Механико-математический факультет МГУ имени М. В. Ломоносова

Примерный конспект курса математического анализа (первый семестр)

А. В. Бегунц, С. В. Шапошников

2019 год

Множества, функции, отношения

Множество — совокупность, набор, собрание некоторых объектов, называемых элементами множества.

Запись $a \in A$ означает, что a принадлежит множеству A.

При задании множеств обычно применяют два подхода:

1) перечисление элементов (например, $A = \{a, b, c\}$);

2) указание характеристического свойства, которым обладают те и только те объекты, которые являются элементами данного множества (например, A состоит из первых трёх букв латинского алфавита); при задании множества через характеристическое свойство часто будем использовать следующую форму записи:

$$A = \{a\colon \text{характеристическое свойство}\}.$$

Каким бы способом ни было задано множество A и каков бы ни был данный объект a, должно быть верно только одно из утверждений $a \in A$ или $a \notin A$.

Парадокс Рассела

Определим множество M характеристическим свойством: это набор всех тех и только тех множеств, которые не содержат себя в качестве элемента. (Вообще говоря, одно множество может содержать другое в качестве элемента, например, $\{1, 2\} \in \{1, \{1, 2\}\}$.) Проверим, содержит ли M себя в качестве элемента.

Если M является своим элементом, то по определению M оно не должно содержать себя в качестве элемента. Противоречие. Если M не является своим элементом, то оно удовлетворяет характеристическому свойству, которым задано множество M, и должно принадлежать M. Опять пришли к противоречию. Итак, мы не можем проверить, является ли M своим элементом или нет.

Другая форма этого парадокса называется «парадоксом брадобрея». Командир некоторой части приказал брадобрею (солдату этой же части) брить тех и только тех солдат, которые не бреются сами. Брадобрей не смог понять, что делать с самим собой. Если он будет брить себя, то окажется, что брадобрей бреется сам и нарушает приказ командира. Если брадобрей не бреет себя, то он является солдатом, который не бреется сам, и опять нарушает приказ.

Как и в предыдущей ситуации, проблема состоит в том, что брадобрей не может проверить, является ли он элементом множества, которое определил командир. Для устранения парадоксальных ситуаций такого вида была создана аксиоматика теории множеств (система аксиом Цермело — Френкеля ZF или ZFC), обсуждение которой выходит за рамки настоящего курса. Отметим, что все действия и способы определения множеств в нашем курсе согласуются с аксиоматикой теории множеств.

Подножества, утверждения, пары

Множество A является *подмножеством* множества B (пишут $A \subset B$), если из того, что $a \in A$, следует, что $a \in B$. Отметим, что $A \subset A$ и из включений $A \subset B$ и $B \subset C$ следует, что $A \subset C$. Говорят, что множества A и B *равны* (пишут $A = B$), если верны включения $A \subset B$ и $B \subset A$.

Важный пример множества — *пустое множество* $\varnothing$, про которое можно утверждать, что ни один объект не является его элементом. Пустое множество $\varnothing$ является подмножеством всякого множества, поскольку невозможно указать элемент, который есть в пустом множестве, но отсутствует в каком-то другом.

Основные утверждения настоящего курса имеют вид «если P, то Q» (краткая запись: «$P \Rightarrow Q$») и «P тогда и только тогда, когда Q» (кратко: «$P \Leftrightarrow Q$»), где P и Q — некоторые высказывания, называемые *посылкой* и *заключением* соответственно. Если справедливо утверждение «$P \Rightarrow Q$», то говорят также, что «Q *необходимо* для P» и «Q является *свойством* P», а «P *достаточно* для Q» и «P является *признаком* Q». Если справедливо утверждение «$P \Leftrightarrow Q$», то говорят, что «P *необходимо и достаточно* для Q», а само утверждение называют *критерием*. Чтобы опровергнуть утверждение вида «$P \Rightarrow Q$», достаточно предъявить *контрпример*, при котором верна посылка P и неверно заключение Q. Доказательство того, что контрпример не существует (см. рассуждение выше о пустом множестве), равносильно доказательству самого утверждения.

Неупорядоченной парой элементов x и y называется множество $\{x, y\}$. *Упорядоченной парой* элементов x и y, в которой x является первым элементом, а y — вторым элементом, называется множество $\{x, \{x, y\}\}$. Далее упорядоченную пару будем обозначать через (x, y). *Декартовым произведением* множеств X и Y называется множество $X \times Y = \{(x, y)\colon x \in X \text{ и } y \in Y\}$.

Функции и операции

Функцией f из множества X в множество Y (далее пишем $f\colon X \to Y$) называется зависимость (правило, закон), сопоставляющая каждому элементу $x \in X$ ровно один элемент $y \in Y$, обозначаемый через $y = f(x)$ (также пишут $f\colon x \mapsto y$) и называемый *значением* функции f на *аргументе* x (или *в точке* x).

Множество X называется *областью определения* функции f, а подмножество множества Y, состоящее из всех значений функции f на аргументах, пробегающих множество X, называется *областью значений* функции f.

Графиком функции f называется множество

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

Таким образом, $\Gamma_f \subset X \times Y$. Обратно, подмножество $\Gamma \subset X \times Y$ является графиком некоторой функции из X в Y тогда и только тогда, когда для всякого элемента $x \in X$ найдётся такой элемент y, что $(x, y) \in \Gamma$, причём если $(x, y) \in \Gamma$ и $(x, z) \in \Gamma$, то $y = z$.

Используя график, можно определить функцию на языке теории множеств. Говорят, что задана функция из X в Y, если задано такое подмножество $\Gamma \subset X \times Y$, что для всякого элемента $x \in X$ найдётся такой элемент y, что $(x, y) \in \Gamma$ и если $(x, y) \in \Gamma$ и $(x, z) \in \Gamma$, то $y = z$. Для всякой пары $(x, y) \in G$ элемент y обозначают через $f(x)$. Множество всех функций $f\colon X \to Y$ обозначается через Y^X.

Композицией функций $f\colon X \to Y$ и $g\colon Z \to X$ называется функция $f \circ g\colon Z \to Y$, заданная соотношением $f \circ g(z) = f(g(z))$. Сопоставление паре функций f и g новой функции $f \circ g$ называют *операцией композиции*.

Теорема 1. *Операция композиции ассоциативна*:

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

$\triangleright$ По определению справедливы равенства

$$f \circ (g \circ h)(x) = f(g \circ h(x)) = f(g(h(x))) = f \circ g(h(x)) = (f \circ g) \circ h(x). \qquad \triangleleft$$

Говорят, что на множестве X задана алгебраическая операция $\circ$, если задана функция $\circ\colon X \times X \to X$. Операция $\circ$ называется *ассоциативной*, если для всех $a, b, c \in X$ верно равенство $(a \circ b) \circ c = a \circ (b \circ c)$. Операция $\circ$ называется *коммутативной*, если для всех $a, b \in X$ верно равенство $a \circ b = b \circ a$.

Например, на множестве X^X определена ассоциативная операция — композиция, причём в общем случае эта операция не будет коммутативной.

В некоторых случаях удобно задавать операцию с помощью *таблицы Кэли*, которая содержит результаты применения операции к каждой паре элементов множества. Например, на множестве $\{0,1\}$ определим операции $\oplus$ и $\odot$ так:

$\oplus$	0	1
0	0	1
1	1	0

$\odot$	0	1
0	0	0
1	0	1

Тогда операции $\oplus$ и $\odot$ ассоциативны и коммутативны. Кроме того, выполняется свойство *дистрибутивности*: $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$.

Биекции

Функция $f\colon X \to Y$ называется *инъекцией*, если из равенства $f(x) = f(z)$ следует, что $x = z$.

Функция $f\colon X \to Y$ называется *сюръекцией*, если для всякого $y \in Y$ найдётся хотя бы один такой элемент $x \in X$, что $y = f(x)$.

Функция $f\colon X \to Y$ называется *биекцией*, если f одновременно инъекция и сюръекция.

Теорема 2. (i) *Если $f\colon X \to Y$ и $g\colon Y \to Z$ являются биекциями, то композиция $g \circ f$ является биекцией $X \to Z$.*

(ii) *Если функция $f\colon X \to Y$ является биекцией, то определена функция $f^{-1}\colon Y \to X$, сопоставляющая каждому элементу y такой элемент x, что $f(x) = y$. Более того, f^{-1} является биекцией.*

$\triangleright$ Обоснуем пункт (i). Поскольку g и f инъективные функции, равенство $g(f(x)) = g(f(z))$ влечёт равенство $f(x) = f(z)$, откуда, в свою очередь, получаем $x = z$. Следовательно, композиция $g \circ f$ является инъекцией. Поскольку g и f сюръективные функции, для всякого $z \in Z$ найдётся такой элемент $y \in Y$, что $z = g(y)$, а для этого y найдётся такой элемент x, что $y = f(x)$ и $z = g(f(x))$. Следовательно, композиция $g \circ f$ является сюръекцией.

Обоснуем пункт (ii). Поскольку f является сюръекцией, для всякого $y \in Y$ найдётся такой $x \in X$, что $f(x) = y$, а так как f является инъекцией, такой x найдётся ровно один. Таким образом, сопоставление каждому y такого x, что $f(x) = y$, является функцией. Биективность f^{-1} следует из того, что f является функцией. $\triangleleft$

Функцию f^{-1} называют *обратной функцией* к функции f.

Множество G с заданной на нём операцией $\circ$ называется *группой*, если

1) операция $\circ$ ассоциативна;

2) существует такой элемент e (единица группы), что для всякого $g \in G$ выполнено равенство $g \circ e = e \circ g = g$;

3) для всякого $g \in G$ существует такой (обратный) элемент g^{-1}, что $g \circ g^{-1} = g^{-1} \circ g = e$.

Если операция коммутативна, то группа называется *коммутативной* (или *абелевой*).

Подгруппой группы G с операцией $\circ$ называется подмножество $H \subset G$, которое само является группой с операцией $\circ$. Группа G с операцией $\circ$ и группа F с операцией $*$ *изоморфны*, если существует такая биекция $f\colon G \to F$, что $f(g \circ h) = f(g) * f(h)$.

Множество $G(X)$ биекций $X \to X$ с операцией композиции является группой, в которой роль единицы играет является биекция $e(x) = x$ (тождественное отображение).

Теорема 3. *Всякая группа X с операцией $*$ изоморфна некоторой подгруппе группы биекций $G(X)$.*

$\triangleright$ Изоморфизмом является сопоставление всякому элементу $z \in X$ биекции $f_z(x) = z * x$. $\triangleleft$

Для всякого множества X обозначим множество всех его подмножеств через 2^X. Покажем, что существует биекция между 2^X и $\{0,1\}^X$. Действительно, искомая биекция сопоставляет всякому подмножеству $A \subset X$ функцию

$$I_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

которую называют *индикатором* множества A.

Операции с множествами

Пересечением множеств A и B называется множество

$$A \cap B = \{x \colon x \in A \text{ и } x \in B\}.$$

Объединением множеств A и B называется множество

$$A \cup B = \{x \colon x \in A \text{ или } x \in B\}.$$

Разностью множеств A и B называется множество

$$A \setminus B = \{x \colon x \in A \text{ и } x \notin B\}.$$

Всегда можно предполагать, что рассматриваемые множества являются подмножествами некоторого фиксированного множества X. Таким образом, на множестве 2^X всех подмножеств множества X заданы операции $\cap$, $\cup$ и $\setminus$. Более того, эти операции связаны с умножением и сложением индикаторов:

$$I_{A\cap B} = I_A \odot I_B,\ I_{A\cup B} = I_A \oplus I_B \oplus (I_A \odot I_B),\ I_{A\setminus B} = I_A \odot (1 \oplus I_B)$$

(операции $\oplus$ и $\odot$ были определены выше).

Теорема 4. *Для любых множеств справедливы следующие равенства.*

(i) $A \cap B = B \cap A$, $A \cap (B \cap C) = (A \cap B) \cap C$.

(ii) $A \cup B = B \cup A$, $A \cup (B \cup C) = (A \cup B) \cup C$.

(iii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(iv) (Формулы Моргана) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus B)$,
$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus B)$.

$\triangleright$ Доказательство всех пунктов проводятся по одной схеме: покажем, что индикаторы множеств, стоящих справа и слева в рассматриваемых равенствах, равны. Ограничимся обоснованием двух равенств. Докажем первое равенство из (iii):

$$\begin{aligned} I_{A\cap(B\cup C)} &= I_A \odot I_{B\cup C} = I_A \odot (I_B \oplus I_C \oplus (I_B \odot I_C)) = \\ &= (I_A \odot I_B) \oplus (I_A \odot I_C) \oplus \big((I_A \odot I_B) \odot (I_A \odot I_C)\big) = \\ &= I_{A\cap B} \oplus I_{A\cap C} \oplus (I_{A\cap B} \odot I_{A\cap C}) = I_{(A\cap B)\cup(A\cap C)}. \end{aligned}$$

Докажем первое равенство из (iv):

$$\begin{aligned} I_{A\setminus(B\cap C)} &= I_A \odot (1 \oplus I_{B\cup C}) = I_A \odot (1 \oplus I_B) \odot (1 \oplus I_C) = \big(I_A \odot (1 \oplus I_B)\big) \odot \big(I_A \odot (1 \oplus I_C)\big) = \\ &= I_{A\setminus B} \odot I_{A\setminus C} = I_{(A\setminus B)\cup(A\setminus B)}. \quad \triangleleft \end{aligned}$$

Пусть дан произвольный набор множеств (конечный или бесконечный). *Объединением* данного набора множеств называется множество, состоящее из тех и только тех элементов, которые принадлежат хотя бы одному из множеств, входящих в набор. *Пересечением* данного набора множеств называется множество, состоящее из тех и только тех элементов, которые принадлежат всем множествам, входящим в набор.

Рассмотрим несколько иной по стилю (но не по содержанию) способ доказательства равенств для множеств. Докажем, что верны обобщающие формулы Моргана равенства:

$$A \setminus \bigcup_{s\in S} B_s = \bigcap_{s\in S} (A \setminus B_s), \quad A \setminus \bigcap_{s\in S} B_s = \bigcup_{s\in S} (A \setminus B_s).$$

Действительно,

$$x \in A \setminus \bigcup_{s\in S} B_s \Leftrightarrow x \in A \text{ и } x \notin \bigcup_{s\in S} B_s.$$

Заметим, что $x \notin \bigcup\limits_{s\in S} B_s \Leftrightarrow x \notin B_s$ для всякого $s \in S$.

Поэтому $x \in A \setminus \bigcup\limits_{s\in S} B_s \Leftrightarrow x \in A$ и $x \notin B_s$ для всякого $s \in S$,
а это означает, что $x \in \bigcap\limits_{s\in S} (A \setminus B_s)$.

В этом рассуждении мы фактически подменяем операции с множествами правилами логики, в то время как в доказательстве теоремы мы подменили операции с множествами арифметическими операциями с индикаторами.

Отношения

Всякое подмножество R декартова произведения $X \times Y$ называется *отношением.* График функции $X \to Y$ называется *функциональным отношением.*

Непустое подмножество $R \subset X \times X$ называют *отношением эквивалентности*, если для всяких $x, y, z \in X$ выполняются условия:

1) $(x, x) \in R$ (*рефлексивность*);

2) если $(x, y) \in R$, то $(y, x) \in R$ (*симметричность*);

3) если $(x, y) \in R$ и $(y, z) \in R$, то $(x, z) \in R$ (*транзитивность*).

Далее вместо $(x, y) \in R$ пишем $x \sim y$ и говорим, что x эквивалентен y.

Множество $R(a) = \{x \in X \colon x \sim a\}$ называется *классом эквивалентности* с представителем a. В силу свойства 3) все элементы $R(a)$ эквивалентны друг другу.

Теорема 5. *Если классы эквивалентности пересекаются по некоторому элементу, то эти классы совпадают.*

$\triangleright$ Предположим, что $c \in R(a) \cap R(b)$. Тогда $x \sim c$ и $c \sim b$ для всякого $x \in R(a)$. Значит, $x \sim b$ и $x \in R(b)$, т. е. $R(a) \subset R(b)$. Аналогично доказывается, что $R(b) \subset R(a)$. $\triangleleft$

Таким образом, множество X является объединением попарно непересекающихся классов эквивалентности. Множество всех классов эквивалентности называют *фактор-множеством* по данному отношению эквивалентности и обозначают через $X/\sim$.

Говорят, что задано *разбиение* $\{X_s\}_{x \in S}$ множества X, если задан такой набор непустых попарно непересекающихся множеств X_s, что

$$X = \bigcup_{s \in S} X_s.$$

Справедливо следующее утверждение.

Теорема 6. *Пусть задано некоторое разбиение множества X. Отношение R, состоящее из пар (x, y), в которых элементы x и y принадлежат одному и тому же множеству разбиения, является отношением эквивалентности.*

Итак, всякое отношение эквивалентности задаёт разбиение множества, а всякое разбиение множества задаёт отношение эквивалентности. Обычно отношение эквивалентности используют для отождествления каких-то элементов множества. Например, чтобы из отрезка получить окружность, можно отождествить его концы, т. е. задать отношение эквивалентности: всякая отличная от конца отрезка точка эквивалентна только себе самой, а концы отрезка ещё и эквивалентны друг другу. В этом случае окружность — фактор-множество по данному отношению эквивалентности.

Непустое подмножество $R \subset X \times X$ называют *отношением частичного порядка*, если для всяких $x, y, z \in X$ выполняются условия:

1) $(x, x) \in R$;

2) если $(x, y) \in R$ и $(y, x) \in R$, то $x = y$;

3) если $(x, y) \in R$ и $(y, z) \in R$, то $(x, z) \in R$.

Далее вместо $(x, y) \in R$ пишем $x \leqslant y$.

В качестве примера зададим на множестве 2^X отношение сравнения: $A \leqslant B$ тогда и только тогда, когда $A \subset B$. Заметим, что не все подмножества сравнимы друг с другом.

Если отношение частичного порядка R таково, что $(x, y) \in R$ или $(y, x) \in R$ для всех элементов x и y множества X, то отношение R называют *отношением линейного порядка* на множестве X.

Множество натуральных чисел. Индукция

Основываясь на системе аксиом теории множеств, можно построить знакомое со школы множество натуральных чисел

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

Мы не будем обсуждать здесь такое построение, а введём его аксиоматически.

Аксиомы Пеано

Множество $\mathbb{N}$ называется *множеством натуральных чисел*, если оно удовлетворяет следующим аксиомам:

1) для каждого элемента $n \in \mathbb{N}$ существует единственный элемент, называемый следующим и обозначаемый через $n + 1$;

2) существует единственный элемент 1, называемый единицей, который не является следующим ни для какого элемента этого множества;

3) всякий отличный от единицы элемент является следующим ровно для одного элемента множества $\mathbb{N}$;

4) (*аксиома индукции*) если подмножество $M \subset \mathbb{N}$ содержит 1 и вместе со всяким своим элементом n содержит следующий элемент $n + 1$, то $M = \mathbb{N}$.

Четвёртая аксиома фактически утверждает, что в определённом смысле множество натуральных чисел — наименьшее из множеств, удовлетворяющих аксиомам 1), 2), 3).

Функция $f\colon \mathbb{N} \to A$ называется *последовательностью элементов* множества A. Вместо $f(n)$ далее пишем a_n, A_n и т. п.

Принцип математической индукции

Пусть имеется произвольная последовательность утверждений $A_1, A_2, \ldots, A_n, \ldots$ Если утверждение A_1 истинно и для всякого n верно следствие $A_n \Rightarrow A_{n+1}$, то все утверждения A_n истинны.

Поясним равносильность аксиомы индукции и принципа математической индукции. Обозначим через M множество всех таких номеров n, что утверждение A_n истинно. Тогда запись $1 \in M$ равносильна тому, что A_1 — истинно, а следствие

$$n \in M \Rightarrow (n + 1) \in M$$

равносильно тому, что $A_n \Rightarrow A_{n+1}$. Истинность всех утверждений A_n равносильна равенству $M = \mathbb{N}$. Таким образом, принцип математической индукции следует из аксиомы индукции. Пусть теперь выполнен принцип математической индукции. Для всякого множества $M \subset \mathbb{N}$ определим последовательность утверждений A_n: «элемент n принадлежит множеству M». Как и выше, запись $1 \in M$ равносильна истинности утверждения A_1, следствие $n \in M \Rightarrow (n + 1) \in M$ равносильно следствию $A_n \Rightarrow A_{n+1}$, а истинность всех утверждений A_n равносильна равенству $M = \mathbb{N}$. Таким образом, аксиома индукции следует из принципа математической индукции.

Рассмотрим несколько примеров применения принципа математической индукции — доказательство утверждений методом математической индукции.

Теорема 7 (неравенство Бернулли). *Для всех $x \geqslant -1$ и $n \in \mathbb{N}$ выполнено неравенство*

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

$\triangleright$ При $n = 1$ неравенство верно: $(1 + x)^1 \geqslant 1 + 1 \cdot x$ (оно обращается в равенство). Пусть для некоторого $n \in \mathbb{N}$ неравенство уже доказано; докажем его для $n + 1$. Поскольку $x \geqslant -1$ и $x^2 \geqslant 0$, получаем $(1 + x)^{n+1} = (1 + x)(1 + x)^n \geqslant (1 + x)(1 + nx) = 1 + (n + 1)x + nx^2 \geqslant 1 + (n + 1)x$. $\triangleleft$

Теорема 8 (бином Ньютона). *Пусть $n \in \mathbb{N}$. Для любых a, b выполняется равенство*

$$(a + b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}, \quad C_n^k = \frac{n!}{k!(n - k)!}$$

(*считаем, что $a^0 = 1$ для всех a*).

$\triangleright$ Доказываем по индукции. При $n = 1$ утверждение верно:

$$(a + b)^1 = 1 \cdot a + 1 \cdot b = C_1^0 a^0 b^1 + C_1^1 a^1 b^0.$$

Предположим, что для некоторого $n \in \mathbb{N}$ неравенство уже доказано и докажем его для $n+1$. Заметим, что по предположению индукции выполнено равенство

$$(a+b)^{n+1} = (a+b)\Big(\sum_{m=0}^{n} C_n^m a^m b^{n-m}\Big).$$

Выпишем коэффициент, который будет стоять перед слагаемым $a^m b^{n-m+1}$ после раскрытия скобок. Такое слагаемое получается только при умножении $a^{m-1}b^{n-(m-1)}$ на a или при умножении $a^m b^{n-m}$ на b. Следовательно, искомый коэффициент равен сумме соответствующих коэффициентов $C_n^{m-1} + C_n^m$, которая равна C_{n+1}^m. $\lhd$

Положим в формуле бинома $a = 1$, $b = x$ и $n = \alpha$:

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

Заметим, что для натуральных показателей α эту сумму можно считать бесконечной; в самом деле, при $k \geqslant \alpha + 1$ коэффициенты при x^k обнуляются. Оказывается, что при $|x| < 1$ такая формула с бесконечной суммой верна не только для натуральных показателей, но и для целых, рациональных и вообще любых вещественных показателей. Например, при $\alpha = -1$ получаем сумму бесконечной геометрической прогрессии

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \ldots + (-1)^n x^n + \ldots$$

В общем случае доказательство формулы бинома с произвольным показателем будет приведено в конце нашего курса.

С помощью аксиомы индукции можно давать индуктивные определения. Рассмотрим, как это делается на примере операции сложения на множестве натуральных чисел.

В силу определения натуральных чисел мы умеем добавлять 1 ко всякому числу n. Предположим, что мы уже знаем, что такое $n+m$ для некоторого m. Тогда полагаем по определению

$$n + (m+1) = (n+m) + 1.$$

Таким образом, мы умеем добавлять 1, и если умеем добавлять m, то умеем добавлять $(m+1)$. По аксиоме индукции мы умеем добавлять произвольное m. Следующее утверждение формализует эти нестрогие рассуждения.

Предложение 1. *Для всякого $n \in \mathbb{N}$ существует единственная функция $f_n\colon \mathbb{N} \to \mathbb{N}$, для которой выполнены условия $f_n(1) = n+1$ и $f_n(m+1) = f_n(m)+1$.*

$\rhd$ Обоснуем единственность. Пусть две функции f_n и g_n удовлетворяют условиям утверждения. Тогда $f_n(1) = g_n(1)$ и если $f_n(m) = g_n(m)$, то

$$f_n(m+1) = f_n(m) + 1 = g_n(m) + 1 = g_n(m+1).$$

По аксиоме индукции множество тех m, для которых выполнено равенство $f_n(m) = g_n(m)$, совпадает с $\mathbb{N}$.

Теперь докажем существование функции f_n. Будем рассуждать по индукции. Положим $f_1(m) = m+1$. Проверим свойства: $f_1(1) = 1+1$, $f_1(m+1) = (m+1)+1 = f_1(m)+1$. Предположим, что для некоторого n требуемая функция f_n существует. Положим

$$f_{n+1}(m) = f_n(m) + 1.$$

Проверим свойства: $f_{n+1}(1) = f_n(1) + 1 = (n+1) + 1$,

$$f_{n+1}(m+1) = f_n(m+1) + 1 = (f_n(m)+1) + 1 = f_{n+1}(m) + 1.$$

По аксиоме индукции множество тех n, для которых существует функция f_n, совпадает с $\mathbb{N}$. $\lhd$

Теперь мы можем положить по определению: $n + m = f_n(m)$.

Аналогичным образом можно определить умножение натуральных чисел и опять же по индукции показать, что выполнены известные со школы свойства операций сложения и умножения.

На множестве натуральных чисел можно задать линейный порядок:

$$n \leqslant m \ \Leftrightarrow\ n = m \ \text{ или } \ m = n + k.$$

Используя свойства сложения и аксиому индукции, можно показать, что это действительно отношение линейного порядка. Кроме того, на множестве натуральных чисел существует единственное отношение линейного порядка, при котором $n \leqslant n+1$.

Иногда удобнее использовать *полную индукцию.*

Предложение 2. *Аксиома индукции равносильна следующему утверждению (аксиоме полной индукции): если подмножество $M \subset \mathbb{N}$ содержит 1 и для всех $n \in \mathbb{N}$ верно следствие*

$$\{k \in \mathbb{N} \colon k \leqslant n\} \subset M \Rightarrow (n+1) \in M,$$

то $M = \mathbb{N}$.

$\triangleright$ Пусть верна аксиома индукции и множество M удовлетворяет условию аксиомы полной индукции. Покажем, что $M = \mathbb{N}$. Рассмотрим подмножество M' множества M, состоящее из тех и только тех его элементов n, для которых множество M содержит все натуральные числа k, меньшие или равные n. Заметим, что $1 \in M'$. Более того, если $n \in M'$, то по определению M' все натуральные числа k, меньшие или равные n, принадлежат M. Множество M удовлетворяет условию аксиомы полной индукции, поэтому $(n+1) \in M$. Следовательно, все числа k, меньшие или равные $n+1$, принадлежат M, и по определению M' число $n+1$ принадлежит M'. Таким образом, для всех $n \in \mathbb{N}$ имеет место следствие $n \in M' \Rightarrow (n+1) \in M'$. Согласно аксиоме индукции получаем $M' = \mathbb{N}$, поэтому и $M = \mathbb{N}$.

Пусть теперь верна аксиома полной индукции и множество M удовлетворяет условию аксиомы индукции. Тогда $1 \in M$ и для всех $n \in \mathbb{N}$ получаем

$$\{k \in \mathbb{N} \colon k \leqslant n\} \subset M \Rightarrow n \in M \Rightarrow (n+1) \in M,$$

поэтому в силу аксиомы полной индукции имеем $M = \mathbb{N}$. $\triangleleft$

Принцип полной математической индукции

Пусть дана последовательность утверждений

$$A_1, A_2, \ldots, A_n, \ldots$$

Предположим, что A_1 истинно и для всякого n из истинности всех утверждений A_k с номерами $k \leqslant n$ следует истинность утверждения A_{n+1}. Тогда все утверждения $A_1, A_2, \ldots, A_n, \ldots$ истинны.

Полностью аналогично тому, как это сделано выше, можно показать равносильность аксиомы полной индукции и принципа полной математической индукции. Таким образом, в силу предложения 2 оба рассмотренных принципа математической индукции равносильны между собой.

Типичным примером применения принципа полной математической индукции является доказательство основной теоремы арифметики: *всякое натуральное число, большее единицы, либо является простым, либо раскладывается в произведение простых чисел, причём единственным образом с точностью до порядка множителей.* Докажем лишь существование разложения. Для $n = 2$ утверждение верно. Предположим, что утверждение верно для всех $k \leqslant n$, и докажем его для $n+1$. Если $n+1$ является простым числом, то утверждение доказано. Если $n+1$ не является простым числом, то найдутся такие натуральные числа n_1 и n_2, что $n_1 > 1$, $n_2 > 1$ и $n+1 = n_1 \cdot n_2$. Поскольку $n_1 \leqslant n$ и $n_2 \leqslant n$, к ним применимо предположение индукции. Следовательно, n_1 и n_2 либо являются простыми, либо раскладываются в произведение простых множителей, а значит, раскладывается в произведение простых множителей и число $n+1$.

Как показывает следующее утверждение, аксиома индукции тесно связана с отношением порядка.

Теорема 9. *Аксиома индукции равносильна тому, что в каждом непустом подмножестве множества натуральных чисел есть минимальный элемент, т. е. элемент этого подмножества, меньший всякого другого элемента этого подмножества.*

$\triangleright$ Предположим, что во всяком непустом подмножестве множества $\mathbb{N}$ есть минимальный элемент. Пусть для некоторого подмножества M множества $\mathbb{N}$ выполнено, что если $1 \in M$ и $n \in M$, то $(n+1) \in M$. Предположим, что $\mathbb{N} \setminus M \neq \varnothing$. Существует минимальный элемент $n \in \mathbb{N} \setminus M$, причём $n \neq 1$. Пусть $m+1 = n$. Тогда $m \in M$ и, следовательно, $m+1 \in M$, но $m+1 = n \notin M$. Противоречие. Следовательно, $M = \mathbb{N}$.

Предположим теперь, что выполнена аксиома индукции. Нам будет удобнее использовать полную индукцию. Пусть A — непустое подмножество множества $\mathbb{N}$. Предположим, что в A нет минимального элемента. Тогда $1 \in M = \mathbb{N} \setminus A$ и из того, что $k \notin M$ при $k \leqslant n$ следует, что

$n+1 \in M$, иначе $n+1$ было бы минимальным элементом в A. По аксиоме индукции $M = \mathbb{N}$, но это противоречит непустоте множества A. Следовательно, в A есть минимальный элемент. $\lhd$

Множество целых чисел $\mathbb{Z}$ есть объединение множеств $\mathbb{N}$, $\{0\}$ и $\{-n\colon n \in \mathbb{N}\}$. Операции сложения, умножения и отношение порядка переносятся с $\mathbb{N}$ на $\mathbb{Z}$.

На декартовом произведении $\mathbb{Z} \times \mathbb{N}$ введём следующее отношение эквивалентности:

$$(m, n) \sim (p, q) \iff mq = np.$$

Множество классов эквивалентности обозначается через $\mathbb{Q}$ и называется *множеством рациональных чисел.* Вместо (m, n) далее пишем $\frac{m}{n}$.

Операции и отношение порядка на множестве $\mathbb{Q}$ определяются следующим образом:

$$\frac{m}{n} + \frac{p}{q} = \frac{mq + np}{nq}, \quad \frac{m}{n} \cdot \frac{p}{q} = \frac{mp}{nq}, \quad \frac{m}{n} \leqslant \frac{p}{q} \iff mq \leqslant np.$$

Можно показать, что определения сложения, умножения и отношения порядка корректны, т. е. не зависят от выбора представителей классов эквивалентности.

Конечные и бесконечные множества. Счётные множества

Будем говорить, что два множества *равномощны*, если существует биекция из множества A в множество B (пишем $A \sim B$).

Можно показать, что:

(1) $A \sim A$;

(2) если $A \sim B$, то $B \sim A$;

(3) если $A \sim B$ и $B \sim C$, то $A \sim C$.

Лемма 1. *Пусть $a, b \in A$, $a \neq b$. Тогда множества $A \setminus \{a\}$ и $A \setminus \{b\}$ равномощны.*

$\rhd$ Пусть $D = A \setminus \{a, b\}$. Тогда получаем $A \setminus \{a\} = D \cup \{b\}$ и $A \setminus \{b\} = D \cup \{a\}$. Определим функцию f из множества $A \setminus \{a\}$ в множество $A \setminus \{b\}$ следующим образом: $f(x) = x$ при $x \in D$ и $f(b) = a$. Функция f — искомая биекция. $\lhd$

Положим $A_n = \{k \in \mathbb{N}\colon k \leqslant n\}$. Это множество, состоящее из первых n натуральных чисел.

Теорема 10. *Если множества A_n и A_m равномощны, то верно равенство $n = m$.*

$\rhd$ Воспользуемся индукцией по n. Пусть $n = 1$ и $f\colon A_1 \to A_m$ — биекция, где $A_1 = \{1\}$. Если $m > 1$, то $f(1) \neq 1$ или $f(1) \neq m$, что противоречит сюръективности f. Следовательно, $m = 1$.

Предположим, что утверждение уже доказано для n и докажем его для $n+1$. Пусть $f\colon A_{n+1} \to A_m$ является биекцией. Тогда эта же самая функция устанавливает биекцию между множествами A_n и $A_m \setminus \{f(n+1)\}$. Остаётся заметить, что множества $A_m \setminus \{f(n+1)\}$ и A_{m-1} равномощны, а затем применить предположение индукции. $\lhd$

Множество A называется *конечным*, если оно или пусто, или равномощно множеству A_n. Говорят, что множество A *состоит из n элементов*, если A равномощно множеству A_n. Согласно доказанному выше утверждению, это определение корректно, т. е. не может оказаться так, что одно и то же множество одновременно состоит из n и m элементов для различных n и m.

Множество называется *бесконечным*, если оно не является конечным.

Покажем, что множество натуральных чисел $\mathbb{N}$ бесконечно. Для любой функции $f\colon A_n \to \mathbb{N}$ натуральное число

$$M = f(1) + f(2) + \ldots + f(n)$$

не лежит в области её значений (так как $f(k) < M$ для всякого $k \leqslant n$), а следовательно, такая функция не может быть сюръекцией и уж тем более биекцией.

Еще один важный пример бесконечного множества — множество всех простых чисел. Предположим, что это множество конечно и $p_1, p_2, \ldots, p_N$ — все простые числа. Тогда, следуя Евклиду, рассмотрим число $p_1 \cdot p_2 \cdots p_N + 1$. Оно не является простым, но и не делится ни на какое простое число. Противоречие. Следовательно, множество всех простых чисел бесконечно.

Счётные множества

Множество A называется *счётным*, если A равномощно множеству $\mathbb{N}$.

Теорема 11. (i) *Всякое подмножество множества натуральных чисел конечно или счётно.*

(ii) *Множество* $\mathbb{N} \times \mathbb{N}$ *является счётным.*

$\triangleright$ Рассмотрим пункт (i). Пусть $B \subset \mathbb{N}$. Можно сразу предполагать, что B — бесконечное множество. Пусть n_1 — минимальный элемент множества B. Тогда $n_1 \geqslant 1$. Пусть n_2 — минимальный элемент множества $B \setminus \{n_1\}$. По построению выполнено неравенство $n_1 < n_2$ и, следовательно, $n_2 \geqslant 2$. Предположим, что уже построены числа $n_1 \geqslant 1$, $n_2 \geqslant 2, \ldots, n_k \geqslant k$ и пусть n_{k+1} — наименьший элемент множества $B \setminus \{n_1, n_2, \ldots, n_k\}$. По построению имеем $n_{k+1} > n_k$ и, следовательно, $n_{k+1} \geqslant k+1$. Продолжая построение, получаем последовательность номеров n_k. Если в множестве B есть элемент b, который не встречается среди n_k, то $n_k < b$ для всех k, что невозможно, так как $n_{b+1} \geqslant b+1 > b$. Следовательно, $B = \{n_k\}$, причём функция $k \to n_k$ устанавливает биекцию между $\mathbb{N}$ и B.

В пункте (ii) искомая биекция задается формулой

$$(k, l) \to \frac{(k+l-1)(k+l)}{2} - l + 1.$$

Мы пересчитываем элементы таблицы $\mathbb{N} \times \mathbb{N}$ по диагоналям, сопоставляя $(1,1) \mapsto 1$, $(1,2) \mapsto 2$, $(2,1) \mapsto 3$ и т. д. $\triangleleft$

Следствие 1. *Всякое подмножество счётного множества либо конечно, либо счётно.*

Конечные или счётные множества называют *не более чем счётными.*

Предложение 3. (i) *Если* $f\colon B \to A$ *— инъекция и* A *— счётное множество, то* B *не более чем счётно.*

(ii) *Если* $f\colon B \to A$ *— сюръекция и* B *— не более чем счётное множество, то* A *не более чем счётно.*

$\triangleright$ Обоснуем пункт (i). Поскольку $f\colon B \to f(B)$ — биекция и $f(B)$ не более чем счётно, заключаем, что и B не более чем счётно.

Рассмотрим пункт (ii). Для каждого $a \in A$ выберем один такой элемент b, что $f(b) = a$. Полученная функция из A в B является инъекцией, так как если $b_1 = b_2$, то $f(b_1) = f(b_2)$. Согласно пункту (i) заключаем, что A не более чем счётно. $\triangleleft$

Отметим, что если A не более чем счётно, то всегда существует сюръекция из $\mathbb{N}$ в A.

Следствие 2. (i) *Объединение конечного или счётного набора конечных или счётных множеств является конечным или счётным множеством.*

(ii) *Декартово произведение конечного набора конечных или счётных множеств является конечным или счётным множеством.*

$\triangleright$ Пункт (ii) сводится к пункту (i). Докажем утверждение пункта (i). Пусть задан набор не более чем счётных множеств A_s, причём индекс s принадлежит не более чем счётному множеству S. Существуют сюръекции $F_s\colon \mathbb{N} \to A_s$ и $q\colon \mathbb{N} \to S$. Тогда отображение $(k, m) \to F_{q(k)}(m)$ является сюръекцией $\mathbb{N} \times \mathbb{N}$ на $\bigcup\limits_{s \in S} A_s$. Следовательно, множество $\bigcup\limits_{s \in S} A_s$ не более чем счётно. $\triangleleft$

Приведённое доказательство является формализацией следующего эвристического рассуждения: запишем элементы объединения $\bigcup\limits_{s \in S} A_s$ в таблицу, в которой в s-й строке выписаны элементы множества A_s, и пересчитаем эту таблицу так же, как это было сделано для декартова произведения $\mathbb{N} \times \mathbb{N}$.

Пример Кантора несчётного множества

Покажем, что множество $\{0, 1\}^\infty$ всех бесконечных последовательностей из нулей и единиц не является счётным. Предположим противное: пусть множество этих последовательностей можно занумеровать натуральными числами. Построим новую последовательность, в которой на k-м месте стоит 1, если у k-й последовательности на k-м месте стоит 0, и стоит 0, если у k-й последовательности на k-м месте стоит 1. Новая последовательность заведомо не была занумерована, так как для всякого k она отличается от k-й последовательности элементом на k-м месте.

Обобщением примера Кантора является следующее утверждение.

Теорема 12. *Для всякого множества A не существует биекции между множеством A и множеством всех его подмножеств 2^A.*

$\triangleright$ Предположим, что биекция $f : A \to 2^A$ существует. Рассмотрим множество M, состоящее из всех тех элементов $a \in A$, для которых $a \notin f(a)$. Поскольку M — подмножество A, имеем $M = f(b)$ для некоторого $b \in A$. Если $b \in M$, то $b \notin f(b) = M$. Если же $b \notin M = f(b)$, то $b \in M$ по определению множества M. В обоих случаях пришли к противоречию. Значит, наше предположение о существовании биекции неверно. $\triangleleft$

Отметим, что доказательство почти дословно повторяет рассуждение из парадокса Рассела.

Будем говорить, что множество A *по мощности не больше* множества B, а множество B *по мощности не меньше* множества A, если существует инъекция из A в B. Если дополнительно известно, что не существует биекции из A в B, то говорим, что A *по мощности строго меньше* множества B, а множество B *по мощности строго больше* множества A.

Существует естественное инъективное отображение множества A в множество 2^A, сопоставляющее каждому элементу $a \in A$ множество $\{a\}$. Таким образом, всякое множество A по мощности строго меньше множества 2^A. Отметим также, что выше была установлена равномощность множеств 2^X и $\{0,1\}^X$. Таким образом, множество всех функций на X, принимающих два значения, по мощности строго больше множества X.

Теорема Кантора — Бернштейна

Теорема 13. *Если множество A равномощно подмножеству множества B и множество B равномощно подмножеству множества A, то эти множества равномощны.*

$\triangleright$ Заменяя B на равномощное ему подмножество множества A, можно считать, что само множество B является подмножеством множества A. Тогда условие теоремы переформулируется следующим образом: если даны три множества $A_2 \subset A_1 \subset A_0$ и множества A_2 и A_0 равномощны, то множества A_1 и A_0 также равномощны. Пусть функция f устанавливает биекцию между множествами A_0 и A_2. Для каждого натурального n определим множества $A_{2n+2} = f(A_{2n})$ и $A_{2n+1} = f(A_{2n-1})$. Заметим, что для всякого n множество $C_{2n} = A_{2n} \setminus A_{2n+1}$ равномощно множеству $C_{2n+2} = A_{2n+2} \setminus A_{2n+3}$. Положим $C_{2n+1} = A_{2n+1} \setminus A_{2n+2}$. Справедливы равенства

$$A_0 = C_0 \cup C_1 \cup C_2 \ldots \cup D \quad \text{и} \quad A_1 = C_1 \cup C_2 \cup C_3 \cup \ldots \cup D,$$

где $D = \bigcap\limits_{n \in \mathbb{N}} A_n$ и все множества C_n и D попарно не пересекаются. Биекцию из A_0 в A_1 установим следующим образом: элементы множеств C_{2n+1} и D оставим на месте (тождественное отображение), а элементам каждого из множеств C_{2n} взаимно однозначно сопоставим элементы множества C_{2n+2}. $\triangleleft$

Теорема Кантора — Бернштейна утверждает, что если множество A по мощности не больше множества B и множество B по мощности не больше множества A, то эти множества равномощны. Если X — некоторое множество, то на множестве 2^X определено отношение эквивалентности: $A \sim B$ тогда и только тогда, когда A равномощно B. На множестве классов эквивалентности определено отношение частичного порядка: $A \leqslant B$ тогда и только тогда, когда A по мощности не больше B. Замечательная теорема Цермело утверждает, что на самом деле этот порядок линеен. Доказательство опирается на так называемую аксиому выбора.

Вещественные числа и принципы полноты

Полем называется множество с операциями умножения и сложения, которое является абелевой группой по сложению (от перестановки мест слагаемых сумма не меняется), после отбрасывания нуля становится абелевой группой по умножению, и на котором выполняется условие дистрибутивности:

$$a(b+c) = (b+c)a = ab + ac = ba + ca.$$

Важными примерами являются поле рациональных чисел $\mathbb{Q}$ и поле $\mathbb{Z}_p$ вычетов по простому модулю p. Поле $\mathbb{Z}_p$ — множество классов эквивалентности на $\mathbb{Z}$ относительно следующего отношения эквивалентности: $m \sim n$ тогда и только тогда, когда $m - n$ делится на p. Класс эквивалентности с представителем m обозначается через $[m]_p$. Операции сложения и умножения задаются следующим образом:

$$[m]_p + [n]_p = [m+n]_p, \quad [n]_p \cdot [m]_p = [n \cdot m]_p.$$

Существенное отличие поля $\mathbb{Q}$ от поля $\mathbb{Z}_p$ состоит в том, что поле рациональных чисел имеет характеристику нуль, т. е. сумма $1 + 1 + \ldots + 1$ никогда не обращается в нуль, а поле $\mathbb{Z}_p$ имеет характеристику p, так как при сложении p единиц получается нуль, причём при сложении меньшего количества единиц нуль не получится.

Поле называется *упорядоченным*, если на нём задано отношение линейного порядка, удовлетворяющее следующим условиям:

1) $a \leqslant b \Leftrightarrow a + c \leqslant b + c$ для всякого c;

2) если $c \geqslant 0$, то $a \leqslant b \Rightarrow ac \leqslant bc$.

Покажем, что упорядоченное поле всегда имеет характеристику нуль. Сначала докажем, что $0 < 1$. Предположим, что $0 > 1$. Тогда, вычитая из правой и левой части 1, получаем, что $-1 > 0$. Следовательно, неравенство $-1 > 0$ можно умножить на -1. Имеем $1 > 0$, что противоречит нашему предположению $0 > 1$. Остаётся заметить, что поскольку $0 < 1$, получаем $0 < 1 < 1 + 1$ и т. д., т. е. $0 < 1 + 1 + \ldots + 1$ для любого числа единиц.

В поле характеристики нуль подмножество из элементов, представляющихся в виде конечной суммы единиц, отождествляется с множеством натуральных чисел, причём элементу, являющемуся суммой n единиц, соответствует натуральное число n, которое и используется для его обозначения. Элементы вида $m \cdot n^{-1}$, где n^{-1} — обратный элемент к n, отождествляются с рациональными числами.

Будем говорить, что множество A лежит *левее* множества B, а B *правее* множества A, если $a \leqslant b$ для всех $a \in A$ и $b \in B$. Говорят, что элемент c *разделяет* множества A и B, если A левее множества $\{c\}$, а B правее множества $\{c\}$. Таким образом, элемент c разделяет множества A и B, если $a \leqslant c \leqslant b$ для всяких $a \in A$ и $b \in B$.

Аксиоматическое определение множества $\mathbb{R}$

Множество вещественных (или действительных) чисел $\mathbb{R}$ — упорядоченное поле, для элементов которого выполняется *аксиома полноты*:

всякие два непустых подмножества $A, B \subset \mathbb{R}$, одно из которых лежит левее другого, можно разделить некоторым элементом $c \in \mathbb{R}$.

В качестве простейшего примера применения аксиомы полноты докажем разрешимость уравнения $x^2 = 2$.

Обозначим через A множество всех положительных чисел a, удовлетворяющих неравенству $a^2 < 2$, и обозначим через B множество всех положительных чисел b, удовлетворяющих неравенству $b^2 > 2$. Заметим, что $1 \in A$ и $2 \in B$. Кроме того, множество A лежит левее множества B. Действительно, неравенства $0 < b^2 - a^2 = (b - a)(b + a)$ и $b + a > 0$ влекут неравенство $a < b$. По аксиоме полноты существует число c, разделяющее множества A и B. Отметим, что $1 < c < 2$. Покажем, что $c^2 = 2$. Предположим, что $c \in A$, т. е. $c^2 < 2$. Пусть $\varepsilon \in (0; 1)$. Заметим, что

$$(c + \varepsilon)^2 = c^2 + \varepsilon(2c + \varepsilon) < c^2 + 5\varepsilon.$$

Положим $\varepsilon = (2 - c^2)/5 > 0$. Имеем $(c + \varepsilon)^2 < 2$ и $(c + \varepsilon) \in A$, что противоречит определению разделяющего числа c. Значит, $c \notin A$. Аналогично доказывается, что $c \notin B$. Таким образом, $c^2 = 2$.

Можно показать, что построенный корень c уравнения $x^2 = 2$ не принадлежит множеству рациональных чисел.

Конкретной реализацией (*моделью*) множества действительных чисел является множество бесконечных десятичных дробей. *Положительной бесконечной десятичной дробью* называется последовательность вида $a_0, a_1 a_2 a_3 \ldots a_k \ldots$, где $a_0 \in \mathbb{N} \cup \{0\}$ и $a_k \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, причём хотя бы один элемент последовательности отличен от нуля; последовательности, в которых с некоторого места стоят одни девятки, запрещены.

На положительных бесконечных десятичных дробях отношение порядка определяется так: $a_0, a_1a_2a_3\ldots \leqslant b_0, b_1b_2b_3\ldots$ тогда и только тогда, когда $a_0 \leqslant b_0$ и для каждого k из равенств $a_0 = b_0, \ldots, a_{k-1} = b_{k-1}$ следует неравенство $a_k \leqslant b_k$.

Нулевой бесконечной десятичной дробью (или просто *нулём*) называется последовательность $0{,}00\ldots$, все элементы которой равны 0. Наконец, *отрицательные бесконечные десятичные дроби* получаются приписыванием знака минус положительным бесконечным десятичным дробям.

Введённое отношение порядка естественным образом продолжается до отношения порядка на множестве всех бесконечных десятичных дробей.

Теорема 14. *На множестве бесконечных десятичных дробей выполняется аксиома полноты.*

$\triangleright$ Пусть заданы непустые подмножества A и B множества десятичных дробей, причём A левее B. Если в A нет положительных элементов, а в B нет отрицательных, то эти множества можно разделить нулём. Рассмотрим случай, когда в A есть хотя бы один положительный элемент. По условию в этом случае B состоит исключительно из положительных дробей. Будем строить разделяющую A и B бесконечную десятичную дробь $c_0, c_1c_2c_3\ldots$. Поскольку B состоит из положительных элементов, существует число b_0, наименьшее из всех натуральных чисел, с которых начинаются дроби из B. Полагаем $c_0 = b_0$. Теперь рассмотрим только те десятичные дроби из B, которые начинаются с b_0, и у них найдём наименьшую первую цифру после запятой: b_1. Полагаем $c_1 = b_1$. Повторяем процедуру с дробями из B, которые начинаются с b_0, b_1 и находим наименьшую вторую цифру после запятой b_2, полагаем $c_2 = b_2$, и т. д. Построенная бесконечная десятичная дробь $c_0, c_1c_2c_3\ldots$, очевидно, не является запрещённой и не больше всякого элемента из B. Покажем, что она не меньше всякого элемента из A. Предположим противное: пусть в A нашлась такая десятичная дробь $a_0, a_1a_2\ldots a_k\ldots$, что $c_0 = a_0, \ldots, c_{k-1} = a_{k-1}$ и $c_k < a_k$. По построению найдётся дробь из B, которая начинается с $c_0, c_1c_2\ldots c_k$. Следовательно, эта дробь также меньше дроби $a_0, a_1a_2\ldots a_k\ldots$, что противоречит условию: A левее B. Случай, когда в B есть хотя бы один отрицательный элемент, разбирается аналогично. $\triangleleft$

Принципы полноты

Аксиома полноты допускает различные эквивалентные формулировки, называемые *принципами полноты.* Мы не будем доказывать равносильность принципов полноты, а выведем их из аксиомы полноты.

Пусть A — непустое подмножество $\mathbb{R}$. Число b называется *верхней гранью* множества A, если $a \leqslant b$ для всех $a \in A$. Если у множества есть хотя бы одна верхняя грань, то говорят, что множество A *ограничено сверху. Точной верхней гранью* множества A (обозначение: $\sup A$) называется наименьшая из верхних граней множества A. Число b называется *нижней гранью* множества A, если $a \geqslant b$ для всех $a \in A$. Если у множества есть хотя бы одна нижняя грань, то говорят, что множество A *ограничено снизу. Точной нижней гранью* множества A (обозначение: $\inf A$) называется наибольшая из нижних граней множества A.

Поскольку не у всякого ограниченного множества есть наименьший (наибольший) элемент, существование точных граней требует обоснования.

Теорема 15 (принцип полноты Вейерштрасса)**.** *Всякое непустое ограниченное сверху* (*снизу*) *множество имеет точную верхнюю* (*точную нижнюю*) *грань.*

$\triangleright$ Пусть A — непустое ограниченное сверху множество, а B — множество верхних граней множества A. Множества A, B непустые и первое из них лежит левее второго. По аксиоме полноты существует разделяющий их элемент. Обозначим его через c. Этот элемент является верхней гранью множества A и не больше всякого элемента множества B. Таким образом, c — точная верхняя грань множества A. $\triangleleft$

Обсудим важное следствие принципа полноты Вейерштрасса.

Теорема 16 (аксиома Архимеда)**.** *Для всякого вещественного числа a существует такое натуральное число n, что $a < n$.*

$\triangleright$ Предположим противное. Тогда множество натуральных чисел $\mathbb{N}$ ограничено сверху. Пусть $A = \sup \mathbb{N}$. Число $A - 1$ не является верхней гранью множества $\mathbb{N}$. Следовательно, найдётся такое $n \in \mathbb{N}$, что $A-1<n$. Прибавляя единицу к правой и левой частям неравенства, получаем $A<n+1$, а это противоречит определению точной верхней грани. $\triangleleft$

Отметим, что часто удобно называть множество вещественных чисел числовой прямой, а его элементы — точками этой прямой. Мы будем использовать эту наглядную геометрическую интерпретацию, но не будем опираться на геометрические аксиомы.

Пусть $a<b$. Множества

$$[a;b]=\{x\colon a\leqslant x\leqslant b\},\quad (a;b)=\{x\colon a<x<b\},\quad [a;b)=\{x\colon a\leqslant x<b\},\quad (a;b]=\{x\colon a<x\leqslant b\}$$

называются отрезком, интервалом и полуинтервалами соответственно. Величина $b-a$ называется длиной отрезка, интервала или полуинтервала. Расстояние между точками x и y числовой прямой обозначается через $|x-y|$ и равно $x-y$ при $x\geqslant y$ и $y-x$ при $x\leqslant y$. Таким образом, это просто длина отрезка с концами в этих точках.

Множества

$$[a;+\infty)=\{x\colon x\geqslant a\},\quad (-\infty;a]=\{x\colon x\leqslant a\},\quad (a;+\infty)=\{x\colon x>a\},\quad (-\infty;a)=\{x\colon x<a\}$$

называются лучами, причём лучи, включающие точку a, называются замкнутыми, а не включающие — открытыми (смысл этих названий станет ясен при изучении топологии вещественной прямой).

Важную роль играют так называемые *неравенства треугольника*:

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

Первое неравенство следует из второго, а второе обосновывается с помощью равенства

$$|x|=\max\{x,-x\}.$$

Пусть $0<a<b$. Тогда найдётся такое натуральное число n, что $an>b$. Пусть n_0 — наименьшее из таких натуральных чисел. Тогда отрезок $[0;a]$ укладывается n_0-1 раз в отрезке $[0;b]$. Таким образом, аксиома Архимеда позволяет измерять отрезки.

Теорема 17 (принцип полноты Кантора о вложенных отрезках)**.** *Всякая последовательность вложенных отрезков $\{[a_n;b_n]\}$ ($[a_n;b_n]\supset[a_{n+1};b_{n+1}]$ для всех $n\in\mathbb{N}$) имеет общую точку, т. е. такую точку, которая принадлежит каждому из отрезков $[a_n;b_n]$. Кроме того, если для всякого $\varepsilon>0$ найдётся отрезок, длина которого меньше ε, то общая точка только одна.*

$\triangleright$ Обозначим через A множество всех левых концов a_n, а через B множество всех правых концов b_n. Множество A лежит левее множества B. Действительно, если бы какой-то правый конец лежал левее какого-то левого конца, то соответствующие им отрезки не пересекались бы, но по условию из любых двух отрезков один является подмножеством другого. По аксиоме полноты существует такая точка c, что $a_n\leqslant c\leqslant b_m$ для всех n,m, в частности, точка c принадлежит каждому отрезку $[a_n;b_n]$. Наконец, заметим, что если есть две различные общие точки всех отрезков, то длина любого отрезка не меньше расстояния между этими точками. $\triangleleft$

Отметим, что на самом деле в доказательстве принципа Кантора мы использовали лишь то, что отрезки попарно пересекаются.

Следствие 3. *Всякий отрезок не является счётным множеством.*

$\triangleright$ Предположим противное. Тогда точки отрезка можно занумеровать натуральными числами. Разделим отрезок на три отрезка. Обязательно найдётся отрезок, в котором нет точки с номером один. Этот отрезок разделим опять на три части и выберем ту часть, в которой нет точки с номером два и т. д. Получим систему вложенных отрезков с таким свойством: n-й отрезок не содержит первые n точек. Следовательно, общая точка этих отрезков не имеет номера. Противоречие. $\triangleleft$

Отметим, что из несчётности отрезка следует несчётность множества действительных чисел. Более того, можно показать, что множества $[a;b]$, $(a;b]$, $[a;b)$, $(a;b)$, $\{0,1\}^\infty$, $\mathbb{R}$ равномощны. Множества, равномощные множеству действительных чисел, называются *континуальными* или имеющими *мощность континуума*.

В 1877 г. основателем современной теории множеств Георгом Кантором было выдвинута континуум-гипотеза — предположение о том, что любое бесконечное подмножество континуального множества является либо счётным, либо континуальным. Другими словами, континуум-гипотеза предполагает, что мощность континуума — наименьшая, превосходящая мощность счётного множества, и множеств «промежуточных» мощностей между счётным множеством и континуальным множеством нет.

Континуум-гипотеза стала первой из двадцати трёх математических проблем, о которых Д. Гильберт доложил на II Международном Конгрессе математиков в Париже в 1900 г., поэтому континуум-гипотеза известна также как первая проблема Гильберта. В середине XX в. установлена независимость этой гипотезы от стандартной аксиоматики теории множеств. Однако при выполнении заданий нашего курса следует пытаться находить решения, не апеллирующие к континуум-гипотезе.

Предел последовательности

Напомним, что задана *последовательность* элементов некоторого множества A, если задана функция $f\colon \mathbb{N} \to A$; вместо $f(n)$ пишем a_n. Для последовательности используют обозначения $\{a_n\}_{n=1}^{\infty}$ или $\{a_n\}$. Отметим, что несмотря на внешнюю схожесть, запись $\{a_n\}$ означает не множество, состоящее из элемента (или элементов) a_n, а именно сокращённую запись обозначения последовательности $\{a_n\}_{n=1}^{\infty}$. Если $A = \mathbb{R}$, то говорят, что задана *числовая последовательность*. Именно числовые последовательности обсуждаются в этом разделе.

Определение предела последовательности

Число a называется *пределом последовательности* $\{a_n\}$, если для всякого $\varepsilon > 0$ найдётся такой номер N, что для всех $n > N$ имеет место неравенство $|a_n - a| < \varepsilon$. Иными словами, номер N таков, что из неравенства $n > N$ следует неравенство $|a_n - a| < \varepsilon$.

Используются обозначения $\lim\limits_{n\to\infty} a_n = a$ и $a_n \to a$ при $n \to \infty$.

Для сокращения записи иногда применяются кванторы $\forall$ — «для всякого» и $\exists$ — «существует». Например, определение предела последовательности можно записать так:

$$\lim_{n\to\infty} a_n = a, \text{ если } \forall\,\varepsilon > 0\ \exists\, N\ \forall\, n \in \mathbb{N} \quad n > N \Rightarrow |a_n - a| < \varepsilon.$$

Через $U_\varepsilon(a)$ будем обозначать интервал $(a-\varepsilon; a+\varepsilon)$ с центром в точке a и «радиусом» ε; через $U(a)$ будем обозначать произвольный интервал, содержащий точку a.

Теорема 18. *Следующие утверждения равносильны.*

(i) $\lim\limits_{n\to\infty} a_n = a$.

(ii) *Для всякого $\varepsilon > 0$ все члены последовательности $\{a_n\}$, кроме, быть может, конечного числа, принадлежат $U_\varepsilon(a)$.*

(iii) *Для всякого интервала $U(a)$ все члены последовательности $\{a_n\}$, кроме, быть может, конечного числа, принадлежат $U(a)$.*

$\triangleright$ Выведем (ii) из (i). По определению для всякого $\varepsilon > 0$ найдётся такой номер N, что для всех $n > N$ выполняется неравенство $|a_n - a| < \varepsilon$, что равносильно условию $a_n \in U_\varepsilon(a)$. Таким образом, все элементы последовательности $\{a_n\}$, кроме, быть может, каких-то элементов с номерами от 1 до N, принадлежат $U_\varepsilon(a)$. Для вывода (i) из (ii) достаточно взять в определении предела в качестве N наибольший из номеров членов последовательности, которые не принадлежат $U_\varepsilon(a)$. Для доказательства равносильности (ii) и (iii) достаточно заметить, что для всякого интервала $U(a)$ существует такое $\varepsilon > 0$, что $U_\varepsilon(a) \subset U(a)$. $\triangleleft$

Теорема 19 (единственность предела)**.** *Если $\lim\limits_{n\to\infty} a_n = a$ и $\lim\limits_{n\to\infty} a_n = b$, то $a = b$.*

$\triangleright$ В случае различных a и b можно указать непересекающиеся интервалы $U(a)$ и $U(b)$. По определению предела последовательности в каждом из этих интервалов содержатся все элементы последовательности, кроме, быть может, конечного числа. Это противоречит тому, что интервалы не пересекаются. $\triangleleft$

Свойства предела

Последовательность $\{a_n\}$ называется *ограниченной*, если существует такое число $C > 0$, что $|a_n| \leqslant C$ при всех $n \in \mathbb{N}$, т. е. все члены последовательности лежат в отрезке $[-C; C]$.

Теорема 20 (ограниченность сходящейся последовательности)**.** *Если последовательность имеет предел, то она ограниченна.*

$\triangleright$ Пусть $\lim\limits_{n\to\infty} a_n = a$. Поскольку $a \in (-|a|-1; |a|+1)$, по определению предела последовательности найдётся такой номер N, что $a_n \in (-|a|-1; |a|+1)$ для всех $n > N$. Положим

$$C = \max\{|a|+1; |a_1|, |a_2|, \dots, |a_N|\},$$

Тогда все члены последовательности $\{a_n\}$ принадлежат отрезку $[-C; C]$. $\triangleleft$

Лемма 2 (об отделимости). *Если* $\lim\limits_{n\to\infty} a_n = a$ *и* $a > 0$, *то существует такое число* N, *что* $a_n > a/2 > 0$ *для всех* $n > N$.

$\triangleright$ Достаточно положить в определении предела последовательности $\varepsilon = a/2$. $\triangleleft$

Теорема 21 (арифметика предела). *Если* $\lim\limits_{n\to\infty} a_n = a$ *и* $\lim\limits_{n\to\infty} b_n = b$, *то* $\lim\limits_{n\to\infty} (a_n \pm b_n) = a \pm b$, $\lim\limits_{n\to\infty} a_n b_n = ab$ *и в случае, когда* $b \neq 0$, $b_n \neq 0$, *справедливо равенство* $\lim\limits_{n\to\infty} a_n/b_n = a/b$.

$\triangleright$ Все доказательства однотипны, поэтому докажем лишь последнее утверждение. Можно считать, что $b > 0$ (случай $b < 0$ полностью аналогичен). По лемме об отделимости найдётся такой номер N_1, что $b_n > b/2$. Кроме того, для всякого $\varepsilon > 0$ найдётся такой номер N_2, что $|a_n - a| < \varepsilon$ и $|b_n - b| < \varepsilon$ для всех $n > N_2$. Заметим, что

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| = \left|\frac{a_n b - b_n a}{b_n b}\right| = \left|\frac{(a_n - a)b + a(b - b_n)}{b_n b}\right|.$$

Тогда для всех $n > \max\{N_1, N_2\}$ справедливы неравенства

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| \leqslant \frac{|a_n - a|b + |a||b - b_n|}{|b_n|b} \leqslant \frac{2b + 2|a|}{b^2}\varepsilon. \quad \triangleleft$$

Из теоремы об арифметике предела следует свойство *линейности предела*: если $\lim\limits_{n\to\infty} a_n = a$ и $\lim\limits_{n\to\infty} b_n = b$, то для всех $\lambda, \mu \in \mathbb{R}$ справедливо равенство $\lim\limits_{n\to\infty} (\lambda a_n + \mu b_n) = \lambda a + \mu b$.

Теорема 22 (переход к пределу в неравенстве). *Пусть* $\lim\limits_{n\to\infty} a_n = a$ *и* $\lim\limits_{n\to\infty} b_n = b$. *Если существует такой номер* N, *что* $a_n \leqslant b_n$ *для всех* $n > N$, *то* $a \leqslant b$.

$\triangleright$ Если $a > b$, то $(a_n - b_n) \to (a - b) > 0$ и найдётся такой номер N_1, что для всех $n > N_1$ верно неравенство $a_n - b_n > 0$. Следовательно, для всех $n > \max\{N, N_1\}$ одновременно верны неравенства $a_n > b_n$ и $a_n \leqslant b_n$, что невозможно. $\triangleleft$

Теорема 23 (о зажатой последовательности). *Если* $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n = a$ *и* $a_n \leqslant c_n \leqslant b_n$ *для всех* $n > N$, *то* $\lim\limits_{n\to\infty} c_n = a$.

$\triangleright$ По условию $a_n \leqslant c_n \leqslant b_n$, поэтому если $a_n, b_n \in U(a)$, то также $c_n \in U(a)$. Следовательно, если окрестности $U(a)$ принадлежат все члены последовательностей $\{a_n\}$ и $\{b_n\}$, кроме конечного числа, то этой же окрестности принадлежат все члены последовательности $\{c_n\}$, кроме конечного числа. $\triangleleft$

Теорема 24 (Вейерштрасс). *Монотонная ограниченная последовательность сходится.*

$\triangleright$ Пусть последовательность $\{a_n\}$ неубывающая и ограниченная. Тогда существует $A = \sup\limits_{n\in\mathbb{N}}\{a_n\}$. По определению точной верхней грани число $A - \varepsilon$ не является верхней гранью при $\varepsilon > 0$. Следовательно, найдётся такой номер N, что $a_N > A - \varepsilon$. В силу монотонности справедливо неравенство $a_n > A - \varepsilon$ для всех $n > N$. $\triangleleft$

Число e

Рассмотрим последовательность $\{x_n\}$, где $x_n = \left(1 + \dfrac{1}{n}\right)^n$. Раскроем скобки, используя бином Ньютона:

$$x_n = \sum_{k=0}^{n} C_n^k n^{-k} = \sum_{k=0}^{n} \frac{1}{k!}\Big(1 - \frac{1}{n}\Big)\cdots\Big(1 - \frac{k-1}{n}\Big).$$

Заметим, что из этого представления следует возрастание последовательности $\{x_n\}$. Действительно,

$$x_n < \sum_{k=0}^{n} \frac{1}{k!}\Big(1 - \frac{1}{n+1}\Big)\cdots\Big(1 - \frac{k-1}{n+1}\Big) < \sum_{k=0}^{n+1} \frac{1}{k!}\Big(1 - \frac{1}{n+1}\Big)\cdots\Big(1 - \frac{k-1}{n+1}\Big) = x_{n+1}.$$

Применяя неравенство $k! \geqslant 2^{k-1}$, $k \in \mathbb{N}$, получаем

$$x_n \leqslant \sum_{k=0}^{n} \frac{1}{k!} \leqslant 1 + 1 + \frac{1}{2} + \frac{1}{4} + \ldots < 3.$$

По теореме Вейерштрасса существует предел $\lim\limits_{n\to\infty} x_n$. Этот предел называют числом e. Определение числа e позволяет найти его значение с любой наперёд заданной точностью. Можно показать, что $e = 2{,}718281828459045\ldots$. Кроме того, число e иррационально, см. задачу 5.7.

Отметим, что данная последовательность и её предел исторически появились в задаче о сложных процентах. Предположим, что в банк под 100% годовых положили один рубль. Тогда в конце года на счету будет уже два рубля. Представим теперь, что весь год разбили на 100 периодов и в конце каждого из них начислили 1% от имеющейся к концу периода суммы. Тогда в конце года на счету будет уже $\left(1 + \frac{1}{100}\right)^{100}$ рублей. Эта сумма уже больше двух рублей и естественным образом появляется вопрос о том, к чему будет приближаться величина $\left(1 + \frac{1}{n}\right)^{n}$ при стремлении n к бесконечности. Ответом, как мы уже знаем, является число e, поэтому исходная сумма увеличится почти в три раза.

Подпоследовательности. Критерий Коши. Ряды

Пусть дана последовательность $\{a_n\}$. Если задана возрастающая последовательность натуральных чисел

$$n_1 < n_2 < \ldots < n_k < n_{k+1} < \ldots,$$

то соответствующая ей последовательность $\{a_{n_k}\}$ называется *подпоследовательностью* последовательности $\{a_n\}$.

Теорема 25 (Больцано). *Во всякой ограниченной последовательности есть сходящаяся подпоследовательность.*

$\triangleright$ Пусть все члены некоторой последовательности $\{a_n\}$ принадлежат отрезку $[c; d]$. Разделим отрезок $[c; d]$ пополам и выберем ту половину, в которой лежит бесконечно много членов последовательности. Эту половину вновь разделим пополам и опять выберем половину, в которой лежит бесконечно много членов последовательности. В результате построения получаем последовательность вложенных отрезков $[c_k; d_k]$, причём $d_k - c_k = (d - c)/2^{k-1}$, $[c_1; d_1] = [c; d]$, и в каждом из отрезков лежит бесконечно много членов последовательности $\{a_n\}$. По теореме о вложенных отрезках существует точка a, принадлежащая всем построенным отрезкам. Построим теперь последовательность возрастающих номеров n_k следующим образом: $n_1 = 1$ и, если уже построили n_k, то находим такой номер $n_{k+1} > n_k$, что $a_{n_{k+1}} \in [c_{k+1}; d_{k+1}]$ (элемент $a_{n_{k+1}}$ найдётся, так как на этом отрезке лежит бесконечно много членов последовательности). Поскольку $|a_{n_k} - a| \leqslant d_k - c_k$ и $(d_k - c_k) \to 0$, получаем $a_{n_k} \to a$. $\triangleleft$

Частичным пределом последовательности называется предел некоторой её подпоследовательности.

Теорема 26. *Если последовательность сходится к некоторому числу, то всякая её подпоследовательность сходится к этому же числу.*

$\triangleright$ Пусть $a_n \to a$ и $\{a_{n_k}\}$ — подпоследовательность последовательности $\{a_n\}$. По индукции можно показать, что $n_k \geqslant k$. Запишем определение предела для $\{a_n\}$: для всякого $\varepsilon > 0$ найдётся такой номер N, что $|a_n - a| < \varepsilon$ для всех $n > N$. Итак, если $k > N$, то $n_k \geqslant k > N$ и $|a_{n_k} - a| < \varepsilon$. $\triangleleft$

Итак, если последовательность сходится, то её предел является единственным частичным пределом. Если ограниченная последовательность не сходится, то у неё есть не менее двух различных частичных пределов. Найдём границы множества частичных пределов ограниченной последовательности.

Теорема 27. *Пусть последовательность $\{a_n\}$ ограниченна. Тогда*

(i) *последовательность верхних граней «хвостов» $M_n = \sup\limits_{k>n} a_k$ является невозрастающей и сходится к некоторому числу M;*

(ii) *последовательность нижних граней «хвостов»* $m_n = \inf\limits_{k>n} a_k$ *является неубывающей и сходится к некоторому числу* m;

(iii) m *и* M *— частичные пределы последовательности* $\{a_n\}$;

(iv) *все частичные пределы лежат между* m *и* M.

$\triangleright$ Первые два утверждения немедленно следуют из определений точных граней и теоремы Вейерштрасса. Докажем (iii). Построим подпоследовательность, сходящуюся к числу M. Пусть элемент a_{n_k} уже построен. Найдём такой номер $n_{k+1} > n_k$, что $a_{n_{k+1}} \geqslant M_{n_k} - \frac{1}{k+1}$. Отсюда вытекает, что $a_{n_k} \to M$. Аналогично строим подпоследовательность, сходящуюся к m. Утверждение (iv) вытекает из неравенств $m_{n_{k-1}} \leqslant a_{n_k} \leqslant M_{n_{k-1}}$. $\triangleleft$

Пределы $\lim\limits_{n\to\infty} M_n$ и $\lim\limits_{n\to\infty} m_n$ называют *верхним* и *нижним* пределом последовательности $\{a_n\}$ и обозначают через $\varlimsup\limits_{n\to\infty} a_n$ и $\varliminf\limits_{n\to\infty} a_n$ соответственно. Для ограниченной последовательности $\{a_n\}$ можно показать, что $\lim\limits_{n\to\infty} a_n = a$ тогда и только тогда, когда $\varlimsup\limits_{n\to\infty} a_n = \varliminf\limits_{n\to\infty} a_n = a$. Следовательно, если ограниченная последовательность не сходится, то у неё есть не менее двух частичных пределов.

Последовательность $\{a_n\}$ называется *последовательностью Коши* или *фундаментальной последовательностью*, если для всякого $\varepsilon > 0$ найдётся такой номер N, что для всех $m, n > N$ выполняется неравенство $|a_n - a_m| < \varepsilon$.

Лемма 3. *Всякая фундаментальная последовательность ограниченна.*

$\triangleright$ Пусть последовательность $\{a_n\}$ фундаментальна. Тогда существует такой номер N, что для всех $n > N$ верно неравенство $|a_n - a_{N+1}| < 1$, и, в частности, $|a_n| \leqslant |a_{N+1}| + 1$. Положим

$$C = \max\{|a_1|, |a_2|, \ldots, |a_N|, |a_{N+1}| + 1\}.$$

Тогда $|a_n| \leqslant C$ для всех $n \in \mathbb{N}$. $\triangleleft$

Лемма 4. *Если у фундаментальной последовательности* $\{a_n\}$ *есть сходящаяся подпоследовательность* $\{a_{n_k}\}$, *которая сходится к некоторому числу* a, *то и вся фундаментальная последовательность сходится к этому числу* a.

$\triangleright$ По определению фундаментальной последовательности для всякого $\varepsilon > 0$ существует такой номер N, что для всех $n, m > N$ верно неравенство $|a_n - a_m| < \varepsilon$. Пусть $k > N$, тогда $n_k \geqslant k > N$ и $|a_n - a_{n_k}| < \varepsilon$. Устремим в последнем неравенстве $k \to \infty$ и получим $|a_n - a| \leqslant \varepsilon$. $\triangleleft$

Теорема 28 (критерий Коши)**.** *Последовательность действительных чисел сходится тогда и только тогда, когда она является последовательностью Коши.*

$\triangleright$ Если последовательность $\{a_n\}$ сходится к a, то её фундаментальность следует из неравенства

$$|a_n - a_m| \leqslant |a_n - a| + |a - a_m|.$$

Обратно, если известно, что последовательность фундаментальна, то она ограниченна и по теореме Больцано из неё можно выбрать сходящуюся подпоследовательность. Следовательно, фундаментальная последовательность сходится. $\triangleleft$

Будем говорить, что последовательность $\{a_n\}$ *стремится к* $+\infty$, если для всякого $M > 0$ найдётся такой номер N, что для всех $n > N$ верно неравенство $a_n > M$. В этом случае будем использовать обозначение $\lim\limits_{n\to\infty} a_n = +\infty$. Последовательность $\{a_n\}$ *стремится к* $-\infty$, если $\{(-a_n)\}$ стремится к $+\infty$. В этом случае будем писать $\lim\limits_{n\to\infty} a_n = -\infty$. Наконец, последовательность $\{a_n\}$ *стремится к* ∞, если $\{|a_n|\}$ стремится к $+\infty$. Такие последовательности называют *бесконечно большими*. Отметим, что несмотря на использование записи вида $\lim\limits_{n\to\infty} a_n = \infty$ бесконечно большие последовательности не принято называть сходящимися или имеющими предел. В частности, критерий Коши относится именно к последовательностям, сходящимся к числу, т. е. имеющим конечный предел. Например, последовательность $\{a_n\}$, где $a_n = n$, стремится к $+\infty$ и не является фундаментальной.

Если последовательность не является ограниченной, то в ней есть подпоследовательность, которая стремится к плюс или минус бесконечности, поэтому если допускать в качестве частичных пределов символы «$+\infty$» и «$-\infty$», то по теореме Больцано множество частичных пределов всякой последовательности *непусто*.

Ряды

Пусть задана числовая последовательность $\{a_n\}$. *Рядом* с членами a_n называется выражение

$$a_1 + a_2 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n.$$

Выражения $S_N = \sum\limits_{n=1}^{N} a_n$ называют частичными суммами ряда $\sum\limits_{n=1}^{\infty} a_n$. Говорят, что *ряд сходится*, если существует (конечный) предел его частичных сумм. Этот предел называют суммой ряда. В остальных случаях говорят, что ряд расходится. Отметим, что в некоторых случаях удобно начинать суммирование с нуля и рассматривать ряд $\sum_{n=0}^{\infty} a_n$. Рассмотрим сумму членов бесконечной геометрической прогрессии: $\sum_{n=0}^{\infty} q^n$. Частичная сумма S_N этого ряда равна $1+q+q^2+\ldots+q^{N-1}$. Если $q = 1$, то $S_N = N$, поэтому ряд расходится. При $q \neq 1$ по формуле суммы членов геометрической прогрессии получаем $S_N = \frac{1-q^N}{1-q}$. Отсюда следует, что ряд сходится только при $|q| < 1$, и его сумма равна $1/(1-q)$.

Полезно иметь ввиду, что на сходимость ряда (но не на сумму) не влияет изменение конечного числа слагаемых (в частности, добавление или исключение конечного числа членов). Кроме того, если ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится и его сумма равна A, то

$$\sum_{n=N+1}^{\infty} a_n = A - \sum_{n=1}^{N} a_n \to 0.$$

Конечно, к рядам примени́м критерий Коши (что мы обсудим позже), но во многих случаях можно обойтись несколькими элементарными наблюдениями.

Необходимое условие сходимости

Если ряд сходится, то последовательность его членов стремится к нулю. Это следует из того, что $a_n = S_n - S_{n-1}$, $n \geqslant 2$.

Теорема Вейерштрасса

Если члены ряда неотрицательны, то его частичные суммы образуют неубывающую последовательность и, следовательно, для его сходимости необходимо и достаточно ограниченности последовательности частичных сумм.

Последнее наблюдение позволяет доказывать сходимость или расходимость рядов с неотрицательными членами, сравнивая их с некоторыми эталонными рядами, из которых самым популярным является сумма бесконечной геометрической прогрессии. Например, если $0 \leqslant a_n \leqslant Cq^n$, $n \in \mathbb{N}$, и $0 < q < 1$, то ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится. Следующее утверждение является примером более тонкого сравнения рядов.

Теорема 29 (признак Коши)**.** *Пусть $a_n \geqslant 0$ и последовательность $\{a_n\}$ невозрастающая. Ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится тогда и только тогда, когда сходится ряд $\sum\limits_{n=1}^{\infty} 2^n a_{2^n}$.*

$\triangleright$ Утверждение следует из двойного неравенства

$$a_2 + 2a_4 + 4a_8 + \ldots + 2^n a_{2^{n+1}} \leqslant a_2 + \ldots + a_{2^{n+1}} \leqslant a_1 + 2a_2 + 4a_4 + \ldots + 2^n a_{2^n}. \qquad \triangleleft$$

Следствие 4. *Ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ сходится тогда и только тогда, когда $p > 1$.*

$\triangleright$ По признаку Коши рассматриваемый ряд сходится тогда и только тогда, когда сходится ряд $\sum\limits_{n=1}^{\infty} (2^{1-p})^n$, а это сумма геометрической прогрессии со знаменателем 2^{1-p}. $\triangleleft$

Расходящийся ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ называется *гармоническим.* Напомним, что средним гармоническим положительных чисел a, b называют частное $2 : \left(\frac{1}{a} + \frac{1}{b}\right)$. Можно проверить, что каждое слагаемое этого ряда, начиная со второго, является средним гармоническим своих соседей.

Переформулируем критерий Коши для последовательностей.

Теорема 30 (критерий Коши сходимости ряда). *Ряд* $\sum\limits_{n=1}^{\infty} a_n$ *сходится тогда и только тогда, когда для всякого* $\varepsilon > 0$ *найдётся такой номер* N, *что для всех* $n > m > N$ *выполнено неравенство* $\Bigl|S_n - S_m\Bigr| = \Bigl|\sum\limits_{k=m+1}^{n} a_k\Bigr| < \varepsilon$.

Следствие 5. *Если сходится ряд* $\sum\limits_{n=1}^{\infty} |a_n|$, *то сходится и ряд* $\sum\limits_{n=1}^{\infty} a_n$.

$\triangleright$ Достаточно применить неравенство

$$|a_{m+1} + \ldots + a_n| \leqslant |a_{m+1}| + \ldots + |a_n|$$

и критерий Коши. $\triangleleft$

Абсолютная и условная сходимость

Если сходится ряд $\sum\limits_{n=1}^{\infty} |a_n|$, то говорят, что ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится *абсолютно*; в этом случае верно неравенство $\bigl|\sum\limits_{n=1}^{\infty} a_n\bigr| \leqslant \sum\limits_{n=1}^{\infty} |a_n|$.

В случае когда ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится, а ряд $\sum\limits_{n=1}^{\infty} |a_n|$ расходится, говорят, что ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится *условно*. Примером условно сходящегося ряда является ряд Лейбница $\sum\limits_{n=1}^{\infty} \frac{(-1)^n}{n}$. Для доказательства его сходимости применим критерий Коши. Пусть $m < n$, тогда

$$\Bigl|\frac{(-1)^{m+1}}{m+1} + \frac{(-1)^{m+2}}{m+2} + \ldots + \frac{(-1)^n}{n}\Bigr| = \frac{1}{m+1} - \frac{1}{m+2} + \frac{1}{m+3} - \frac{1}{m+4} + \ldots =$$
$$= \frac{1}{m+1} + \Bigl(-\frac{1}{m+2} + \frac{1}{m+3}\Bigr) + \Bigl(-\frac{1}{m+4} + \frac{1}{m+5}\Bigr) + \ldots \leqslant \frac{1}{m+1} + \frac{1}{n}.$$

Таким образом, если для всякого $\varepsilon > 0$ взять какое-либо $N \geqslant \frac{1}{2\varepsilon}$, то тогда для всяких $n, m > N$ будет верна оценка

$$\Bigl|\sum_{k=1}^{n} \frac{(-1)^k}{k} - \sum_{k=1}^{m} \frac{(-1)^k}{k}\Bigr| \leqslant \frac{1}{m+1} + \frac{1}{n} < \frac{1}{2N} \leqslant \varepsilon.$$

Важно отметить, что всякую последовательность $\{x_n\}$ можно представить в виде последовательности частичных сумм некоторого ряда:

$$x_n = x_1 + (x_2 - x_1) + (x_3 - x_2) + \ldots + (x_n - x_{n-1}),$$

т. е. x_n — частичная сумма ряда $\sum\limits_{n=1}^{\infty} (x_n - x_{n-1})$, где $x_0 = 0$. Таким образом, сходимость ряда $\sum\limits_{n=1}^{\infty} (x_n - x_{n-1})$ равносильна сходимости последовательности $\{x_n\}$.

Топология вещественной прямой

Окрестностью точки $a \in \mathbb{R}$ называют всякий интервал, содержащий точку a. Внутри всякой окрестности точки a всегда можно выбрать симметричную окрестность

$$U_r(a) = \{x \in \mathbb{R} \colon |x - a| < r\}.$$

Множество $U'_r(a) = U_r(a) \setminus \{a\}$ называется *проколотой окрестностью* точки a. *Открытым множеством* в $\mathbb{R}$ называют всякое множество, которое вместе с любой своей точкой содержит некоторую её окрестность. Например: интервал, вся числовая прямая, пустое множество. *Замкнутым множеством* в $\mathbb{R}$ называют всякое множество, дополнение к которому является открытым множеством. Например: точка, отрезок, вся числовая прямая, пустое множество.

Теорема 31. *Объединение любого набора и пересечение конечного набора открытых множеств является открытым множеством. Пересечение любого набора и объединение конечного набора замкнутых множеств является замкнутым множеством.*

$\triangleright$ Пусть $V = \bigcup\limits_{s\in S} V_s$, где V_s — открытые множества. Если $a \in V$, то найдётся такое V_s, что $a \in V_s$. По определению открытого множества найдётся окрестность $U(a) \subset V_s$. По определению объединения множеств получаем $U(a) \subset V$. Таким образом, V — открытое множество. Пусть теперь $V = \bigcap\limits_{n=1}^{N} V_n$, где V_n — открытые множества. Если $a \in V$, то $a \in V_n$ для каждого $n = 1, 2, \ldots, N$. Для каждого n существуют окрестности $U_{r_n}(a) \subset V_n$. Положим $r = \min\{r_1, r_2, \ldots, r_N\}$. Тогда $U_r(a) \subset V_n$ для всех $n \in \mathbb{N}$, а следовательно, $U_r(a) \subset V$. Таким образом, V — открытое множество. Утверждения об объединении и пересечении замкнутых множеств выводятся из уже доказанных утверждений для открытых множеств с помощью формул Моргана. $\triangleleft$

Открытые множества на числовой прямой допускают следующее описание.

Теорема 32. *Всякое открытое множество в $\mathbb{R}$ или пусто, или совпадает с $\mathbb{R}$, или является объединением не более чем счётного набора попарно непересекающихся интервалов и лучей.*

$\triangleright$ Пусть V — непустое открытое множество на числовой прямой, $V \neq \mathbb{R}$ и $a \in V$. Рассмотрим множество $D = \{x\colon (x; a) \subset V\}$. Если D не ограничено снизу, то $(-\infty; a) \subset V$. Если D ограничено снизу, то положим $\mu = \inf D$. Из определения точной нижней грани следует, что $(\mu; a) \subset V$. Аналогично рассматриваем множество $E = \{x\colon (a; x) \subset V\}$: если E не ограничено сверху, то $(a; +\infty) \subset V$. Если E ограничено сверху, то положим $\sigma = \sup E$. Из определения точной верхней грани следует, что $(a; \sigma) \subset V$. Таким образом, для всякой точки $a \in V$ существует такой содержащий a интервал или луч $(\mu; \sigma)$, что всякий интервал (луч) $(\alpha; \beta) \supset (\mu; \sigma)$, $(\alpha; \beta) \neq (\mu; \sigma)$, не содержится в множестве V. Найдём такие интервалы для всякой точки $a \in V$. По построению, если эти интервалы пересекаются, то они совпадают. Следовательно, различные интервалы из этого набора не могут пересекаться. Наконец, заметим, что на числовой прямой можно расположить лишь не более чем счётный набор попарно непересекающихся интервалов. $\triangleleft$

Следствие 6. *Числовую прямую нельзя представить в виде объединения непустых непересекающихся открытых (замкнутых) множеств.*

$\triangleright$ Если два отрытых множества не пересекаются, то концы интервалов, образующих одно из них, не принадлежат другому. Если прямая является объединением двух непересекающихся множеств, одно из которых открыто, то концы интервалов, образующих это открытое множество, лежат в его дополнении. Следовательно, второе множество не может быть открытым. $\triangleleft$

Точка a называется *внутренней* точкой множества $A \subset \mathbb{R}$, если она входит в это множество с некоторой своей окрестностью. Множество открыто тогда и только тогда, когда все его точки внутренние.

Точка a называется *предельной* точкой множества $A \subset \mathbb{R}$, если во всякой её окрестности лежит бесконечно много точек множества A. Всякая внутренняя точка множества является его предельной точкой.

Точка a называется *граничной* точкой множества $A \subset \mathbb{R}$, если во всякой её окрестности есть как точки множества A, так и точки из дополнения к множеству A. Граничная точка множества не обязана принадлежать этому множеству. Если множество состоит из одной точки, то эта точка является граничной.

Теорема 33. *Точка a является предельной для множества A тогда и только тогда, когда для всякого $r > 0$ проколотая окрестность $U'_r(a)$ содержит хотя бы одну точку множества A.*

$\triangleright$ Если точка a является предельной для множества A, то всякая окрестность $U_r(a)$ содержит бесконечно много точек множества A, а проколотая окрестность $U'_r(a)$ получается из $U_r(a)$ исключением только одной точки. Пусть теперь известно, что во всякой проколотой окрестности точки a есть хотя бы одна точка множества A. Предположим, что a не является предельной точкой множества A. Тогда найдётся окрестность $U_r(a)$, в которой лежит не более чем конечное число точек множества A. Значит, уменьшая r, можно найти окрестность точки a, в которой точек из A, отличных от a, просто нет, что невозможно по условию. $\triangleleft$

Теорема 34 (Больцано)**.** *У всякого бесконечного ограниченного множества есть предельная точка.*

$\triangleright$ Поскольку множество бесконечно, найдётся последовательность $\{a_n\}$ его элементов, в которой все члены попарно различны. Эта последовательность ограниченна и по теореме Больцано существует подпоследовательность $\{a_{n_k}\}$, которая сходится к некоторому числу a. Из определения предела последовательности и того, что все члены подпоследовательности являются попарно различными элементами множества A, следует, что a — предельная точка. $\triangleleft$

Следствие 7. *У всякого бесконечного несчётного множества есть предельная точка.*

$\triangleright$ Действительно, найдётся отрезок $[n; n+1]$, где $n \in \mathbb{Z}$, в котором лежит бесконечно много точек данного множества. По теореме Больцано на этом отрезке есть предельная точка данного множества. $\triangleleft$

Выше мы описали структуру открытых множеств на числовой прямой. Теперь опишем структуру замкнутых множеств.

Теорема 35. *Следующие утверждения равносильны.*

(i) *Множество A замкнуто.*

(ii) *Множество A содержит все свои граничные точки.*

(iii) *Множество A содержит все свои предельные точки.*

(iv) *Всякая сходящаяся последовательность элементов множества A сходится к элементу множества A.*

$\triangleright$ Из (i) следует (ii), так как точка из дополнения к замкнутому множеству A входит в дополнение с некоторой окрестностью, что невозможно для граничной точки. Поскольку всякая предельная точка множества является либо его внутренней либо его граничной точкой, из (ii) следует (iii). Из (iii) следует (iv). Действительно, если последовательность сходится, то либо в этой последовательности бесконечно много членов совпадает с пределом, и значит, предел лежит в A, либо предел является предельной точкой множества A. Наконец, из (iv) следует (i). Действительно, если точка a из дополнения к A не лежит в этом дополнении вместе с некоторой окрестностью, то во всякой сколь угодно малой окрестности есть точка множества A. Следовательно, существует последовательность элементов множества A, сходящаяся к a. Тогда $a \in A$, что противоречит исходному предположению о том, что a принадлежит дополнению A. $\triangleleft$

Важный пример замкнутого множества — множество частичных пределов последовательности. Пусть L — множество частичных пределов последовательности $\{a_n\}$. Предположим, что последовательность l_k элементов L сходится к некоторому числу l. Поскольку l_k — частичные пределы последовательности $\{a_n\}$, существует такая возрастающая последовательность номеров n_k, что $|l_k - a_{n_k}| < 1/k$. Из оценки $|a_{n_k} - l| \leqslant 1/k + |l_k - l|$ следует сходимость последовательности $\{a_{n_k}\}$ к l. Таким образом, $l \in L$, и замкнутость L доказана.

Теорема 36 (Бэр)**.** *Если числовая прямая представлена в виде объединения не более чем счётного набора замкнутых множеств, то хотя бы одно из этих множеств содержит интервал.*

$\triangleright$ Предположим, что числовая прямая является объединением не более чем счётного набора замкнутых множеств F_n, ни одно из которых никакого интервала не содержит. Дополнение к множеству F_1 открыто и непусто, следовательно, содержит интервал, внутри которого можно выбрать отрезок. Таким образом, найдётся отрезок $[a_1; b_1]$, который не имеет общих точек с множеством F_1. Интервал $(a_1; b_1)$ не содержится в множестве F_2, а значит, пересечение интервала $(a_1; b_1)$ с дополнением к множеству F_2 открыто и непусто. Это пересечение содержит интервал, внутри которого можно выбрать отрезок. Таким образом, внутри отрезка $[a_1; b_1]$ найдётся отрезок $[a_2; b_2]$, который не имеет общих точек с множеством F_2. Продолжая построение, получим последовательность вложенных отрезков, у которых есть общая точка. Эта точка не принадлежит ни одному из множеств F_n, хотя по условию она должна принадлежать хотя бы одному из них. Противоречие. $\triangleleft$

Из теоремы Бэра, в частности, следует, что множество иррациональных чисел нельзя представить в виде объединения не более чем счётного объединения замкнутых множеств. Если бы это было возможно, то вместе с одноточечными множествами рациональных чисел эти множества дали бы в объединении всю числовую прямую. По теореме Бэра хотя бы одно из этих множеств должно было бы содержать интервал, что невозможно. Из теоремы Бэра можно вывести несчётность множества вещественных чисел. Действительно, если множество вещественных чисел не более чем счётно, то оно есть объединение не более чем счётного набора замкнутых одноточечных множеств. По теореме Бэра хотя бы одно из этих одноточечных множеств должно содержать целый интервал, что невозможно.

Утверждение теоремы Бэра справедливо и при замене числовой прямой на отрезок; доказательство потребует некоторых технических (но не идейных) изменений. Справедливо также и более общее утверждение. Пусть замкнутое множество F является объединением не более чем счётного набора замкнутых множеств F_n. Тогда найдётся такой интервал $(\alpha;\beta)$ и номер n, что $F \cap (\alpha;\beta) \neq \varnothing$ и $F \cap (\alpha;\beta) \subset F_n$.

Отметим, что теорему Бэра можно использовать для доказательства непустоты пересечения бесконечной последовательности открытых множеств: если для каждого n открытое множество V_n *всюду плотно* в $\mathbb{R}$, т. е. во всяком интервале есть точка этого множества, то пересечение $\bigcap_n V_n$ непусто и даже всюду плотно в $\mathbb{R}$.

Компакты и их свойства

Говорят, что набор множеств $\{U_\alpha\}$ образует *покрытие* множества A, если $A \subset \bigcup_\alpha U_\alpha$ (т. е. каждый элемент множества A принадлежит хотя бы одному множеству U_α). *Компактом* называется такое множество, что из всякого его покрытия открытыми множествами можно выбрать конечное подпокрытие. Иными словами, если компакт K покрыт набором открытых множеств $\{U_\alpha\}$, то существуют такие α_1, …, α_n, что множество K покрыто набором множеств $\{U_{\alpha_1}, \ldots, U_{\alpha_n}\}$.

Теорема 37 (Борель — Гейне — Лебег)**.** *Всякий отрезок является компактом.*

$\triangleright$ Предположим, что некоторый отрезок не является компактом. Это означает, что существует покрытие отрезка открытыми множествами, из которого нельзя выбрать конечное подпокрытие. Разделим отрезок пополам. Хотя бы одна из получившихся половин не допускает конечного подпокрытия, иначе весь отрезок имел бы конечное подпокрытие. Выберем такую половину, снова разделим полученный отрезок пополам и выберем его половину, не допускающую конечного подпокрытия. Продолжая построение, получим последовательность вложенных отрезков, длины которых стремятся к нулю, и для каждого из которых нельзя выбрать конечное подпокрытие. По принципу вложенных отрезков у этой последовательности имеется общая точка. Эта точка покрыта некоторым открытым множеством. Длины отрезков стремятся к нулю, поэтому рано или поздно отрезки будут содержаться внутри этого открытого множества, что по построению невозможно. $\triangleleft$

Теорема 38. *Справедливы следующие утверждения.*

(i) *Компактное множество ограниченно.*

(ii) *Компактное множество замкнуто.*

(iii) *Замкнутое подмножество компакта есть компакт.*

$\triangleright$ Обоснуем (i). Пусть K — компакт. Тогда $K \subset \bigcup_{n=1}^{\infty}(-n;n)$. Выберем конечное подпокрытие: $(-n_k;n_k)$, где $k = 1, 2, \ldots, N$. Положим $C = \max\{n_1, n_2, \ldots, n_N\}$. Тогда $K \subset (-C;C)$.

Перейдём к утверждению (ii). Пусть K — компакт и $a \in \mathbb{R} \setminus K$. Тогда открытые множества $U_n = \mathbb{R} \setminus [a - 1/n; a + 1/n]$ покрывают K. Выбираем конечное подпокрытие $U_{n_1}, \ldots, U_{n_N}$. Пусть $C = \max\{n_1, n_2, \ldots, n_N\}$. Тогда интервал $(a - 1/C; a + 1/C)$ лежит в дополнении к K.

Докажем утверждение (iii). Пусть K — компакт и $F \subset K$ — замкнутое множество. Для всякого покрытия $\{V_s\}_{s\in S}$ множества F открытыми множествами V_s эти множества вместе с открытым множеством $\mathbb{R}\setminus F$ покрывают K. Выберем конечное подпокрытие $V_{s_1}, \ldots, V_{s_N}$ и, быть может, $\mathbb{R}\setminus F$. Тогда открытые множества $V_{s_1}, \ldots, V_{s_N}$ покрывают F. $\triangleleft$

Из этой теоремы и компактности отрезка следует критерий компактности множества на числовой прямой.

Следствие 8. *Множество на числовой прямой компактно тогда и только тогда, когда оно ограниченно и замкнуто.*

Следствие 9. *Множество на числовой прямой является компактом тогда и только тогда, когда из всякой последовательности элементов этого множества можно выбрать подпоследовательность, сходящуюся к элементу этого множества.*

$\triangleright$ Если множество компактно, то оно ограниченно и замкнуто. Следовательно, по теореме Больцано из любой последовательности его элементов можно выбрать сходящуюся подпоследовательность, а в силу замкнутости предел этой подпоследовательности принадлежит компакту.

Докажем теперь, что из существования в любой последовательности элементов данного множества сходящейся подпоследовательности к элементу этого множества следует компактность этого множества. Если множество не является ограниченным, то можно предъявить стремящуюся к бесконечности последовательность его элементов, из которой нельзя выбрать сходящуюся к конечному числу подпоследовательность. Следовательно, множество ограниченно. Всякая сходящаяся последовательность элементов этого множества сходится к элементу этого множества, поэтому данное множество замкнуто. Как уже отмечалось выше, замкнутое и ограниченное множество на числовой прямой является компактом. $\triangleleft$

Важным примером компакта является *множество Кантора.* Разделим отрезок $[0;1]$ на три части точками $1/3$ и $2/3$ и исключим интервал $(1/3;2/3)$, оставив два отрезка $\Delta_{1,1} = [0;1/3]$ и $\Delta_{1,2} = [2/3;1]$. Пусть $F_1 = \Delta_{1,1} \cup \Delta_{1,2}$. Предположим, что уже построены отрезки $\Delta_{n,k}$. Разделим каждый из них на три части и исключим средний интервал. Объединяем получившуюся систему отрезков $\Delta_{n+1,s}$ и обозначаем объединение через F_{n+1}. Множество Кантора C определяется равенством $C = \bigcap_n F_n$. Это множество обладает многими замечательными свойствами, из которых отметим следующие:

1) C — компакт;

2) C — континуальное множество, состоящее из тех и только тех точек отрезка $[0;1]$, которые в троичной системе счисления можно записать, не используя единицу (но допуская бесконечные последовательности нулей или двоек);

3) суммарная длина выброшенных интервалов равна единице, т. е. в определённом смысле длина множества C равна нулю.

Выше мы показали, что всякое открытое множество является не более чем счётным объединением попарно непересекающихся интервалов. Поскольку всякое замкнутое множество является дополнением к такой системе интервалов, может сложится ложное мнение, что всякое замкнутое множество является не более чем счётным набором точек и отрезков. Множество Кантора разрушает эту иллюзию: в нём нет внутренних точек, но множество его предельных точек совпадает с самим множеством C.

Предел функции

Пусть $D \subset \mathbb{R}$ и $f\colon D \to \mathbb{R}$. Напомним, что при такой записи мы предполагаем, что функция f определена на множестве D. Пусть a — предельная точка множества D. Если множество D не ограничено сверху, то в качестве точки a допускается $+\infty$, а если D не ограничено снизу, то в качестве a допускается $-\infty$.

Определение предела функции по Гейне

Число b называется *пределом* функции f при $x \to a$ (иногда добавляют: по множеству D), если для всякой последовательности $\{x_n\}$ элементов множества D, для которой $x_n \to a$ и $x_n \neq a$, последовательность $\{f(x_n)\}$ сходится к b. Обозначение:

$$\lim_{x\to a} f(x) = b.$$

Из свойств предела последовательности следуют *свойства предела функции.*

1) *Единственность предела.* Если $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{x\to a} f(x) = c$, то $b = c$.

Для свойств 2—4 предположим, что функции f, g и h определены на множестве D и a — предельная точка этого множества.

2) *Арифметика предела.* Пусть $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{x\to a} g(x) = c$. Тогда $\lim\limits_{x\to a}(f(x) \pm g(x)) = b \pm c$, $\lim\limits_{x\to a} f(x)g(x) = bc$, а в случае, когда $c \neq 0$ и $g(x) \neq 0$ при $x \in D$, имеем $\lim\limits_{x\to a} f(x)/g(x) = b/c$.

Из теоремы об арифметике предела следует свойство *линейности предела*: если $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{x\to a} g(x) = c$, то для всех $\lambda, \mu \in \mathbb{R}$ справедливо равенство $\lim\limits_{x\to a}(\lambda f(x) + \mu g(x)) = \lambda b + \mu c$.

3) *Переход к пределу в неравенствах.* Если в некоторой проколотой окрестности точки a выполнено неравенство $f \leqslant g$ и существуют пределы $\lim\limits_{x\to a} f(x) = b$, $\lim\limits_{x\to a} g(x) = c$, то $b \leqslant c$.

4) *Предел зажатой функции.* Если в некоторой проколотой окрестности точки a выполнено неравенство $f \leqslant h \leqslant g$ и $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = b$, то $\lim\limits_{x\to a} h(x) = b$.

5) *Предел композиции.* Пусть $f\colon D \to E$ и $g\colon E \to \mathbb{R}$. Предположим, что a — предельная точка множества D и существует предел $\lim\limits_{x\to a} f(x) = b$, где b — предельная точка множества E, причём $f(x) \neq b$ в некоторой проколотой окрестности точки a. Тогда если $\lim\limits_{y\to b} g(y) = c$, то $\lim\limits_{x\to a} g(f(x)) = c$.

Доказываются все перечисленные утверждения однотипно путём сведения к пределу последовательности, поэтому рассмотрим лишь доказательство утверждения 5.

$\triangleright$ Пусть $x_n \to a$ и $x_n \neq a$. Поскольку $f \neq b$ в некоторой проколотой окрестности точки a, найдётся такой номер N, что $f(x_n) \neq b$ для всех $n > N$. Последовательность $f(x_{N+1}), f(x_{N+2}), \ldots$ сходится к b и не один из её членов не равен b, поэтому последовательность $g(f(x_{N+1})), g(f(x_{N+2})), \ldots$ сходится к c. Отбрасывание или добавление конечного числа членов последовательности не влияет на её сходимость, следовательно, вся последовательность $\{g(f(x_n))\}$ также сходится к c. Таким образом, мы по определению предела показали, что $\lim\limits_{x\to a} g(f(x))$. $\triangleleft$

Определение предела функции по Коши

Пусть, как и выше, точка a является предельной точкой множества D. Число b называется пределом функции f при $x \to a$ (по множеству D), если для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для всех $x \in D$ из неравенства $0 < |x - a| < \delta$ следует неравенство $|f(x) - b| < \varepsilon$.

С использованием кванторов это определение можно записать так: $\lim\limits_{x\to a} f(x) = b$, если
$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, x \in D \quad 0 < |x - a| < \delta \Rightarrow |f(x) - b| < \varepsilon.$$
Напомним, что записи $0 < |x - a| < \delta$ и $x \in U'_\delta(a)$ равносильны.

Если множество D не ограничено сверху, то можно определить предел функции f при $x \to +\infty$ по множеству D следующим образом. Число b называется пределом функции f при $x \to +\infty$ (по множеству D), если для всякого $\varepsilon > 0$ найдётся такое $M > 0$, что для всех $x \in D$ из неравенства $x > M$ следует неравенство $|f(x) - b| < \varepsilon$. Аналогично, если множество D не ограничено снизу, определяется предел функции при $x \to -\infty$.

Теорема 39. *Определения предела по Коши и по Гейне равносильны.*

$\triangleright$ Пусть функция f имеет предел по Коши при $x \to a$ (по множеству D), и этот предел равен b. Рассмотрим произвольную последовательность $\{x_n\}$ элементов множества D, для которой $x_n \to a$ и $x_n \neq a$, и покажем, что последовательность $\{f(x_n)\}$ сходится к b. По определению Коши для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для всех $x \in D$ из неравенства $0 < |x - a| < \delta$ следует неравенство $|f(x) - b| < \varepsilon$. Поскольку $x_n \to a$ и $x_n \neq a$, существует такое число N, что при всех $n > N$ выполнено неравенство $0 < |x_n - a| < \delta$, а значит, при всех $n > N$ выполнено неравенство $0 < |f(x_n) - b| < \varepsilon$. Это означает, что последовательность $\{f(x_n)\}$ сходится к b, т. е. функция f имеет предел по Гейне при $x \to a$ (по множеству D), и этот предел равен b.

Пусть теперь функция f имеет предел по Гейне при $x \to a$ (по множеству D), и этот предел равен b. Предположим, что число b не является пределом функции f по Коши при $x \to a$ (по множеству D). Это означает, что существует такое $\varepsilon > 0$, что для любого $\delta > 0$ найдётся такое значение $x \in D$, что $0 < |x - a| < \delta$, но $|f(x) - b| \geqslant \varepsilon$. Будем брать в качестве δ числа вида $1/n$, $n \in \mathbb{N}$, и обозначать соответствующее значение x через x_n. Тогда получим $0 < |x_n - a| < 1/n$, $n \in \mathbb{N}$, поэтому $x_n \to a$ и $x_n \neq a$. С другой стороны, $|f(x_n) - b| \geqslant \varepsilon$, $n \in \mathbb{N}$, поэтому последовательность $\{f(x_n)\}$ не сходится к числу b, а это противоречит определению предела по Гейне. $\triangleleft$

Приведём примеры утверждений, которые естественнее доказываются в терминах определения Коши.

Отделимость. Если $\lim\limits_{x\to a} f(x) = b > 0$, то найдётся такая проколотая окрестность $U'_\delta(a)$, что для всех $x \in U'_\delta(a) \cap D$ справедливо неравенство $f(x) > b/2 > 0$.

▷ Положим в определении предела $\varepsilon = b/2$. Тогда найдётся такое $\delta > 0$, что для всех $x \in U'_\delta(a) \cap D$ имеем $|f(x) - b| < b/2$. Следовательно, для всех $x \in U'_\delta(a) \cap D$ получаем

$$f(x) = f(x) - b + b > -\frac{b}{2} + b = \frac{b}{2}.$$ ◁

Ограниченность. Если существует конечный предел $\lim\limits_{x\to a} f(x)$, то найдутся такие проколотая окрестность $U'_\delta(a)$ и число $C > 0$, что для всех $x \in U'_\delta(a) \cap D$ справедливо неравенство $|f(x)| \leqslant C$.

Доказательство полностью аналогично предыдущему.

Первый замечательный предел:

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

▷ Это равенство выводится из неравенств $x \cos x \leqslant \sin x \leqslant x$ при $0 \leqslant x < \pi/2$ и теоремы о пределе зажатой функции. Неравенство $\sin x \leqslant x$ следует из того, что дуга длины $2x$ не короче стягивающей её хорды длины $2 \sin x$. Неравенство $x \cos x \leqslant \sin x$ следует из неравенства $x \leqslant \operatorname{tg} x$, которое, в свою очередь, является выражением того факта, что сумма длин отрезков касательных, проведенных к концам дуги длины $2x$, равна $2 \operatorname{tg} x$ и не короче этой дуги. ◁

Теоремы о существовании предела функции.

Если a — предельная точка множества $D^- = D \cap (-\infty; a)$ и существует предел функции f по множеству D^- при $x \to a$ (сокращённо пишут $x \to a - 0$ или $x \to a-$), то этот предел называют *пределом слева* функции f в точке a. Аналогично определяется предел справа.

Пусть a является предельной точкой для множеств D^- и D^+. Функция f имеет предел в точке a равный b тогда и только тогда, когда пределы функции f слева и справа в точке a существуют и равны b.

Теорема 40 (Вейерштрасс)**.** *Если функция $f\colon D \to \mathbb{R}$ неубывающая и ограниченная, а точка a — предельная для множества $D^- = D \cap (-\infty; a)$, то существует предел слева*

$$\lim_{x\to a-0} f(x) = \sup_{D^-} f(x).$$

▷ Пусть $\sup_{D^-} f(x) = L$. Для всякого $\varepsilon > 0$ найдётся точка $x_0 \in D^-$, для которой $f(x_0) > L - \varepsilon$. Поскольку f неубывающая, неравенство $f(x) > L - \varepsilon$ выполняется для всех точек $x \in (x_0; a) \cap D^-$. Кроме того, $f(x) \leqslant L$ при $x \in D^-$ по определению величины L. Таким образом, для всех $x \in (x_0; a) \cap D^-$ выполнено неравенство $|f(x) - L| < \varepsilon$. ◁

Аналогично можно показать, что если функция $f\colon D \to \mathbb{R}$ неубывающая и ограниченная, а точка a — предельная для множества $D^+ = D \cap (a; +\infty)$, то существует предел справа

$$\lim_{x\to a+0} f(x) = \inf_{D^+} f(x).$$

Следствие 10. *Если на интервале определена монотонная функция, то в каждой точке интервала у этой функции существуют пределы слева и справа.*

Теорема 41 (критерий Коши)**.** *Пусть $f\colon D \to \mathbb{R}$ и a — предельная точка множества D. Предел $\lim\limits_{x\to a} f(x)$ существует тогда и только тогда, когда для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что из неравенств $0 < |x - a| < \delta$ и $0 < |y - a| < \delta$ следует неравенство $|f(x) - f(y)| < \varepsilon$ (т. е. выполняется условие Коши).*

▷ Из неравенства $|f(x) - f(y)| \leqslant |f(x) - b| + |b - f(y)|$, где $b = \lim\limits_{x\to a} f(x)$, следует необходимость. Докажем достаточность условия Коши. Пусть $x_n \to a$ и $x_n \neq a$. Тогда последовательность $\{f(x_n)\}$ удовлетворяет условию Коши и поэтому сходится. Пусть помимо x_n задана другая последовательность $y_n \to a$, $y_n \neq a$. Тогда смешанная последовательность $z_n : x_1, y_1, x_2, y_2, \ldots$ сходится к a и $z_n \neq a$. Следовательно, последовательность $\{f(z_n)\}$ сходится, а значит, к тому же самому пределу сходятся её подпоследовательности $\{f(x_n)\}$ и $\{f(y_n)\}$. Итак, мы показали существование предела по Гейне. ◁

В заключение опишем абстрактную конструкцию, обобщающую уже рассмотренные нами понятия предела последовательности и предела функции.

Предел по базе

Пусть задано непустое множество X. *Базой* называется такой непустой набор $\mathcal{B}$ подмножеств множества X, что

1) в этом наборе нет пустого множества;

2) для всяких двух множеств $U, V \in \mathcal{B}$ найдётся третье множество $W \in \mathcal{B}$, для которого справедливо включение $W \subset U \cap V$.

Пусть $f\colon X \to \mathbb{R}$. Число b называется *пределом функции f по базе* $\mathcal{B}$ (пишем $\lim\limits_{\mathcal{B}} f(x) = b$), если для всякого $\varepsilon > 0$ найдётся такой элемент U базы $\mathcal{B}$, что для всех $x \in U$ справедливо неравенство $|f(x) - b| < \varepsilon$.

Примеры баз:

1) множества вида $\{n \in \mathbb{N}\colon n > N\}$; предел по такой базе — предел последовательности.

2) проколотые окрестности точки a; предел по этой базе — предел функции.

В качестве иллюстрации работы с пределом по базе докажем три полезных утверждения.

I (*Единственность предела*). Если $\lim\limits_{\mathcal{B}} f(x) = b$ и $\lim\limits_{\mathcal{B}} f(x) = c$, то $b = c$.

$\triangleright$ Пусть $\varepsilon > 0$. По определению предела по базе найдётся элемент U базы $\mathcal{B}$, на котором $|f(x) - b| < \varepsilon$, и элемент V базы $\mathcal{B}$, на котором $|f(x) - c| < \varepsilon$. Существует элемент $W \subset U \cap V$ базы $\mathcal{B}$. Поскольку $W \neq \varnothing$, существует $x \in W$. Для этого x имеем

$$|b - c| \leqslant |b - f(x)| + |f(x) - c| < 2\varepsilon.$$

Если $|b - c| > 0$, то при $\varepsilon = |b - c|/2$ приходим к противоречию. Следовательно, $b = c$. $\triangleleft$

II (*Линейность предела*). Если $\lim\limits_{\mathcal{B}} f(x) = b$ и $\lim\limits_{\mathcal{B}} g(x) = c$, то $\lim\limits_{\mathcal{B}}(\lambda f(x) + \mu g(x)) = \lambda b + \mu c$ для всех $\lambda, \mu \in \mathbb{R}$.

$\triangleright$ Для всякого $\varepsilon > 0$ найдём такие элементы U и V базы $\mathcal{B}$, что $|f(x) - b| \leqslant \varepsilon$ при $x \in U$ и $|g(x) - c| \leqslant \varepsilon$ при $x \in V$. Пусть $W \in \mathcal{B}$ и $W \subset U \cap V$. Тогда на этом элементе базы справедливо неравенство $|\alpha f(x) + \beta g(x) - \alpha b - \beta c| < \varepsilon(|\alpha| + |\beta|)$. $\triangleleft$

III (*Монотонность предела*). Пусть существуют пределы $\lim\limits_{\mathcal{B}} f(x) = b$, $\lim\limits_{\mathcal{B}} g(x) = c$ и на некотором элементе W базы $\mathcal{B}$ выполняется неравенство $f \leqslant g$. Тогда $b \leqslant c$.

$\triangleright$ Для всякого $\varepsilon > 0$ найдётся такой элемент U базы $\mathcal{B}$, что на нём выполняется неравенство $b - \varepsilon < f(x) \leqslant g(x) < c + \varepsilon$. Значит, $b < c + 2\varepsilon$. Если $b > c$, то при $\varepsilon = (b - c)/2$ получаем неравенство $b < b$, что невозможно. Следовательно, $b \leqslant c$. $\triangleleft$

Для сокращения записи удобно использовать «о-символику». Пусть задана база $\mathcal{B}$ и на некотором её элементе выполнено равенство $h(x) = \alpha(x) g(x)$. Тогда если $\lim\limits_{\mathcal{B}} \alpha(x) = 0$, то пишут $h = o(g)$ по базе $\mathcal{B}$; если функция $\alpha(x)$ ограниченна на некотором элементе базы $\mathcal{B}$, то пишут $h = O(g)$ по базе $\mathcal{B}$; если $\lim\limits_{\mathcal{B}} \alpha(x) = 1$, то пишут $h \sim g$ по базе $\mathcal{B}$.

Будем говорить, что функция f *стремится к* $+\infty$ по базе $\mathcal{B}$, если для всякого $M > 0$ найдётся такой элемент U базы $\mathcal{B}$, что на нём выполняется неравенство $f(x) > M$. В этом случае будем использовать обозначение $\lim\limits_{\mathcal{B}} f(x) = +\infty$. Функция f *стремится к* $-\infty$ по базе $\mathcal{B}$, если $(-f)$ стремится к $+\infty$; в этом случае будем писать $\lim\limits_{\mathcal{B}} f(x) = -\infty$. Наконец, функция f *стремится к* ∞ по базе $\mathcal{B}$, если $|f|$ стремится к $+\infty$ по базе $\mathcal{B}$. Такие функции называют *бесконечно большими* по базе $\mathcal{B}$. Подчеркнём, что, как и в случае последовательностей, несмотря на использование записи вида $\lim\limits_{\mathcal{B}} f = \infty$ бесконечно большие функции не называют сходящимися к бесконечности (тем не менее, иногда говорят, что они имеют бесконечный предел). Мы будем считать, что утверждение «функция имеет предел» означает стремление функции к конечному числу.

Непрерывные функции

Определение и локальные свойства

Функция f непрерывна в точке a по множеству D, если f определена на D, $a \in D$ и для всякого $\varepsilon > 0$ найдётся такое число $\delta > 0$, что для всех $x \in D$ из неравенства $|x-a| < \delta$ следует неравенство $|f(x) - f(a)| < \varepsilon$, т. е.

$$\forall\, \varepsilon > 0\; \exists\, \delta > 0\; \forall\, x \in D \quad |x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Напомним, что записи $|x-a| < \delta$ и $x \in U_\delta(a)$ равнозначны.

Предложение 4. *Следующие утверждения равносильны.*

(i) *Функция f непрерывна в точке a по множеству D.*

(ii) *Для всякой последовательности $x_n \in D$, сходящейся к точке a, последовательность $\{f(x_n)\}$ сходится к $f(a)$.*

(iii) *Точка a — изолированная точка множества D или a — предельная точка множества D и $\lim\limits_{x \to a} f(x) = f(a)$.*

Следствие 11. *Если функции f и g непрерывны в точке a по множеству D, то функции $f \pm g$ и $f \cdot g$ непрерывны в точке a по множеству D. Если $g(x) \neq 0$ при $x \in D$, то f/g непрерывна в точке a по множеству D.*

$\triangleright$ Если $x_n \to a$ при $n \to \infty$, то по условию имеем $f(x_n) \to f(a)$ и $g(x_n) \to g(a)$. По арифметике предела последовательности получаем $f(x_n) \pm g(x_n) \to f(a) \pm g(a)$, $f(x_n)g(x_n) \to f(a)g(a)$, и если $g(x) \neq 0$ при $x \in D$, то $f(x_n)/g(x_n) \to f(a)/g(a)$. $\triangleleft$

Пусть X — непустое подмножество множества $\mathbb{R}$.

Следствие 12. *Если функции $f\colon D \to X$ и $g\colon X \to \mathbb{R}$ непрерывны в точке $a \in D$ по множеству D и в точке $f(a)$ по множеству X соответственно, то функция $g(f(x))$ непрерывна в точке a по множеству D.*

$\triangleright$ Пусть $x_n \in D$ и $x_n \to a$ при $n \to \infty$. Тогда $f(x_n) \to f(a)$ и $g(f(x_n)) \to g(f(a))$. $\triangleleft$

Следствие 13 (локальная ограниченность). *Если функция f непрерывна в точке a по множеству D, то найдутся такие окрестность $U_\delta(a)$ и число $C > 0$, что для всех $x \in U_\delta(a) \cap D$ верно неравенство $|f(x)| \leqslant C$.*

$\triangleright$ Положим $\varepsilon = 1$. Тогда по определению непрерывности найдётся такое $\delta > 0$, что для всех $x \in D$ из неравенства $|x-a| < \delta$ следует неравенство $|f(x) - f(a)| < 1$. Следовательно, для всех $x \in U_\delta(a) \cap D$ получаем $|f(x)| \leqslant |f(x) - f(a)| + |f(a)| \leqslant C$, где $C = |f(a)| + 1$. $\triangleleft$

Следствие 14 (об отделимости). *Если функция f непрерывна в точке a и $f(a) > 0$, то найдётся такая окрестность $U_\delta(a)$, что для всех $x \in U_\delta(a)$ верно неравенство $f(x) > \frac{f(a)}{2}$.*

Точки разрыва

Точки разрыва функции — это точки области определения, в которых функция не является непрерывной. Если в точке a функция f не является непрерывной по множеству D, то a — предельная точка D, в которой либо предел $\lim\limits_{x \to a} f(x)$ существует, но не равен $f(a)$, либо предела $\lim\limits_{x \to a} f(x)$ не существует. В первом случае говорят, что a является *точкой устранимого разрыва* функции f. Рассмотрим второй случай. Если точка является предельной для множеств $D \cap (-\infty; a)$ и $D \cap (a; +\infty)$, то можно говорить об односторонних пределах функции f. Если односторонние пределы функции f в точке a существуют, но различны, то говорят, что a является *точкой разрыва первого рода*. Если хотя бы одного одностороннего предела не существует (в частности, если функция стремится к бесконечности при стремлении x к точке a слева или справа), то a называют *точкой разрыва второго рода*.

Предложение 5. *Монотонная на интервале функция может иметь на этом интервале разрывы только первого рода, причём множество этих точек разрыва не более чем счётно.*

$\triangleright$ Пусть f — неубывающая функция на интервале $(a;b)$. Рассмотрим произвольную точку $x_0 \in (a;b)$. По теореме Вейерштрасса

$$\lim_{x\to x_0-0} f(x) = \sup_{x<x_0} f(x) = A \leqslant f(x_0) \leqslant B = \inf_{x>x_0} f(x) = \lim_{x\to x_0+0} f(x).$$

Если x_0 — точка разрыва функции f, то хотя бы одно из неравенств $A \leqslant f(x_0)$ и $f(x_0) \leqslant B$ строгое и $A < B$. Таким образом, x_0 — точка разрыва первого рода. Поставим точке x_0 в соответствие интервал $(A;B)$. Получаем, что разным точкам разрыва сопоставлены попарно непересекающиеся интервалы. На числовой прямой можно расположить не более чем счётное множество непустых попарно непересекающихся интервалов, поэтому множество точек разрыва не более чем счётно. $\triangleleft$

Пусть теперь функция f определена на всей числовой прямой. Если функция f ограниченна на множестве E, то *колебанием* функции f на множестве E называется величина

$$\omega(f,E) = \sup_{x,y\in E} |f(x) - f(y)|.$$

В случае, когда функция f неограниченна на множестве E, то полагаем по определению $\omega(f,E) = +\infty$. Далее всегда считаем, что $+\infty$ больше всякого вещественного числа.

Пусть функция f ограниченна в некоторой окрестности $U_r(a)$ точки a. Функция $\delta \mapsto \omega(f, U_\delta(a))$ неубывающая и неотрицательная на интервале $(0;r)$. По теореме Вейерштрасса существует предел

$$\omega_f(a) = \lim_{\delta\to 0+} \omega(f, U_\delta(a)),$$

который называют *колебанием* функции f в точке a. Если во всякой окрестности точки a функция f неограниченна, то полагаем по определению $\omega_f(a) = +\infty$.

Предложение 6. *Непрерывность функции f в точке a равносильна равенству $\omega_f(a) = 0$.*

$\triangleright$ Пусть функция f непрерывна в точке a. Докажем, что тогда $\omega_f(a) = 0$. Для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что $|f(x) - f(a)| < \varepsilon$ при $x \in U_\delta(a)$, в частности, $\omega(f, U_\delta(a)) \leqslant 2\varepsilon$. Следовательно, $\omega_f(a) \leqslant 2\varepsilon$ и в силу произвольности ε справедливо равенство $\omega_f(a) = 0$.

Пусть теперь $\omega_f(a) = 0$. Покажем, что функция f непрерывна в точке a. По определению $\omega_f(a)$ для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что $\omega(f, U_\delta(a)) < \varepsilon$, в частности, $|f(x) - f(a)| < \varepsilon$ при $x \in U_\delta(a)$. $\triangleleft$

Следовательно, множество точек разрыва функции f можно представить в виде $\bigcup_n \{x\colon \omega_f(x) \geqslant 1/n\}$.

Предложение 7. *Для любого натурального числа n множество $\{x\colon \omega_f(x) \geqslant 1/n\}$ замкнуто.*

$\triangleright$ Покажем, что дополнение к этому множеству открыто. Предположим, что $\omega_f(x_0) < 1/n$. По определению предела найдётся такая окрестность $U_\delta(x_0)$, что $\omega(f, U_\delta(x_0)) < 1/n$. Тогда для всякого $x \in U_\delta(x_0)$ выполняется неравенство $\omega_f(x) < 1/n$. $\triangleleft$

Следовательно, множество точек разрыва функции f является объединением не более чем счётного набора замкнутых множеств. Более того, для всякого множества, которое является объединением не более чем счётного набора замкнутых множеств, можно предъявить функцию с таким множеством точек разрыва. Тем самым, далеко не каждое множество может быть множеством точек разрыва некоторой функции, равно как и не каждое множество может быть множеством точек непрерывности. Так, существует функция, разрывная в каждой точке числовой прямой: функция Дирихле $D(x) = I_{\mathbb{Q}}$, равная 1, если число x рационально и равная 0 в противном случае. Существует функция, у которой множество точек разрыва — все рациональные числа: функция Римана $R(x) = 1/n$, если $x = m/n$, где m/n — несократимая дробь, $R(0) = 1$, и $R(x) = 0$ во всех иррациональных точках. Однако не существует функции, у которой множество точек разрыва — все иррациональные точки числовой прямой.

Глобальные свойства непрерывных функций

Функция непрерывна на множестве D, если она непрерывна в каждой точке множества D по множеству D.

Теорема 42 (Вейерштрасс). *Если функция f непрерывна на компакте K, то она ограниченна на K, т. е. существует такое число $C > 0$, что $|f(x)| \leqslant C$ для всех $x \in K$.*

$\triangleright$ *Первый способ.* Предположим, что функция f не является ограниченной. Тогда для всякого натурального числа n существует такое $x_n \in K$, что $|f(x_n)| \geqslant n$. Поскольку K — компакт, существует подпоследовательность $\{x_{n_k}\}$, сходящаяся к некоторому $x_0 \in K$. В силу непрерывности функции f последовательность $f(x_{n_k})$ сходится к $f(x_0)$, но по построению $|f(x_{n_k})| \geqslant n_k \to \infty$. Противоречие.

Второй способ. Поскольку функция f непрерывна в каждой точке a компакта K, существуют такие окрестность $U(a)$ и число $C(a) > 0$, что $|f(x)| \leqslant C(a)$ при $x \in U(a) \cap K$. Окрестности $\{U(a)\}$ покрывают K. По определению компакта можно выбрать конечное подпокрытие $U(a_1), \ldots, U(a_N)$ с соответствующими числами $C_1, \ldots, C_N$. Положив $C = \max\{C_1, C_2, \ldots, C_N\}$, получаем $|f(x)| \leqslant C$ для всех $x \in K$. $\triangleleft$

Теорема 43 (Вейерштрасс). *Если функция f непрерывна на компакте K, то существуют такие $x_m, x_M \in K$, что*

$$f(x_m) = \inf_{x\in K} f(x), \quad f(x_M) = \sup_{x\in K} f(x).$$

$\triangleright$ Рассмотрим только случай точной нижней грани. Положим $A = \inf\limits_{x\in K} f(x)$.

Первый способ. Для всякого натурального числа n существует такое число $x_n \in K$, что $A \leqslant f(x_n) \leqslant A + 1/n$. Значит, $f(x_n) \to A$ при $n \to \infty$. Пользуясь компактностью K, выберем подпоследовательность x_{n_k}, сходящуюся к некоторому x_0. Функция f непрерывна, поэтому $f(x_{n_k}) \to f(x_0)$ и, значит, $f(x_0) = A$.

Второй способ. Предположим противное: пусть $A < f(x)$ для всякого $x \in K$. Тогда функция $g(x) = 1/(f(x) - A)$ непрерывна на компакте K. По предыдущей теореме эта функция ограниченна, т. е. $g(x) \leqslant C$ для некоторой положительной константы C. Следовательно, $f(x) \geqslant A + 1/C$ для всех $x \in K$, что противоречит определению точной нижней грани. $\triangleleft$

Теорема 44 (Коши). *Пусть функция f непрерывна на отрезке $[a; b]$.*

(i) *Если $f(a)f(b) < 0$, то $f(c) = 0$ для некоторого $c \in [a; b]$.*

(ii) *Если $A = f(a)$, $B = f(b)$ и $A \leqslant C \leqslant B$, то $f(c) = C$ для некоторого $c \in [a; b]$.*

$\triangleright$ Докажем сначала утверждение (i).

Первый способ. Разделим отрезок $[a; b]$ пополам. Если оказалось так, что $f((a + b)/2) = 0$, то положим $c = (a + b)/2$. Если же значение $f((a + b)/2)$ отлично от нуля, то f принимает значения разных знаков на концах или отрезка $[a; (a+b)/2]$, или отрезка $[(a+b)/2; b]$. Пусть $[a_1; b_1]$ — тот из этих отрезков, на концах которого f принимает значения разных знаков. Повторим рассуждения с отрезком $[a_1; b_1]$ вместо $[a; b]$ и т. д. Возможны две ситуации: или на некотором шаге построения середина очередного отрезка является искомым числом c, или построена система вложенных отрезков $[a_n; b_n]$, длины которых стремятся к нулю и для которых выполняется неравенство $f(a_n)f(b_n) < 0$. Пусть c — общая точка отрезков $[a_n; b_n]$. Тогда $a_n \to c$ и $b_n \to c$ при $n \to \infty$. Функция f непрерывна, поэтому $f(a_n)f(b_n) \to f^2(c)$. Получаем $f^2(c) \leqslant 0$ и, следовательно, $f(c) = 0$.

Второй способ. Предположим, что такой точки c нет. Тогда множества

$$F_- = \{x\colon\ f(x) \leqslant 0\} \quad \text{и} \quad F_+ = \{x\colon\ f(x) \geqslant 0\}$$

замкнуты, не пусты и не пересекаются, а в объединении дают весь отрезок $[a; b]$. Это невозможно. Противоречие.

Теперь докажем утверждение (ii). Если $A < C < B$, то для функции $g(x) = f(x) - C$ выполняется условие $g(a)g(b) < 0$ и по пункту (i) найдётся $c \in [a; b]$, для которого $g(c) = f(c) - C = 0$. $\triangleleft$

Промежутком называется всякое подмножество числовой прямой, которое вместе с любыми двумя своими точками содержит отрезок, их соединяющий. Можно показать, что промежутками являются только пустое множество, одноточечные множества, интервалы, полуинтервалы, отрезки, лучи и вся числовая прямая.

Следствие 15. *Если функция $f : I \to \mathbb{R}$ непрерывна на непустом промежутке I, то $f(I)$ — промежуток.*

$\triangleright$ Если f принимает на I одно и то же значение, то $f(I)$ — одноточечное множество. Пусть теперь в $f(I)$ есть различные элементы y_1 и y_2. Тогда найдутся такие x_1 и x_2 из I, что $y_1 = f(x_1)$,

$y_2 = f(x_2)$ и отрезок с концами x_1 и x_2 лежит в I. По теореме о промежуточном значении для всякого значения y, лежащего между y_1 и y_2, найдётся такая точка x из отрезка с концами x_1 и x_2, что $f(x) = y$. Таким образом, отрезок с концами y_1 и y_2 лежит в $f(I)$. ◁

Отметим, что обратное утверждение, вообще говоря, неверно. Примером является функция $f(x) = \sin 1/x$ при $x \neq 0$ и $f(0) = 0$.

Теорема 45. *Если функция f монотонна на промежутке I и $f(I)$ — промежуток, то функция f непрерывна.*

▷ Предположим, что функция f неубывающая и в некоторой точке $x_0 \in I$ разрывна. Пусть x_0 — внутренняя точка промежутка I. Случай, когда x_0 — граничная точка I, рассматривается аналогично. Имеем

$$\lim_{x\to x_0-0} f(x) = \sup_{x<x_0} f(x) = A \leqslant f(x_0) \leqslant B = \inf_{x>x_0} f(x) = \lim_{x\to x_0+0} f(x).$$

Поскольку x_0 — точка разрыва, имеем $A < f(x_0)$ или $f(x_0) < B$. Предположим, что $A < f(x_0)$ (случай $f(x_0) < B$ разбирается аналогично). Заметим, что $f(x) \leqslant A$ при $x < x_0$ и $f(x) \geqslant f(x_0)$ при $x \geqslant x_0$. Следовательно, $(A; f(x_0)) \cap f(I) = \varnothing$. Пусть теперь $x_1 < x_0$. Тогда получаем включения $[f(x_1); f(x_0)] \subset f(I)$ и $(A; f(x_0)) \subset [f(x_1); f(x_0)]$. Противоречие. ◁

Теорема 46 (об обратной функции)**.** *Пусть f — строго монотонная и непрерывная функция на промежутке I. Тогда*

(i) *множество $J = f(I)$ является промежутком;*

(ii) *$f\colon I \to J$ является биекцией;*

(iii) *f^{-1} является непрерывной строго монотонной функцией.*

▷ Первые два пункта следуют из строгой монотонности и непрерывности функции f. Предположим теперь, что f возрастает. Пусть $y_1 < y_2$ и $y_1 = f(x_1)$, $y_2 = f(x_2)$. Если $x_1 \geqslant x_2$, то $f(x_1) \geqslant f(x_2)$, а это противоречит неравенству $y_1 < y_2$. Следовательно, $x_1 < x_2$ и f^{-1} возрастает. Аналогично можно показать, что если f убывает, то f^{-1} убывает. Таким образом, f^{-1} строго монотонна и по построению f^{-1} отображает промежуток J на промежуток I. По доказанному выше функция f^{-1} непрерывна. ◁

Применяя теорему об обратной функции к $f(x) = x^{2n-1}$, где $n \in \mathbb{N}$, получаем существование возрастающей непрерывной функции $f^{-1}(y) = y^{\frac{1}{2n-1}}$. Аналогичным образом обосновывается существование непрерывной строго возрастающей функции $y^{\frac{1}{2n}}$ на $[0; +\infty)$, которая является обратной к $f(x) = x^{2n}$ на $[0; +\infty)$.

Построение показательной и логарифмической функций

Пусть $a > 1$. Будем предполагать, что нам известно определение a^x для рациональных x и уже известны следующие свойства:

$$a^0 = 1, \quad a^{x+y} = a^x \cdot a^y, \quad x < y \Longleftrightarrow a^x < a^y$$

для всех рациональных x, y.

Нам потребуются два вспомогательных утверждения.

Предложение 8. *Пусть функция f определена на множестве $\mathbb{Q} \cap [a; b]$ и существует такое число $M > 0$, что для всех $x, y \in \mathbb{Q} \cap [a; b]$ верно неравенство*

$$|f(x) - f(y)| \leqslant M|x - y|.$$

Тогда существует единственная непрерывная функция $\widetilde{f}$, для которой справедливо равенство $f = \widetilde{f}$ на $\mathbb{Q} \cap [a; b]$.

▷ Пусть последовательность рациональных чисел r_n сходится к некоторому $x \in [a; b]$. Для всякого $\varepsilon > 0$ найдётся такой номер N, что для всех $n, m > N$ выполнено неравенство $|r_n - r_m| < \varepsilon$. Следовательно, для этих номеров n, m имеем $|f(r_n) - f(r_m)| < M\varepsilon$ и, значит, последовательность $f(r_n)$ фундаментальна и сходится. Тогда для двух последовательностей рациональных чисел $r_n \to x$ и $s_n \to x$ при $n \to \infty$ получаем $|f(r_n) - f(s_n)| \leqslant M|r_n - s_n| \to 0$, т. е. предел последовательности $f(r_n)$ совпадает с пределом последовательности $f(s_n)$. Положим $\widetilde{f}(x) = \lim\limits_{n\to\infty} f(r_n)$. Тогда $\widetilde{f} = f$ на $\mathbb{Q} \cap [a; b]$ и $|\widetilde{f}(x) - \widetilde{f}(y)| \leqslant M|x - y|$ для всех $x, y \in [a; b]$, что влечёт непрерывность $\widetilde{f}$.

Докажем единственность. Предположим, что две непрерывные функции $\widetilde{f}_1$ и $\widetilde{f}_2$ совпадают на $\mathbb{Q} \cap [a;b]$. Для произвольного $x \in [a;b]$ найдём какую-нибудь последовательность рациональных чисел r_n, сходящуюся к x. В силу непрерывности получаем

$$\widetilde{f}_1(x) = \lim_{n\to\infty} \widetilde{f}_1(r_n) = \lim_{n\to\infty} \widetilde{f}_2(r_n) = \widetilde{f}_2(x).$$

Таким образом, функции $\widetilde{f}_1$ и $\widetilde{f}_2$ совпадают. $\lhd$

Предложение 9. *Пусть $N \in \mathbb{N}$. Существует такое положительное число $C(N)$, что для всех $x, y \in \mathbb{Q} \cap [-N; N]$ выполнено неравенство*

$$|a^x - a^y| \leqslant C(N)|x - y|.$$

$\rhd$ Будем предполагать, что $x < y$. Если $y - x \geqslant 1$, то

$$a^y - a^x \leqslant a^y \leqslant a^N \leqslant a^N(y - x).$$

Пусть теперь $y - x < 1$. Найдётся такое натуральное число n, что $\frac{1}{n+1} \leqslant y - x \leqslant 1/n$. Имеем

$$a^y - a^x = a^x(a^{y-x} - 1) \leqslant a^N(a^{\frac{1}{n}} - 1) \leqslant a^N(a-1)\frac{1}{n} \leqslant 2a^N(a-1)\frac{1}{n+1} \leqslant 2a^N(a-1)(y-x).$$

Здесь мы использовали неравенство $a^{\frac{1}{n}} - 1 \leqslant (a-1)\frac{1}{n}$, которое является следствием неравенства Бернулли. Искомое число $C(N)$ равно максимуму из чисел a^N и $2(a-1)a^N$. $\lhd$

Пусть $E \subset \mathbb{R}$ и определена функция $f : E \to \mathbb{R}$. Если существует такое число $C > 0$, что для всех $x, y \in E$ верно неравенство

$$|f(x) - f(y)| \leqslant C|x - y|,$$

то говорят, что функция f удовлетворяет на множестве E *условию Липшица* с константой C.

Применим доказанные утверждения. На всяком отрезке $[-N; N]$ существует единственная непрерывная функция g_N, совпадающая с a^x в рациональных точках. Если $M > N$, то на отрезке $[-N; N]$ появляются две функции g_M и g_N, которые совпадают в силу единственности. Для всякого x положим $g(x) = g_N(x)$, где $[-N; N]$ — произвольный отрезок такого вида, содержащий точку x. Непрерывная функция g определена на $\mathbb{R}$ и совпадает с a^x на $\mathbb{Q}$. Используя арифметику предела, можно показать, что g является строго возрастающей функцией и $g(x + y) = g(x) \cdot g(y)$. Остаётся положить по определению $a^x = g(x)$ для всех вещественных чисел x.

Таким образом, мы построили возрастающую и непрерывную на $\mathbb{R}$ функцию a^x. По теореме об обратной функции существует обратная к a^x строго возрастающая и непрерывная функция, которую называют логарифмической и обозначают $\log_a x$. При $0 < a < 1$ положим

$$a^x = \left(\frac{1}{a}\right)^{-x}$$

и получим убывающую и непрерывную на $\mathbb{R}$ функцию, которая также имеет обратную, обозначаемую через $\log_a x$.

Если функция f положительна и задана функция g, то по определению полагают

$$f(x)^{g(x)} = e^{g(x)\ln f(x)}.$$

Второй замечательный предел:

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

$\rhd$ Пусть $x > 1$. Существует натуральное число n, для которого $n \leqslant x < n + 1$. Тогда

$$\left(1 + \frac{1}{n+1}\right)^n \leqslant \left(1 + \frac{1}{x}\right)^x \leqslant \left(1 + \frac{1}{n}\right)^{n+1}.$$

В силу определения числа e для всякого $\varepsilon > 0$ найдётся такое N, что для всех $n > N$ правая часть меньше $e + \varepsilon$, а левая часть больше $e - \varepsilon$. Следовательно, для всех $x > N + 1$ величина $\left(1 + \frac{1}{x}\right)^x$ отличается от e меньше чем на ε, т. е. выполняется определение предела по Коши. $\lhd$

Равномерная непрерывность и равномерная сходимость

Функция f *равномерно непрерывна* на множестве D, если для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для всех $x, y \in D$ из неравенства $|x - y| < \delta$ следует неравенство $|f(x) - f(y)| < \varepsilon$.

Теорема 47 (Кантор)**.** *Если функция f непрерывна на компакте K, то f равномерно непрерывна на K.*

$\triangleright$ *Первый способ.* Предположим противное: существует такое $\varepsilon > 0$, что для всякого $n \in \mathbb{N}$ найдутся $x_n, y_n \in K$, удовлетворяющие неравенствам $|x_n - y_n| < 1/n$ и $|f(x_n) - f(y_n)| \geqslant \varepsilon$. Выберем в $\{x_n\}$ сходящуюся к некоторому x_0 подпоследовательность $\{x_{n_k}\}$. Последовательность $\{y_{n_k}\}$ тоже сходится к x_0. В силу непрерывности функции f разность $f(x_{n_k}) - f(y_{n_k})$ стремится к нулю. Противоречие.

Второй способ. Пусть $\varepsilon > 0$. В силу непрерывности f для всякой точки $a \in K$ существует такая окрестность $U_\delta(a)$, что $\omega(f, U_\delta(a)) < \varepsilon$, причём δ зависит от a. Окрестности $U_{\delta/3}(a)$ покрывают K и из них можно выбрать конечное подпокрытие: $U_{\delta_1/3}(a_1), \ldots, U_{\delta_N/3}(a_N)$. Положим $\delta = \frac{1}{3}\min\{\delta_1, \delta_2, \ldots, \delta_N\}$. Если $|x - y| < \delta$ и $x \in U_{\delta_k/3}(a_k)$, то $x, y \in U_{\delta_k}(a_k)$ и $|f(x) - f(y)| < \varepsilon$. $\triangleleft$

Запишем определение равномерной непрерывности в терминах колебания $\omega(f, U_\delta(x))$. Функция f равномерно непрерывна на D, если для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для всякого $x \in D$ верно неравенство $\omega(f, U_\delta(x)) < \varepsilon$. В силу монотонности по δ функции $\omega(f, U_\delta(x))$ последнее неравенство выполняется для всех $\delta' \in (0; \delta]$. Таким образом, функция f равномерно непрерывна на D, если $\lim\limits_{\delta \to 0} \sup_{x \in D} \omega(f, U_\delta(x)) = 0$. Непрерывность функции f на множестве D равносильна равенству $\lim\limits_{\delta \to 0} \omega(f, U_\delta(x)) = 0$ для всякого $x \in D$. Таким образом, равномерная непрерывность означает, что колебание функции f на $U_\delta(x)$ приближается к нулю в определённом смысле с одной скоростью сразу во всех точках $x \in D$ (говорят, что сходится к нулю равномерно), а обычная непрерывность означает, что колебание функции $\omega(f, U_\delta(x))$ приближается к нулю в каждой точке, но, возможно, с разными скоростями в разных точках (говорят, что сходится поточечно).

Рассмотрим пример: $f(x) = \sin(1/x)$ на $D = (0; 1)$. В каждой точке интервала $(0; 1)$ данная функция непрерывна, но она не является равномерно непрерывной на $(0; 1)$. Действительно, поскольку сколь угодно близко к нулю функция принимает значения ± 1, справедливо равенство $\sup_{x \in (0;1)} \omega(f, U_\delta(x)) = 2$ для всякого $\delta > 0$.

Равномерная сходимость, с который мы столкнулись при обсуждении равномерной непрерывности, в завуалированном виде встречалась нам в теории предела функции. Рассмотрим в качестве примера следующую задачу. Пусть непрерывная функция $f\colon [1; +\infty) \to \mathbb{R}$ такова, что для всякого $x \geqslant 0$ числовая последовательность $f(x), f(x+1), \ldots, f(x+n), \ldots$ сходится к нулю при $n \to \infty$. Верно ли, что $\lim\limits_{x \to +\infty} f(x) = 0$? Ответ отрицателен.

Действительно, будем строить функцию f следующим образом. На отрезке $[n; n + 1/n]$ график функции f представляет собой боковые стороны равнобедренного треугольника с основанием $[n; n + 1/n]$ и высотой 1, а на отрезке $[n + 1/n; n + 1]$ функция f равна нулю. Пусть $x \geqslant 1$. Если дробная часть x не меньше $1/N$, то $f(x + n) = 0$ для всех $n > N$. Однако предел функции f при $x \to +\infty$ не равен нулю.

Интересно, что из сходимости последовательности $f(nx)$ к нулю для всякого $x \geqslant 1$ и в самом деле следует, что $\lim\limits_{x \to +\infty} f(x) = 0$.

Две рассмотренные выше задачи можно переформулировать следующим образом. Задана последовательность функций $f_n(x) = f(x + n)$ (или $f_n(x) = f(nx)$). Известно, что $f_n(x) \to 0$ для всякого $x \geqslant 1$ (т. е. последовательность функций сходится поточечно к нулю). Верно ли, что $\sup_{x \geqslant 1} |f_n(x)| \to 0$ (т. е. последовательность функций сходится равномерно к нулю)?

Последовательность функций $\{f_n\}$ сходится *поточечно* на множестве D к функции f, если в каждой точке x множества D числовая последовательность $\{f_n(x)\}$ сходится к значению $f(x)$. Можно построить пример, показывающий, что поточечная сходимость не сохраняет непрерывность. Однако непрерывность не теряется совсем, множество точек непрерывности предельной функции всюду плотно. Поэтому, например, функцию Дирихле нельзя получить как поточечный предел непрерывных функций.

Покажем, что поточечный предел последовательности непрерывных функций имеет точку непрерывности.

Теорема 48. *Пусть* $\{f_n\}$ *— последовательность непрерывных на* $\mathbb{R}$ *функций. Предположим, что для каждого* $x \in \mathbb{R}$ *последовательность* $\{f_n(x)\}$ *сходится к* $f(x)$. *Тогда функция* f *имеет хотя бы одну точку непрерывности.*

$\triangleright$ Предположим, что у функции f нет точек непрерывности. Тогда вся числовая прямая является объединением замкнутых множеств $\{x\colon \omega_f(x) \geqslant 1/k\}$. Следовательно, хотя бы одно из этих множеств по теореме Бэра содержит отрезок $[\alpha;\beta]$. Итак, пусть $\omega_f(x) \geqslant 1/k$ при $x \in [\alpha;\beta]$. Пусть $\varepsilon > 0$. Рассмотрим набор зам-кнутых множеств

$$F_N = \{x \in [\alpha;\beta]\colon |f_n(x) - f_m(x)| \leqslant \varepsilon \,\text{для всех}\, n, m > N\}.$$

Тогда получаем $[\alpha;\beta] = \cup_N F_N$. Следовательно, по теореме Бэра существует отрезок $[\gamma;\delta] \subset F_N$, $\gamma < \delta$. Заметим, что для всех точек x отрезка $[\gamma;\delta]$ имеем $|f_n(x) - f(x)| \leqslant \varepsilon$. Все функции f_n непрерывны, поэтому при малых ε это неравенство противоречит условию $\omega_f(x) \geqslant 1/k$. $\triangleleft$

Поскольку доказательство теоремы фактически без изменений можно повторить с заменой множества $\mathbb{R}$ на всякий отрезок, функция f имеет всюду плотное в $\mathbb{R}$ множество точек непрерывности.

Обсудим более детально определение поточечной сходимости. Последовательность функций $\{f_n\}$ сходится поточечно на множестве D к функции f, если для каждой точки $x \in D$ и всякого $\varepsilon > 0$ существует такой номер N, что при всех $n > N$ верно неравенство $|f_n(x) - f(x)| < \varepsilon$. Заметим, что номер N не обязан быть одним и тем же для разных значений x. Если же для всякого $\varepsilon > 0$ существует такой номер N, что при всех $n > N$ неравенство $|f_n(x) - f(x)| < \varepsilon$ выполняется *для всех* $x \in D$, то говорят, что последовательность функций $\{f_n\}$ сходится к функции f *равномерно* на множестве D. Иными словами, равномерная сходимость последовательности $\{f_n\}$ к функции f на множестве D означает, что числовая последовательность $\{\sup\limits_D |f_n(x) - f(x)|\}$ сходится к нулю:

$$\lim_{n\to\infty} \sup_D |f_n(x) - f(x)| = 0.$$

Равномерная сходимость влечёт поточечную (обратное неверно) и сохраняет непрерывность.

Теорема 49. *Если последовательность непрерывных функций сходится равномерно, то её предел является непрерывной функцией.*

$\triangleright$ Пусть последовательность функций $\{f_n\}$ сходится равномерно на D к функции f. Утверждение следует из неравенств

$$|f(x) - f(a)| \leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \leqslant \\ \leqslant |f_n(x) - f_n(a)| + 2\sup_D |f - f_n|. \quad \triangleleft$$

Функциональный ряд $\sum\limits_{n=1}^{\infty} f_n$ *сходится поточечно* на множестве D, если для всякого $x \in D$ соответствующий числовой ряд $\sum\limits_{n=1}^{\infty} f_n(x)$ сходится. Функциональный ряд $\sum\limits_{n=1}^{\infty} f_n$ называется *равномерно сходящимся на множестве* D, если последовательность $\{S_n\}$ его частичных сумм равномерно сходится на D. Из определения и доказанной выше теоремы следует, что равномерно сходящийся на D функциональный ряд из непрерывных на D функций сходится к непрерывной функции.

Для доказательства равномерной сходимости ряда часто используют следующее простое наблюдение.

Теорема 50 (признак Вейерштрасса)**.** *Пусть* $\{f_n\}$ *— последовательность ограниченных функций на* D. *Если существует такой сходящийся числовой ряд* $\sum\limits_{n=1}^{\infty} a_n$, *что* $|f_n(x)| \leqslant a_n$ *для всех* $x \in D$ *и* $n \in \mathbb{N}$ *(мажорантный ряд), то ряд* $\sum\limits_{n=1}^{\infty} f_n$ *сходится равномерно на* D.

$\triangleright$ При каждом $x \in D$ числовой ряд $\sum\limits_{n=1}^{\infty} f_n(x)$ сходится абсолютно, поскольку $|f_n(x)| \leqslant a_n$ и ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится. Положим $f(x) = \sum\limits_{n=1}^{\infty} f_n(x)$. Имеем

$$\Big|f(x) - \sum_{n=1}^{N} f_n(x)\Big| \leqslant \sum_{n=N+1} |f_n(x)| \leqslant \sum_{n=N+1} a_n.$$

Следовательно,

$$\sup_{x\in D}\Big|f(x) - \sum_{n=1}^{N} f_n(x)\Big| \leqslant \sum_{n=N+1} a_n \to 0, \quad N \to \infty. \qquad \triangleleft$$

Наконец, обсудим вопрос о равномерном приближении на отрезке непрерывной функции многочленом.

Теорема 51 (теорема Вейерштрасса). *Пусть функция f непрерывна на отрезке $[a;b]$. Тогда для всякого $\varepsilon > 0$ существует такой многочлен P, что $\max\limits_{x\in[a;b]} |P(x) - f(x)| < \varepsilon$.*

Известно множество доказательств этой теоремы, в том числе конструктивные доказательства, когда многочлен P предъявляется явно (см., например, задачи 10.6 и 10.7). Однако мы рассмотрим другое доказательство[1], основанное исключительно на методах и идеях, которые уже подробно обсуждались в настоящем курсе.

Лемма 5. *Пусть $\alpha < u < v < \beta$. Тогда для всякого $\gamma > 0$ существует такой многочлен $q(x)$, что*

1) $0 \leqslant q(x) \leqslant 1$ *всюду на* $[\alpha;\beta]$;
2) $1-\gamma \leqslant q(x) \leqslant 1$ *при* $x \in [\alpha;u]$;
3) $q(x) \leqslant \gamma$ *при* $x \in [v;\beta]$.

$\triangleright$ Пусть $r(x) = kx + b$, где коэффициенты k и b выбраны таким образом, чтобы функция r возрастала и были верны неравенства $0 < r(x) < 1/2$ при $x \in [\alpha;u]$ и $1/2 < r(x) < 1$ при $x \in [v;\beta]$. Положим

$$q_n(x) = (1 - (r(x))^n)^{2^n}, \quad \text{где } n \in \mathbb{N}.$$

Тогда $0 \leqslant q_n(x) \leqslant 1$ всюду на $[\alpha;\beta]$. Если $x \in [\alpha;u]$, то в силу неравенства Бернулли и монотонности r получаем:

$$q_n(x) = (1 - (r(x))^n)^{2^n} \geqslant 1 - (2r(x))^n \geqslant 1 - (2r(u))^n.$$

Поскольку $2r(u) < 1$, последовательность многочленов $\{q_n(x)\}$ сходится равномерно к 1 на отрезке $[\alpha;u]$. Пусть теперь $x \in [v;\beta]$. В силу неравенства Бернулли имеем $1 + (2r(x))^n \leqslant (1 + (r(x))^n)^{2^n}$, поэтому верна следующая цепочка неравенств:

$$q_n(x) \leqslant (2r(x))^{-n}(1 + (2r(x))^n)(1 - (r(x))^n)^{2^n} \leqslant (2r(x))^{-n}(1 + (r(x))^n)^{2^n}(1 - (r(x))^n)^{2^n}.$$

Следовательно, $q_n(x) \leqslant (2r(x))^{-n}(1 - (r(x))^{2n})^{2^n} \leqslant (2r(v))^{-n}$. Поскольку $2r(v) > 1$, последовательность многочленов $\{q_n(x)\}$ сходится равномерно к нулю на $[v;\beta]$. Таким образом, для всякого $\gamma > 0$ существует такой номер n, что многочлен q_n удовлетворяет условиям 1), 2) и 3). $\triangleleft$

Теперь перейдём к доказательству теоремы Вейерштрасса.

$\triangleright$ Пусть функция f непрерывна на $[a;b]$ и положим

$$M = \max_{x\in[a;b]} |f|.$$

Доопределим f при $x > b$ значением $f(b)$ до непрерывной на луче $[a;+\infty)$ функции. Пусть $\varepsilon > 0$. Через E обозначим множество таких $x > a$, что найдётся многочлен P, отличающийся на отрезке $[a;x]$ от функции f менее чем на ε, т. е.

$$\max_{t\in[a;x]} |P(t) - f(t)| < \varepsilon.$$

Функция f непрерывна в точке a, поэтому существует такое значение $x > a$, что при $t \in [a;x]$ имеем $|f(a) - f(t)| < \varepsilon$, т. е. на отрезке $[a;x]$ в качестве многочлена P можно взять константу $f(a)$.

[1] Rudolf Vyborny. The Weierstrass theorem on polynomial approximation // Mathematica Bohemica, vol. 130, no. 2, pp. 161–166 (2005).

Следовательно, E непусто. Если множество E не ограничено сверху, то теорема уже доказана: достаточно рассмотреть отрезок $[a;x]$, где $x \in E$ и $x > b$. Пусть теперь множество E ограничено сверху и $c = \sup E$. Покажем, что это невозможно. Пусть $\gamma > 0$. Найдём такое число $\delta > 0$, что $|f(x) - f(c)| < \gamma$ при $x \in [c - 2\delta, c + 2\delta]$. Существует такой многочлен P, что

$$\max_{[a;c-\delta]} |P(t) - f(t)| < \varepsilon.$$

Пусть q — многочлен из леммы и $\alpha = a$, $u = c - 2\delta$, $v = c - \delta$ и $\beta = c + \delta$. Рассмотрим многочлен $Q(x) = (1 - q(x))f(c) + q(x)P(x)$. Заметим, что

$$|Q(x) - f(x)| \leqslant (1 - q(x))|f(c) - f(x)| + q(x)|P(x) - f(x)|.$$

Если $x \in [a; c - 2\delta]$, то

$$|Q(x) - f(x)| \leqslant 2M\gamma + \max_{[a;c-\delta]} |P(t) - f(t)|.$$

Если $x \in [c - 2\delta; c - \delta]$, то

$$|Q(x) - f(x)| \leqslant |f(c) - f(x)| + \max_{[a;c-\delta]} |P(t) - f(t)| < \gamma + \max_{[a;c-\delta]} |P(t) - f(t)|.$$

Если $x \in [c - \delta; c + \delta]$, то

$$|Q(x) - f(x)| \leqslant |f(c) - f(x)| + (M + \max_{[c-\delta;c+\delta]} |P|)\gamma \leqslant (1 + M + \max_{[c-\delta;c+\delta]} |P|)\gamma.$$

Выбирая достаточно малое положительное число γ, получаем оценку $|Q(x) - f(x)| < \varepsilon$ для всех $x \in [a; c + \delta]$. Следовательно, $c + \delta \in E$, а это противоречит тому, что $c = \sup E$. Таким образом, множество E не ограничено сверху. $\lhd$

Дифференцируемые функции. Пример Вейерштрасса

Пусть функция f определена в окрестности точки a. Функция f называется *дифференцируемой* в точке a, если существуют такие число A и функция α, что $\lim\limits_{h \to 0} \alpha(h) = 0$ и для всех h из некоторой окрестности нуля верно равенство

$$f(a + h) - f(a) = Ah + \alpha(h)h.$$

Важно иметь ввиду, что A и α зависят от точки a и правильнее писать $A(a)$ и $\alpha(a, h)$.

Функция $l : \mathbb{R} \to \mathbb{R}$ называется *линейной*, если $l(\lambda x) = \lambda l(x)$ при всех $\lambda, x \in \mathbb{R}$ и $l(x_1 + x_2) = l(x_1) + l(x_2)$ при всех $x_1, x_2 \in \mathbb{R}$. Можно показать, что все линейные функции имеют вид $l(x) = cx$, где c — произвольная вещественная постоянная. В частности, в отличие от школьной программы, функция $y = cx + b$ является линейной лишь при $b = 0$.

Линейная функция $h \mapsto Ah$ называется *дифференциалом* функции f в точке a и обозначается через $df(h)$ или $df(a, h)$. Мы покажем, что в фиксированной точке a число A и, соответственно, дифференциал данной функции определены однозначно.

Рассмотрим несколько примеров.

1) Пусть $f(x) = x$. Тогда

$$f(a + h) - f(a) = (a + h) - a = h = 1 \cdot h + 0,$$

поэтому $dx(h) = h$ и $\alpha(a, h) = 0$.

2) Пусть $f(x) = x^2$. Тогда

$$f(a + h) - f(a) = (a + h)^2 - a^2 = 2ah + h^2,$$

поэтому $d(x^2)(a, h) = 2ah$ и $\alpha(a, h) = h$.

3) Пусть $f(x) = x^3$. Тогда

$$f(a + h) - f(a) = (a + h)^3 - a^3 = 3a^2h + (3ah + h^2)h,$$

поэтому $d(x^3)(a, h) = 3a^2h$ и $\alpha(a, h) = 3ah + h^2$.

Предел

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h},$$

если он существует, называется *производной* функции f в точке a и обозначается через $f'(a)$ или $\frac{df(a)}{dx}$.

Предложение 10. *Функция f дифференцируема в точке a тогда и только тогда, когда в этой точке существует производная $f'(a)$, причём в случае дифференцируемости дифференциал имеет вид $df(a,h) = f'(a)h$.*

$\triangleright$ Равенство

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a)$$

равносильно тому, что

$$\frac{f(a+h)-f(a)}{h} - f'(a) = \alpha(h) \quad \text{и} \quad \lim_{h\to 0} \alpha(h) = 0. \qquad \triangleleft$$

Аналитическая и геометрическая интерпретация

Непрерывность функции f в точке a означает, что в малой окрестности точки a значения функции f мало отличаются от $f(a)$, т. е. в некоторой окрестности точки a выполнено равенство $f(x) = f(a) + \alpha(x)$, где $\lim\limits_{x\to a} \alpha(x) = 0$.

Дифференцируемость функции f в точке a означает, что в окрестности этой точки функция f приближается функцией $y = f'(a)(x-a) + f(a)$, т. е. $f(x) = f(a) + f'(a)(x-a) + \alpha(x)(x-a)$, где $\lim\limits_{x\to a} \alpha(x) = 0$. Фактически это и есть главное отличие дифференцируемых функций от всех остальных: локально их график выглядит как часть прямой.

Пусть функция f дифференцируема в точке a. *Касательной* в точке a к графику функции f называется прямая, задаваемая уравнением

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

Сразу отметим, что касательная сколь угодно близко к точке касания может пересекать график функции. Например, график функции $f(x) = x^2 \sin 1/x$ при $x \neq 0$ и $f(0) = 0$ в точке $a = 0$ имеет касательную $y = 0$.

Из определения производной немедленно следует, что секущая $y = \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$ стремится к касательной при $b \to a$ в том смысле, что коэффициент при $(x-a)$ стремится к $f'(a)$. Поэтому говорят, что касательная — предельное положение секущей.

Непрерывность и дифференцируемость

Предложение 11. *Если функция f дифференцируема в точке a, то f непрерывна в точке a.*

$\triangleright$ Действительно, имеем

$$\lim_{x\to a} f(x) = \lim_{h\to 0} f(a+h) = \lim_{h\to 0}(f(a) + f'(a)h + \alpha(a,h)h) = f(a). \qquad \triangleleft$$

Однако непрерывная функция может быть не дифференцируема, например, непрерывная функция $f(x) = |x|$ не дифференцируема в нуле. Более того, существуют непрерывные функции, у которых нет производной ни в одной точке числовой прямой.

Теорема 52 (пример Вейерштрасса)**.** *Функция*

$$f(x) = \sum_{n=1}^{\infty} 2^{-n} \sin(8^n x)$$

непрерывна на $\mathbb{R}$, но нигде не дифференцируема.

$\triangleright$ Поскольку $|2^{-n}\sin(8^n x)| \leqslant 2^{-n}$ и ряд $\sum\limits_{n=1}^{\infty} 2^{-n}$ сходится, по признаку Вейерштрасса исходный функциональный ряд сходится к f равномерно на $\mathbb{R}$, поэтому функция f является всюду непрерывной. Теперь покажем, что f нигде не дифференцируема. Зафиксируем $x \in \mathbb{R}$. Для всякого натурального числа N имеем

$$f\Big(x \pm \frac{\pi}{2\cdot 8^N}\Big) - f(x) = \sum_{n=1}^{N} \frac{1}{2^n}\bigg(\sin 8^n\Big(x \pm \frac{\pi}{2\cdot 8^N}\Big) - \sin\Big(8^n x\Big)\bigg).$$

Отметим, что слагаемые с номерами $n > N$ равны нулю, так как $\sin(8^n x \pm 8^{n-N}\pi/2) = \sin(8^n x)$. Оценим отдельно слагаемое с номером $n = N$. Поскольку

$$\Big|\sin\Big(8^N x \pm \frac{\pi}{2}\Big) - \sin\Big(8^N x\Big)\Big| = \sqrt{2}\,\Big|\cos\Big(8^N x \pm \frac{\pi}{4}\Big)\Big|,$$

выбором $+$ или $-$ перед $\pi/4$ у аргумента косинуса можно добиться выполнения неравенства $\left|\sin\left(8^N x \pm \frac{\pi}{2}\right) - \sin\left(8^N x\right)\right| \geqslant 1$. Применяя неравенство $|\sin x| \leqslant |x|$, оценим остальные слагаемые:
$$\sum_{n=1}^{N-1} \frac{1}{2^n}\left|\sin 8^n\left(x \pm \frac{\pi}{2\cdot 8^N}\right) - \sin\left(8^n x\right)\right| \leqslant \sum_{n=1}^{N-1} \frac{1}{2^{n-1}} \sin\left(\frac{\pi 8^n}{4\cdot 8^N}\right) \leqslant \frac{\pi}{2\cdot 8^N}\sum_{n=1}^{N-1} 4^n \leqslant \frac{\pi}{6\cdot 2^N}.$$
Применяя неравенство $|a+b| \geqslant |a| - |b|$, получаем оценку
$$\left|f\left(x \pm \frac{\pi}{2\cdot 8^N}\right) - f(x)\right| \geqslant \left(1 - \frac{\pi}{6}\right)\frac{1}{2^N}.$$
Таким образом,
$$\left(\frac{\pi}{2\cdot 8^N}\right)^{-1}\left|f\left(x \pm \frac{\pi}{2\cdot 8^N}\right) - f(x)\right| \geqslant \left(1 - \frac{\pi}{6}\right)\frac{2^{2N+1}}{\pi},$$
где правая часть неравенства стремится к $+\infty$ при $N \to \infty$. Функция f не имеет производной в точке x, иначе правая часть последнего неравенства ограниченна, что несовместимо с неограниченностью левой части неравенства. $\lhd$

Важнейшие свойства дифференцируемых функций

1) *Линейность.* Пусть функции f и g дифференцируемы в точке a. Тогда для любых постоянных λ, μ функция $\lambda f + \mu g$ дифференцируема в точке a и $(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$.

2) *Правило Лейбница.* Пусть функции f и g дифференцируемы в точке a. Тогда функция fg дифференцируема в точке a и $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

3) *Производная частного.* Пусть функции f и g дифференцируемы в точке a, причём в некоторой окрестности точки a функция g отлична от нуля. Тогда функция $\frac{f}{g}$ дифференцируема в точке a и $\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{(g(a))^2}$.

4) *Производная сложной функции.* Пусть функция f дифференцируема в точке $g(a)$ и функция g дифференцируема в точке a. Тогда функция $f \circ g$ дифференцируема в точке a и $(f \circ g)'(a) = f'(g(a))g'(a)$.

5) *Производная обратной функции.* Пусть f — непрерывная и строго монотонная функция, отображающая интервал I на интервал J. Предположим, что f дифференцируема в точке $a \in I$ и $f'(a) \neq 0$. Тогда обратная функция f^{-1} дифференцируема в точке $f(a)$ и $(f^{-1})'(a) = 1/f'(f^{-1}(a))$.

Свойства 1, 2, 3 и 5 выводятся напрямую из определения производной, например:
$$\lim_{x\to a} \frac{f(x)g(x) - f(a)g(a)}{x-a} = \lim_{x\to a} g(x) \cdot \lim_{x\to a} \frac{f(x) - f(a)}{x-a} + f(a)\lim_{x\to a} \frac{g(x) - g(a)}{x-a} = g(a)f'(a) + f(a)g'(a).$$

Несколько сложнее обосновывается правило 4.

$\rhd$ Функция f дифференцируема в точке $g(a)$, поэтому
$$f(g(a) + t) - f(g(a)) = f'(g(a))t + \alpha(t)t, \quad \text{где } \lim_{t\to 0}\alpha(t) = 0.$$
Далее считаем, что $\alpha(0) = 0$ и функция α непрерывна в нуле. Полагая $t = g(a+h) - g(a)$, приходим к равенству
$$f(g(a+h)) - f(g(a)) = f'(g(a))(g(a+h) - g(a)) + \alpha(g(a+h) - g(a))(g(a+h) - g(a)).$$
Функция g дифференцируема в точке a, поэтому
$$g(a+h) - g(a) = g'(a)h + \beta(h)h, \quad \text{где } \lim_{h\to 0}\beta(h) = 0.$$
Тогда
$$f(g(a+h)) - f(g(a)) = f'(g(a))g'(a)h + \gamma(h)h,$$
где
$$\gamma(h) = f'(g(a))\beta(h) + \alpha(g(a+h) - g(a))(g'(a) + \beta(h)).$$
Получаем $\lim\limits_{h\to 0}\gamma(h) = 0$. $\lhd$

Поскольку в точке x дифференциал имеет вид $df(h) = f'(x)h$ и $dx(h) = h$, его можно записать в следующем виде: $df = f'(x)\,dx$. Полезно переписать свойства 1—5 в терминах дифференциалов.

1) *Линейность*: $d(\alpha f + \beta g) = \alpha df + \beta dg$.

2) *Правило Лейбница*: $d(fg) = fdg + gdf$.

3) *Дифференциал частного*: $d\left(\frac{f}{g}\right) = \frac{gdf - fdg}{g^2}$.

4) *Дифференциал сложной функции*: $df(g) = df(dg) = f'(g)\,dg$.

Таким образом, равенство $df = f'(x)\,dx$ верно независимо от того, считаем ли мы x функцией или независимой переменной. Это наблюдение называют *инвариантностью первого дифференциала*. Проиллюстрируем инвариантность следующим образом. Пусть задана биективная и дифференцируемая функция $g\colon \mathbb{R}_t \to \mathbb{R}_x$, причём обратная функция тоже дифференцируема (говорят, что задан *диффеоморфизм*). Фактически $x = g(t)$ — замена переменной x на переменную t, переход к новой координате. Если на $\mathbb{R}_x$ была определена дифференцируемая функция $f(x)$, то естественным образом в новой системе координат этой функции соответствует дифференцируемая функция $f(g(t))$ и её дифференциал вычисляется подстановкой в дифференциал функции f вместо x функции $g(t)$, т. е. форма первого дифференциала не меняется при заменах координат.

5) *Дифференциал обратной функции*:

$$df^{-1} = (df)^{-1},$$

где $(df)^{-1}$ — линейная функция, обратная к линейной функции df. Действительно, функция $x = y/k$ является обратной функцией к $y = kx$ ($k \neq 0$). Аналогичным образом, функция $h \mapsto \frac{1}{f'(a)}h$ является обратной к функции $h \mapsto f'(a)h$.

Проиллюстрируем правила дифференцирования на примере экспоненты и логарифма.

Напомним второй замечательный предел:

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

Заменяя x на $-x$, можно показать, что при $x \to -\infty$ предел функции $\left(1+\frac{1}{x}\right)^x$ также равен e. Более того, заменяя x на $1/x$, заключаем, что справедливо равенство

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

Функция $\ln x$ (логарифм по основанию e) непрерывна на своей области определения, поэтому

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = \lim_{x\to 0}\ln\left((1+x)^{1/x}\right) = \ln e = 1.$$

Заменяя $1+x$ на e^x и пользуясь непрерывностью экспоненты, получаем

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

Вычисленные нами пределы равносильны второму замечательному пределу, поэтому их также относят к списку замечательных. Такое название связано с тем, что именно с помощью этих пределов обосновываются формулы дифференцирования всех основных элементарных функций. Например,

$$(e^x)' = \lim_{h\to 0}\frac{e^{x+h}-e^x}{h} = e^x\lim_{h\to 0}\frac{e^h-1}{h} = e^x.$$

По правилу дифференцирования обратной функции находим

$$(\ln x)' = 1/x.$$

Полезно иметь в виду следующую *таблицу производных*:

1) $(\mathrm{const})' = 0$; 2) $(x^\alpha)' = \alpha x^{\alpha-1}$;

3) $(a^x)' = a^x \ln a$; $(e^x)' = e^x$; 4) $(\log_a x)' = \frac{1}{x\ln a}$; $(\ln x)' = \frac{1}{x}$;

5) $(\sin x)' = \cos x$; 6) $(\cos x)' = -\sin x$;

7) $(\operatorname{tg} x)' = 1 + \operatorname{tg}^2 x = \frac{1}{\cos^2 x}$; 8) $(\operatorname{ctg} x)' = -1 - \operatorname{ctg}^2 x = -\frac{1}{\sin^2 x}$;

9) $(\arcsin x)' = -(\arccos x)' = \dfrac{1}{\sqrt{1-x^2}}$;

10) $(\operatorname{arctg} x)' = -(\operatorname{arcctg} x)' = \dfrac{1}{1+x^2}$.

Основные теоремы о дифференцируемых функциях

Следующий список утверждений является квинтэссенцией дифференциального исчисления.

Теорема 53 (Ферма). *Пусть функция f определена в некоторой окрестности точки a, которая является точкой локального максимума (минимума) функции f, т. е. существует такая окрестность $U(a)$, что $f(x) \leqslant f(a)$ $(f(x) \geqslant f(a))$ для всех $x \in U(a)$. Если f дифференцируема в точке a, то $f'(a) = 0$.*

$\triangleright$ Пусть a — точка локального максимума функции f.

Если $x \in U(a)$ и $x < a$, то $f(x) - f(a) \leqslant 0$ и $x - a < 0$. Следовательно,

$$f'(a) = \lim_{x \to a-0} \frac{f(x) - f(a)}{x - a} \geqslant 0.$$

Если $x \in U(a)$ и $x > a$, то $f(x) - f(a) \leqslant 0$ и $x - a > 0$. Следовательно,

$$f'(a) = \lim_{x \to a-0} \frac{f(x) - f(a)}{x - a} \leqslant 0.$$

Из неравенств $f'(a) \leqslant 0$ и $f'(a) \geqslant 0$ заключаем, что $f'(a) = 0$. $\triangleleft$

Физическая интерпретация: при движении по прямой в точке максимального удаления обязательно должна быть остановка.

Геометрическая интерпретация: если график функции f имеет в точке локального максимума (минимума) касательную, то эта касательная горизонтальна.

Теорема 54 (Ролль). *Если функция f непрерывна на отрезке $[a; b]$ и дифференцируема на интервале $(a; b)$, причём $f(a) = f(b)$, то найдётся точка $c \in (a; b)$, в которой $f'(c) = 0$.*

$\triangleright$ Если максимум и минимум достигается в концах отрезка, то функция есть константа. Если максимум или минимум достигается внутри отрезка, то по теореме Ферма в этой точке производная функции равна нулю. $\triangleleft$

Физическая интерпретация: если при движении по прямой материальная точка вернулась в то положение, в котором была в начальный момент времени, то обязательно был момент, когда скорость равнялась нулю (момент разворота).

Геометрическая интерпретация: если в концах отрезка значения функции совпадают, а график имеет касательную во всякой внутренней точке этого отрезка, то среди касательных обязательно есть горизонтальная.

Теорема 55 (Лагранж). *Если функция f непрерывна на отрезке $[a; b]$ и дифференцируема на интервале $(a; b)$, то найдётся такая точка $c \in (a; b)$, что $f(b) - f(a) = f'(c)(b - a)$.*

$\triangleright$ Применяем теорему Ролля к функции

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

(вычитаем из функции секущую). $\triangleleft$

Физическая интерпретация: при движении по прямой всегда есть момент времени, в который мгновенная скорость равна средней скорости.

Геометрическая интерпретация: если график имеет касательную во всякой внутренней точке отрезка, то среди касательных есть такая, которая параллельна секущей, соединяющей точки графика, соответствующие концам отрезка.

Теорема 56 (Коши). *Если функции f и g непрерывны на отрезке $[a; b]$ и дифференцируемы на интервале $(a; b)$, то найдётся такая точка $c \in (a; b)$, что*

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

$\triangleright$ Применяем теорему Ролля к функции

$$F(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x) = = \begin{vmatrix} f(b) - f(a) & f(x) \\ g(b) - g(a) & g(x) \end{vmatrix}. \quad \triangleleft$$

Физическая и геометрическая интерпретация теоремы Коши. Пусть задана зависимость $(f(x), g(x))$ движения точки по плоскости от времени $x \in [a; b]$. Тогда если получаемая кривая (траектория движения точки) имеет вектор скорости при всех $x \in (a; b)$, то существует момент времени, в который вектор скорости коллинеарен вектору полного перемещения.

Важным применением теоремы Коши является правило Лопиталя.

Теорема 57 (Бернулли — Лопиталь). *Предположим, что функции f и g дифференцируемы на интервале $(a; b)$, причём $g'(x) \neq 0$ при $x \in (a; b)$. Пусть также выполняется одно из следующих двух условий:*

1) $\lim\limits_{x \to b-} f(x) = \lim\limits_{x \to b-} g(x) = 0$ *или* 2) $\lim\limits_{x \to b-} g(x) = \infty$.

Тогда если $\lim\limits_{x \to b-} \frac{f'(x)}{g'(x)} = A$, *то* $\lim\limits_{x \to b-} \frac{f(x)}{g(x)} = A$.

$\triangleright$ Пусть $x, y \in (a; b)$. По теореме Коши найдётся такое число c, лежащее строго между x и y, что

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}.$$

Пусть выполнено второе условие: $\lim\limits_{x \to b-} g(x) = \infty$. Будем считать, что $a < y < c < x < b$. Выбирая y достаточно близко к b, можно приблизить отношение $\frac{f'(c)}{g'(c)}$ к A с любой наперёд заданной точностью. Теперь зафиксируем y и устремим x к b. Тогда выражения $\frac{g(y)}{g(x)}$ и $\frac{f(y)}{g(x)}$ приближаются к нулю, а вся правая часть приближается к A. Аналогичным образом проводятся рассуждения, когда выполнено первое условие. Формализация проведённых рассуждений на языке $\varepsilon - \delta$ предлагается читателю в качестве полезного упражнения. $\triangleleft$

Производная дифференцируемой функции не обязана быть непрерывной функцией (например, функция $f(x) = x^2 \sin(1/x)$ при $x \neq 0$ и $f(0) = 0$ дифференцируема на всей числовой прямой, а её производная разрывна в нуле). Однако у производной обязательно есть точка непрерывности.

Теорема 58. *Пусть функция f всюду дифференцируема на интервале I. Тогда её производная $f'(x)$ имеет всюду плотное в $\mathbb{R}$ множество точек непрерывности.*

$\triangleright$ Функция $f'(x)$ является поточечным пределом последовательности непрерывных функций $g_n(x) = n(f(x + 1/n) - f(x))$. $\triangleleft$

Кроме того, производная принимает все свои промежуточные значения.

Теорема 59 (Дарбу). *Пусть функция f дифференцируема на некотором интервале I и $a, b \in I$, $a < b$. Тогда для всякого числа C, лежащего между $A = f'(a)$ и $B = f'(b)$ найдётся точка $c \in [a; b]$, в которой $f'(c) = C$.*

$\triangleright$ Пусть $A < C < B$. Рассмотрим функцию $F(x) = f(x) - Cx$. Поскольку $F'(a) < 0$ и $F'(b) > 0$, на некотором интервале $(a; a + \delta)$ верно неравенство $f(x) < f(a)$ и на некотором интервале $(b; b - \delta)$ верно неравенство $f(x) < f(b)$. Следовательно, на концах отрезка не может достигаться минимум функции F. Следовательно, минимум достигается внутри отрезка и в этой точке по теореме Ферма функция F' равна нулю. $\triangleleft$

Таким образом, у производной не может быть устранимых разрывов или разрывов первого рода.

Формула Тейлора. Ряд Тейлора

Понятие производной n-го порядка вводится индуктивно. Пусть в каждой точке некоторой окрестности $U(a)$ точки a существует производная функции f. Тогда определена функция $f' : U(a) \to \mathbb{R}$. Если эта функция дифференцируема в точке a, то её производная называется *второй производной* функции f в точке a и обозначается через $f''(a)$. Далее, если в окрестности точки a определена функция $f^{(n-1)}$, то значение $(f^{(n-1)})'(a)$ (если оно существует) называется *производной порядка n* в точке a и обозначается через $f^{(n)}(a)$. Например, производная n-го порядка функции x^m, где $m \in \mathbb{N}$, равна 0 при $n > m$, равна $m!$ при $n = m$ и равна $m(m-1)\cdots(m-n+1)x^{m-n}$ при $n < m$. Удобно считать, что функция f является своей «нулевой» производной, т. е. $f^{(0)} = f$.

Теорема 60 (обобщение правила Лейбница)**.** *Пусть функции f и g дифференцируемы n раз в точке a. Тогда функция fg имеет производную порядка n в точке a и справедливо равенство*

$$(fg)^{(n)}(a) = \sum_{k=0}^{n} C_n^k f^{(k)}(a) g^{(n-k)}(a).$$

$\triangleright$ Воспользуемся методом математической индукции. Для $n = 1$ требуемое утверждение — обычное правило Лейбница. Предположим, что для некоторого n утверждение уже доказано; тогда для $n+1$ имеем

$$(fg)^{(n+1)} = ((fg)^{(n)})' = \Big(\sum_{k=0}^{n} C_n^k f^{(k)} g^{(n-k)}\Big)' =$$

$$= f^{(n+1)}g + fg^{(n+1)} + \sum_{k=0}^{n-1} C_n^k f^{(k+1)} g^{(n-k)} + \sum_{k=1}^{n} C_n^k f^{(k)} g^{(n-k+1)}.$$

Заменяя в первой сумме индекс суммирования k на $k+1$ и учитывая, что $C_n^{k-1} + C_n^k = C_{n+1}^k$, получаем требуемое равенство для $(fg)^{n+1}$. $\triangleleft$

Например,

$$(x^2e^x)^{(n)} = \sum_{k=0}^{n} C_n^k (x^2)^{(k)} (e^x)^{(n-k)} = (C_n^0 x^2 + 2C_n^1 x + 2C_n^2)e^x.$$

Пусть многочлен P записан в виде

$$P(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \ldots + c_n(x-a)^n.$$

Дифференцируя k раз многочлен P и подставляя $x = a$, получаем

$$c_k = \frac{P^{(k)}(a)}{k!}.$$

Это наблюдение позволяет написать многочлен, имеющий заданные значения первых n производных в данной точке.

Пусть в окрестности точки a задана функция f, имеющая в точке a все производные до порядка n включительно. Многочлен

$$T_n(x) = T_n(x; f, a) = \sum_{k=0}^{n} \frac{f^k(a)}{k!}(x-a)^k$$

называется *многочленом Тейлора* порядка n функции f в точке a. Заметим, что это просто многочлен степени не выше n, у которого производные до порядка n включительно совпадают в точке a с производными функции f. Естественно ожидать, что функция f и её многочлен Тейлора в точке a мало отличаются при $x \to a$.

Теорема 61. *Пусть функция f имеет в точке a все производные до порядка n включительно. Тогда*

$$f(x) - T_n(x) = \alpha(x)(x-a)^n, \quad \lim_{x\to a} \alpha(x) = 0.$$

$\triangleright$ По правилу Лопиталя и определению производной получаем

$$\lim_{x\to a} \frac{f(x) - T_n(x)}{(x-a)^n} = \lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a) - f^{(n)}(a)(x-a)}{n!(x-a)} =$$

$$= \frac{1}{n!}\Big(\lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{(x-a)} - f^{(n)}(a)\Big) = 0. \quad \triangleleft$$

Напомним введённые в главе 8 обозначения («o-символику»). Пусть в некоторой проколотой окрестности точки a выполнено равенство

$$h(x) = \alpha(x)g(x).$$

Тогда

если $\lim\limits_{x\to a} \alpha(x) = 0$, то пишут $h = o(g)$ при $x \to a$;

если $\alpha(x)$ ограниченна, то пишут $h = O(g)$ при $x \to a$;

если $\lim\limits_{x\to a} \alpha(x) = 1$, то пишут $h \sim g$ при $x \to a$.

Например, утверждение теоремы можно переписать в следующем виде:

$$f(x) = T_n(x) + o((x-a)^n) \quad \text{при } x \to a.$$

Это равенство называют формулой Тейлора с остаточным членом *в форме Пеано.*

Выпишем формулу Тейлора с остаточным членом в форме Пеано для элементарных функций в точке $a = 0$:

1) $e^x = \sum\limits_{k=0}^{n} \dfrac{x^k}{k!} + o(x^n);$

2) $\sin x = \sum\limits_{k=0}^{n} \dfrac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2n+2});$

3) $\cos x = \sum\limits_{k=0}^{n} \dfrac{(-1)^k x^{2k}}{(2k)!} + o(x^{2n+1});$

4) $\ln(1+x) = \sum\limits_{k=1}^{n} \dfrac{(-1)^{k-1} x^k}{k} + o(x^n);$

5) $(1+x)^\alpha = \sum\limits_{k=0}^{n} \dfrac{\alpha(\alpha-1)\cdots(\alpha-k+1)x^k}{k!} + o(x^n).$

Представляет интерес получение более точного вида остаточного члена

$$R_n(x) = R_n(x; f, a) = f(x) - T_n(x; f, a).$$

Теорема 62. *Пусть функция f имеет на отрезке $[x; a]$ непрерывные производные до порядка n включительно и на интервале $(x; a)$ производную порядка $n+1$. Пусть, далее, функция g непрерывна на $[x; a]$ и дифференцируема на $(x; a)$, причём $g'(t) \neq 0$ при $t \in (x; a)$. Тогда*

$$R_n(x; f, a) = \frac{f^{(n+1)}(c)}{n!\, g'(c)}(g(x) - g(a))(x-c)^n,$$

где c — некоторая точка интервала $(x; a)$.

$\triangleright$ Рассмотрим функцию

$$F(t) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k.$$

Заметим, что

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

По теореме Коши имеет место равенство

$$\frac{F(a) - F(x)}{g(a) - g(x)} = \frac{F'(c)}{g'(c)},$$

подставляя в которое выражение для $F'(c)$, получаем требуемое соотношение. $\triangleleft$

Рассмотрим важные частные случаи.

Остаточный член в форме Коши. Пусть $g(t) = t - x$. Тогда $g'(c) = 1$, $g(x) - g(a) = x - a$ и

$$R_n(x; f, a) = \frac{f^{(n+1)}(c)}{n!}(x-c)^n(x-a).$$

Остаточный член в форме Лагранжа. Пусть $g(t) = (t-x)^{n+1}$. Тогда $g'(c) = (n+1)(c-x)^n$, $g(x) - g(a) = -(a-x)^{n+1}$ и

$$R_n(x; f, a) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Пусть функция f имеет в точке a производные любого порядка (говорят, что функция *бесконечно дифференцируема* в этой точке). Полезно иметь ввиду следующее утверждение.

Теорема 63 (Борель)**.** *Для всякой числовой последовательности $\{c_n\}$ существует такая бесконечно дифференцируемая на всей числовой прямой функция f, что $f^{(k)}(0) = c_k$ для всякого k.*

Доказательство мы оставляем в качестве задачи для самостоятельного решения.

Если функция f бесконечно дифференцируема в точке a, то ряд

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!}$$

называется *рядом Тейлора* функции f в точке a.

Может оказаться, что такой ряд сходится лишь при $x = a$ или сходится, но не к функции f. Если в некоторой окрестности $U(a)$ ряд Тейлора функции f сходится к f, то функцию f называют *аналитической* в окрестности $U(a)$.

Функция $f(x) = e^{-1/x}$ при $x > 0$ и $f(x) = 0$ при $x \leqslant 0$ является классическим примером Коши функции, у которой ряд Тейлора в нуле не сходится к f ни в какой окрестности нуля.

Теорема 64. *Пусть функция f бесконечно дифференцируема на интервале $(a-r; a+r)$, причём $|f^{(n)}(x)| \leqslant CM^n n!$ для всех $x \in (a-r; a+r)$ и некоторых чисел $C > 0$ и $M > 0$. Если $M < 1/r$, то ряд Тейлора функции f в точке a сходится к f на интервале $(a-r; a+r)$ равномерно.*

$\triangleright$ По формуле Тейлора с остаточным членом в форме Лагранжа для всякого $x \in (a-r; a+r)$ найдётся такая точка c между ними, что

$$\left|f(x) - \sum_{n=0}^{N} \frac{f^{(n)}(a)(x-a)^n}{n!}\right| = \left|\frac{f^{(n+1)}(c)(x-a)^{n+1}}{(n+1)!}\right|.$$

Остаётся заметить, что правая часть оценивается сверху выражением $C(Mr)^n$ и $Mr < 1$. $\triangleleft$

Доказанная теорема позволяет обосновать сходимость рядов Тейлора функций e^x, $\sin x$, $\cos x$ и получить равенства

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}, \quad \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!},$$

в которых ряды сходятся равномерно на всяком отрезке числовой прямой.

Сложнее разобраться с рядами Тейлора для функций $\ln(1+x)$ и $(1+x)^\alpha$. Рассмотрим $f(x) = (1+x)^\alpha$ при $\alpha \notin \mathbb{N} \cup \{0\}$. Заметим, что

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

Полученное выше достаточное условие сходимости ряда Тейлора позволяет обосновать равенство

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$$

лишь в малой окрестности нуля. Однако можно показать, что при $|x| < 1$ ряд сходится и его сумма равна $(1+x)^\alpha$. Для обоснования сходимости на интервале $(-1; 1)$ вместо остаточного члена в форме Лагранжа следует использовать остаточный член в форме Коши. Пусть $x \neq 0$. Остаточный член в форме Коши имеет вид

$$\frac{f^{(n+1)}(c)(x-c)^n x}{n!} = \frac{\alpha(\alpha-1)\cdots(\alpha-n)(1+c)^{\alpha-1}x}{n!} \left(\frac{x-c}{1+c}\right)^n,$$

где c лежит между 0 и x. Если $x > 0$, то

$$|x-c| = |x||1-cx^{-1}| \leqslant |x|(1+c).$$

Если $x < 0$, то

$$|x-c| = |x||1-cx^{-1}| \leqslant |x|(1-|c|) = |x||1+c|.$$

Таким образом, верна оценка

$$\left|\frac{x-c}{1+c}\right| \leqslant |x|,$$

из которой немедленно следует стремление остаточного члена к нулю при $|x| < 1$. Аналогичным образом можно исследовать сходимость ряда Тейлора функции $\ln(1+x)$.

Исследование функций: монотонность, экстремумы, выпуклость

Применим полученные выше результаты к исследованию функций.

Монотонность

Исследование дифференцируемых функций на монотонность основано на следующем утверждении.

Предложение 12. *Пусть функция f дифференцируема на интервале $(a;b)$.*

(i) *Функция f неубывающая (невозрастающая) на $(a;b)$ тогда и только тогда, когда $f' \geqslant 0$ $(f' \leqslant 0)$ на $(a;b)$.*

(ii) *Если $f' > 0$ $(f' < 0)$ на $(a;b)$, то f возрастает (убывает) на $(a;b)$.*

(iii) *$f' = 0$ на $(a;b)$ тогда и только тогда, когда $f = \mathrm{const}$ на $(a;b)$.*

$\triangleright$ Сразу отметим, что (iii) следует из (i). Докажем (i). Если функция f неубывающая (невозрастающая), то отношение

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

неотрицательно (неположительно) для всех точек x_1 и x_2 интервала $(a;b)$. Следовательно, предел при $x_1 \to x_2$, равный $f'(x_2)$, также неотрицателен (неположителен).

В обратную сторону утверждение следует из теоремы Лагранжа. Для всяких x_1 и x_2 верно равенство

$$f(x_1) - f(x_2) = f'(c)(x_1 - x_2),$$

где точка c лежит между x_1 и x_2. Если $f' \geqslant 0$, то $f(x_2) \geqslant f(x_1)$ при $x_2 > x_1$ и функция f неубывающая. Если $f' \leqslant 0$, то при $x_2 > x_1$ получаем $f(x_2) \leqslant f(x_1)$, а значит, функция f невозрастающая. Для доказательства пункта (ii) достаточно заметить, что в случае строгой положительности f' или строгой отрицательности f' неравенство между $f(x_1)$ и $f(x_2)$ будет строгим. $\triangleleft$

Экстремумы

Точка a называется точкой *локального максимума* (*минимума*) функции f, если f определена в некоторой окрестности $U(a)$ точки a и $f(a) \geqslant f(x)$ $(f(a) \leqslant f(x))$ для всех $x \in U(a)$. Если при $x \neq a$ неравенства строгие, то говорят о *строгом минимуме* или *строгом максимуме* функции f. Точки локального максимума или минимума называются точками *локального экстремума* функции, а точки строго локального максимума или минимума называются точками *строгого локального экстремума* функции.

Если функция f дифференцируема в точке a локального экстремума, то по теореме Ферма $f'(a) = 0$, т. е. равенство нулю производной является необходимым условием локального экстремума.

Используя достаточные условия монотонности, можно сформулировать достаточное условие экстремума, например: если функция определена и непрерывна в окрестности $U(a)$, дифференцируема в проколотой окрестности $U'(a)$, причём $f' \geqslant 0$ при $x < a$ и $f' \leqslant 0$ при $x > a$, то точка a является точкой локального максимума.

Формула Тейлора позволяет получить достаточные условия экстремума в терминах производных высокого порядка. Предположим, что f имеет n производных в окрестности точки a, причём

$$f(a) = f'(a) = \ldots = f^{(n-1)}(a) = 0 \quad \text{и} \quad f^{(n)}(a) \neq 0.$$

Тогда по формуле Тейлора получаем

$$f(x) - f(a) = \left(\frac{f^{(n)}(a)}{n!} + \alpha(x)\right)(x-a)^n, \quad \lim_{x \to a} \alpha(x) = 0.$$

Если $n = 2k$, то в достаточно малой окрестности точки a разность $f(x) - f(a)$ имеет тот же знак, что и $f^{(n)}(a)$. Если $n = 2k+1$, то разность $f(x) - f(a)$ справа и слева от a имеет различные знаки. Таким образом, имеет место следующее утверждение.

Теорема 65. *Предположим, что функция f имеет n производных в окрестности точки a и*
$$f(a) = f'(a) = \ldots = f^{(n-1)}(a) = 0 \quad \text{и} \quad f^{(n)}(a) \neq 0.$$
Если $n = 2k$ и $f^{(n)}(a) < 0$ ($f^{(n)}(a) > 0$), то a — точка локального максимума (минимума). Если $n = 2k+1$, то точка a не является точкой локального экстремума функции f.

Выпуклость

Будем говорить, что функция f *выпукла* на интервале $(a;b)$, если для всех $x_1, x_2 \in (a;b)$ и всякого $\lambda \in [0;1]$ выполняется неравенство
$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda)f(x_2).$$
Если для всех $x \neq y$ и $\lambda \in (0;1)$ неравенство строгое, то функция f называется *строго выпуклой*. Наконец, если функция $(-f)$ выпукла (строго выпукла), то говорят, что f *вогнута* (*строго вогнута*).

Заметим, что точка с координатами
$$(\lambda x_1 + (1-\lambda)x_2, \lambda f(x_1) + (1-\lambda)f(x_2))$$
лежит на хорде
$$y = \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) + f(x_1),$$
стягивающей точки $(x_1, f(x_1))$ и $(x_2, f(x_2))$, а
$$(\lambda x_1 + (1-\lambda)x_2, f(\lambda x_1 + (1-\lambda)x_2))$$
есть точка на дуге графика функции $y = f(x)$, которую стягивает эта хорда. Таким образом, геометрический смысл выпуклости состоит в том, что любая хорда лежит не ниже дуги графика, которую она стягивает.

Предложение 13. *Функция f выпукла на интервале $(a;b)$ тогда и только тогда, когда для любых точек $x_1 < x < x_2$ из этого интервала выполнено неравенство*
$$\frac{f(x) - f(x_1)}{x - x_1} \leqslant \frac{f(x_2) - f(x)}{x_2 - x}.$$
$\triangleright$ Обозначим выражение $\lambda x_1 + (1-\lambda)x_2$ через x. Получаем
$$\lambda = \frac{x_2 - x}{x_2 - x_1}, \quad 1 - \lambda = \frac{x - x_1}{x_2 - x_1}.$$
Неравенство $f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda)f(x_2)$ равносильно неравенству
$$f(x) \leqslant \frac{x_2 - x}{x_2 - x_1}f(x_1) + \frac{x - x_1}{x_2 - x_1}f(x_2),$$
которое, в свою очередь, можно записать в виде
$$\frac{f(x) - f(x_1)}{x - x_1} \leqslant \frac{f(x_2) - f(x)}{x_2 - x}. \qquad \triangleleft$$

Таким образом, если функция f выпукла на интервале $(a;b)$, то для любых четырёх точек $x_1 < x_2 < x_3 < x_4$ из этого интервала наклон хорды, соответствующей точкам x_1, x_3, не превосходит наклона хорды, соответствующей точкам x_2, x_4, т. е. *наклоны хорд графика выпуклой функции не убывают.*

Теорема 66. *Пусть функция f выпукла на интервале $(a;b)$. Тогда для всякого отрезка $[c;d] \subset (a;b)$ существует такое число L, что для всех $x, y \in [c;d]$ выполнено неравенство*
$$|f(x) - f(y)| \leqslant L|x - y|,$$
в частности, функция f непрерывна на $(a;b)$.

$\triangleright$ Пусть $a < u < c$ и $d < v < b$. Тогда для всех $c < x < y < d$ выполняется неравенство
$$A = \frac{f(c) - f(u)}{c - u} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(v) - f(d)}{v - d} = B,$$
из которого следует искомое неравенство для $L = \max\{|A|, |B|\}$. Непрерывность функции f следует из доказанного неравенства. $\triangleleft$

Следующее утверждение содержит равносильные определения выпуклости для дифференцируемых функций.

Теорема 67. *Пусть функция f дифференцируема на $(a;b)$. Тогда следующие утверждения эквивалентны:*

(i) f — *выпукла (строго выпукла) на* $(a;b)$;

(ii) f' — *неубывающая (возрастающая) на* $(a;b)$;

(iii) $f(x) \underset{(>)}{\geqslant} f'(z)(x-z)+f(z)$ *для всех* $x,z\in(a;b)$, $x\neq z$.

$\triangleright$ Выведем (ii) из (i). Пусть $s<x_1<u<v<x_2<t$. Наклоны хорд не убывают, поэтому

$$\frac{f(x_1)-f(s)}{x_1-s}\leqslant\frac{f(u)-f(x_1)}{u-x_1}\underset{(<)}{\leqslant}\frac{f(x_2)-f(v)}{x_2-v}\leqslant\frac{f(t)-f(x_2)}{t-x_2}.$$

Устремляя $s\to x_1$ и $t\to x_2$, заключаем, что $f'(x_1)\leqslant f'(x_2)$ для выпуклой функции f и $f'(x_1)<f'(x_2)$ для строго выпуклой функции f.

Покажем, что (iii) следует из (ii). По теореме Лагранжа справедливо равенство $f(x)-f(z)=f'(c)(x-z)$, где c лежит между x и z. Следовательно,

$$f(x)-(f'(z)(x-z)+f(z))=(f'(c)-f'(z))(x-z)\geqslant 0,$$

где последнее неравенство выполняется в силу монотонности f', причём в случае строгой монотонности f' последнее неравенство строгое.

Выведем (i) из (iii). Заметим, что утверждение (iii) означает, что наклон хорды, стягивающей точки $(x,f(x))$ и $(z,f(z))$, где $x<z$, ограничен снизу значением производной $f'(x)$, а сверху — значением производной $f'(z)$. Пусть $x_1<x<x_2$. Тогда

$$\frac{f(x)-f(x_1)}{x-x_1}\leqslant f'(x)\leqslant\frac{f(x_2)-f(x)}{x_2-x},$$

т. е. наклоны хорд не убывают и f выпуклая функция. Отметим, что в случае строгого неравенства в утверждении (iii) и здесь получаем строгое неравенство между наклонами хорд, а значит, и строгую выпуклость. $\triangleleft$

Третье утверждение теоремы имеет наглядный геометрический смысл: график выпуклой дифференцируемой функции лежит не ниже всякой своей касательной.

Второе утверждение этой теоремы позволяет доказывать выпуклость дважды дифференцируемых функций.

Следствие 16. *Пусть функция f дважды дифференцируема на $(a;b)$.*

(i) *Функция f выпукла на $(a;b)$ тогда и только тогда, когда $f''\geqslant 0$ на $(a;b)$.*

(ii) *Если $f''>0$ на $(a;b)$, то f строго выпукла на $(a;b)$.*

Приведём полезное утверждение, доказательство которого является упражнением на метод математической индукции.

Теорема 68 (неравенство Йенсена)**.** *Пусть функция f выпукла на интервале $(a;b)$. Тогда для всех чисел $x_1,\ldots,x_n\in(a;b)$ и всех чисел $\lambda_1,\ldots,\lambda_n\in[0;1]$, для которых $\lambda_1+\ldots+\lambda_n=1$, выполняется неравенство*

$$f(\lambda_1x_1+\ldots+\lambda_nx_n)\leqslant\lambda_1f(x_1)+\ldots+\lambda_nf(x_n).$$

Рассмотрим пример. Применяя неравенство Йенсена к выпуклой функции $-\ln x$ и полагая $\lambda_1=\ldots=\lambda_n=1/n$, получаем для произвольных положительных чисел $x_1,\ldots,x_n$ известное неравенство между средним арифметическим и средним геометрическим:

$$\frac{x_1+\ldots+x_n}{n}\geqslant\sqrt[n]{x_1\cdots x_n}.$$

Отметим, что неравенство обращается в равенство тогда и только тогда, когда $x_1=x_2=\ldots=x_n$.

Говорят, что дифференцируемая в точке x_0 функция f имеет *перегиб* в точке x_0, если существует такое число $\delta>0$, что функция f выпукла (вогнута) на интервале $(x_0-\delta;x_0)$ и вогнута (выпукла) на интервале $(x_0;x_0+\delta)$. Для дифференцируемой в окрестности точки x_0 функции f наличие перегиба в точке x_0 означает изменение взаимного расположения графика функции и касательной к нему при переходе через эту точку.

В заключение рассмотрим простейший вариант знаменитого метода касательных Ньютона.

Метод касательных Ньютона

Предположим, что на интервале $(a;b)$ задана непрерывная функция f, причём в точках $u < v$ этого интервала значения f имеют разные знаки. Как найти решение уравнения $f(x) = 0$? Одно из доказательств теоремы о промежуточном значении (именно эта теорема гарантирует существование решения уравнения $f(x) = 0$) основано на *методе бисекций*, состоящего в том, что отрезок $[u;v]$ делится пополам, значение f в середине отрезка сравнивается с нулём, и в случае, когда оно отлично от нуля, делится пополам та половина, на концах которой значения функции f имеют разные знаки, и т. д. За n шагов мы будем знать решение с точностью до $(v-u)/2^n$. Оказывается, что при некоторых дополнительных условиях на функцию f можно добиться того, что за n шагов решение будет известно с точностью до $C/2^{2^n}$, где C — некоторая константа, не зависящая от n.

Пусть функция f дважды непрерывно дифференцируема на интервале $(a;b)$, причём для всех $x \in (a;b)$ и некоторых положительных чисел α и β выполнены неравенства $f'(x) \geqslant \alpha > 0$ и $\beta \geqslant f''(x) > 0$. Предположим, что существует такая точка $c \in (a;b)$, что $f(c) = 0$. При этих условиях функция f строго возрастает и, следовательно, такая точка c только одна.

Пусть $x_1 > c$. Зададим последовательность $\{x_n\}$ рекуррентно: если значение x_n уже определено, то x_{n+1} — точка пересечения касательной $y = f'(x_n)(x-x_n) + f(x_n)$ с осью Ox, т. е.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Предложение 14. *Во введённых обозначениях последовательность $\{x_n\}$ сходится к числу c.*

$\triangleright$ Покажем, что $x_n > c$ для всех n. По условию $x_1 > c$. Предположим, что для некоторого n это неравенство доказано и докажем его для $n+1$. График строго выпуклой функции лежит не ниже всякой своей касательной и имеет с ней единственную общую точку. Поскольку $f(x_n) > 0$, получаем $f(x_{n+1}) > 0$, что немедленно влечёт неравенство $x_{n+1} > c$.

Поскольку $x_n > c$, получаем $f(x_n) > 0$. Кроме того, по условию $f'(x_n) > 0$. Следовательно, $x_{n+1} < x_n$ и последовательность $\{x_n\}$ убывает. По теореме Вейерштрасса $\{x_n\}$ сходится к некоторому значению z. Устремляя $n \to \infty$ в рекуррентном соотношении, определяющем x_n, получаем $f(z) = 0$ и, значит, $z = c$. $\triangleleft$

Оценим скорость сходимости x_n к c. По формуле Тейлора с остаточным членом в форме Лагранжа имеем

$$f(c) = f(x_n) + f'(x_n)(c-x_n) + \frac{f''(u)(c-x_n)^2}{2},$$

где u лежит между x_n и c. Поскольку $f(c) = 0$, получаем

$$\Big|x_n - \frac{f(x_n)}{f'(x_n)} - c\Big| = \Big|\frac{f''(u)(c-x_n)^2}{2f'(x_n)}\Big|.$$

Заметим, что левая часть равна $|x_{n+1} - c|$. Используя неравенства $f'' \leqslant \beta$ и $f' \geqslant \alpha$, имеем

$$|x_{n+1} - c| \leqslant M|x_n - c|^2, \quad M = \frac{\beta}{2\alpha}.$$

Из этого неравенства следует оценка

$$|x_{n+1} - c| \leqslant M^{-1}(M|x_1 - c|)^{2^n}.$$

Если $M|x_1 - c| < 1/2$, то получаем $|x_{n+1} - c| \leqslant M^{-1}/2^{2^n}$. Последовательность x_n сходится к c, поэтому начиная с некоторого номера N неравенство $M|x_1 - c| < 1/2$ выполнено и можно в качестве x_1 взять x_N.

Типичным примером применения метода касательных Ньютона является итерационная формула Герона для вычисления величины $\sqrt{c}$, $c > 0$:

$$x_{n+1} = \frac{1}{2}\Big(x_n + \frac{c}{x_n}\Big).$$

Здесь метод Ньютона применяется к функции $f(x) = x^2 - c$ и последовательность $\{x_n\}$ сходится к $\sqrt{c}$.

Список литературы

1. Фихтенгольц Г. М. Курс дифференциального и интегрального исчисления. В 3-х тт. — М.: Физматлит, 2001. — Т. 1. — 616 с.

Учебник знаменитого отечественного математика Г. М. Фихтенгольца является одним из самых популярных пособий по математическому анализу и, несмотря на несколько архаичный язык изложения, содержит большое число интереснейших результатов, которые в настоящее время не включаются в курс анализа. Кроме того, в книге разобрано много интересных примеров, знакомство с которыми будет полезно современному студенту.

2. Зорич В. А. Математический анализ. Часть I. Изд. 7-е, новое доп. — М.: МЦНМО, 2015. — 576 с.

Книга представляет собой одно из самых полных и удачных изложений современного курса математического анализа и, несомненно, будет полезна студентам. Обратим особое внимание на задачи и упражнения, которые приводятся после каждого параграфа. Решение этих задач поможет читателю глубоко понять и прочувствовать основные определения и методы математического анализа.

3. Архипов Г. И., Садовничий В. А., Чубариков В. Н. Лекции по математическому анализу. 4-е изд., испр. — М.: Дрофа, 2004. — 640 с.

Учебник представляет собой запись реальных лекций, которые авторы многие годы читали на механико-математическом факультете МГУ имени М. В. Ломоносова, и отличается лаконичностью и отточенностью изложения. Книга содержит интересные дополнительные темы, которые не всегда включаются в курс анализа, например, теорему А. А. Карацубы о быстром умножении двух чисел и критерий Г. Вейля равномерного распределения элементов числовой последовательности.

4. Натансон И. П. Теория функций вещественной переменной. Изд. 2-е, перераб. — М.: ГИТТЛ, 1957. — 552 с.

Книга И. П. Натансона является введением в теорию функций и, безусловно, будет полезным дополнением к базовому курсу анализа. Материал излагается ясно и очень наглядно, проводится большое число красивых фактов, которые в другой литературе не затрагиваются вовсе. Читателю также рекомендуется обратить внимание на задачи и упражнения после каждой главы.

5. Шварц Л. Анализ. I-й том. — М.: Мир, 1972. — 824 с.

Книга известного французского математика Л. Шварца может быть рекомендована как дополнение к базовому курсу математического анализа и потребует от читателя изначально глубокого понимания его основ. Однако знакомство с этим учебником позволит шире взглянуть на изучаемый материал и лучше осознать суть идей и методов, которые встречаются в базовом курсе и часто кажутся набором весьма изобретательных трюков.

6. Демидович Б. П. Сборник задач и упражнений по математическому анализу: Учеб. пособие. — 13-е изд., испр. — М.: Изд-во Московского университета; ЧеРо, 1997. — 624 с.

Это классический сборник задач по всем темам курса математического анализа, который успешно используется для проведения практических занятий на математических факультетах ведущих вузов. Читатель найдёт здесь как простые упражнения алгоритмического характера, так и сложные глубокие утверждения, доказательства которых требуют уверенного владения методами и идеями анализа.

7. Виноградова И. А., Олехник С. Н., Садовничий В. А. Математический анализ в задачах и упражнениях: В 3-х т. Том 1: Дифференциальное и интегральное исчисление. — Новое изд. — М.: Изд-во Московского университета; МЦНМО, 2017. — 412 с.

Пособие содержит отличную подборку задач по всем темам курса математического анализа и основано на обобщении авторами богатого опыта многолетнего ведения практических занятий на механико-математическом факультете МГУ имени М. В. Ломоносова. Отличительной особенностью задачника является наличие в начале каждой главы подробного введения с обзором методов и разбором примеров. Данное пособие может быть рекомендовано для самостоятельного освоения методов решения задач по математическому анализу.

8. Макаров Б. М., Голузина М. Г., Лодкин А. А., Подкорытов А. Н. Избранные задачи по вещественному анализу: Учеб. пособие для вузов. — М.: Наука. Гл. ред. физ.-мат. лит., 1992. — 432 с.

Пособие содержит разнообразные задачи, в основном высокого уровня сложности почти по

всем темам математического анализа. Среди задач встречается ряд важных теорем, примеров и контрпримеров, знание которых существенным образом дополняет стандартный курс анализа. Кроме того, при решении задач из данного сборника можно познакомиться со многими методами и приемами, практически не обсуждаемыми на обычных занятиях по анализу.

9. Верещагин Н.К., Шень А. Лекции по математической логике и теории алгоритмов. Часть 1. Начала теории множеств. — 4-е изд., доп. — М.: МЦНМО, 2012. — 112 с.

Книга представляет собой замечательное по своей простоте и лаконичности введение в теорию множеств. Обсуждаются вполне упорядоченные множества и трансфинитная индукция, аксиома выбора, теорема Цермело и лемма Цорна.

10. Ландау Э. Основы анализа. 2-е изд. — М.: URSS, 2010. — 184 с.

В книге подробно обсуждаются основания математического анализа, в частности, приводится аккуратное построение основных числовых множеств. Здесь можно найти подробное обсуждение аксиоматики Пеано, а также индуктивное построение арифметических операций и отношения порядка на множестве натуральных чисел.

Отметим, что наряду с указанными выше, подойдут и другие издания перечисленных книг. Конечно, охватить всю литературу, которая может быть полезна изучающим курс математического анализа, не представляется возможным.