^n} = 0.$$

Доказательство замечания 3 и его следствий остается читателю в качестве упражнения.

Определение. Пусть $\{x_n\}_{n=1}^{\infty}$ — последовательность в множестве X, $\{n_k\}_{k=1}^{\infty}$ — строго возрастающая последовательность натуральных чисел (то есть $n_k < n_{k+1}$ для всех $k \in \mathbb{N}$). Их композиция ($\{x_n\}$ — внешнее отображение, $\{n_k\}$ — внутреннее) называется *подпоследовательностью* последовательности $\{x_n\}$ и обозначается $\{x_{n_k}\}$.

Условие строгого возрастания последовательности индексов $\{n_k\}$ выражает тот факт, что члены подпоследовательности расположены "в том же порядке", что и члены исходной последовательности. Так, последовательности $\{2, 4, 6, 8, \dots\}$ и $\{1, 4, 9, 16, \dots\}$ — подпоследовательности последовательности $\{1, 2, 3, 4, \dots\}$ (соответственно, $n_k = 2k$ и $n_k = k^2$), а последовательности $\{1, 1, 2, 3, \dots\}$ и $\{2, 1, 3, 4, \dots\}$ — нет.

Замечание 1. Для строго возрастающей последовательности $\{n_k\}$ при всех k будет $n_k \geqslant k$. Действительно, $n_1 \geqslant 1$ (база индукции), и из неравенства $n_k \geqslant k$ следует, что $n_{k+1} \geqslant n_k + 1 \geqslant k + 1$ (индукционный переход).

Лемма 2. *Всякая подпоследовательность последовательности, имеющей предел, стремится к тому же пределу: если $\{x_n\}$ — вещественная последовательность, $\{x_{n_k}\}$ — ее подпоследовательность, $a \in \overline{\mathbb{R}} \cup \{\infty\}$, $x_n \to a$, то $x_{n_k} \to a$.*

Доказательство. Пусть $a \in \mathbb{R}$. Возьмем $\varepsilon > 0$. По определению предела существует такой номер N, что $|x_n - a| < \varepsilon$ для всех $n > N$. Но тогда, если $k > N$, то по замечанию 1 и $n_k > N$, а значит, $|x_{n_k} - a| < \varepsilon$.

В случае бесконечного предела a в доказательстве следует заменить неравенства вида $|x_n - a| < \varepsilon$ на $|x_n| > E$ и т.п. или воспользоваться языком окрестностей. $\Box$

Лемма 3. *Пусть $\{x_n\}$ — вещественная последовательность, $\{x_{n_k}\}$ и $\{x_{m_l}\}$ — ее подпоследовательности, причем объединение множеств их индексов равно $\mathbb{N}$, $a \in \overline{\mathbb{R}} \cup \{\infty\}$. Тогда, если $x_{n_k} \to a$, $x_{m_l} \to a$, то и $x_n \to a$.*

Доказательство. Для определенности ограничимся случаем, когда $a \in \mathbb{R}$. Возьмем $\varepsilon > 0$. По определению предела одной и другой подпоследовательности найдутся такие номера K и L, что

$$|x_{n_k} - a| < \varepsilon \quad \text{для всех } k > K,$$
$$|x_{m_l} - a| < \varepsilon \quad \text{для всех } l > L.$$

Положим $N = \max\{n_K, m_L\}$. Если $n > N$, то либо $n = n_k$ при $k > K$, либо $n = m_l$ при $l > L$. В любом случае $|x_n - a| < \varepsilon$. $\square$

Теорема 4. Принцип выбора Больцано – Вейерштрасса. *Из всякой ограниченной последовательности можно извлечь сходящуюся подпоследовательность.*

Доказательство. Поскольку последовательность $\{x_n\}$ ограничена, все ее члены принадлежат некоторому отрезку $[a, b]$. Обозначим через $[a_1, b_1]$ ту половину отрезка $[a, b]$, которая содержит бесконечно много членов последовательности $\{x_n\}$ (это означает, что $x_n \in [a_1, b_1]$ для бесконечного множества индексов n); если обе половины содержат бесконечно много членов последовательности, то можно взять любую половину. Выберем такой номер n_1, что $x_{n_1} \in [a_1, b_1]$. Далее, обозначим через $[a_2, b_2]$ ту половину отрезка $[a_1, b_1]$, которая содержит бесконечно много членов последовательности $\{x_n\}$, и выберем такой номер $n_2 > n_1$, что $x_{n_2} \in [a_2, b_2]$. Этот процесс продолжим неограниченно: на шаге с номером k обозначаем через $[a_k, b_k]$ ту половину отрезка $[a_{k-1}, b_{k-1}]$, которая содержит бесконечно много членов последовательности $\{x_n\}$, и выбираем такой номер $n_k > n_{k-1}$, что $x_{n_k} \in [a_k, b_k]$. Так мы построим последовательность стягивающихся отрезков $\{[a_k, b_k]\}$ и подпоследовательность $\{x_{n_k}\}$ исходной последовательности. По теореме о стягивающихся отрезках существует такая точка c, что $a_k \to c$ и $b_k \to c$. Но так как $a_k \leqslant x_{n_k} \leqslant b_k$, по теореме о сжатой последовательности и $x_{n_k} \to c$. $\square$

Следующее замечание дополняет принцип выбора для неограниченных последовательностей.

Замечание 2. *Если вещественная последовательность не ограничена сверху (снизу), то из нее можно извлечь подпоследовательность, стремящуюся к плюс бесконечности (минус бесконечности).*

Доказательство. Для определенности докажем утверждение в случае $+\infty$. Так как последовательность $\{x_n\}$ не ограничена сверху, найдется номер n_1, для которого $x_{n_1} > 1$. Далее, найдется номер $n_2 > n_1$, для которого $x_{n_2} > 2$ (иначе $\{x_n\}$ была бы ограничена сверху числом $\max\{x_1, \ldots, x_{n_1}, 2\}$). Этот процесс продолжаем неограниченно: на шаге с номером k найдется номер $n_k > n_{k-1}$, для которого $x_{n_k} > k$. По замечанию 7 к определению бесконечного предела $x_{n_k} \to +\infty$. $\square$

Определение. Точка a называется *частичным пределом* последовательности $\{x_n\}$, если существует подпоследовательность $\{x_{n_k}\}$, стремящаяся к a.

Говоря о вещественных последовательностях, мы имеем в виду их частичные пределы в $\overline{\mathbb{R}}$. Ясно, что предел последовательности является ее частичным пределом. Числа 1 и -1 — частичные пределы последовательности $x_n = (-1)^n$, так как $x_{2k} \to 1$, $x_{2k+1} \to -1$.

Определение. Пусть вещественная последовательность $\{x_n\}$ ограничена сверху. Величина

$$\overline{\lim_{n\to\infty}}\, x_n = \lim_{n\to\infty} \sup_{k\geqslant n} x_k$$

называется *верхним пределом* последовательности $\{x_n\}$.

Пусть вещественная последовательность $\{x_n\}$ ограничена снизу. Величина

$$\underline{\lim_{n\to\infty}}\, x_n = \lim_{n\to\infty} \inf_{k\geqslant n} x_k$$

называется *нижним пределом* последовательности $\{x_n\}$.

Замечание 1. Последовательности $z_n = \sup\limits_{k\geqslant n} x_k$ и $y_n = \inf\limits_{k\geqslant n} x_k$ иногда называют *верхней и нижней огибающими* последовательности $\{x_n\}$.

Если $\{x_n\}$ не ограничена сверху, то $z_n = +\infty$ при всех n, и поэтому полагают $\overline{\lim}\, x_n = +\infty$. По таким же причинам, если $\{x_n\}$ не ограничена снизу, полагают $\underline{\lim}\, x_n = -\infty$.

Замечание 2. *Верхний и нижний пределы вещественной последовательности $\{x_n\}$ существуют в $\overline{\mathbb{R}}$, причем $\underline{\lim}\, x_n \leqslant \overline{\lim}\, x_n$.*

Доказательство. Пусть $\{x_n\}$ ограничена и сверху, и снизу. При переходе к подмножеству супремум не увеличивается, а инфимум не уменьшается, поэтому $\{y_n\}$ возрастает, а $\{z_n\}$ убывает. При всех n

$$y_1 \leqslant y_n \leqslant z_n \leqslant z_1.$$

По теореме о пределе монотонной последовательности $\{x_n\}$ и $\{y_n\}$ сходятся, то есть существуют конечные пределы $\lim y_n = \underline{\lim}\, x_n$

и $\lim z_n = \overline{\lim}\, x_n$. По теореме о предельном переходе в неравенстве $\underline{\lim}\, x_n \leqslant \overline{\lim}\, x_n$.

Если $\{x_n\}$ ограничена сверху, но не снизу, то $\underline{\lim}\, x_n = -\infty$, а последовательность $\{z_n\}$ убывает. Поэтому существует

$$\overline{\lim}\, x_n = \lim z_n \in [-\infty, +\infty).$$

Если $\{x_n\}$ ограничена снизу, но не сверху, то $\overline{\lim}\, x_n = +\infty$, а последовательность $\{y_n\}$ возрастает, поэтому существует

$$\underline{\lim}\, x_n = \lim y_n \in (-\infty, +\infty].$$

Если $\{x_n\}$ не ограничена ни сверху, ни снизу, то по определению $\underline{\lim}\, x_n = -\infty$, $\overline{\lim}\, x_n = +\infty$. $\square$

Теорема 5. О верхнем и нижнем пределе последовательности. *Пусть* $\{x_n\}$ *— вещественная последовательность. Тогда справедливы следующие утверждения.*

1. *Верхний предел — наибольший, а нижний предел — наименьший из частичных пределов* $\{x_n\}$.

2. *Предел* $\{x_n\}$ *в* $\overline{\mathbb{R}}$ *существует тогда и только тогда, когда* $\underline{\lim}\, x_n = \overline{\lim}\, x_n$. *При этом* $\lim x_n$ *равен их общему значению.*

Доказательство. I. Пусть $\{x_n\}$ ограничена и сверху и снизу.

1. Обозначим $b = \overline{\lim}\, x_n$. Докажем, что b — частичный предел $\{x_n\}$, для чего построим подпоследовательность последовательности $\{x_n\}$, стремящуюся к b. При всех n будет $z_n \geqslant b$. По любым $\varepsilon > 0$ и $N \in \mathbb{N}$ найдется номер $n > N$, для которого $x_n > b - \varepsilon$, поскольку

$$z_{N+1} = \sup\{x_{N+1}, x_{N+2}, \dots\} > b - \varepsilon.$$

Теперь подберем такой номер n_1, что $x_{n_1} > b - 1$, затем такой номер $n_2 > n_1$, что $x_{n_2} > b - \frac{1}{2}$, и этот процесс продолжим неограниченно: на k-м шаге, когда номер n_{k-1} уже выбран, найдется такой номер $n_k > n_{k-1}$, что $x_{n_k} > b - \frac{1}{k}$. Таким образом, оказывается построена подпоследовательность $\{x_{n_k}\}$, члены которой удовлетворяют неравенству

$$b - \frac{1}{k} < x_{n_k} \leqslant z_{n_k}.$$

Подпоследовательность $\{z_{n_k}\}$ последовательности $\{z_n\}$, стремящейся к b, тоже стремится к b. По теореме о сжатой последовательности и $x_{n_k} \to b$.

Если $\{x_{m_l}\}$ — подпоследовательность $\{x_n\}$, $x_{m_l} \to \beta$, то, сделав предельный переход в неравенстве $x_{m_l} \leqslant z_{m_l}$, мы получим, что $\beta \leqslant b$, то есть b — наибольший частичный предел $\{x_n\}$.

Аналогично доказывается, что $\underline{\lim}\, x_n$ — наименьший частичный предел $\{x_n\}$.

2. По определению y_n и z_n при всех n

$$y_n \leqslant x_n \leqslant z_n.$$

Если $\underline{\lim}\, x_n = \overline{\lim}\, x_n$, то по теореме о сжатой последовательности существует $\lim x_n$ и $\lim x_n = \underline{\lim}\, x_n = \overline{\lim}\, x_n$.

Если же существует $\lim x_n$, то все частичные пределы и, в частности, верхний и нижний предел, с ним совпадают.

II. Пусть $\{x_n\}$ ограничена сверху, но не снизу. Тогда по определению $\underline{\lim}\, x_n = -\infty$. По замечанию 2 к принципу выбора из $\{x_n\}$ можно выбрать подпоследовательность, стремящуюся к $-\infty$, то есть $-\infty$ — частичный предел $\{x_n\}$. Разумеется, $-\infty$ меньше любого другого частичного предела, если они есть. То, что $\overline{\lim}\, x_n$ — наибольший частичный предел $\{x_n\}$, доказывается, как в части I. Если $\underline{\lim}\, x_n = \overline{\lim}\, x_n$, то есть $z_n \to -\infty$, то и $x_n \to -\infty$, так как $x_n \leqslant z_n$. Обратно, если $\lim x_n = -\infty$, то и $\overline{\lim}\, x_n = -\infty$ как частичный предел.

Случай, когда $\{x_n\}$ ограничена снизу, но не сверху, разбирается аналогично.

III. Если $\{x_n\}$ не ограничена ни сверху, ни снизу, то первое утверждение теоремы очевидно, а второе не реализуется. $\square$

Замечание 3. Принцип выбора Больцано – Вейерштрасса вытекает из теоремы 5: из любой ограниченной последовательности $\{x_n\}$ можно извлечь подпоследовательность, сходящуюся к $\overline{\lim}\, x_n$ (и подпоследовательность, сходящуюся к $\underline{\lim}\, x_n$). Тем не менее, мы предпочли отдельное доказательство принципа выбора, которое проще доказательства теоремы 5.

Замечание 4. Читателю предлагается доказать следующую характеристику конечного верхнего и нижнего предела с помощью неравенств:

$$b = \varlimsup x_n \in \mathbb{R} \iff \begin{cases} \forall \varepsilon > 0 \ \ \exists N \ \ \forall n > N \ \ x_n < b + \varepsilon, \\ \forall \varepsilon > 0 \ \ \forall N \ \ \exists n > N \ \ x_n > b - \varepsilon; \end{cases}$$
$$a = \varliminf x_n \in \mathbb{R} \iff \begin{cases} \forall \varepsilon > 0 \ \ \exists N \ \ \forall n > N \ \ x_n > a - \varepsilon, \\ \forall \varepsilon > 0 \ \ \forall N \ \ \exists n > N \ \ x_n < a + \varepsilon. \end{cases}$$

Определение. Говорят, что последовательность $\{x_n\}$ *сходится в себе*, если для любого положительного числа ε существует такой номер N, что для всех номеров n и l, больших N, выполняется неравенство $|x_n - x_l| < \varepsilon$:

$$\forall \varepsilon > 0 \ \ \exists N \ \ \forall n, l > N \ \ |x_n - x_l| < \varepsilon.$$

Сходящуюся в себе последовательность также называют *фундаментальной последовательностью* или *последовательностью Коши.*

Как видно, в определении сходимости в себе требуется, чтобы члены последовательности с достаточно большими номерами были близки не к точке a, а друг к другу.

Лемма 4.

1. *Сходящаяся в себе последовательность ограничена.*

2. *Если у сходящейся в себе последовательности есть сходящаяся подпоследовательность, то и сама последовательность сходится.*

Доказательство. 1. Пользуясь сходимостью $\{x_n\}$ в себе, подберем такой номер N, что для всех $n, l > N$ будет $|x_n - x_l| < 1$. В частности, тогда $|x_n - x_{N+1}| < 1$ для всех $n > N$. Следовательно, при всех $n > N$

$$|x_n| \leqslant |x_n - x_{N+1}| + |x_{N+1}| < 1 + |x_{N+1}|.$$

Положим
$$R = \max\{|x_1|, \ldots, |x_N|, 1 + |x_{N+1}|\};$$
тогда $|x_n| \leqslant R$ для всех номеров n.

2. Пусть $\{x_n\}$ сходится в себе, $x_{n_k} \to a$. Возьмем $\varepsilon > 0$. По определению предела найдется такой номер K, что $|x_{n_k} - a| < \frac{\varepsilon}{2}$ для

всех $k > K$, а по определению сходимости в себе найдется такой номер N, что $|x_n - x_l| < \frac{\varepsilon}{2}$ для всех $n, l > N$. Покажем, что найденное N — требуемое для ε из определения предела. Пусть $n > N$. Положим $M = \max\{N+1, K+1\}$; тогда $n_M \geqslant n_{N+1} > n_N \geqslant N$ и, аналогично, $n_M > K$. Следовательно,

$$|x_n - a| \leqslant |x_n - x_{n_M}| + |x_{n_M} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

В силу произвольности ε это и означает, что $x_n \to a$. □

Теорема 6. Критерий Больцано – Коши сходимости последовательности. *Сходимость вещественной последовательности равносильна ее сходимости в себе.*

Доказательство. 1. Пусть $\lim x_n = a$. Возьмем $\varepsilon > 0$. По определению предела найдется такой номер N, что $|x_n - a| < \frac{\varepsilon}{2}$ для всех $n > N$. Тогда для всех $n, m > N$

$$|x_n - x_m| \leqslant |x_n - a| + |a - x_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

В силу произвольности ε это и значит, что $\{x_n\}$ сходится в себе.

2. Пусть последовательность $\{x_n\}$ сходится в себе. По пункту 1 леммы 4 она ограничена. По принципу выбора Больцано – Вейерштрасса из нее можно извлечь сходящуюся подпоследовательность, а тогда по пункту 2 леммы 4 она сама сходится. □

Критерий Больцано – Коши бывает удобен тем, что позволяет доказывать существование предела последовательности, не используя само значение предела.

§ 3. Предел функции

Напомним, что ε-окрестностью точки $a \in \mathbb{R}$ называется интервал $(a - \varepsilon, a + \varepsilon)$.

Определение. Пусть $a \in \mathbb{R}$, $D \subset \mathbb{R}$. Точка a называется *предельной точкой* или *точкой сгущения* множества D, если в любой окрестности точки a найдется точка множества D, отличная от a.

Это определение удобно переформулировать с помощью понятия проколотой окрестности точки.

Определение. *Проколотой окрестностью* точки a называется множество
$$\dot{V}_a = V_a \setminus \{a\}.$$

Таким образом, точка a называется предельной точкой множества D, если любая проколотая окрестность точки a имеет с D непустое пересечение.

Замечание 1. Предельная точка множества может принадлежать, а может не принадлежать множеству. Так, каждая точка отрезка $[c, d]$ является предельной для интервала (c, d).

Замечание 2. Если a — предельная точка D, то в любой окрестности точки a (проколотой или нет — неважно) найдется бесконечно много точек множества D.

Доказательство. Пусть в некоторой окрестности V_a точки a лишь конечное число точек D. Перенумеруем их, исключив, быть может, саму точку a:
$$\dot{V}_a \cap D = \{x_1, \ldots, x_N\}.$$
Обозначим
$$r = \min\{|x_1 - a|, \ldots, |x_N - a|\}.$$
Тогда $r > 0$ и $\dot{V}_a(r) \cap D = \varnothing$, что противоречит определению предельной точки. $\square$

Следующее замечание поясняет название “предельная точка”.

Замечание 3. Точка a является предельной для множества D тогда и только тогда, когда существует последовательность $\{x_n\}$ точек множества D, отличных от a, стремящаяся к a.

Доказательство. Пусть a — предельная точка D. Тогда для любого $n \in \mathbb{N}$ в проколотой $\frac{1}{n}$-окрестности a найдется точка x_n множества D. Построенная последовательность $\{x_n\}$ стремится к a, так как $|x_n - a| < \frac{1}{n} \to 0$.

Обратно, если существует последовательность $\{x_n\}$ с перечисленными свойствами, то по определению предела в любой окрестности точки a найдется член этой последовательности (туда попадут даже все члены, начиная с некоторого номера), то есть точка множества D, отличная от a. $\square$

Определение. Если точка a принадлежит множеству D, но не является его предельной точкой, то a называется *изолированной точкой* множества D.

Замечание 4. Определение предельной точки применимо и в случаях $a = \pm\infty$ и $a = \infty$. Напомним, что

$$\dot{V}_{+\infty} = (E, +\infty),\ \dot{V}_{-\infty} = (-\infty, -E),\ \dot{V}_{\infty} = (-\infty, -E) \cup (E, +\infty).$$

Замечания 2 и 3 остаются в силе, требуются лишь небольшие изменения в доказательствах. Например, если предположить, что в некоторой окрестности $+\infty$ содержится лишь конечное число точек $x_1, \dots x_N$ множества $D \subset \mathbb{R}$, то в проколотой окрестности $(R, +\infty)$, где $R = \max\limits_{1 \leqslant i \leqslant N} x_i$, нет точек D. Легко также видеть, что неограниченность D (соответственно, неограниченность сверху, снизу) равносильна тому, что ∞ (соответственно, $+\infty$, $-\infty$) — предельная точка D.

Определение. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $a \in \mathbb{R}$ — предельная точка D, $A \in \mathbb{R}$. Точку A называют *пределом* функции f в точке a и пишут

$$\lim_{x \to a} f(x) = A \quad \text{или} \quad f(x) \underset{x \to a}{\longrightarrow} A,$$

если выполняется одно из следующих утверждений.

1. Определение на ε-языке, или по Коши.

Для любого положительного числа ε существует такое положительное число δ, что для всех точек x множества D, отличных от a и удовлетворяющих неравенству $|x - a| < \delta$, выполняется неравенство $|f(x) - A| < \varepsilon$:

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in D \setminus \{a\} : |x - a| < \delta \ \ |f(x) - A| < \varepsilon.$$

2. Определение на языке окрестностей.

Для любой окрестности V_A точки A существует такая окрестность V_a точки a, что образ пересечения проколотой окрестности $\dot{V}_a$ с множеством D при отображении f содержится в окрестности V_A:

$$\forall V_A \ \ \exists V_a \ \ f\big(\dot{V}_a \cap D\big) \subset V_A.$$

3. Определение на языке последовательностей, или по Гейне.

Для любой последовательности $\{x_n\}$ точек множества D, отличных от a, стремящейся к a, последовательность $\{f(x_n)\}$ стремится к A:

$$\forall\{x_n\}:\ x_n \in D\setminus\{a\}, x_n \to a\quad f(x_n)\to A.$$

Говорят также: “$f(x)$ стремится к A при x, стремящемся к a”. Переменная x в обозначении предела “немая”, поэтому ее можно заменить любой другой буквой или опустить и писать $\lim\limits_a f$.

Равносильность определений предела на ε-языке и на языке окрестностей очевидна: во втором записано то же, что и в первом, только с помощью других обозначений. Определение на языке последовательностей сводит новое понятие предела функции к уже определенному понятию предела последовательности. Равносильность определений предела на ε-языке и на языке последовательностей требует отдельного доказательства и составит утверждение теоремы 1.

Замечание 1. Из условия, что a — предельная точка D, следует, что $\dot V_a\cap D\neq\varnothing$ для любой окрестности точки a и что существует последовательность $\{x_n\}$ со свойствами: $x_n\in D$, $x_n\neq a$, $x_n\to a$.

Замечание 2. Если две функции $f,g\colon D\to\mathbb{R}$ совпадают на множестве $\dot U_a\cap D$, где U_a — какая-нибудь окрестность точки a, и $f(x)\underset{x\to a}{\longrightarrow}A$, то и $g(x)\underset{x\to a}{\longrightarrow}A$. Действительно, если окрестность V_a подходит для V_A в определении предела на языке окрестностей для f, то окрестность $V_a\cap U_a$ подходит для V_A в определении предела для g.

Таким образом, существование предела функции в точке a — *локальное* свойство: оно зависит от поведения функции в любой наперед заданной окрестности точки a.

Замечание 3. Если существует такая окрестность U_a точки a, что $\dot U_a\subset D$, то условие $x\in D$ в определении на ε-языке и условие $x_n\in D$ в определении на языке последовательностей можно опустить, а в определении на языке окрестностей писать $f(\dot V_a)$ вместо $f(\dot V_a\cap D)$.

В самом деле, если $V_a\subset U_a$, то $\dot V_a\cap D=\dot V_a$, а если $x_n\in X$, $x_n\neq a$, $x_n\to a$, то, начиная с некоторого номера, $x_n\in\dot U_a$ и, значит, $x_n\in D$.

Замечание 4. Значение функции f в точке a не участвует в определении предела в точке a. Поэтому для существования и значения предела в точке a несущественно, задана ли функция f в точке a и, если задана, то как именно.

Замечание 5. Определение предела легко обобщается на случаи, когда a и (или) A равны $+\infty$, $-\infty$ или ∞. Требование, что a — предельная точка D, сохраняется и тогда, когда $a = \pm\infty$ или $a = \infty$.

Определение на языке окрестностей вообще не меняется. Определение на языке последовательностей не меняется, за исключением того, что незачем писать $x \neq (\pm)\infty$. Определение на ε-языке записывается в этих случаях по-разному. Приведем три примера, воспользовавшись логической символикой.

1. Говорят, что $\lim\limits_{x\to+\infty} f(x) = \infty$, если

$$\forall E > 0 \ \ \exists \Delta > 0 \ \ \forall x \in D : x > \Delta \ \ |f(x)| > E.$$

2. Говорят, что $\lim\limits_{x\to a} f(x) = -\infty$ $(a \in \mathbb{R})$, если

$$\forall E > 0 \ \ \exists \delta > 0 \ \ \forall x \in D \setminus \{a\} : |x - a| < \delta \ \ f(x) < -E.$$

3. Говорят, что $\lim\limits_{x\to\infty} f(x) = A$ $(A \in \mathbb{R})$, если

$$\forall \varepsilon > 0 \ \ \exists \Delta > 0 \ \ \forall x \in D : |x| > \Delta \ \ |f(x) - A| < \varepsilon.$$

Читателю предлагается самому сформулировать определения в остальных двенадцати случаях.

Замечание 6. Если для любой последовательности $\{x_n\}$ точек со свойствами из определения Гейне ($x_n \in D$, $x_n \neq a$, $x_n \to a$) последовательность $\{f(x_n)\}$ имеет предел, то все эти пределы равны между собой и, таким образом, функция имеет предел в точке a.

Действительно, пусть $x_n, y_n \in D \setminus \{a\}$ для всех $n \in \mathbb{N}$, $x_n \to a$, $y_n \to a$, $f(x_n) \to A$, $f(y_n) \to B$. Построим новую последовательность $\{z_n\}$: $z_{2n-1} = x_n$, $z_{2n} = y_n$. Тогда $z_n \in D \setminus \{a\}$, $z_n \to a$. Поэтому $f(z_n)$ стремится к некоторому пределу. Отсюда $A = B$ как пределы подпоследовательностей последовательности, имеющей предел.

Таким образом, при доказательстве существования предела функции в точке на языке последовательностей достаточно проверить лишь существование всех пределов $\{f(x_n)\}$, где $x_n \in D$, $x_n \neq a$, $x_n \to a$, а их равенство можно не проверять.

Замечание 7. Понятие предела последовательности, обсуждавшееся в § 1 и 2, — частный случай понятия предела функции. В самом деле, для последовательности множество D есть $\mathbb{N}$, а $+\infty$ — предельная точка $\mathbb{N}$ (в записи $n \to \infty$ по традиции опускают знак плюс перед бесконечностью).

Типичный случай, когда ставится вопрос о пределе функции: функция задана по крайней мере в проколотой окрестности точки a или на промежутке вида $\langle c, a)$ или $(a, b\rangle$.

Докажем теперь равносильность данных в начале параграфа определений предела.

Теорема 1. *Определения предела функции по Коши и по Гейне равносильны.*

Доказательство. Для определенности докажем теорему при $a, A \in \mathbb{R}$, оставив другие случаи читателю в качестве упражнения.

1. Пусть A — предел функции f в точке a по Коши; докажем, что тогда A — предел f и по Гейне. Возьмем последовательность $\{x_n\}$ со свойствами из определения Гейне: $x_n \in D$, $x_n \neq a$, $x_n \to a$. Требуется доказать, что $f(x_n) \to A$. Возьмем $\varepsilon > 0$. По определению Коши подберем такое $\delta > 0$, что для всех $x \in D$, для которых $x \neq a$ и $|x - a| < \delta$, будет $|f(x) - A| < \varepsilon$. По определению предела последовательности $\{x_n\}$ для числа δ найдется такой номер N, что при всех $n > N$ верно неравенство $|x_n - a| < \delta$. Но тогда $|f(x_n) - A| < \varepsilon$ для всех $n > N$. В силу произвольности ε это значит, что $f(x_n) \to A$.

2. Пусть A — предел функции f в точке a по Гейне; докажем, что тогда A — предел f и по Коши. Пусть не так, то есть A не есть предел по Коши. Записывая отрицание определения Коши, получаем, что найдется такое число $\varepsilon^* > 0$, что для любого $\delta > 0$ существует x (зависящее от δ) со свойствами:

$$x \in D, \quad x \neq a, \quad |x - a| < \delta, \quad |f(x) - A| \geqslant \varepsilon^*.$$

Следовательно, для каждого $n \in \mathbb{N}$ по числу $\delta = \frac{1}{n}$ найдется такая точка x_n, что

$$x_n \in D, \quad x_n \neq a, \quad |x_n - a| < \frac{1}{n}, \quad |f(x_n) - A| \geqslant \varepsilon^*.$$

По теореме о сжатой последовательности построенная последовательность $\{x_n\}$ стремится к a, так как

$$0 < |x_n - a| < \frac{1}{n}.$$

Тогда по определению Гейне $f(x_n) \to A$. По определению предела последовательности $\{f(x_n)\}$ для числа ε^* найдется такой номер N, что для всех $n > N$ справедливо неравенство $|f(x_n) - A| < \varepsilon^*$, что противоречит построению. $\square$

Установленная равносильность различных определений предела функции позволяет выбирать наиболее удобный способ доказательства теорем о пределах. Как правило, эти теоремы могут быть доказаны двумя способами: повторением рассуждений, которыми доказывались аналогичные теоремы для предела последовательности, или сведением к уже доказанным теоремам для предела последовательности с помощью определения Гейне. Чаще мы будем пользоваться вторым приемом.

Теорема 2. Единственность предела функции. *Функция в данной точке не может иметь более одного предела: если* $f\colon D \subset \mathbb{R} \to \mathbb{R}$, a *— предельная точка* D, $A, B \in \overline{\mathbb{R}}$, $f(x) \xrightarrow[x \to a]{} A$, $f(x) \xrightarrow[x \to a]{} B$, *то* $A = B$.

Доказательство. Возьмем последовательность $\{x_n\}$ со свойствами: $x_n \in D$, $x_n \neq a$, $x_n \to a$. По определению Гейне $f(x_n) \to A$ и $f(x_n) \to B$. В силу единственности предела последовательности $A = B$. $\square$

Замечание 1. В теореме 2 предполагалось, что $A, B \in \overline{\mathbb{R}}$. В то же время, как и для последовательностей, если $f(x) \xrightarrow[x \to a]{} \pm\infty$, то $f(x) \xrightarrow[x \to a]{} \infty$.

Теорема 3. Локальная ограниченность функции, имеющей предел. *Пусть* $f\colon D \subset \mathbb{R} \to \mathbb{R}$, a *— предельная точка* D, $A \in \mathbb{R}$, $f(x) \underset{x\to a}{\longrightarrow} A$. *Тогда существует такая окрестность* V_a *точки* a, *что* f *ограничена в* $V_a \cap D$.

Доказательство. По определению предела для числа $\varepsilon = 1$ найдется такая окрестность V_a точки a, что $|f(x) - A| < 1$ для всех $x \in \dot{V}_a \cap D$. Следовательно, для таких x будет $|f(x)| < |A| + 1$. Если $a \notin D$, то на этом доказательство заканчивается, так как $\dot{V}_a \cap D = V_a \cap D$. Если же $a \in D$, то

$$|f(x)| \leqslant \max\{|A| + 1, |f(a)|\}$$

для всех $x \in V_a \cap D$. $\square$

Замечание 2. Функция, имеющая предел в некоторой точке (и даже во всех точках своей области определения), не обязана быть ограниченной. Таковы, например, функции $f(x) = x$ $(x \in \mathbb{R})$ и $g(x) = \frac{1}{x}$ $(x \in \mathbb{R} \setminus \{0\})$. Поэтому в названии теоремы присутствует слово “локальная”.

Замечание 3. О стабилизации знака функции, имеющей предел. *Если* $f\colon D \subset \mathbb{R} \to \mathbb{R}$, a *— предельная точка* D, $\lim\limits_{x\to a} g(x) = B \in \overline{\mathbb{R}} \setminus \{0\}$, *то существует такая окрестность* V_a *точки* a, *что знаки* $g(x)$ *и* B *совпадают в* $\dot{V}_a \cap D$; *в частности,* $g(x) \neq 0$ *для всех* $x \in \dot{V}_a \cap D$. *Последнее верно и в случае* $\lim\limits_{x\to a} g(x) = \infty$.

Доказательство. Рассмотрим для определенности случай $B > 0$. Если утверждение неверно, то для любого $n \in \mathbb{N}$ существует точка $x_n \in \dot{V}_a(\frac{1}{n}) \cap D$, для которой $g(x_n) \leqslant 0$. Построенная последовательность $\{x_n\}$ стремится к a. По определению предела $g(x_n) \to B$, а по теореме о предельном переходе в неравенстве $B \leqslant 0$, что противоречит условию. $\square$

Теорема 4. Арифметические действия над функциями, имеющими предел. *Пусть* $f, g\colon D \subset \mathbb{R} \to \mathbb{R}$, a *— предельная точка* D, $A, B \in \mathbb{R}$, $f(x) \underset{x\to a}{\longrightarrow} A$, $g(x) \underset{x\to a}{\longrightarrow} B$. *Тогда*

1) $f(x) + g(x) \underset{x\to a}{\longrightarrow} A + B$;

2) $f(x)g(x) \underset{x\to a}{\longrightarrow} AB$;

3) $f(x) - g(x) \underset{x\to a}{\longrightarrow} A - B$;

4) $|f(x)| \underset{x\to a}{\longrightarrow} |A|$;

5) *если, кроме того,* $B \neq 0$, *то* $\dfrac{f(x)}{g(x)} \underset{x\to a}{\longrightarrow} \dfrac{A}{B}$.

Доказательство. С помощью определения на языке последовательностей теорема 4 сводится к теореме 5 § 1. Докажем, например, первое утверждение. Возьмем последовательность $\{x_n\}$ со свойствами: $x_n \in D$, $x_n \neq a$, $x_n \to a$. Тогда по определению Гейне $f(x_n) \to A$, $g(x_n) \to B$. По теореме о пределе суммы для последовательностей $f(x_n) + g(x_n) \to A + B$. В силу произвольности последовательности $\{x_n\}$ это и значит, что $f(x) + g(x) \underset{x\to a}{\longrightarrow} A + B$. При доказательстве утверждения о пределе частного следует еще учесть, что по замечанию 3 существует такая окрестность V_a, что частное $\frac{f}{g}$ определено по крайней мере на множестве $\dot{V}_a \cap D$. $\square$

Замечание 4. Теорема 4 верна и для бесконечных пределов, за исключением неопределенностей вида $\infty - \infty$, $0 \cdot \infty$, $\frac{0}{0}$ и $\frac{\infty}{\infty}$. При доказательстве нужно сослаться на теорему 6 § 1.

Замечание 5. Определение бесконечно малой и бесконечно большой переносится на функции. Так, функция, стремящаяся к нулю в точке a, называется *бесконечно малой* в точке a. Утверждения о том, что произведение бесконечно малой функции на ограниченную есть бесконечно малая, и о связи между бесконечно большими и бесконечно малыми сохраняют силу.

Теорема 5. О пределе композиции. *Пусть* $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $g\colon E \subset \mathbb{R} \to \mathbb{R}$, $f(D) \subset E$, $a, A, B \in \overline{\mathbb{R}} \cup \{\infty\}$ *и выполнены условия:*

1) a *— предельная точка* D, $f(x) \underset{x\to a}{\longrightarrow} A$;

2) A *— предельная точка* E, $g(x) \underset{x\to A}{\longrightarrow} B$;

3) *существует такая окрестность* V_a *точки* a, *что* $f(x) \neq A$ *для любого* $x \in \dot{V}_a \cap D$.

Тогда $(g \circ f)(x) \underset{x\to A}{\longrightarrow} B$.

Доказательство. Возьмем последовательность $\{x_n\}$ со свойствами: $x_n \in D$, $x_n \neq a$, $x_n \to a$, и обозначим $y_n = f(x_n)$. Тогда $y_n \in E$, и по определению Гейне $y_n \to A$. Кроме того, начиная с некоторого номера, $x_n \in V_a$, а тогда $y_n \neq A$ для таких n. Последовательность $\{y_n\}$ удовлетворяет условиям определения Гейне предела функции g. Поэтому $g(y_n) \to B$, то есть $(g \circ f)(x_n) \to B$. В силу произвольности последовательности $\{x_n\}$ это и значит, что $(g \circ f)(x) \underset{x \to A}{\longrightarrow} B$. $\square$

Пример, показывающий, что запрет $f(x)$ равняться A опустить нельзя, будет приведен в следующем параграфе, в замечании к теореме о непрерывности композиции.

Теорема 6. Предельный переход в неравенстве для функций. *Пусть $f, g\colon D \subset \mathbb{R} \to \mathbb{R}$, a — предельная точка D, $f(x) \leqslant g(x)$ для всех $x \in D \setminus \{a\}$, $A, B \in \overline{\mathbb{R}}$, $f(x) \underset{x \to a}{\longrightarrow} A$, $g(x) \underset{x \to a}{\longrightarrow} B$. Тогда $A \leqslant B$.*

Доказательство. Возьмем последовательность $\{x_n\}$ со свойствами: $x_n \in D$, $x_n \neq a$, $x_n \to a$. Тогда по определению Гейне $f(x_n) \to A$, $g(x_n) \to B$. По теореме о предельном переходе в неравенстве для последовательностей $A \leqslant B$. $\square$

Теорема 7. О сжатой функции. *Пусть $f, g, h\colon D \subset \mathbb{R} \to \mathbb{R}$, a — предельная точка D, $f(x) \leqslant g(x) \leqslant h(x)$ для всех $x \in D \setminus \{a\}$, $A \in \mathbb{R}$, $f(x) \underset{x \to a}{\longrightarrow} A$, $h(x) \underset{x \to a}{\longrightarrow} A$. Тогда и $g(x) \underset{x \to a}{\longrightarrow} A$.*

Доказательство. Возьмем последовательность $\{x_n\}$ со свойствами: $x_n \in D$, $x_n \neq a$, $x_n \to a$. Тогда по определению Гейне $f(x_n) \to A$, $h(x_n) \to A$. Кроме того, по условию для всех $n \in \mathbb{N}$

$$f(x_n) \leqslant g(x_n) \leqslant h(x_n).$$

По теореме о сжатой последовательности $g(x_n) \to A$. В силу произвольности последовательности $\{x_n\}$ это и значит, что $g(x) \underset{x \to a}{\longrightarrow} A$. $\square$

Замечание 1. Аналогично доказывается, что если $f(x) \leqslant g(x)$ для всех $x \in D \setminus \{a\}$ и $f(x) \underset{x \to a}{\longrightarrow} +\infty$ ($g(x) \underset{x \to a}{\longrightarrow} -\infty$), то и $g(x) \underset{x \to a}{\longrightarrow} +\infty$ ($f(x) \underset{x \to a}{\longrightarrow} -\infty$).

Замечание 2. В теоремах 6 и 7 и замечании 1 достаточно выполнения неравенств на множестве $\dot{V}_a \cap D$, где V_a — какая-нибудь окрестность точки a.

Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $D_1 \subset D$, a — предельная точка D_1 (а следовательно, и D). Тогда, если предел f в точке a существует и равен A, то предел сужения f на D_1 в точке a также существует и равен A. В самом деле, если соотношение $f(x) \in V_A$ выполняется для всех x из $\dot{V}_a \cap D$, то оно тем более выполняется для всех x из $\dot{V}_a \cap D_1$. Однако, возможна ситуация, когда предел сужения существует, а предел исходной функции — нет.

Определение. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $D_1 \subset D$, a — предельная точка D_1. Предел $\lim\limits_{x\to a} f\big|_{D_1}(x)$ называется *пределом* функции f в точке a *по множеству* D_1.

Предел подпоследовательности $\{x_{n_k}\}$ можно рассматривать как частный случай предела по множеству (а именно, по множеству значений последовательности индексов $\{n_k\}$). Рассмотрим другой типичный случай предела по множеству — односторонний предел функции.

Определение. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $a \in \mathbb{R}$.

1. Если a — предельная точка множества $D_1 = D \cap (-\infty, a)$, то предел функции f в точке a по множеству D_1 называется *левосторонним пределом* функции f в точке a и обозначается $\lim\limits_{x\to a-} f(x)$ или $f(a-)$.

2. Если a — предельная точка множества $D_2 = D \cap (a, +\infty)$, то предел функции f в точке a по множеству D_2 называется *правосторонним пределом* функции f в точке a и обозначается $\lim\limits_{x\to a+} f(x)$ или $f(a+)$.

Вместо $a\pm$ в обозначениях односторонних пределов иногда пишут $a \pm 0$; выражение $a \pm 0$ следует рассматривать как единое обозначение, а не сложение (вычитание).

Запишем определение левостороннего предела на разных языках. Равенство $\lim\limits_{x\to a-} f(x) = A$ $(A \in \mathbb{R})$ означает, что:

1) $\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in D : 0 < a - x < \delta \ \ |f(x) - A| < \varepsilon$;

2) $\forall V_A \ \ \exists \delta > 0 \ \ \forall x \in D : 0 < a - x < \delta \ \ f(x) \in V_A$;

3) $\forall \{x_n\} : \ x_n \in D, x_n < a, x_n \to a \ \ f(x_n) \to A$.

В определении правостороннего предела следует написать соответственно $0 < x - a < \delta$ и $x_n > a$. В записи на языке окрестностей можно ввести специальные обозначения для левой и правой окрестности и проколотой окрестности точки a (промежутков $(a-\delta, a]$, $[a, a+\delta)$, $(a-\delta, a)$, $(a, a+\delta)$), но мы не будем ими пользоваться. Под левосторонним пределом в точке $+\infty$ и правосторонним пределом в точке $-\infty$ понимают обычный предел в этих точках.

Замечание 1. Пусть $f: D \subset \mathbb{R} \to \mathbb{R}$, $a \in \mathbb{R}$, $D_1 = D \cap (-\infty, a)$, $D_2 = D \cap (a, +\infty)$, a — предельная точка D_1 и D_2. Тогда для существования предела f в точке a необходимо и достаточно, чтобы односторонние пределы f в точке a существовали и были равны друг другу.

Доказательство. Ясно, что если предел существует, то односторонние пределы существуют и равны ему. Обратно, если односторонние пределы существуют и равны A, то для любой окрестности точки A найдутся такие $\delta_1, \delta_2 > 0$, что соотношение $f(x) \in V_A$ выполняется для всех $x \in D$: $0 < a - x < \delta_1$ и для всех $x \in D$: $0 < x - a < \delta_2$. Тогда оно выполняется для всех $x \in D$: $0 < |x - a| < \delta$, где $\delta = \min\{\delta_1, \delta_2\}$. $\square$

В условиях замечания 1 предел функции f в точке a называют двусторонним, в отличие от односторонних пределов. Типичный случай, когда выполняются условия замечания: функция задана по крайней мере в проколотой окрестности точки a.

Теорема 8. О пределе монотонной функции.
Пусть $f: D \subset \mathbb{R} \to \mathbb{R}$, $a \in (-\infty, +\infty]$, $D_1 = D \cap (-\infty, a)$, a — *предельная точка* D_1.

1. *Если f возрастает и ограничена сверху на D_1, то существует конечный предел $f(a-)$.*

2. *Если f убывает и ограничена снизу на D_1, то существует конечный предел $f(a-)$.*

Доказательство. Докажем первое утверждение; второе доказывается аналогично. Положим $A = \sup\limits_{x \in D_1} f(x)$. Тогда $A \in \mathbb{R}$ в силу ограниченности функции сверху. Докажем, что $f(a-) = A$. Возьмем $\varepsilon > 0$. По определению верхней грани существует такая точка $x_0 \in D_1$, что $f(x_0) > A - \varepsilon$. Но тогда для всех таких $x \in D_1$, что $x > x_0$, в силу возрастания f

$$A - \varepsilon < f(x_0) \leqslant f(x) \leqslant A < A + \varepsilon.$$

Теперь положим $\delta = a - x_0$ при $a \in \mathbb{R}$ или $\Delta = \max\{x_0, 1\}$ при $a = +\infty$; тогда неравенство из определения предела выполнено для всех таких $x \in D$, что $0 < a - x < \delta$ (соответственно, $x > \Delta$). $\square$

Замечание 2. Аналогично утверждениям 1 и 2 теоремы доказываются следующие утверждения.

3. *Если f возрастает и не ограничена сверху на D_1, то предел $f(a-)$ существует и равен $+\infty$.*

4. *Если f убывает и не ограничена снизу на D_1, то предел $f(a-)$ существует и равен $-\infty$.*

Замечание 3. Сформулируем аналогичную теорему для правостороннего предела.

Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $a \in [-\infty, +\infty)$, $D_2 = D \cap (a, +\infty)$, a — предельная точка D_2.

1. *Если f возрастает и ограничена снизу на D_2, то существует конечный предел $f(a+)$.*

2. *Если f убывает и ограничена сверху на D_2, то существует конечный предел $f(a+)$.*

3. *Если f возрастает и не ограничена снизу на D_2, то предел $f(a+)$ существует и равен $-\infty$.*

4. *Если f убывает и не ограничена сверху на D_2, то предел $f(a+)$ существует и равен $+\infty$.*

Чаще всего встречаются случаи, когда функция f задана по крайней мере на промежутке вида $\langle c, a)$ или $(a, b\rangle$.

Теорема 9. Критерий Больцано – Коши для функций. *Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, a — предельная точка D. Тогда существование конечного предела f в точке a равносильно следующему утверждению:*

для любого положительного числа ε существует такая окрестность V_a точки a, что для любых двух точек $\bar{x}$ и $\bar{\bar{x}}$ множества D, принадлежащих проколотой окрестности $\dot{V}_a$, выполняется неравенство $|f(\bar{x}) - f(\bar{\bar{x}})| < \varepsilon$:

$$\forall\, \varepsilon > 0 \quad \exists V_a \quad \forall\, \bar{x}, \bar{\bar{x}} \in \dot{V}_a \cap D \quad |f(\bar{x}) - f(\bar{\bar{x}})| < \varepsilon. \tag{3}$$

Доказательство. 1. Пусть $\lim\limits_{x\to a} f(x) = A \in \mathbb{R}$. Возьмем $\varepsilon > 0$. По определению предела найдется такая окрестность V_a точки a, что $|f(x) - A| < \frac{\varepsilon}{2}$ для всех $x \in \dot{V}_a \cap D$. Тогда, если $\bar{x}, \bar{\bar{x}} \in \dot{V}_a \cap D$, то

$$|f(\bar{x}) - f(\bar{\bar{x}})| \leqslant |f(\bar{x}) - A| + |A - f(\bar{\bar{x}})| < \varepsilon.$$

В силу произвольности ε условие (3) выполнено.

2. Пусть выполнено условие (3). Докажем существование предела f в точке a на языке последовательностей. Возьмем последовательность $\{x_n\}$ со свойствами: $x_n \in D$, $x_n \neq a$, $x_n \to a$, и докажем, что существует $\lim f(x_n) \in \mathbb{R}$. По $\varepsilon > 0$ подберем окрестность V_a из условия (3). По определению предела $\{x_n\}$ найдется такой номер N, что $x_n \in V_a$ для всех $n > N$; тогда $x_n \in \dot{V}_a \cap D$ для тех же n. По выбору V_a для всех $n, l > N$ будет $|f(x_n) - f(x_l)| < \varepsilon$. Таким образом, последовательность $\{f(x_n)\}$ сходится в себе и, значит, имеет конечный предел. Тогда в силу замечания 6 к определению предела функция f имеет конечный предел в точке a. $\square$

Предлагаем читателю самостоятельно записать критерий Больцано – Коши без использования символа окрестности для случаев $a \in \mathbb{R}$, $a = \pm\infty$, $a = \infty$ и для односторонних пределов.

§ 4. Непрерывные функции

Определение непрерывности функции в точке x_0 — формализация того свойства, что значения функции в точках, близких к x_0, близки к $f(x_0)$.

Определение. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $x_0 \in D$. Функция f называется *непрерывной* в точке x_0, если выполняется одно из следующих утверждений.

1. Предел функции f в точке x_0 существует и равен $f(x_0)$. Это определение применимо, если x_0 — предельная точка D.

2. Для любого положительного числа ε существует такое положительное число δ, что для всех точек x множества D, удовлетворяющих неравенству $|x - x_0| < \delta$, выполняется неравенство $|f(x) - f(x_0)| < \varepsilon$:

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in D : |x - x_0| < \delta \;\; |f(x) - f(x_0)| < \varepsilon.$$

3. Для любой окрестности $V_{f(x_0)}$ точки $f(x_0)$ существует такая окрестность V_{x_0} точки x_0, что образ пересечения окрестности V_{x_0} с множеством D содержится в окрестности $V_{f(x_0)}$:

$$\forall V_{f(x_0)} \quad \exists V_{x_0} \quad f\big(V_{x_0} \cap D\big) \subset V_{f(x_0)}.$$

4. Для любой последовательности $\{x_n\}$ точек множества D, стремящейся к x_0, последовательность $\{f(x_n)\}$ стремится к $f(x_0)$:

$$\forall \{x_n\} : \; x_n \in D, \, x_n \to x_0 \quad f(x_n) \to f(x_0).$$

5. Бесконечно малому приращению аргумента соответствует бесконечно малое приращение функции:

$$\Delta y \underset{\Delta x \to 0}{\longrightarrow} 0.$$

Здесь

$$\Delta x = x - x_0, \qquad \Delta y = f(x) - f(x_0)$$

— приращение аргумента и приращение функции. Это определение применимо, если x_0 — предельная точка D.

Определение 2 — это определение на ε-языке, или по Коши, определение 3 — на языке окрестностей, а определение 4 — на языке последовательностей, или по Гейне.

Замечание 1. Равносильность определений в случае, когда x_0 — предельная точка D, следует из равносильности различных определений предела. Под номерами 2, 3 и 4 на ε-языке, языке окрестностей и языке последовательностей записан тот факт, что число $A = f(x_0)$ является пределом функции f в точке x_0, с одним отличием в каждом случае. В определении на ε-языке опущено условие $x \neq x_0$. Это не меняет определения, так как неравенство $|f(x) - f(x_0)| < \varepsilon$ при $x = x_0$, очевидно, выполнено. В определении на языке окрестностей окрестность V_{x_0} не проколота. Это не меняет определения, так как при $x = x_0$, очевидно, $f(x) \in V_{f(x_0)}$. В определении на языке последовательностей опущено условие $x_n \neq x_0$. Это также не меняет определения, потому что если $x_n = x_0$ при некоторых n, то $f(x_n) = f(x_0)$ при таких n, и неравенство $|f(x_n) - f(x_0)| < \varepsilon$ из определения предела последовательности $\{f(x_n)\}$ для этих n выполняется. Наконец, определение 5 на любом из языков записывается так же, как и определение 1.

Замечание 2. Пусть x_0 — изолированная точка D. Тогда в окрестности $V_{x_0}(\delta)$ при достаточно малом $\delta > 0$ нет точек множества D, отличных от x_0, и, следовательно,

$$f\big(V_{x_0}(\delta) \cap D\big) = \{f(x_0)\} \subset V_{f(x_0)},$$

каковы бы ни были функция f и окрестность $V_{f(x_0)}$. Если $x_n \in D$, $x_n \to x_0$, то, начиная с некоторого номера, $x_n \in V_{x_0}(\delta)$ и, значит, $x_n = x_0$, а тогда $f(x_n) \to f(x_0)$, какова бы ни была функция f. Поэтому определения 2–4 равносильны и в случае изолированной точки x_0.

Согласно этому замечанию, *всякая функция непрерывна в изолированной точке своей области определения.*

Определение. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $x_0 \in D$. Если функция f не является непрерывной в точке x_0, то говорят, что f *разрывна* (терпит разрыв, испытывает разрыв) в точке x_0, а точку x_0 называют *точкой разрыва* функции f.

Определение. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $x_0 \in D$. Если сужение функции f на множество $E_1 = D \cap (-\infty, x_0]$ ($E_2 = D \cap [x_0, +\infty)$) непрерывно в точке x_0, то говорят, что функция f *непрерывна слева (справа)* в точке x_0.

Если x_0 — предельная точка множества E_1 (соответственно, E_2), то непрерывность функции f в точке x_0 слева (справа) означает, что предел f в точке x_0 слева (справа) существует и равен $f(x_0)$. Читателю предлагается самому записать определение односторонней непрерывности другими способами.

Пусть функция задана по крайней мере в окрестности точки x_0. Из замечания 1 к определению односторонних пределов следует, что непрерывность функции f в точке x_0 равносильна ее одновременной непрерывности слева и справа в точке x_0.

Если существуют конечные пределы $f(x_0-)$ и $f(x_0+)$, но не все три числа $f(x_0-)$, $f(x_0+)$, $f(x_0)$ равны между собой, то точку x_0 называют *точкой разрыва первого рода* функции f. Разрыв первого рода еще называют *скачком.* Скачком также называют разность

$f(x_0+) - f(x_0-)$, а разности $f(x_0) - f(x_0-)$ и $f(x_0+) - f(x_0)$ называют скачками функции f слева и справа в точке x_0. В противном случае, то есть если хотя бы один из односторонних пределов в точке разрыва x_0 бесконечен или вовсе не существует, точку x_0 называют *точкой разрыва второго рода* функции f.

Часто встречается ситуация, когда функция f задана по крайней мере в проколотой окрестности точки x_0, но не задана в точке x_0. В этом случае x_0 также называют точкой разрыва функции f (первого рода, если существуют конечные пределы $f(x_0-) \neq f(x_0+)$ и второго рода, если если хотя бы один из односторонних пределов в точке x_0 бесконечен или вовсе не существует), несмотря на то, что f не определена в точке x_0. Эти названия оправданы тем, что, как бы ни доопределить функцию в точке x_0, для новой функции x_0 будет точкой разрыва первого или второго рода соответственно.

Пусть $f(x_0-) = f(x_0+) = A \in \mathbb{R}$, но $f(x_0) \neq A$ или f не определена в точке x_0. Тогда точку x_0 называют *точкой устранимого разрыва* функции f. Если доопределить или переопределить функцию в точке x_0, как говорят, "по непрерывности", т.е положить $f(x_0) = A$, то новая функция (которую обычно обозначают той же буквой, "забывая" об исходной функции) будет непрерывна в точке x_0. Далее удобно считать, что функция, не определенная в точке устранимого разрыва, доопределяется в ней по непрерывности.

Если точка x_0 — один из концов промежутка, на котором определена функция (точнее, $f\colon D \subset \mathbb{R} \to \mathbb{R}$ и существует такой интервал (a, b), что $(a, b) \cap D$ равно $(a, x_0]$ или $[x_0, b)$), то имеет смысл говорить лишь об одном из односторонних пределов в точке x_0. Если при этом x_0 — точка разрыва функции f, то ее также называют точкой разрыва первого или второго рода в зависимости от того, существует ли конечный односторонний предел f в этой точке. В случае существования конечного $f(x_0+)$ или $f(x_0-)$ имеет смысл говорить о скачке f с одной стороны.

Примеры. Следующие примеры иллюстрируют разные возможности, возникающие в точках разрыва.

1. Функция сигнум (знак): $f(x) = \operatorname{sign} x = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0. \end{cases}$

Тогда $f(0+) = 1$, $f(0-) = -1$, и 0 — точка разрыва первого рода, что иллюстрирует рисунок 9.

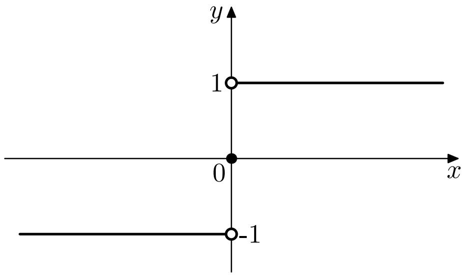

Рис. 9

2. $f(x) = |\operatorname{sign} x| = \begin{cases} 1, & x \neq 0, \\ 0, & x = 0. \end{cases}$
Тогда $f(0+) = f(0-) = 1$, и 0 — точка разрыва первого рода, притом устранимого разрыва (рисунок 10).

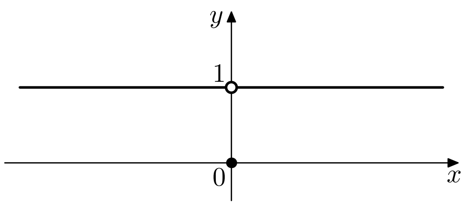

Рис. 10

3. $f(x) = \frac{1}{x}$. Тогда $f(0+) = +\infty$, $f(0-) = -\infty$, и 0 — точка разрыва второго рода.

4. $f(x) = \dfrac{1}{x^2}$. Тогда $f(0+) = f(0-) = +\infty$, и 0 — точка разрыва второго рода.

5. $f(x) = \dfrac{x-1}{x^2-1}$. Тогда f определена на $\mathbb{R} \setminus \{-1, 1\}$, и на области определения $f(x) = \frac{1}{x+1}$. Точка -1 — точка разрыва второго рода, 1 — точка устранимого разрыва. Положив $f(1) = \frac{1}{2}$, получим непрерывную в точке 1 функцию.

6. $f(x) = \sin\frac{1}{x}$. Докажем, что 0 — точка разрыва второго рода. Положим $x_n = \dfrac{1}{2n\pi}$, $y_n = \dfrac{1}{(2n+\frac{1}{2})\pi}$, $n \in \mathbb{N}$. Тогда $x_n, y_n > 0$, $x_n, y_n \to 0$, $f(x_n) = 0 \to 0$, $f(y_n) = 1 \to 1$. Отсюда следует, что правостороннего предела f в точке 0 не существует, так как определение этого предела на языке последовательностей не выполняется. Аналогично доказывается отсутствие левостороннего предела f в точке 0. $\square$

График этой функции изображен на рисунке 11.

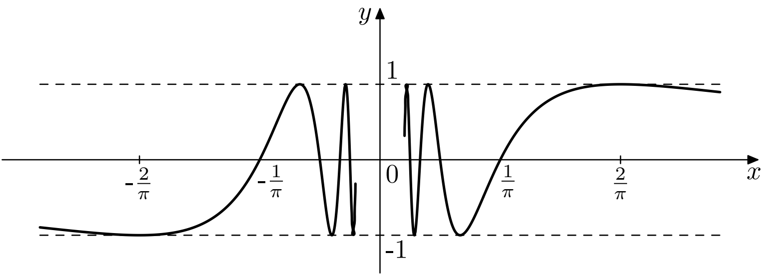

Рис. 11

7. $f(x) = x\sin\frac{1}{x}$. Здесь $\lim\limits_{x\to 0} f(x) = 0$, потому что f есть произведение бесконечно малой на ограниченную. Поэтому 0 — точка устранимого разрыва (рисунок 12).

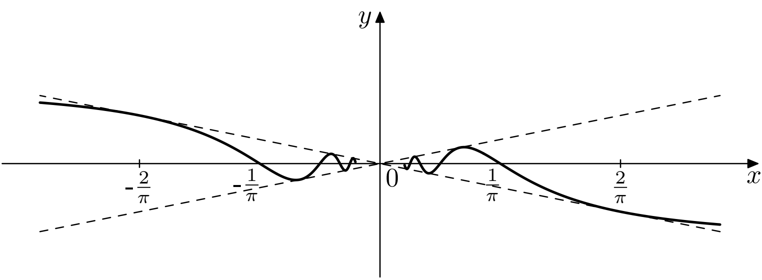

Рис. 12

8. $f(x) = 2^{1/x}$. Тогда $f(0+) = +\infty$, $f(0-) = 0$, и 0 — точка разрыва второго рода (рисунок 13).

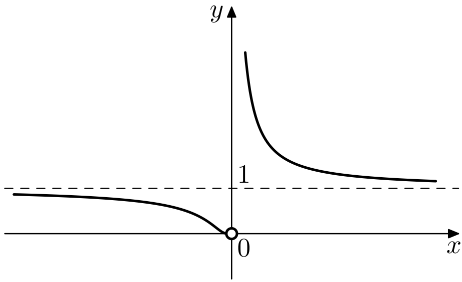

Рис. 13

Эти факты, а также непрерывность функций из примеров 3–8 в остальных точках, следуют из свойств элементарных функций, которые будут установлены в § 5.

9. Функция Дирихле

$$\chi(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q} \end{cases}$$

разрывна в каждой точке, причем все разрывы второго рода.

Действительно, для каждой точки $x_0 \in \mathbb{R}$ можно подобрать две последовательности: рациональных чисел $\{x_n\}$ и иррациональных чисел $\{y_n\}$, такие что $x_n, y_n > x_0$, $x_n, y_n \to x_0$. Например, можно взять $x_n = \frac{[nx_0]+1}{n}$, $y_n = x_n + \frac{\sqrt{2}}{n}$. Тогда $\chi(x_n) = 1 \to 1$, $\chi(y_n) = 0 \to 0$. Поэтому предела χ справа в точке x_0 не существует. Аналогично доказывается отсутствие левостороннего предела. $\square$

10. Функция Римана

$$\psi(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \in \mathbb{Q}, \text{ дробь несократима}, \\ 0, & x \notin \mathbb{Q} \end{cases}$$

непрерывна во всех иррациональных и имеет разрывы первого рода во всех рациональных точках.

Достаточно доказать, что предел ψ в любой точке x_0 равен 0. Возьмем $\varepsilon > 0$ и подберем такой номер N, что $\frac{1}{N} < \varepsilon$. Количество рациональных чисел со знаменателями, меньшими N, в проколотой 2-окрестности точки x_0 конечно; обозначим через r расстояние от x_0 до ближайшего из них. Тогда в проколотой r-окрестности x_0 нет рациональных чисел со знаменателями, меньшими N, то есть все значения функции ψ в этой проколотой окрестности меньше ε. $\square$

Изобразить графики функций Дирихле и Римана не удается из-за хаотического поведения этих функций.

Определение. Функция называется *непрерывной на множестве* D, если она непрерывна в каждой точке множества D.

Множество функций, заданных и непрерывных на множестве D, обозначается $C(D)$. Если $D = \langle a, b\rangle$ — промежуток, то скобки обычно опускают и пишут $C\langle a, b\rangle$.

Теорема 1. Арифметические действия над непрерывными функциями. *Пусть функции* $f, g\colon D \subset \mathbb{R} \to \mathbb{R}$ *непрерывны в точке* $x_0 \in D$. *Тогда функции* $f + g$, $f - g$, fg, $|f|$ *непрерывны в точке* x_0, *а если* $g(x_0) \neq 0$, *то и* $\frac{f}{g}$ *непрерывна в точке* x_0.

Доказательство. Если x_0 — изолированная точка D, то утверждение тривиально. Если же x_0 — предельная точка D, то теоремы о непрерывности следуют из теоремы 4 § 3 о пределах. Если f и g непрерывны в точке x_0, то

$$f(x) \underset{x\to x_0}{\longrightarrow} f(x_0), \qquad g(x) \underset{x\to x_0}{\longrightarrow} g(x_0).$$

Тогда

$$f(x) + g(x) \underset{x\to x_0}{\longrightarrow} f(x_0) + g(x_0),$$

что и означает непрерывность суммы в точке x_0. Рассуждение для других операций проводится аналогично. $\square$

Замечание 1. О стабилизации знака непрерывной функции. *Если функция* $g\colon D \subset \mathbb{R} \to \mathbb{R}$ *непрерывна в точке* $x_0 \in D$, $g(x_0) \neq 0$, *то существует такая окрестность* V_{x_0}, *что знак* $g(x)$ *совпадает со знаком* $g(x_0)$ *для всех* $x \in V_{x_0} \cap D$.

Для изолированной точки x_0 это утверждение тривиально, а для предельной точки следует из замечания о стабилизации знака функции, имеющей предел.

Примеры. Постоянная функция, очевидно, непрерывна на $\mathbb{R}$. Функция $f(x) = x$ непрерывна на $\mathbb{R}$ (в определении на ε-языке достаточно положить $\delta = \varepsilon$, а определение на языке последовательностей выполняется тривиально). Тогда по теореме 1 всякий многочлен, то есть функция вида

$$P(x) = \sum_{k=0}^{n} c_k x^k,$$

непрерывен на $\mathbb{R}$, а всякая рациональная дробь, то есть частное двух многочленов, непрерывна на своей области определения (то есть множестве точек, где знаменатель не обращается в нуль).

Теорема 2. Непрерывность композиции. *Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$, $g\colon E \subset \mathbb{R} \to \mathbb{R}$, $f(D) \subset E$, f непрерывна в точке $x_0 \in D$, g непрерывна в точке $f(x_0)$. Тогда $g \circ f$ непрерывна в точке x_0.*

Доказательство. Пусть $x_n \in D$, $x_n \to x_0$. Обозначим $y_n = f(x_n)$, $y_0 = f(x_0)$; тогда $y_n, y_0 \in E$. По определению непрерывности f в точке x_0 на языке последовательностей $y_n \to y_0$. По определению непрерывности g в точке y_0 на языке последовательностей $g(y_n) \to g(y_0)$, то есть $(g \circ f)(x_n) \to (g \circ f)(x_0)$. Последнее в силу произвольности последовательности $\{x_n\}$ и означает непрерывность $g \circ f$ в точке x_0. $\square$

Замечание 2. Пусть $f(x) = x \sin \frac{1}{x}$, $g(y) = |\operatorname{sign} y|$. Тогда $f(x) \underset{x\to 0}{\longrightarrow} 0$, $g(y) \underset{y\to 0}{\longrightarrow} 0$, но композиция $g \circ f$ не имеет предела в нуле, так как $(g \circ f)(\frac{1}{n\pi}) = 0 \to 0$, а $(g \circ f)(\frac{1}{(n+1/2)\pi}) = 1 \to 1$. Этот пример показывает, что утверждение “если $f(x) \underset{x\to a}{\longrightarrow} A$, $g(x) \underset{x\to A}{\longrightarrow} B$, то $(g \circ f)(x) \underset{x\to a}{\longrightarrow} B$” может не выполняться. Теоремы о пределе и о непрерывности композиции предоставляют два способа исправить это утверждение.

Перейдем к глобальным свойствам непрерывных функций, то есть к свойствам, связанным с непрерывностью на всей области определения.

Теорема 3 (К. Вейерштрасс). О непрерывных функциях. 1. *Непрерывная на отрезке функция ограничена.*

2. *Непрерывная на отрезке функция принимает наибольшее и наименьшее значение.*

Доказательство. Пусть $f \in C[a, b]$.

1. Предположим, что f не ограничена. Тогда для любого n найдется точка $x_n \in [a, b]$, такая что $|f(x_n)| > n$. Пользуясь принципом выбора Больцано – Вейерштрасса, выделим из построенной последовательности $\{x_n\}$ подпоследовательность $\{x_{n_k}\}$, сходящуюся к некоторому пределу c: $x_{n_k} \to c$. По свойству замкнутости отрезка $c \in [a, b]$. Функция f непрерывна на отрезке $[a, b]$ и, в частности, в точке c. По определению непрерывности $f(x_{n_k}) \to f(c)$. Тогда сходящаяся последовательность $\{f(x_{n_k})\}$ ограничена, что противоречит построению, так как $|f(x_{n_k})| > n_k \geqslant k$ при всех $k \in \mathbb{N}$.

2. Для определенности докажем, что f принимает наибольшее значение. Обозначим $M = \sup\limits_{x \in [a,b]} f(x)$. По доказанному утверждению 1 функция f ограничена сверху, то есть $M \in \mathbb{R}$. Допустим, что M не является значением функции. Тогда $f(x) < M$ при всех $x \in [a, b]$, и по теореме 1 функция $\varphi = \frac{1}{M-f}$ непрерывна на $[a, b]$. По утверждению 1 она ограничена сверху: существует такое положительное число μ, что

$$\varphi(x) \leqslant \mu \quad \text{для всех } x \in [a, b].$$

Но тогда

$$f(x) \leqslant M - \frac{1}{\mu} \quad \text{для всех } x \in [a, b],$$

то есть число $M - \frac{1}{\mu}$, меньшее M, является верхней границей функции f, что противоречит определению M. $\square$

Замечание 1. Часто первое утверждение теоремы называют первой, а второе — второй теоремой Вейерштрасса. Вторая теорема Вейерштрасса уточняет первую (если функция имеет наибольшее и наименьшее значение, то она ограничена). Вместе с тем, первая теорема Вейерштрасса допускает обобщение на комплекснозначные функции, а вторая лишена для них смысла.

Замечание 2. Оба условия — и непрерывность функции, и то, что ее область определения есть отрезок, — существенны. Так, функции $f_1(x) = x$ и $f_2(x) = \frac{1}{x}$ непрерывны, но не ограничены, соответственно, на $\mathbb{R}$ и $(0,1]$. Функция

$$f_3(x) = \begin{cases} \frac{1}{x}, & x \in (0,1], \\ 0, & x = 0, \end{cases}$$

задана на $[0,1]$, разрывна в одной точке 0, но не ограничена. Функции f_1, f_2 и f_3 не имеют наибольшего значения. Наибольшего значения не имеет и ограниченная непрерывная функция $f_4(x) = x$ на $[0,1)$.

Обсудим подробнее определение непрерывности. Непрерывность функции $f\colon D \subset \mathbb{R} \to \mathbb{R}$ на множестве D означает, что f непрерывна в каждой точке $\bar{x} \in D$. Запишем это на ε-языке:

$$\forall \bar{x} \in D \;\; \forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall \bar{\bar{x}} \in D : |\bar{x} - \bar{\bar{x}}| < \delta \;\; |f(\bar{x}) - f(\bar{\bar{x}})| < \varepsilon.$$

В этой формуле число δ зависит и от ε, и от точки $\bar{x}$.

Возникает вопрос: можно ли в определении непрерывности подобрать число δ, зависящее только от ε и обслуживающее одновременно все точки $\bar{x}$ множества D?

Определение. Функция $f\colon D \subset \mathbb{R} \to \mathbb{R}$ называется *равномерно непрерывной* на множестве D, если для любого положительного числа ε существует такое положительное число δ, что для всех точек $\bar{x}$, $\bar{\bar{x}}$ множества D, удовлетворяющих неравенству $|\bar{x} - \bar{\bar{x}}| < \delta$, выполняется неравенство $|f(\bar{x}) - f(\bar{\bar{x}})| < \varepsilon$:

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall \bar{x}, \bar{\bar{x}} \in D : |\bar{x} - \bar{\bar{x}}| < \delta \;\; |f(\bar{x}) - f(\bar{\bar{x}})| < \varepsilon.$$

Как видно, в определении равномерной непрерывности точка $\bar{x}$ с квантором общности написана после δ, то есть δ зависит только от ε. В определении требуется, чтобы неравенство $|f(\bar{x}) - f(\bar{\bar{x}})| < \varepsilon$ выполнялось одновременно для всех пар точек, расстояние между которыми меньше δ. Из определений ясно, что всякая равномерно непрерывная функция непрерывна. Обратное неверно, что будет показано на примерах.

Запишем еще отрицание определения равномерной непрерывности функции f на множестве D:

$$\exists \varepsilon^* > 0 \;\; \forall \delta > 0 \;\; \exists \bar{x}_\delta, \bar{\bar{x}}_\delta \in D : |\bar{x}_\delta - \bar{\bar{x}}_\delta| < \delta, \; |f(\bar{x}_\delta) - f(\bar{\bar{x}}_\delta)| \geqslant \varepsilon^*.$$

Примеры.

1. Функция $f(x)=x$ равномерно непрерывна на $\mathbb{R}$: в определении можно взять $\delta=\varepsilon$.

2. Докажем, что функция $g(x)=x^2$ не является равномерно непрерывной на $\mathbb{R}$. Положим $\varepsilon^*=1$ и возьмем $\delta>0$. Пусть $\bar{x}_\delta=\frac{1}{\delta}$, $\bar{\bar{x}}_\delta=\frac{1}{\delta}+\frac{\delta}{2}$. Тогда

$$|\bar{x}_\delta-\bar{\bar{x}}_\delta|=\frac{\delta}{2}<\delta,$$

но

$$|\bar{x}_\delta^2-\bar{\bar{x}}_\delta^2|=\frac{\delta}{2}\left(\frac{2}{\delta}+\frac{\delta}{2}\right)>1=\varepsilon^*,$$

то есть для функции g выполнено отрицание определения равномерной непрерывности.

3. Читателю предлагается самому убедиться, что непрерывные функции $h(x)=\frac{1}{x}$ и $k(x)=\sin\frac{1}{x}$ (последняя к тому же ограничена) не являются равномерно непрерывными на $(0,1]$.

Теорема 5 (Г. Кантор). *Непрерывная на отрезке функция равномерно непрерывна.*

Доказательство. Пусть $f\in C[a,b]$. Предположим, что f не является равномерно непрерывной. Тогда существует такое $\varepsilon^*>0$, что при каждом $n\in\mathbb{N}$ найдутся точки $\bar{x}_n,\bar{\bar{x}}_n\in[a,b]$:

$$\left|\bar{x}_n-\bar{\bar{x}}_n\right|<\frac{1}{n},\qquad \left|\bar{y}_n-\bar{\bar{y}}_n\right|\geqslant\varepsilon^*,$$

где $\bar{y}_n=f(\bar{x}_n)$, $\bar{\bar{y}}_n=f\big(\bar{\bar{x}}_n\big)$. Пользуясь принципом выбора Больцано – Вейерштрасса, выделим из последовательности $\{\bar{x}_n\}$ подпоследовательность $\{\bar{x}_{n_k}\}$, сходящуюся к некоторому пределу $c\in\mathbb{R}$: $\bar{x}_{n_k}\to c$; по свойству замкнутости отрезка $c\in[a,b]$. Тогда и $\bar{\bar{x}}_{n_k}\to c$, так как

$$\left|\bar{\bar{x}}_{n_k}-c\right|\leqslant\left|\bar{\bar{x}}_{n_k}-\bar{x}_{n_k}\right|+\left|\bar{x}_{n_k}-c\right|<\frac{1}{n_k}+\left|\bar{x}_{n_k}-c\right|\to 0.$$

По непрерывности f в точке c

$$\bar{y}_{n_k}\to f(c),\qquad \bar{\bar{y}}_{n_k}\to f(c).$$

Следовательно, $\bar{y}_{n_k}-\bar{\bar{y}}_{n_k}\to 0$ и, начиная с некоторого номера, $\left|\bar{y}_{n_k}-\bar{\bar{y}}_{n_k}\right|<\varepsilon^*$, что противоречит построению. $\square$

Теорема 6 (Б. Больцано, О. Коши). О промежуточном значении непрерывной функции. *Пусть функция f непрерывна на $[a,b]$. Тогда для любого числа C, лежащего между $f(a)$ и $f(b)$, найдется такое $c \in (a,b)$, что $f(c) = C$.*

Доказательство. 1. Пусть числа $f(a)$ и $f(b)$ разных знаков: $f(a)f(b) < 0$; докажем, что существует такая точка $c \in (a,b)$, что $f(c) = 0$. Не умаляя общности, будем считать, что $f(a) < 0 < f(b)$; второй случай рассматривается аналогично. Рассмотрим середину отрезка $[a,b]$ — точку $\frac{a+b}{2}$. Если $f(\frac{a+b}{2}) = 0$, то теорема доказана — можно положить $c = \frac{a+b}{2}$. Иначе положим

$$[a_1, b_1] = \begin{cases} [\frac{a+b}{2}, b], & \text{если } f(\frac{a+b}{2}) < 0, \\ [a, \frac{a+b}{2}], & \text{если } f(\frac{a+b}{2}) > 0. \end{cases}$$

В обоих случаях $f(a_1) < 0 < f(b_1)$. Продолжим этот процесс построения промежутков. Если на некотором шаге функция f обратится в 0 в середине отрезка, то доказательство на этом закончится. Иначе будет построена последовательность вложенных отрезков $\{[a_n, b_n]\}$, таких что $f(a_n) < 0 < f(b_n)$. При этом отрезки стягивающиеся, так как $b_n - a_n = \frac{b-a}{2^n} \to 0$. По теореме о стягивающихся отрезках существует единственная точка c, принадлежащая одновременно всем отрезкам $[a_n, b_n]$, при этом $a_n \to c$ и $b_n \to c$. По теореме о предельном переходе в неравенстве $f(c) \leqslant 0 \leqslant f(c)$, то есть $f(c) = 0$. Следовательно, $c \in (a,b)$, и точка c — требуемая.

2. Докажем теорему в общем случае. Пусть $\varphi = f - C$. Тогда $\varphi \in C[a,b]$ как разность непрерывных функций, $\varphi(a)\varphi(b) < 0$. По доказанному существует такая точка $c \in (a,b)$, что $\varphi(c) = 0$, то есть $f(c) = C$. $\square$

Замечание 1. Теорему Больцано – Коши можно сформулировать и так: *если непрерывная на промежутке функция принимает какие-то два значения, то она принимает все значения, лежащие между ними.*

Здесь существенно и то, что функция непрерывна, и то, что она задана на промежутке. Функция sign, заданная на $\mathbb{R}$, разрывна в нуле. Она принимает значения -1 и 1, но из чисел между -1 и 1 только 0 является значением функции. Сужение функции sign на $\mathbb{R}\setminus\{0\}$ непрерывно, но не принимает значений, лежащих между -1 и 1.

Замечание 2. Иногда первую часть теоремы — утверждение о том, что непрерывная функция, принимающая на концах промежутка значения разных знаков, имеет на этом промежутке корень, называют первой теоремой Больцано – Коши, а теорему в общем случае — второй теоремой Больцано – Коши.

Замечание 3. Метод половинного деления, использованный при доказательстве первой части теоремы, позволяет приближенно находить корни уравнений.

Лемма 1. Характеристика промежутков. *Пусть $E \subset \mathbb{R}$. Тогда следующие утверждения равносильны.*

1. *E — промежуток (возможно, вырожденный).*
2. *Для любых x, y, принадлежащих E $(x < y)$, $[x, y] \subset E$.*

Доказательство. Второе утверждение следует из первого тривиально. Докажем обратный переход. Пусть $E \neq \varnothing$. Обозначим $m = \inf E$, $M = \sup E$. Ясно, что $E \subset [m, M]$. Докажем, что $(m, M) \subset E$. Пусть $m < z < M$. Тогда по определению граней существуют точки $x, y \in E$: $x < z < y$. По условию $z \in E$. $\square$

Замечание 4. Из курса геометрии известно понятие выпуклого множества. Множество (на прямой, на плоскости, в трехмерном пространстве) называется *выпуклым,* если вместе с любыми своими двумя точками оно содержит весь отрезок, их соединяющий. Лемма 1 утверждает, что на прямой выпуклыми являются промежутки, и только они.

Теоремы Вейерштрасса и Больцано – Коши позволяют делать выводы о множестве значений непрерывной функции.

Теорема 7. О сохранении промежутка. *Множество значений непрерывной на промежутке функции есть промежуток.*

Доказательство. Пусть $f \in C\langle a, b\rangle$,

$$m = \inf_{x\in\langle a,b\rangle} f(x), \qquad M = \sup_{x\in\langle a,b\rangle} f(x)$$

$(m, M \in \overline{\mathbb{R}})$. По теореме 6 множество $E = f(\langle a, b\rangle)$ выпукло, а по лемме 1 E — промежуток, то есть $f(\langle a, b\rangle) = \langle m, M\rangle$. $\square$

Замечание 1. Утверждение теоремы 7 формулируют короче: *непрерывный образ промежутка — промежуток.*

Замечание 2. Промежуток $\langle m, M\rangle$ может быть другого типа, нежели $\langle a, b\rangle$.

Так, функция синус отображает промежутки $\mathbb{R}$ и $[0, 2\pi)$ на отрезок $[-1, 1]$, а интервал $(0, \pi)$ — на полуинтервал $(0, 1]$.

Следствие 1. О сохранении отрезка. *Непрерывный образ отрезка — отрезок.*

Доказательство. Действительно, множество $f([a, b])$ — промежуток по теореме 7, а по теореме Вейерштрасса оно имеет наибольший и наименьший элемент. $\square$

Замечание 3. Наибольшее и наименьшее значения не обязательно достигаются на концах отрезка.

Теорема о сохранении промежутка, вообще говоря, не допускает обращения. Так, множество значений разрывной функции

$$f(x) = \begin{cases} x, & x \in [0, 1], \\ 0, & x \in (1, 2], \end{cases}$$

есть отрезок $[0, 1]$. Однако, для монотонной функции обратное утверждение верно.

Теорема 8. О разрывах и непрерывности монотонной функции. *Пусть $f\colon \langle a, b\rangle \to \mathbb{R}$, f монотонна. Тогда справедливы следующие утверждения.*

1. *f не может иметь разрывов второго рода.*

2. *Непрерывность f равносильна тому, что ее множество значений — промежуток.*

Доказательство. Пусть для определенности f возрастает.

1. Пусть $x_0 \in (a, b\rangle$, $x_1 \in \langle a, x_0)$. Тогда $f(x_1) \leqslant f(x) \leqslant f(x_0)$ для всех $x \in (x_1, x_0)$, поэтому f возрастает и ограничена сверху на $\langle a, x_0)$. По теореме о пределе монотонной функции существует конечный предел $f(x_0-)$, причем по теореме о предельном переходе в неравенстве $f(x_1) \leqslant f(x_0-) \leqslant f(x_0)$. Аналогично доказывается, что для любой точки $x_0 \in \langle a, b)$ существует конечный предел $f(x_0+)$, причем $f(x_0) \leqslant f(x_0+) \leqslant f(x_2)$ для всех $x_2 \in (x_0, b\rangle$.

2. Ввиду следствия о сохранении промежутка остается доказать достаточность. Пусть $f(\langle a,b\rangle)$ — промежуток. Докажем непрерывность f слева в любой точке $x_0 \in (a,b\rangle$ от противного. Пусть $f(x_0-) < f(x_0)$ (существование конечного левостороннего предела уже доказано). Возьмем $y \in (f(x_0-), f(x_0))$. Тогда если $a < x_1 < x_0$, то $y \in [f(x_1), f(x_0)]$. Следовательно, $y \in f(\langle a,b\rangle)$, то есть y — значение функции. С другой стороны, для всех $x \in \langle a, x_0)$ будет $f(x) \leqslant f(x_0-) < y$, а для всех $x \in [x_0, b\rangle$ будет $f(x) \geqslant f(x_0) > y$, то есть функция не принимает значение y. Полученное противоречие доказывает, что $f(x_0-) = f(x_0)$. Аналогично получается непрерывность f справа в любой точке $x_0 \in \langle a,b)$. $\square$

Теорема 9. О существовании и непрерывности обратной функции. *Пусть $f \in C\langle a,b\rangle$, f строго монотонна,*

$$m = \inf_{x\in\langle a,b\rangle} f(x), \qquad M = \sup_{x\in\langle a,b\rangle} f(x).$$

Тогда справедливы следующие утверждения.

1. *f обратима, $f^{-1}\colon \langle m,M\rangle \to \langle a,b\rangle$ — биекция.*
2. *f^{-1} строго монотонна одноименно с f.*
3. *f^{-1} непрерывна на $\langle m,M\rangle$.*

Доказательство. Пусть для определенности f строго возрастает.

1. Если $x_1, x_2 \in \langle a,b\rangle$, $x_1 < x_2$, то $f(x_1) < f(x_2)$; следовательно, f обратима. По следствию о сохранении промежутка $f(\langle a,b\rangle) = \langle m,M\rangle$. По общим свойствам обратного отображения f^{-1} — биекция $\langle m,M\rangle$ и $\langle a,b\rangle$.

2. Докажем, что f^{-1} строго возрастает. Если $y_1, y_2 \in \langle m,M\rangle$, $y_1 < y_2$, то $y_1 = f(x_1)$, $y_2 = f(x_2)$, где $x_1, x_2 \in \langle a,b\rangle$, $x_1 = f^{-1}(y_1)$, $x_2 = f^{-1}(y_2)$. При этом $x_1 < x_2$, так как возможность $x_1 \geqslant x_2$ исключена в силу строгого возрастания f.

3. Возрастающая функция f^{-1} задана на промежутке $\langle m,M\rangle$, а ее множество значений — промежуток $\langle a,b\rangle$. По теореме 8 она непрерывна. $\square$

Замечание 1. Для обратимости строго монотонной функции и строгой монотонности обратной функции непрерывность не нужна.

Замечание 2. Сформулируем еще несколько фактов без доказательства.

1. Множество точек разрыва монотонной функции не более чем счетно.

2. Если функция задана на промежутке, непрерывна и обратима, то она строго монотонна и, следовательно, обратная функция непрерывна.

3. Существует обратимая функция $f\colon \mathbb{R} \to \mathbb{R}$, непрерывная в точке 0, но такая, что f^{-1} разрывна в точке $f(0)$.

§ 5. Элементарные функции

Основными элементарными называют следующие функции.

1. Постоянная: $x \mapsto c$, $c \in \mathbb{R}$.

2. Степенная: $x \mapsto x^\alpha$, $\alpha \in \mathbb{R}$.

3. Показательная: $x \mapsto a^x$, $a > 0$, $a \neq 1$.

4. Логарифм: $x \mapsto \log_a x$, $a > 0$, $a \neq 1$.

5–8. Тригонометрические: синус, косинус, тангенс, котангенс — sin, cos, tg, ctg.

9–12. Обратные тригонометрические: арксинус, арккосинус, арктангенс, арккотангенс — arcsin, arccos, arctg, arcctg.

Функции, которые получаются из основных элементарных функций с помощью конечного числа арифметических действий и операций композиции, называются *элементарными* (без добавления прилагательного “основные”).

В школе изучались, как правило, элементарные функции, но не только они. Функции, заданные разными формулами на разных частях области определения, как, например, sign, могут не быть элементарными. Список основных элементарных функций можно без ущерба сократить, так как, например, $\cos x = \sin(\frac{\pi}{2} - x)$, а $\arccos x = \frac{\pi}{2} - \arcsin x$. Мы не ставим цели минимизировать список основных элементарных функций.

Далее мы дадим четкие определения основных элементарных функций (что редко делают в школе) и исследуем их свойства.

Постоянная функция. Функция $x \mapsto c$, как уже отмечалось, непрерывна на $\mathbb{R}$.

Степенная функция. Мы определим степень x^α при различных x и α, последовательно усложняя вид α. Это позволит нам рассматривать две функции: степенную и показательную.

Степенную функцию с показателем α, которая x сопоставляет x^α, будем обозначать e_α: $e_\alpha(x) = x^\alpha$. Заранее отметим, что области определения степенных функций могут быть различны при различных показателях.

При $\alpha = 1$, очевидно, $e_1 = \mathrm{id}_{\mathbb{R}}$. Как уже отмечалось, e_1 непрерывна на $\mathbb{R}$.

При $\alpha = n \in \mathbb{N}$ по определению

$$x^n = \underbrace{x \cdot \ldots \cdot x}_{n \text{ раз}}, \qquad x \in \mathbb{R}.$$

Следовательно, функция e_n непрерывна на $\mathbb{R}$ как произведение непрерывных.

При $\alpha = -n$, где $n \in \mathbb{N}$, полагаем

$$x^{-n} = \frac{1}{x^n}, \qquad x \in \mathbb{R} \setminus \{0\}.$$

Следовательно, функция e_{-n} непрерывна на $\mathbb{R} \setminus \{0\}$ как частное непрерывных.

При $\alpha = 0$ по определению полагаем $x^0 = 1$ при всех $x \neq 0$; тогда можно в соответствии с общим соглашением доопределить по непрерывности $x^0 = 1$ и при $x = 0$.

Если $n \in \mathbb{N}$, n нечетно, то функция e_n строго возрастает на $\mathbb{R}$, $\sup\limits_{x\in\mathbb{R}} e_n(x) = +\infty$, $\inf\limits_{x\in\mathbb{R}} e_n(x) = -\infty$; по теореме о сохранении промежутка $e_n(\mathbb{R}) = \mathbb{R}$. Если $n \in \mathbb{N}$, n четно, то функция e_n строго возрастает на $\mathbb{R}_+$, $\sup\limits_{x\in\mathbb{R}_+} e_n(x) = +\infty$, $\min\limits_{x\in\mathbb{R}_+} e_n(x) = 0$; по теореме о сохранении промежутка $e_n(\mathbb{R}_+) = \mathbb{R}_+$. По теореме о существовании и непрерывности обратной функции существует и непрерывна функция

$$e_{1/n} = \begin{cases} e_n^{-1}, & n \text{ нечетно}, \\ \left(e_n\big|_{\mathbb{R}_+}\right)^{-1}, & n \text{ четно}. \end{cases}$$

Эту функцию также называют *корнем n-й степени,* так что записи $e_{1/n}(x)$, $x^{1/n}$ и $\sqrt[n]{x}$ означают одно и то же. Итак,

$$e_{1/n}\colon \mathbb{R} \xrightarrow{\text{на}} \mathbb{R}, \qquad n \text{ нечетно},$$
$$e_{1/n}\colon \mathbb{R}_+ \xrightarrow{\text{на}} \mathbb{R}_+, \quad n \text{ четно};$$

$e_{1/n}$ строго возрастает и непрерывна.

Теперь определим x^α при рациональном $\alpha = r$, то есть при $r = \frac{p}{q}$, где $p = \mathbb{Z}$, $q \in \mathbb{N}$, дробь $\frac{p}{q}$ несократима. Полагаем

$$x^r = (x^p)^{1/q},$$

для всех тех x, для которых правая часть имеет смысл. Другими словами, $e_r = e_{1/q} \circ e_p$. Таким образом, x^r определено в следующих случаях:

$$x > 0, \quad r \text{ любое},$$
$$x = 0, \quad r \geqslant 0,$$
$$x < 0, \quad q \text{ нечетно}.$$

Функция e_r непрерывна на своей области определения; она строго возрастает на $[0, +\infty)$ при $r > 0$, строго убывает на $(0, +\infty)$ при $r < 0$.

Графики степенных функций в первой четверти при различных α изображены на рисунке 14. На нем у каждого графика функции $y = x^\alpha$ подписано значение α.

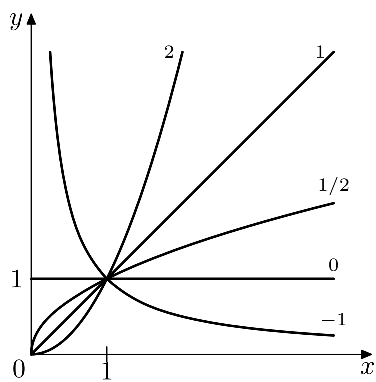

Рис. 14

Чтобы закончить определение степенной функции, нам потребуется степень с иррациональным показателем. Далее мы одновременно определим и ее, и показательную функцию.

Показательная функция. Положим $0^x = 0$ для всех положительных x.

Пусть $a > 0$. Мы хотим определить a^x для всех $x \in \mathbb{R}$. Пока что a^x определено для $x \in \mathbb{Q}$. Обозначим эту функцию $a^{(\cdot)}\big|_{\mathbb{Q}}$; при $a \neq 1$ назовем ее показательной функцией рационального аргумента. Перечислим без доказательства ее свойства, известные из школьного курса. Эти свойства легко проверяются по определению. В них $r, s \in \mathbb{Q}$.

1. Если $r < s$, то $a^r < a^s$ при $a > 1$ и $a^r > a^s$ при $0 < a < 1$.
2. $a^{r+s} = a^r a^s$.
3. $(a^r)^s = a^{rs}$.
4. $(ab)^r = a^r b^r$.

Определение. Пусть $a > 0$, $x \in \mathbb{R}$. Положим

$$a^x = \lim_{r \to x} a^r\big|_{\mathbb{Q}}.$$

При $a > 0$, $a \neq 1$ функция $\exp_a$, действующая по формуле

$$\exp_a x = a^x, \qquad x \in \mathbb{R},$$

называется *показательной функцией с основанием* a.

Чтобы определение имело смысл, необходимо доказать, что предел существует и что для рациональных x новое определение a^x совпадает с уже имеющимся.

Лемма 1. *Пусть* $a > 0$, $\{r_n\}$ — *последовательность рациональных чисел,* $r_n \to 0$. *Тогда* $a^{r_n} \to 1$.

Доказательство. При $a = 1$ лемма очевидна, так как $a^{r_n} = 1$ при всех n.

Пусть $a > 1$. Докажем лемму сперва в частном случае $r_n = \frac{1}{n}$. Поскольку $a^{1/n} > 1$, имеем $a^{1/n} = 1 + \alpha_n$, где $\alpha_n > 0$. Тогда по неравенству Бернулли

$$a = (1 + \alpha_n)^n \geqslant 1 + n\alpha_n,$$

откуда $0 < \alpha_n < \frac{a-1}{n}$. Значит, $\alpha_n \to 0$, что равносильно $a^{1/n} \to 1$.

Далее, по доказанному

$$a^{-1/n} = \frac{1}{a^{1/n}} \to \frac{1}{1} = 1.$$

Пусть теперь $\{r_n\}$ — произвольная последовательность из условия леммы. Возьмем $\varepsilon > 0$ и, пользуясь определением предела, подберем такой номер N_0, что

$$1 - \varepsilon < a^{-1/N_0} < a^{1/N_0} < 1 + \varepsilon. \tag{4}$$

Поскольку $r_n \to 0$, найдется такой номер N, что для всех $n > N$ будет $-\frac{1}{N_0} < r_n < \frac{1}{N_0}$. В силу строгой монотонности показательной функции рационального аргумента

$$1 - \varepsilon < a^{-1/N_0} < a^{r_n} < a^{1/N_0} < 1 + \varepsilon$$

для всех $n > N$. Это и означает, что $a^{r_n} \to 1$.

Если $0 < a < 1$, то $\frac{1}{a} > 1$, и по доказанному

$$a^{r_n} = \frac{1}{(1/a)^{r_n}} \to \frac{1}{1} = 1. \quad \square$$

Лемма 2. *Пусть $a > 0$, $x \in \mathbb{R}$, $\{r_n\}$ — последовательность рациональных чисел, $r_n \to x$. Тогда существует конечный предел последовательности $\{a^{r_n}\}$.*

Доказательство. При $a = 1$ лемма очевидна. Пусть $a > 1$.

Возьмем какую-либо возрастающую последовательность $\{s_n\}$ рациональных чисел, стремящуюся к x. Например, можно взять последовательность десятичных приближений к x с недостатком: $s_n = \frac{[10^n x]}{10^n}$. Тогда $x - \frac{1}{10^n} < s_n \leqslant x$, поэтому $s_n \to x$. Докажем, что последовательность $\{s_n\}$ возрастает. Неравенство $s_n \leqslant s_{n+1}$ равносильно $10[A] \leqslant [10A]$, где $A = 10^n x$. Но $10[A]$ — целое число, не превосходящее $10A$, поэтому $10[A] \leqslant [10A]$.

Последовательность $\{a^{s_n}\}$ возрастает и ограничена сверху числом $a^{[x]+1}$. Следовательно, $\{a^{s_n}\}$ сходится к некоторому пределу L. Но тогда

$$a^{r_n} = a^{r_n - s_n} a^{s_n} \to L,$$

потому что $a^{r_n - s_n} \to 1$ по лемме 1.

Если $0 < a < 1$, то $\frac{1}{a} > 1$, и по доказанному $(\frac{1}{a})^{r_n} \to L$, причем $L > 0$. Тогда

$$a^{r_n} = \frac{1}{(1/a)^{r_n}} \to \frac{1}{L}. \quad \square$$

Из леммы 2 вытекает корректность определения a^x. В самом деле, согласно замечанию 6 к определению предела по Гейне, предел существует. Если же $x \in \mathbb{Q}$, то, полагая $r_n = x$ при всех n, получаем, что новое определение a^x совпадает со старым.

Отметим отдельно, что $1^x = 1$ при всех $x \in \mathbb{R}$.

Установим несколько свойств показательной функции.

E1. *Функция* $\exp_a$ *строго возрастает на* $\mathbb{R}$ *при* $a > 1$ *и строго убывает на* $\mathbb{R}$ *при* $0 < a < 1$.

Доказательство. Пусть $a > 1$, $x < y$. Докажем, что $a^x < a^y$. Возьмем рациональные числа $\bar{r}$ и $\bar{\bar{r}}$, такие что

$$x < \bar{r} < \bar{\bar{r}} < y,$$

и две последовательности рациональных чисел $\{\bar{r}_n\}$ и $\{\bar{\bar{r}}_n\}$, такие что

$$\bar{r}_n < x < y < \bar{\bar{r}}_n, \quad \bar{r}_n \to x, \; \bar{\bar{r}}_n \to y.$$

Тогда в силу строгой монотонности показательной функции рационального аргумента

$$a^{\bar{r}_n} < a^{\bar{r}} < a^{\bar{\bar{r}}} < a^{\bar{\bar{r}}_n}.$$

По теореме о предельном переходе в неравенстве

$$a^x \leqslant a^{\bar{r}} < a^{\bar{\bar{r}}} \leqslant a^y.$$

Случай $0 < a < 1$ разбирается аналогично или переходом к основанию $\frac{1}{a}$, как в лемме 2. $\square$

E2. $a^{x+y} = a^x a^y$. *В частности,* $a^{-x} = \dfrac{1}{a^x}$.

Для доказательства надо взять две последовательности рациональных чисел $\{\bar{r}_n\}$ и $\{\bar{\bar{r}}_n\}$, стремящиеся к x и y, и перейти к пределу в равенстве

$$a^{\bar{r}_n + \bar{\bar{r}}_n} = a^{\bar{r}_n} a^{\bar{\bar{r}}_n},$$

которое для рациональных показателей известно. □

Свойство E2 иногда называют основным свойством степени.

E3. *Показательная функция непрерывна на* $\mathbb{R}$.

Доказательство. Непрерывность показательной функции в нуле доказывается на языке последовательностей, как лемма 1. Пусть $a > 1$, $\{x_n\}$ — последовательность вещественных чисел, $x_n \to 0$. Возьмем $\varepsilon > 0$ и зафиксируем номер N_0, для которого выполняется неравенство (4). Тогда найдется такой номер N, что для всех $n > N$ будет $-\frac{1}{N_0} < x_n < \frac{1}{N_0}$. В силу строгой монотонности показательной функции (свойства E1)

$$1 - \varepsilon < a^{-1/N_0} < a^{x_n} < a^{1/N_0} < 1 + \varepsilon$$

для таких n. Это и означает, что $a^{x_n} \to 1$. Случай $0 < a < 1$ разбирается аналогично или переходом к основанию $\frac{1}{a}$.

Непрерывность в произвольной точке x_0 следует из доказанной непрерывности в нуле:

$$a^{x_0 + \Delta x} - a^{x_0} = a^{x_0}(a^{\Delta x} - 1) \to 0. \quad \square$$

E4. $(a^x)^y = a^{xy}$.

Доказательство. Возьмем две последовательности рациональных чисел $\{x_n\}$ и $\{y_m\}$: $x_n \underset{n\to\infty}{\longrightarrow} x$, $y_m \underset{m\to\infty}{\longrightarrow} y$. По известному свойству степени с рациональным показателем $(a^{x_n})^{y_m} = a^{x_n y_m}$. Зафиксируем m и устремим n к ∞. По определению показательной функции $a^{x_n y_m} \underset{n\to\infty}{\longrightarrow} a^{x y_m}$ и $a^{x_n} \underset{n\to\infty}{\longrightarrow} a^x$, а по непрерывности степенной функции с рациональным показателем $(a^{x_n})^{y_m} \underset{n\to\infty}{\longrightarrow} (a^x)^{y_m}$. Поэтому $(a^x)^{y_m} = a^{x y_m}$. Осталось устремить m к ∞ и воспользоваться непрерывностью показательной функции. □

E5. $(ab)^x = a^x b^x$.

Для доказательства надо сделать предельный переход в равенстве для степеней с рациональным показателем.

E6. $\exp_a\colon \mathbb{R} \xrightarrow{\text{на}} (0, +\infty)$.

Доказательство. Пусть $a > 1$. Функция $\exp_a$ строго возрастает, поэтому существуют пределы $\lim\limits_{x\to\pm\infty} a^x$. По неравенству Бернулли ($a = 1 + \alpha$, $\alpha > 0$)

$$a^n = (1+\alpha)^n \geqslant 1 + n\alpha \to +\infty, \qquad a^{-n} = \frac{1}{a^n} \to 0.$$

Значит, по свойству сохранения промежутка $\exp_a(\mathbb{R}) = \langle 0, +\infty)$. Кроме того, значение 0 не принимается в силу строгой монотонности: если $a^{x_0} = 0$, то $a^x < 0$ при $x < x_0$, чего быть не может.

Доказательство при $0 < a < 1$ проводится аналогично. $\square$

Логарифм. Мы доказали, что функция $\exp_a$ — биекция между $\mathbb{R}$ и $(0, +\infty)$.

Определение. Пусть $a > 0$, $a \neq 1$. Функция, обратная к $\exp_a$, называется *логарифмом по основанию* a и обозначается $\log_a$.

Из теоремы о существовании и непрерывности обратной функции следует, что

$$\log_a\colon (0, +\infty) \xrightarrow{\text{на}} \mathbb{R},$$

функция $\log_a$ непрерывна, строго возрастает при $a > 1$ и строго убывает при $0 < a < 1$;

$$\lim_{x\to+\infty} \log_a x = \begin{cases} +\infty, & a > 1, \\ -\infty, & 0 < a < 1, \end{cases} \qquad \lim_{x\to 0+} \log_a x = \begin{cases} -\infty, & a > 1, \\ +\infty, & 0 < a < 1. \end{cases}$$

Графики показательной функции и логарифма при $a > 1$ изображены на рисунке 15a, а при $0 < a < 1$ — на рисунке 15b.

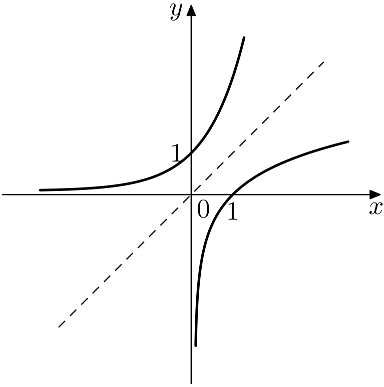

Рис. 15a

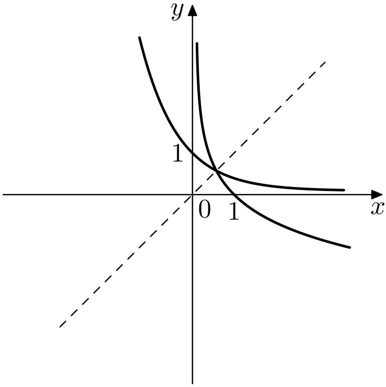

Рис. 15b

Пусть $a > 0$, $a \neq 1$. По определению обратной функции $\log_a x$ — это такое число y, что $a^y = x$. Другими словами, чтобы доказать равенство $\log_a x = y$, следует проверить, что $a^y = x$. Докажем этим приемом три свойства логарифма.

В формулировках свойств будет предполагаться, что $a, b > 0$, $a, b \neq 1$.

Л1. $\log_a(xy) = \log_a x + \log_a y$ *при всех* $x, y > 0$.

Доказательство. По основному свойству степени

$$a^{\log_a x + \log_a y} = a^{\log_a x} a^{\log_a y} = xy. \quad \square$$

Л2. $\log_a x^\alpha = \alpha \log_a x$ *при всех* $x > 0$, $\alpha \in \mathbb{R}$.
В частности, $\log_a \dfrac{1}{x} = -\log_a x$.

Доказательство. По свойству E4

$$a^{\alpha \log_a x} = (a^{\log_a x})^\alpha = x^\alpha. \quad \square$$

Л3. $\log_a x = \dfrac{\log_b x}{\log_b a}$ *при всех* $x > 0$.

В частности, $\log_a b = \dfrac{1}{\log_b a}$.

Доказательство. По свойству Е4

$$b^{\log_a x \log_b a} = \left(b^{\log_b a}\right)^{\log_a x} = a^{\log_a x} = x,$$

то есть $\log_b a \cdot \log_a x = \log_b x$. $\square$

Наиболее удобен в использовании, и это будет видно далее, логарифм по основанию e. Он называется *натуральным* логарифмом и обозначается символом ln. Показательная функция с основанием e называется еще *экспонентой* и обозначается exp. Часто используются логарифмы по основаниям 10 и 2 — *десятичные* и *двоичные* логарифмы, что связано с использованием десятичной и двоичной систем счисления. Десятичный логарифм обозначается символом lg. Перечисленные обозначения логарифмов используются чаще всего, но некоторые авторы предпочитают другие обозначения (например, log для натурального или даже lg для двоичного логарифма).

Свойство Л3 позволяет выражать логарифмы по любому основанию через логарифмы по одному конкретному основанию. Например, можно выразить все логарифмы через натуральные:

$$\log_a x = \frac{\ln x}{\ln a}.$$

Формула

$$\lg x = \lg e \cdot \ln x$$

позволяет свести вычисление десятичных логарифмов к вычислению натуральных. Здесь $\lg e$ — коэффициент перехода, который вычисляется один раз:

$$\lg e = 0,43429\ldots$$

Степенная функция (продолжение). При всех $x > 0$, $\alpha \in \mathbb{R}$ по свойству E4 верна формула

$$x^\alpha = e^{\alpha \ln x}.$$

Поэтому *степенная функция* e_α *непрерывна* на $(0, +\infty)$ при всех $\alpha \in \mathbb{R}$ (ранее это было установлено при рациональных α). Если α иррационально, то

$$e_\alpha \colon [\,0, +\infty) \xrightarrow{\text{на}} [\,0, +\infty), \quad \alpha > 0,$$

$$e_\alpha \colon (0, +\infty) \xrightarrow{\text{на}} (0, +\infty), \quad \alpha < 0.$$

Непрерывность e_α в нуле при $\alpha > 0$ также имеет место: если $x_n > 0$, $x_n \to 0$, то $y_n = \ln x_n \to -\infty$ и $e_\alpha(x_n) = e^{\alpha y_n} \to 0 = e_\alpha(0)$.

Замечание 1. Обозначения $\log_a$ для логарифма и exp для экспоненты являются общепринятыми, в отличие от обозначений e_α и $\exp_a$ для степенной и показательной функции. Мы ввели последние два обозначения, чтобы различать функции и их значения в точке x, обозначаемые x^α и a^x. Можно было обойтись и без новых символов, используя запись $(\cdot)^\alpha$ и $a^{(\cdot)}$, но эти обозначения имеют свои неудобства.

Замечание 2. Существуют разные соглашения, касающиеся определения степени. Некоторые авторы считают, что степенная функция $e_\alpha(x) = x^\alpha$ определена только при $x > 0$ (что не мешает им на соседней странице использовать запись $(-1)^n$); другие делают оговорку для целых α. Третьи различают, например, символы $\sqrt[3]{x}$ и $x^{1/3}$ и считают, что первый определен для всех x, а второй — только для положительных. Мы считаем степень определенной на самом широком множестве, на котором ее можно разумно определить (ограничиваясь вещественными числами), и не видим никаких математических оснований сужать область определения степени.

Синус и косинус. Мы будем пользоваться школьным определением косинуса и синуса как абсциссы и ординаты точки единичной окружности, а также всеми тригонометрическими формулами, выведенными на его основе. Полнота этого определения зависит от того, насколько строго определено соответствие между вещественными числами (точками числовой прямой) и точками единичной окружности (“углами”, “поворотами” и т.п.). Обратив внимание на имеющийся в школьном определении пробел, мы сейчас только скажем, что есть несколько принципиальных возможностей его ликвидировать (не опираясь, разумеется, на следствия этого определения, чтобы не попасть в порочный круг). Эти возможности обсуждаются при изучении интегралов и рядов.

Лемма 3. *Если* $0 < x < \frac{\pi}{2}$, *то*

$$\sin x < x < \operatorname{tg} x.$$

Доказательство. Изобразим единичную окружность и угол в x радиан (рисунок 16).

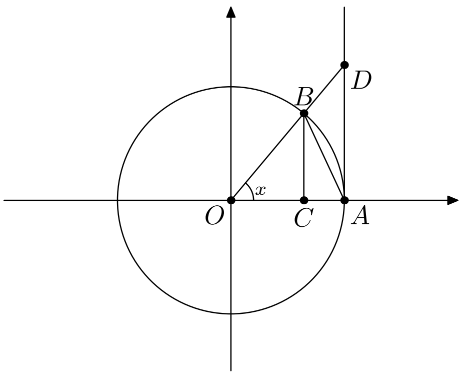

Рис. 16

На рисунке

$$\triangle OAB \subset \text{сект.}OAB \subset \triangle OAD.$$

Поэтому площади фигур связаны неравенством

$$S_{\triangle OAB} < S_{\text{сект.}OAB} < S_{\triangle OAD}.$$

Учитывая, что

$$S_{\triangle OAB} = \frac{1}{2}|OA||BC|,\ S_{\text{сект.}OAB} = \frac{1}{2}|OA|^2 x,\ S_{\triangle OAB} = \frac{1}{2}|OA||AD|,$$
$$|OA| = 1, \quad |BC| = \sin x, \quad |AD| = \operatorname{tg} x,$$

получаем требуемое. $\square$

Следствие 1. *Неравенство $|\sin x| \leqslant |x|$ выполняется при всех $x \in \mathbb{R}$, причем равенство имеет место только при $x = 0$.*

Доказательство. При $x \in (0, \frac{\pi}{2})$ строгое неравенство доказано в лемме. Если $x \geqslant \frac{\pi}{2}$, то

$$|\sin x| \leqslant 1 < \frac{\pi}{2} \leqslant x,$$

и неравенство установлено при всех $x > 0$. Если же $x < 0$, то $-x > 0$, и по доказанному

$$|\sin x| = |\sin(-x)| < |-x| = |x|.$$

Следствие 2. *Функции синус и косинус непрерывны на $\mathbb{R}$.*

Доказательство. Для любой точки $x_0 \in \mathbb{R}$ имеем:

$$|\sin x - \sin x_0| = \left|2\sin\frac{x - x_0}{2}\cos\frac{x + x_0}{2}\right| \leqslant$$

$$\leqslant 2 \cdot \frac{|x - x_0|}{2} \cdot 1 = |x - x_0| \underset{x \to x_0}{\longrightarrow} 0. \quad \square$$

Непрерывность косинуса доказывается аналогично или с помощью формулы приведения $\cos x = \sin\left(\frac{\pi}{2} - x\right)$, уже доказанной непрерывности синуса и теоремы о непрерывности композиции.

Графики синуса и косинуса изображены на рисунках 17 и 18.

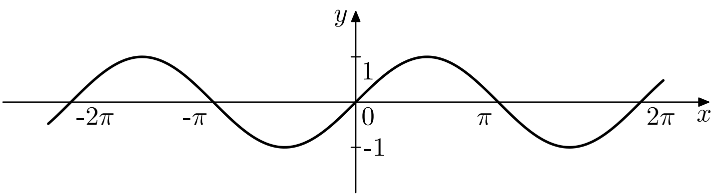

Рис. 17

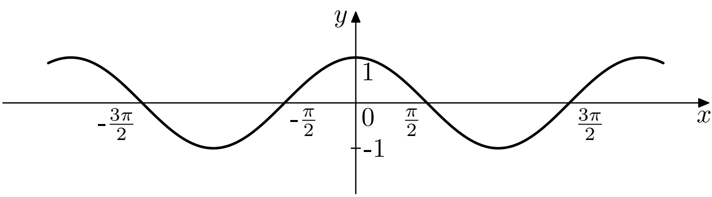

Рис. 18

Тангенс и котангенс. *Функции*

$$\operatorname{tg} x = \frac{\sin x}{\cos x}, \qquad x \in \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi : k \in \mathbb{Z} \right\},$$
$$\operatorname{ctg} x = \frac{\cos x}{\sin x}, \qquad x \in \mathbb{R} \setminus \left\{ k\pi : k \in \mathbb{Z} \right\}$$

непрерывны на своих областях определения по теореме о непрерывности частного.

Графики тангенса и котангенса изображены на рисунках 19 и 20.

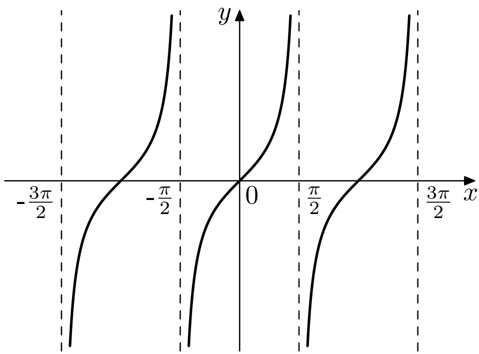

Рис. 19

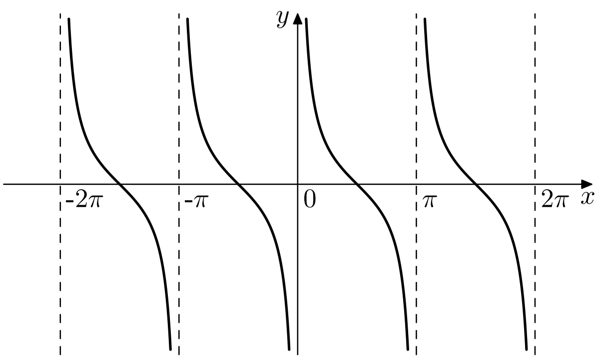

Рис. 20

Арксинус. Функция

$$\sin\colon \mathbb{R} \xrightarrow{\text{на}} [-1, 1]$$

не является обратимой, так как принимает свои значения более одного раза (даже бесконечно много раз). Но сужение синуса на отрезок $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$:

$$\sin\big|_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]}\colon \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \xrightarrow{\text{на}} [-1, 1]$$

строго возрастает, и потому обратимо.

Определение. Функция, обратная к сужению синуса на отрезок $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, называется *арксинусом:*

$$\arcsin = \left(\sin\big|_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]}\right)^{-1}.$$

По теореме о существовании и непрерывности обратной функции

$$\arcsin\colon [-1, 1] \xrightarrow{\text{на}} \left[-\frac{\pi}{2}, \frac{\pi}{2}\right],$$

функция арксинус строго возрастает и непрерывна. По определению обратной функции, если $x \in [-1, 1]$, то равенство $y = \arcsin x$ равносильно тому, что y $\in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ и $\sin y = x$.

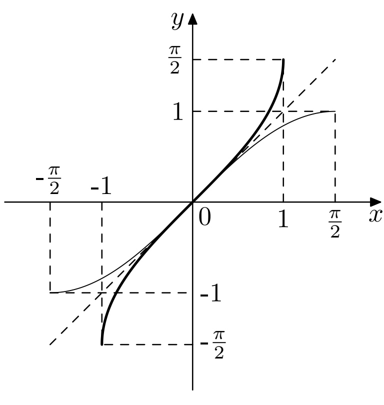

Рис. 21

Арккосинус. Функция

$$\cos\colon \mathbb{R} \xrightarrow{\text{на}} [-1,1]$$

не является обратимой. Но сужение косинуса на отрезок $[0,\pi]$:

$$\cos\big|_{[0,\pi]}\colon [0,\pi] \xrightarrow{\text{на}} [-1,1]$$

строго убывает, и потому обратимо.

Определение. Функция, обратная к сужению косинуса на отрезок $[0,\pi]$, называется *арккосинусом:*

$$\arccos = \left(\cos\big|_{[0,\pi]}\right)^{-1}.$$

По теореме о существовании и непрерывности обратной функции

$$\arccos\colon [-1,1] \xrightarrow{\text{на}} [0,\pi],$$

функция арккосинус строго убывает и непрерывна. По определению обратной функции, если $x \in [-1,1]$, то равенство $y = \arccos x$ равносильно тому, что y $\in [0,\pi]$ и $\cos y = x$.

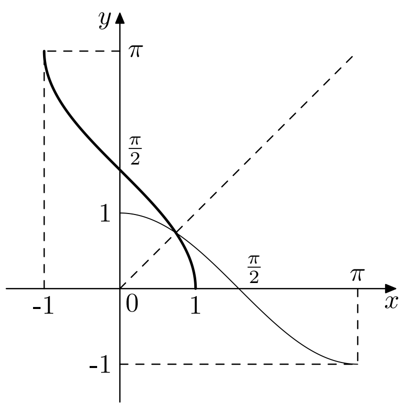

Рис. 22

Докажем тождество

$$\arcsin x + \arccos x = \frac{\pi}{2}, \qquad x \in [-1, 1].$$

Обозначим $y = \frac{\pi}{2} - \arccos x$ и докажем, что $y = \arcsin x$. Действительно, $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, так как $\arccos x \in [0, \pi]$. С другой стороны,

$$\sin y = \cos \arccos x = x. \quad \square$$

Арктангенс. Функция тангенс не является обратимой. Но сужение тангенса на интервал $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$:

$$\operatorname{tg}\big|_{\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)} \colon \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \xrightarrow{\text{на}} \mathbb{R}$$

строго возрастает, и потому обратимо.

Определение. Функция, обратная к сужению тангенса на интервал $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, называется *арктангенсом:*

$$\operatorname{arctg} = \left(\operatorname{tg}\big|_{\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)}\right)^{-1}.$$

По теореме о существовании и непрерывности обратной функции

$$\operatorname{arctg} \colon \mathbb{R} \xrightarrow{\text{на}} \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

функция арктангенс строго возрастает и непрерывна. По определению обратной функции, если $x \in \mathbb{R}$, то равенство $y = \operatorname{arctg} x$ равносильно тому, что $y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ и $\operatorname{tg} y = x$.

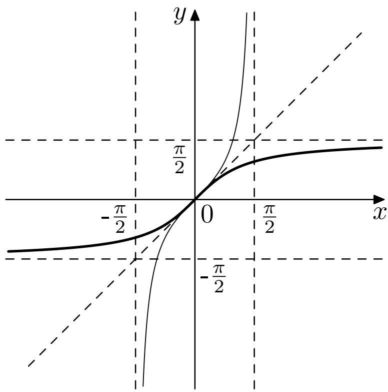

Рис. 23

Арккотангенс. Функция котангенс не является обратимой. Но сужение котангенса на интервал $(0, \pi)$:

$$\operatorname{ctg}\big|_{(0,\pi)} \colon (0, \pi) \xrightarrow{\text{на}} \mathbb{R}$$

строго убывает, и потому обратимо.

Определение. Функция, обратная к сужению котангенса на интервал $(0, \pi)$, называется *арккотангенсом:*

$$\operatorname{arcctg} = \left(\operatorname{ctg}\big|_{(0,\pi)}\right)^{-1}.$$

По теореме о существовании и непрерывности обратной функции

$$\operatorname{arcctg} \colon \mathbb{R} \xrightarrow{\text{на}} (0, \pi),$$

функция арккотангенс строго убывает и непрерывна. По определению обратной функции, если $x \in \mathbb{R}$, то равенство $y = \operatorname{arcctg} x$ равносильно тому, что у $\in (0, \pi)$ и $\operatorname{ctg} y = x$.

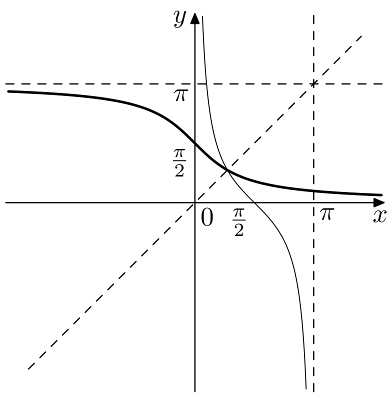

Рис. 24

Графики обратных тригонометрических функций изображены на рисунках 21–24.

Докажем тождество

$$\operatorname{arctg} x + \operatorname{arcctg} x = \frac{\pi}{2}, \qquad x \in \mathbb{R}.$$

Обозначим $y = \frac{\pi}{2} - \operatorname{arcctg} x$ и докажем, что $y = \operatorname{arctg} x$. Действительно, y $\in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, так как $\operatorname{arcctg} x \in (0, \pi)$. С другой стороны,

$$\operatorname{tg} y = \operatorname{ctg} \operatorname{arcctg} x = x. \quad \square$$

Итак, мы определили двенадцать основных элементарных функций и доказали их непрерывность. Ввиду того, что арифметические операции и композиция не выводят из класса непрерывных функций, верна следующая теорема.

Теорема 1. *Все элементарные функции непрерывны на своих областях определения.*

§ 6. Замечательные пределы и сравнение функций

Замечательными пределами называют пять равенств, часто использующихся при раскрытии неопределенностей.

З1. $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1.$

Доказательство. В силу леммы 3 § 5 при всех $x \in \left(0, \frac{\pi}{2}\right)$

$$\cos x < \frac{\sin x}{x} < 1. \tag{5}$$

Так как все три части неравенства (5) — четные функции, неравенство верно и при $x \in \left(-\frac{\pi}{2}, 0\right)$. При $x \to 0$ левая часть (5) стремится к 1 в силу непрерывности косинуса. По теореме о сжатой функции получаем требуемое. $\square$

Полученный результат можно сформулировать и так: если доопределить функцию $\varphi(x) = \frac{\sin x}{x}$ в нуле единицей, то получившаяся функция будет непрерывна на $\mathbb{R}$. Здесь и далее удобно пользоваться этим соглашением и доопределять функции по непрерывности.

Следствие 1.

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\frac{1}{2},\qquad \lim_{x\to 0}\frac{\operatorname{tg} x}{x}=1,$$
$$\lim_{x\to 0}\frac{\arcsin x}{x}=1,\qquad \lim_{x\to 0}\frac{\operatorname{arctg} x}{x}=1.$$

Доказательство. Пользуясь замечательным пределом для синуса, а также непрерывностью косинуса и теоремой об арифметических действиях над функциями, имеющими предел, находим:

$$\frac{1-\cos x}{x^2}=\frac{1}{2}\left(\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2\xrightarrow[x\to 0]{}\frac{1}{2},$$
$$\frac{\operatorname{tg} x}{x}=\frac{1}{\cos x}\,\frac{\sin x}{x}\xrightarrow[x\to 0]{}1.$$

Для вычисления третьего предела сделаем замену $x=\sin y$:

$$\lim_{x\to 0}\frac{\arcsin x}{x}=\lim_{y\to 0}\frac{y}{\sin y}=1.$$

Замену можно обосновать, например, так. Функция $f(x)=\arcsin x$ непрерывна в точке 0, $f(0)=0$, а функция $g(y)=\frac{y}{\sin y}$ непрерывна в точке 0, $g(0)=1$. По теореме о непрерывности композиции функция $g(f(x))=\frac{\arcsin x}{x}$ непрерывна в точке 0, и $g(f(0))=1$.

Последний предел вычисляется аналогично. □

При обосновании замены переменной вместо теоремы о непрерывности композиции можно было воспользоваться теоремой о пределе композиции или языком последовательностей, как будет сделано далее, при вычислении пределов 4 и 5.

32. $\displaystyle\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e.$

Доказательство. Напомним, что число e определялось как предел последовательности:

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

Разница между этим и доказываемым равенствами в том, что теперь речь идет о пределе не последовательности, а функции $f(x) = \left(1+\frac{1}{x}\right)^x$, заданной на $\mathbb{R}\setminus[-1,0]$: аргумент x не обязан принимать натуральные и даже положительные значения.

Для доказательства воспользуемся языком последовательностей. Возьмем последовательность $\{x_n\}$, стремящуюся к бесконечности, и докажем, что

$$f(x_n) \to e. \tag{6}$$

1. Пусть сначала $x_n \in \mathbb{N}$ для всех n. Возьмем $\varepsilon > 0$ и по определению числа e подберем такой номер K, что для всех номеров (то есть натуральных чисел) $k > K$ будет $|f(k) - e| < \varepsilon$. Но, начиная с некоторого номера, $x_n > K$, а тогда $|f(x_n) - e| < \varepsilon$, что и означает выполнение (6).

2. Пусть $x_n \to +\infty$. Тогда, начиная с некоторого номера, $x_n \geqslant 1$, поэтому, не уменьшая общности, можно считать, что $x_n \geqslant 1$ при всех n. Уменьшая или увеличивая основание и показатель степени, получим неравенства

$$\left(1+\frac{1}{[x_n]+1}\right)^{[x_n]} \leqslant \left(1+\frac{1}{x_n}\right)^{x_n} \leqslant \left(1+\frac{1}{[x_n]}\right)^{[x_n]+1},$$

которые перепишем в виде

$$\frac{f\big([x_n]+1\big)}{1+\frac{1}{[x_n]+1}} \leqslant f(x_n) \leqslant \left(1+\frac{1}{[x_n]}\right) f\big([x_n]\big). \tag{7}$$

Так как $\big\{[x_n]\big\}$ и $\big\{[x_n]+1\big\}$ — последовательности натуральных чисел, стремящиеся к $+\infty$, то по доказанному $f\big([x_n]\big) \to e$ и $f\big([x_n]+1\big) \to e$. Следовательно, крайние части в (7) стремятся к e, а тогда по теореме о сжатой последовательности и $f(x_n)$ стремится к e.

3. Пусть $x_n \to -\infty$; тогда $y_n = -x_n \to +\infty$ и $y_n - 1 \to +\infty$. По доказанному

$$f(x_n) = \left(1+\frac{1}{-y_n}\right)^{-y_n} = \left(\frac{y_n}{y_n-1}\right)^{y_n} = \left(1+\frac{1}{y_n-1}\right) f(y_n-1) \to e.$$

4. Пусть, наконец, $\{x_n\}$ — произвольная бесконечно большая последовательность, причем $x_n \notin [-1,0\,]$ при всех $n \in \mathbb{N}$. Если число

отрицательных (положительных) членов последовательности $\{x_n\}$ конечно, то $x_n \to +\infty$ $(-\infty)$, и соотношение $f(x_n) \to e$ уже доказано. Если же в последовательности бесконечно много и положительных, и отрицательных членов, то разобьем натуральный ряд на две подпоследовательности $\{n_k\}$ и $\{m_l\}$: $x_{n_k} > 0$, $x_{m_l} < -1$. По доказанному $f(x_{n_k}) \to e$ и $f(x_{m_l}) \to e$. По лемме 3 § 2 о подпоследовательностях $f(x_n) \to e$. $\square$

Замечание 1. Заменяя x на $\frac{1}{x}$, второй замечательный предел можно записать и так:

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

33. $\displaystyle\lim_{x\to 0}\frac{\log_a(1+x)}{x} = \frac{1}{\ln a}, \quad a > 0,\ a \neq 1.$

В частности, $\displaystyle\lim_{x\to 0}\frac{\ln(1+x)}{x} = 1.$

Доказательство. Так как $\log_a(1+x) = \frac{\ln(1+x)}{\ln a}$, достаточно доказать равенство для натурального логарифма. Имеем

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

Во втором равенстве мы воспользовались непрерывностью логарифма в точке e и теоремой о пределе композиции (или о непрерывности композиции, если доопределить $(1+x)^{1/x}$ значением e при $x = 0$). $\square$

34. $\displaystyle\lim_{x\to 0}\frac{(1+x)^\alpha - 1}{x} = \alpha, \quad \alpha \in \mathbb{R}.$

Доказательство. При $\alpha = 0$ доказываемое равенство тривиально. Пусть $\alpha \neq 0$. Возьмем последовательность $\{x_n\}$: $x_n \to 0$, $x_n \neq 0$. Не уменьшая общности, можно считать, что $|x_n| < 1$. Тогда в силу непрерывности и строгой монотонности степенной функции $y_n = (1+x_n)^\alpha - 1 \to 0$, $y_n \neq 0$. При этом

$$\alpha\ln(1+x_n) = \ln(1+y_n).$$

Пользуясь замечательным пределом для логарифма, находим

$$\frac{(1+x_n)^\alpha - 1}{x_n} = \frac{y_n}{x_n} = \frac{y_n}{\ln(1+y_n)}\,\alpha\,\frac{\ln(1+x_n)}{x_n} \to \alpha. \quad \square$$

35. $\lim\limits_{x\to 0} \dfrac{a^x - 1}{x} = \ln a, \ \ a > 0.$

В частности, $\lim\limits_{x\to 0} \dfrac{e^x - 1}{x} = 1.$

Доказательство. При $a = 1$ доказываемое равенство тривиально. Пусть $a \neq 1$. Возьмем последовательность $\{x_n\}$: $x_n \to 0$, $x_n \neq 0$. Тогда в силу непрерывности и строгой монотонности показательной функции $y_n = a^{x_n} - 1 \to 0$, $y_n \neq 0$. При этом

$$x_n \ln a = \ln(1 + y_n).$$

Пользуясь замечательным пределом для логарифма, находим

$$\frac{a^{x_n} - 1}{x_n} = \frac{y_n}{x_n} = \frac{y_n}{\ln(1 + y_n)} \ln a \to \ln a. \quad \square$$

Цель следующей серии определений — придать четкий смысл высказываниям типа “одна функция стремится к нулю (бесконечности) быстрее другой”, “две функции стремятся к нулю (бесконечности) с одинаковой скоростью” и т.п.

Определение. Пусть $f, g\colon D \subset \mathbb{R} \to \mathbb{R}$, x_0 — предельная точка D и существуют функция $\varphi\colon D \to \mathbb{R}$ и окрестность V_{x_0} точки x_0, такие что

$$f(x) = \varphi(x)g(x) \text{ для всех } x \in \dot{V}_{x_0} \cap D. \tag{8}$$

1. Если φ ограничена на $\dot{V}_{x_0} \cap D$, то говорят, что функция f *ограничена по сравнению* с g при $x \to x_0$, и пишут

$$f(x) = O(g(x)), \qquad x \to x_0. \tag{9}$$

2. Если $\varphi(x) \underset{x\to x_0}{\longrightarrow} 0$, то говорят, что функция f — *бесконечно малая по сравнению* с g при $x \to x_0$, и пишут

$$f(x) = o(g(x)), \qquad x \to x_0. \tag{10}$$

3. Если $\varphi(x) \underset{x\to x_0}{\longrightarrow} 1$, то говорят, что функции f и g *эквивалентны* или *асимптотически равны* при $x \to x_0$, и пишут

$$f(x) \sim g(x), \qquad x \to x_0. \tag{11}$$

Определение. Пусть $f, g\colon D \subset \mathbb{R} \to \mathbb{R}$. Если существует число $C > 0$, такое что $|f(x)| \leqslant C|g(x)|$ для всех $x \in D$, то говорят, что функция f *ограничена по сравнению* с g на множестве D, и пишут

$$f(x) = O(g(x)), \qquad x \in D. \tag{12}$$

Определение. Если $f(x) = O(g(x))$ и $g(x) = O(f(x))$ (при $x \to x_0$ или $x \in D$), то говорят, что функции f и g *сравнимы* (при $x \to x_0$ или $x \in D$ соответственно), и пишут $f \asymp g$.

Читают формулы (9), (10) и (12), соответственно, так: "$f(x)$ есть O-большое от $g(x)$ при x, стремящемся к x_0", "$f(x)$ есть o-малое от $g(x)$ при x, стремящемся к x_0", "$f(x)$ есть O-большое от $g(x)$ при x, принадлежащем D". O-символы были введены в математику Э. Ландау. Соотношения с O-символами и символами $\sim$ и $\asymp$ называют асимптотическими. Если ясно, о какой точке x_0 идет речь, то запись $x \to x_0$, а иногда и обозначение аргумента x опускают и пишут, например, $f = o(g)$.

Обсудим подробнее свойства асимптотических соотношений.

Замечание 1. 1. Соотношение (9) равносильно следующему: существуют число $C > 0$ и окрестность V_{x_0} точки x_0, такие что

$$|f(x)| \leqslant C|g(x)|, \qquad x \in \dot{V}_{x_0} \cap D. \tag{13}$$

2. Соотношение (10) равносильно следующему: для любого $\varepsilon > 0$ существует окрестность U_{x_0} точки x_0, такая что

$$|f(x)| \leqslant \varepsilon|g(x)|, \qquad x \in \dot{U}_{x_0} \cap D. \tag{14}$$

Доказательство. Эти утверждения сразу следуют из определений; тем не менее, приведем подробное доказательство.

1. Если выполнено соотношение (9), то существуют окрестность V_{x_0} и функция φ из определения. В силу ограниченности φ существует такая постоянная $C > 0$, что $|\varphi(x)| \leqslant C$ для всех $x \in \dot{V}_{x_0} \cap D$. Из равенства (8) следует (13).

Обратно, пусть существуют такие число $C > 0$ и окрестность V_{x_0}, что выполнено (13). Заметим, что для всех $x \in \dot{V}_{x_0} \cap D$, если $g(x) = 0$, то и $f(x) = 0$. Положим при $x \in D$

$$\varphi(x) = \begin{cases} \frac{f(x)}{g(x)}, & g(x) \neq 0, \\ 0, & g(x) = 0. \end{cases} \tag{15}$$

Тогда на $\dot{V}_{x_0} \cap D$ будет $f = \varphi g$ и $|\varphi| \leqslant C$, то есть выполнено (9).

2. Пусть выполнено (10), то есть существуют окрестность V_{x_0} и функция φ из определения. Возьмем $\varepsilon > 0$. Так как $\varphi(x) \underset{x \to x_0}{\longrightarrow} 0$, найдется такая окрестность U_{x_0}, содержащаяся в V_{x_0}, что $|\varphi| \leqslant \varepsilon$ на $\dot{U}_{x_0} \cap D$. Значит, верно соотношение (14).

Обратно, если V_{x_0} — окрестность, подобранная по числу $\varepsilon = 1$, а функция φ определяется равенством (15), то $f = \varphi g$ на $\dot{V}_{x_0} \cap D$, и по условию $\varphi(x) \underset{x \to x_0}{\longrightarrow} 0$. Значит, верно соотношение (10). $\square$

Замечание 2. Пусть существует такая окрестность U_{x_0} точки x_0, что g не обращается в нуль в $\dot{U}_{x_0} \cap D$. Тогда определение можно упростить.

1. Соотношение (9) равносильно следующему: существует такая окрестность V_{x_0} точки x_0, что функция $\frac{f}{g}$ ограничена в $\dot{V}_{x_0} \cap D$.

2. Соотношение (10) равносильно тому, что

$$\frac{f(x)}{g(x)} \underset{x \to x_0}{\longrightarrow} 0.$$

3. Соотношение (11) равносильно тому, что

$$\frac{f(x)}{g(x)} \underset{x \to x_0}{\longrightarrow} 1.$$

Для доказательства достаточно обозначить частное $\frac{f}{g}$ через φ.

Замечание 2 удобно при проверке асимптотических соотношений на практике, поскольку для элементарных функций дополнительное условие из замечания обычно выполняется.

Замечание 3. 1. *Асимптотическое равенство функций является отношением эквивалентности (в смысле § 4 введения).*

2. *Соотношения $f \sim g$, $f = g + o(g)$ и $f = g + o(f)$ равносильны.*

3. *Если $f = o(g)$, то $f = O(g)$.*

4. *Если $\alpha \neq 0$, $f \sim \alpha g$, то $f \asymp g$.*

Читатель легко докажет эти факты сам. Утверждения, обратные к третьему и четвертому, неверны (см. далее формулу (16)).

Замечание 4. Найденные замечательные пределы можно записать в виде асимптотических равенств: при $x \to 0$

$$\sin x \sim \operatorname{tg} x \sim \arcsin x \sim \operatorname{arctg} x \sim \ln(1+x) \sim x,$$
$$1 - \cos x \sim \frac{x^2}{2},$$
$$a^x - 1 \sim x \ln a, \quad (1+x)^\alpha - 1 \sim \alpha x$$

или (см. замечание 3)

$$\begin{aligned} &\sin x = x + o(x), && \operatorname{tg} x = x + o(x),\\ &\arcsin x = x + o(x), && \operatorname{arctg} x = x + o(x),\\ &\ln(1+x) = x + o(x), && \cos x = 1 - \frac{x^2}{2} + o(x^2),\\ &a^x = 1 + x\ln a + o(x), && (1+x)^\alpha = 1 + \alpha x + o(x). \end{aligned}$$

Вот еще несколько примеров использования новых символов:

$$\begin{aligned} &x = o(x^2), \quad x \to \infty, && x^2 = o(x), \quad x \to 0,\\ &\sin x = O(x), \quad x \in \mathbb{R}, && \sin x = O(x), \quad x \to 0,\\ &x = O(\sin x), \quad x \to 0, && x \neq O(\sin x), \quad x \in \mathbb{R},\\ &x \asymp x\left(2 + \sin\frac{1}{x}\right), \quad x \to 0. \end{aligned} \tag{16}$$

Мы привыкли к тому, что всякое равенство может быть прочитано как слева направо, так и справа налево: если “a равно b”, то и “b равно a”. Равенства с O-символами этим свойством не обладают. Например, равенство $\sin x = O(x)$ ($x \in \mathbb{R}$) верно, а равенство $O(x) = \sin x$ ($x \in \mathbb{R}$) не допускает однозначного истолкования: не всякая функция, являющаяся $O(x)$, есть $\sin x$. Кроме того, из $a = c$, $b = c$ следует, что $a = b$, но из равенств $\sin x = O(x)$, $x = O(x)$ нелепо выводить, что $\sin x = x$.

Этот кажущийся парадокс вызван только выбором обозначений. На самом деле, $O(g)$ и $o(g)$ — это множества функций, удовлетворяющих определению, а равенства $f = O(g)$ и $f = o(g)$ означают включения $f \in O(g)$ и $f \in o(g)$. Так, $O(1)$ означает класс ограниченных функций (на множестве D или $\dot{V}_{x_0} \cap D$), а $o(1)$ — класс бесконечно малых в точке x_0 функций. Равенства вида $o(g) = O(g)$ означают включение левой части в правую. Однако, по традиции в соотношениях с O-символами используют знак равенства, что бывает удобно, так как, например, позволяет переносить O-члены из одной части равенства в другую.

Для асимптотических соотношений справедливы формулы:

$$o(g) + o(g) = o(g), \quad o(g) - o(g) = o(g) \text{ (а не 0)},$$
$$2O(g) = O(g), \qquad O(g)O(h) = O(gh)$$

и т.п. Например, второе равенство следует понимать так: если $f_1 = o(g)$ и $f_2 = o(g)$, то и $f_1 - f_2 = o(g)$ (ясно, что функции f_1 и f_2 не обязаны уничтожаться). Читатель сам при необходимости разберется с подобными равенствами.

Следующая теорема описывает схему применения асимптотических равенств для вычисления пределов.

Теорема 1. Замена на эквивалентную при вычислении пределов. *Пусть* $f, \widetilde{f}, g, \widetilde{g}\colon D \subset \mathbb{R} \to \mathbb{R}$, x_0 *— предельная точка* D,

$$f(x) \sim \widetilde{f}(x),\ g(x) \sim \widetilde{g}(x), \qquad x \to x_0.$$

Тогда справедливы следующие утверждения.

1. $\lim\limits_{x\to x_0} f(x)g(x) = \lim\limits_{x\to x_0} \widetilde{f}(x)\widetilde{g}(x)$.

2. *Если* x_0 *— предельная точка области определения* $\frac{f}{g}$, *то* $\lim\limits_{x\to x_0} \frac{f(x)}{g(x)} = \lim\limits_{x\to x_0} \frac{\widetilde{f}(x)}{\widetilde{g}(x)}$.

(В обоих утверждениях пределы одновременно существуют или нет и, если существуют, то равны.)

Замечание 1. Если $g(x) \not\equiv 0$ в $\dot{V}_{x_0} \cap D$, то и $\widetilde{g}(x) \not\equiv 0$ в $\dot{\widetilde{V}}_{x_0} \cap D$, и обратно. Поэтому точка x_0 одновременно является или не является предельной для областей определения $\frac{f}{g}$ и $\frac{\widetilde{f}}{\widetilde{g}}$.

Доказательство. По определению эквивалентных функций существуют окрестности U_{x_0}, V_{x_0} и функции φ, ψ, стремящиеся к 1 при $x \to x_0$, такие что

$$f = \varphi\widetilde{f} \quad \text{на } \dot{U}_{x_0} \cap D, \qquad\qquad g = \psi\widetilde{g} \quad \text{на } \dot{V}_{x_0} \cap D.$$

Тогда на множестве $\dot{W}_{x_0} \cap D$, где $W_{x_0} = U_{x_0} \cap V_{x_0}$, верны оба равенства. Значит, на $\dot{W}_{x_0} \cap D$

$$fg = (\varphi\psi)(\widetilde{f}\widetilde{g}).$$

Следовательно, если $\lim\limits_{x \to x_0} \widetilde{f}(x)\widetilde{g}(x)$ существует и равен A, то по теореме о пределе произведения $\lim\limits_{x \to x_0} f(x)g(x)$ существует и равен A; верно и обратное. Доказательство для предела частного проводится аналогично (с той разницей, что может понадобиться еще сузить окрестность, чтобы φ и ψ не обращались в ней в нуль). $\square$

Замечание 2. Из условий теоремы не следует, что

$$\lim_{x \to x_0} (f(x) \pm g(x)) = \lim_{x \to x_0} (\widetilde{f}(x) \pm \widetilde{g}(x)).$$

Например, если $x_0 = +\infty$, $f(x) = x + 1$, $\widetilde{f}(x) = g(x) = \widetilde{g}(x) = x$, то

$$\lim_{x \to x_0} (f(x) - g(x)) = 1, \qquad \lim_{x \to x_0} (\widetilde{f}(x) - \widetilde{g}(x)) = 0.$$

Другими словами, при отыскании предела множители в произведении и частном заменять на эквивалентные можно, а слагаемые в сумме и разности, вообще говоря, нельзя.

Пример. Найдем $\lim\limits_{x \to 0} \dfrac{\ln(1 + x + x^2) + \arcsin 3x - 5x^3}{\sin 2x + \operatorname{tg}^2 x}$.

Учитывая, что при $x \to 0$

$$\ln(1 + x + x^2) = x + x^2 + o(x + x^2) = x + o(x),$$
$$\arcsin 3x = 3x + o(x), \qquad 5x^3 = o(x),$$
$$\sin 2x = 2x + o(x), \qquad \operatorname{tg}^2 x = o(x),$$

находим:

$$\lim_{x \to 0} \frac{\ln(1 + x + x^2) + \arcsin 3x - 5x^3}{\sin 2x + \operatorname{tg}^2 x} = \lim_{x \to 0} \frac{4x + o(x)}{2x + o(x)} = 2.$$

Замечание 3. Для нахождения пределов функций вида f^g $(f>0)$ бывает удобным преобразование

$$(f(x))^{g(x)} = e^{g(x)\ln f(x)}.$$

Пусть $\lim\limits_{x\to a} g(x)\ln f(x) = A$. Тогда по свойствам экспоненты

$$\lim_{x\to a}(f(x))^{g(x)} = \begin{cases} e^A, & A\in\mathbb{R}, \\ +\infty, & A=+\infty, \\ 0, & A=-\infty. \end{cases}$$

Этим приемом задача сводится к нахождению предела произведения.

Рассмотрим два асимптотических равенства:

$$\cos x \sim 1, \qquad x\to 0,$$
$$\cos x \sim 1-\frac{x^2}{2}, \qquad x\to 0,$$

очевидно, верных (предел отношения левой и правой частей равен 1). Однако, второе из них содержит в каком-то смысле больше информации о функции косинус. Дело в том, что погрешность второго равенства бесконечно мала по сравнению с погрешностью первого:

$$\lim_{x\to 0}\frac{1-x^2/2-\cos x}{1-\cos x} = 1-\lim_{x\to 0}\frac{x^2/2}{1-\cos x} = 0.$$

Поэтому разумно ввести следующее понятие.

Пусть $f\sim g$, $f\sim h$. Если $f-h=o(f-g)$, то говорят, что *асимптотическое равенство $f\sim h$ точнее,* чем $f\sim g$.

Пусть $f\colon D\subset\mathbb{R}\to\mathbb{R}$, x_0 — предельная точка D и задана конечная или счетная система функций $\{g_k\}_{k=0}^N$ $(N\in\mathbb{N})$ или $\{g_k\}_{k=0}^\infty$, $g_k\colon D\to\mathbb{R}$, каждая из которых бесконечно мала по сравнению с предыдущей: при всех $k\in[0:N-1]$ или $k\in\mathbb{Z}_+$

$$g_{k+1}(x) = o(g_k(x)), \qquad x\to x_0.$$

Большую роль в анализе играют асимптотические формулы вида

$$f(x) = \sum_{k=0}^{n} c_k g_k(x) + o(g_n(x)), \quad x \to x_0.$$

Эти равенства тем точнее, чем больше n. Особенно часто встречаются случаи, когда $g_k(x) = (x - x_0)^k$ ($x_0 \in \mathbb{R}$) или $g_k(x) = x^{-k}$ ($x_0 = \infty$).

Если $f(x) \sim c(x - x_0)^k$, где $c \neq 0$, то функция $c(x - x_0)^k$ называется *главной степенной частью* f при $x \to x_0$.

Не всякая функция f допускает асимптотическое разложение по заданной системе функций, но если такое асимптотическое разложение есть, то оно единственно.

Теорема 2. Единственность асимптотического разложения. *Пусть $D \subset \mathbb{R}$, x_0 — предельная точка D, $n \in \mathbb{Z}_+$, $f, g_k\colon D \to \mathbb{R}$ ($k \in [\,0 : n]$), при всех $k \in [\,0 : n-1]$*

$$g_{k+1}(x) = o(g_k(x)), \quad x \to x_0,$$

и для любой окрестности V_{x_0} существует точка $t \in \dot{V}_{x_0} \cap D$, в которой $g_n(t) \neq 0$. Тогда, если асимптотическое разложение функции f по системе $\{g_k\}$ существует, то оно единственно: из равенств

$$f(x) = \sum_{k=0}^{n} c_k g_k(x) + o(g_n(x)), \quad x \to x_0, \tag{17}$$

$$f(x) = \sum_{k=0}^{n} d_k g_k(x) + o(g_n(x)), \quad x \to x_0 \tag{18}$$

следует, что $c_k = d_k$ при всех $k \in [\,0 : n]$.

Доказательство. По индукции заключаем, что

$$g_k(x) = o(g_l(x)), \qquad x \to x_0,\ l < k.$$

Обозначим

$$E_k = \{x \in D : g_k(x) \neq 0\}, \qquad k \in [\,0 : n].$$

Если бы функция g_k тождественно обращалась в нуль на множестве вида $\dot{U}_{x_0} \cap D$, то и функция $g_n = \varphi_k g_k$, где φ_k — функция из определения символа o, обращалась бы тождественно в нуль на множестве $\dot{V}_{x_0} \cap D$, что противоречит условию. Следовательно, x_0 — предельная точка каждого E_k.

Допустим, что $c_k = d_k$ не при всех $k \in [\,0 : n]$. Положим

$$m = \min\{k \in [\,0 : n] : c_k \neq d_k\}.$$

Из разложений (17) и (18) следует, что

$$f(x) = \sum_{k=0}^{m} c_k g_k(x) + o(g_m(x)), \quad x \to x_0, \tag{19}$$

$$f(x) = \sum_{k=0}^{m} d_k g_k(x) + o(g_m(x)), \quad x \to x_0. \tag{20}$$

Вычтя (20) из (19), найдем

$$0 = (c_m - d_m) g_m(x) + o(g_m(x)), \qquad x \to x_0.$$

Поделив на $g_m(x)$ при $x \in E_m$ и перейдя к пределу по множеству E_m, получим $c_m = d_m$, что противоречит определению m. $\square$

Замечание 1. Случай, когда g_n — тождественный нуль на множестве вида $\dot{V}_{x_0} \cap D$, тривиален. Действительно, тогда все g_k при $k < n$ и f обладают тем же свойством, поэтому асимптотическое разложение имеет место с любыми коэффициентами c_k.

Один частный случай асимптотического разложения — известное из школьного курса понятие наклонной асимптоты функции (графика функции). Напомним определения асимптот.

Определение. Пусть $x_0 \in \mathbb{R}$, функция f задана по крайней мере на $\langle a, x_0)$ или $(x_0, b\rangle$ и действует в $\mathbb{R}$. Прямая $x = x_0$ называется *вертикальной асимптотой* функции f, если $f(x_0+)$ или $f(x_0-)$ равны $-\infty$ или $+\infty$.

Определение. Пусть $\langle a, +\infty) \subset D \subset \mathbb{R}$, $f\colon D \to \mathbb{R}$, $\alpha, \beta \in \mathbb{R}$. Прямая $y = \alpha x + \beta$ называется *наклонной асимптотой* функции f при $x \to +\infty$, если

$$f(x) = \alpha x + \beta + o(1), \qquad x \to +\infty. \tag{21}$$

Аналогично определяется наклонная асимптота при $x \to -\infty$ функции, заданной по крайней мере на $(-\infty, b\rangle$.

Горизонтальная асимптота — частный случай наклонной при $\alpha = 0$. Прямая $y = \beta$ — горизонтальная асимптота функции f при $x \to \pm\infty$ тогда и только тогда, когда $f(x) \underset{x\to\pm\infty}{\longrightarrow} \beta$.

Теорема 3. Уравнение наклонной асимптоты. *Пусть $\langle a, +\infty) \subset D \subset \mathbb{R}$, $f\colon D \to \mathbb{R}$, $\alpha, \beta \in \mathbb{R}$. Прямая $y = \alpha x + \beta$ — асимптота f при $x \to +\infty$ тогда и только тогда, когда*

$$\alpha = \lim_{x\to+\infty} \frac{f(x)}{x}, \qquad \beta = \lim_{x\to+\infty} (f(x) - \alpha x). \tag{22}$$

Доказательство. Если прямая $y = \alpha x + \beta$ — асимптота f, то

$$f(x) = \alpha x + \beta + \varphi(x), \qquad \varphi(x) \underset{x\to+\infty}{\longrightarrow} 0. \tag{23}$$

Поэтому

$$\frac{f(x)}{x} - \alpha = \frac{\beta + \varphi(x)}{x} \underset{x\to+\infty}{\longrightarrow} 0,$$
$$f(x) - \alpha x = \beta + \varphi(x) \underset{x\to+\infty}{\longrightarrow} \beta.$$

Обратно, если выполнены равенства (22), то, обозначив

$$\varphi(x) = f(x) - \alpha x - \beta,$$

мы получим (23), что равносильно (21). $\square$

Второй частный случай асимптотических разложений — равенства вида

$$f(x) = f(x_0) + A(x - x_0) + o(x - x_0), \qquad x \to x_0,$$

изучение которых составляет дифференциальное исчисление.

Пример. Пусть $f(x) = \sqrt{1+x}$, $x \to 0$. Ясно, что

$$f(x) = 1 + o(1).$$

Вычислив замечательный предел, мы уточнили это равенство:

$$f(x) = 1 + \frac{1}{2}x + o(x).$$

В свою очередь,

$$f(x) - 1 - \frac{x}{2} = \frac{1 + x - (1 + x/2)^2}{\sqrt{1+x} + 1 + x/2} = \frac{-x^2/4}{\sqrt{1+x} + 1 + x/2} \sim -\frac{x^2}{8},$$

откуда

$$f(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + o(x^2).$$

Читателю предлагается уточнить и это равенство и доказать, что

$$f(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + o(x^3).$$

Общий способ получения асимптотических равенств — формула Тейлора — будет рассмотрен в следующей главе.

ГЛАВА 3. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ВЕЩЕСТВЕННОЙ ПЕРЕМЕННОЙ

§ 1. Введение

Основная идея дифференциального исчисления заключается в том, что в определенных ситуациях исследование поведения функции можно свести к изучению многочлена, который “хорошо приближает” эту функцию. Не уточняя пока смысл сказанного, отметим преимущества такого подхода.

1) Нахождение значений многочленов требует лишь операций сложения и умножения. Поэтому, аппроксимируя функцию многочленами, мы получаем удобный способ приближенного вычисления ее значений.

2) Проанализировать поведение многочленов первой или второй степени не составляет труда, поскольку их графики легко строятся. Если такие многочлены достаточно хорошо приближают функцию в окрестности какой-либо точки, то мы сможем получить информацию о некоторых локальных свойствах этой функции.

3) В § 6 главы 2 было введено понятие степенной главной части бесконечно малой функции. Главную часть многочлена вблизи его корня можно выделить алгебраически, разложив многочлен на множители. Для функций более общего вида такой способ не годится. Однако, если мы знаем “достаточно хорошее” локальное приближение функции многочленом, то главные части функции и многочлена будут одинаковы. Эти соображения оказываются полезными для получения асимптотических разложений и раскрытия неопределенностей.

Формализуем поставленную задачу. Пусть $E \subset \mathbb{R}$, $a \in E$ — предельная точка множества E, $f\colon E \to \mathbb{R}$, $n \in \mathbb{Z}_+$. Необходимо построить многочлен $p(x) = \sum\limits_{k=0}^{n} c_k (x-a)^k$ степени не выше n, для которого

$$p(a) = f(a) \quad \text{и} \quad f(x) = p(x) + o\big((x-a)^n\big) \quad (x \to a). \qquad (*)$$

Это значит, что погрешность приближения f многочленом p стремится к нулю быстрее старшей степени p, когда $x \to a$.

Замечание. *Если многочлен p, удовлетворяющий условию $(*)$, существует, то он единственный.*

Доказательство. Пусть q — другой многочлен степени не выше n, удовлетворяющий условию $(*)$. Положим

$$r(x) = p(x) - q(x) = \sum_{k=0}^{n} r_k (x-a)^k.$$

Тогда $r(x) = o\big((x-a)^n\big)$ при $x \to a$. Допустим, что $r \not\equiv 0$, то есть не все коэффициенты r_k равны нулю. Тогда мы можем записать $r(x) = r_m(x-a)^m + \ldots + r_n(x-a)^n$, где $m \in \mathbb{Z}_+$ и $r_m \neq 0$. Отсюда

$$\sum_{k=m}^{n} r_k (x-a)^{k-m} = \frac{r(x)}{(x-a)^m} = o\big((x-a)^{n-m}\big) \quad (x \to a)\,,$$

что невозможно, так как при $x \to a$ правая часть стремится к нулю, а левая — к $r_m \neq 0$. $\square$

Следует отметить, что единственность многочлена p можно также получить из единственности асимптотического разложения (см. § 6 главы 2).

Определение 1. Многочлен p степени не выше n, удовлетворяющий условию $(*)$, называется *многочленом Тейлора функции f порядка n в точке a* и обозначается $T_{a,n}f$.

Из определения очевидно, что существование $T_{a,0}f$ равносильно непрерывности функции f в точке a. В случае $n \in \mathbb{N}$ вопрос о существовании многочлена Тейлора $T_{a,n}f$ более сложен. Вначале мы изучим его для $n = 1$. Это приведет нас к понятиям *дифференцируемости* и *производной*, которые будут подробно обсуждаться в трех следующих параграфах. Затем мы рассмотрим задачу в общем случае. Будет показано, что для широкого класса функций f многочлен Тейлора существует, а его коэффициенты выражаются через производные старших порядков функции f в точке a. Кроме того, мы укажем способ записи и оценки погрешности $f - T_{a,n}f$, который позволяет использовать многочлен Тейлора для приближенного вычисления значений функции f.

При изучении многочлена $T_{a,n}f$ мы будем накладывать на множество E и точку a следующее дополнительное условие:

$$\exists \delta > 0\colon\ (a-\delta, a+\delta) \cap E \text{ — невырожденный промежуток.} \qquad (+)$$

Это ограничение является более жестким, чем в определении $T_{a,n}f$. Если нас интересует поведение функции f вблизи точки a, то при таком соглашении можно считать множество E промежутком. Поэтому в дальнейшем мы будем, как правило, рассматривать функции, заданные на промежутке.

Производная тесно связана с такими геометрическими характеристиками графиков функций, как монотонность, экстремумы, выпуклость. Два последних параграфа главы посвящены рассмотрению этих свойств и их приложений.

§ 2. Дифференцируемость функции в точке

Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in \langle A, B\rangle$.

Определение 1. Дифференцируемость функции в точке. Функция f называется *дифференцируемой в точке* a, если существует такое число k, что

$$f(x) = f(a) + k(x-a) + o(x-a) \quad (x \to a).$$

Коэффициент k называется *производной* f *вточке* a и обозначается $f'(a)$ (используются также обозначения $\dfrac{df}{dx}(a)$ и $Df(a)$).

Фактически дифференцируемость f означает существование у нее многочлена Тейлора первого порядка в точке a. Отсюда вытекает единственность производной, поскольку $f'(a)$ является коэффициентом $T_{a,1}f$ при $x-a$.

Число $f(x)-f(a)$ называется *приращением функции* f *в точке* a, соответствующим приращению аргумента $h = x-a$, и обозначается $(\Delta_a f)(h)$ или просто $\Delta_a f(h)$. Тогда определение 1 можно записать в виде

$$\Delta_a f(h) = f'(a)\cdot h + o(h) \quad (h \to 0).$$

Отбрасывая в правой части этого соотношения $o(h)$, мы получим приближенное равенство $f(x)-f(a) \approx f'(a)(x-a)$. Если $f'(a) \neq 0$, то при $x \to a$ относительная погрешность такого приближения бесконечно мала, так как $\lim\limits_{x\to a} \dfrac{o(x-a)}{f'(a)(x-a)} = 0$. Таким образом, линейная функция $h \mapsto f'(a)\cdot h$ дает главную часть приращения $\Delta_a f(h)$.

Эта функция называется *дифференциалом* f *в точке* a и обозначается $d_a f(h)$.

Рассмотрим семейство функций $l_k(x) = f(a)+k(x-a)$, где k принимает произвольные вещественные значения. Их графики представляют собой прямые, проходящие через точку $\big(a, f(a)\big)$. Заметим, что

$$f(x) - l_k(x) = \big(f'(a) - k\big)(x - a) + o(x - a) \quad (x \to a).$$

Если $k = f'(a)$, то разность $f(x) - l_k(x)$ стремится к нулю быстрее $x - a$ при $x \to a$, а для других $k \in \mathbb{R}$ она имеет тот же порядок малости, что и $x - a$. Поэтому прямая

$$y = f(a) + f'(a)(x - a)$$

дает наименьшее отклонение от графика функции f вблизи точки a. Назовем эту прямую *касательной* к графику f в точке a. Ее уравнение можно записать в виде $y - f(a) = d_a f(x - a)$. Если через α обозначить угол наклона касательной к оси OX, то $f'(a) = \operatorname{tg}\alpha$. Геометрический смысл введенных понятий ясен из рисунка 25.

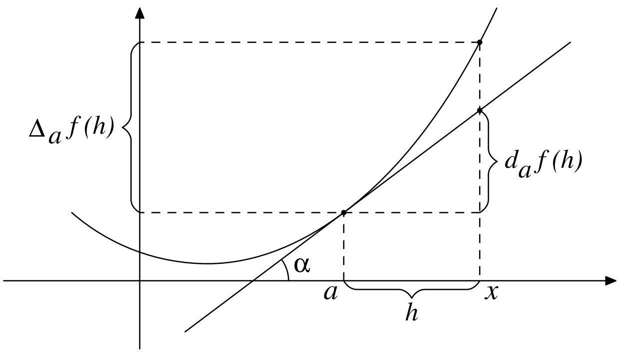

Рис. 25

Докажем теперь критерий дифференцируемости функции в точке.

Теорема 1. Критерий дифференцируемости. *Предположим, что* $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in \langle A, B\rangle$, $k \in \mathbb{R}$. *Тогда равносильны следующие утверждения.*

1) *f дифференцируема в точке a и $f'(a) = k$.*

2) *Предел $\lim\limits_{x\to a} \dfrac{f(x)-f(a)}{x-a}$ существует и равен k.*

3) *Существует функция $F\colon \langle A, B\rangle \to \mathbb{R}$, непрерывная в точке a, для которой $F(a) = k$ и $f(x) - f(a) = F(x)(x-a)$ при $x \in \langle A, B\rangle$.*

Замечания

1) Утверждение 2 удобно тем, что оно дает *способ вычисления* производной. Кроме того, представление производной как предела позволяет определить *бесконечные производные* и *односторонние производные*. Это будет сделано ниже.

2) Утверждение 3, дающее явное выражение приращения функции через приращение аргумента, будет использоваться в § 3 при выводе правил дифференцирования.

3) Характеристики дифференцируемости, сформулированные в утверждениях 2 и 3, нельзя перенести на случай функций нескольких переменных, тогда как определение 1 имеет многомерный аналог. Именно поэтому мы не использовали утверждение 2 в качестве *определения* дифференцируемости, хотя для практических целей оно удобнее.

Доказательство проведем по схеме 1) $\Rightarrow$ 2) $\Rightarrow$ 3 $\Rightarrow$ 1).

1) $\Rightarrow$ 2). Условие 1) означает, что $f(x) - f(a) = k(x-a) + o(x-a)$ при $x \to a$. Поэтому

$$\frac{f(x)-f(a)}{x-a} = k + \frac{o(x-a)}{x-a} \to k \quad (x \to a).$$

2) $\Rightarrow$ 3). Положим $F(x) = \dfrac{f(x)-f(a)}{x-a}$ при $x \neq a$ и $F(a) = k$. В силу 2) $\lim\limits_{x\to a} F(x) = F(a)$, то есть F непрерывна в точке a. Остальные свойства F очевидны.

3) $\Rightarrow$ 1). В силу 3)

$$f(x) = f(a) + F(x)(x-a) = f(a) + k(x-a) + (F(x)-k)(x-a).$$

Так как F непрерывна в точке a, то $F(x) \to F(a) = k \quad (x \to a)$. Поэтому $(F(x)-k)(x-a) = o(x-a)$ при $x \to a$, и мы получаем утверждение 1). $\square$

Дробь $\dfrac{f(x)-f(a)}{x-a}$ называется *разностным отношением.* Оно имеет следующий физический смысл. Если функция f описывает зависимость координаты движущейся материальной точки от времени, то разностное отношение дает среднюю скорость точки на отрезке времени между моментами a и x. Тогда $f'(a)$ можно понимать как *мгновенную* скорость точки в момент времени a.

Трактовка производной как предела разностного отношения позволяет ввести несколько важных понятий.

1) *Бесконечные производные.* Будем говорить, что $f'(a) = \pm\infty$, если $\lim\limits_{x\to a}\dfrac{f(x)-f(a)}{x-a} = \pm\infty$. Аналогично определяется равенство $f'(a) = \infty$. Отметим, что если функция f имеет *бесконечную* производную в точке a, то она *не дифференцируема* в этой точке.

2) *Односторонние производные.* Обозначим

$$f'_{\pm}(a) = \lim_{x\to a\pm}\frac{f(x)-f(a)}{x-a}.$$

Если $f'_-(a)$ и $f'_+(a)$ существуют, то они называются соответственно левой и правой производными f в точке a. Дифференцируемость f в точке $a \in (A, B)$ означает, что $f'_-(a)$ и $f'_+(a)$ конечны и равны между собой. Их общее значение в этом случае есть $f'(a)$. Для доказательства достаточно воспользоваться замечанием об односторонних пределах (см. § 3 главы 2).

3) *Касательная как предел секущих.* Пусть $b \in \langle A, B\rangle$. Рассмотрим прямую, проходящую через точки $\bigl(a, f(a)\bigr)$ и $\bigl(b, f(b)\bigr)$ (назовем ее *секущей*). Она задается уравнением

$$y = f(a) + k(b)(x-a), \text{ где } k(b) = \frac{f(b)-f(a)}{b-a}.$$

Если f дифференцируема в точке a, то $\lim\limits_{b\to a} k(b) = f'(a)$. Поэтому при $b \to a$ секущая, вращаясь вокруг точки $\bigl(a, f(a)\bigr)$, приближается к прямой $y = f(a) + f'(a)(x-a)$, то есть к касательной. В случае $f'(a) = \pm\infty$ угол наклона секущей стремится к $\frac{\pi}{2}$ соответственно справа или слева, а она сама — к вертикальному положению (см. рисунок 26а). Наконец, если $f'_-(a)$ и $f'_+(a)$ существуют и различны, то предельное положение секущей зависит от того, с какой стороны

b приближается к a (слева или справа). В этом случае график f имеет излом (см. рисунок 26b).

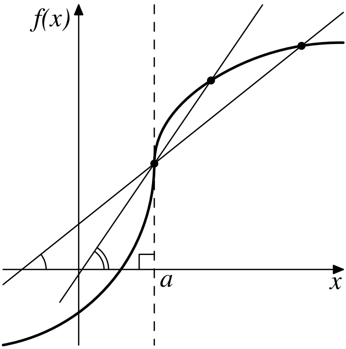

Рис. 26a

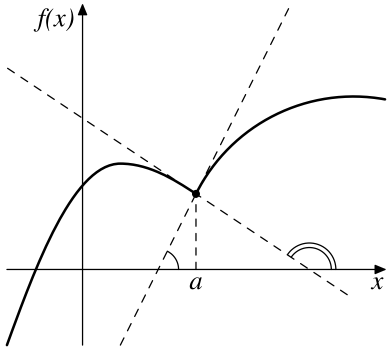

Рис. 26b

Приведем теперь несколько примеров исследования функции на дифференцируемость.

1) *Функция* $f(x) = x^2$ *дифференцируема в любой точке* $a \in \mathbb{R}$ *и* $f'(a) = 2a$. Действительно, $\lim\limits_{x\to a} \dfrac{x^2 - a^2}{x - a} = \lim\limits_{x\to a} (x + a) = 2a$.

2) *Функция* $f(x) = \frac{1}{x}$ *дифференцируема в любой точке* $a \in \mathbb{R} \setminus \{0\}$ *и* $f'(a) = -\frac{1}{a^2}$. Если $a \in \mathbb{R} \setminus \{0\}$, то

$$\lim_{x\to a} \frac{\frac{1}{x} - \frac{1}{a}}{x - a} = \lim_{x\to a} \frac{a - x}{xa\,(x - a)} = -\lim_{x\to a} \frac{1}{xa} = -\frac{1}{a^2}.$$

3) *Пусть* $f(x) = x^2 \sin\frac{1}{x}$ *при* $x \neq 0$ *и* $f(0) = 0$. *Тогда* f *дифференцируема в точке 0 и* $f'(0) = 0$. Так как $f(0) = 0$, то

$$f'(0) = \lim_{x\to 0} \frac{f(x)}{x} = \lim_{x\to 0} x \sin\frac{1}{x} = 0$$

(см. пример 7 в § 4 главы 2).

4) *Функция* $f(x) = \sqrt[3]{x}$ *не дифференцируема в точке 0, но* $f'(0) = +\infty$. Действительно, $f'(0) = \lim\limits_{x\to 0} \dfrac{\sqrt[3]{x}}{x} = \lim\limits_{x\to 0} x^{-2/3} = +\infty$.

5) *У функции* $f(x) = |x|$ *не существует в точке 0 производной (даже бесконечной), но она имеет в нуле конечные левую и правую*

производные. Так как $f'_{\pm}(0) = \lim\limits_{x\to 0\pm} \dfrac{|x|}{x}$, то $f'_{-}(0) = -1$ и $f'_{+}(0) = 1$. Поскольку $f'_{-}(0) \neq f'_{+}(0)$, у функции f не существует производной в нуле.

6) *Пусть $f(x) = x\sin\frac{1}{x}$ при $x \neq 0$ и $f(0) = 0$. Эта функция не имеет в нуле даже односторонних производных.* Действительно, разностное отношение $\dfrac{f(x) - f(0)}{x}$ равно $\sin\dfrac{1}{x}$, поэтому оно не имеет в нуле предела ни слева, ни справа (см. § 4 главы 2).

В заключение обсудим взаимосвязь между дифференцируемостью и непрерывностью.

Замечания

1) *Если функция f дифференцируема в точке a, то она непрерывна в точке a.* Действительно,

$$f(x) - f(a) = f'(a)(x-a) + o(x-a) \to 0 \quad (x \to a).$$

2) Обратить замечание 1 нельзя: примеры 4 – 6 показывают, что из непрерывности не следует дифференцируемость. Более того, справедлив более сильный факт, впервые установленный Вейерштрассом: *существует непрерывная на $\mathbb{R}$ функция, не дифференцируемая ни в одной точке.* Мы не будем доказывать это утверждение.

§ 3. Правила дифференцирования

Производную дифференцируемой функции можно находить, вычисляя предел ее разностного отношения (см. § 2). Указанный способ дает *принципиальную возможность* продифференцировать любую функцию, однако для *практического применения* он неудобен. В этом параграфе мы установим связь операции дифференцирования с другими операциями, производимыми над функциями, — композицией и арифметическими действиями. Кроме того, мы докажем теорему о дифференцировании обратной функции. Эти правила позволяют выразить производную сложно устроенной функции через производные функций более простого вида. Отсюда, в частности, вытекает, что производная любой элементарной функции может быть выражена через производные основных элементарных функций, определенных в § 5 главы 2. Таблица этих производных выводится в конце параграфа.

Теорема 1. Дифференцирование композиции. *Пусть* $f\colon \langle A,B\rangle \to \langle C,D\rangle$, $g\colon \langle C,D\rangle \to \mathbb{R}$, $a \in \langle A,B\rangle$. *Если f дифференцируема в точке a и g дифференцируема в точке $f(a)$, то $g\circ f$ дифференцируема в точке a и*

$$\big(g\circ f\big)'(a) = g'\big(f(a)\big)\cdot f'(a).$$

Замечания

1) Для *линейных* функций f и g утверждение теоремы очевидно, поскольку их композиция будет снова линейной функцией, угловой коэффициент которой есть произведение угловых коэффициентов f и g, то есть их производных.

2) Правило дифференцирования композиции легко обобщается на случай нескольких функций: производная композиции равна произведению производных всех функций в соответствующих точках. Например, для трех функций формула примет вид

$$\big(h\circ g\circ f\big)'(a) = h'\big((g\circ f)(a)\big)\cdot g'\big(f(a)\big)\cdot f'(a).$$

Поэтому теорему 1 называют еще *правилом цепочки.* Предлагаем читателю самостоятельно сформулировать обобщение теоремы на случай конечного числа функций.

3) Правило цепочки можно также сформулировать с помощью дифференциалов: *дифференциал композиции есть композиция дифференциалов*, то есть $d_a(f\circ g) = d_{f(a)}g\circ d_a f$. Действительно, для любого $h\in\mathbb{R}$

$$\begin{aligned} d_a(f\circ g)(h) = (f\circ g)'(a)\cdot h =& g'\big(f(a)\big)\cdot f'(a)\cdot h = \\ &= d_{f(a)}g\big(d_a f(h)\big) = \big(d_{f(a)}g\circ d_a f\big)(h). \end{aligned}$$

Такая редакция теоремы 1 допускает обобщение на случай функций нескольких переменных.

Доказательство. Воспользуемся третьим утверждением критерия дифференцируемости. Существует функция $F\colon \langle A,B\rangle \to \mathbb{R}$, непрерывная в точке a, такая что

$$F(a) = f'(a) \quad \text{и} \quad f(x) - f(a) = F(x)(x-a) \ \text{ при } \ x\in\langle A,B\rangle.$$

Кроме того, найдется функция $G\colon \langle C, D\rangle \to \mathbb{R}$, непрерывная в точке $b = f(a)$, для которой

$$G(b) = g'(b) \quad \text{и} \quad g(y) - g(b) = G(y)(y-b) \ \text{ при } \ y \in \langle C, D\rangle.$$

Подставив в последнее уравнение $y = f(x)$, мы получим

$$\begin{aligned}\bigl(g\circ f\bigr)(x) - \bigl(g\circ f\bigr)(a) =& G\bigl(f(x)\bigr)\cdot\bigl(f(x)-f(a)\bigr) = \\ =& G\bigl(f(x)\bigr)\cdot F(x)(x-a) = H(x)(x-a),\end{aligned}$$

где $H(x) = G\bigl(f(x)\bigr)\cdot F(x)$ $\ (x \in \langle A, B\rangle)$. Так как f и F непрерывны в точке a, а G непрерывна в точке $f(a)$, то H непрерывна в точке a. Снова применяя критерий дифференцируемости, мы получим, что $g\circ f$ дифференцируема в точке a и

$$\bigl(g\circ f\bigr)'(a) = H(a) = G(f(a))\cdot F(a) = g'\bigl(f(a)\bigr)\cdot f'(a). \quad \square$$

Теорема 2. Дифференцирование и арифметические операции. *Пусть $f, g\colon \langle A, B\rangle \to \mathbb{R}$, $a \in \langle A, B\rangle$, f и g дифференцируемы в точке a. Тогда справедливы следующие утверждения.*

1) *Для любых $\alpha, \beta \in \mathbb{R}$ функция $\alpha f + \beta g$ дифференцируема в точке a и*

$$\bigl(\alpha f + \beta g\bigr)'(a) = \alpha f'(a) + \beta g'(a).$$

2) *Функция $f\cdot g$ дифференцируема в точке a и*

$$\bigl(f\cdot g\bigr)'(a) = f'(a)\cdot g(a) + g'(a)\cdot f(a).$$

3) *Если $g(a) \neq 0$, то $\dfrac{f}{g}$ дифференцируема в точке a и*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)\cdot g(a) - g'(a)\cdot f(a)}{g^2(a)}.$$

Замечания

1) Утверждение 1 называется *линейностью операции дифференцирования.* По индукции это правило очевидным образом обобщается на случай нескольких функций.

2) Утверждение 2 также можно доказать для произведения нескольких функций. Например, для трех функций формула примет вид

$$\big(f\cdot g\cdot h\big)'(a)=f'(a)\cdot g(a)\cdot h(a)+f(a)\cdot g'(a)\cdot h(a)+f(a)\cdot g(a)\cdot h'(a).$$

Предлагаем читателю написать общую формулировку.

3) В условиях утверждения 3 функция g не обращается в ноль в некоторой проколотой окрестности точки a, то есть существует такое число $\delta>0$, что $g(x)\neq 0$ для любого $x\in(a-\delta,a+\delta)\cap\langle A,B\rangle$. Это следует из замечания о стабилизации знака непрерывной функции (см. § 4 главы 2), так как $g(a)\neq 0$, а дифференцируемость g в точке a влечет ее непрерывность. Таким образом, область определения функции $\dfrac{f}{g}$ и точка a удовлетворяют соглашению (+), принятому в § 1.

Доказательство.

1) Воспользуемся выражением производной через предел. В силу линейности предела мы получим

$$\big(\alpha f+\beta g\big)'(a)=\lim_{x\to a}\frac{(\alpha f(x)+\beta g(x))-(\alpha f(a)+\beta g(a))}{x-a}=$$
$$=\alpha\lim_{x\to a}\frac{f(x)-f(a)}{x-a}+\beta\lim_{x\to a}\frac{g(x)-g(a)}{x-a}=\alpha f'(a)+\beta g'(a).$$

2) Докажем вначале равенство $\left(f^2\right)'(a)=2f(a)\cdot f'(a)$, соответствующее случаю $f=g$. Если $h(t)=t^2$, то $f^2(x)=\big(h\circ f\big)(x)$. По теореме 1 функция f^2 дифференцируема в точке a и

$$\left(f^2\right)'(a)=h'\big(f(a)\big)\cdot f'(a)=2f(a)\cdot f'(a)$$

(см. пример 1 в § 2). В общем случае

$$\begin{aligned}\big(fg\big)'(a)&=\left(\tfrac14\left((f+g)^2-(f-g)^2\right)\right)'(a)=\\&=\tfrac12\left(f(a)+g(a)\right)\cdot\left(f'(a)+g'(a)\right)-\\&-\tfrac12\big(f(a)-g(a)\big)\cdot\big(f'(a)-g'(a)\big)=f'(a)\cdot g(a)+g'(a)\cdot f(a)\end{aligned}$$

(мы использовали первое утверждение теоремы и частный случай второго утверждения, доказанный выше).

3) Проверим вначале равенство $\left(\dfrac{1}{g}\right)'(a) = -\dfrac{g'(a)}{g^2(a)}$, соответствующее случаю $f \equiv 1$. Если $h(t) = \dfrac{1}{t}$, то $\dfrac{1}{g(x)} = \big(h \circ g\big)(x)$, откуда

$$\left(\frac{1}{g}\right)'(a) = h'\big(g(a)\big) \cdot g'(a) = -\frac{g'(a)}{g^2(a)}$$

(см. пример 2 в § 2). В общем случае

$$\left(\frac{f}{g}\right)'(a) = \left(f \cdot \frac{1}{g}\right)'(a) =$$
$$= f'(a) \cdot \frac{1}{g(a)} + f(a) \cdot \left(-\frac{g'(a)}{g^2(a)}\right) = \frac{f'(a) \cdot g(a) - g'(a) \cdot f(a)}{g^2(a)}$$

(мы использовали второе утверждение теоремы и частный случай третьего утверждения). $\square$

Замечание. Теорему 2 можно доказать и иначе, с помощью третьего утверждения критерия дифференцируемости. Предлагаем читателю сделать это в качестве упражнения.

Теорема 3. Дифференцирование обратной функции. *Пусть f — строго монотонная непрерывная функция на $\langle A, B\rangle$, $a \in \langle A, B\rangle$, f дифференцируема в точке a, $f'(a) \neq 0$. Тогда обратная к f функция f^{-1} дифференцируема в точке $f(a)$ и*

$$\left(f^{-1}\right)'\big(f(a)\big) = \frac{1}{f'(a)}.$$

Замечания

1) Геометрический смысл теоремы ясен из рисунка 27а. Так как график f^{-1} получается из графика f симметрией относительно прямой $y = x$, то

$$\left(f^{-1}\right)'\big(f(a)\big) = \operatorname{tg}\beta = \operatorname{ctg}\alpha = \frac{1}{f'(a)}.$$

Заметим, что в случае $f'(a) = 0$ касательная к графику f^{-1} в точке $f(a)$ направлена вертикально, что соответствует бесконечной производной f^{-1} в точке $f(a)$ (рисунок 27b).

2) Используя дифференциалы, правило дифференцирования обратной функции можно записать в виде $d_{f(a)} f^{-1} = \big(d_a f\big)^{-1}$. Действительно, для любого $h \in \mathbb{R}$

$$\big(d_a f\big)^{-1}(h) = \frac{1}{f'(a)} \cdot h = \big(f^{-1}\big)'\big(f(a)\big) \cdot h = \big(d_{f(a)} f^{-1}\big)(h).$$

В такой редакции утверждение теоремы справедливо и для функций нескольких переменных.

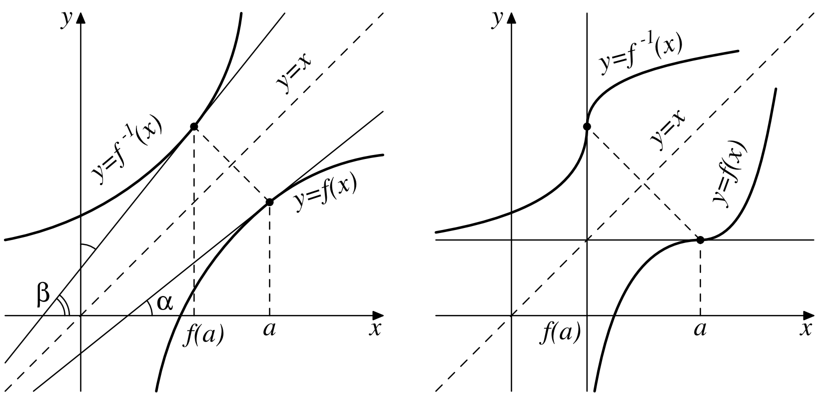

Рис. 27a Рис. 27b

Доказательство. Пусть $g = f^{-1}$, $b = f(a)$. По теореме 9 § 4 главы 2 множество значений f есть промежуток (обозначим его $\langle C, D\rangle$) и g непрерывна на $\langle C, D\rangle$. В силу утверждения 3 критерия дифференцируемости существует функция $F\colon \langle A, B\rangle \to \mathbb{R}$, непрерывная в точке a, для которой

$$F(a) = f'(a) \quad \text{и} \quad f(x) - f(a) = F(x)(x - a) \quad (x \in \langle A, B\rangle).$$

Так как $f(x) \neq f(a)$ при $x \neq a$, то $F(x) \neq 0$ при $x \neq a$. Кроме того, $F(a) = f'(a) \neq 0$. Поэтому $F \neq 0$ всюду на $\langle A, B\rangle$ и, следовательно, функция $\dfrac{1}{F}$ определена на $\langle A, B\rangle$ и непрерывна в точке a. Положив $x = g(y)$ $(y \in \langle C, D\rangle)$, мы получим

$$y - b = f(x) - f(a) = F(x)(x - a) = F\bigl(g(y)\bigr) \cdot \bigl(g(y) - g(b)\bigr),$$

откуда

$$g(y) - g(b) = H(y)(y - b), \quad \text{где } H(y) = \frac{1}{F\bigl(g(y)\bigr)}\,.$$

Функция H определена на $\langle C, D\rangle$ и непрерывна в точке b по теореме о непрерывности композиции (см. § 4 главы 2). Отсюда в силу критерия дифференцируемости

$$g'(b) = H(b) = \frac{1}{F\bigl(g(b)\bigr)} = \frac{1}{F(a)} = \frac{1}{f'(a)}\,. \quad \square$$

Перейдем к вычислению производных основных элементарных функций. В приведенной ниже таблице для каждой функции $f(x)$ указывается формула для ее производной $f'(x)$ и область определения производной (она обозначается $D_{f'}$). Не следует путать $D_{f'}$ с областью определения *формулы*, которой f' задается. Так, функция $\frac{1}{x}$ определена на $\mathbb{R} \setminus \{0\}$, но с производной $f(x) = \ln x$ она совпадает только на $(0, +\infty)$. Более строгой была бы запись $f'(x) = \frac{1}{x}\Big|_{(0,+\infty)}$, но мы будем писать просто $f'(x) = \frac{1}{x}$, подразумевая при этом, что $x \in D_{f'}$.

Подробного описания требует область задания производной степенной функции $f(x) = x^\alpha$, так как она зависит от α. Случай $\alpha = 0$, соответствующий постоянной функции, разобран в таблице отдельно. Для $\alpha \neq 0$ возможны четыре ситуации.

a) $\alpha \in \mathbb{N}$. Тогда f' определена на $\mathbb{R}$. В случае $\alpha = 1$ мы считаем $x^{\alpha-1} = 1$ даже при $x = 0$, то есть $x' \equiv 1$.

b) $\alpha \in \mathbb{Z} \setminus \mathbb{N}$. Тогда f' определена на $\mathbb{R} \setminus \{0\}$.

c) $\alpha = \frac{m}{2n+1}$, где $m \in \mathbb{Z}$, $n \in \mathbb{N}$. При $\alpha \geqslant 1$ функция f' определена на $\mathbb{R}$, а при $\alpha < 1$ — на $\mathbb{R} \setminus \{0\}$.

d) $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ или $\alpha = \frac{2m+1}{2n}$, где $m \in \mathbb{Z}$, $n \in \mathbb{N}$. При $\alpha > 1$ функция f' определена на $[0, +\infty)$, а при $\alpha < 1$ — на $(0, +\infty)$.

Таблица производных

№	Функция $f(x)$	Производная $f'(x)$	Множество $D_{f'}$
1	1	0	$\mathbb{R}$
2	e^x	e^x	$\mathbb{R}$
	$b^x \quad (b>0)$	$b^x \ln b$	
3	$\ln x$	$\dfrac{1}{x}$	$(0,+\infty)$
	$\log_b x \ \ (b>0\,,\ b\neq 1)$	$\dfrac{1}{x\ln b}$	
4	x^α	$\alpha x^{\alpha-1}$	См. описание выше
5	$\sin x$	$\cos x$	$\mathbb{R}$
6	$\cos x$	$-\sin x$	$\mathbb{R}$
7	$\operatorname{tg} x$	$\dfrac{1}{\cos^2 x}$	$\mathbb{R}\setminus\{\frac{\pi}{2}+\pi k:\ k\in\mathbb{Z}\}$
8	$\operatorname{ctg} x$	$-\dfrac{1}{\sin^2 x}$	$\mathbb{R}\setminus\{\pi k:\ k\in\mathbb{Z}\}$
9	$\arcsin x$	$\dfrac{1}{\sqrt{1-x^2}}$	$(-1,1)$
10	$\arccos x$	$-\dfrac{1}{\sqrt{1-x^2}}$	$(-1,1)$
11	$\operatorname{arctg} x$	$\dfrac{1}{1+x^2}$	$\mathbb{R}$
12	$\operatorname{arcctg} x$	$-\dfrac{1}{1+x^2}$	$\mathbb{R}$

Докажем формулы для производных, приведенные в таблице. При этом мы будем использовать замечательные пределы, выведенные в главе 2.

1) Пусть $f(x)\equiv 1$, $a\in\mathbb{R}$. Тогда $f'(a)=\lim\limits_{x\to a}\dfrac{1-1}{x-a}=0$.

2) Достаточно рассмотреть $f(x)=b^x \quad (b>0)$. Пусть $a\in\mathbb{R}$. Тогда

$$f'(a)=\lim_{x\to a}\frac{b^x-b^a}{x-a}=b^a\cdot\lim_{x\to a}\frac{b^{x-a}-1}{x-a}=b^a\ln b.$$

3) Достаточно рассмотреть $f(x) = \log_b x$ $(b > 0,\ b \neq 1)$. Пусть $a > 0$. Тогда

$$f'(a) = \lim_{x\to a} \frac{\log_b x - \log_b a}{x-a} = \lim_{x\to a} \frac{\log_b \frac{x}{a}}{x-a} = \lim_{x\to a} \frac{\frac{x}{a} - 1}{(x-a)\ln b} = \frac{1}{a\ln b}.$$

4) Пусть $f(x) = x^\alpha$, где $\alpha \neq 0$. Для α возможны четыре ситуации, описанные выше. В случаях a) – c) областью определения f является $\mathbb{R}$ или $\mathbb{R} \setminus \{0\}$, а в случае d) — $[0, +\infty)$ или $(0, +\infty)$ (см. § 5 главы 2). Равенства

$$f'(a) = \lim_{x\to a} \frac{x^\alpha - a^\alpha}{x-a} =$$
$$= a^\alpha \cdot \lim_{x\to a} \frac{\left(\frac{x}{a}\right)^\alpha - 1}{x-a} = a^\alpha \cdot \lim_{x\to a} \frac{\alpha\left(\frac{x}{a} - 1\right)}{x-a} = \alpha \cdot a^{\alpha-1}$$

справедливы при $a \in \mathbb{R}\setminus\{0\}$ в случаях a) – c) и при $a > 0$ в случае d). Пусть теперь $a = 0$. Заметим, что при $\alpha < 0$ не определено $f(0)$, а значит, и $f'(0)$. Если $\alpha > 0$, то

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0} \frac{x^\alpha}{x} = \lim_{x\to 0} x^{\alpha-1}.$$

Этот предел равен нулю при $\alpha > 1$, единице при $\alpha = 1$ и бесконечности при $\alpha \in (0,1)$. Поэтому $f'(0) = 0 = \alpha \cdot 0^{\alpha-1}$ при $\alpha > 1$, $f'(0) = 1$ при $\alpha = 1$, а при $\alpha \in (0,1)$ функция f не дифференцируема в нуле.

5) Пусть $f(x) = \sin x$, $a \in \mathbb{R}$. Тогда

$$f'(a) = \lim_{x\to a} \frac{\sin x - \sin a}{x-a} =$$
$$= \lim_{x\to a} \frac{2\sin\frac{x-a}{2}\cos\frac{x+a}{2}}{x-a} = \lim_{x\to a} \frac{2\,\frac{x-a}{2}\cos a}{x-a} = \cos a.$$

6) В силу теорем 1 и 2

$$(\cos x)' = \left(\sin(\tfrac{\pi}{2} - x)\right)' = \cos(\tfrac{\pi}{2} - x) \cdot (\tfrac{\pi}{2} - x)' = -\sin x.$$

7) По теореме 2

$$(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x}\right)' =$$
$$= \frac{(\sin x)' \cdot \cos x - (\cos x)' \cdot \sin x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

8) В силу теорем 1 и 2

$$\big(\operatorname{ctg} x\big)' = \big(\operatorname{tg}(\tfrac{\pi}{2} - x)\big)' = \frac{1}{\cos^2(\frac{\pi}{2} - x)} \cdot (\tfrac{\pi}{2} - x)' = -\frac{1}{\sin^2 x}.$$

9) Пусть $f(x) = \arcsin x$, $a \in (-1, 1)$. Положим $g(y) = \sin y$, $b = \arcsin a$. Так как $b \in (-\frac{\pi}{2}, \frac{\pi}{2})$, то $g'(b) = \cos b > 0$. По теореме 3

$$f'(a) = \frac{1}{g'(b)} = \frac{1}{\cos b} = \frac{1}{\sqrt{1 - \sin^2 b}} = \frac{1}{\sqrt{1 - a^2}}.$$

10) Так как $\arccos x = \frac{\pi}{2} - \arcsin x$, то

$$(\arccos x)' = -(\arcsin x)' = -\frac{1}{\sqrt{1 - x^2}}.$$

11) Пусть $f(x) = \operatorname{arctg} x$, $a \in \mathbb{R}$. Положим $g(y) = \operatorname{tg} y$, $b = \operatorname{arctg} a$. Так как $b \in (-\frac{\pi}{2}, \frac{\pi}{2})$, то $g'(b) = \dfrac{1}{\cos^2 b} \neq 0$. По теореме 3

$$f'(a) = \frac{1}{g'(b)} = \cos^2 b = \frac{1}{\operatorname{tg}^2 b + 1} = \frac{1}{a^2 + 1}.$$

12) Так как $\operatorname{arcctg} x = \frac{\pi}{2} - \operatorname{arctg} x$, то

$$(\operatorname{arcctg} x)' = -(\operatorname{arctg} x)' = -\frac{1}{x^2 + 1}.$$

Доказательство приведенных в таблице формул завершено. $\square$

§ 4. Теоремы о среднем и их приложения

В этом параграфе рассматриваются теоремы, связывающие значения функций на концах промежутка со значениями их производных в некоторых "средних" точках, лежащих внутри промежутка. Кроме того, мы выведем несколько важных следствий этих теорем, в частности, правило Лопиталя, полезное при вычислении пределов.

Докажем вначале необходимое условие того, что в заданной внутренней точке промежутка функция достигает своего наибольшего или наименьшего значения. Этот факт будет использоваться как при доказательстве теорем о среднем, так и при изучении локальных экстремумов функции (см. § 7).

Теорема 1 (П. Ферма́). *Пусть* $a \in (A, B)$, $f\colon \langle A, B\rangle \to \mathbb{R}$ — *дифференцируемая в точке* a *функция. Если* $f(a) = \max\limits_{\langle A,B\rangle} f$ *или* $f(a) = \min\limits_{\langle A,B\rangle} f$, *то* $f'(a) = 0$.

Замечание. Геометрический смысл теоремы состоит в том, что в точке a касательная к графику функции горизонтальна (см. рисунок 28). В условиях теоремы существенно, что точка a лежит *внутри* промежутка. Например, для функции $f(x) = x$, заданной на $[0, 1]$, утверждение неверно, так как $\max\limits_{[0,1]} f = f(1)$, но $f'(1) = 1 \neq 0$.

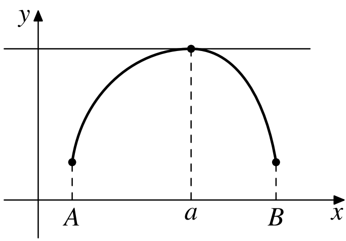

Рис. 28a

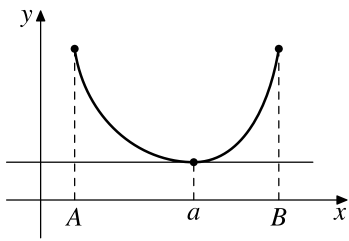

Рис. 28b

Доказательство. Пусть для определенности $f(a) = \max\limits_{\langle A,B\rangle} f$. Тогда $f(x) - f(a) \leqslant 0$ при любом $x \in \langle A, B\rangle$. Если $x > a$, то $\dfrac{f(x) - f(a)}{x - a} \leqslant 0$, откуда

$$f'_+(a) = \lim_{x\to a+} \frac{f(x) - f(a)}{x - a} \leqslant 0.$$

Рассматривая аналогичным образом левосторонний предел, мы получим

$$f'_-(a) = \lim_{x\to a-} \frac{f(x) - f(a)}{x - a} \geqslant 0.$$

Дифференцируемость f в точке a влечет равенства

$$f'_+(a) = f'(a) = f'_-(a)$$

(см. § 2). Поэтому $f'(a) \leqslant 0$ и $f'(a) \geqslant 0$, то есть $f'(a) = 0$. $\square$

Перейдем теперь к теоремам о среднем. Начнем с определения дифференцируемости функции на промежутке.

Определение 1. Функция $f\colon \langle A, B\rangle \to \mathbb{R}$ называется *дифференцируемой на* $\langle A, B\rangle$, если она дифференцируема в любой точке промежутка $\langle A, B\rangle$.

Теорема 2 (М. Ролль). *Пусть функция f непрерывна на $[A, B]$ и дифференцируема на (A, B). Если $f(A) = f(B)$, то существует точка $c \in (A, B)$, для которой $f'(c) = 0$.*

Геометрический смысл теоремы также заключается в том, что касательная к графику f в точке c горизонтальна.

Замечание. Утверждение теоремы может быть неверным, если f не дифференцируема в некоторой точке (A, B) (рисунок 29a) или разрывна на одном из концов $[A, B]$ (рисунок 29b).

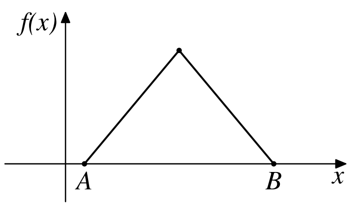

Рис. 29a

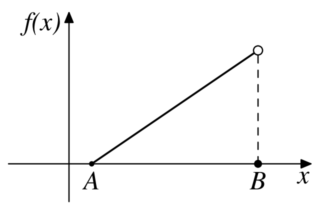

Рис. 29b

Доказательство. По теореме Вейерштрасса (см. § 4 главы 2) найдутся такие $a, b \in [A, B]$, что $f(a) = \min\limits_{[A,B]} f$, $f(b) = \max\limits_{[A,B]} f$. Пусть вначале a и b — концевые точки $[A, B]$. Так как $f(A) = f(B)$, то $\min\limits_{[A,B]} f = \max\limits_{[A,B]} f$. Поэтому f постоянна на $[A, B]$, откуда $f'(c) = 0$ для *любой* точки $c \in (A, B)$. Предположим теперь, что a или b лежит в (A, B). Тогда по теореме Ферма мы получим соответственно $f'(a) = 0$ или $f'(b) = 0$. $\square$

Из критерия дифференцируемости в точке a вытекает равенство $\lim\limits_{x\to a} \dfrac{f(x)-f(a)}{x-a} = f'(a)$. Оно означает, что $\dfrac{f(x)-f(a)}{x-a} \approx f'(a)$ при x, близких к a. Оказывается, у этого соотношения есть глобальный аналог. Даже при *фиксированном* x разностное отношение $\dfrac{f(x)-f(a)}{x-a}$ есть производная f, но не в точке a, а в некоторой "средней" точке, лежащей между a и x. Сформулируем этот важный факт строго.

Теорема 3 (Ж.-Л. Лагранж). *Пусть функция f непрерывна на $[A,B]$ и дифференцируема на (A,B). Тогда существует точка $c \in (A,B)$, для которой $\dfrac{f(B)-f(A)}{B-A} = f'(c)$.*

Замечание. Геометрический смысл теоремы Лагранжа таков: в некоторой точке $c \in (A,B)$ касательная к графику f параллельна хорде, соединяющей концы графика f (см. рисунок 30).

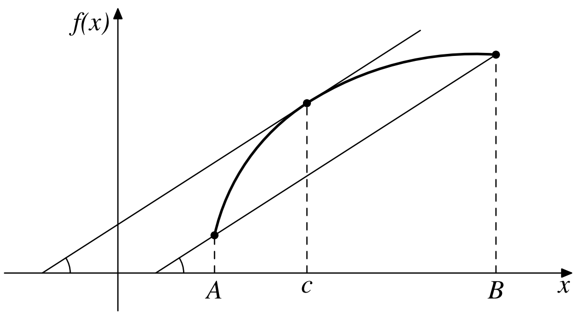

Рис. 30

Доказательство. Рассмотрим функцию

$$g(x) = f(x) - \frac{f(B)-f(A)}{B-A}(x-A).$$

Ясно, что $g(A) = g(B) = f(A)$. Применяя теорему Ролля к функции g, найдем точку $c \in (A,B)$, для которой $g'(c) = 0$. Последнее равенство эквивалентно $f'(c) = \dfrac{f(B)-f(A)}{B-A}$. $\square$

Замечания

1) Точка c, вообще говоря, не единственна. Например, для функции $f(x) = x$ в качестве c можно взять *любую* точку из (A, B).

2) Теорема Ролля является частным случаем теоремы Лагранжа.

3) Утверждение теоремы Лагранжа можно записать в виде

$$f(B) - f(A) = f'(c)(B - A).$$

Это равенство называют *формулой конечных приращений*, поскольку оно выражает приращение функции f на отрезке $[A, B]$ через приращение ее аргумента.

4) Теорему Лагранжа часто используют в следующей симметричной форме: *существует такое $\theta \in (0, 1)$, что*

$$\frac{f(B) - f(A)}{B - A} = f'\big(A + \theta(B - A)\big).$$

Для доказательства достаточно положить $\theta = \dfrac{c - A}{B - A}$, где c — точка из утверждения теоремы. Преимущество такой записи состоит в том, что она справедлива и для функций, заданных на $[B, A]$ при $B < A$. Действительно, точка $A + \theta(B - A)$ лежит между A и B *при любых* A и B $(A \neq B)$, а левая часть равенства не меняется при перестановке A и B.

Следствие 1. Оценка конечных приращений. *Пусть функция f непрерывна на $[A, B]$ и дифференцируема на (A, B). Если существуют $m, M \in \mathbb{R}$: $m \leqslant f'(x) \leqslant M$ при всех $x \in (A, B)$, то*

$$m\,(B - A) \leqslant f(B) - f(A) \leqslant M\,(B - A).$$

В частности, если $|f'(x)| \leqslant M$ при всех $x \in (A, B)$, то

$$|f(B) - f(A)| \leqslant M\,(B - A).$$

Доказательство. Так как по формуле конечных приращений $f(B) - f(A) = f'(c)\,(B - A)$, следствие 1 вытекает из неравенства $m \leqslant f'(c) \leqslant M$. $\square$

Следствие 2. *Пусть f дифференцируема на $\langle A, B\rangle$ и существует такое $M > 0$, что $|f'(x)| \leqslant M$ при всех $x \in \langle A, B\rangle$. Тогда f равномерно непрерывна на $\langle A, B\rangle$.*

Доказательство. Для любого $\varepsilon > 0$ положим $\delta = \frac{\varepsilon}{M}$. Если $x, y \in \langle A, B\rangle$ и $|x - y| < \delta$, то в силу следствия 1

$$|f(x) - f(y)| \leqslant M\,|x - y| < M\delta = \varepsilon. \quad \square$$

Теорема 4 (О. Коши). *Пусть функции f и g непрерывны на $[A, B]$ и дифференцируемы на (A, B). Предположим, что $g'(x) \neq 0$ при всех $x \in (A, B)$. Тогда существует точка $c \in (A, B)$, для которой*

$$\frac{f(B) - f(A)}{g(B) - g(A)} = \frac{f'(c)}{g'(c)}.$$

Замечания

1) Теорема Лагранжа получается из теоремы Коши, если взять $g(x) = x$.

2) В условиях теоремы 4 значения функции g в точках A и B различны, иначе по теореме Ролля нашлась бы точка $a \in (A, B)$, в которой $g'(a) = 0$.

3) Утверждение теоремы Коши можно записать в симметричной форме: *существует такое $\theta \in (0, 1)$, что*

$$\frac{f(B) - f(A)}{g(B) - g(A)} = \frac{f'\big(A + \theta(B - A)\big)}{g'\big(A + \theta(B - A)\big)}.$$

В такой редакции теорема справедлива и при $B < A$ (см. замечание 4 к теореме Лагранжа).

Доказательство. Положим

$$h(x) = f(x) - \frac{f(B) - f(A)}{g(B) - g(A)}\,(g(x) - g(A)).$$

Ясно, что $h(A) = h(B) = f(A)$. Применяя теорему Ролля к функции h, найдем точку $c \in (A, B)$, для которой $h'(c) = 0$. Последнее равенство эквивалентно $f'(c) = \dfrac{f(B)-f(A)}{g(B)-g(A)} \cdot g'(c)$, что и доказывает теорему. $\square$

Рассмотрим теперь применение производной к вычислению пределов. Мы приведем правило Лопиталя раскрытия неопределенностей в двух вариантах. Первый из них относится к неопределенностям вида $\frac{0}{0}$.

Теорема 5. Правило Лопиталя для бесконечно малых функций. *Пусть $-\infty \leqslant a < b \leqslant +\infty$, функции f и g дифференцируемы на (a,b), причем $g'(x) \neq 0$ для всех $x \in (a,b)$. Предположим, что выполнены следующие условия:*

1) $f(x) \to 0$ и $g(x) \to 0$ при $x \to a$;

2) предел $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)}$ существует и равен $L \in \overline{\mathbb{R}}$.

Тогда предел $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ также существует и равен L.

Замечания

1) В данной редакции правило Лопиталя служит для нахождения *правостороннего предела*. Предлагаем читателю самостоятельно сформулировать аналогичные утверждения для *левостороннего предела* и *двустороннего предела*.

2) В условиях теоремы $g(x) \neq 0$ для всех $x \in (a,b)$, то есть функция $\dfrac{f}{g}$ определена на (a,b), как и $\dfrac{f'}{g'}$. Это будет установлено в процессе доказательства.

Доказательство. Рассмотрим два случая.

1) $a \in \mathbb{R}$. Доопределим функции f и g нулем в точке a. Тогда f и g станут непрерывными на $[a,b)$. Пусть $x \in (a,b)$. Существует такое $c(x) \in (a,b)$, что

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'\big(c(x)\big)}{g'\big(c(x)\big)},$$

причем $g(x) \neq 0$ (см. теорему Коши и замечание к ней). Так как $a < c(x) < x$, то $\lim\limits_{x\to a} c(x) = a$ и $c(x) \neq a$ при $x \neq a$. По теореме о пределе композиции (см. § 3 главы 2)

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'\big(c(x)\big)}{g'\big(c(x)\big)} = \lim_{x\to a} \left(\frac{f'}{g'} \circ c\right)(x) = L.$$

2) $a = -\infty$. В силу локальности предела можно считать, что $b \in \mathbb{R}$. Функции $\varphi(t) = f(\ln t)$ и $\psi(t) = g(\ln t)$ определены и дифференцируемы на $(0, e^b)$, причем

$$\varphi'(t) = f'(\ln t) \cdot \frac{1}{t} \quad \text{и} \quad \psi'(t) = g'(\ln t) \cdot \frac{1}{t}$$

(см. теорему 1 § 3). Отсюда $\psi'(t) \neq 0$ при всех $t \in (0, e^b)$. По теореме о пределе композиции

$$\frac{\varphi'(t)}{\psi'(t)} = \frac{f'}{g'}\big(\ln t\big) \to L \quad \text{при} \quad t \to 0+\,.$$

Так как $\varphi(t) \to 0$ и $\psi(t) \to 0$ при $t \to 0+$, функции φ и ψ удовлетворяют условиям теоремы 5, которая для промежутка $(0, e^b)$ уже доказана. Применяя ее, мы получим

$$\lim_{x\to-\infty} \frac{f(x)}{g(x)} = \lim_{t\to0+} \frac{\varphi(t)}{\psi(t)} = \lim_{t\to0+} \frac{\varphi'(t)}{\psi'(t)} = L.$$

Заметим также, что $g(\ln t) = \psi(t) \neq 0$ для $t \in (0, e^b)$, поэтому $g(x) \neq 0$ для всех $x \in (a, b)$. $\square$

Сформулируем теперь второй вариант правила Лопиталя, относящийся к неопределенностям вида $\frac{\infty}{\infty}$.

Теорема 6. Правило Лопиталя для бесконечно больших функций. *Пусть $-\infty \leqslant a < b \leqslant +\infty$, функции f и g дифференцируемы на (a, b) , причем $g'(x) \neq 0$ для всех $x \in (a, b)$. Предположим, что выполнены следующие условия:*

1) $\lim\limits_{x\to a} g(x) = \infty$;

2) *предел* $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)}$ *существует и равен* $L \in \overline{\mathbb{R}}$.

Тогда предел $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ *также существует и равен* L.

Замечания к теореме 5 справедливы и для теоремы 6. Отметим, что в теореме 6 функция f *не предполагается* бесконечно большой, хотя на практике правило Лопиталя обычно применяют при наличии неопределенности.

Доказательство теоремы 6 мы проведем только для *конечного* предела L, оставляя случаи $L = \pm\infty$ читателю в качестве упражнения.

Доказательство. Предположим вначале, что $L = 0$. Пусть $\varepsilon > 0$. Так как $\lim\limits_{t\to a} \dfrac{f'(t)}{g'(t)} = 0$, то найдется такое $\sigma \in (a, b)$, что $\left|\dfrac{f'(t)}{g'(t)}\right| < \dfrac{\varepsilon}{2}$ при всех $t \in (a, \sigma)$. Фиксируем $y \in (a, \sigma)$, и пусть $x \in (a, y)$. По теореме Коши для некоторого $t \in (x, y)$ выполняется равенство $\dfrac{f(x) - f(y)}{g(x) - g(y)} = \dfrac{f'(t)}{g'(t)}$, откуда $\left|\dfrac{f(x) - f(y)}{g(x) - g(y)}\right| < \dfrac{\varepsilon}{2}$. Заметим, что

$$\lim_{x\to a} \frac{g(x) - g(y)}{g(x)} = 1 \quad \text{и} \quad \lim_{x\to a} \frac{f(y)}{g(x)} = 0,$$

поскольку $\lim\limits_{x\to a} g(x) = \infty$, а y фиксировано. Поэтому найдется такое $\delta \in (a, y)$, что

$$0 < \frac{g(x) - g(y)}{g(x)} < \frac{3}{2} \quad \text{и} \quad \left|\frac{f(y)}{g(x)}\right| < \frac{\varepsilon}{4} \quad \text{при любых } x \in (a, \delta).$$

Тогда для всех $x \in (a, \delta)$

$$\left|\frac{f(x)}{g(x)}\right| = \left|\frac{f(x) - f(y)}{g(x) - g(y)} \cdot \frac{g(x) - g(y)}{g(x)} + \frac{f(y)}{g(x)}\right| \leqslant$$
$$\leqslant \left|\frac{f(x) - f(y)}{g(x) - g(y)}\right| \cdot \frac{g(x) - g(y)}{g(x)} + \left|\frac{f(y)}{g(x)}\right| < \frac{\varepsilon}{2} \cdot \frac{3}{2} + \frac{\varepsilon}{4} = \varepsilon.$$

Это означает, что $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = 0 = L$.

Пусть теперь L — произвольное вещественное число. Положим $h(x) = f(x) - L\,g(x)$. Тогда

$$\lim_{x\to a} \frac{h'(x)}{g'(x)} = \lim_{x\to a} \left(\frac{f'(x)}{g'(x)} - L\right) = 0,$$

и по доказанному выше

$$\lim_{x\to a} \frac{f(x)}{g(x)} - L = \lim_{x\to a} \frac{f(x) - L\,g(x)}{g(x)} = \lim_{x\to a} \frac{h(x)}{g(x)} = 0. \quad \square$$

Замечания

1) В случае $L > 0$ правило Лопиталя имеет следующую физическую интерпретацию: *если одна материальная точка движется быстрее другой приблизительно в L раз, то и расстояния, пройденные этими точками за равное время, будут различаться примерно в L раз.*

2) Обратить утверждения теорем 5 и 6 нельзя. Рассмотрим функции $f(t) = t + \sin t$ и $g(t) = t$. Тогда

$$\lim_{t\to+\infty} \frac{f(t)}{g(t)} = \lim_{t\to+\infty} \left(1 + \frac{\sin t}{t}\right) = 1,$$

но отношение $\dfrac{f'(t)}{g'(t)}$, равное $1 + \cos t$, не имеет предела на $+\infty$. Этот пример также имеет физическую интерпретацию. Если колесо катится без проскальзывания с единичной скоростью, то в момент времени t абсциссы оси колеса и ниппеля на его ободе равны соответственно $g(t)$ и $f(t)$. За большой промежуток времени ось и ниппель пройдут примерно равные расстояния, то есть отношение этих расстояний будет близко к единице. Тем не менее скорость оси колеса постоянна и равна 1, а горизонтальная составляющая скорости ниппеля колеблется между 0 и 2.

Выведем с помощью правила Лопиталя несколько равенств, играющих важную роль в математическом анализе.

1) *Если* $\alpha > 0$, *то* $\ln x = o\big(x^\alpha\big)$ *при* $x \to +\infty$. Действительно, положим $f(x) = \ln x$, $g(x) = x^\alpha$. Тогда f и g — бесконечно большие функции на $+\infty$, $g'(x) = \alpha x^{\alpha-1} \neq 0$ на $(0, +\infty)$. Так как

$$\lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to+\infty} \frac{1/x}{\alpha x^{\alpha-1}} = \lim_{x\to+\infty} \frac{1}{\alpha x^\alpha} = 0,$$

то и $\lim\limits_{x\to+\infty} \dfrac{\ln x}{x^\alpha} = 0$, что и требовалось доказать. $\square$

2) *Если* $\alpha > 0$, $b > 1$, *то* $x^\alpha = o\big(b^x\big)$ *при* $x \to +\infty$. Пусть вначале $\alpha = 1$. Полагая $f(x) = x$ и $g(x) = b^x$, мы получим

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = \lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to+\infty} \frac{1}{b^x \ln b} = 0.$$

В общем случае положим $c = b^{1/\alpha}$. Так как $c > 1$, то

$$\frac{x^\alpha}{b^x} = \left(\frac{x}{c^x}\right)^\alpha \to 0 \quad \text{при} \quad x \to +\infty$$

(мы воспользовались уже доказанным частным случаем и теоремой о пределе композиции). $\square$

Заметим, что общий случай можно получить и непосредственно по правилу Лопиталя, применяя его *несколько раз*. Например,

$$\lim_{x\to+\infty} \frac{x^2}{b^x} = \lim_{x\to+\infty} \frac{2x}{b^x \ln b} = \lim_{x\to+\infty} \frac{2}{b^x (\ln b)^2} = 0.$$

Примеры 1 и 2 показывают, что степенная функция растет на $+\infty$ быстрее логарифма, но медленнее показательной функции. Этот факт бывает полезен при вычислении пределов.

3) $\lim\limits_{x\to0+} x^x = 1$. Воспользоваться правилом Лопиталя напрямую здесь нельзя. Заметим, что $\ln x^x = x \ln x = \frac{\ln x}{1/x}$. Теперь мы получили неопределенность $\frac{\infty}{\infty}$, и можно применить теорему 6. Так как

$$\lim_{x\to0+} \frac{\ln x}{1/x} = \lim_{x\to0+} \frac{1/x}{-1/x^2} = -\lim_{x\to0+} x = 0,$$

то $\lim\limits_{x\to0+} x^x = 1$. $\square$

§ 5. Производные высших порядков

До сих пор мы говорили про первую производную функции. Введем теперь понятие производной произвольного порядка.

Пусть $E \subset \mathbb{R}$, $f\colon E \to \mathbb{R}$. Определим E_1 как множество точек E, удовлетворяющих условию (+) из § 1, в которых функция f дифференцируема. Мы будем называть E_1 *множеством дифференцируемости* f, а функцию $x \mapsto f'(x)$, определенную на E_1, — *производной* f. Обозначим через E_2 множество всех точек $a \in E_1$, удовлетворяющих условиям:

1) существует такое $\delta > 0$, что $E_1 \cap (a-\delta, a+\delta) = E \cap (a-\delta, a+\delta)$;

2) функция f' дифференцируема в точке a.

Пусть $a \in E_2$. Число $(f')'(a)$ называется *второй производной* f в точке a и обозначается $f''(a)$ или $f^{(2)}(a)$ (используются также обозначения $\dfrac{d^2 f}{dx^2}(a)$ и $D^2 f(a)$), а функция f называется *дважды дифференцируемой в точке* a. Функция f'', определенная на E_2 соотношением $x \mapsto f''(x)$, называется *второй производной* f.

Предположим, что мы уже определили функции $f^{(2)}, \ldots, f^{(n-1)}$, заданные соответственно на $E_2, \ldots, E_{n-1}$. Обозначим через E_n множество всех точек $a \in E_{n-1}$, удовлетворяющих условиям:

1) существует такое $\delta > 0$, что $E_{n-1} \cap (a-\delta, a+\delta) = E \cap (a-\delta, a+\delta)$;

2) функция $f^{(n-1)}$ дифференцируема в точке a.

Пусть $a \in E_n$. Число $\left(f^{(n-1)}\right)'(a)$ называется *n-й производной* f в точке a и обозначается $f^{(n)}(a)$ (используются также обозначения $\dfrac{d^n f}{dx^n}(a)$ и $D^n f(a)$), а f называется *n раз дифференцируемой в точке* a. Функция $f^{(n)}$, определенная на E_n соотношением $x \mapsto f^{(n)}(x)$, называется *n-й производной* f. Таким образом, по индукции мы определили n-ю производную f для произвольного $n \in \mathbb{N}$. Договоримся также под нулевой производной понимать саму функцию, то есть считать $f^{(0)} = f$ и $E_0 = E$.

Положим для $a \in E$ и $n \in \mathbb{N}$

$$f_+^{(n)}(a) = \left(f\,\Big|_{[a,+\infty)\cap E}\right)^{(n)}(a) \quad \text{и} \quad f_-^{(n)}(a) = \left(f\,\Big|_{(-\infty,a]\cap E}\right)^{(n)}(a).$$

Заметим, что при $n = 1$ мы получим определение односторонних производных, введенное в § 2.

Определим теперь понятие *дифференциала n-го порядка* функции f. Пусть $h \in \mathbb{R}$. Обозначим $g_1(x) = d_x f(h)$, $x \in E_1$. Для $a \in E_2$ положим $d_a^2 f(h) = d_a g_1(h)$. Тогда $d_a^2 f(h)$ называется *вторым дифференциалом f в точке a на приращении* h. Пусть $d_a^2 f, \ldots, d_a^{n-1} f$ уже определены. Положим $g_{n-1}(x) = d_x^{n-1} f(h)$, $x \in E_{n-1}$. Для $a \in E_n$ число $d_a g_{n-1}(h)$ называется *n-м дифференциалом f в точке a на приращении* h и обозначается $d_a^n f(h)$. Таким образом, по индукции мы определили дифференциал функции f произвольного порядка n. Для $n = 1$ мы будем вместо $d_a^1 f$ писать просто $d_a f$, как в § 2.

Покажем, что

$$d_a^n f(h) = f^{(n)}(a) \cdot h^n \quad \text{при всех} \quad n \in \mathbb{N} \ \text{и} \ a \in E_n.$$

Для $n = 1$ это следует из определения $d_a f$. Пусть для $n-1$ формула доказана. Тогда

$$\begin{aligned} d_a^n f(h) =& d_a g_{n-1}(h) = g'_{n-1}(a) \cdot h = \\ &= \left(f^{(n-1)}(x) \cdot h^{n-1}\right)'\Big|_{x=a} \cdot h = f^{(n)}(a) \cdot h^n. \end{aligned}$$

Введем теперь понятие n-кратной дифференцируемости функции на множестве. Пусть $F \subset E$. Назовем функцию f *n раз дифференцируемой на F*, если она n раз дифференцируема в любой точке F (то есть $F \subset E_n$). Если при этом $f^{(n)}$ непрерывна на F, то f называется *n раз непрерывно дифференцируемой на F*.

Опишем важные классы функций, связанные с дифференцируемостью. Пусть $n \in \mathbb{Z}_+$, $E \subset \mathbb{R}$. Обозначим через $C^n(E)$ множество функций, определенных и n раз непрерывно дифференцируемых на E. Для $n = 0$ класс $C^n(E)$ есть множество непрерывных на E функций. В этом случае вместо $C^0(E)$ мы будем писать просто $C(E)$, что согласуется с обозначением § 4 главы 2. Кроме того, положим $C^\infty(E) = \bigcap\limits_{n=0}^{\infty} C^n(E)$. Этот класс состоит из функций, имеющих на E производные любого порядка.

Из определения ясно, что классы C^n уменьшаются с ростом n, то есть

$$C^n(E) \supset C^{n+1}(E) \ \text{ и } \ C^n(E) \supset C^\infty(E) \ \text{ при всех } \ n \in \mathbb{Z}_+.$$

Заметим, что эти включения строгие. Действительно, положим $f_n(x) = x^{n+1/3}$. Тогда функция

$$f_n^{(n)}(x) = \left(n + \tfrac{1}{3}\right) \cdot \left(n - \tfrac{2}{3}\right) \cdot \ldots \cdot \tfrac{4}{3} \sqrt[3]{x}$$

определена и непрерывна на $\mathbb{R}$, но не дифференцируема в нуле. Поэтому $f_n \in C^n(\mathbb{R})$, но $f_n \notin C^{n+1}(\mathbb{R})$ и, тем более, $f_n \notin C^\infty(\mathbb{R})$.

Заметим также, что класс *дифференцируемых на E* функций строго шире класса *непрерывно дифференцируемых на E* функций. Например, если $f(x) = x^2 \sin \frac{1}{x}$ при $x \neq 0$ и $f(0) = 0$, то f дифференцируема на $\mathbb{R}$ (см. § 2), но $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ разрывна в нуле, так как не имеет там предела. Модифицируя этот пример, можно показать, что класс n раз дифференцируемых на E функций строго шире, чем $C^n(E)$.

В заключение рассмотрим связь производных старших порядков с арифметическими операциями.

Теорема 1. *Пусть $E \subset \mathbb{R}$, $a \in E$, $n \in \mathbb{N}$. Если f и g n раз дифференцируемы в точке a, то справедливы следующие утверждения.*

1) *Для любых $\alpha, \beta \in \mathbb{R}$ функция $\alpha f + \beta g$ n раз дифференцируема в точке a и*

$$(\alpha f + \beta g)^{(n)}(a) = \alpha f^{(n)}(a) + \beta g^{(n)}(a).$$

2) *Функция $f \cdot g$ n раз дифференцируема в точке a и*

$$(f \cdot g)^{(n)}(a) = \sum_{k=0}^{n} C_n^k\, f^{(k)}(a)\, g^{(n-k)}(a).$$

Первое утверждение непосредственно вытекает из линейности операции дифференцирования (см. § 2). Предлагаем читателю проверить его самостоятельно. Выражение для $(f \cdot g)^{(n)}$ называется *формулой Лейбница.* Ее доказательство проводится по индукции и почти идентично доказательству бинома Ньютона, поэтому мы не будем его приводить.

§ 6. Формула Тейлора

В этом параграфе мы решим задачу о существовании многочлена Тейлора, которая была поставлена в § 1. Будет доказано, что

$$T_{a,n}f(x) = f(a) + \frac{f'(a)}{1!}\,(x-a) + \ldots + \frac{f^{(n)}(a)}{n!}\,(x-a)^n$$

(при определенных условиях на функцию f, которые обсуждаются ниже). Кроме того, мы в различных ситуациях изучим величину

$$R_{a,n}f(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k\,,$$

называемую *остатком в формуле Тейлора.* В заключение будут получены тейлоровские разложения для основных элементарных функций.

Пусть $a \in \mathbb{R}$, $n \in \mathbb{N}$. Рассмотрим вначале случай, когда функция есть многочлен степени не выше n. Обозначим его через p и запишем в виде

$$p(x) = \sum_{k=0}^{n} c_k\,(x-a)^k, \text{ где } c_0, \ldots, c_n \in \mathbb{R}.$$

Ясно, что $c_0 = p(a)$. Как выразить остальные коэффициенты c_k в терминах p? Ответ на этот вопрос дает следующая теорема.

Теорема 1. Формула Тейлора для многочленов. *Пусть $n \in \mathbb{Z}_+$, p — многочлен степени не выше n. Тогда при любых $a, x \in \mathbb{R}$*

$$p(x) = \sum_{k=0}^{n} \frac{p^{(k)}(a)}{k!}\,(x-a)^k.$$

Доказательство. Покажем вначале, что

$$\left((x-a)^k\right)^{(m)}\Big|_{x=a} = \begin{cases} 0, & \text{если} \quad m \neq k \\ m!, & \text{если} \quad m = k. \end{cases}$$

Действительно, при $m \leqslant k$ мы получим

$$\left((x-a)^k\right)^{(m)}\Big|_{x=a} =$$
$$= k(k-1)\ldots(k-m+1)(x-a)^{k-m}\Big|_{x=a} = \begin{cases} 0, & \text{если} \quad m < k \\ m!, & \text{если} \quad m = k. \end{cases}$$

Если же $m > k$, то $\left((x-a)^k\right)^{(m)} = 0$ при всех x, поскольку функция $\left((x-a)^k\right)^{(k)}$ постоянна.

Докажем теперь утверждение теоремы. Разложив многочлен p по степеням $x - a$, представим его в виде $p(x) = \sum\limits_{k=0}^{n} c_k\,(x-a)^k$. Тогда

$$p^{(m)}(a) = \sum_{k=1}^{n} c_k \left((x-a)^k\right)^{(m)}\Big|_{x=a} = c_m \cdot m!$$

при всех $m = 1, \ldots, n$, откуда $c_m = \dfrac{p^{(m)}(a)}{m!}$. $\square$

Таким образом,

$$T_{a,n}p(x) = p(x) = \sum_{k=0}^{n} \frac{p^{(k)}(a)}{k!}\,(x-a)^k$$

при всех $x \in \mathbb{R}$. Полученное соотношение наводит на мысль, что и для функций f более общего вида должно выполняться равенство $T_{a,n}f(x) = \sum\limits_{k=0}^{n} \dfrac{f^{(k)}(a)}{k!}\,(x-a)^k$. Чтобы его правая часть имела смысл, нужно потребовать n-кратную дифференцируемость f в точке a. Оказывается, что при данном условии многочлен Тейлора функции f действительно существует и вычисляется по указанной формуле. В этом заключается *локальный вариант формулы Тейлора.* Для его доказательства нам потребуется следующая лемма.

Лемма 1. *Пусть $E \subset \mathbb{R}$, $a \in E$, $g\colon E \to \mathbb{R}$, $n \in \mathbb{N}$. Предположим, что функция g n раз дифференцируема в точке a и*

$$g(a) = g'(a) = \ldots = g^{(n)}(a) = 0.$$

Тогда $g(x) = o\left((x-a)^n\right)$ при $x \to a$.

Доказательство. Докажем лемму индукцией по n.

База индукции. Пусть $n = 1$. Так как $g(a) = g'(a) = 0$, то по определению дифференцируемости

$$g(x) = g(a) + g'(a)(x-a) + o(x-a) = o(x-a) \text{ при } x \to a.$$

Индукционный переход. Пусть для $n \in \mathbb{N}$ лемма доказана. Докажем ее для $n+1$. Предположим, что g $n+1$ раз дифференцируема в точке a и $g(a) = g'(a) = \ldots = g^{(n+1)}(a) = 0$. Тогда $g'(a) = \ldots = (g')^{(n)}(a) = 0$, и, применяя индукционное предположение к функции g', мы получим равенство

$$g'(x) = o\left((x-a)^n\right) \quad (x \to a).$$

Существует такое $\sigma > 0$, что $E \cap (a-\sigma, a+\sigma)$ является промежутком и функция g n раз дифференцируема на нем. По определению o найдется $\delta \in (0, \sigma)$:

$$|g'(x)| \leqslant \varepsilon\,|x-a|^n \text{ при } x \in E,\ |x-a| < \delta.$$

Пусть $x \in E$, $|x-a| < \delta$. По формуле конечных приращений (см. § 4) существует $\theta \in (0,1)$, для которого

$$g(x) = g(x) - g(a) = g'\bigl(a + \theta\,(x-a)\bigr) \cdot (x-a).$$

Заметим, что $|a + \theta\,(x-a) - a| = \theta\,|x-a| < \theta\delta < \delta$, откуда

$$|g'\bigl(a + \theta\,(x-a)\bigr)| \leqslant \varepsilon\,(\theta\,|x-a|)^n \leqslant \varepsilon\,|x-a|^n.$$

Тогда

$$|g(x)| = |g(x) - g(a)| = |g'\bigl(a + \theta\,(x-a)\bigr)| \cdot |x-a| \leqslant \varepsilon\,|x-a|^{n+1}.$$

Поэтому $g(x) = o\bigl((x-a)^{n+1}\bigr)$ при $x \to a$, и лемма доказана. $\square$

Перейдем к локальному варианту формулы Тейлора.

Теорема 2. Формула Тейлора – Пеано. *Пусть $E \subset \mathbb{R}$, $a \in E$, $f\colon E \to \mathbb{R}$, $n \in \mathbb{N}$. Предположим, что f n раз дифференцируема в точке a. Тогда*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k + o\,((x-a)^n) \quad (x \to a).$$

Доказательство. Положим

$$p(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k.$$

По теореме 1 $p(x)$ также равно $\sum\limits_{k=0}^{n} \dfrac{p^{(k)}(a)}{k!}\,(x-a)^k$. Приравнивая коэффициенты этих многочленов при $(x-a)^k$, мы получим

$$p^{(k)}(a) = f^{(k)}(a) \ \text{ для всех } \ k = 0, \ldots, n.$$

Положим $g(x) = f(x) - p(x)$. Тогда $g(a) = g'(a) = \ldots = g^{(n)}(a) = 0$. По лемме 1 $g(x) = o\,((x-a)^n)$ при $x \to a$, то есть

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k + o\,((x-a)^n) \quad (x \to a). \quad \square$$

Замечания

1) Если $a = 0$, то полученное равенство называют еще *формулой Маклорена.*

2) Формулу Тейлора – Пеано можно также записать с помощью дифференциалов старших порядков:

$$f(x) = \sum_{k=0}^{n} \frac{d_a^k f(x-a)}{k!} + o\left((x-a)^n\right).$$

Здесь под $d_a^0 f$ мы понимаем функцию, тождественно равную $f(a)$.

Теорема 2 утверждает, что $T_{a,n}f(x)$ “хорошо приближает” $f(x)$ при $x \approx a$. Однако формула Тейлора – Пеано не дает никакой оценки погрешности $R_{a,n}f(x)$ такой аппроксимации при конкретном x, что делает ее непригодной для приближенных вычислений. Мы получим формулу для $R_{a,n}f(x)$, из которой можно будет судить о малости остатка.

Пусть $a, x \in \mathbb{R}$, $a \neq x$. Обозначим через $\Delta_{a,x}$ отрезок с концами a и x, а через $\widetilde{\Delta}_{a,x}$ — интервал с концами a и x.

Теорема 3. Формула Тейлора – Лагранжа. *Пусть $n \in \mathbb{Z}_+$, функция f $n+1$ раз дифференцируема на $\langle A, B\rangle$, $a, x \in \langle A, B\rangle$, $a \neq x$. Тогда найдется точка $c \in \widetilde{\Delta}_{a,x}$, для которой*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Замечания

1) Остаток $R_{a,n}f(x) = f(x) - T_{a,n}f(x)$ в формуле Тейлора – Лагранжа напоминает по виду слагаемое с номером $n+1$ в многочлене Тейлора, но $f^{(n+1)}$ вычисляется не в точке a, а в некоторой точке c, лежащей между a и x. Следует отметить, что эта точка c *зависит от* x, поэтому остаток, вообще говоря, не есть многочлен.

2) Утверждение теоремы можно переписать в следующем виде: *существует такое $\theta \in (0,1)$, что*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}\left(a + \theta\,(x-a)\right)}{(n+1)!}(x-a)^{n+1}.$$

3) Условия теоремы 3 можно ослабить. Достаточно было бы предположить, что $f \in C^n(\Delta_{a,x})$ и $f^{(n)}$ дифференцируема на $\widetilde{\Delta}_{a,x}$. Это будет ясно из доказательства.

Доказательство. Для $t \in \Delta_{a,x}$ положим

$$\varphi(t) = (x-t)^{n+1}, \quad F(t) = R_{t,n}f(x) = f(x)-f(t)-\sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k.$$

Тогда при $t \in \widetilde{\Delta}_{a,x}$

$$\begin{aligned}
F'(t) &= -f'(t) - \sum_{k=1}^{n}\left(\frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(k)}(t)}{k!}k(x-t)^{k-1}\right) = \\
&= -f'(t) + \sum_{k=1}^{n}\frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} - \sum_{k=1}^{n}\frac{f^{(k+1)}(t)}{k!}(x-t)^k = \\
&= -f'(t) + \sum_{m=0}^{n-1}\frac{f^{(m+1)}(t)}{m!}(x-t)^m - \sum_{k=1}^{n}\frac{f^{(k+1)}(t)}{k!}(x-t)^k = \\
&= -f'(t) + f'(t) - \frac{f^{(n+1)}(t)}{n!}(x-t)^n = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n
\end{aligned}$$

(после замены индекса $m = k-1$ в первой сумме становится очевидным, что она отличается от второй одним слагаемым). Применяя к функциям F и φ теорему Коши о среднем (см. § 4), мы найдем точку $c \in \widetilde{\Delta}_{a,x}$, для которой

$$\begin{aligned}
\frac{F(a)}{(x-a)^{n+1}} = \frac{F(a)-F(x)}{\varphi(a)-\varphi(x)} = \frac{F'(c)}{\varphi'(c)} &= \\
&= \frac{-f^{(n+1)}(c)(x-c)^n}{-n!(n+1)(x-c)^n} = \frac{f^{(n+1)}(c)}{(n+1)!}.
\end{aligned}$$

Домножая обе части равенства на $(x-a)^{n+1}$, мы получим искомую формулу. $\square$

Замечание. Теорема 2 будет следовать из формулы Тейлора – Лагранжа, если в формулировке теоремы наложить на функцию f следующее более сильное условие: *существует такое $\delta > 0$, что*

f n раз дифференцируема на $E \cap (a-\delta, a+\delta)$ и $f^{(n)}$ непрерывна в точке a. Действительно, пусть $x \in E \cap (a-\delta, a+\delta)$. В силу условия (+) из § 1 можно считать, что $\Delta_{a,x} \subset E$. По теореме 3

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}\big(c(x)\big)}{n!}(x-a)^n =$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}\big(c(x)\big) - f^{(n)}(a)}{n!}(x-a)^n,$$

где точка $c(x)$ лежит между a и x. Так как $|c(x)-a| < |x-a|$, то $c(x) \to a$ при $x \to a$, а тогда $f^{(n)}\big(c(x)\big) \to f^{(n)}(a)$ ввиду непрерывности $f^{(n)}$ в точке a. Поэтому

$$R_{a,n}f(x) = \frac{f^{(n)}\big(c(x)\big) - f^{(n)}(a)}{n!}(x-a)^n = o\big((x-a)^n\big) \quad (x \to a). \quad \square$$

Выведем теперь разложения Тейлора – Пеано для простейших элементарных функций.

Теорема 4. *Пусть $n \in \mathbb{Z}_+$. При $x \to 0$ справедливы следующие утверждения.*

1) $e^x = \sum\limits_{k=0}^{n} \dfrac{x^k}{k!} + o\big(x^n\big) = 1 + \dfrac{x}{1!} + \dfrac{x^2}{2!} + \ldots + \dfrac{x^n}{n!} + o\big(x^n\big).$

2) $\sin x = \sum\limits_{k=0}^{n} (-1)^k \dfrac{x^{2k+1}}{(2k+1)!} + o\big(x^{2n+2}\big) =$

$$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o\big(x^{2n+2}\big).$$

3) $\cos x = \sum\limits_{k=0}^{n} (-1)^k \dfrac{x^{2k}}{(2k)!} + o\big(x^{2n+1}\big) =$

$$= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + o\big(x^{2n+1}\big).$$

4) $\ln(1+x) = \sum\limits_{k=1}^{n} (-1)^{k+1} \dfrac{x^k}{k} + o\big(x^n\big) =$

$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{n+1} \frac{x^n}{n} + o\big(x^n\big).$$

5) *Пусть* $\alpha \in \mathbb{R}$. *Положим* $C_\alpha^k = \dfrac{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-k+1)}{k!}$ *при* $k \in \mathbb{Z}_+$. *Тогда*

$$(1+x)^\alpha = \sum_{k=0}^{n} C_\alpha^k x^k + o\bigl(x^n\bigr) = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \ldots + \\ + \frac{\alpha(\alpha-1)\cdot\ldots\cdot(\alpha-n+1)}{n!}x^n + o\bigl(x^n\bigr).$$

Замечания

1) Если $\alpha \in \mathbb{Z}_+$ и $\alpha \geqslant k$, то C_α^k является биномиальным коэффициентом, что и оправдывает выбранное обозначение. Величины C_α^k определены при любых $\alpha \in \mathbb{R}$ и называются *обобщенными биномиальными коэффициентами.*

2) При $\alpha \in \mathbb{N}$ функция $f(x) = (1+x)^\alpha$ является многочленом степени α, поэтому $f \equiv T_{0,\alpha}f$. В силу предыдущего замечания это тождество есть не что иное, как бином Ньютона. Поэтому формулу Тейлора – Пеано для $(1+x)^\alpha$ называют *биномиальным разложением.*

Доказательство.

1) Если $f(x) = e^x$, то $f^{(k)}(x) = e^x$ и $f^{(k)}(0) = 1$ для любых $k \in \mathbb{N}$ и $x \in \mathbb{R}$. Поэтому $T_{0,n}f(x) = \sum\limits_{k=0}^{n} \dfrac{x^k}{k!}$, и требуемое равенство получается из теоремы 2.

2) Пусть $f(x) = \sin x$. Докажем вначале, что

$$f^{(m)}(x) \equiv \sin\bigl(x + \tfrac{\pi m}{2}\bigr) \quad \text{для любых} \quad m \in \mathbb{Z}_+.$$

При $m = 0$ формула очевидна. Пусть для некоторого $m \in \mathbb{Z}_+$ она уже доказана. Тогда

$$f^{(m+1)}(x) = \bigl(\sin\bigl(x + \tfrac{\pi m}{2}\bigr)\bigr)' = \cos\bigl(x + \tfrac{\pi m}{2}\bigr) = \sin\bigl(x + \tfrac{\pi m}{2} + \tfrac{\pi}{2}\bigr),$$

и индукционный переход завершен. Теперь $f^{(m)}(0) = \sin\frac{\pi m}{2}$, откуда

$$f^{(2k)}(0) = 0 \quad \text{и} \quad f^{(2k+1)}(0) = (-1)^k \ \ (k \in \mathbb{Z}_+).$$

Поэтому

$$T_{0,2n+1}f(x) = T_{0,2n+2}f(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!},$$

что и доказывает 2).

3) Предлагаем читателю провести доказательство самостоятельно. *Указание:* воспользуйтесь тождеством $\left(\cos x\right)^{(m)} \equiv \cos\left(x + \frac{\pi m}{2}\right)$ $(m \in \mathbb{Z}_+)$, которое получается путем дифференцирования формулы для $\left(\sin x\right)^{(m)}$.

4) Пусть $f(x) = \ln(1+x)$. Докажем вначале по индукции, что при всех $k \in \mathbb{N}$

$$f^{(k)}(x) = (-1)^{k+1} \frac{(k-1)!}{(x+1)^k} \quad (x > -1).$$

Для $k = 1$ утверждение очевидно. Пусть для некоторого $k \in \mathbb{N}$ данная формула верна. Тогда

$$f^{(k+1)}(x) = \left(f^{(k)}\right)'(x) =$$
$$= (-1)^{k+1} (k-1)! \frac{(-k)}{(x+1)^{k+1}} = (-1)^{k+2} \frac{k!}{(x+1)^{k+1}},$$

то есть равенство верно и для $k+1$. Поэтому

$$f^{(k)}(0) = (-1)^{k+1} (k-1)! \text{ при } k \in \mathbb{N}, \quad f(0) = 0,$$

откуда $T_{0,n}f(x) = \sum\limits_{k=1}^{n} (-1)^{k+1} \dfrac{x^k}{k}$.

5) Пусть $\alpha \in \mathbb{R}$, $f(x) = (1+x)^\alpha$. Тогда

$$f(0) = 1 \quad \text{и} \quad f^{(k)}(x) = \alpha(\alpha-1)\ldots(\alpha-k+1)(1+x)^{\alpha-k} \quad (k \in \mathbb{N}).$$

Поэтому $T_{0,n}f(x) = \sum\limits_{k=0}^{n} \dfrac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!} x^k$. $\square$

Формула Тейлора – Пеано часто используется для раскрытия неопределенностей. Отметим, что асимптотические равенства, полученные в главе 2, являются частными случаями разложений Тейлора – Пеано. Например, соотношение $\lim\limits_{x\to 0}\dfrac{\ln(1+x)}{x}=1$ эквивалентно $\ln(1+x)=x+o(x)\ \ (x\to 0)$, а это есть четвертое утверждение теоремы 4 при $n=1$. Увеличивая n, мы получим более точные разложения. Для примера вычислим $\lim\limits_{x\to 0}\dfrac{\ln(1+x)-\sqrt{1+2x}}{x^3}$. Пользуясь соотношениями 4 и 5 теоремы 4, запишем формулы для $\ln(1+x)$ и $\sqrt{1+2x}$ при $n=1,2,3$:

$$
\begin{aligned}
&\ln(1+x)=x+o(x), && \sqrt{1+2x}=x+o(x);\\
&\ln(1+x)=x-\frac{x^2}{2}+o\big(x^2\big), && \sqrt{1+2x}=x-\frac{x^2}{2}+o\big(x^2\big);\\
&\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}+o\big(x^3\big), && \sqrt{1+2x}=x-\frac{x^2}{2}+\frac{x^3}{2}+o\big(x^3\big).
\end{aligned}
$$

Подстановка этих равенств при $n=1$ и $n=2$ под знак предела дает нам соответственно $\lim\limits_{x\to 0}\dfrac{o(x)}{x^3}$ и $\lim\limits_{x\to 0}\dfrac{o\big(x^2\big)}{x^3}$, то есть неопределенность $\frac{0}{0}$ в обоих случаях сохраняется. Если же взять $n=3$, то мы получим

$$
\lim_{x\to 0}\frac{\ln(1+x)-\sqrt{1+2x}}{x^3}=\lim_{x\to 0}\frac{-\frac{x^3}{6}+o\big(x^3\big)}{x^3}=-\frac{1}{6}.
$$

Использовать разложения с бóльшим n здесь не имеет смысла, это лишь усложнит вычисления.

Как уже отмечалось, для приближенного вычисления значений функции по формуле Тейлора важно знать величину остатка. Найти его точно мы не можем, поскольку не знаем “среднюю” точку c из теоремы 3. Для *оценки* остатка используется следующее утверждение.

Лемма 2. *Пусть функция f удовлетворяет условиям теоремы 3. Предположим, что существует такое число $M>0$, что $\big|f^{(n+1)}(t)\big|\leqslant M$ для любого $t\in\widetilde{\Delta}_{a,x}$. Тогда*

$$
|R_{a,n}f(x)|\leqslant\frac{M\,|x-a|^{n+1}}{(n+1)!}.
$$

Доказательство. По формуле Тейлора – Лагранжа

$$R_{a,n}f(x) = \frac{f^{(n+1)}(c)}{(n+1)!}\,(x-a)^{n+1}$$

при некотором $c \in \widetilde{\Delta}_{a,x}$. Утверждение леммы вытекает из неравенства $\left|f^{(n+1)}(c)\right| \leqslant M$. $\square$

Несмотря на свою простоту, лемма 2 является эффективным инструментом для оценки остатка в формуле Тейлора. Проиллюстрируем это на примерах. Пусть $n \in \mathbb{Z}_+$ и $x \in \mathbb{R}$.

1) *Если* $f(t) = e^t$, *то* $\left|R_{0,n}f(x)\right| \leqslant \dfrac{e^{\max\{x,0\}} \cdot |x|^{n+1}}{(n+1)!}$.

Действительно, $\sup\limits_{t\in\widetilde{\Delta}_{0,x}} \left|f^{(n+1)}(t)\right| = \sup\limits_{t\in\widetilde{\Delta}_{0,x}} e^t = e^{\max\{x,0\}}$. $\square$

2) *Если* $f(t) = \sin t$ *или* $f(t) = \cos t$, *то* $\left|R_{0,n}f(x)\right| \leqslant \dfrac{|x|^{n+1}}{(n+1)!}$.

Для $f(t) = \sin t$ мы получим

$$\left|f^{(n+1)}(t)\right| = \left|\cos\left(t + \frac{\pi n}{2}\right)\right| \leqslant 1 \quad (t \in \mathbb{R}),$$

так что утверждение непосредственно следует из леммы 2. Случай $f(t) = \cos t$ разбирается аналогично. $\square$

Замечание. В главе 2 отмечалось, что $\lim\limits_{n\to\infty} \dfrac{|x|^n}{n!} = 0$ при любом $x \in \mathbb{R}$. Поэтому для функций f, рассмотренных в примерах 1 и 2, $R_{0,n}f(x) \to 0$ при $n \to \infty$, то есть $\lim\limits_{n\to\infty} T_{0,n}f(x) = f(x)$. Этот факт окажется полезным в теории степенных рядов.

В заключение обоснуем некоторые свойства числа Непера e, сформулированные без доказательства в § 2 главы 2.

Теорема 5.

1) *Для любого* $n \in \mathbb{N}$ *существует такое* $c \in (0,1)$, *что*

$$e = 1 + \frac{1}{1!} + \ldots + \frac{1}{n!} + \frac{e^c}{(n+1)!}.$$

2) *Число* e *иррационально.*

Доказательство.

1) Пусть $f(x) = e^x$. Записывая $f(1)$ по формуле Тейлора – Лагранжа, мы получим требуемое равенство.

2) В главе 2 было доказано, что $e \in (2,3)$. Предположим, что e рационально. Тогда $e = \frac{m}{n}$ при некоторых $m, n \in \mathbb{N}$, причем $n \geqslant 2$, поскольку $e \notin \mathbb{Z}$. В силу утверждения 1 существует такое $c \in (0,1)$, что

$$\frac{m}{n} - 1 - \frac{1}{1!} - \ldots - \frac{1}{n!} = \frac{e^c}{(n+1)!}\,.$$

Умножая это равенство на $n!$, мы получим

$$\frac{e^c}{n+1} = m\,(n-1)! - n! - \frac{n!}{1!} - \ldots - \frac{n!}{n!} \in \mathbb{Z}\,.$$

Поэтому $\dfrac{e^c}{n+1} \in \mathbb{N}$ и, значит, $e^c \geqslant n+1$. Но это невозможно, так как $n+1 \geqslant 3$, а $e^c < e < 3$. Полученное противоречие завершает доказательство. $\square$

Замечание. Из утверждения 1 теоремы 5 вытекает двойное неравенство

$$0 < e - 1 - \frac{1}{1!} - \ldots - \frac{1}{n!} < \frac{3}{(n+1)!}\,.$$

Иными словами, справедливо приближенное равенство

$$e \approx 1 + \frac{1}{1!} + \ldots + \frac{1}{n!}\,,$$

причем абсолютная погрешность такой аппроксимации меньше, чем $\dfrac{3}{(n+1)!}$. Так как последовательность $\dfrac{3}{(n+1)!}$ быстро стремится к нулю при $n \to \infty$, указанная формула позволяет приближенно вычислить число e с любой требуемой точностью.

§ 7. Монотонность и экстремумы функций

Важным приложением дифференциального исчисления является исследование функций и построение их графиков. Поэтому оставшаяся часть главы посвящена изучению геометрических свойств функций. Вначале мы разберем понятия *монотонности* и *экстремума* функции, тесно связанные с первой производной. Грубо говоря, характер монотонности определяется по знаку производной, а в точках экстремума производная меняет знак. Ниже мы придадим этим высказываниям точный смысл. Понятия *выпуклости* и *вогнутости* функции, за которые отвечает вторая производная, будут подробно рассмотрены в следующем параграфе.

Наиболее важным приложением первой производной является проверка монотонности на промежутке дифференцируемой функции. Сформулируем утверждение, позволяющее осуществлять эту проверку.

Теорема 1. Условия монотонности функции на промежутке. *Пусть функция f непрерывна на $\langle A, B\rangle$ и дифференцируема на (A, B). Справедливы следующие утверждения.*

1) *Функция f возрастает на $\langle A, B\rangle$ тогда и только тогда, когда $f'(x) \geqslant 0$ для любого $x \in (A, B)$.*

2) *Если $f'(x) > 0$ при всех $x \in (A, B)$, то f строго возрастает на $\langle A, B\rangle$.*

Замечания

1) Первое утверждение теоремы называется *критерием возрастания f на $\langle A, B\rangle$*. Второе утверждение дает лишь *достаточное условие* строгого возрастания f на $\langle A, B\rangle$. Необходимым оно не является. Например, функция $f(x) = x^3$ строго возрастает на $\mathbb{R}$, но $f'(0) = 0$.

2) Поменяв в формулировке теоремы знаки неравенств на противоположные, мы получим соответственно *критерий убывания* и *достаточное условие строгого убывания* f на $\langle A, B\rangle$. Для их доказательства достаточно применить теорему 1 к функции $(-f)$.

Доказательство. Проверим вначале необходимость в утверждении 1. Пусть $a, x \in (A, B)$, $x \neq a$. Если $x > a$, то $f(x) \geqslant f(a)$, откуда $\dfrac{f(x) - f(a)}{x - a} \geqslant 0$. Если же $x < a$, то $f(x) \leqslant f(a)$, и снова $\dfrac{f(x) - f(a)}{x - a} \geqslant 0$. Поэтому

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \geqslant 0 \quad \text{при любом } \ a \in (A, B).$$

Достаточность в первом утверждении и второе утверждение докажем параллельно. Пусть $a, b \in \langle A, B\rangle$, $b > a$. Воспользуемся формулой конечных приращений (см. § 4). Существует такое $c \in (a, b)$, что $f(b) - f(a) = f'(c)\,(b - a)$. В условиях первого утверждения $f'(c) \geqslant 0$, откуда $f(b) \geqslant f(a)$, и мы получаем возрастание f на $\langle A, B\rangle$. Из условий второго утверждения вытекает неравенство $f'(c) > 0$, откуда $f(b) > f(a)$, то есть f строго возрастает на $\langle A, B\rangle$. $\square$

Следствие 1. **Критерий постоянства функции.** *Пусть функция f непрерывна на $\langle A, B\rangle$ и дифференцируема на (A, B). Тогда равносильны следующие утверждения.*

1) *f постоянна на $\langle A, B\rangle$.*

2) *$f'(x) = 0$ при всех $x \in (A, B)$.*

Доказательство. Импликация 1) $\Rightarrow$ 2) очевидна. Докажем переход 2) $\Rightarrow$ 1). Пусть $f'(x) = 0$ для любого $x \in (A, B)$. Тогда при всех $x \in (A, B)$ выполняются неравенства $f'(x) \geqslant 0$ и $f'(x) \leqslant 0$. Первое из них по теореме 1 дает возрастание, а второе — убывание f на $\langle A, B\rangle$. Поэтому f постоянна на $\langle A, B\rangle$. $\square$

Следствие 2. *Пусть функции f и g непрерывны на $[A, B\rangle$ и дифференцируемы на (A, B). Если $f(A) = g(A)$ и $f'(x) > g'(x)$ при всех $x \in (A, B)$, то $f(x) > g(x)$ при любом $x \in (A, B\rangle$.*

Доказательство. Пусть $h = f - g$. Так как функция h удовлетворяет условиям второго утверждения теоремы 1, то она строго возрастает на $[A, B\rangle$. Поэтому при любом $x \in (A, B\rangle$ выполняется неравенство $h(x) > h(A) = 0$, откуда $f(x) > g(x)$. $\square$

Предлагаем читателю сформулировать аналогичное утверждение для случая, когда f и g совпадают не на левом, а на правом конце промежутка.

Следствие 2 используется для доказательства неравенств. Разберем два полезных примера.

1) $\cos x > 1 - \dfrac{x^2}{2}$ *при всех* $x > 0$.

Действительно, положим $f(x) = \cos x$, $g(x) = 1 - \dfrac{x^2}{2}$. Тогда

$$f'(x) - g'(x) = x - \sin x > 0 \quad \text{при} \;\; x > 0$$

(см. § 5 главы 2). Применяя следствие 2 к промежутку $[0,+\infty)$, мы получим требуемое неравенство. $\square$

2) $\sin x > x - \dfrac{x^3}{6}$ *при всех* $x > 0$.

Положим $f(x) = \sin x$, $g(x) = x - \dfrac{x^3}{6}$. Тогда

$$f'(x) - g'(x) = \cos x - 1 + \frac{x^2}{2} > 0 \quad \text{при} \;\; x > 0$$

(см. предыдущий пример). Осталось применить следствие 2 к промежутку $[0,+\infty)$. $\square$

Другой важной геометрической характеристикой графика функции является *локальный экстремум*, или просто *экстремум*. Перейдем к изучению этого понятия.

Определение 1. Экстремум функции. Предположим, что $E \subset \mathbb{R}$, $f\colon E \to \mathbb{R}$, $a \in E$.

1) Пусть существует $\delta > 0$: при всех $x \in (a-\delta, a+\delta) \cap E$ выполняется неравенство $f(x) \geqslant f(a)$. Тогда a называется *точкой минимума* f. Если же $f(x) \leqslant f(a)$ при всех $x \in (a-\delta, a+\delta) \cap E$, то a называется *точкой максимума* f.

2) Пусть существует $\delta > 0$: при всех $x \in (a-\delta, a+\delta) \cap E \setminus \{a\}$ выполняется неравенство $f(x) > f(a)$. Тогда a называется *точкой строгого минимума* f. Если же при всех $x \in (a-\delta, a+\delta) \cap E \setminus \{a\}$ $f(x) > f(a)$, то a называется *точкой строгого максимума* f.

3) Если a является точкой минимума или максимума функции f, то a называется *точкой экстремума* f.

Не следует думать, что в точке минимума реализуется наименьшее значение функции. Оно будет наименьшим лишь *локально*, то есть в некоторой окрестности точки минимума. Рассмотрим для примера функцию $f(x) = x^2\,(x+1)$ и точку $a = 0$. Ясно, что $f(x) > f(0)$ при условии $0 < |x| < 1$, поэтому a является точкой строгого минимума. Тем не менее $\inf\limits_{\mathbb{R}} f = -\infty$, поскольку $\lim\limits_{x \to -\infty} f(x) = -\infty$. Аналогичное замечание справедливо и для точек максимума.

В определении 1 на функцию f не накладывалось никаких ограничений. В дальнейшем мы обычно будем предполагать, что f

дифференцируема в точке a, поскольку условия экстремума (см. теоремы 2 – 5) записываются в терминах производных. Кроме того, поскольку понятие экстремума является локальным, мы можем ограничиться рассмотрением функций, заданных на промежутке (это объяснялось в § 1).

Обратимся теперь к задаче *поиска* точек экстремума. Решение этой задачи производится в два этапа. Вначале с помощью *необходимого условия экстремума* мы отбрасываем точки, в которых экстремума заведомо не может быть. Затем оставшиеся точки исследуются с помощью *достаточного условия экстремума*. Перейдем к реализации этой схемы.

Теорема 2. Необходимое условие экстремума. *Пусть* $f\colon \langle A,B\rangle \to \mathbb{R}$, $a \in (A,B)$, f *дифференцируема в точке* a. *Если* a *является точкой экстремума* f, *то* $f'(a) = 0$.

Точка a, удовлетворяющая условию $f'(a) = 0$, называется *стационарной* для функции f.

Доказательство. Будем для определенности считать, что a — точка минимума f. Тогда найдется такое $\delta > 0$, что

$$[a-\delta, a+\delta] \subset (A,B) \quad \text{и} \quad f(x) \geqslant f(a) \ \text{при всех} \ x \in [a-\delta, a+\delta].$$

Применяя к функции $f\big|_{[a-\delta,a+\delta]}$ теорему Ферма (см. § 4), мы получим равенство $f'(a) = 0$. $\square$

Замечания

1) Условие стационарности точки не гарантирует, что в ней есть экстремум. Например, для $f(x) = x^3$ точка 0 является стационарной, но f не имеет экстремума в нуле. Таким образом, теорема 2 не дает достаточного условия экстремума.

2) Экстремумы могут быть в точках, где функция не дифференцируема. Так, функция $f(x) = \sqrt[3]{x^2}$ имеет минимум в точке 0, но не дифференцируема в ней.

3) Утверждение теоремы неверно для концевых экстремумов. Например, функция f, определенная на отрезке $[0,1]$ формулой $f(x) = x$, имеет минимум в точке 0, но $f'(0) = 1$.

4) Назовем точку $a \in (A,B)$ *критической* (или *подозрительной на экстремум*), если либо a стационарна для f, либо f не диффе-

ренцируема в точке a. Теорема 2 утверждает, что все принадлежащие (A, B) точки экстремума f лежат в множестве ее критических точек.

Важным приложением теоремы 2 является *задача о наибольшем и наименьшем значениях* функции на отрезке. Пусть f — непрерывная на $[A, B]$ функция. По теореме Вейерштрасса (см. § 4 главы 2) f достигает на $[A, B]$ наибольшего и наименьшего значений. Для их нахождения необходимо выполнить следующие действия.

1) Найти множество C критических точек f.

2) Вычислить величины

$$M = \max\{f(A), f(B), \max_C f\} \ \text{ и } \ m = \min\{f(A), f(B), \min_C f\}.$$

Тогда $M = \max\limits_{[A,B]} f$ и $m = \min\limits_{[A,B]} f$. Действительно, если $\max\limits_{[A,B]} f = f(a)$, то a является точкой максимума функции f. В силу замечания 4 либо $a \in \{A, B\}$, либо $a \in C$. Поэтому $M \geqslant \max\limits_{[A,B]} f$, а обратное неравенство очевидно. Рассуждения для m проводятся аналогично. Если множество C конечно (что обычно и бывает в практических приложениях), то вычисление M и m не представляет трудностей.

В задаче о наибольшем и наименьшем значениях не требовалось искать точки экстремума, достаточно было знать критические точки функции. Если же мы хотим построить график функции, необходимо знать точки экстремума и их тип. Рассмотрим *достаточные условия* экстремума, позволяющие исследовать критические точки. Первое условие основано на изучении знака производной в окрестности точки.

Теорема 3. Достаточное условие экстремума в терминах первой производной. *Пусть* $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in (A, B)$, *и* $\delta > 0$ *таково, что* $(a-\delta, a+\delta) \subset \langle A, B\rangle$. *Предположим, что функция* f *непрерывна в точке* a *и дифференцируема на* $(a-\delta, a) \cup (a, a+\delta)$.

1) *Если* $f'(x) < 0$ *при* $x \in (a-\delta, a)$ *и* $f'(x) > 0$ *при* $x \in (a, a+\delta)$, *то* a *— точка строгого минимума* f.

2) *Если* $f'(x) > 0$ *при* $x \in (a-\delta, a)$ *и* $f'(x) < 0$ *при* $x \in (a, a+\delta)$, *то* a *— точка строгого максимума* f.

Замечание. Если функция f' не меняет знак на множестве $(a-\delta, a) \cup (a, a+\delta)$, то f не имеет экстремума в точке a. Действительно, в этом случае функция f строго монотонна на $(a-\delta, a+\delta)$ по теореме 1.

Доказательство. Докажем только первое утверждение. По теореме 1 функция f строго убывает на $(a-\delta, a]$ и строго возрастает на $[a, a+\delta)$. Тогда $f(x) > f(a)$ при $x \in (a-\delta, a)$ и при $x \in (a, a+\delta)$. Поэтому a — точка строгого минимума f. $\square$

Теорема 3 дает лишь достаточное условие экстремума, оно *не является необходимым.* Поясним это на примере. Положим

$$f(x) = x^2\left(2+\sin\tfrac{1}{x}\right) \text{ при } x \neq 0 \quad \text{и} \quad f(0) = 0.$$

Функция f дифференцируема на $\mathbb{R}$ (см. § 2). Очевидно, что 0 — точка строгого минимума f. Заметим, однако, что

$$f'(x) = 2x\left(2+\sin\tfrac{1}{x}\right) - \cos\tfrac{1}{x} \text{ при } x \neq 0.$$

Так как $2x\left(2+\sin\frac{1}{x}\right) \to 0$ при $x \to 0$, а $\cos\frac{1}{x}$ сколь угодно близко к нулю принимает значения 1 и -1, то f' не сохраняет знак ни в одной окрестности точки 0.

Другое достаточное условие, использующее вторую производную функции, получается из формулы Тейлора.

Теорема 4. Достаточное условие экстремума в терминах второй производной. *Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in (A, B)$. Предположим, что функция f дважды дифференцируема в точке a, причем $f'(a) = 0$. Тогда справедливы следующие утверждения.*

1) *Если $f''(a) > 0$, то a — точка строгого минимума f.*

2) *Если $f''(a) < 0$, то a — точка строгого максимума f.*

Доказательство. Ограничимся проверкой первого утверждения. Применяя формулу Тейлора – Пеано (см. теорему 2 § 6) и используя условие $f'(a) = 0$, мы получим

$$\begin{aligned} f(x) - f(a) = f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + o\big((x-a)^2\big) &= \\ = \frac{f''(a)}{2}(x-a)^2\big(1+o(1)\big) \quad (x \to a)&. \end{aligned}$$

Так как $1+o(1) \to 1$ при $x \to a$, то найдется такое $\delta > 0$, что

$$(a-\delta, a+\delta) \subset (A, B) \quad \text{и} \quad 1+o(1) > 0 \text{ при всех } x \in (a-\delta, a+\delta).$$

Тогда $f(x) > f(a)$ для любого $x \in (a-\delta, a+\delta) \setminus \{a\}$, то есть a — точка строгого минимума f. □

Замечание. Если $f''(a) = 0$, то теорема 4 не дает возможности проверить наличие у f экстремума в точке a. Например, в точке $a = 0$ функция $f(x) = x^4$ имеет строгий минимум, а функция $f(x) = x^3$ не имеет экстремума.

Обобщая метод доказательства теоремы 4, можно получить более общее правило исследования точек с помощью старших производных. Сформулируем его без доказательства.

Теорема 5. Связь экстремума со старшими производными. *Пусть* $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in (A, B)$, $n \in \mathbb{N}$. *Предположим, что функция* f n *раз дифференцируема в точке* a, *причем* $f'(a) = \ldots = f^{(n-1)}(a) = 0$, *а* $f^{(n)}(a) \neq 0$. *Тогда справедливы следующие утверждения.*

1) *Если* n *нечетно, то* f *не имеет экстремума в точке* a.

2) *Пусть* n *четно. Если* $f^{(n)}(a) > 0$, *то* f *имеет в точке* a *строгий минимум, а если* $f^{(n)}(a) < 0$, *то строгий максимум.*

На практике теорема 3 используется чаще, чем теоремы 4 и 5, поскольку ее условия на функцию менее ограничительны и она не требует трудоемкого вычисления старших производных. Недостаток теоремы 3 состоит в том, что ее нельзя обобщить на функции нескольких переменных, тогда как теорема 4 имеет многомерный аналог.

Изложенная выше теория позволяет получить информацию о поведении производной функции, дифференцируемой на промежутке. Мы докажем несколько свойств производной, которые окажутся полезными при изучении неопределенного интеграла. Начнем с теоремы о нулях производной.

Теорема 6 (Г. Дарбу). О нулях производной. *Пусть* f — *дифференцируемая функция на* $\langle A, B\rangle$. *Предположим, что точки* $a, b \in \langle A, B\rangle$ *таковы, что* $f'(a) < 0$, $f'(b) > 0$. *Тогда существует точка* c, *лежащая между* a *и* b, *для которой* $f'(c) = 0$.

Замечание. Если бы функция f' предполагалась непрерывной, то теорема Дарбу была бы следствием теоремы Больцано – Коши (см. § 4 главы 2). В § 5 показано, однако, что f' может быть разрывной. Тем не менее *утверждение* теоремы Больцано – Коши справедливо и для f'.

Доказательство. Пусть для определенности $a < b$. Поскольку f непрерывна на $[a,b]$, по теореме Вейерштрасса найдется такая точка $c \in [a,b]$, что $f(c) = \min\limits_{[a,b]} f$. Если $c \in (a,b)$, то по теореме Ферма мы получим требуемое равенство $f'(c) = 0$. Поэтому достаточно доказать, что *точка* c *лежит в интервале* (a,b). Так как $\lim\limits_{x\to a+} \dfrac{f(x)-f(a)}{x-a} = f'(a) < 0$, то существует такое $\delta \in (0, b-a)$, что при $x \in (a, a+\delta)$ выполняется неравенство $\dfrac{f(x)-f(a)}{x-a} < 0$, то есть $f(x) < f(a)$. Поэтому $\min\limits_{[a,b]} f < f(a)$ и, значит, $c \neq a$. Аналогично показывается, что $c \neq b$. $\square$

Из теоремы Дарбу вытекают несколько полезных следствий.

Следствие 1. Монотонность функции с ненулевой производной. *Пусть f — дифференцируемая функция на $\langle A, B\rangle$. Если $f'(x) \neq 0$ при всех $x \in \langle A, B\rangle$, то f строго монотонна на $\langle A, B\rangle$.*

Это утверждение уточняет условие строгой монотонности функции, сформулированное в теореме 1.

Доказательство. Заметим, что f' сохраняет знак на $\langle A, B\rangle$. Действительно, если бы для некоторых $a, b \in \langle A, B\rangle$ выполнялись неравенства $f'(a) < 0$ и $f'(b) > 0$, то по теореме Дарбу f' обратилась бы в некоторой точке в ноль, что невозможно. Положительность f' на $\langle A, B\rangle$ влечет по теореме 1 строгое возрастание f на $\langle A, B\rangle$, а отрицательность — строгое убывание. $\square$

Следствие 2. О сохранении промежутка. *Пусть f — дифференцируемая функция на $\langle A, B\rangle$. Тогда образ f' есть промежуток.*

Доказательство. Пусть $a, b \in \langle A, B\rangle$. Будем для определенности считать, что $f'(a) \leqslant f'(b)$. Достаточно доказать, что отрезок $[f'(a), f'(b)]$ содержится в образе f' (см. лемму 1 § 4 главы 2). Пусть $y \in \big(f'(a), f'(b)\big)$. Рассмотрим функцию $F(x) = f(x) - y \cdot x$. Так как $F'(x) = f'(x) - y$, то $F'(a) < 0$ и $F'(b) > 0$. По теореме Дарбу найдется такое $c \in \langle A, B\rangle$, что $F'(c) = 0$, то есть $f'(c) = y$. Поэтому y лежит в образе f'. $\square$

Следствие 3. О скачках производной. *Пусть f — дифференцируемая функция на $\langle A, B\rangle$, $a \in \langle A, B\rangle$. Тогда у функции f' в точке a не может быть скачка ни слева, ни справа.*

Таким образом, производная дифференцируемой функции может иметь разрывы только второго рода.

Доказательство. Предположим, что $a \in \langle A, B)$ и f' имеет в точке a скачок справа. Это означает, что предел $L = \lim\limits_{x\to a+} f'(x)$ конечен и не совпадает с $f'(a)$. Будем для определенности считать, что $L < f'(a)$. Пусть $y \in \bigl(L, f'(a)\bigr)$. Найдется такое $\delta > 0$, что

$$[a, a+\delta) \subset \langle A, B) \quad \text{и} \quad f'(x) < y \ \text{ при любых } \ x \in (a, a+\delta).$$

Тогда $f'(a) > y$ и $f'(x) < y$ при всех $x \in (a, a+\delta)$. Таким образом, образ $f'\,\bigr|_{[a,a+\delta)}$ не является промежутком, что противоречит следствию 2. Случай левого скачка разбирается аналогично. $\square$

В заключение этого параграфа рассмотрим *параметрически заданные функции.* Пусть $\varphi, \psi\colon \langle A, B\rangle \to \mathbb{R}$. Будем понимать систему

$$\begin{cases} x = \varphi(t) \\ y = \psi(t) \end{cases} \quad (t \in \langle A, B\rangle)$$

как отображение промежутка $\langle A, B\rangle$ в плоскость $\mathbb{R}^2$. Если трактовать параметр t как время, то система описывает движение точки на плоскости. Можно ли исключить из этой системы t и однозначно выразить y через x? Иными словами, является ли кривая $\bigl\{\bigl(x(t), y(t)\bigr)\colon t \in \langle A, B\rangle\bigr\}$ графиком некоторой зависимости y от x? Ответ на этот вопрос отрицательный. Рассмотрим для примера $\varphi(t) = \cos t$, $\psi(t) = \sin t$ при $t \in [0, 2\pi]$. Такая система описывает движение точки по окружности, а окружность не является графиком какой-либо функции. Заметим, однако, что если функция φ *строго монотонна на* $\langle A, B\rangle$, то выразить y через x можно. Действительно, в этом случае φ имеет обратную функцию γ, определенную на $\varphi\bigl(\langle A, B\rangle\bigr)$. Из первого уравнения системы получаем $t = \gamma(x)$, поэтому $y = \psi\bigl(\gamma(x)\bigr)$, $x \in \varphi\bigl(\langle A, B\rangle\bigr)$. Положим $f = \psi \circ \gamma$. Мы будем называть f *параметрически заданной функцией.*

Предположим теперь, что функции φ и ψ *дифференцируемы на* $\langle A, B\rangle$. Что можно сказать о дифференцируемости f и как вычислить ее производную? Сформулируем достаточное условие существования и дифференцируемости параметрически заданной функции.

Теорема 7. О параметрически заданных функциях. *Пусть φ и ψ — дифференцируемые на $\langle A, B\rangle$ функции, причем $\varphi'(t) \neq 0$ при всех $t \in \langle A, B\rangle$. Тогда*

1) *существует единственная функция $f\colon \varphi\bigl(\langle A, B\rangle\bigr) \to \mathbb{R}$, для которой при всех $t \in \langle A, B\rangle$ выполняется равенство $f\bigl(\varphi(t)\bigr) = \psi(t)$;*

2) *функция f дифференцируема на $\varphi\bigl(\langle A, B\rangle\bigr)$ и*

$$f'\bigl(\varphi(t)\bigr) = \frac{\psi'(t)}{\varphi'(t)} \quad \text{при любом} \quad t \in \langle A, B\rangle.$$

Доказательство.

1) Так как $\varphi' \neq 0$ на $\langle A, B\rangle$, то по следствию 1 теоремы Дарбу функция φ строго монотонна на $\langle A, B\rangle$. Пусть $E = \varphi\bigl(\langle A, B\rangle\bigr)$, $\gamma\colon E \to \langle A, B\rangle$ — обратная функция к φ. Если $f = \psi \circ \gamma$, то

$$f\bigl(\varphi(t)\bigr) = \psi\Bigl(\gamma\bigl(\varphi(t)\bigr)\Bigr) = \psi(t) \ \text{ при всех } \ t \in \langle A, B\rangle.$$

Обратно, пусть $f\bigl(\varphi(t)\bigr) \equiv \psi(t)$. Полагая $t = \gamma(x)$, мы получим

$$f(x) = f\Bigl(\varphi\bigl(\gamma(x)\bigr)\Bigr) = \psi\bigl(\gamma(x)\bigr) = \bigl(\psi \circ \gamma\bigr)(x) \quad \text{при любом} \quad x \in E,$$

то есть $f = \psi \circ \gamma$. Таким образом, функция f единственна и имеет вид $\psi \circ \gamma$.

2) Дифференцируемость функции f на E следует из равенства $f = \psi \circ \gamma$ в силу теорем 1 и 3 § 3. Дифференцируя соотношение $f\bigl(\varphi(t)\bigr) = \psi(t)$, мы получим $f'\bigl(\varphi(t)\bigr) \cdot \varphi'(t) = \psi'(t)$, откуда и вытекает искомая формула для f'. $\square$

§ 8. Выпуклые функции

Перейдем к рассмотрению других важных геометрических характеристик поведения функции — *выпуклости* и *вогнутости*.

Параграф состоит из трех частей. Вначале мы дадим определение выпуклости и вогнутости и выведем из него ряд утверждений геометрического характера, полезных для построения графиков функций. Затем мы обсудим связь выпуклости с дифференциальным исчислением, которая позволит нам исследовать функции на выпуклость аналитическими средствами. Заключительная часть параграфа будет посвящена доказательству неравенств при помощи выпуклости. Мы выведем ряд классических неравенств, имеющих важное значение в различных областях математики. Перейдем к реализации этой схемы.

Рассмотрим функцию $f\colon \langle A,B\rangle \to \mathbb{R}$.

Определение 1.

1) Пусть для любых $a,b\in\langle A,B\rangle$ и $\lambda\in(0,1)$ справедливо неравенство

$$f\bigl(\lambda a+(1-\lambda)b\bigr)\leqslant \lambda f(a)+(1-\lambda)f(b).$$

Тогда f называется *выпуклой на* $\langle A,B\rangle$.

2) Пусть для любых $a,b\in\langle A,B\rangle$ $(a\neq b)$ и $\lambda\in(0,1)$ верно неравенство

$$f\bigl(\lambda a+(1-\lambda)b\bigr)< \lambda f(a)+(1-\lambda)f(b).$$

Тогда f называется *строго выпуклой на* $\langle A,B\rangle$.

Замечание. Определение строгой выпуклости не изменится при перестановке a и b, поэтому можно не умаляя общности считать $a<b$, что мы и будем делать в дальнейшем. Отметим также, что при $a=b$ неравенство в 1) превращается в равенство.

Выясним геометрический смысл выпуклости функции. Пусть $a,b\in\langle A,B\rangle$, $a<b$. Положим $x=\lambda a+(1-\lambda)b$. Когда λ пробегает интервал $(0,1)$, x пробегает все точки (a,b). Решив уравнение относительно λ, мы получим

$$\lambda=\frac{b-x}{b-a}\quad\text{и}\quad 1-\lambda=\frac{x-a}{b-a}.$$

Поэтому неравенство из определения строгой выпуклости перепишется в виде

$$f(x)<\frac{b-x}{b-a}\,f(a)+\frac{x-a}{b-a}\,f(b).$$

График правой части этого неравенства на $[a, b]$ есть хорда, соединяющая точки $\big(a, f(a)\big)$ и $\big(b, f(b)\big)$. Таким образом, строгая выпуклость функции означает, что любая хорда, соединяющая две точки ее графика, лежит строго выше самого графика, за исключением концевых точек (см. рисунок 31a). Если же каждая хорда лежит *не ниже* графика, то мы получим нестрогое неравенство, соответствующее выпуклости.

Заменяя в определении 1 знаки неравенств на противоположные, мы получим определения *вогнутой* и *строго вогнутой* функции. В литературе используются также термины *выпуклость вниз* и *выпуклость вверх* как синонимы соответственно выпуклости и вогнутости. Геометрический смысл вогнутости ясен из рисунка 31b. Заметим, что вогнутость функции f равносильна выпуклости $-f$. Поэтому в дальнейшем мы будем говорить только про выпуклость, предоставляя читателю разобрать вогнутость в качестве упражнения.

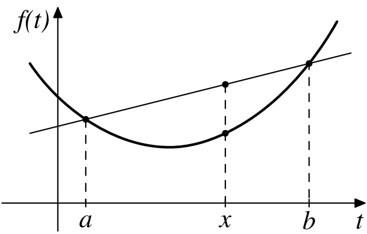

Рис. 31a

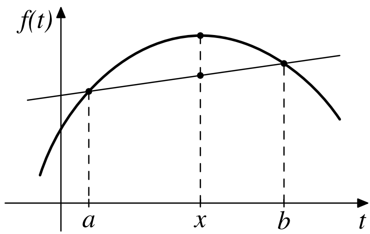

Рис. 31b

Замечание. Непосредственно из определения вытекают следующие правила для арифметических действий над выпуклыми функциями. *Пусть f и g — выпуклые функции на $\langle A, B\rangle$. Тогда*

1) *функция $f + g$ выпукла на $\langle A, B\rangle$;*

2) *при любом $\alpha > 0$ функция αf выпукла на $\langle A, B\rangle$;*

3) *при любом $\alpha < 0$ функция αf вогнута на $\langle A, B\rangle$.*

Обсуждение геометрических приложений выпуклости мы начнем с наглядного геометрического критерия строгой выпуклости.

Лемма 1. О трех хордах. *Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$. Равносильны следующие утверждения.*

1) *Функция f строго выпукла на $\langle A, B\rangle$.*

2) *Для любых $a, b, c \in \langle A, B\rangle$, $a < c < b$ справедливо неравенство*

$$\frac{f(c)-f(a)}{c-a} < \frac{f(b)-f(c)}{b-c}.$$

3) *Для любых $a, b, c \in \langle A, B\rangle$, $a < c < b$ справедливы неравенства*

$$\frac{f(c)-f(a)}{c-a} < \frac{f(b)-f(a)}{b-a} < \frac{f(b)-f(c)}{b-c}.$$

Замечания

1) Геометрический смысл двойного неравенства в 3) ясен из рисунка 32. Пусть $M = \big(a, f(a)\big)$, $N = \big(b, f(b)\big)$, $K = \big(c, f(c)\big)$. Если хорды MK, KN и MN наклонены к оси OX под углами α, β и γ соответственно, то утверждение 3 означает, что $\operatorname{tg}\alpha < \operatorname{tg}\gamma < \operatorname{tg}\beta$. Второе утверждение леммы — упрощенный вариант третьего, относящийся только к хордам MK и KN. Тем не менее утверждения 2 и 3 равносильны.

2) Если бы неравенства для хорд в 2) и 3) были нестрогими, мы бы получили *критерий выпуклости f на $\langle A, B\rangle$*. Его доказательство получается из доказательства леммы 1 заменой строгих неравенств нестрогими.

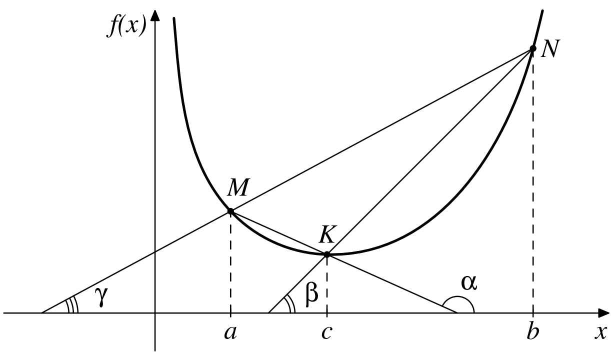

Рис. 32

Доказательство. Равносильность утверждений 1 – 3 мы докажем по схеме 1) ⇒ 3) ⇒ 2) ⇒ 1). Переход 3) ⇒ 2) тривиален. Проверим две другие импликации.

1) $\Rightarrow$ 3) Пусть $a, b, c \in \langle A, B\rangle$, $a < c < b$. Положим $\lambda = \dfrac{b-c}{b-a}$. Тогда $c = \lambda a + (1-\lambda)b$, и из строгой выпуклости f вытекает неравенство $f(c) < \lambda f(a) + (1-\lambda)f(b)$. Преобразуем его двумя способами. Во-первых,

$$f(c) - f(b) < \lambda\bigl(f(a) - f(b)\bigr) \Leftrightarrow$$
$$f(b) - f(c) > \frac{b-c}{b-a}\bigl(f(b) - f(a)\bigr) \Leftrightarrow \frac{f(b)-f(a)}{b-a} < \frac{f(b)-f(c)}{b-c},$$

что дает правое неравенство в 3). Во-вторых,

$$f(c) - f(a) < (1-\lambda)\bigl(f(b) - f(a)\bigr) \Leftrightarrow$$
$$f(c) - f(a) < \frac{c-a}{b-a}\bigl(f(b) - f(a)\bigr) \Leftrightarrow \frac{f(c)-f(a)}{c-a} < \frac{f(b)-f(a)}{b-a},$$

и левое неравенство в 3) также доказано.

2) $\Rightarrow$ 1) Пусть $a, b \in \langle A, B\rangle$, $a < b$, $\lambda \in (0, 1)$. Обозначим $c = \lambda a + (1-\lambda)b$. Тогда

$$\lambda = \frac{b-c}{b-a} \quad \text{и} \quad 1 - \lambda = \frac{c-a}{b-a},$$

откуда $\dfrac{c-a}{1-\lambda} = \dfrac{b-c}{\lambda}$. Используя 2), мы получим

$$\frac{f(c)-f(a)}{c-a} < \frac{f(b)-f(c)}{b-c} \Leftrightarrow \frac{f(c)-f(a)}{1-\lambda} < \frac{f(b)-f(c)}{\lambda} \Leftrightarrow$$
$$\lambda\bigl(f(c) - f(a)\bigr) < (1-\lambda)\bigl(f(b) - f(c)\bigr) \Leftrightarrow f(c) < \lambda f(a) + (1-\lambda)f(b),$$

что и означает строгую выпуклость f на $\langle A, B\rangle$. $\square$

Из леммы вытекает полезное следствие, которое мы будем использовать при доказательстве геометрических свойств выпуклых функций.

Следствие. *Пусть* $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in \langle A, B\rangle$. *Обозначим* $F(x) = \dfrac{f(x)-f(a)}{x-a}$. *Тогда справедливы следующие утверждения.*

1) *Если* f *выпукла на* $\langle A, B\rangle$, *то* F *возрастает на* $\langle A, B\rangle \setminus \{a\}$.

2) *Если f строго выпукла на $\langle A, B\rangle$, то F строго возрастает на $\langle A, B\rangle \setminus \{a\}$.*

Доказательство. Ограничимся проверкой второго утверждения. Пусть $x, y \in \langle A, B\rangle \setminus \{a\}$, $x < y$. Докажем неравенство $F(x) < F(y)$. Пользуясь утверждением 3 леммы 1, мы получим

$$\frac{f(x)-f(a)}{x-a} < \frac{f(y)-f(a)}{y-a} \quad \text{при} \quad a < x < y\,;$$
$$\frac{f(a)-f(x)}{a-x} < \frac{f(a)-f(y)}{a-y} \quad \text{при} \quad x < y < a\,;$$
$$\frac{f(a)-f(x)}{a-x} < \frac{f(y)-f(a)}{y-a} \quad \text{при} \quad x < a < y\,.$$

Каждое из этих неравенств эквивалентно $F(x) < F(y)$. $\square$

Замечание. Из рисунка 31а видно, что *продолжение за точки $\big(a, f(a)\big)$ и $\big(b, f(b)\big)$ хорды графика выпуклой функции f, соединяющей эти точки, лежит ниже графика f.* Предлагаем читателю вывести этот факт из леммы 1. *Указание:* используйте лемму для точек a, b и x, где $x \notin [a, b]$.

Пример. *Функция $f(x) = x^2$ строго выпукла на $\mathbb{R}$.* Действительно, пусть $a, b, c \in \mathbb{R}$, $a < c < b$. Тогда

$$\begin{aligned}\frac{f(c)-f(a)}{c-a} - \frac{f(b)-f(c)}{b-c} &= \frac{c^2-a^2}{c-a} - \frac{b^2-c^2}{b-c} = \\ &= c+a-(b+c) = a-b < 0,\end{aligned}$$

откуда по лемме 1 и вытекает строгая выпуклость. $\square$

Разобранный пример служит лишь для иллюстрации леммы о трех хордах. На практике выпуклость исследуется с помощью производных (см. теоремы 4 и 5).

До сих пор при изучении выпуклости мы не накладывали на функцию качественных условий, таких, как непрерывность или дифференцируемость. Оказывается, что выпуклые функции обязаны быть “достаточно хорошими”. Сформулируем теорему, объясняющую точный смысл сказанного.

Теорема 1. Об односторонних производных. *Пусть функция f выпукла на $\langle A, B\rangle$, $a \in \langle A, B\rangle$. Тогда справедливы следующие утверждения.*

1) *Для $a < B$ существует $f'_+(a) \in [-\infty, +\infty)$, а для $a > A$ существует $f'_-(a) \in (-\infty, +\infty]$.*

2) *Если $a \in (A, B)$, то $f'_+(a)$ и $f'_-(a)$ конечны и $f'_-(a) \leqslant f'_+(a)$.*

Доказательство. Рассмотрим функцию $F\colon \langle A, B\rangle \setminus \{a\} \to \mathbb{R}$, действующую по формуле $F(x) = \dfrac{f(x) - f(a)}{x - a}$.

1) Пусть $a < B$. По следствию леммы 1 функция F возрастает на (a, B). В силу теоремы о пределе монотонной функции (см. § 3 главы 2) существует $\lim\limits_{x \to a+} F(x) \in [-\infty, +\infty)$, который и есть $f'_+(a)$. Для $f'_-(a)$ утверждение доказывается аналогично.

2) Пусть $x, y \in \langle A, B\rangle$, $x < a < y$. Тогда по следствию леммы 1 $F(x) < F(y)$. Устремляя в этом неравенстве x и y к a соответственно слева и справа, мы получим $f'_-(a) \leqslant f'_+(a)$. Поскольку $f'_-(a) > -\infty$ и $f'_+(a) < +\infty$, обе производные оказываются конечными. $\square$

Следствие. Непрерывность выпуклых функций. *Если функция f выпукла на $\langle A, B\rangle$, то она непрерывна на (A, B).*

Доказательство. Пусть $a \in (A, B)$. Из существования конечных производных $f'_\pm(a)$ следует непрерывность f в точке a слева и справа (см. § 2). Поэтому f непрерывна в точке a. $\square$

Пример. Положим

$$f(x) = -\sqrt{1 - x^2} \ \text{ при } \ x \in (-1, 1] \quad \text{и} \quad f(-1) = 1.$$

Выпуклость f на $[-1, 1]$ ясна из геометрических соображений, поскольку график $f\big|_{(-1,1]}$ представляет собой полуокружность. Несложно проверить, что $f'_+(-1) = -\infty$, $f'_-(1) = +\infty$ и f разрывна в точке -1. Таким образом, в концевых точках промежутка выпуклая функция может вести себя хуже, чем в его внутренних точках.

Наряду с хордами для изучения выпуклых функций используется еще один полезный инструмент — *опорная прямая.* Рассмотрим это понятие подробнее.

Определение 2. Опорная прямая. Пусть $f\colon \langle A,B\rangle \to \mathbb{R}$, $a \in \langle A,B\rangle$, $k \in \mathbb{R}$, $g(x) = f(a) + k\,(x-a)$.

1) Прямая, задаваемая уравнением $y = g(x)$, называется *опорной для функции f в точке a*, если $f(x) \geqslant g(x)$ при всех $x \in \langle A,B\rangle$.

2) Прямая, задаваемая уравнением $y = g(x)$, называется *строго опорной для функции f в точке a*, если $f(x) > g(x)$ для любого $x \in \langle A,B\rangle \setminus \{a\}$.

Таким образом, опорной для f в точке a считается прямая, проходящая через точку $\big(a, f(a)\big)$, которая лежит *не выше графика f*. Строго опорная прямая лежит *ниже графика функции f* во всех точках, кроме $\big(a, f(a)\big)$.

Изучим взаимосвязь выпуклости с опорными прямыми. Докажем вначале лемму о существовании опорной прямой для выпуклой функции.

Лемма 2. Описание опорных прямых выпуклой функции. *Пусть $f\colon \langle A,B\rangle \to \mathbb{R}$, $a \in (A,B)$, $k \in \mathbb{R}$. Предположим, что прямая ℓ задается уравнением $y = f(a) + k\,(x-a)$. Справедливы следующие утверждения.*

1) *Пусть f выпукла на $\langle A,B\rangle$. Прямая ℓ является опорной для f в точке a тогда и только тогда, когда $k \in \big[f'_-(a), f'_+(a)\big]$.*

2) *Пусть f строго выпукла на $\langle A,B\rangle$. Прямая ℓ является строго опорной для функции f в точке a тогда и только тогда, когда $k \in \big[f'_-(a), f'_+(a)\big]$.*

Геометрический смысл леммы иллюстрирует рисунок 33.

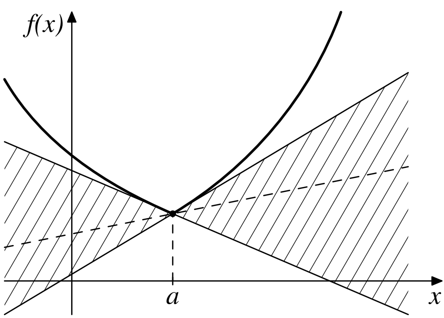

Рис. 33

Замечания

1) Из теоремы 1 вытекает существование для $a \in (A, B)$ конечных производных $f'_{\pm}(a)$ и неравенство $f'_-(a) \leqslant f'_+(a)$. Поэтому в любой точке $a \in (A, B)$ у выпуклой функции *существует* опорная прямая, а у строго выпуклой — строго опорная прямая. Это верно и для концевой точки a промежутка $\langle A, B \rangle$, если *потребовать* существование в ней конечной односторонней производной (см. доказательство леммы).

2) Если функция f *дифференцируема* в точке $a \in (A, B)$, то ее опорная прямая в точке a *единственна и совпадает с касательной в точке* a. Действительно, в этом случае $f'_-(a) = f'_+(a) = f'(a)$, поэтому прямая ℓ будет опорной только при условии $k = f'(a)$. Значит, опорная прямая единственна, а ее уравнением будет уравнение касательной $y = f(a) + f'(a)\,(x - a)$.

Доказательство. Мы ограничимся доказательством второго утверждения.

Необходимость. Пусть ℓ является строго опорной прямой функции f в точке a. По определению правой производной

$$f(x) - f(a) = f'_+(a)\,(x - a) + o(x - a) \ \text{ при } \ x \to a+\,.$$

Поэтому

$$0 < f(x) - f(a) - k\,(x - a) = (x - a) \cdot \bigl(f'_+(a) - k + o(1)\bigr) \quad (x \to a+),$$

откуда $f'_+(a) - k + o(1) > 0$ при $x \in (a, B)$. Устремляя x к a справа, мы получим $f'_+(a) \geqslant k$. Аналогичные рассуждения для $x \in (A, a)$ приводят к неравенству $f'_-(a) \leqslant k$.

Достаточность. Пусть $k \in \bigl[f'_-(a), f'_+(a)\bigr]$. Рассмотрим функцию $F(x) = \dfrac{f(x) - f(a)}{x - a}$. В силу следствия леммы 1 F строго возрастает на $(a, B\rangle$, поэтому

$$f'_+(a) = \inf_{(a,B)} F < F(x) \ \text{ при всех } \ x \in (a, B\rangle.$$

Тогда для $x \in (a, B\rangle$ мы получим

$$f(x) - f(a) = F(x)\,(x - a) > f'_+(a)\,(x - a) \geqslant k\,(x - a)\,,$$

откуда $f(x) > f(a) + k\,(x - a)$. Справедливость этого неравенства для $x \in \langle A, a)$ проверяется аналогично. Поэтому на $\langle A, B\rangle \setminus \{a\}$ график f лежит выше ℓ. $\square$

Покажем теперь, что утверждение леммы 2 можно обратить.

Лемма 3. Характеристика выпуклости с помощью опорных прямых. *Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$. Тогда справедливы следующие утверждения.*

1) *Если в каждой точке $a \in \langle A, B\rangle$ функция f имеет опорную прямую, то она выпукла на промежутке $\langle A, B\rangle$.*

2) *Если в каждой точке $a \in \langle A, B\rangle$ функция f имеет строго опорную прямую, то она строго выпукла на $\langle A, B\rangle$.*

Доказательство. Мы вновь ограничимся доказательством второго, более сильного утверждения. Пусть $a, b, c \in \langle A, B\rangle$, $a < c < b$. Положим $F(x) = \dfrac{f(x) - f(c)}{x - c}$. В точке c у функции f есть строго опорная прямая $y = f(c) + k\,(x - c)$. Пусть $x \neq c$. Тогда

$$0 < f(x) - f(c) - k\,(x - c) = (x - c)\big(F(x) - k\big)\,,$$

откуда $F(x) > k$ при $x > c$ и $F(x) < k$ при $x < c$. Так как $a < c$, а $b > c$, то $F(a) < F(b)$, то есть

$$\frac{f(c) - f(a)}{c - a} < \frac{f(b) - f(c)}{b - c}.$$

По лемме 1 мы получаем строгую выпуклость f на $\langle A, B\rangle$. $\square$

Комбинируя леммы 2 и 3, можно получить *критерий выпуклости* функции в терминах опорной прямой. Мы докажем его только для *дифференцируемых* функций, поскольку в такой редакции он нагляднее и удобнее для практических целей.

Теорема 2. Критерий выпуклости в терминах касательных. *Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, f дифференцируема на $\langle A, B\rangle$. Справедливы следующие утверждения.*

1) *Функция f выпукла на $\langle A, B\rangle$ тогда и только тогда, когда для любых $a, x \in \langle A, B\rangle$ выполняется неравенство $f(x) \geqslant T_{a,1}f(x)$.*

2) *Функция f строго выпукла на $\langle A, B\rangle$ тогда и только тогда, когда для любых различных $a, x \in \langle A, B\rangle$ выполняется неравенство $f(x) > T_{a,1}f(x)$.*

Таким образом, график выпуклой функции можно представлять себе как *верхнюю огибающую* семейства ее касательных (см. рисунок 34a). График строго выпуклой функции пересекается с каждой касательной только в точке касания (рисунок 34b).

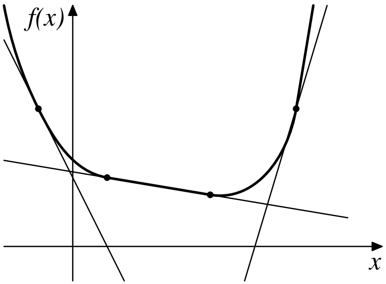

Рис. 34a

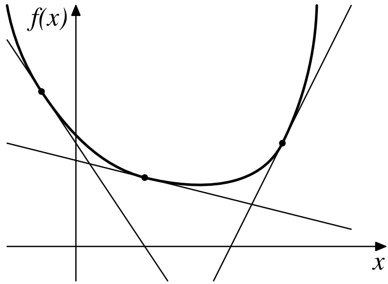

Рис. 34b

Доказательство. Мы ограничимся доказательством второго утверждения. Необходимость вытекает из замечания 2 к лемме 2. Проверим достаточность. Поскольку неравенство $f(x) > T_{a,1}f(x)$ верно при всех различных $a, x \in \langle A, B \rangle$, касательная $y = T_{a,1}f(x)$ является строго опорной прямой для f в любой точке $a \in \langle A, B \rangle$. Поэтому строгая выпуклость f на $\langle A, B \rangle$ следует из леммы 3. $\square$

Выпуклость и вогнутость используются при построении графиков функций, имеющих *наклонные асимптоты* (см. определение в § 6 главы 2). Рисунок 35 позволяет предположить, что выпуклая функция приближается к своей асимптоте сверху, а вогнутая — снизу. Докажем, что так и есть на самом деле.

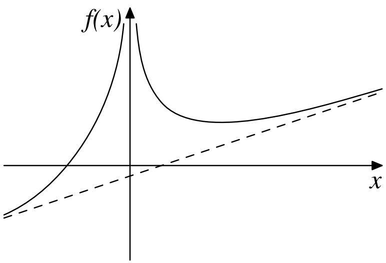

Рис. 35a

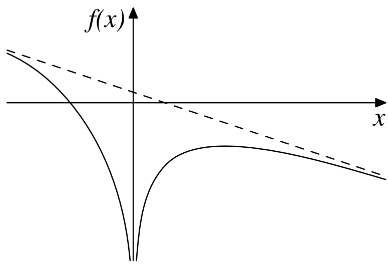

Рис. 35b

Теорема 3. Выпуклость и асимптота. *Пусть функция* $f\colon (A, +\infty) \to \mathbb{R}$ *имеет при* $x \to +\infty$ *асимптоту* $y = kx + b$. *Справедливы следующие утверждения.*

1) *Если* f *выпукла на* $(A, +\infty)$, *то* $f(x) \geqslant kx + b$ *при всех* $x > A$.

2) *Если* f *строго выпукла на* $(A, +\infty)$, *то* $f(x) > kx + b$ *при всех* $x > A$.

Замечания

1) Утверждения теоремы верны и для асимптот на $-\infty$.

2) Поскольку $f(x) - kx - b \to 0$ при $x \to +\infty$, мы можем трактовать асимптоту выпуклой функции как ее *опорную прямую в* $+\infty$. Таким образом, теорема 3 дополняет лемму 2 для функций, заданных на неограниченных промежутках.

Доказательство. Мы вновь ограничимся доказательством второго утверждения. Если мы проверим *строгое убывание функции* $g(x) = f(x) - kx$ *на* $(A, +\infty)$, то по теореме о пределе монотонной функции при всех $x \in (A, +\infty)$ будет выполняться неравенство $g(x) > \lim\limits_{t\to+\infty} g(t) = b$, которое и означает $f(x) > kx + b$.

Докажем строгое убывание g. Пусть $A < x < y$. Положим $F(t) = \dfrac{f(t) - f(x)}{t - x}$. Заметим, что $\lim\limits_{t\to+\infty} F(t) = k$. Действительно, если $k = 0$, то $\lim\limits_{t\to+\infty} f(t) = b$, и равенство очевидно, а при $k \neq 0$ $\lim\limits_{t\to+\infty} f(t) = \infty$, откуда

$$\lim_{t\to+\infty} F(t) = \lim_{t\to+\infty} \frac{f(t)}{t} = k$$

(см. § 6 главы 2). В силу следствия леммы 1 F строго возрастает на $(x, +\infty)$. Тогда по теореме о пределе монотонной функции $F(t) < k$ при всех $t > x$ и, в частности, $F(y) < k$. Поэтому

$$f(y) - f(x) < k\,(y - x) \Leftrightarrow f(y) - ky < f(x) - kx \Leftrightarrow g(y) < g(x),$$

откуда и следует строгое убывание g. $\square$

Полученные ранее критерии выпуклости удобны с геометрической точки зрения, но они не годятся для практического использования. Чтобы исследовать функцию на выпуклость, необходимо иметь *аналитический критерий*. Оказывается, для *дифференцируемой* функции выпуклость эквивалентна монотонности ее производной. Сформулируем это утверждение.

Теорема 4. Критерий выпуклости в терминах первой производной. *Пусть* $f\colon \langle A, B\rangle \to \mathbb{R}$, f *непрерывна на* $\langle A, B\rangle$ *и дифференцируема на* (A, B). *Справедливы следующие утверждения.*

1) f *выпукла на* $\langle A, B\rangle$ *тогда и только тогда, когда* f' *возрастает на* (A, B).

2) f *строго выпукла на* $\langle A, B\rangle$ *тогда и только тогда, когда* f' *строго возрастает на* (A, B).

Доказательство. Ограничимся проверкой утверждения 2.

Необходимость. Пусть $x, y \in (A, B)$, $x < y$. Достаточно доказать двойное неравенство

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

Положим $F(t) = \dfrac{f(t) - f(x)}{t - x}$ $(t \neq x)$. По следствию леммы 1 функция F строго возрастает на $(x, y]$, поэтому

$$f'(x) = \lim_{t\to x+0} F(t) = \inf_{(x,y]} F < F(y) = \frac{f(y) - f(x)}{y - x}.$$

Правое неравенство проверяется аналогично.

Достаточность. Пусть $a, b, c \in \langle A, B\rangle$, $a < c < b$. По лемме 1 достаточно проверить, что

$$\frac{f(c) - f(a)}{c - a} < \frac{f(b) - f(c)}{b - c}.$$

По теореме Лагранжа (см. § 4) найдутся такие $\alpha \in (a, c)$ и $\beta \in (c, b)$, что

$$f'(\alpha) = \frac{f(c) - f(a)}{c - a} \quad \text{и} \quad f'(\beta) = \frac{f(b) - f(c)}{b - c}.$$

Так как f' строго возрастает и $\alpha < \beta$, то $f'(\alpha) < f'(\beta)$, откуда и вытекает искомое неравенство. $\square$

Комбинируя теорему 4 с *условиями монотонности* из § 7, мы получим характеристику выпуклости в терминах второй производной.

Теорема 5. Условия выпуклости в терминах второй производной. *Пусть* $f\colon \langle A, B\rangle \to \mathbb{R}$, *$f$ непрерывна на* $\langle A, B\rangle$ *и дважды дифференцируема на* (A, B). *Справедливы следующие утверждения.*

1) *f выпукла на* $\langle A, B\rangle$ *тогда и только тогда, когда* $f''(x) \geqslant 0$ *при всех* $x \in (A, B)$.

2) *Если* $f''(x) > 0$ *при всех* $x \in (A, B)$, *то f строго выпукла на* $\langle A, B\rangle$.

Доказательство.

1) По теореме 4 выпуклость f равносильна возрастанию f' на (A, B), а оно, согласно критерию монотонности, эквивалентно условию $f''(x) \geqslant 0$ при всех $x \in (A, B)$.

2) Если $f''(x) > 0$ при всех $x \in (A, B)$, то f' строго возрастает на (A, B) (теорема 1 § 7), откуда по теореме 4 вытекает строгая выпуклость f. $\square$

С выпуклостью связано еще одно геометрическое понятие — *точка перегиба*. Рассмотрим его подробнее.

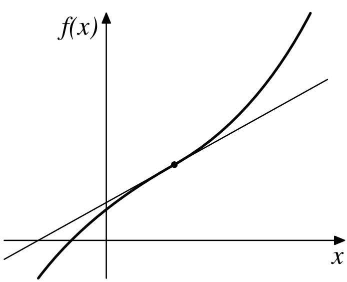

Рис. 36a

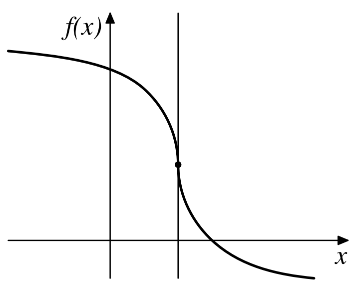

Рис. 36b

Определение 3. Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in (A, B)$. Предположим, что выполнены следующие условия:

1) существует такое $\delta > 0$, что $(a-\delta, a+\delta) \subset (A, B)$ и f имеет разный характер выпуклости на $(a-\delta, a]$ и $[a, a+\delta)$ (то есть на одном из промежутков f выпукла, на другом — вогнута);

2) f непрерывна в точке a;

3) существует $f'(a) \in \overline{\mathbb{R}}$.

Тогда a называется *точкой перегиба* f.

Таким образом, в точке перегиба график функции меняет характер выпуклости и переходит с одной стороны касательной на другую. Рисунок 36 поясняет смысл сказанного.

Обратимся к задаче *поиска* точек перегиба.

Теорема 6. Необходимое условие перегиба.
Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in (A, B)$. *Предположим, что функция* f *дважды дифференцируема в точке* a. *Если* a *является точкой перегиба* f, *то* $f''(a) = 0$.

Замечание. Теорема не дает достаточного условия перегиба. Например, функция $f(x) = x^4$ строго выпукла на $\mathbb{R}$, но $f''(0) = 0$.

Доказательство. Будем для определенности считать, что f вогнута слева от a и выпукла справа от a. Выберем столь малое $\delta > 0$, что функция f дифференцируема на $(a-\delta, a+\delta)$, вогнута на $(a-\delta, a]$ и выпукла на $[a, a+\delta)$. Тогда f' убывает на $(a-\delta, a]$ и возрастает на $[a, a+\delta)$. Значит, a является точкой минимума f', откуда $f''(a) = 0$ по теореме 2 § 7. $\square$

Теорема 7. Достаточное условие перегиба.
Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, $a \in (A, B)$, f *непрерывна в точке* a *и имеет в ней производную из* $\overline{\mathbb{R}}$. *Пусть существует такое* $\delta > 0$, *что* f *дважды дифференцируема на* $(a-\delta, a) \cup (a, a+\delta)$ *и выполнено одно из следующих условий:*

1) $f'' > 0$ *на* $(a-\delta, a)$ *и* $f'' < 0$ *на* $(a, a+\delta)$;

2) $f'' < 0$ *на* $(a-\delta, a)$ *и* $f'' > 0$ *на* $(a, a+\delta)$.

Тогда a *— точка перегиба* f.

Доказательство. Будем для определенности считать, что выполнено условие 1). По теореме 6 функция f выпукла на $(a-\delta, a]$ и вогнута на $[a, a+\delta)$. Тем самым выполнено первое требование в определении 3, а два других следуют непосредственно из условия. $\square$

В качестве приложения теорем 5 – 7 изучим характер выпуклости простейших элементарных функций. Теперь мы имеем возможность строго обосновать, почему их графики выглядят привычным для нас образом.

Примеры

1) При $b > 0$, $b \neq 1$ функция $x \mapsto b^x$ строго выпукла на $\mathbb{R}$.

2) Функция $x \mapsto \log_b x$ строго выпукла на $(0, +\infty)$ при $b \in (0, 1)$ и строго вогнута на $(0, +\infty)$ при $b > 1$.

3) Функция $x \mapsto \sin x$ строго выпукла на промежутках вида $[\pi(2k-1), 2\pi k]$ и строго вогнута на $[2\pi k, \pi(2k+1)]$ $(k \in \mathbb{Z})$. Точки πk при $k \in \mathbb{Z}$ являются точками перегиба функции.

4) Функция $x \mapsto \cos x$ строго вогнута на промежутках вида $[\pi(2k - \frac{1}{2}), \pi(2k + \frac{1}{2})]$ и строго выпукла на $[\pi(2k + \frac{1}{2}), \pi(2k + \frac{3}{2})]$ $(k \in \mathbb{Z})$. Точки $\pi\big(k + \frac{1}{2}\big)$ при $k \in \mathbb{Z}$ являются точками перегиба функции.

5) Функция $x \mapsto \operatorname{tg} x$ строго вогнута на промежутках вида $\big(\pi\big(k - \frac{1}{2}\big), \pi k\big]$ и строго выпукла на $\big[\pi k, \pi\big(k + \frac{1}{2}\big)\big)$ $(k \in \mathbb{Z})$. Точки πk при $k \in \mathbb{Z}$ являются точками перегиба функции.

6) Функция $x \mapsto \operatorname{ctg} x$ строго выпукла на промежутках вида $\big(\pi k, \pi\big(k + \frac{1}{2}\big)\big]$ и строго вогнута на $\big[\pi\big(k + \frac{1}{2}\big), \pi(k+1)\big)$ $(k \in \mathbb{Z})$. Точки $\pi\big(k + \frac{1}{2}\big)$ при $k \in \mathbb{Z}$ являются точками перегиба функции.

7) Функция $x \mapsto \arcsin x$ строго вогнута на $[-1, 0]$ и строго выпукла на $[0, 1]$. В точке 0 функция имеет перегиб.

8) Функция $x \mapsto \arccos x$ строго выпукла на $[-1, 0]$ и строго вогнута на $[0, 1]$. В точке 0 функция имеет перегиб.

9) Функция $x \mapsto \operatorname{arctg} x$ строго выпукла на $(-\infty, 0]$ и строго вогнута на $[0, +\infty)$. В точке 0 функция имеет перегиб.

10) Функция $x \mapsto \operatorname{arcctg} x$ строго вогнута на $(-\infty, 0]$ и строго выпукла на $[0, +\infty)$. В точке 0 функция имеет перегиб.

11) Функция $x \mapsto x^\alpha$ строго выпукла на $(0, +\infty)$ при $\alpha < 0$ или $\alpha > 1$ и строго вогнута на $(0, +\infty)$ при $\alpha \in (0, 1)$.

Замечание. В последнем примере мы рассмотрели для простоты промежуток $(0, +\infty)$, поскольку на нем x^α определена *при любом* $\alpha \in \mathbb{R}$. Для некоторых α результат можно усилить. Например, функция x^4 строго выпукла на $\mathbb{R}$, а $x^{1/3}$ строго выпукла на $(-\infty, 0]$, строго вогнута на $[0, +\infty)$ и имеет перегиб в нуле. Предлагаем читателю самостоятельно рассмотреть *все* случаи, предварительно вспомнив материал § 3.

Все сформулированные утверждения вытекают из теорем 5 и 7. Разберем в качестве образца третий пример. Если $f(x) = \sin x$, то $f''(x) = -\sin x$, поэтому $f''(x) > 0$ при $x \in (2\pi k - \pi, 2\pi k)$ и $f''(x) < 0$ при $x \in (2\pi k, 2\pi k + \pi)$, где $k \in \mathbb{Z}$. Так как f непрерывна на концах указанных промежутков, то по теореме 5 она строго выпукла на $[2\pi k - \pi, 2\pi k]$ и строго вогнута на $[2\pi k, 2\pi k + \pi]$. В точках вида $\pi k \ \ (k \in \mathbb{Z})$ функция f дифференцируема и меняет характер выпуклости, поэтому по теореме 7 она имеет в этих точках перегиб. Заметим также, что $f''(x) \neq 0$ при $x \notin \{\pi k\colon\ k \in \mathbb{Z}\}$, и по теореме 6 в точке x не может быть перегиба. Это полностью согласуется с предыдущими рассуждениями. $\square$

Перейдем к последней части параграфа, в которой мы докажем ряд важных неравенств с помощью выпуклости. Опишем общую схему рассуждений. Определение и геометрические следствия выпуклости и вогнутости функции записываются *с помощью неравенств*. Если мы установим выпуклость на основании теоремы 5, то эти неравенства тем самым *будут доказаны*. Применяя данную схему для различных выпуклых и вогнутых функций, можно получать содержательные результаты. Проиллюстрируем сказанное двумя примерами.

Примеры

1) $\sin x > \frac{2}{\pi}\, x$ *при всех* $x \in \left(0, \frac{\pi}{2}\right)$.

Действительно, положим $f(x) = \sin x$. Как показано выше, функция f строго вогнута на $\left[0, \frac{\pi}{2}\right]$. Поэтому внутренние точки хорды $y = \frac{2}{\pi}\, x$, соединяющей $(0,0)$ и $\left(\frac{\pi}{2}, 1\right)$, лежат строго ниже графика f, а это и нужно доказать. $\square$

2) *Если* $\alpha > 1$, *то* $(1+x)^\alpha > 1 + \alpha x$ *при всех* $x \geqslant -1$, $x \neq 0$.

Здесь мы используем другой способ доказательства неравенств, основанный на теореме 2. Положим $f(x) = (1+x)^\alpha$. Так как

$$f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2} > 0 \ \text{ для всех } \ x > -1,$$

то f строго выпукла на $(-1, +\infty)$, и по теореме 2 $\ f(x) > T_{0,1} f(x)$ при любых $x \in (-1, +\infty) \setminus \{0\}$. По теореме 4 § 6 $\ T_{0,1} f(x) = 1 + \alpha x$, что доказывает искомое неравенство для $x > -1$, а при $x = -1$ оно очевидно. Полученный результат является *обобщением неравенства Бернулли* на случай нецелых $\alpha > 1$. $\square$

Распоряжаясь подходящим образом параметрами, фигурирующими в определениях выпуклости и вогнутости, можно получать различные классические неравенства. Перейдем к их рассмотрению.

Определение 4. Числа p и q из $(1, +\infty)$, связанные соотношением $\frac{1}{p} + \frac{1}{q} = 1$, называются *сопряженными показателями.*

Сопряженные показатели будет использоваться в неравенствах Юнга и Гёльдера, которые мы докажем ниже.

Теорема 8. Неравенство Юнга. *Пусть $x, y \geqslant 0$, p и q — сопряженные показатели. Тогда $xy \leqslant \dfrac{x^p}{p} + \dfrac{x^q}{q}$, причем равенство реализуется тогда и только тогда, когда $x^p = y^q$.*

Доказательство. При нулевом x или y неравенство Юнга очевидно, а равенством в этом случае оно будет только при $x = y = 0$. Поэтому будем считать $x, y > 0$. Положим $f(t) = \ln t$. Так как f строго вогнута на $(0, +\infty)$, то

$$f\big(\lambda x^p + (1-\lambda)\, y^q\big) > \lambda f\big(x^p\big) + (1-\lambda) f\big(y^q\big) \quad (\lambda \in (0,1)),$$

причем равенство достигается тогда и только тогда, когда $x^p = y^q$. Поскольку показатели p и q сопряженные, мы можем выбрать $\lambda = \frac{1}{p}$ и $1 - \lambda = \frac{1}{q}$. Тогда

$$\ln\left(\tfrac{1}{p}\, x^p + \tfrac{1}{q}\, y^q\right) \geqslant \tfrac{1}{p}\, \ln x^p + \tfrac{1}{q}\, \ln y^q,$$

то есть $\ln\left(\frac{1}{p}\, x^p + \frac{1}{q}\, y^q\right) \geqslant \ln xy$, что и дает неравенство Юнга. $\square$

Следующие два неравенства, которые нам предстоит доказать, имеют наглядную геометрическую интерпретацию. Поэтому до их обсуждения мы должны определить некоторые операции над векторами в $\mathbb{R}^n$. Пусть $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ — векторы из $\mathbb{R}^n$.

1) *Сложение векторов*: $x + y = (x_1 + y_1, \ldots, x_n + y_n)$.
2) *Умножение вектора на число*: $\lambda x = (\lambda x_1, \ldots, \lambda x_n)$ $(\lambda \in \mathbb{R})$.
3) *Скалярное произведение векторов*: $x \cdot y = x_1 y_1 + \ldots + x_n y_n$.
4) *Длина вектора*: $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$.

Замечания

1) В курсе аналитической геометрии операции векторной алгебры в $\mathbb{R}^2$ и в $\mathbb{R}^3$ определяют без использования координат, а приведенные выше формулы доказывают. Таким образом, введенные нами операции над векторами в $\mathbb{R}^n$ при $n = 2$ и $n = 3$ согласуются с определениями, известными из геометрии.

2) Для $p \geqslant 1$ можно определить *p-норму вектора* x по формуле

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}.$$

Заметим, что $\|x\|_2$ есть длина вектора x, поэтому p-норму называют еще *обобщенной длиной*. Мы будем использовать p-нормы в неравенствах Минковского и Гёльдера (теоремы 9 и 10).

Напомним еще определения *коллинеарности* и *сонаправленности* векторов, также относящиеся к курсу геометрии.

Определение 5. Пусть $x, y \in \mathbb{R}^n$.

1) Векторы x и y называются *коллинеарными*, если либо один из них нулевой, либо существует $\lambda \in \mathbb{R} \setminus \{0\}$, для которого $x = \lambda y$.

2) Векторы x и y называются *сонаправленными*, если либо один из них нулевой, либо существует $\lambda > 0$, для которого $x = \lambda y$.

Теорема 9. Неравенство Минковского для неотрицательных чисел. *Пусть $n \in \mathbb{N}$, $p \geqslant 1$. Предположим, что векторы $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ имеют неотрицательные координаты. Тогда*

$$\left(\sum_{k=1}^{n} (x_k + y_k)^p\right)^{1/p} \leqslant \left(\sum_{k=1}^{n} x_k^p\right)^{1/p} + \left(\sum_{k=1}^{n} y_k^p\right)^{1/p}.$$

Равенство реализуется тогда и только тогда, когда либо $p = 1$, либо x и y сонаправлены.

Замечание. Для сонаправленных векторов неравенство Минковского, очевидно, превратится в равенство. Содержательным является обратное утверждение.

Доказательство. Достаточно рассмотреть случай, когда $p > 1$ и оба вектора ненулевые, в противном случае утверждение очевидно. Положим $X = \|x\|_p$ и $Y = \|y\|_p$. Заметим, что $X > 0$ и $Y > 0$. Функция $f(x) = x^p$ строго выпукла на $[0, +\infty)$, поскольку она непрерывна в нуле и

$$f''(x) = p(p-1)x^{p-2} > 0 \text{ при } x > 0$$

(см. теорему 5). Тогда для $\lambda \in (0,1)$ и $k \in \{1, \dots, n\}$

$$f\left(\lambda \frac{x_k}{X} + (1-\lambda)\frac{y_k}{Y}\right) \leqslant \lambda f\left(\frac{x_k}{X}\right) + (1-\lambda) f\left(\frac{y_k}{Y}\right),$$

причем равенство реализуется только в случае $\dfrac{x_k}{X} = \dfrac{y_k}{Y}$. Если $\lambda = \dfrac{X}{X+Y}$ и $1 - \lambda = \dfrac{Y}{X+Y}$, то

$$f\left(\frac{x_k + y_k}{X+Y}\right) \leqslant \frac{X}{X+Y} f\left(\frac{x_k}{X}\right) + \frac{Y}{X+Y} f\left(\frac{y_k}{Y}\right),$$

откуда

$$\frac{(x_k + y_k)^p}{(X+Y)^p} \leqslant \frac{X}{X+Y} \left(\frac{x_k}{X}\right)^p + \frac{Y}{X+Y} \left(\frac{y_k}{Y}\right)^p.$$

Сложим эти неравенства для $k = 1, \dots, n$. Так как

$$\sum_{k=1}^{n} \frac{x_k^p}{X^p} = \frac{1}{\|x\|_p^p} \sum_{k=1}^{n} x_k^p = 1 \quad \text{и} \quad \sum_{k=1}^{n} \frac{y_k^p}{Y^p} = 1,$$

мы получим

$$\frac{\|x+y\|_p^p}{(X+Y)^p} = \sum_{k=1}^{n} \frac{(x_k + y_k)^p}{(X+Y)^p} \leqslant$$

$$\leqslant \frac{X}{X+Y} \sum_{k=1}^{n} \frac{x_k^p}{X^p} + \frac{Y}{X+Y} \sum_{k=1}^{n} \frac{y_k^p}{Y^p} = \frac{X}{X+Y} + \frac{Y}{X+Y} = 1,$$

откуда $\|x+y\|_p \leqslant X+Y = \|x\|_p + \|y\|_p$, а это и есть неравенство Минковского. Равенство получится только при условии, что *во всех* сложенных неравенствах было равенство. Это означает, что

$$\frac{x_k}{X} = \frac{y_k}{Y} \text{ при } k=1,\ldots,n,$$

то есть $x = \frac{X}{Y}\,y$. $\square$

Следствие. Неравенство Минковского в $\mathbb{R}^n$. *Пусть $n \in \mathbb{N}$, $p \geqslant 1$, $x=(x_1,\ldots,x_n)$ и $y=(y_1,\ldots,y_n)$ — векторы из $\mathbb{R}^n$. Тогда*

$$\left(\sum_{k=1}^n |x_k+y_k|^p\right)^{1/p} \leqslant \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p\right)^{1/p}.$$

В случае $p>1$ равенство реализуется тогда и только тогда, когда x и y сонаправлены.

Доказательство. Будем считать, что оба вектора ненулевые и $p>1$, иначе утверждение очевидно. Применяя теорему 9 к векторам с координатами $|x_k|$ и $|y_k|$ $(k=1,\ldots,n)$, мы получим

$$\left(\sum_{k=1}^n |x_k+y_k|^p\right)^{1/p} \leqslant \left(\sum_{k=1}^n \bigl(|x_k|+|y_k|\bigr)^p\right)^{1/p} \leqslant$$
$$\leqslant \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p\right)^{1/p}.$$

Первый переход превращается в равенство тогда и только тогда, когда x_k и y_k имеют одинаковый знак при любых $k=1,\ldots,n$. Равенство во втором переходе по теореме 9 означает существование $\lambda>0$, при котором $|x_k| = \lambda\,|y_k|$ для всех $k=1,\ldots,n$. Так как знаки x_k и y_k одинаковы, мы получаем $x_k = \lambda\, y_k$ $(k=1,\ldots,n)$, то есть сонаправленность x и y. $\square$

Замечания

1) При $p=1$ равенство достигается не только для сонаправленных векторов x и y. Предлагаем читателю самостоятельно сформулировать условие, при котором реализуется равенство.

2) Неравенство Минковского можно записать в краткой форме

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

При $p = 2$ числа $\|x\|_p$, $\|y\|_p$ и $\|x+y\|_p$ являются длинами сторон треугольника с вершинами в точках 0, x и $x+y$, поэтому неравенство превращается в *неравенство треугольника*, знакомое нам в $\mathbb{R}^2$ и в $\mathbb{R}^3$ из геометрии. Равенство в нем достигается, когда вершины треугольника лежат на одной прямой, что и означает сонаправленность x и y. Таким образом, неравенство Минковского можно понимать как обобщенное неравенство треугольника.

Теорема 10. Неравенство Гёльдера для неотрицательных чисел. *Пусть p и q — сопряженные показатели, $n \in \mathbb{N}$. Если векторы $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ имеют неотрицательные координаты, то*

$$\sum_{k=1}^{n} x_k\, y_k \leqslant \left(\sum_{k=1}^{n} x_k^p\right)^{1/p} \cdot \left(\sum_{k=1}^{n} y_k^q\right)^{1/q}.$$

Равенство реализуется тогда и только тогда, когда

$$x_k^p\, \|y\|_q^q = y_k^q\, \|x\|_p^p \ \text{ при всех } \ k = 1, \ldots, n.$$

Замечание. Равенство в неравенстве Гёльдера достигается тогда и только тогда, когда сонаправлены векторы $(x_1^p, \ldots, x_n^p)$ и $(y_1^q, \ldots, y_n^q)$. Действительно, необходимость вытекает непосредственно из теоремы. Достаточность проверяется прямым вычислением, и мы предлагаем читателю доказать ее самостоятельно. *Указание:* воспользуйтесь равенством $\frac{q}{p} = q\,(1 - \frac{1}{q}) = q - 1$.

Доказательство. Здесь мы воспользуемся выпуклостью не напрямую, а посредством неравенства Юнга. Будем считать векторы x и y ненулевыми, иначе утверждение теоремы очевидно. Пусть a и b — векторы с координатами соответственно $a_k = \dfrac{x_k}{\|x\|_p}$ и $b_k = \dfrac{y_k}{\|y\|_q}$ $(k = 1, \ldots, n)$. Заметим, что

$$\sum_{k=1}^{n} a_k^p = \frac{1}{\|x\|_p^p} \sum_{k=1}^{n} x_k^p = 1 \quad \text{и} \quad \sum_{k=1}^{n} b_k^q = 1.$$

По теореме 8 неравенства $a_k\, b_k \leqslant \dfrac{a_k^p}{p} + \dfrac{b_k^q}{q}$ выполняются для всех $k = 1, \ldots, n$, причем равенство в них достигается лишь при условии $a_k^p = b_k^q$. Суммируя эти неравенства, мы получим

$$\sum_{k=1}^{n} a_k\, b_k \leqslant \frac{1}{p} \sum_{k=1}^{n} a_k^p + \frac{1}{q} \sum_{k=1}^{n} b_k^q = \frac{1}{p} + \frac{1}{q} = 1\,,$$

откуда $\sum\limits_{k=1}^{n} x_k\, y_k \leqslant \|x\|_p \cdot \|y\|_q$, что и дает неравенство Гёльдера. Равенство в нем реализуется тогда и только тогда, когда оно достигается *во всех сложенных неравенствах*. Поэтому при любых $k \in \{1, \ldots, n\}$ $a_k^p = b_k^q$, то есть $x_k^p\, \|y\|_q^q = y_k^q\, \|x\|_p^p$. $\square$

Следствие 1. Неравенство Гёльдера в $\mathbb{R}^n$. *Пусть $n \in \mathbb{N}$, $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ — векторы из $\mathbb{R}^n$, p и q — сопряженные показатели. Тогда*

$$\left|\sum_{k=1}^{n} x_k\, y_k\right| \leqslant \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \cdot \left(\sum_{k=1}^{n} |y_k|^q\right)^{1/q}.$$

Доказательство. Применяя теорему 10 к векторам с координатами $|x_k|$ и $|y_k|$ $(k = 1, \ldots, n)$, мы получим

$$\left|\sum_{k=1}^{n} x_k\, y_k\right| \leqslant \sum_{k=1}^{n} |x_k| \cdot |y_k| \leqslant \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \cdot \left(\sum_{k=1}^{n} |y_k|^q\right)^{1/q}. \quad \square$$

Следствие 2. Неравенство Коши в $\mathbb{R}^n$. *Пусть $n \in \mathbb{N}$, $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ — векторы из $\mathbb{R}^n$. Тогда*

$$\left|\sum_{k=1}^{n} x_k\, y_k\right| \leqslant \sqrt{\sum_{k=1}^{n} x_k^2} \cdot \sqrt{\sum_{k=1}^{n} y_k^2}\,.$$

Равенство достигается тогда и только тогда, когда векторы x и y коллинеарны.

Заметим, что для коллинеарных векторов неравенство Коши превращается в равенство. Содержательным является обратное утверждение.

Доказательство. Вновь можно считать векторы x и y ненулевыми. Заметим, что $\left|\sum\limits_{k=1}^{n} x_k\, y_k\right| \leqslant \sum\limits_{k=1}^{n} |x_k| \cdot |y_k|$, причем равенство означает, что все произведения $x_k \cdot y_k$ имеют одинаковый знак. Таким образом, либо для всех $k = 1, \ldots, n$ знаки x_k и y_k одинаковы, либо для всех k они противоположны. Применяя неравенство Гёльдера с показателями $p = q = 2$, мы получим

$$\sum_{k=1}^{n} |x_k| \cdot |y_k| \leqslant \sqrt{\sum_{k=1}^{n} x_k^2} \cdot \sqrt{\sum_{k=1}^{n} y_k^2}\,,$$

причем в случае равенства выполняются соотношения $x_k^2 = \lambda^2 y_k^2$ $(k = 1, \ldots, n)$, где $\lambda = \dfrac{\|x\|_2}{\|y\|_2}$. Если при всех $k = 1, \ldots, n$ знаки x_k и y_k одинаковы, то $x_k = \lambda\, y_k$, а если противоположны, то $x_k = -\lambda\, y_k$. В любом случае получается коллинеарность x и y. $\square$

Замечание. В $\mathbb{R}^2$ и $\mathbb{R}^3$ скалярное произведение векторов x и y определяется по формуле

$$x \cdot y = |x| \cdot |y| \cdot \cos\alpha,$$

где α — угол между x и y. Отсюда видно, что $|x \cdot y| \leqslant |x| \cdot |y|$, а равенство достигается только при $\alpha = 0$ или $\alpha = \pi$, то есть для коллинеарных векторов. Следствие 2 утверждает, что эти факты верны в $\mathbb{R}^n$ *для любого* $n \in \mathbb{N}$, а его доказательство проводится аналитически, без использования углов. Более того, неравенство Коши позволяет нам *корректно определить* угол $\alpha \in [0, \pi]$ между векторами x и y по формуле $\cos\alpha = \dfrac{x \cdot y}{|x| \cdot |y|}$.

Чтобы выпуклость стала еще более мощным инструментом для работы с неравенствами, перепишем ее определение в более общей форме.

Теорема 11. Неравенство Иенсена. *Пусть* $f\colon \langle A, B\rangle \to \mathbb{R}$, $n \in \mathbb{N}$. *Предположим, что* $x_1, \ldots, x_n \in \langle A, B\rangle$, *а* $\lambda_1, \ldots, \lambda_n$ —

положительные числа, удовлетворяющие условию $\lambda_1+\ldots+\lambda_n=1$. *Тогда справедливы следующие утверждения.*

1) *Если* f *выпукла на* $\langle A,B\rangle$, *то*

$$f\big(\lambda_1x_1+\ldots+\lambda_nx_n\big)\leqslant\lambda_1f(x_1)+\ldots+\lambda_nf(x_n).$$

2) *Если* f *строго выпукла на* $\langle A,B\rangle$, *то либо*

$$f\big(\lambda_1x_1+\ldots+\lambda_nx_n\big)<\lambda_1f(x_1)+\ldots+\lambda_nf(x_n),$$

либо $x_1=\ldots=x_n$ *(в последнем случае реализуется равенство).*

Замечания

1) Теорема будет верна и для вогнутых функций, если в ее формулировке заменить знаки неравенств на противоположные.

2) При $n=2$ неравенство Иенсена превращается в неравенство из определения выпуклости или строгой выпуклости.

3) В условиях теоремы точка $c=\lambda_1x_1+\ldots+\lambda_nx_n$ лежит в $\langle A,B\rangle$, а если среди чисел $x_1,\ldots,x_n$ есть различные, то $c\in(A,B)$. Действительно, пусть для определенности $x_1\leqslant\ldots\leqslant x_n$. Если $x_1<x_n$, то

$$c<x_n(\lambda_1+\ldots+\lambda_n)=x_n\quad\text{и}\quad c>x_1(\lambda_1+\ldots+\lambda_n)=x_1,$$

откуда $c\in(x_1,x_n)\subset(A,B)$. В противном случае $x_1=\ldots=x_n$, поэтому $c=x_n\in\langle A,B\rangle$.

Доказательство. Будем считать, что не все числа $x_1,\ldots,x_n$ одинаковые, иначе в неравенстве Иенсена реализуется равенство, и оба утверждения теоремы очевидны. Пусть $c=\lambda_1x_1+\ldots+\lambda_nx_n$. В силу замечания 3 точка c лежит в (A,B). Тогда по лемме 2 функция f имеет в точке c опорную прямую, а если f строго выпукла, то эта прямая будет строго опорной. Пусть она задается уравнением $y=kx+b$. Перейдем к доказательству утверждений теоремы.

1) По определению опорной прямой $f(c)=kc+b$ и $f(x_j)\geqslant kx_j+b$ при всех $j=1,\ldots,n$. Поэтому

$$\begin{aligned}f\big(\lambda_1x_1+\ldots+\lambda_nx_n\big)&=f(c)=\\&=kc+b=k\left(\lambda_1x_1+\ldots+\lambda_nx_n\right)+b\left(\lambda_1+\ldots+\lambda_n\right)=\\&=\lambda_1\left(kx_1+b\right)+\ldots+\lambda_n\left(kx_n+b\right)\leqslant\lambda_1f(x_1)+\ldots+\lambda_nf(x_n).\end{aligned}$$

2) Если прямая строго опорная, то в неравенствах $f(x_j) \geqslant kx_j+b$ $(j = 1, \ldots, n)$ равенство достигается только в случае $x_j = c$. Так как среди точек $x_1, \ldots, x_n$ есть различные, то хотя бы одно из этих неравенств будет строгим. Поэтому строгим окажется и итоговое неравенство. $\square$

В качестве приложения неравенства Иенсена рассмотрим *неравенство Коши для средних.* Пусть $x_1, \ldots, x_n \in \mathbb{R}$.

Определение 6. Среднее арифметическое и среднее геометрическое.

1) Величина $\frac{1}{n}(x_1 + \ldots + x_n)$ называется *средним арифметическим чисел* $x_1, \ldots, x_n$.

2) Если $x_1, \ldots, x_n \geqslant 0$, то $\sqrt[n]{x_1 \cdot \ldots \cdot x_n}$ называется *средним геометрическим* $x_1, \ldots, x_n$.

Теорема 12. Неравенство Коши для средних. *Для любых неотрицательных чисел* $x_1, \ldots, x_n$ *справедливо неравенство*

$$\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \leqslant \frac{1}{n}(x_1 + \ldots + x_n),$$

причем равенство реализуется только при условии $x_1 = \ldots = x_n$.

Доказательство. Если хотя бы одно из чисел равно нулю, то утверждение очевидно. Пусть $x_k > 0$ при всех $k = 1, \ldots, n$. Положим $f(x) = \ln x$. В силу вогнутости f из неравенства Иенсена мы получим

$$f\left(\frac{x_1 + \ldots + x_n}{n}\right) \geqslant \frac{f(x_1) + \ldots + f(x_n)}{n} \Leftrightarrow$$

$$\ln\left(\frac{x_1 + \ldots + x_n}{n}\right) \geqslant \frac{1}{n}\ln(x_1 \cdot \ldots \cdot x_n) \Leftrightarrow$$

$$\frac{x_1 + \ldots + x_n}{n} \geqslant \sqrt[n]{x_1 \cdot \ldots \cdot x_n},$$

а равенство достигается только при $x_1 = \ldots = x_n$. $\square$

ГЛАВА 4. ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ВЕЩЕСТВЕННОЙ ПЕРЕМЕННОЙ

§ 1. Первообразная функция и неопределенный интеграл

В предыдущей главе мы изучили операцию дифференцирования, сопоставляющую функции ее производную, и вывели правила, позволяющие вычислить производную любой элементарной функции. Здесь мы рассмотрим *обратную задачу*, в которой производная известна, а функцию требуется найти.

Определение 1. Первообразная функции на промежутке. Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$. Функция $F\colon \langle A, B\rangle \to \mathbb{R}$ называется *первообразной* f, если F дифференцируема на $\langle A, B\rangle$ и $F'(x) = f(x)$ при всех $x \in \langle A, B\rangle$.

Для функции $f\colon \langle A, B\rangle \to \mathbb{R}$ обозначим через $\mathcal{P}_f\big(\langle A, B\rangle\big)$ множество всех ее первообразных на $\langle A, B\rangle$. Чтобы изучить понятие первообразной, необходимо обсудить следующие вопросы, связанные с $\mathcal{P}_f\big(\langle A, B\rangle\big)$.

1) Для каких функций f множество $\mathcal{P}_f\big(\langle A, B\rangle\big)$ непусто?

2) Как описать множество $\mathcal{P}_f\big(\langle A, B\rangle\big)$, если оно непусто?

3) Как вычислять первообразные функции f на $\langle A, B\rangle$, если они существуют?

Наиболее простым является вопрос 2), поэтому рассмотрим его первым. Мы докажем, что если первообразная f на $\langle A, B\rangle$ существует, то она единственна с точностью до постоянного слагаемого.

Теорема 1. Описание первообразных функции на промежутке. *Пусть f, F и G — функции, заданные на $\langle A, B\rangle$. Предположим, что $F \in \mathcal{P}_f\big(\langle A, B\rangle\big)$. Тогда равносильны следующие утверждения.*

1) $G \in \mathcal{P}_f\big(\langle A, B\rangle\big)$.

2) *Существует такое число c, что $G(x) = F(x) + c$ при всех $x \in \langle A, B\rangle$.*

Иными словами, $\mathcal{P}_f\big(\langle A, B\rangle\big) = \{F + c\colon\ c \in \mathbb{R}\}$. Таким образом, для описания $\mathcal{P}_f\big(\langle A, B\rangle\big)$ достаточно найти *какую-либо одну* первообразную f.

Доказательство.

1) $\Rightarrow$ 2) Пусть $H = G - F$. Тогда

$$H'(x) = G'(x) - F'(x) = f(x) - f(x) = 0 \text{ при всех } x \in \langle A, B\rangle.$$

В силу критерия постоянства функции (см. § 7 главы 3) функция H постоянна на $\langle A, B\rangle$, а это и нужно доказать.

2) $\Rightarrow$ 1) $G'(x) = F'(x) = f(x)$ при всех $x \in \langle A, B\rangle$. $\square$

Полученное описание $\mathcal{P}_f\big(\langle A, B\rangle\big)$ позволяет нам дать следующее определение.

Определение 2. Неопределенный интеграл функции на промежутке. Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, $F \in \mathcal{P}_f\big(\langle A, B\rangle\big)$. Множество функций $\{F(x)+c\colon\ c \in \mathbb{R}\}$ называется *неопределенным интегралом f на $\langle A, B\rangle$* и обозначается $\int f(x)\,dx$.

Фактически $\int f(x)\,dx$ есть множество $\mathcal{P}_f\big(\langle A, B\rangle\big)$ всех первообразных f на $\langle A, B\rangle$, но запись $\{F(x) + c\colon\ c \in \mathbb{R}\}$ удобнее, поскольку она дает явное описание этого множества. Для краткости принято писать просто

$$\int f(x)\,dx = F(x) + c,$$

опуская фигурные скобки и понимая под c произвольное вещественное число.

Обозначение $\int f(x)\,dx$ требует пояснений. Оно используется по аналогии с обозначением для *определенного интеграла*, который тесно связан с неопределенным. Грубо говоря, определенный интеграл есть предел сумм вида $\sum\limits_k f(x_k)\,\Delta x_k$, где Δx_k — "малые приращения" аргумента f. Таким образом, выражение $f(x)\,dx$ похоже по виду на слагаемые в этих суммах, а значок $\int$ напоминает латинскую букву "S", указывающую на операцию суммирования. Для неопределенного интеграла символ dx не несет никакой смысловой нагрузки, как и аргумент функции f, и можно было бы писать просто $\int f$. Тем не менее запись $\int f(x)\,dx$ удобна по двум причинам.

1°. При наличии у функции f параметров символ dx позволяет указать, что переменной является именно x. Например, запись $\int x^y\,dx$ означает, что подынтегральная функция имеет вид $x \mapsto x^y$, а y является параметром. Выражение $\int x^y\,dy$ имеет совсем другой смысл, здесь мы интегрируем функцию $y \mapsto x^y$, а параметром является x.

2°. Трактуя *dx как дифференциал*, можно сделать более простыми и наглядными формулы *замены переменной* и *интегрирования по частям*, которые будут рассмотрены ниже.

Хотя неопределенный интеграл является не функцией, а множеством функций, на него можно распространить некоторые операции, введенные ранее для функций.

1) *Дифференцирование неопределенного интеграла.* Пусть функция $f\colon \langle A, B\rangle \to \mathbb{R}$ имеет первообразную на $\langle A, B\rangle$. Так как для любой первообразной F функции f равенство $F' = f$ справедливо на $\langle A, B\rangle$, мы можем положить

$$\left(\int f(x)\,dx\right)' = f(x) \quad \text{при} \quad x \in \langle A, B\rangle.$$

2) *Арифметические действия с неопределенными интегралами.* В § 2 главы 2 для числовых множеств определялись операции сложения и умножения на число. Дадим аналогичное определение для множеств функций, заданных на $\langle A, B\rangle$. Пусть D и E — множества функций, действующих из $\langle A, B\rangle$ в $\mathbb{R}$, $h\colon \langle A, B\rangle \to \mathbb{R}$. Положим

$$\begin{aligned} D + F &= \{f + g\colon f \in D, g \in E\},\\ E + h &= \{f + h\colon f \in E\},\\ \lambda E &= \{\lambda f\colon f \in E\} \quad (\lambda \in \mathbb{R}). \end{aligned}$$

Так как неопределенные интегралы тоже представляют собой множества функций, мы можем применить к ним эти правила. Пусть $f, g\colon \langle A, B\rangle \to \mathbb{R}$, F, G — их первообразные на $\langle A, B\rangle$, $H\colon \langle A, B\rangle \to \mathbb{R}$. Предлагаем читателю проверить справедливость следующих равенств:

$$\int f(x)\,dx + \int g(x)\,dx = \{F(x) + G(x) + c\colon\ c \in \mathbb{R}\}\ ,$$

$$\int f(x)\,dx + H(x) = \{F(x) + H(x) + c\colon\ c \in \mathbb{R}\}\ ,$$

$$\lambda \int f(x)\,dx = \{\lambda F(x) + c\colon\ c \in \mathbb{R}\} \quad (\lambda \neq 0)\ .$$

Перейдем теперь к вопросу *о существовании* первообразной. Задача описания функций, имеющих первообразную, слишком сложна. Мы рассмотрим по отдельности необходимое и достаточное условия существования первообразной. Начнем с необходимого условия. В § 7 главы 3 было показано, что производная функции, дифференцируемой на промежутке, не может иметь скачков ни в одной из точек этого промежутка (следствие 3 теоремы Дарбу). Значит, этому требованию должна удовлетворять и функция, имеющая первообразную на промежутке. Например, функция

$$f(x) = \begin{cases} 1, & x \geqslant 0 \\ -1, & x < 0 \end{cases}$$

имеет разрыв первого рода в нуле, поэтому у нее не существует первообразной на $\mathbb{R}$. С другой стороны, разрывы второго рода у производной могут быть (см. замечание к теореме Дарбу). Поэтому непрерывность функции не является необходимым условием существования у нее первообразной.

Рассмотрим теперь достаточное условие существования первообразной. Оказывается, что любая непрерывная на промежутке функция имеет на нем первообразную. Как показано выше, необходимым это условие не будет.

Теорема 2. Достаточное условие существования первообразной. *Если функция f непрерывна на $\langle A, B\rangle$, то у нее на $\langle A, B\rangle$ есть первообразная.*

Доказательство теоремы основано на понятии определенного интеграла, поэтому мы не будем приводить его здесь.

Поскольку элементарные функции непрерывны на своих областях определения, с помощью теоремы 2 можно доказывать существование их первообразных. Однако применять теорему непосредственно можно только к тем элементарным функциям, которые определены *на промежутке*. Чтобы обойти эту трудность, обобщим определение 1.

Определение 3. Первообразная функции на множестве. Пусть $E \subset \mathbb{R}$, $f\colon E \to \mathbb{R}$. Если функция F дифференцируема на E и $F'(x) = f(x)$ при всех $x \in E$, то F называется *первообразной f на множестве E.*

Покажем, что задача нахождения первообразной f на E сводится к поиску первообразной на промежутке. Если f имеет на E первообразную F, то множество E будет для F множеством дифференцируемости, поэтому для любой точки $a \in E$ существует такое $\delta > 0$, что $E \cap (a - \delta, a + \delta)$ является невырожденным промежутком (см. § 5 главы 3). Будем предполагать это условие выполненным. Назовем промежуток $\langle A, B\rangle \subset E$ *максимальным для* E, если не существует промежутка $\langle A_1, B_1\rangle \subset E$, для которого $\langle A, B\rangle \subsetneqq \langle A_1, B_1\rangle$. Заметим, что *если промежутки* $\langle A, B\rangle$ *и* $\langle C, D\rangle$ *максимальны для* E *и различны, то* $\langle A, B\rangle \cap \langle C, D\rangle = \varnothing$. Действительно, если $\langle A, B\rangle \cap \langle C, D\rangle \neq \varnothing$, то $\langle A, B\rangle \cup \langle C, D\rangle$ будет промежутком, лежащим в E и содержащим $\langle A, B\rangle$ и $\langle C, D\rangle$. Поэтому $\langle A, B\rangle = \langle A, B\rangle \cup \langle C, D\rangle = \langle C, D\rangle$ ввиду максимальности $\langle A, B\rangle$ и $\langle C, D\rangle$.

Лемма 1. *Пусть F — дифференцируемая на E функция. Тогда справедливы следующие утверждения.*

1) *Множество E есть объединение своих максимальных промежутков.*

2) *Пусть $G = F\big|_{\langle A,B\rangle}$, где $\langle A, B\rangle$ — максимальный промежуток для E. Тогда $F'(a) = G'(a)$ для любого $a \in \langle A, B\rangle$.*

Доказательство.

1) Достаточно проверить, что E содержится в объединении своих максимальных промежутков, поскольку обратное включение очевидно. Пусть $a \in E$. Обозначим через Ω множество всех таких невырожденных промежутков Δ, что $\Delta \subset E$ и $a \in \Delta$. Заметим, что $\Omega \neq \varnothing$, поскольку $E \cap (a - \delta, a + \delta) \in \Omega$ при некотором $\delta > 0$. Положим $D = \bigcup\{\Delta\colon\ \Delta \in \Omega\}$. Докажем, что D есть промежуток. Если $x \in D$, то $x \in \Delta$ при некотором $\Delta \in \Omega$. Тогда отрезок с концами a и x содержится в Δ, а значит, и в D. Поэтому D является промежутком. Максимальность D следует из его определения.

2) Пусть $a \in \langle A, B\rangle$. Существует такое $\delta > 0$, что $E \cap (a-\delta, a+\delta)$ есть невырожденный промежуток (обозначим его через Δ). Тогда $\Delta \subset \langle A, B\rangle$, иначе $\langle A, B\rangle \subsetneqq \langle A, B\rangle \cup \Delta \subset E$, что противоречит максимальности $\langle A, B\rangle$. Поэтому $F\,|_\Delta \equiv G\,|_\Delta$, откуда $F'(a) = G'(a)$. $\square$

Лемма 1 показывает, что для нахождения первообразной некоторой функции на E достаточно найти ее на каждом максимальном промежутке. В частности, если функция f непрерывна на E, то у нее есть первообразная на E. Однако описание $\mathcal{P}_f\big(E\big)$ будет более сложным, поскольку постоянные слагаемые на разных максимальных промежутках могут не совпадать. Пусть $F \in \mathcal{P}_f\big(E\big)$ и $G\colon E \to \mathbb{R}$. Справедливо следующее утверждение: *G является первообразной f на E тогда и только тогда, когда функция $G - F$ постоянна на каждом максимальном промежутке в E.*

Приведем пример, иллюстрирующий сказанное. Пусть $f(x) = \frac{1}{x}$. Функция f определена на множестве $E = \mathbb{R}\backslash\{0\}$, имеющем два максимальных промежутка: $(-\infty, 0)$ и $(0, +\infty)$. Покажем, что функция $F(x) = \ln|x|$ является первообразной f на E. Действительно,

$$\big(F\,\big|_{(0,+\infty)}\big)'\,(x) = \big(\ln x\big)' = \frac{1}{x}\,,$$
$$\big(F\,\big|_{(-\infty,0)}\big)'\,(x) = \big(\ln(-x)\big)' = \frac{1}{-x}\cdot(-1) = \frac{1}{x}\,.$$

Поэтому

$$\int \frac{dx}{x} = \ln|x| + \begin{cases} c_1, & x > 0 \\ c_2, & x < 0 \end{cases} \quad (c_1, c_2 \in \mathbb{R}).$$

Тем не менее мы будем по-прежнему писать $\displaystyle\int \frac{dx}{x} = \ln|x| + c$, подразумевая под c не константу, а кусочно-постоянную функцию.

Обратимся теперь к задаче *вычисления* первообразных элементарных функций. Прежде всего выпишем таблицу простейших интегралов, к которым с помощью различных приемов будут сводиться другие, более сложные интегралы.

1) $\int a\,dx = ax + c \quad (a \in \mathbb{R})$.

2) $\displaystyle\int x^a\,dx = \frac{x^{a+1}}{a+1} + c \quad (a \neq -1)$.

3) $\displaystyle\int \frac{dx}{x} = \ln|x| + c$.

4) $\int e^x\,dx = e^x + c$.

5) $\displaystyle\int a^x\,dx = \frac{a^x}{\ln a} + c \quad (a > 0,\ a \neq 1)$.

6) $\int \sin x\,dx = -\cos x + c$.

7) $\int \cos x\,dx = \sin x + c$.

8) $\displaystyle\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + c$.

9) $\displaystyle\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + c$.

10) $\displaystyle\int \frac{dx}{x^2+a^2} = \frac{1}{a}\operatorname{arctg}\frac{x}{a} + c \quad (a \neq 0)$.

11) $\displaystyle\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + c \quad (a \neq 0)$.

12) $\displaystyle\int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin\frac{x}{a} + c \quad (a > 0)$.

13) $\displaystyle\int \frac{dx}{\sqrt{x^2+a}} = \ln\left|x + \sqrt{x^2+a}\right| + c \quad (a \in \mathbb{R})$.

Для проверки любого из этих равенств надо продифференцировать его правую часть и убедиться в том, что получилась подынтегральная функция, стоящая в левой части.

Замечания

1) Все равенства, кроме 11) и 13), фактически получаются из таблицы § 3 главы 3 перестановкой столбцов, содержащих f и f'. Функции, стоящие в правой части равенств 11) и 13), принято называть соответственно "высоким логарифмом" и "длинным логарифмом".

2) Буква c в приведенной таблице имеет следующий смысл: если подынтегральная функция определена на промежутке, то c является произвольным вещественным числом, а для функций, заданных на объединении максимальных промежутков, c обозначает функцию, постоянную на каждом из таких промежутков.

В главе 3 мы выяснили, что операция дифференцирования любой элементарной функции сопоставляет элементарную функцию. Верно ли это для операции интегрирования, обратной по отношению к дифференцированию? Поставим задачу более формально. Пусть f — элементарная функция, определенная на некотором промежутке $\langle A, B\rangle$. Мы уже знаем, что f имеет на $\langle A, B\rangle$ первообразную. Обязана ли она быть *элементарной*? Ответ на этот вопрос отрицательный. Если функция f имеет элементарную первообразную, то интеграл $\int f(x)\,dx$ называют *берущимся*, в противном случае — *неберущимся*. Ниже мы поговорим о берущихся и неберущихся интегралах более подробно.

Поскольку не все интегралы от элементарных функций берутся, не существует и универсального способа вычисления первообразных этих функций, в отличие от правил дифференцирования, доказанных в § 3 главы 3. Тем не менее мы выведем несколько свойств неопределенного интеграла, позволяющих в ряде случаев первообразные находить.

Теорема 3. Линейность неопределенного интеграла. *Пусть функции $f, g\colon \langle A, B\rangle \to \mathbb{R}$ имеют на $\langle A, B\rangle$ первообразные. Тогда для любых $\alpha, \beta \in \mathbb{R} \setminus \{0\}$*

$$\int \bigl(\alpha f(x) + \beta g(x)\bigr)\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx.$$

Напомним, что операции сложения и умножения на ненулевые числа для неопределенных интегралов были введены выше.

Доказательство. Пусть F, G — первообразные на $\langle A, B\rangle$ функций f, g соответственно. В правой части стоит множество функций $\{\alpha F(x) + \beta G(x) + c\colon c \in \mathbb{R}\}$. Так как

$$\bigl(\alpha F(x) + \beta G(x) + c\bigr)' = \alpha F'(x) + \beta G'(x) = \alpha f(x) + \beta g(x),$$

то это множество совпадает с $\int \bigl(\alpha f(x) + \beta g(x)\bigr)\,dx$. $\square$

Докажем теперь формулу замены переменной в интеграле, которая является основным инструментом нахождения первообразных.

Теорема 4. Замена переменной в неопределенном интеграле. *Пусть $f\colon \langle A, B\rangle \to \mathbb{R}$, F — первообразная f на $\langle A, B\rangle$, $\varphi\colon \langle C, D\rangle \to \langle A, B\rangle$ — дифференцируемая на $\langle C, D\rangle$ функция. Тогда*

$$\int f\bigl(\varphi(x)\bigr) \cdot \varphi'(x)\,dx = F\bigl(\varphi(x)\bigr) + c.$$

Доказательство. По правилу цепочки (см. § 3 главы 3)

$$\big(F\big(\varphi(x)\big)+c\big)' = F'\big(\varphi(x)\big)\cdot\varphi'(x) = f\big(\varphi(x)\big)\cdot\varphi'(x),$$

что и доказывает теорему. □

Утверждение теоремы можно трактовать следующим образом. Положим $y=\varphi(x)$. Тогда $dy=\varphi'(x)\,dx$, где dx — приращение аргумента функции φ (см. § 2 главы 3). Формальный переход от x к y в левой части равенства дает $\int f(y)\,dy$. Вычисляя этот интеграл по переменной y, мы получим $F(y)+c$, то есть $F\big(\varphi(x)\big)+c$. Эти рассуждения объясняют смысл термина “замена переменной” и позволяют упростить практическое применение формулы, хотя они основаны лишь на интуитивном понимании символа dx как дифференциала.

В теореме 4 мы выражали первообразную функции $\big(f\circ\varphi\big)\cdot\varphi'$ через первообразную f. Если функция φ имеет обратную, то можно эту формулу использовать и в другую сторону, для нахождения первообразной f. Сформулируем теорему в такой редакции.

Следствие. *Пусть в условиях теоремы* 4 φ *имеет обратную функцию* ψ, *определенную на* $\langle A,B\rangle$. *Если функция* G *является первообразной* $\big(f\circ\varphi\big)\cdot\varphi'$, *то*

$$\int f(x)\,dx = G\big(\psi(x)\big)+c.$$

Доказательство. Пусть F — первообразная f на $\langle A,B\rangle$. Тогда функция $F\big(\varphi(y)\big)$ является первообразной для $f\big(\varphi(y)\big)\cdot\varphi'(y)$ по теореме 4, а функция $G(y)$ — по условию. Поэтому разность $F\big(\varphi(y)\big)-G(y)$ постоянна на $\langle C,D\rangle$. Подставляя $y=\psi(x)$ при $x\in\langle A,B\rangle$, мы получим постоянство функции $F(x)-G\big(\psi(x)\big)$ на $\langle A,B\rangle$. По теореме 1 $G\circ\psi$ является первообразной f на $\langle A,B\rangle$, то есть $\int f(x)\,dx = G\big(\psi(x)\big)+c$. □

Второй вариант формулы замены переменной тоже будет нагляднее, если трактовать dx как дифференциал. Полагая $x=\varphi(y)$, мы получим $dx=d\varphi(y)=\varphi'(y)\,dy$. Поэтому

$$\int f(x)\,dx = \int f\big(\varphi(y)\big)\cdot\varphi'(y)\,dy = G(y)+c = G\big(\psi(x)\big)+c.$$

Примеры

1) *Вычислить интеграл* $\int \operatorname{ctg} x\,dx$.
Применим теорему 4. Полагая $y = \sin x$, мы получим

$$\int \operatorname{ctg} x\,dx = \int \frac{\cos x\,dx}{\sin x} = \\ = \int \frac{d\sin x}{\sin x} = \int \frac{dy}{y} = \ln|y| + c = \ln|\sin x| + c.$$

2) *Вычислить интеграл* $\displaystyle\int \frac{dx}{1+\sqrt{x}}$.
Воспользуемся следствием теоремы 4. Положим $y = \sqrt{x}$. Тогда $x = y^2$, откуда $dx = 2y\,dy$ и

$$\int \frac{dx}{1+\sqrt{x}} = \int \frac{2y\,dy}{1+y} = \\ = 2\int \left(1 - \frac{1}{1+y}\right) dy = 2\int dy - 2\int \frac{d(y+1)}{y+1} = \\ = 2y - 2\ln|y+1| + c = 2\sqrt{x} - \ln(\sqrt{x}+1)^2 + c,$$

поскольку $d(y+1) = (y+1)'\,dy = dy$.

Докажем теперь *формулу интегрирования по частям*, которая дает еще один важный прием вычисления первообразных.

Теорема 5. Формула интегрирования по частям. *Пусть* $f, g \in C^1\big(\langle A, B\rangle\big)$. *Тогда*

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int g(x)f'(x)\,dx.$$

Замечания

1) Согласно правилу сложения функции с интегралом, сформулированному выше, правая часть равенства есть

$$\{f(x)g(x) - H(x) + c\colon\ c \in \mathbb{R}\},$$

где H — первообразная функции $g(x)f'(x)$.

2) Подынтегральные функции непрерывны и, в силу теоремы 2, имеют первообразные.

3) Если трактовать символ dx как дифференциал, то утверждение теоремы можно записать в следующем виде:

$$\int f(x)\,dg(x) = f(x)g(x) - \int g(x)\,df(x).$$

Доказательство. Пусть H — первообразная функции $g \cdot f'$. Тогда

$$\begin{aligned}\big(f(x)g(x) - H(x)\big)' =& \big(f(x)g(x)\big)' - H'(x) = \\ =& f(x)g'(x) + g(x)f'(x) - g(x)f'(x) = f(x)g'(x).\end{aligned}$$

Таким образом, множество $\{f(x)g(x) - H(x) + c\colon\ c \in \mathbb{R}\}$ совпадает с $\int f(x)g'(x)\,dx$. Отсюда в силу замечания 1 и вытекает искомое равенство. □

Примеры

1) *Вычислить интеграл* $\int \ln x\,dx$.

Пусть $f(x) = \ln x$ и $g(x) = x$. Тогда $df(x) = (\ln x)'\,dx = \frac{dx}{x}$ и $dg(x) = dx$. По теореме 5 мы получаем

$$\int \ln x\,dx = x\ln x - \int x \cdot \frac{dx}{x} = x\ln x - \int dx = x\ln x - x + c.$$

2) *Пусть* $a \in \mathbb{R}$, $a \neq 0$, $I_n = \int \frac{dx}{(x^2+a)^n}$ $(n \in \mathbb{N})$. *Выразить интеграл* I_{n+1} *через* I_n *для произвольного натурального* n.

Обозначим $f(x) = \dfrac{1}{(x^2+a)^n}$ и $g(x) = x$. Тогда

$$df(x) = \left(\frac{1}{(x^2+a)^n}\right)' dx = -\frac{2nx}{(x^2+a)^{n+1}}\,dx \quad \text{и} \quad dg(x) = dx.$$

По теореме 5

$$\begin{aligned}I_n =& \frac{x}{(x^2+a)^n} + 2n\int \frac{x^2}{(x^2+a)^{n+1}}\,dx = \\ =& \frac{x}{(x^2+a)^n} + 2n\int \frac{x^2+a-a}{(x^2+a)^{n+1}}\,dx = \\ =& \frac{x}{(x^2+a)^n} + 2n\int \frac{dx}{(x^2+a)^n} - 2na\int \frac{dx}{(x^2+a)^{n+1}} = \\ =& \frac{x}{(x^2+a)^n} + 2nI_n - 2naI_{n+1}\,,\end{aligned}$$

откуда $2naI_{n+1} = (2n-1)I_n + \dfrac{x}{(x^2+a)^n}$.

Замечание. В примере 2 мы получили равенство

$$2na \int \frac{dx}{(x^2+a)^{n+1}} = (2n-1) \int \frac{dx}{(x^2+a)^n} + \frac{x}{(x^2+a)^n} \quad (n \in \mathbb{N}),$$

которое называется *формулой понижения.* Применяя эту формулу n раз, мы сможем выразить интеграл $\displaystyle\int \frac{dx}{(x^2+a)^{n+1}}$ через табличный интеграл $\displaystyle\int \frac{dx}{x^2+a}$.

Приведенные выше примеры демонстрируют различные методы вычисления первообразных, но не дают представления о том, какие интегралы берутся, а какие — нет. Изучим этот вопрос более систематически. Существует несколько классов элементарных функций, для которых доказано, что интеграл от каждой функции из этого класса берется, и разработаны методы вычисления интегралов. Наиболее важным является *класс рациональных функций*, то есть функций, представимых в виде отношения двух многочленов. Рассмотрим его подробнее.

Теорема 6. *Любая рациональная функция имеет элементарную первообразную.*

Мы не будем приводить строгого доказательства теоремы, но укажем способ поиска первообразной. Назовем рациональную функцию *простейшей дробью*, если она выражается одной из следующих формул:

1) $\dfrac{a}{(x+p)^n}$, где $n \in \mathbb{N}$, $a, p \in \mathbb{R}$;

2) $\dfrac{ax+b}{(x^2+px+q)^n}$, где $n \in \mathbb{N}$, $a, b, p, q \in \mathbb{R}$, причем $p^2 < 4q$.

В курсе алгебры доказывается, что любую рациональную функцию можно представить в виде суммы многочлена и нескольких простейших дробей. Интегралы от многочленов и простейших дробей первого рода вычисляются непосредственно по таблице. Для простейших дробей второго рода мы вначале кратко опишем алгоритм поиска первообразной, а затем поясним его на примере.

1°. Если коэффициент p отличен от нуля, то от него можно избавиться, собрав в знаменателе полный квадрат и выполнив замену переменной $y = x + \frac{p}{2}$. В случае $p = 0$ запишем искомый интеграл в виде

$$a \int \frac{x\,dx}{(x^2+q)^n} + b \int \frac{dx}{(x^2+q)^n}.$$

2°. Интеграл $\displaystyle\int \frac{x\,dx}{(x^2+q)^n}$ можно вычислить с помощью замены $y = x^2 + q$, так как $dy = 2x\,dx$.

3°. Применяя к интегралу $I_n = \displaystyle\int \frac{dx}{(x^2+q)^n}$ доказанную выше формулу понижения $n-1$ раз, мы сведем его к интегралу I_1, который является табличным.

Пример. *Вычислить интеграл* $\displaystyle\int \frac{2x\,dx}{(x^2+2x+2)^2}$.

Применим к этому интегралу описанный выше алгоритм. Вначале положим $y = x + 1$. Тогда $dy = dx$, и

$$\int \frac{2x\,dx}{(x^2+2x+2)^2} = \int \frac{\big(2(x+1)-2\big)\,dx}{\big((x+1)^2+1\big)^2} =$$
$$= \int \frac{(2y-2)\,dy}{(y^2+1)^2} = 2\int \frac{y\,dy}{(y^2+1)^2} - 2\int \frac{dy}{(y^2+1)^2}.$$

Для первого интеграла используем замену $t = y^2 + 1$. Так как $dt = 2y\,dy$, то

$$2\int \frac{y\,dy}{(y^2+1)^2} = \int \frac{dt}{t^2} = -\frac{1}{t} + c = -\frac{1}{y^2+1} + c = -\frac{1}{x^2+2x+2} + c.$$

Для второго интеграла воспользуемся формулой понижения при $n = 1$ и $a = 1$:

$$2\int \frac{dy}{(y^2+1)^2} = \int \frac{dy}{y^2+1} + \frac{y}{y^2+1} =$$
$$= \operatorname{arctg} y + \frac{y}{y^2+1} + c = \operatorname{arctg}(x+1) + \frac{x+1}{x^2+2x+2} + c.$$

Вычитая второе равенство из первого, мы получим ответ:

$$\int \frac{2x\,dx}{(x^2+2x+2)^2} = -\operatorname{arctg}(x+1) - \frac{x+2}{x^2+2x+2} + c.$$

Другие классы берущихся интегралов обычно сводятся к интегралам от рациональных функций с помощью подходящих подстановок. Мы не будем их подробно описывать.

В заключение скажем несколько слов про *неберущиеся интегралы.* Через такие интегралы выражаются многие *специальные функции*, имеющие важное значение в математике. Выпишем некоторые из них:

1) *интегральный синус* $\displaystyle\int \frac{\sin x}{x}\,dx$;
2) *интегральный косинус* $\displaystyle\int \frac{\cos x}{x}\,dx$;
3) *интегральный логарифм* $\displaystyle\int \frac{dx}{\ln x}$;
4) *интегральная показательная функция* $\displaystyle\int \frac{e^x}{x}\,dx$;
5) *интегралы Френеля* $\int \sin x^2\,dx$ *и* $\int \cos x^2\,dx$;
6) *интеграл вероятностей* $\int e^{-x^2}\,dx$.

Доказательство того, что эти и некоторые другие интегралы не берутся, проводится с помощью алгебраической техники, выходящей за рамки нашего курса.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ